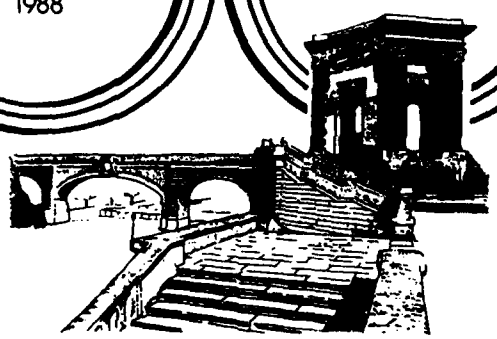
R&D No. 5983-EE-02

ESSDERC

MONTPELLIER
FRANCE
SEPTEMBER 13-16,
1988



18th European Solid State Device Research Conference

EDITED BY:
J.-P. NOUGIER
D. GASQUET

les éditions de physique

JOURNAL DE PHYSIQUE

Tome 49 Colloque C4, supplément au n° 9 Septembre 1988

ESSDERC 88

18th European Solid State Device Research Conference

September 13-16 1988
Montpellier (France)

Edited by : J.-P. NOUGIER
D. GASQUET

les éditions de physique

Avenue du Hoggar,
Zone Industrielle de Courtabœuf,
B.P. 112,
F-91944 Les Ulis Cedex, France

THE UNIVERSITY OF MONTPELLIER

short historical survey through the ages

In 1989, the University of Montpellier will celebrate its 700 th birthday. Indeed it was officially created by Pope Nicolas IV on October 26, 1289. The papal bull stated : "...Therefore, we wish that the precious pearl of science propagates through all countries and, mainly, in those which are recognized as being suited to the birth of that science. And Montpellier, illustrious and famous town, is renown as being fitted to study in a wonderful way. Hence our feeling is that, it is of public interest to fix there an intellectual centre, where wise men might collect the desired fruit that the science's supreme Leader would generate. That is the reason why, we decide that this town will set up, henceforth, a **Studium generale** where masters might teach, and students might learn attending lectures in these Faculties..."

Thus bringing together, in a single University, the Faculties of Medecine, of Law and of Arts, Pope Nicolas IV put the Schools of Montpellier at the same level as those of Paris and Bologna.

Indeed, the 3 Faculties were born far before 1289, and the creation of the University was but an official recognition of a preexisting situation. The oldest faculty is, officially, the Faculty of Law, the birth of which is considered as being in the year 1160 when an Italian from Piacenza came there, but one might infer that this was so because this faculty had some renown already. The Faculty of Medecine was born on August 17th, 1220, but it has been proved that it was already famous in the German town of Mayence in 1137. The faculty of Arts was officially born on March 20th, 1242.

Following the French Revolution of 1789, the University of Montpellier was closed, as was every French University, in 1792. Nevertheless, the Faculty of Medecine was created again as a "School of Health", as early as 1794, so were Paris and Strasbourg.

- The Faculty of Medecine was restored in 1803.
- The Faculty of Litterature (ancient faculty of Arts) was restored in 1808, and transformed into a Faculty of Litterature and Human Sciences in 1957.
- The Faculty of Law was restored in 1878 and became the Faculty of Law and Economic Sciences in 1957.
- In the meantime, the Faculty of Sciences had been created in 1809, as a developping branch of the ancient Faculty of Arts.

Other institutions were created :

- A school of Pharmacology, created in 1803 and transformed into a faculty in 1920.
- The Agronomy High School (Engineering school) in 1873.
- The Chemistry High School (Engineering school) in 1857.
- The University Institute of Technology (Technicians and Assistant Engineers) in 1968.
- The Science Institute for Engineers (Engineer school inside the faculty) in 1971.

This short overview clearly exhibits the dynamism of the University of Montpellier, which was split into 3 sub-Universities in 1968, and which, in spite of its old age, with its 47 000 students, is permanently recovering youth through unceasing struggles and fights against routine and facility, these two poisons of University spirit.

PREFACE

It is a great pleasure and a great honour to welcome the European Solid State Device Research Conference, for the second time in Montpellier.

The call for papers of the ESSDERC 88 was a great success, since 387 papers were submitted, among which 171 were selected for presentation at the conference. The accepted papers include the names of 565 different authors, scattered among 19 different countries. Besides the normal contributions, 14 invited papers have been selected among more than 60 proposals of the program committee and of the representatives of the ESSDERC.

These few statistics clearly exhibit the vitality and the dynamism of the field of Research in Devices, covered by the ESSDERC Conference. It also emphasises the efforts made by European countries in that domain. Now, not only Germany, Great Britain and France bring an important contribution to that conference, as had been the case during the last ten to fifteen years, but more and more European countries become deeply involved, scientifically and economically, in that strategic domain. From this point of view, two features are highly significant :

- Other European countries are now becoming included in the "ESSDERC circuit", starting from Italy (ESSDERC 87 held in Bologna).
- At the ESSDERC 78, also held in Montpellier, exactly 10 years ago, 136 papers had been selected, among 182 submitted : hence twice more papers have been submitted at the ESSDERC 88 than at the ESSDERC 78, which clearly shows the increasing growth of the field, and the rejection rate was at that time 25% instead of 56% now, which could indicate an increasing quality of the accepted papers.

Also, within this ten years period, two important events, related to the Conference, took place :

- The Montpellier ESSDERC 78 Conference was the first time when proceedings of contributed papers, and not just of invited papers, were published. At that time, only 52 papers had been published, of the 136 presented, after reviewing, during the Conference, the 55 papers submitted for publication. However, this attempt was abandoned in the subsequent Conferences, except in a few cases (Lille ESSDERC 84 for example).
- At the Bologna ESSDERC 87 Conference, the proceedings of all the contributed papers were for the first time provided to the participants at the beginning of the conference. This is considered as being a major improvement, which was tried to be kept in the Montpellier ESSDERC 88 Conference. The Organizors, as well as the Program Committee, strongly hope that this tradition, now established, will be kept in the future.

The ESSDERC Conference covers a broad spectrum of scientific topics, within the general scope of the physics and of the technology of logic and analog solid state devices, including fabrication, modeling, characterization.
Therefore, it is intended :

- to provide an international forum for scientists and professionals wishing to broaden their scientific culture.
- to encourage the exchange of technical information beyond the limits of specific areas.
- to encourage cooperation between scientists and professionals of different institutions and countries.

We hope that the Montpellier ESSDERC 88 Conference will help people in doing so.

Finally, I would like to thank all those who made a considerable effort in contributing to the organization of the Conference, in selecting papers, planning programmes, chairing sessions, preparing invited talks. A special acknowledgement is due to the Organizations, listed in a separate page, which have sponsored and supported the Conference.

Prof. J.P. NOUGIER
Chairman of the Conference

SUPPORTING ORGANIZATION

Direction des Recherches et Etudes Techniques (DRET)
Conseil Général du Languedoc Roussillon
Montpellier Languedoc Roussillon Technopôle
Centre National de la Recherche Scientifique (CNRS)
Siemens France
Centre National d'Etude des Télécommunications (CNET)
United State Army (USA-RDSG, London)
Conseil Général de l'Hérault
Bendix
International Business Machines (IBM France)
S.G.S. Thomson
Matra (France)
Société Anonyme des Télécommunications (SAT)

SPONSORS

European Physical Society

ORGANIZING COMMITEE

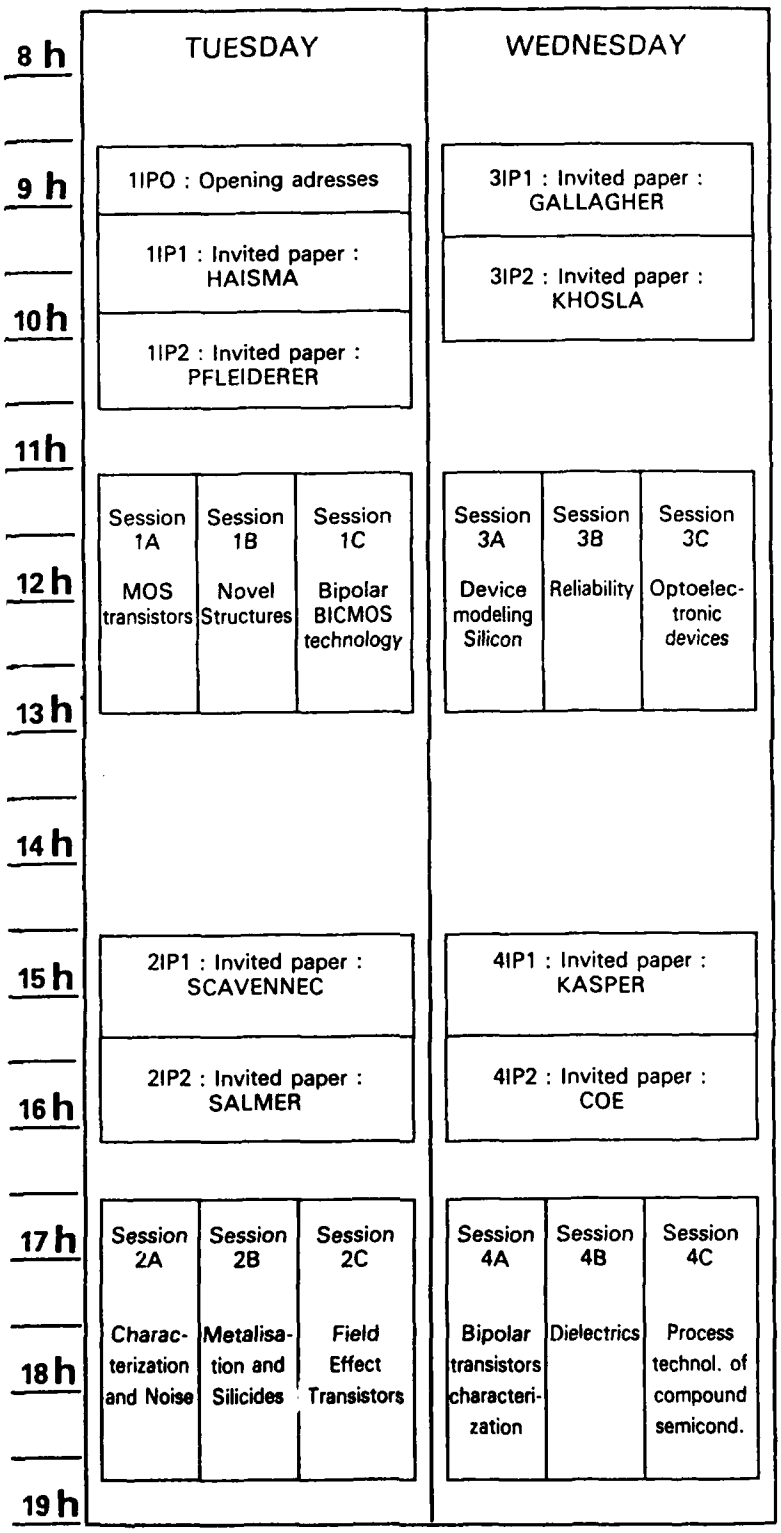
J.P.	NOUGIER	Chairman, Director of the C.E.M.
M.	SAVELLI	Vice Chairman (Finances), C.E.M.
G.	LECOY	Vice Chairman (Program), C.E.M.
J.P.	GASQUET	Secretary, C.E.M.
J.P.	NOBLANC	Chairman ESSDERC/ESSIRC Steer. Comm.
G.	SONCINI	Chairman ESSDERC 87
A.	HEUBERGER	Chairman ESSDERC 89

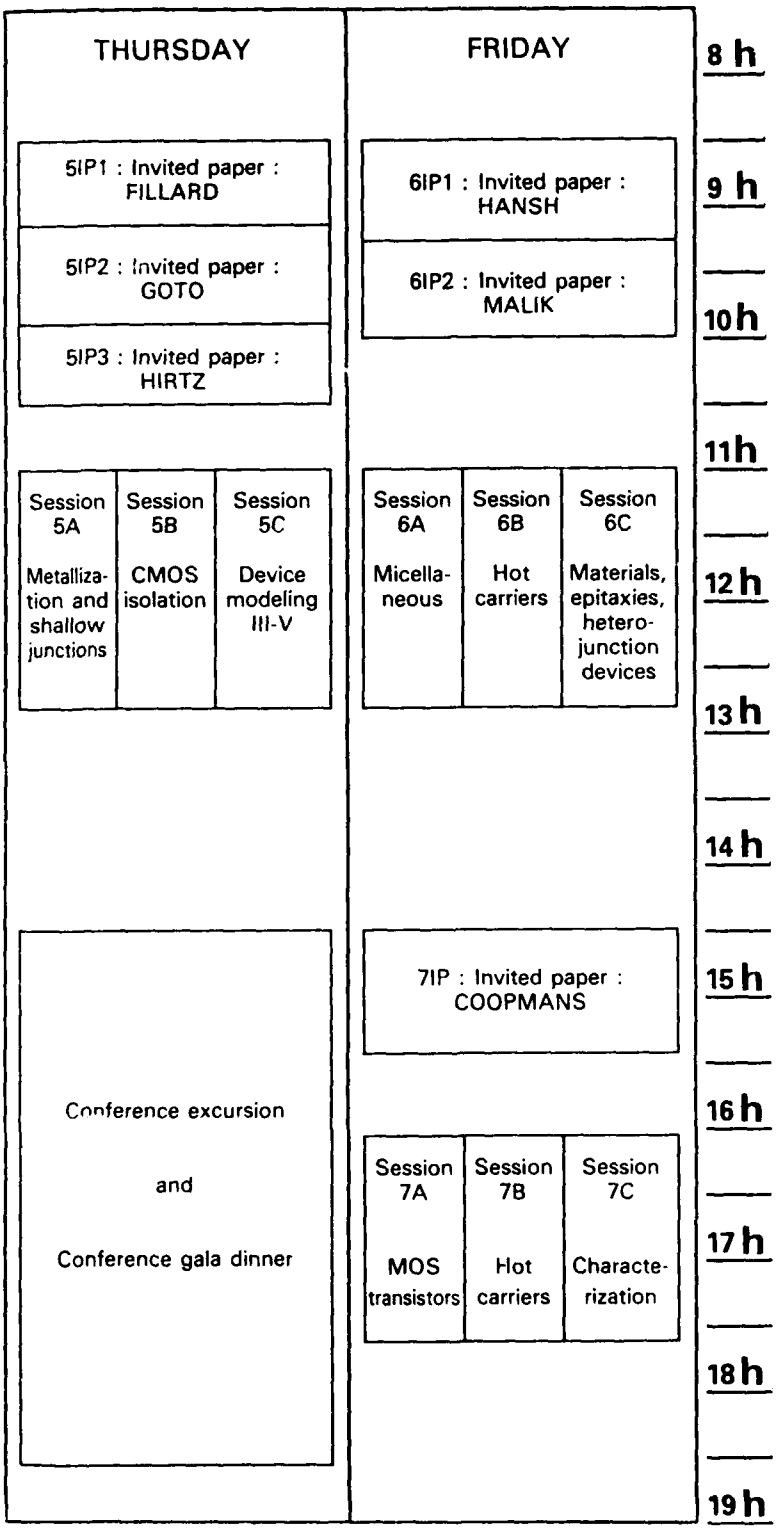
PROGRAM COMMITTEE

L.	BALDI	SGS Microelettron.	Italy
P.	BALK	Tech. Univ. Aachen	F.R.G.
E.	CONSTANT	Univ. Lille	France
G.	DECLERCK	IMEC Louvain	Belgium
H.C.	DE GRAAF	Philips Eindhoven	The Netherlands
J.P.	DUCHEMIN	Thomson DHM	France
V.	GHERGIA	CSELT Torino	Italy
M.	ILEGEMS	Lausanne Inst. Techn.	Switzerland
M.	MONTIER	Thomson Semic.	France
J.P.	NOBLANC	CNET Bagneux	France
J.P.	NOUGIER	CEM Univ. Montpellier	France
J.	ROBERTSON	Edimburgh Univ.	Great-Britain
M.	ROCCHI	LEP	France
H.	RUPPRECHT	Fraunhofer Freiburg	F.R.G.
F.	VAN DE WIELE	Univ. Cath. Louvain	Belgium
P.	WEISSGLAS	Inst. Micr. Techn.	Sweden
A.	WIEDER	SIEMENS, Munich	F.R.G.
B.	WILSON	Plessey	Great-Britain
C.	WOOD	GEC Wembley	Great-Britain

REPRESENTATIVES

U.S.A. : F. CAPASSO, Bell. Labs, Murray Hill
Japan : M. INOUE, Osaka Institute of Technology
China : Q.-Y. TONG, Nanjing Institute of Technology


8 h
9 h
10 h
11 h
12 h
13 h
14 h
15 h
16 h
17 h
18 h
19 h
TUESDAY
WEDNESDAY
1IPO : Opening adresses
1IP1 : Invited paper : HAISMA
1IP2 : Invited paper : PFLEIDERER
3IP1 : Invited paper : GALLAGHER
3IP2 : Invited paper : KHOSLA
Session 1A MOS transistors
Session 1B Novel Structures
Session 1C Bipolar BICMOS technology
Session 3A Device modeling Silicon
Session 3B Reliability
Session 3C Optoelectronic devices
2IP1 : Invited paper : SCAVENNEC
2IP2 : Invited paper : SALMER
4IP1 : Invited paper : KASPER
4IP2 : Invited paper : COE
Session 2A Characterization and Noise
Session 2B Metalisation and Silicides
Session 2C Field Effect Transistors
Session 4A Bipolar transistors characterization
Session 4B Dielectrics
Session 4C Process technol. of compound semicond.


THURSDAY
FRIDAY
8 h
5IP1 : Invited paper : FILLARD
6IP1 : Invited paper : HANSH
9 h
5IP2 : Invited paper : GOTO
6IP2 : Invited paper : MALIK
10h
5IP3 : Invited paper : HIRTZ
11h
Session 5A Metallization and shallow junctions
Session 5B CMOS isolation
Session 5C Device modeling III-V
Session 6A Micellaneous
Session 6B Hot carriers
Session 6C Materials, epitaxies, heterojunction devices
12 h
13 h
14 h
7IP : Invited paper : COOPMANS
15 h
Conference excursion
and
Conference gala dinner
16 h
Session 7A MOS transistors
Session 7B Hot carriers
Session 7C Characterization
17 h
18 h
19 h

CONTENTS*

*Italic titles have not been followed by a written contribution in these proceedings

Session 1B Room B : Novel structures
Chairman : J. HAÏSMA (Philips, The Netherlands)

Session 1C Room C : Bipolar, BICMOS : technology
Chairman : A. WIEDER (Siemens, F.R.G.)

Session 2C Room C : Field effect transistors
Chairman : M. ILEGEMS (E.P.L., Switzerland)

SESSION 3 (Wednesday morning)
Session 3IP Room A : Invited papers
Chairman : J. MAGARSHAK (Thomson, France)

Session 3A Room A : Device modeling
Chairman : H.C. DE GRAAFF (Philips, Eindhoven)

Session 3B Room B : Reliability
Chairman : F. NEPPL (Siemens, F.R.G.)

SESSION 4 (Wednesday afternoon)
Session 4IP Room A : Invited papers
Chairman : L. TREITINGER (Siemens, F.R.G.)

Session 4A Room A : Bipolar transistors : characterization
Chairman : R.E. OAKLEY (Plessey, Great-Britain)

Session 4B Room B : Dielectrics
Chairman : P. BALK (Aachen Univ., F.R.G.)

SESSION 5 (Thursday morning)
Session 5IP Room A : Invited papers
Chairman : J.P. DUCHEMIN (Thomson, France)

Session 5A Room A : Metallisation and shallow junctions
Chairman : R.F. DE KEERSMAECKER (IMEC, Belgium)

Session 5B Room B : CMOS isolation
Chairman : PFLEIDERER (Siemens, F.R.G.)

Session 5C Room C : Device modeling
Chairman : E. CONSTANT (C.H.S., France)

Session 6B Room B : Hot carriers
Chairman : L. BALDI (SGS Thomson, Italy)

Session 6C Room C : Materials, epitaxies, HJD
Chairman : B.L. WILSON (Plessey, Great-Britain)

SESSION 1

Session 1IP Room A : Invited papers

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

SOI TECHNOLOGIES : THEIR PAST, PRESENT AND FUTURE

J. HAISMA

Philips Research Laboratories, PO Box 80.000, NL-5600 JA Eindhoven, The Netherlands

Résumé - Les technologies de réalisation de silicium sur isolant (SOI) sont sommairement décrites avec quelques propriétés spécifiques. La réalisation de SOI ne requiert pas seulement un silicium parfaitement réalisé mais aussi des couches adjacentes non perturbées par l'action de la récristallisation et un substrat qui est resté en bon état.
Les problèmes expérimentaux établis par les techniques SOI sont indiqués et discutés.
Le passé, le présent et l'avenir de SOI seront évalués sans vouloir être complet.

Abstract - Technologies applied for silicon on insulator (SOI) are briefly described as well as their specific properties. Conditions of realizing SOI not only imply obtaining a perfect monocrystalline silicon layer but also an undamaged stack of (SOI) layers and a defect-free substrate.
Technological problems encountered by applying the various SOI techniques are described and discussed in some detail.
Past, present and future of SOI will be evaluated in some aspects.

1 - MOTIVES

Silicon-on-sapphire (SOS) can certainly be entitled as one of the progenitors of silicon-on-insulator (SOI), an heteroepitaxial approach of the latter subject. It had then probably as a first incentive to make available monocrystalline silicon on top of an insulator giving in this way a radiation-hard material which had to remain operative under circumstances of severe radiation impact. SOS has had a flourishing time of research in the early and mid seventies. Deceptions have, however, arisen due to the mediocre quality of the silicon brought about that time. Only recently this circumstance has been improved[1] by careful materials engineering.
In the meantime researchers have explored new issues[2,3,4,5] for the realization of SOI but now directed towards the more generally progressing development of the IC technology.
These novel goals imply a higher operating speed, lower dynamic power consumption, greater packing density, ameliorated radiation tolerances, some simpler fabrication sequences, more layout flexibility and immunity to latch-up for CMOS circuits.

The incentives and potential advantages of SOI and device applications have been well described and formulated in the past[6].

The state of the art of SOI technologies and the aims for the future, provided with some marginal notes, will be the subject matter of this paper.

2 - STATE OF THE ART OF EXISTING SOI TECHNOLOGIES

The list below shows the subjects which have to be treated to give a more or less complete impression of the technologies nowadays pursued.

I. Zone-Melting Recrystallization (ZMR)
   - laser recrystallization (Argon-ion, CO_2);
   - electron-beam recrystallization
   - recrystallization by lamp annealing
   - recrystallization by graphite-strip heater

II. Separation by IMplanted OXygen (SIMOX);

III. Full Insulation by Porous Oxidized Silicon (FIPOS);

IV. Lateral Solid-Phase Epitaxial Growth (L-SPEG);

V. Wafer bonding and wafer thinning;

VI. Heteroepitaxial SOI.

The subjects mentioned are reviewed following this scheme.

I. Laser & electron-beam recrystallization; lamp & graphite-strip heater recrystallization

Laser physics has evaluated to mature laser tools for industrial applications. At least two types are available as continuous wave lasers with a sufficiently high power output to this aim: the Argon-ion laser in the visible (more than 20 Watts C.W.) and the CO_2 laser at 10,6 µm wavelength (normally about 50 Watts C.W. but if requested more than 1 kWatt). Both laser types have in SOI their own merits.

The Argon-ion laser tube has a reasonable short life (2000 hours) and the power is somewhat limited e.g. to form a broader linespot. The CO_2 laser has a long life due to the fact that it is a continous gas-flowing system but CO_2-laser radiation is hardly absorbed by undoped silicon at low temperatures. Therefore it has to start heating up the polycrystalline silicon by absorption in the sandwiching SiO_2 layers. Only from 800°C on silicon is absorbing 10.6 µm radiation by its free electrons. So the two laser systems have intrinsic differences and the achievements with them cannot be exchanged without due consideration.

By far the main part of laser-recrystallization research has been done with the Argon-ion laser and often very successfully. Nevertheless, no example is known so far having reached a real industrial level. This is probably due to the fact that specific wafer engineering is required in order to obtain a result sufficiently outstanding to be mentioned fully reproducible and adapted to the IC's of today. Or in other words, the adventitious circumstances are so specific that for the moment no direct transfer from research to industry seems to be undertaken. The probable reasons will be discussed in more detail below.

Electron-beam recrystallization is another possibility for local ZMR. Electron-beam heating has the advantage that all materials whereupon the electron beam impinges, are fully absorbing. Therefore the input energy is primarily dissipated in the top layer.

Furthermore an electron beam can easily be directed/transformed in shape/scanned by electromagnetic means. Beam shaping and scan velocity can be mastered in a most flexible way. The energy can be made sufficiently large for line scanning, bringing about larger area recrystallization in one run. Nowadays vacuum has to be considered as mastered. So electron-beam recrystallization is certainly one of the candidates for an SOI technology for the future. However, electron-beam systems adapted to an SOI technology have not been made commercially available and this fact surely has hampered a larger scale evaluation with it.

Laser recrystallization and electron-beam recrystallization are typical local-area recrystallization technologies, limiting a high production of throughput if wafers have to be treated at full width.

This limitation is removed by the next two technologies: ZMR by lamp or graphite-strip heating. Both technologies need a homogeneous substrate temperature in the order of 1100°C, normally supplied from below the wafer. As soon as the bias temperature has been reached a line source (focussed lamp light or heated graphite strip) melts the polysilicon layer over the entire width of the wafer in one run. Lamp heating itself is very clean (the carbon substrate holder might contaminate), carbon-strip heating gives somewhat contamination due to evaporation of carbon. Device-quality SOI has been realised this way. Due to the high substrate temperature, the thermal gradient of the zone-melting silicon layer is small which is favourable for a defect-free silicon in selected areas. Optimum engineering entrains the defects to predetermined line sections; often only isolated defects occur.

II. SIMOX (Separation by Implanted Oxygen)

The incentive of ZMR is to transform polysilicon to monocrystalline silicon. Another approach is to start from perfect single-crystalline material and to create a buried electrically insulating layer below the surface with the restriction that the top silicon layer remains in a single-crystalline state during the treatments. This challenge has been met. Standard silicon wafers are implanted with a high dose of oxygen ions (dose e.g.: 2.5×10^{18} O^+ cm^{-2}; energy: 200 keV) at a temperature between 500 and 650°C, keeping the wafer out of the amorphization range due to in-situ annealing. Nowadays high-current oxygen implanters are available to realize such a high dose in a reasonable time (but still hours of implantation).

Implantation is followed by a high-temperature anneal step in excess of 1300°C to restore the implantation-damaged superficial Si layer and to eliminate precipitates of SiO_2 by dissolving them and diffusing the oxygen atoms into the buried SiO_2 layer. So after successful annealing an SOI layer remains saturated with oxygen atoms up to the solid solubility limit.

Secondly, dislocations are formed during that post-implantation anneal step, having a density of 10^9 per cm^2. The oxygen precipates, however, can be removed (dissolved) during the high-temperature anneal, the dislocations cannot.

Even at temperatures only a few degrees below the melting point of silicon, the dislocation

density hardly diminishes. In such a case it can be said: prevention is better than cure. And research tends in this direction. Several solutions have been indicated recently. Channeling implants at low dose rates drastically reduces implantation damage. Sequential subcritical dose implantation followed by an anneal step in an intermittent way changes the damage profile leading to a low dislocation-content SIMOX. Cavities, probably filled with oxygen, caused by high dose, high beam current implantation into unprotected bare Si wafers, may play a role in defect confining during a high-temperature anneal step. Two MeV high energy oxygen implantations seem to be a good energy range for low-density-dislocation achievements. These four means reduce the dislocation density to 10^5 per cm^2 or less.

SIMOX is one of the promising technologies for 2-dimensional planar SOI. High-dose implantation and high-temperature annealing are processing steps compatible with existing IC clean-room technologies.

SIMOX is a most adapted technology for very thin, (100 - 200 nm) SOI layers, the essential prerequisite for advanced CMOS technologies, notably if SOI layers thinner than the depletion layer are required[7].

SIMOX challenges a dislocation-preventable realization scheme to be established.

Nevertheless, MOS technology applied to high-temperature annealed samples, show Hall mobilities comparable to bulk silicon, an acceptable low density of interface states and low leakage currents.

III. FIPOS (Full Insulation by Porous Oxidized Silicon)

Since a long time it is well known that the anodic dissolution of doped, respectively undoped silicon, or differently doped silicon, in aquous HF solutions can be quite different. This is the basic phenomenon behind porous silicon, the starting material of FIPOS. Under influence of an electric field, p-doped silicon (e.g.) is very sensitive to HF etching, contrary to n-doped silicon. Curious enough, the HF etching occurs locally in pores (porosive etching) and not isotropically as one would expect from pure electrochemical reasons. Now, if porous silicon, having a high surface to volume ratio below an unetchable silicon layer, is brought about, SOI can be realized by applying a three-step annealing/oxidation process occurring in the pores:

- preoxidation at 300°C in dry oxygen; result: oxygen atoms adhere to the walls of the pores;
- complete oxidation of the pores at 800°C; and
- densification of the obtained silicon oxide at 1050 - 1090°C in wet oxygen, often followed by an annealstep in nitrogen.

In this way islands of SOI have been realized and IC's have been made in them with good electrical properties.

So, FIPOS is certainly a candidate for SOI. However, this is not the final verdict in this matter because porous silicon as a phenomenon is still a subject of research. The buried siliconoxide layer does not have properties of a top quality insulating layer. The SOI islands suffer from stress induced by the buried oxide layer, in some cases even causing dislocations in the SOI. The island surface, so far realized, is rather small, often stripe-type

and not exceeding a length of a few hundreds of microns. The islands, however, are fully isolated. FIPOS awaits further evaluation if not a next invention.

Besides this, wet chemical silicon engineering is not a most adapted clean-room technology. CMOS processes have been applied to test FIPOS material. Good electrical characteristics are obtained, comparable to bulk silicon values.

IV. L-SPEG (Lateral Solid-Phase Epitaxial Growth)

Homoepitaxy is a most perfectly adapted way of doing epitaxy. Lattice constant, dilatation and atomic bonding forces between substrate and "epimaterial" are completely tuned. Only as soon as homoepitaxy is required in a patterned structure of holes in an insulator, problems arise due to sidewall effects. But these difficulties can be overcome.

Moreover, in the case of silicon homoepitaxy, the native oxide layer has to disappear before deposition of amorphous silicon or direct epitaxial growth of silicon occurs. And besides this, the substrate surface has to be cleaned up to the highest levels of chemical cleanliness. But nowadays these painstaking circumstances have been mastered as has been shown in ultra-high vacuum techniques.

Lateral-SPEG reaches much farther than homoepitaxy alone. In such a case one wants to realize homoepitaxy to be laterally grown in amorphous silicon over an insulating layer by starting the lateral growth on top of planarized, homoepitaxially grown seeding spots or lines.

This all together constitutes the basic procedure for lateral-SPEG. It seems so simple but some steps are critical.

L-SPEG is a low-temperature procedure, realization below 650°C is feasible. However, the width of overgrowth in amorphous silicon is limited due to polycrystalline growth arising from nuclei in the not yet overgrown region; order of magnitude of the epitaxially grown stripes: a few microns. Only in heavily doped silicon the extensions can be made much larger but this is not a suitable specification for IC applications.

V. Wafer bonding and wafer thinning

Silicon technology is renowned for its constraints against procedures which contaminate. Most wafer-bonding procedures, e.g. by solder or spin-on glasses fall under this head. However, Van der Waals bonding in its most simple form, does not. This bonding occurs already if two flat polished material planes, having a high surface finish, are brought in "optical" contact. By this simple clean handling, dipole forces are induced into the two surfaces at short distance expelling the layer of air in between. These forces are strong under shear and pressure, soft under pulling. They can be strengthened by an anneal step causing chemical bonding at the interface. This way also silicon wafers can be bonded as well as silicon wafers having on top a siliconoxide layer. In the latter case an insulating layer is buried, the necessary condition for SOI. The next step now is to thin one of the wafers (the active wafer) to the thickness above the buried siliconoxide layer required. This thinning is normally done by introducing an automatic stopper built in into the active wafer. In this respect the etching of highly-doped silicon having on top a lowly-doped epilayer, is well known. Etching stops as soon as the lowly-doped epilayer is reached. However, the highly-

doped substrate has some detrimental effects on the epilayer.

Now recently we suggested a procedure of Van der Waals bonding and thinning without an automatic stopper. In this case two thermally oxidized standard wafers are Van der Waals bonded. The handling wafer of the two is mechanically polished on both sides before bonding up to an "optical precision" concerning flatness and homogeneity of thickness (submicron precision). Mechanical, i.e. standard optical polishing, gives a high geometrical precision but subsurface damage. Standard silicon-wafer polishing is done by tribochemical polishing (Syton polishing = in a slurry of SiO_2 grains in a NaOH solution). This gives a lower geometrical precision but a damage-free (sub)surface. The active wafer now is mechanically polished somewhat above the thickness of the damaged layer plus the required SOI thickness. The final treatment then is a short Syton polishing step to remove the thin subsurface damaged layer. One stops polishing as soon as the required SOI thickness is reached. So doing wafer bonding and thinning is reduced to the most elemental steps of wafer manufacturing. Only the geometrical precisions have to be upgraded.

Wafer bonding and thinning is a technology which requires polishing facilities which are not always available in research environments. This technology has therefore so far received limited attention, although it is compatible with "silicon-wafer-manufacturer" techniques.It could even be envisaged in how far it constitutes a pure wafer-manufacturer technology.

VI. Heteroepitaxial SOI

Heteroepitaxially grown silicon on sapphire, SOS, initiated once SOI but today silicon on an amorphous insulator surpasses SOS by far.

This is understandable as constraints for heteroepitaxial growth are rather severe:

- the lattice contants have to be matched,
- the dilatations as a function of temperature have to be adapted and
- there has to be susceptibility for atomic bonding.

Silicon is very sensitive to active hydrogen and oxygen atoms which implies that only top-level-clean environments are appropriate to this aim. Chemical Vapour Deposition (CVD) and Molecular Beam Epitaxy (MBE) are technologies basically well prepared for it.

Sapphire as substrate is a candidate and SOS is still investigated, even applied for somewhat exotic applications.

Spinel ($MgAl_2O_4$) has been used in a CVD experiment[8] but it seems not to have resulted in a breakthrough.

Zirconia (i.e. ZrO_2, stabilized to a cubic phase by the addition of Y_2O_3) is another candidate[9] which can synthetically be grown up to large diameters. Its lattice constant has somewhat to be adapted to silicon but mainly dilatation of zirconia ($\approx 110 \times 10^{-7}/°C$) differs too much from that of silicon ($\approx 35 \times 10^{-7}/°C$). Therefore a modified, more adapted zirconia crystal might be a worthwhile research subject to this end.

So to say, a well insulating, monocrystalline substrate which has the right chemical and physical properties and which is IC compatible, has still to be found.

3 - ADVENTITIOUS MATTERS AND CONSTRAINTS

First goal in SOI is to realize perfect silicon on top of an insulator. The conditions, however, under which SOI realizations have to occur are rather critical and strict. That is why one has to keep count of a larger number of circumstances which will briefly be discussed below.

I. Adventitious matters

- Silicon technology can be divided into two main streams of activities: CMOS and bipolar. SOI specifications for these two categories diverge. CMOS is expected to become optimized in submicron-thick insulated silicon layers. This condition is not guaranteed by all the methods described.
  Bipolar circuits on the contrary are often scheduled on thicker SOI layers of, say, a few microns thickness. SOI for bipolar circuits, however, seems to be less in demand than for CMOS.
  In special cases bipolar (power) switching can be combined with CMOS-SOI control circuitry (intelligent power) to guarantee electrical insulation and noise immunity.
- Given the technology, small stripes, larger-area SOI, alternating defect-free/defect-containing areas, and full wafer scale SOI can be realized.
- All materials and layers used in SOI have to be IC compatible. If spin-on glasses and BPSG (boro-phosphate-silicate glass) layers are applied, whatever may be the reason, the applicability in an IC production environment is restricted. Carbon contamination of the graphite-strip heater are in many cases not acceptable.
- All SOI processing steps have to be IC compatible. Extreme anneal steps (>1200°C) are not considered to be standard IC achievements.
- In local-area recrystallization processing, scan velocity plays an important role. An industrially acceptable production of throughput has to be feasible.
- In the whole processing scheme two-dimensional or three-dimensional circuit goals constitute basic conditions for the method applied. A proper choice what technology will be the most suited one, has to be made beforehand.
- The necessary energy density is a decisive parameter in the optimization of defect-free SOI. It has to be sufficiently low to prevent an intermixed environmental stack of layers on top of a defect-full substrate.
- In some technologies it is not straightforward to couple the input power homogeneously into the to be recrystallized layer, seen over the whole width of a wafer. For example, laser power input has shown to be sensitive to small thickness variations, probably in the capping layer, whatever may be the cause (mode competition?).
- For the future of SOI it is a necessity to widen the "operating window" of several methods described to become an advisible method for real SOI propositions.
- In some cases it has to be envisaged whether the technology is an IC-environmental technology or a wafer-manufacturer technology (SIMOX; wafer bonding and thinning).

II. Technological, experimental and industrial constraints

As soon as SOI has been realized (first aim) the perfectness of the silicon layer (second aim) is subject of research as well as the "new" conditions of its adjacent layers (third aim).

In this respect not all the technologies described have the same behaviour. Let us take as an example zone-melting recrystallization.

Generally speaking, grain boundaries are the main class of defects encountered. They arise automatically from an inhomogeneously diverging solidifying meltfront. They cannot be cured, they have to be prevented or confined in such a way that they occur only in predetermined areas where their existence is not detrimental for the functioning of the devices. This can be brought about by three technological issues:

- adapted heat transfer to the substrate (heat sink structure)
- adapted reflective and antireflective patterning of the capping layer and
- beam shaping (applies only in some cases of ZMR).

Best results are obtained if the temperature gradient is as small as possible but this condition pushes the substrate temperature to values which are considered to be too high in IC processing (>1300°C).

The crystallographic homogeneity of recrystallized SOI layers is often poor. X-ray analyses and RBS measurements show in the SOI a plurality of crystallographic planes. In other cases of small-angle grain boundaries the axis vectors are in-plane rotated over a small angle (normally smaller than 5°).

The crystallographic inhomogeneity can be resolved by seeding the crystallizing front from the underlying silicon substrate. However, seeding disturbs the heat transfer locally so severely that careful engineering is necessary to obtain an acceptable result.

SOI realized by applying locally a too high energy density gives rise to melting phenomena in the stack of layers. This entails often mass flow in the recrystallizing silicon layer which results in an SOI layer of inhomogeneous thickness.

Structuring of the insulating layer gives rise to a non-flat SOI layer which deforms easily during recrystallization.

In many cases SOI-treated wafers show an unevenly high warpage or bow; up to values of more than 100 µm. If this has happened an enormous amount of plastic deformation has occurred. And plastic deformation in monocrystalline materials is the sure proof of many defects present, often complete networks of dislocations to relieve the stress or strain have arisen.

In brief, SOI engineering entails a complex of phenomena which all have to be within the limits of acceptance of advanced silicon technologies. This all together makes SOI really a harder goal to meet than many researchers have supposed it was.

4. PROSPECTS

SOI in the technological sense is in full pace of exploration towards an applicable addition to existing IC technologies. The table gives concisely a survey of what is technologically

Table: SOI technologies: figures of merit

SOI Technology	Usable SOI area	3-D candidate	SOI perfectness	Particularities
Laser recrystallization	small	yes	entrainment defects	Technology to be applied during IC processing; careful engineering required to prevent device degrading; submicron MOS; low throughput.
Electron-beam recrystallization	small	yes	entrainment defects	Technology to be applied during IC processing; vacuum environment; careful engineering required to prevent device degrading; submicron MOS; low throughput; pseudo line scan.
Graphite-strip heater recrystallization	larger up to full scale	no	carbon contamination; can be made defect free	Technology to be applied in early stage of IC processing; higher throughput; lower defect density; submicron MOS; wafer warpage
Tungsten-halogen lamps recrystallization	larger up to full scale	no	can be made defect free	Technology to be applied in early stage of IC processing; well suited for thicker SOI layers (> 10 μm); higher throughput; wafer warpage.
SOI by O^+ ion implantation *and annealing; SIMOX*	full *wafer*	*no*	dislocations to be pre-*vented; oxide precipi-*tates to be annealed	A technology for wafer suppliers before IC processing; *most adapted to existing silicon technology; submicron MOS.*
Full insulation of porous oxidized SI; FIPOS	small islands fully isolated	no	not dislocation free	A technology with a limited applicability due to the smallness of the islands; a technology before IC processing; submicron and thicker layers.
Lateral solid-phase epitaxial growth; L-SPEG	small stripes	yes	twins	A technology with a limited applicability due to the smallness of the stripes; processing IC compatible; submicron Si-layer.
SOI by wafer bonding and thinning	full wafer	yes	comparable to bulk silicon	Advanced technology of wafer production; thicker SOI layers (1 - 5 μm); technology for wafer suppliers.
Heteroepitaxy	substrate scale	probably yes	so far not defect free *due to lattice mismatch* and dilatation	Technology to be evaluated; IC compatibility dependent on substrate material.

going on and some characteristics belonging to it.

All technologies described are in a state of evaluation; no one so far is settled.

A precondition for applications is that IC aim and SOI technology to be applied, are compatibly adapted.

A first goal to meet in SOI is to realize perfect 2-dimensional SOI in a well" behaving" stack of SOI layers leaving the substrate undamaged.

A second goal might become a 3-dimensional project but this seems seriously harder to realize.

Besides this, it has to be evaluated what SOI technologies are specific wafer-manufacturer technologies and which are not.

Generally speaking, many open questions remain to be solved.

Could one expect that some wafer-manufacturer technologies of today e.g. Syton polishing, will be applied in an IC environment for planarization of half-finished wafers to the benefit of some SOI technologies?

Will high dose ion-implantation facilities for SIMOX wafers once move to an IC environment?

Or will SOI sooner become a serious starting point for multifunctional systems where incompatible technologies, systems and materials of today can be combined in an IC acceptable way?

That is what SOI researchers and IC manufacturers have to establish in the future.

1. M.A. Parker, R. Sinclair, T.W. Sigmon; Appl. Phys. Lett. 47 (1985)626.

2. G.W. Cullen, editor; Single-crystal Silicon on non-single-crystal Insulators; J. Cryst. Growth 63 (1983)429-590.

3. J.M. Poate, J.W. Mayer; Laser Annealing of Semiconductors; Academic Press; New York, 1982.

4. S. Furukawa, editor; Silicon-on-Insulator; its Technology and Applications; D. Reidel Publ. Comp., Dordrecht, 1985.

5. A. Chiang, M.W. Geis, L. Pfeiffer; Semiconductor-on-Insulator and Thin Film Transistor Technology; Materials Research Society Symposia Proceedings, Volume 53; Pittsburgh 1986.

6. S.L. Partridge in: Solid State Devices 1985, Edited by P. Balk and O.G. Folberth; Elsevier Science Publishers B.V.; 1986 - Amsterdam; pp 115.

7. J.P. Colinge in: 1988, European Silicon on Insulator workshop; from material to devices; Meylan, France, March 15th - 17th 1988, Edited by D. Bensahel & G. Bomchil; paper F-06.

8. M. Ihara, Microelectronic Engineering 1 (1983)161.

9. M.H. Manasevit, I. Golecki, L.A. Moudy, J.J. Yang, J.E. Mee; J. Electrochem. Soc. 130 (1983)95.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CMOS-TECHNOLOGY - STATUS, TRENDS AND APPLICATIONS

F. NEPPL and H.-J. PFLEIDERER

Siemens AG, Corporate Research and Development, Microelectronics, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

Abstract

Todays CMOS technology is reviewed with respect to process architecture, acquired CMOS specific device know-how, design and applications. The process and device related problems of further miniaturization are discussed. Trends of innovative device design and technology concepts able to overcome these problems are summarized. Finally examples are given for future CMOS applications.

1. INTRODUCTION

In the last decade CMOS has turned from a specialty technology for battery driven devices and military applications into the main stream technology for high density VLSI circuits. By now it has widely replaced the earlier dominating NMOS-technology for new designs. Also world wide CMOS sales now exceed NMOS sales significantly. Mainly two CMOS advantages are responsible for this development: low power dissipation and high noise margin. The significantly lower power dissipation improves power/delay product and allows higher packing density and chip complexity without the penalty of excessive chip temperature or cooling cost. The increased noise margin makes CMOS circuits more resistant to variations of supply voltage, temperature and process variations. Additionally the importance of the initial CMOS inherent disadvantages of higher process complexity and larger chip size have continuously decreased: CMOS process design became smarter and simultaneously NMOS technology had to implement more sophisticated circuit techniques (dynamic NMOS) and processes to reduce power dissipation to an acceptable level.

This paper summarizes the current status of CMOS technology and its applications, discusses the problems of further scaling and circuit integration and outlines trends and extensions to overcome these problems.

2. CURRENT STATUS OF CMOS TECHNOLOGY

2.1 PROCESS ARCHITECTURE

Most of todays modern CMOS-processes have 1.0µm design rules at least in some critical levels like gate, contacts and metal 1. This requires 5:1 or 10:1 optical stepper lithography.
A typical process concept contains epitaxy on highly doped substrate and twin tub for sufficient latch-up resistance, LOCOS isolation, about 20 nm of gate oxide, n^+-poly- or polycide-gates, LDD for hot carrier resistant n-

channel devices, reflowed isolation oxide for pre-planarization, tapered contact holes and diffusion barriers for low resistive and highly reliable contacts to Si and double level metallization for high packing density.
The importance of the question whether n-well or p-well is the appropriate choice for main stream CMOS has been largely reduced by true twin tub concepts that allow to switch from n-well to p-well concept by simply changing the substrate from p- to n-type without changing the process sequence and implantation doses and hence the device parameters.
Modern CMOS processes have 10 to 12 photolithography steps including double level metallization and passivation. This is significantly lower than initial conservative CMOS approaches with about 15 masks and is close to the 10 masks required by a competitive NMOS process with buried contact and depletion transistors. Examples of process simplifications are self-aligned twin tub generation /1/, elimination of field and channel implant masks by appropriate adjustment of the well concentrations /2/, elimination of the LDD implant mask by new sidewall spacer techniques /3/ and selective oxidation of highly doped regions to protect these regions against subsequent S/D implantations in a self-aligned way.
For process and device optimization extensive use is made of process and device modeling to accelerate development and to reduce the expense of experimental wafer processing.

2.2 CMOS SPECIFIC DEVICE KNOW-HOW

Much of the present status of CMOS technology is the result of intensive research in general MOS technology like gate oxide quality, hot carrier degradation and interconnect systems. In this section only the CMOS specific areas of latch-up and the compensated p-channel transistor will be reviewed.
The CMOS-inherent latch-up effect /4/ is the possibility of triggering the parasitic n^+pnp^+-SCRs present in every CMOS inverter. A lumped element model of the situation is given in Fig. 1. Triggering of the SCR is possible if the product of the current gains of the two coupled parasitic bipolar transistors involved exceeds 1. It was demonstrated, however, that for $\beta_{lat} \times \beta_{vert} > 1$ lower β-product does not necessarily imply higher latch-up resistance (see Fig. 2). Particularly for high β_{vert} the critical or trigger current is mainly determined by the shunt resistances involved and becomes independent of the β-product /5/.
As is seen in Fig. 2 epi on highly doped substrate increases the critical current by more than an order of magnitude. This is a consequence of the drastical reduction of the substrate resistance. The characteristics of the MOS devices at the surface are not affected since the resistivity of the thin epi layer itself is still high, i.e. the doping level is low. Generally, thinner epi makes shunting more effective. If the epi layer is too thin, however, the highly doped substrate may partly compensate the well doping and thus favor emitter/collector punch-through of the vertical bipolar transistor.

Fig.1

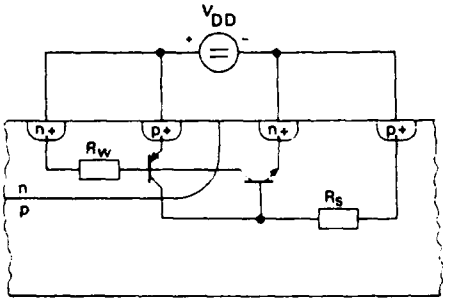


Fig.2

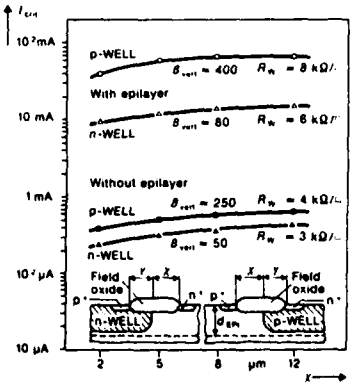


Fig. 1: Lumped element model of parasitic SCR in n-well CMOS with R_s = substrate resistance and R_w = well resistance.

Fig. 2: Critical current of the parasitic SCR versus base width x of the parasitic lateral bipolar transistor for different technologies /15/.

Low well resistance also increases the critical current. Aside of the reduced shunt resistance, the corresponding higher doping level additionally degrades the current gain of the vertical bipolar transistor. In standard well constructions, however, excessive well doping increases body factor and junction capacitances and reduces the junction breakdown voltage of the MOS devices in the well and thus deteriorates circuit performance.
This shows, that doping levels, epi thickness and well depths have to be adjusted carefully to obtain optimum overall device characteristics.
In addition to these bulk related effects, the parasitic field MOSFETS at the surface were shown to influence the turn-on of the two parasitic bipolar transistors and hence reduce the critical current and the latch-up resistance /6/. Finally, also dynamic triggering of latch-up during power-on and noise pulses could be explained by a combined effort of experiments /7/ and simulations /8/.
Initially, the two parasitic bipolar transistors were simply decoupled by separating n- and p-channel MOS devices by 20µm and more. This was a significant limitation for shrinking CMOS circuits. Based on the understanding of the essential parameters and their interdependences, even with conventional approaches careful process design currently allows minimum n^+/p^+-spacings of 5 - 6µm without latch-up risk. This implies that pad cells with drivers and ESD protection circuits, which are particularly effective current injecters, are specially separated from the adjacent logic by guardrings.
The second CMOS specific device problem is the understanding and design optimization of the "normally off" buried p-channel transistor, which results when in case of n^+-poly gates (Φ_{MS} = -0.3 V) the threshold voltage is adjusted to a reasonable value around -0.8 V. Of particular concern are subthreshold behavior or punch-through /9/ and $U_T(L)$ /10/ of this MOSFET type.

At the compensated surface, the total space charge region underneath the gate is a superposition of the depletion zone of the channel junction and of the gate induced depletion zone leading to a minimum of potential energy for holes (Fig. 3). The energy difference between this minimum und the hole source potential is the potential barrier ΔV the holes have to surmount in the sub-threshold region. Since the minimum of the potential barrier is located away from the surface (Fig. 3), the gate voltage control of ΔV is relatively weak. This results in a relatively high sub-threshold swing and makes the buried p-channel transistors particularly prone to punch-through. The situation is discussed in detail in /9/. Figs. 4 + 5 summarize the essential results of this study in terms of the potential barrier ΔV. Fig 4 demonstrates how the drain voltage control of the barrier and hence the susceptability to punch-through is reduced when the gate oxide thickness is reduced. Fig. 5 shows the decrease of ΔV with increasing channel junction depth d and surface p-concentration N_A. From the variety of dependences, guidelines for the design of punch-through resistant buried p-channel transistors down to the sub-micron regime could be derived: low gate oxide thickness, high n-well concentration, low surface concentration N_A and low channel junction depth d. The latter, however, has technological limitations

Fig.3

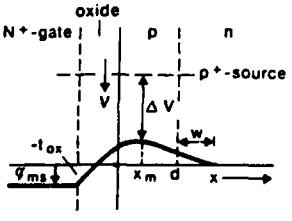


Fig.4

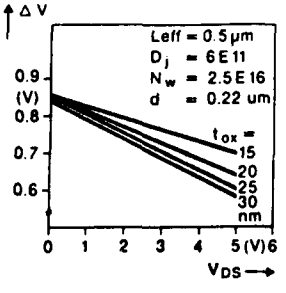


Fig.5

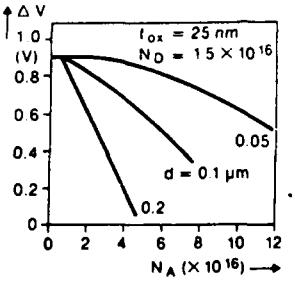


Fig. 3: Band diagram of buried channel structure of the pMOSFET with n^+-poly gate /9/.

Fig. 4: Gate oxide thickness dependence of the drain voltage induced lowering of the potential barrier ΔV for holes at zero gate voltage /9/.

Fig. 5: Influence of channel junction depth d and n-well concentration N_A on the potential barrier ΔV /9/.

in terms of implantation energy and annealing cycle. The doping conditions imply either relatively high threshold voltages or a very high concentration of centers for impurity scattering. Both effects limit the p-channel transistor current.

2.3 CMOS DESIGN

Generally, switching from NMOS to CMOS brought design simplifications because in many cases straight forward CMOS design could replace rather complicated (e.g. dynamic) NMOS logic. In particular, there is no permanent concern about dimensioning the load and drive transistors because there is no trade-off between power consumption and speed like in NMOS. CMOS, instead, widely allows the use of standard transistors. Only a few critical data paths need to be optimized. This design simplification was also the basis for easier design automation and hence the MOS semi-custom approach with gate arrays and standard cells.
One of the main design innovations of CMOS is the simplicity of implementing analog functions.This is demonstrated in Fig. 6 for an operational amplifier. The comparison with an equivalent NMOS version clearly shows the design simplification in case of CMOS. Therefore, CMOS favors the combination of digital signal processing with A/D converters or other analog functions on a single chip. This is important because, even though signal processing now is mostly digital, interfacing with e.g. sensors is mostly analog. Together with the low power consumption these design advantages make CMOS the favorable technology for further system integration.

Fig.6

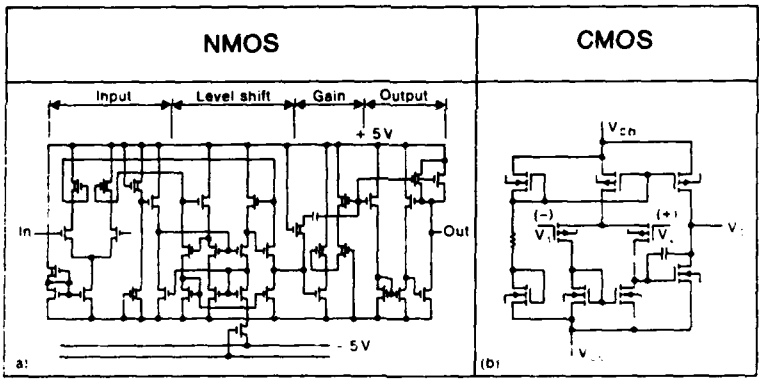


Fig. 6: Comparison of operational amplifier circuits in NMOS and CMOS.

2.4 CMOS APPLICATIONS

Because of the advantages discussed above, CMOS has become the technology of choice for a wide range of applications. This includes microcontrollers, microprocessors, microcomputers, components for signal processing and telecommunication, EPROMs, SRAMs and DRAMs.
Improvements of defect densities, lithography and device know-how as well as the reduction of power density have increased the integration density of CMOS in production to over one million transistors per cm^2. The most prominent representative of this category of chips is the 1 Mbit DRAM now available on the market from several companies. The SIEMENS 1 Mbit DRAM, e.g., has about 2 x 10^6 MOS devices on a chip area of 54.6 mm^2, and a typical access time of 100 ns.
A completely different example is standard-cell logic. This type of chip naturally is not optimized with respect to high density of transistor functions. Instead, its integration density is predominantly determined by the metal pitches. Chip sizes of up to 200 mm^2 with up to several hundred thousand transistors but a dense net of interconnects are currently realized.
Finally, CMOS is increasingly replacing TTL standard logic.

3. THE FUTURE OF CMOS

3.1 TRENDS OF MINIMUM FEATURE SIZE AND INTEGRATION DENSITY

The urge for higher circuit speed, increased system integration and reduced cost per function keeps CMOS driven towards further scaling. Integration complexity is expected to increase from todays 10E6 to 10E7 to over 10E9 components per chip in the year 2000.
Many predictions are based on the future development of DRAMs which are considered to be technology leaders with respect to minimum feature size and integration density and for which in the past a certain regularity of the appearance of new generations was observed. Today, 4 Mbit DRAMs with minimum feature size of 0.8µm are already on their way into pilot line production. Furthermore, 16Mbit DRAMs with minimum feature size of 0.6µm are scheduled for the first half of the nineties and research aims already at 0.3 to 0.4µm feature size for 64 Mbit DRAMs.
First experiments even demonstrate the feasibility of 0.1µm MOSFETs which provide gate delays of 19 psec at room temperature and 13 psec at 77 K /11/. Fig. 7 exhibits the basic device structure and 0.1µm transistor characteristics.

Fig.7

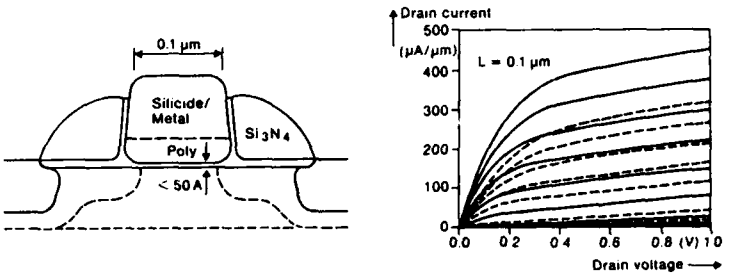


Fig. 7: Cross section and J-V characteristics of 0.1µm nMOSFET at 77 K (-) and 300 K (---) /11/.

This last example shows that MOS still has considerable potential for further scaling. Many process and device related problems need to be solved, however, before such technologies - if ever - can be utilized in a production like environment.

3.2 PROCESSING RELATED PROBLEMS

To fullfill the increasing requirements with respect to defect density, first of all special attention has to be paid to the quality of the starting material which increasingly also affects device reliability.
An indispensable precondition for further scaling of CMOS is the ability to reproducibly generate structures of a few tenths of a micron. Down to ≈0.5µm optical lithography with deep UV in combination with sophisticated resist techniques will dominate. With X-ray, electron beam and ion beam lithography there exist techniques down to less than a tenth of a micron. Electron-beam and ion-beam lithography, however, have a low throughput and are therefore not yet suitable for high volume production. X-ray lithography provides sufficiently high throughput, but might be limited by mask tolerances which are transfered 1:1 onto the chip.
Although reactive ion etching provides excellent line width control, radiation damage during etching might affect the quality of extremely thin gate oxides (≤ 10 nm) and of the Si/SiO_2-interface unacceptably. Reactive ion stream etching might reduce radiation damage sufficiently /12/. Radiation damage during sputter deposition and plasma enhanced deposition is also a concern and needs to be minimized.
Finally more attention has to be paid to device degradation by electrostatic potentials which build up during plasma processing and ion implantation /13/.
Only when solutions to these processing related problems are found, sub-0.5um-technologies will eventually go into production.

3.3 DEVICE RELATED PROBLEMS

Further scaling of CMOS below 1.0µm requires significant innovations in device design.
Sufficiently latch up resistant circuits with n^+/p^+-spacings below 5µm ask for special technological measures: the retrograde well structure /14/ provides the beneficial lower well resistance without affecting the device characteristics at the surface. Its realization, however, needs the not yet established MeV-implantation. Schottky-S/D /15/ or silicided S/D reduce the emitter efficiency of the parasitic bipolar transistors and thus their β-product. Finally, trench isolation of the wells in combination with epi on highly doped substrate was shown to effectively decouple the bipolar transistors because the collector current of the vertical bipolar transistor is shunted to the substrate and does not reach the base of the lateral bipolar transistor /16/. Holding voltages above 10 V can be achieved for n^+/p^+-spacings as low as 2µm if the trench depth comes close to the epi thickness (Fig. 8). Control of the trench side wall isolation is necessary, however, before this technique can be successfully implemented.
For intra-well device isolation, bird's beak free techniques, which additionally provide a higher degree of planarization, have to be adopted. Candidates are BOX /17/ or refill of patterned oxide by selective Si-epitaxy /18/.

Fig.8

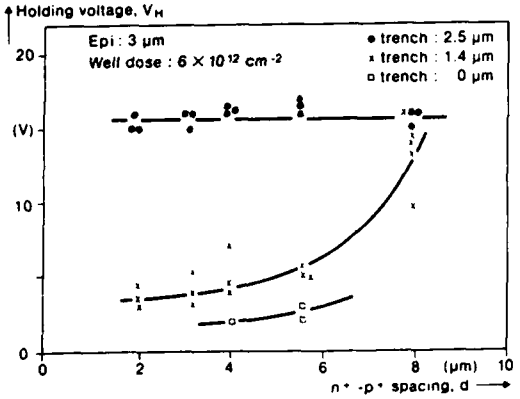


Fig.9

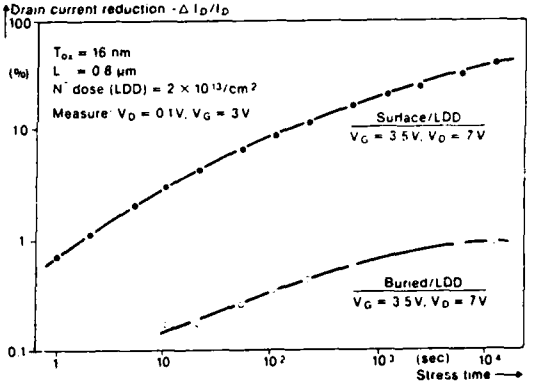


Fig. 8: Holding voltage V_H versus n^+/p^+-spacing d with trench depth as parameter /16/.

Fig. 9: Drain current degradation of surface channel and buried channel nMOSFETs /19/.

Sub-micron CMOS will feature MOSFETS with gate lengths of a few tenths of a micron, gate oxide thickness of a few nm only and source/drain junction depths <0.2µm. These MOSFETs eventually will have to be operated at a supply voltage <5 V.
A crucial problem is the hot carrier degradation of future n-channel MOSFETs. The standard LDD-technique might not be able to guarantee sufficient transistor lifetime. More sophisticated LDD structures like buried LDD provide better reliability. An alternative approach is to adopt the buried channel structure for the n-channel and thus shift the location of maximum hot carrier generation into the bulk and away from the endangered gate oxide. For an acceptably low threshold voltage, this requires a gate material with higher work function than n^+-poly. Fig. 9 shows more than an order of magnitude reduction of the degradation rate for buried n-channels /19/. It has to be noted, that there is a trade-off with respect to the short channel characteristics. Technologically, however, the buried n-channel is easier to control than the p-channel, because As instead of the fast diffusing B is used for surface compensation.

Also the p-channel MOSFETs become increasingly prone to hot carrier degradation and will require LDD techniques. The short channel characteristics of the p-MOSFETs can be improved by changing the work function of the gate material to higher values, i.e. from n+-poly to silicide, metal or p+-polycide gate. Thereby the p-channel type is changed from buried to surface type. Fig. 10 demonstrates the improvement of the short channel characteristics in terms of the sub-threshold swing obtained for silicide and p+-polycide gate, respectively /20/. Furthermore, U_T of the surface channel transistor with p+-polycide gate can be chosen low (e.g. 0.6 V) without a trade-off with respect to punch-through. This results in a 15-20% higher current drive than for the standard buried p-channel device.
These examples show that the gate work function can improve the characteristics of both types of MOSFETs and therefore might become a significant parameter for optimization of sub-micron devices.
SALICIDE will be widely used to minimize parasitic series resistances of MOSFETs. Most techniques use a uniform metal deposition (e.g. Ti, Co, Pt) with subsequent selective silicide reaction with underlying Si /21/. Recently selective CVD of silicide with the advantage of minimum Si consumption and planarization capability has also been demonstrated for this application /22/. In combination with the "strap" technique /23/ SALICIDE allows a significant reduction of the junction capacitances, adds additional wiring flexibility and offers the possibility of direct contacts between source/drain and the gate level for both transistor types, analog to the buried contact in NMOS. Therefore, SALICIDE is one of the key process innovations for improved packing density and speed.

Fig.10

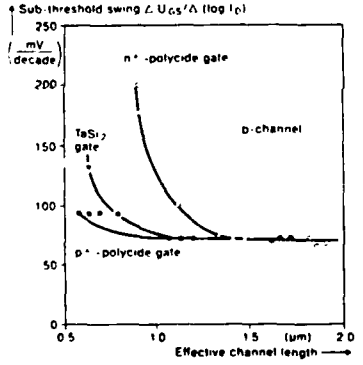


Fig.11

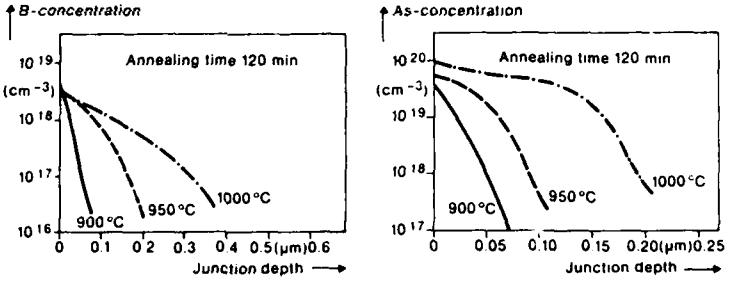


Fig. 10: Subthreshold swing of pMOSFETs with different gate electrodes. Channel implants were adjusted for constant threshold voltage /20/.

Fig. 11: SIMS profiles of As and B doped source/drain regions fabricated by outdiffusion of implanted $TaSi_2$ /25/.

Junction depths ≤0.2µm can be obtained by rapid thermal annealing and/or preamorphization of Si by Si or Ge implantation /24/. The fabrication of extremely shallow junctions ≤ 0.1µm was demonstrated by outdiffusion of implanted silicide /25/. Fig. 11 shows examples for boron and arsenic.
Scaling below 1.0µm raises particular problems with respect to the interconnect system. For non-scaled thickness, interconnect lifetime was estimated to decrease with the fifth power of the scaling factor /26/ which makes electromigration failures an increasingly important issue. Since wiring capacitances no longer scale with line width and are dominated by the metal thickness dependent fringing capacitances, reducing the metal thickness is essential to reduce the total capacitive load. This makes the electromigration problem even worse. In sub-micron metallizations, stress migration was demonstrated to become an additional failure mode. Its importance increases with decreasing line width /27/. Consequently the introduction of a metal "stronger" than aluminum is mandatory.

As interconnect cross sections decrease and average interconnect lengths rather increase, interconnect resistivity is no longer negligible with respect to voltage drops and RC delays. Therefore particularly low resistive interconnect materials are needed.
Al/Ti multisandwich structures /28/ provide a promising compromise between both requirements. Tungsten may be restricted to short and medium range interconnects because of its relatively high resistivity. High temperature superconductors were largely eliminated as a possible solution to the interconnect problem in CMOS technology /29/. One of the reasons is that at the required temperatures of $\leq$ 80 K the established Al-technology again fullfills all requirements.
Contact and via filling e.g. by CVD-tungsten will be generally used to avoid metal step coverage problems in small contacts. The standard planarization of the inter metal dielectric including PECVD-oxide, polyimide or spin-on-glass auxiliary layers and etch-back might continue to dominate for a while because reduced metal thickness and superconformal PECVD-oxides alleviate the actual problems.
Deep sub-micron optical lithography for fine line interconnects, however, requires additional efforts with respect to global planarization to account for the limited depth of focus of high resolution lenses.

3.4 FURTHER PERSPECTIVES OF CMOS

With the device innovations discussed above, standard CMOS has considerable potential for further scaling. In addition, however, conceptual innovations are discussed to drive CMOS towards higher performance and packing density.
For higher speed, operating CMOS at 77 K is often discussed. A summary of performance improvements, necessary device optimizations and remaining problems is given e.g. in a special issue of IEEE Trans. Electron Devices /30/. Advantages include high transistor current, extremely low latch up risk, low metal resistance and better reliability with respect to thermally activated degradation mechanisms. This, e.g., dramatically improves metallization reliability. The significantly increased hot carrier induced degradation at low temperatures, however, requires a reduced supply voltage. Therefore, the possible speed advantage is partly lost again. An overall increase of circuit speed of a factor of 1.5 to 2 over room temperature CMOS is predicted. Cost and inconvenience of cooling, however, might prevent this approach to be generally adopted.

SOI structures on Si wafer substrates are of increasing interest. They avoid the substrate cost penalty of earlier SOS and maintain the device and circuit advantages of this technology, i.e. low capacitances, low sub-threshold swing,higher packing density and radiation hardness. Technologies used are oxygen implantation of Si (SIMOX), oxidation of porous Si, recrystallization of poly-Si /31/ and lateral growth of Si-epitaxy /32/. The applicability of these techniques in production of highly integrated circuits has to be proven yet.

BICMOS technology combines the low power dissipation of CMOS with the high speed and the inherent analog advantages of bipolar devices. From a processing point of view, there is no reservation to integrate both technologies in one line, since many process steps can be used simultaneously for CMOS and bipolar, anyway, and bipolar specific process steps can be inserted as modules. High speed SRAMs are the field of application for BICMOS which offers high performance bipolar transistors. Since CMOS and bipolar components have to be optimized simultaneously, process complexity is rather high and the number of masks may increase to 20. An alternative approach uses mostly CMOS process steps for the fabrication of the bipolar transistors and requires only 1 to 3 additional masks. Nevertheless, this type of BICMOS also provides significantly better current drive (and hence speed) and easier implementation of analog functions than pure CMOS. Such BICMOS with at most moderate extra cost might be a candidate for replacing pure CMOS as a main stream technology in the future.
To account for the increasing development cost and for the simultaneous need for application specific technologies, process architecture is generally forced towards a modular scheme with one basic CMOS process that can be supplemented by certain application specific modules. BICMOS is an outstanding example. But also process optimization with respect to e.g. DRAMs, SRAMs or E^2PROMs will increasingly follow this concept.

3.5 DESIGN CONSIDERATIONS AND APPLICATIONS OF FUTURE CMOS

Reduction of supply voltage is often discussed for mitigation of hot carrier problems and of excessive power dissipation of sub-micron chips. 0.5µm technology is considered to be a critical process generation for this change. Interfacing with existing 5 V circuitry can be solved in principle by on-chip voltage reduction. The major concern is the increase of gate delay caused by a V_{DD}-reduction which would have to be overcompensated to make scaling still advantageous with respect to higher speed.
A general trend of applications is towards more system integration to reduce cost, system delay and reliability problems related to the number of packages and pins involved. In view of increasing chip complexity, extensive use of automatic design and shrinkability of optimized components is mandatory.
CMOS chips for small and medium size computer systems have to provide a data processing rate of more than 10 MIPS. CMOS components for broad band ISDN have to meet the 140 or 560 Mbit/sec requirement of this technology. Applications like high definition TV ask for signal processing with up to 50 MHz band width and high density memories with low access time. All these applications need the most progressive technology available and partly also new chip architecture.
The most aggressive requirements might come from the attempt to functionally simulate the human brain by realizing "artificial neuron nets" for extensive parallel and associative processing. This concept is advantageous e.g. for fast recognition of image and speech patterns. Even simple implementations with 1oo neurons require the packing density of 1.0µm technology, already. More useful systems with 10^4 neurons would need 0.1µm technology and thus the most progressive MOS technology demonstrated so far.

4. CONCLUSION

This summary shows that CMOS is a main stream technology with considerable potential of technological development and applications. Since one approaches the fundamental physical limits, however, the cost of progress is increasing significantly. Particularly the step from feasibility to production seems to become increasingly difficult, time consuming and expensive. On the other hand, speed improvement by scaling is becoming less than predicted by simple scaling. Consequently the price for a given improvement of performance is rising dramatically and is even about to exceed the possibilities of individual companies. To push deep sub-micron CMOS technology nevertheless, government-funded cooperations like SEMATEC in USA or JESSI in Europe have been started. These efforts show that CMOS is considered a key technology for further technological and economical development.

REFERENCES

/1/ L.C. Parillo et al., IEDM Technial Digest (1980) p. 752
/2/ M.L. Chen et al., IEDM Technical Digest (1986) p. 256
/3/ L.C. Parillo et al., IEDM Technical Digest (1986) p. 244
/4/ R.R. Troutman, "Latch up in CMOS Technology", Kluwer Academic Publishers, Boston (1986)
/5/ D. Takacs et al., IEDM Technical Digest (1983) p. 159
/6/ D. Takacs et al., IEEE Trans. Electron Devices, Vol. ED-31, No. 3 (1984) p. 279
/7/ W. Reczek et al., ESSDERC 1988
/8/ M. Strzempa-Depre, IEEE Trans. Electron Devices, ED, to be published
/9/ F.M. Klaassen et al., IEDM Technical Digest (1984) p. 613
/10/ J.S.T. Huang et al., IEEE Trans. Electron Devices, Vol. ED-31, No. 12, (1984) p. 1889
/11/ G.A. Sai-Halasz et al., IEDM Technical Digest (1987) p. 397
/12/ C. Takahashi et al., 1988 Symposium on VLSI Technology, Digest of Technical Papers, p. 83

/13/ K. Tsunokuni et al., 19th Conf. on Solid State Devices and Materials, Tokyo, Extended Abstracts (1987) p. 195
/14/ R.D. Rung et al., IEEE Trans. Electron Devices, ED-28, (1981) p. 1115
/15/ M. Sugino et al., IEEE Trans. Electron Devices, ED-30, (1983) p. 110
/16/ Y. Niiton et al., IEDM Technical Digest (1985) p. 509
/17/ G. Fuse et al., IEDM Technical Digest (1987) p. 732
/18/ J.O. Borland, IEDM Technical Digest (1987) p. 12
/19/ M. Nakahara et al., IEDM Technical Digest (1985) p. 238
/20/ H. Gierisch et al., ESSDERC, Europhysics Conf. Abstracts (1985), Vol. 9H, p. 40
/21/ K. Ohyu et al., 19th Conf. on Solid State Devices and Materials, Toyko, Extended Abstracts (1987) p. 411
/22/ K. Hieber et al., Thin Solid Films, 140 (1986) p. 131
/23/ D.C. Chen et al., IEDM Technical Digest (1984) p. 118
/24/ K. Yamada et al., Jpn. J. Appl. Phys., 22-1 (1983) p. 157
/25/ H. Gierisch et al., J. Vac. Sci. Tech, B5 (1987) p. 508
/26/ M.H. Woods et al., IEDM Technical Digest (1984) p. 50
/27/ S. Magumi et al., 25th Annual Proceedings Reliability Physics Symposium (1987) p. 15
/28/ P. Kücher et al., Proc. Int. Symp. on Trends and New Applications in Thin Films, Vol. 1 (1987) p. 137
/29/ H. Föll, Mikroelektronik, No. 4 (1988) to be published
/30/ IEEE Trans. Electron Devices, ED-34, No. 1
/31/ S.L. Partridge, IEDM Technical Digest (1986) p. 428
/32/ T. Kubata et al., IEDM Techical Digest (1987) p. 344

Session 1A Room A : M.O.S.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A HIGH PERFORMANCE LIQUID-NITROGEN CMOS SRAM TECHNOLOGY

J.Y.-C. SUN, S. KLEPNER, Y. TAUR, H. HANAFI, P. RESTLE, T. BUCELOT, K. PETRILLO, R. DENNARD, S. SCHUSTER, T. CHAPPELL, B. CHAPPELL and D. HEIDEL

IBM Research Division, Thomas J. Watson Research Center, PO Box 218, Yorktown Heights, NY 10598, U.S.A.

Abstract - A 3.5 ns ECL-compatible 64Kb liquid-nitrogen CMOS (LN-CMOS) SRAM technology with 2.5V power-supply voltage is described. Key features of this high performance 0.5μm-channel LN-CMOS SRAM technology optimized for 77K operation include 0.6μm optical lithography for the gate level, dual polysilicon work functions, retrograde n-well, low resistance arsenic and boron source/drain diffusions, self-aligned titanium silicide, and two-level metal interconnects. For the first time, the leverage of liquid nitrogen CMOS with 2.3X chip level performance improvement at 77K over room temperature CMOS is demonstrated.

1 - INTRODUCTION

Low temperature or liquid-nitrogen temperature operation of CMOS VLSI circuits becomes increasingly attractive as the channel length is scaled down to 0.5μm and below /1/. Liquid-nitrogen CMOS (LN-CMOS) offers speed, density, and reliability for high performance VLSI logic systems. The ultimate potential of CMOS at 77K cannot be realized by straight-forward cooling of a CMOS chip with devices designed for room temperature operation, especially in the sub 0.5 μm channel length regime and in CMOS chips with n^+-polysilicon gates for both types of FETs. The concept of LN-CMOS device design and process optimization for best performance, power, and reliability has been reported /1/ with roughly a factor of two ring oscillator speed improvement. In this paper, a high performance 0.5μm LN-CMOS SRAM technology with n^+ and p^+ polysilicon gates, self-aligned silicide, two layers of metal wiring and buried contacts will be described. The fabrication, testing, and device and circuit performance of a 64Kb SRAM chip will be discussed.

2 - CHIP FABRICATION AND OPTIMIZATION

Key features essential to this optimized LN-CMOS SRAM device design and fabrication process include dual work function (n^+ and p^+) polysilicon gates for surface channel FETs, abrupt n^+ and p^+ source and drain junctions for low device series resistance, deep n-channel threshold-adjust implant for higher carrier mobility, and reduced power-supply voltage (2.5 V) for better performance and reliability at lower power as proposed in /1/. The schematic of the 0.5μm LN-CMOS device cross section is shown in Fig. 1, indicating the p/p^+ epi, retrograde n-well, semi-recessed oxide isolation, dual work function polysilicon gates, 12.5 nm gate oxide, 0.15-0.2μm source/drain junction depth, self-aligned silicide, nitride passivation, polyimide, and two levels of aluminum interconnects. In this 64Kb SRAM, the second level aluminum is used to reduce the word line resistance.

The ECL-compatible CMOS SRAM chip was designed with a conservative 1μm ground rules and 0.5μm effective channel length. The 0.6μm polysilicon gate was defined optically with excellent line width control and exposure latitude by an I-line stepper using a contrast-enhanced layer over standard resist. Fig. 2 shows 0.5μm resist lines and spaces with vertical profiles after developing. Excellent resolution (0.5μm) and line width bias ($\leq$ 0.1μm) are achieved across a 1 cm^2 field.

After the reactive ion etching of the polysilicon gate, a spacer consisting of thermal oxide and thin nitride was formed on the gate edge before source/drain ion implantation. Fig. 3 shows the SEM cross section photograph of the polysilicon gate, spacer, and self-aligned silicide in the SRAM memory array area. This spacer, with a slight oxide tapering at the gate edge, played a key role in (i) reducing the gate-drain overlap capacitance ($C_{gd} \leq 0.3fF/\mu m$), (ii) preventing the silicide from bridging the gate and the source/drain dif-

fusion, and (iii) improving the reliability by reducing the electric field at the gate edge. After the spacer was formed, degenerate doping of the n^+ and p^+ polysilicon gates and low resistance shallow n^+ and p^+ junctions were achieved simultaneously with the As and B ion implantation and drive-in steps, respectively. No extra mask steps are necessary for this dual work function polysilicon gate process.

The 64Kb SRAM circuit was originally designed for room temperature operation as described elsewhere /2/. It uses a four-transistor SRAM cell with buried contacts. The chip incorporates CMOS asynchronous receivers capable of interfacing low voltage ECL signal levels at high speeds with good tolerance to parametric variations. No special circuit optimization was made specifically for liquid-nitrogen temperature operation. SRAM chips were also fabricated with an optimized 0.5μm room temperature n-poly, n-well CMOS technology (RT-CMOS) /3,2/ and compared with the LN-CMOS chips.

3 - RESULTS

An automated full-wafer probe station was developed and used for all device parametric and chip functional tests at 77K and above. The probe station with a liquid-nitrogen cooled chuck is enclosed in a nitrogen filled metal box to avoid condensation on the wafer.

Fig. 4 shows the measured threshold voltage at V_{DS} = 50 mV for n-channel FETs at room temperature and 77K from this LN-CMOS technology. The roll-off is slightly worse at 77K than at 300K, but the excellent sub-threshold turn off slope (roughly 22 mV/dec) of LN-CMOS devices allows the threshold voltage and power supply voltage to be reduced for higher performance and reliability. The shift in threshold voltage from room temperature to 77K is roughly 200 mV. Similar results were obtained for p-channel FETs. Table I summarizes the measured device characteristics and parametric data of LN-CMOS at 77K as well as those of the RT-CMOS at 300K. The RT-CMOS transistors have n^+-poly gates. The threshold voltages of RT-CMOS p-channel FETs are -0.7V and -1.3V at 300K and 77K, respectively. The large shift is due to carrier freeze-out in the buried-channel p-type device, which is avoided in the optimized LN-CMOS by making both n and p-channel transistors surface-channel mode with the dual polysilicon approach. The n and p-channel threshold voltages of LN-CMOS devices are symmetrical, independent of temperature. By carefully controlling the annealing process after the p-type polysilicon is formed on p-channel devices, there was no boron penetration from the polysilicon through the thin gate oxide to the channel region. Note that the transconductance of n-channel transistors improves more upon cooling compared with p-channel transistors, which should be taken into account to optimize the circuit performance.

The LN-CMOS transistors are not completely turned off at room temperature, which may present some problems for circuits with dynamic nodes to operate adequately during room temperature testing with enough noise margin. However, there is a one-to-one correlation of device characteristics between room temperature and 77K. Most parametric tests can be done at room temperature, which facilitates process monitoring and reduces test cost. On the other hand, circuit and system functional tests may be performed at a reduced power-supply voltage and below room temperature, e.g., -55C, as discussed later.

A device model suitable for circuit simulation was developed and calibrated with measured hardware data. It takes into account mobility, resistance, saturation velocity, and operating temperature. Using the model we computed the nominal delay per stage as a function of capacitance loading at 77K for a typical 3-way NAND circuit (fan -in = fan - out = 3) with 0.5μm LN-CMOS transistors. The results are shown in Fig. 5. Also shown for comparison is the data at 22C and 85C for the 0.5μm RT-CMOS technology which is optimized for room temperature operation at 3.3V as previously described in Table I. With 15μm wide CMOS transistors, the slope (loading sensitivity) is 0.54ns/pF for the LN-CMOS operated at 77K and 2.5V. In contrast, the loading sensitivities are 1.13ns/pF and 1.32ns/pF for RT-CMOS at 22C and 85C, respectively. The circuit performance of the RT-CMOS NAND circuit degrades by 17% from 22C to 85C which is the normal chip temperature for room temperature systems. The improvement factor between RT-CMOS at 85C and LN-CMOS at 77K is 2.43X, which comes from the higher transconductance of optimized LN-CMOS devices and the reduced voltage swing.

The access time of this LN-CMOS 64Kb SRAM is 3.5ns at 2.5V and 77K, representing the fastest SRAM chip at this density. Fig. 6 shows that there is little improvement in access time when the power supply is increased above 2.5V, due to velocity saturation and device series resistance. Also shown in Fig. 6 is the access time of the RT-CMOS SRAM chip optimized for room temperature operation. A factor of 1.4X or 1.77X speed improvement is observed between 27C and 77K or 85C and 77K, respectively, if the RT-CMOS chip is cooled to 77K and operated at 3.3V. However, it must be emphasized that the power-supply voltage for RT-CMOS RAM chip must be reduced in order to maintain adequate hot carrier reliability at 77K /4/. On an equal or comparable hot carrier reliability basis, the supply voltage for 77K operation of the RT-CMOS RAM chip should be 2.6V instead of 3.3V. In Fig. 6, the nominal supply voltage with comparable hot carrier reliability is marked by an arrow for each technology and operating temperature. Therefore under comparable hot carrier reliability constraints, only 1.4X access time improvement is achieved by operating the RT-CMOS chip (optimized for RT) at 77K, much less than the

2.3X gain realized by the optimized LN-CMOS. Our data in Fig. 6 clearly demonstrated the advantages of device and process optimization for LN-CMOS to achieve the best speed and power-delay product while maintaining appropriate reliability at liquid-nitrogen temperature.

The RAM access time decreases monotonically as temperature is lowered, as shown in Fig. 7. The access time increases by roughly 20% as channel length is increased from 0.55μm to 0.75μm. The temperature sensitivity of the LN-CMOS RAM access time is 8.5ps/K between 75K and 130K, and roughly 13ps/K near room temperature, respectively. In comparison, the temperature sensitivity of the RT-CMOS RAM access time is roughly 20ps/K between 25C and 85C. This lower temperature sensitivity is one of the major advantages of LN-CMOS. Combining the more stable temperature due to liquid-nitrogen cooling and the smaller temperature sensitivity, LN-CMOS systems tend to have better overall system performance.

Also shown in Fig. 7 is the temperature dependence of the chip power. The chip power decreases as temperature is increased from 77K as a result of less current and lower speed, which is consistent with the temperature dependence of device transconductance. It reaches a minimum near 225K and than increases with temperature afterwards. This increase in chip power at T>250K is due to dc power from higher sub-threshold current. The threshold voltage is too low to have the device turned off completely at $V_{GS} = 0$. The chip is also operational at room temperature, although with reduced noise margin and higher power. Although simple functional tests can be done at room temperature, from noise margin, chip power and testing cost point of view, it may be desirable to do complete functional tests of the chip at -55C.

4 - CONCLUSION

A 3.5 ns ECL-compatible 64Kb liquid-nitrogen CMOS(LN-CMOS) SRAM technology with 2.5 V power-supply voltage is reported. A factor of 2.3X improvement in access time of this LN-CMOS SRAM chip over its room temperature counterpart (RT-CMOS, 3.3V, 85C) was achieved by a dual work function (n^+ and p^+) polysilicon process optimized for 77K operation. Key features of this high performance 0.5μm-channel LN-CMOS SRAM technology include 0.6μm optical lithography for the gate level, dual polysilicon work functions, retrograde n-well, low resistance arsenic and boron source/drain diffusions, self-aligned titanium silicide, and two-level metal interconnects. The device design, chip fabrication, and low temperature testing issues have been discussed. This LN-CMOS SRAM is the fastest RAM chip at the 64Kb level of integration for any technology.

REFERENCES

/1/ J.Y.-C. Sun, *et al.*, *IEEE Trans. Electron. Dev.*, **ED-34**, (1987) 19.
/2/ T.I. Chappell, *et al.*, 1988 *VLSI Circuit Symposium*, (1988).
/3/ L.K. Wang, *et al.*, *Proc. Symp. VLSI Tech.* (1986) 13.
/4/ T.H. Ning, *et al.*, *IEEE Trans. Electron. Dev.*, **ED-26**, (1979) 346.

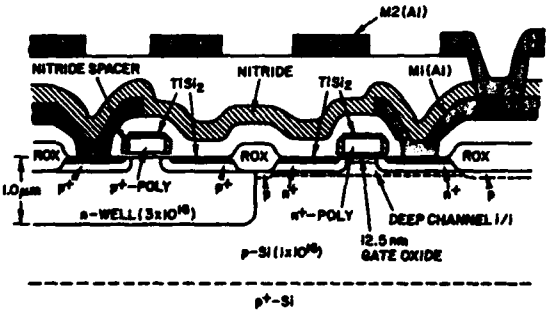


Fig. 1: Schematics of LN-CMOS device cross section

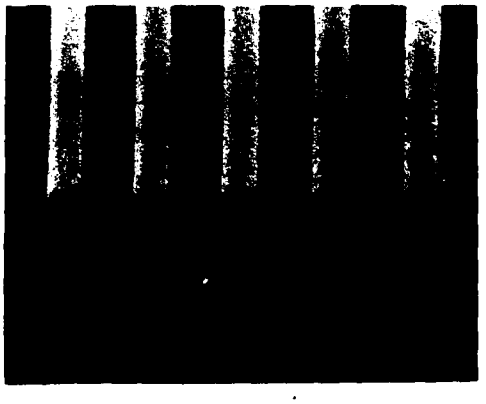


Fig. 2: SEM photograph of 0.5μm resist lines and spaces exposed and developed with a contrast enhancement layer.

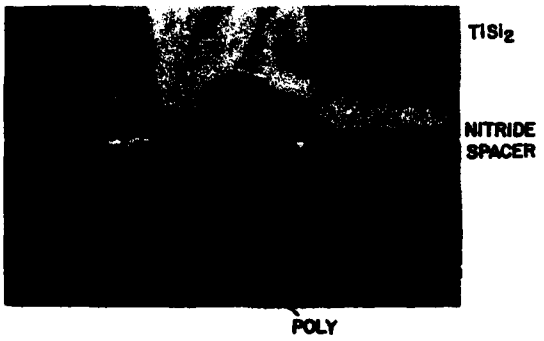


Fig. 3: SEM micrograph of the poly gate, nitride spacer, and salicide in the memory array area.

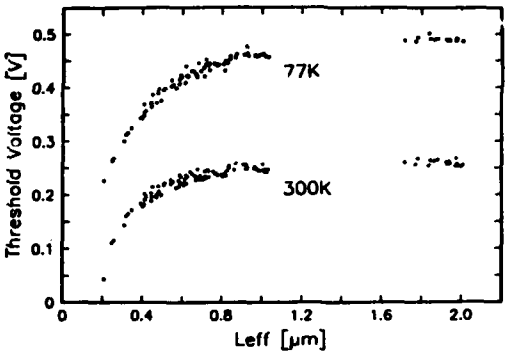


Fig. 4: Threshold voltage vs. channel length at room temperature and 77K for the LN-CMOS n-channel transistor. (V_{DS} = 50 mV).

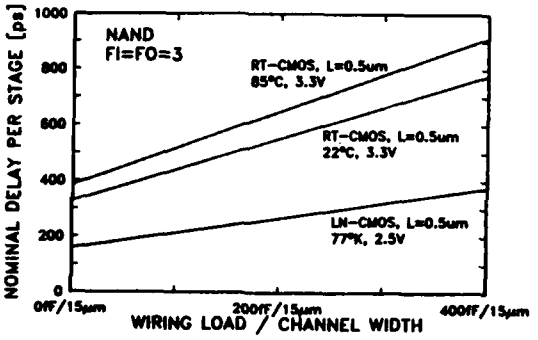


Fig. 5: Simulated and measured stage delay of a 3-way NAND circuit, comparing RT-CMOS (optimized for room temperature operation) and LN-CMOS (optimized for liquid-nitrogen temperature operation).

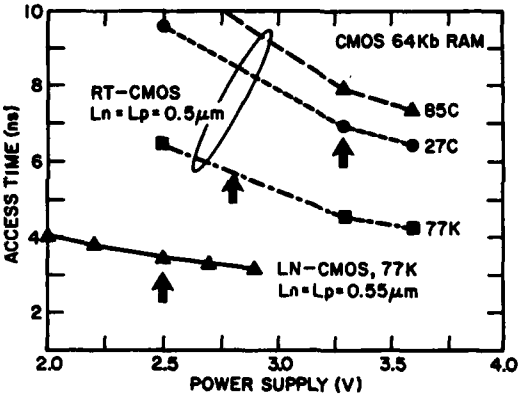


Fig. 6: The RAM access time vs. power supply voltage. Arrows (↑) indicate the nominal supply voltages with comparable hot carrier reliability.

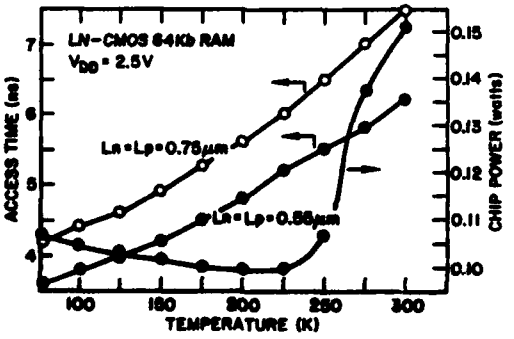


Fig. 7: Dependence of RAM access time and power on temperature and channel length.

TABLE I. Summary of device and parametric data

	LN-CMOS 77K		RT-CMOS 300K	
Parameter	n-ch	p-ch	n-ch	p-ch
G_m (mS/mm)	210	90	120	70
Threshold (V)	0.4	−0.4	0.7	−0.7
S/D Resistance (Ω-μm)	400	1600	800	1600
Peak μ_{FE} (cm^2/V-s)	2700	420	400	140
Turn-Off Slope (mV/dec)	21	22	80	90
$TiSi_2$ R_s (Ω/□)	1	0.9	4	3.5
Power Supply (V)	2.5		3.3	
Gate Oxide (nm)	12.2		12.2	
Channel Length (μm)	0.55		0.5	

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CHARACTERISATION OF NARROW-SPACED ISOLATION IN A TWIN RETROGRADE WELL SUBMICRON CMOS PROCESS

P.A. van der PLAS, P.H.J. SPIJKERS and F.M. KLAASSEN

Philips Research Laboratories, PO Box 80000, NL-5600 JA Eindhoven, The Netherlands

Résumé

Des petites dimensions entre les régions N^+ et P^+ ont été realisé utilisant une technologie de puits jumeaux rétrogrades en combinaison avec une technique d'isolation avancée. Dans cet article, nous avons démontré à l'aide de simulations, et de résultats éxperimentaux, la réduction requise des courants de fuite a l'intérieur des transistors de champs. Les résultats montrent qu'un espace de 2.5 μm séparant les régions N^+ et P^+ peut être réalisé.

Abstract

Small N^+-P^+ spacings have been realised using twin retrograde well technology in combination with an advanced isolation scheme.In this paper suppression of field transistor leakage currents is demonstrated by simulation and experimental results. The results show that an N^+-P^+ spacing of 2.5 μm can be realised without increased narrow-width effects in adjacent transistors.

1 - Introduction

To fully utilize the potential of submicron CMOS transistors, narrow-spaced isolation is needed. To achieve this, bird's beak formation has been reduced by using suppressed LOCOS (SLOCOS) technology [1]. To decraese leakage currents between active MOSFETs, highly doped field transistors are required. However, lateral diffusion of the highly doped channel stopper into the active area has be to avoided to eliminate narrow-width effects. Also lateral diffusion of the N-well has to be reduced in order to achieve small N^+- P^+ spacings.

High energy implantations for simultaneous formation of wells and effective self-aligned channel stoppers have been introduced after field oxide formation [2]. Due to the absence of lateral diffusion narrow-width effects in adjacent MOSFETs have been reduced considerably and a very small N^+-P^+ spacing of 2.5 μm has been realised. Owing to this small spacing a small and fast full CMOS 1-Mbit SRAM has been fabricated successfully [3].

In addition, the low temperature budget of the process enables the use of a thin epi-layer for effective latch-up suppression.

2 - Process summary

The field devices were made in a similar technology as was used for the fabrication of 1M-SRAMs [4]:

(1) 600 nm of field oxide was grown using SLOCOS technology [1].

(2) P^+ was implanted at 600 keV with various doses to form the N-well.
The implant was locally masked by 2.6 μm thick resist.

(3) B^+ was implanted at 400 keV with various doses to form the P-well using an inverse of the N-well mask.

(4) The field oxide was planarised to 550 nm.

(5) Formation of LDD transistors with silicidated poly and source/drain areas.

The field oxide thickness under poly was 520 nm. The depth of the N^+- and P^+-junctions was 0.25 μm and 0.3 μm respectively.

3 - Device simulations

Since the field transistor current is essentially controlled by 2-D potential barrier lowering, the relevant parameters like dopant concentration at the interface and parasitic channel length have been varied using a device simulator [5], which enables simulation of the non-planar LOCOS structure. The field oxide morphology and doping profiles used here have been derived from SEM and 2-D process modelling data. The structure used to investigate narrow N^+-P^+ spacing is given in Fig. 1. In Fig. 2 a three dimensional plot of the doping profiles used is given. Figs. 3 and 4 show the current of the N^+-Nwell and P^+-Pwell field transistors as a function of channel length and doping. For a current level of 10^{-9} A per cm gate width, simulations indicate that a minimum value of 1.5 μm for the N^+-P^+ spacing is feasible. Using conventional well technology, the above value can hardly be reduced below 3.5 μm [6].

4 - Experimental results

The testmodule for measuring the leakage currents from N^+ to Nwell and P^+ to Pwell has alternating polysilicon gates and junctions adjacent to the field oxide edges to make the module worst case with respect to punch-through. In Figs. 5 and 6 measured subtreshold characteristics of different parasitic field transistors are given. Both figures show that a threshold voltage well above 10 V can be realised for small spacings. For the N^+-Nwell and P^+-Pwell field transistors the channel lengths have been varied in steps of 0.25 μm. Typical results have been summarized in Fig. 7 for various implantation doses. Using the previous current criterium, for X_n and X_p minimum values of 0.95 μm (P=5e16) and 0.45 μm (N=8e16) have been found. These values differ by less than 0.1 μm from the simulated values.

Fig. 8 shows the experimental spread for N^+-Nwell field transistors over a wafer for various channel lengths. The figure indicates that for a Pwell concentration of 5e16 B/cm^3 a channel length of 1.25 μm is required, while simulation results (see Fig. 2) indicate that a smaller spacing can be used. The spread is caused by misalignment, variation of implantation angle and shadowing by the high energy implantation mask [7]. Similar results have been found for the P^+-Pwell field transistors. Therefore the simulated value of 1.5 μm for the N^+-P^+ spacing has to increase to 2.5 μm in practice.

Fig. 9 shows the threshold shift of NMOSTs of various widths, demonstrating the absence of narrow width effects due to absence of lateral diffusion of the channelstopper. Furthermore the small influence of the Nwell dose on the body factor of the PMOST transistor, as is shown in Fig. 10, demonstrates the possibility of independent optimisation of field and active MOSFETs.

5 - Conclusions

The results demonstrate that twin retrograde well technology enables excellent transistor performance combined with small spaced field isolation. Simulations indicate that this technology allows to reduce the N^+-P^+ spacing down to 1.5 μm . Experimental results show that process spread requires an increase of the N^+-P^+ spacing to 2.5 μm.

Literature

1 P.A. van der Plas, W.C.E. Snels, A. Stolmeijer, H.J. den Blanken, R. de Werdt
Symposium on VLSI technology 1987, p.19-20, Karuizawa, Japan

2 R.D. Rung, C.J. Dell'oca, L.G. Walker, IEEE Elec. Dev., vol. ed-28, no 10, october 1981

3 F. List, S. Chu, J. Dikken, C. Hartgring, S. Bell, J. Raemaekers, B. Walsh, R. Salters
Proceedings of ISSCC 1988, feb. 1988, page 178-179, San Fransisco (USA)

4 R. de Werdt et. al., Techn. Dig. IEDM 1987, page 532-535, Washington 1987

5 SEMMY, a 2-D general purpose PHILIPS proprietary device simulator

6 F.M. Klaassen, ESSDERC 1986, p.113, Inst. Phys. Conf. Series 82

7 M.G. Pitt, P.A. van der Plas, to be presented at ESSDERC 1988

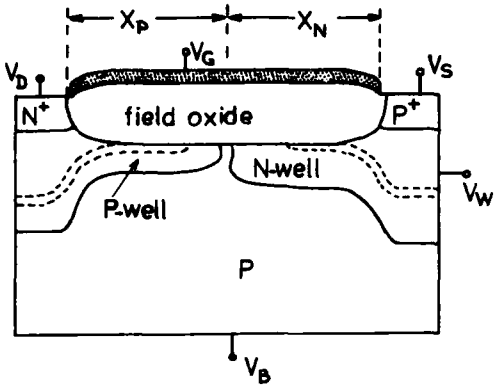


Fig.1 - Cross section of simulated isolation structure. The dashed lines indicate the maximum concentration in the wells. The extension of the well masks over the field oxide edge is called X_n and X_p.

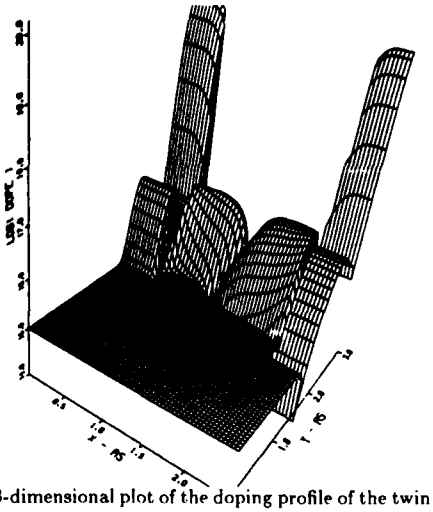


Fig.2 - 3-dimensional plot of the doping profile of the twin retrograde well structure.

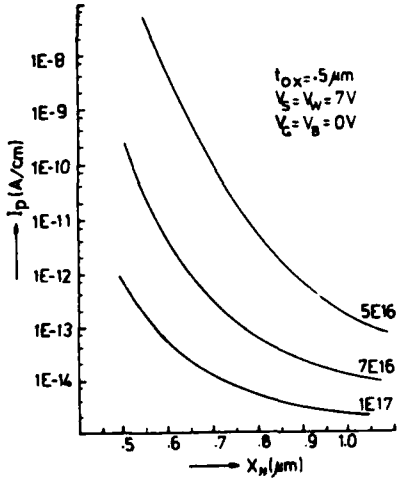


Fig.3 - Current of N^+-Nwell field transistor vs X_p for various doping concentrations.

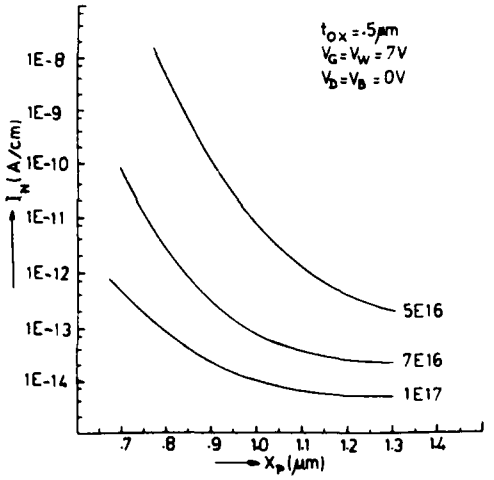


Fig.4 - Current of P^+-Pwell field transistor vs X_n for various doping concentrations.

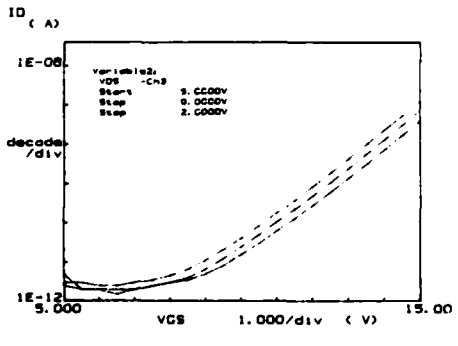


Fig.5 - Measured characteristics of N^+ - *Nwell* field transistor for X_n=1 μm (W = 100 μm).

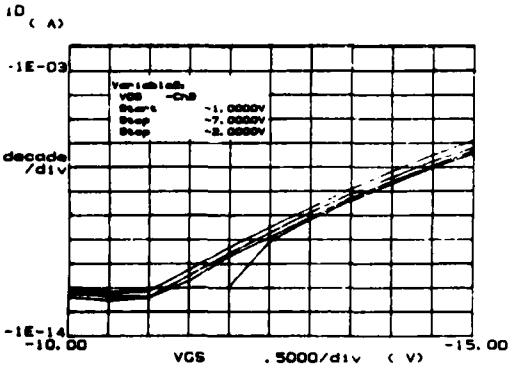


Fig.6 - Measured characteristics of P^+-Pwell field transistor for X_p 0.5 μm (W = 100 μm).

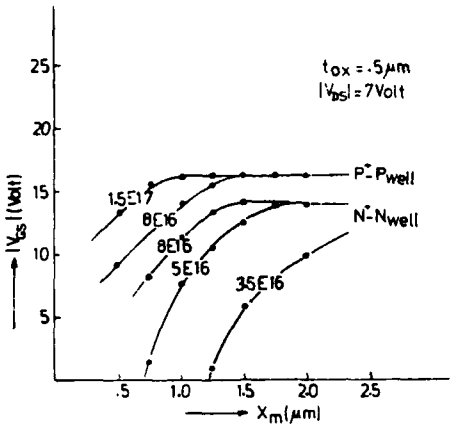


Fig.7 - Gate voltage Vgs, corresponding to 1 nA parasitic current per cm gate width, vs the extension of the well mask Xm, for various doping densities at the interface.

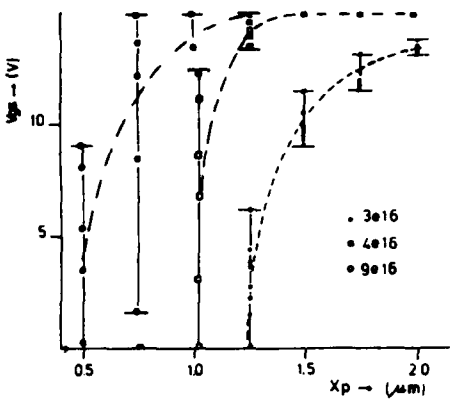


Fig.8 - Variation of gate voltage, corresponding to 1 nA parasitic current per cm gate width, vs X_p for various doping densities at the interface.

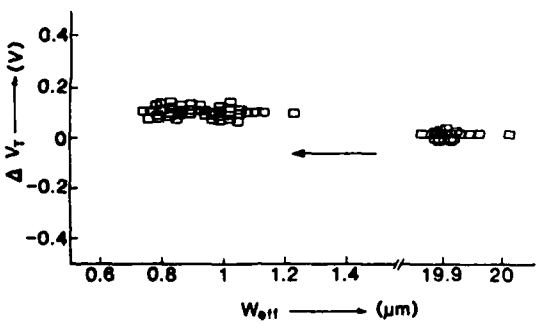


Fig.9 - The threshold shift ΔV_t vs the effective channel width W_{eff} for a NMOST of 20 μm channel length.

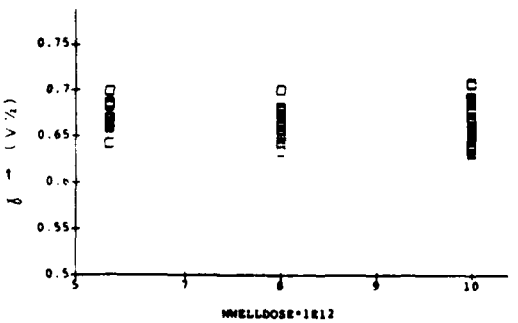


Fig. 10 - The body factor γ of a PMOST vs the Nwell dose for implantation energies of 850 keV (W/L=10/1).

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

DEVICE CHARACTERISATION OF A HIGH DENSITY HALF-MICRON CMOS PROCESS

P.H. WOERLEE, C.A.H. JUFFERMANS, H. LIFKA, A.J. WALKER, T. POORTER, H.J.H. MERKS-EPPINGBROEK and F.M. OUDE LANSINK

Philips Research Laboratories, NL-5600 JA Eindhoven, The Netherlands

Abstract The device characterisation and the ringoscillator performance of a high density half-micron CMOS process were studied. A novel recessed field isolation technology , twin retrograde wells, N^+ poly-silicon gate material and lightly doped drain structures for both the n- and p-channel devices were used in the device fabrication. The device properties and the ringoscillator performance will be presented.

1. INTRODUCTION

It is expected that the scaling of CMOS towards half-micron dimensions will lead to increases in packing density and circuit performance. However the performance increase of the scaled transistors can be degraded by several parasitic effects such as a reduced mobility /1/, series resistances /2/ and a lower power supply voltage. Furthermore the use of advanced technology which is needed for the realisation of a high packing density can influence the device behaviour. In this paper a study on the transistor properties and the ringoscillator performance of a high density half-micron CMOS process will be presented. The aim of the work was to characterize devices which were fabricated using advanced front-end technology. A scaled recessed LOCOS technology and twin retrograde wells were used during the device fabrication. The device design, the fabrication technology, the device characterisation and the ringoscillator performance will be presented in the following sections.

2. DEVICE DESIGN

The devices were designed for a 3.3 volt power supply voltage and a threshold voltage of 0.8 volt. The gate oxide thickness was 12.5 nm. The minimum effective channel length was 0.35 μm for the n-channel transistors and 0.4 μm for the p-channel transistors. N^+ poly-silicon gate material was used to obtain a high current-drive capability for the p-channel devices. Consequently the p-channel device is of the buried channel type. It is very sensitive to the drain induced barrier lowering (DIBL) effect. Simulations showed that for optimum device design the threshold counter doping junction depth and source-drain junction depths should be less than 0.08 μm and 0.15 μm respectively. The shallow channel junction depth was realized by the use of an anti punch through implant (P,As). Retrograde wells were used so that the device isolation and transistors could be optimized independently of each other. Furthermore an improved control of the threshold voltage of the p-channel devices can be obtained. An LDD structure was used for both the n- and p-channel devices for three reasons. Firstly for the reduction of the hot carrier degradation effects. Secondly the shallow N^- or P^- junction depth reduces the short channel and DIBL effects. Thirdly the N^+ and P^+ doping profiles can be optimized for a salicidation process. The N^- and P^- dosis were 4 $10^{13} cm^{-2}$ and the oxide spacer length was 0.15 μm.

3. FABRICATION PROCESS

Direct write electron beam lithography was used for the fine line patterning. A fully recessed LOCOS technique was used for the field isolation. The recessed LOCOS technology leads to a nearly planar surface after field oxidation. This is beneficial for future optical lithography and the reduction of step coverage problems. The recessed isolation process is based on a suppressed LOCOS scheme in which an oxynitride/nitride layer is used as an oxidation mask /3/. The recessed oxide is obtained by a two step oxidation process. First a thin (0.3 μm) oxide is grown at 1000° C in a wet oxygen ambient.This oxide layer is etched in a standard wet chemical HF etch. Thereafter a second oxidation is done (0.55 μm thickness). In figure 1 a scanning electron microscope micrograph of the field isolation structure after the second oxidation is shown. A nearly planar field oxide can be observed. A lateral encroachment of about 0.1 μm and an active area pitch below 1.2 μm were obtained. Formation of crystal defects was extensively studied. Wright defect etching and cross section transmission electron microscopy did not reveal any defects.

After removal of the oxynitride/nitride stack and cleaning steps a sacrifice oxidation is done. Thereafter the retrograde wells and the anti punch through implantations were done. Then 12.5 nm of gate oxide is grown in a dilluted dry O_2/N_2 mixture (10 percent O_2 by volume) followed by a blanket threshold implant and deposition of 0.3 μm polycrystalline silicon which was doped by an arsenic implantation ($10^{16} cm^{-2}$). After poly-Si etching the N^- implantation were done followed by a thermal drive step ($950°C$) to obtain sufficiently gradual source-drain junction profiles. The P^- implant was done either before or after the N^- drive step. The low temperature LDD variant leads to shallow junctions and a reduced lateral diffusion. Thisway DIBL phenomena and short channel effects are reduced. After LPCVD oxide deposition and spacer etching the N^+ and P^+ source-drain implantations were done followed by a final high temperature activation step ($900°C$). Then a 0.3 μm LPCVD oxide layer is deposited, contact windows are opened and Al/Si is deposited. The final step is a post metallisation anneal at $450°C$ in a N_2/H_2 mixture.

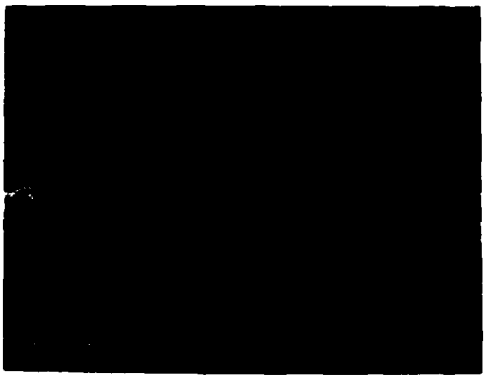


Fig.1 Scanning electron microscope micrograph of a cross section of the recessed field oxide after the second oxidation step. The oxidation mask width was 0.6 μm, the open spaces are 0.7 and 1.4 μm respectively .

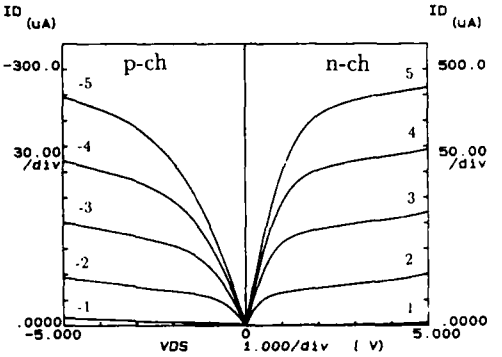


Fig.2 $I_{ds} - V_{ds}$ characteristics of n- and p-channel transistors with an effective gate length of 0.35 μm. The gate width was 0.7 μm. The gate bias is given in the figure. The p-channel device with a high temperature LDD drive step is shown .

4. DEVICE CHARACTERISATION

In figure 2 the $I_{ds} - V_{ds}$ characteristics of half-micron n- and p-channel transistors processed with the recessed field isolation are shown. The effective gate length and gate width of these devices were 0.35 μm and 0.7 μm respectively. Good device characteristics and a high current drive capability can be observed. The maximum transconductance in the saturation region is 150 $\mu S/\mu m$ and 90 $\mu S/\mu m$ for the n- and p-channel devices respectively. The ratio between the current drive capabilities of the p- and n-channel devices is 0.6.

In figure 3 the subthreshold characteristics of narrow width transistors are shown. The drain bias was varied between 0.1 and 4.1 volt. The off currents ($V_{gs} = 0V, V_{ds} = 4V$) are well below 10 pA/μm. The characteristics show no sign of side wall inversion, i.e. no double threshold can be observed. The subthreshold slopes of the n- and p-channel devices are 88 mV/dec and 100 mV/dec respectively. In figure 4 the linear ($V_{ds} = 0.1V$) and saturation ($V_{ds} = 3.7V$) threshold voltages are plotted for n- and p-channel devices versus the effective channel length. The threshold definition was set at a source-drain current of $W/L \cdot 10^{-7} A$. Data for the p-channel devices with the low temperature LDD drive step were plotted in this figure. The short channel effects for these n- and p-channel devices are sufficiently small for effective channel lengths down to 0.35 μm. P-channel devices processed with the high temperature P^- drive step showed much stronger short channel effects (see fig. 3) and an unacceptably large lateral diffusion. Narrow width effects on the threshold voltage were small, the threshold voltage variations were less than 50 mV for a variation of channel width from 0.7 μm to 8 μm. The absence of the narrow width effect is due to the high channel doping levels of the scaled devices and the use of retrograde wells. The body coefficient of the n-channel devices was found to vary between 0.2 and 0.6 $V^{1/2}$ for effective channel lengths between 0.35 μm and 0.9 μm. For the p-channel devices values between 0.35 and 0.6 $V^{1/2}$ were found.

The hot carrier degradation of the n- and p-channel devices was extensively studied. The n-channel devices were stressed at V_{ds} above 6.0 V and V_{gs} such that the substrate current is maximum. As a lifetime criterion a change in the maximum gain factor in the linear region of 10 percent was taken. The maximum drain bias for a device lifetime of 10 years as function of the effective channel length is plotted in figure 5. For a half-micron gate length the maximum drain bias is 4.2 V which is well above 3.7 V, the maximum value for a 3.3 V power supply voltage.

The worst case degradation of the p-channel devices occurs under different stress conditions. It has been reported recently /4/ that the strongest degradation of sub-micron p-channel devices occurs when the gate current is maximum i.e. for V_{gs} close to the threshold voltage. Hence the p-channel devices were stressed with V_{gs} such that the gate

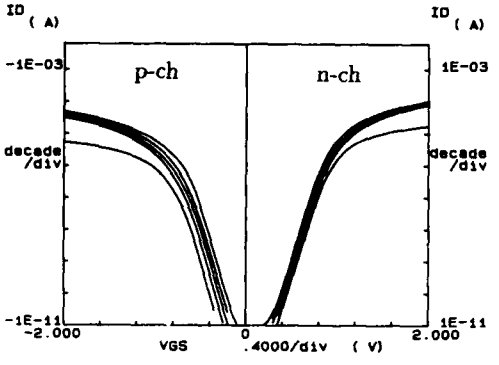


Fig.3 Subthreshold characteristics of the n- and p-channel transistors shown in figure 2. The drain bias is varied between 0.1 and 4.1 V .

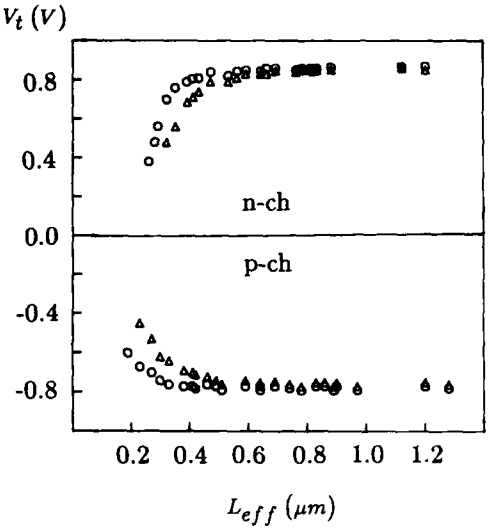


Fig.4 The linear (○) and the saturation threshold voltage (△) plotted versus the effective channel length for the n-channel and the p-channel devices The p-channel devices with the low temperature LDD drive step are shown .

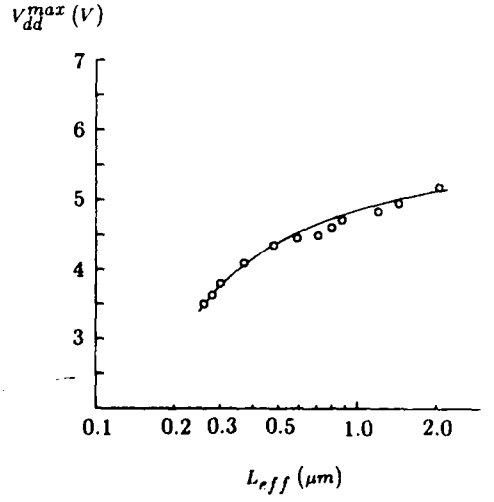


Fig.5 Maximum drain bias for a 10 year device lifetime versus the effective channel length for the n-channel devices .

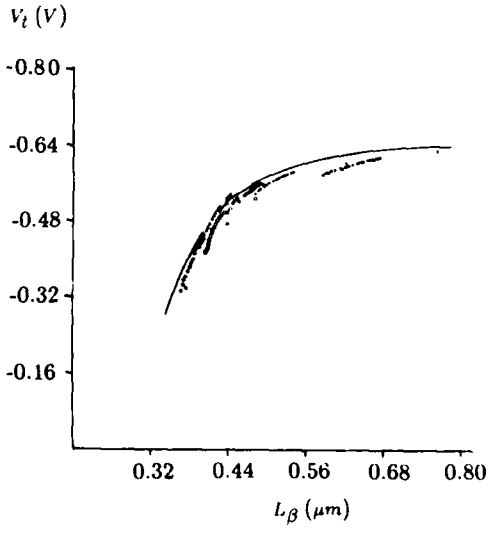


Fig.6 Variation of the saturated threshold voltage and the effective channel length of p-channel devices during stress. The solid curve is the V_t-L_β dependence of the unstressed devices.

current is maximal and with V_{d_s} between -6 V and -8 V. In figure 6 changes in the saturation threshold voltage versus the effective channel length L_β during the stress period are plotted for devices with the low temperature LDD drive step. L_β is determined from the gain factor in the linear region. It appears that during stress devices follow the same V_t - L_β curve as the unstressed devices with different channel lengths do. This indicates that variations in the threshold voltage during stress are mainly due to channel length shortening effects.Hence the devices with shorter channel lengths are most affected. The lifetime criterion of the p-channel devices under the above stress conditions was taken as a 100 mV shift in the reverse saturation threshold voltage ($I_{d_s} = 10^{-7}A, V_{d_s} = -3.7V$). For a 10 year device lifetime a maximum drain bias of -4.2 V was found for devices with a half-micron gate length .

5. RINGOSCILLATOR PERFORMANCE

Unloaded 51-stage ringoscillators with half-micron gate length n- and p-channel transistors were fabricated. The width of the active area was $1\mu m$ and square contact hole of 1 μm size were used. A plot of the delay per stage versus the power per stage for power supply voltages between 2 and 6 volt is given in figure 7. For a power supply voltage of 3.3 V the delay per stage is 120 ps, the power dissipation is 13 μW per stage. A relatively large source-drain junction area was used in the present design, so the stage delay can be improved considerably. The power-delay product at 3.3 V equals 1.6 fJ. This power-delay product is one of the lowest values reported thus far.

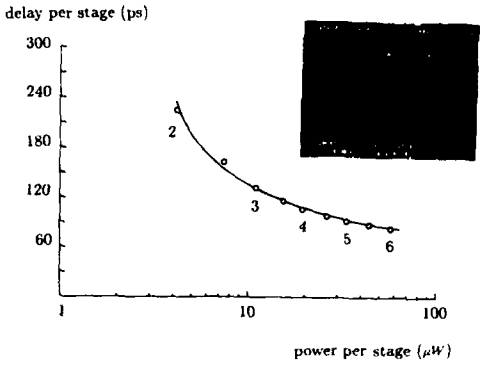


Fig.7 The delay per stage versus the power dissipation per stage for a 51 stage ringoscillator. Values for the power supply voltage are plotted in the figure. The insert shows an oscilloscope trace of the ringoscillator output at a power supply voltage of 3.3 V.

6. CONCLUSIONS

The device characterisation of a 3.3 V high packing density half-micron CMOS process has been described. An advanced recessed LOCOS technology and twin retrograde wells were succesfully used in the device fabrication. A good device behaviour has been found. A high performance and a good control of short channel effects were observed. The hot carrier degradation of the devices was studied and a 10 years device lifetime for a power supply voltage of 4.2 volt was found. The stage delay of a 51 stage ringoscillator was studied and both high speed and a very low power-delay product were observed. The observed device behaviour shows that very high performance circuits can be expected for half-micron CMOS.

References

/1/ G. Baccarani and M.R. Wordeman, IEEE Trans. Electron Dev., 30, pp 1295 (1983)
/2/ K.K. Ng and W.T. Lynch, IEEE Trans. Electron Dev., 34, pp 503 (1987).
/3/ P.A. van der Plas, W.C.E. Snels, A. Stolmeijer, H.J. den Blanken and R. de Werdt, Dig. of Tech. Papers 1987 symp. on VLSI technol., pp 19 (1987)
/4/ M. Koyanagi, A.G. Lewis,J. Zhu, R.A. Martin, T.Y. Huang and J.Y. Chen IEDM Tech. Dig. , pp 722, (1986)

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

GENERATION LIFETIME AND HOLD TIME OF SMALL MOS DEVICES

N.O. PEARCE[1], A.R. PEAKER and B. HAMILTON

Centre for Electronic Materials and the Department of Electrical Engineering and Electronics, University of Manchester Institute of Science and Technology, PO Box 88, GB-Manchester M60 1QD, Great-Britain

Resumé - Une nouvelle méthode pour la charactérisation de transientes d'inversion dans des dispositifs métal-oxide-semiconducteur est reportée. L'utilisation de la Spectroscopie Isothermique des Transientes de Niveaux Profonds permet de séparer les courants de génération d'origines différentes. La combinasion de graphes d'activation avec des graphiques 'Zerbst' donne une charactérisation complète de dispositifs de dimensions allant jusqu'à moins d'un micromètre.

Abstract - A new method for characterising inversion transients in metal-oxide-semiconductor devices is reported. The technique uses Isothermal Deep Level Transient Spectroscopy to separate generation currents from various sources. The combination of activation graphs with Zerbst plots gives a complete characterisation of devices with dimensions down to sub-micron levels.

1 - INTRODUCTION

The rate at which inversion layer charge builds up in an MOS structure is a crucial parameter which critically affects the performance of many MOS devices. In 1966 Zerbst[1] described a technique which allowed a separation of the generation currents which create the inversion layer into those components which depend on the depletion width from those which do not. Using this analysis Zerbst calculated a generation lifetime (from the depletion component) and a surface recombination velocity (from the components which did not depend on the depletion width). Many variations of the original Zerbst method have been proposed. A very comprehensive review of these was published in 1985 by Kang and Schroder[2]. Some of the modifications speed up the measurement, some simplify the calculation, some undertake a more rigorous analysis but in the main the fundamental concept of separation into two components remains.

The last five years has seen remarkable progress in MOS devices and the silicon from which they are made. The use of epitaxial layers and improved gettering techniques result in the contribution of minority carriers from the depletion region being insignificant. This means that the hold time of the best MOS structures manufactured today is not dependent on deep states in the depletion region. However, contributions from carrier diffusion from the bulk material and from the region where the surface depletion region edge meets the surface are increasingly important. The latter is particularly so in very small devices. The bulk contribution is not depletion width dependent and so the Zerbst analysis embraces it into the surface term. However, the contribution from the depletion edge at the surface is depletion width dependent (but not linearly) and so distorts the Zerbst curve. Consequently, additional measurements are needed to propertly characterise the generation processes. In addition, the instrumentation normally used for the Zerbst measurement only permits the use of special test capacitors with fairly large areas, and not of the normal micron-sized FET gates in which the relative importance of the various generation mechanisms is expected to be very different. In particular, surface generation from the lateral depletion region becomes of increasing importance as the device size is reduced. It is also interesting to consider that the total number of bulk traps or interface states present in a high quality sub-micron device is, on average, of the order of two or three. Thus the trap population, whatever its origin, will be subject to statistical fluctuation, the effects of which can never be appreciated on large test capacitors.

Recently we described a technique which uses a Deep Level Transient Spectroscopy (DLTS) system to overcome some of these limitations[3]. In this method, DLTS scans of the capacitance change are obtained and an Arrhenius plot then yields activation energies characteristic of the generation mechanisms. This method also has advantages over single-shot Zerbst in terms of its high sensitivity, principally because the data are a result of signal averaging in a relatively noise-free part of the spectrum. Essentially, the DLTS output is the difference between two successive measurements of the device capacitance. This means that flicker noise and d.c. offset are subtracted out, enabling devices with very small quiescent capacitances to be measured. Hence the technique can be used to directly measure micron-sized devices. Unfortunately it was necessary to considerably simplify the analysis because, as the temperature is varied, some of the macroscopic parameters of the specimen (such as the equilibrium inversion capacitance C_f and the carrier concentration n_i) change. The present paper describes the use of an isothermal variant of the DLTS technique, for which the temperature is maintained constant throughout a scan of the system's rate window.

([1]) present address: Bio-Rad Laboratories,780 Montague Expressway, Suite #201, San Jose, CA95131, USA.

2 - THE ISOTHERMAL TECHNIQUE

A family of DLTS peaks obtained on an MOS test capacitor by the I-DLTS measurements is shown in Figure 1. Unlike normal DLTS, the signal ΔC is plotted against the logarithm of the rate window, with temperature as a parameter. There are several additional advantages in measuring emission transients using the isothermal variety of DLTS. In particular, stressing of the sample resulting from the application of temperature cycles is minimised because it is not necessary to take the sample repeatedly through a wide range of temperatures. Such considerations are of particular importance in MOS devices where temperature cycling to cryogenic temperatures often results in degradation.

Essentially, the technique consists in scanning the system rate window over a wide range whilst maintaining the temperature constant. The measurement is then repeated at other fixed temperatures in order to obtain a family of DLTS peaks. These are used to produce an activation plot but since each DLTS peak is a measurement of an unchanging capacitance transient, it is possible, by manipulating the data constituting one peak, to replot the original transient. This exercise can in fact be carried out with a high degree of accuracy. It might seem paradoxical that such a reconstruction should be undertaken but it gives rise to more noise-free data than a single direct measurement of the transient. The rate window is defined by two sampling periods, the mid point of these occurring at times t_1 and t_2 after the device has been switched into depletion. The ratio between these two times is kept constant at 2 (i.e. $t_2/t_1=2$). The sampling times are swept across the transient starting at the lowest values of the sample times t_1 and t_2. This corresponds to the highest rate window. An Arrhenius plot can be constructed from the position of the DLTS peaks and information concerning the generation mechanisms obtained in the same way as described previously using the temperature scanning technique[3].

In order to reconstruct the C-t transient from the DLTS peaks, we must convert the $C(t_1) - C(t_2)$ vs. rate window data which are produced by the DLTS measurement, into the actual values of the capacitance as a function of time. To do this it is important that the first measurement, corresponding to the shortest sampling times, should give rise to a signal which is negligible compared to the quiescent capacitance of the device. The reason for this is that the change in capacitance between t=0 and $t=t_1(1)$, where $t_1(1)$ is the shortest sampling time used in the scan, is not known and is assumed to be zero. In practice this condition is readily satisfied within the range of normal measurement temperatures. The driver software is written in such a way that the times t_1 and t_2 double every 12 points, this number being chosen to generate an appropriate number of data points. This means that at any measurement point other than the first twelve points, the time of the second sampling time, t_2, is equal to the value of the first sample, t_1, twelve points back. This may be expressed more conveniently in mathematical terms, as follows, where i corresponds to the i^{th} data point of the scan:

$$t_2(i) = t_1(i-12) \quad \text{for } i>12 \qquad (2)$$

This relationship between the two samples is used to construct the capacitance transient from the DLTS result. The time at the mid-point of the first sampling gate t_1 is the time t(i) of the C-t pair, and the capacitance C(i) is extracted in the following way from the DLTS data: first, the capacitance difference between the value at t = 0 and at the time $t = t_1$ is computed for all points; then, this capacitance is added to the difference between the maximum (saturated) value of the capacitance, and the previously determined value of the equilibrium inversion capacitance C_f. In the case of the first twelve points for which the signal ΔC should be negligibly small, the values of C(i) are set equal to ΔC, which means, in effect, that we are ignoring the capacitance change between t=0 and $t=t_1$ for the first 12 points:

$$C(i) = \Delta C(i) \qquad \text{for } 1<i<12 \qquad (3)$$

It follows that for all the remaining points:

$$C(i) = \Delta C(i) + C(i-12) \quad \text{for } i > 12 \qquad (4)$$

As both of the samples reach the flat, equilibrium part of the C-t curve. C(i) will saturate. This saturated value C_{sat} is compared to the measured inversion capacitance C_f, and the difference, $dC = C_f - C_{sat}$, is added to all the C(i) values obtained from equations (3) and (4). The array elements C(i) now hold the actual capacitance of the device at the time $t(i) = t_1(i)$ following the return to the quiescent reverse bias after the accumulation pulse, thus completing the conversion of the DLTS data to the capacitance transient.

3 - EXPERIMENTAL ILLUSTRATION

The Arrhenius plot, derived from the data in Figure 1, is shown in Figure 2. It yields a straight line with an activation energy of 429 meV. This low energy eliminates the possibility that the minority carriers originate in the bulk and diffuse to the depletion region. In such cases the process is activated at near band gap energy.

If the dominant generation mechanism is due to generation via deep states within the depletion layer an energy of slightly more than half band gap (650meV) is observed[3]. In the case shown it is considerably less and it is likely that the generation is dominated by surface phenomena. Carrying out a reconstruction of the C-t transient from one of the I-DLTS peaks and performing the Zerbst analysis, as shown in Figure 3, provides support for this interpretation. Indeed, apart from the later stages of the recovery, the Zerbst plot exhibits a constant generation current, independent of the width of the depletion layer. This does not fit the model for deep level generation within the depletion layer, but is consistent with surface generation in a large area device. The value obtained for the carrier generation lifetime from this plot is meaningless since the Zerbst plot is virtually flat. The fit to the C-t plot was computed by solving Zerbst's equation using the values for the surface generation velocity obtained from the Zerbst analysis. Not surprisingly, the transient response time is solely dependent on the value of the intercept of the Zerbst plot with the vertical axes.

From the Zerbst plot, it is immediately evident that depletion layer generation is not responsible for the inversion. However, both the constant generation currents and the high temperature at which the measurements were obtained, make a diffusion component from the bulk the favourite contender. However, in this case we know that this is not possible because of the small value of the activation energy. Although Schroder and Nathanson[4] showed that surface generation, especially from a lateral depletion region, could play an important role as a source of minority carriers, diffusion from the bulk and deep level generation in the depletion layer are often assumed to be the only important sources of minority carriers to feed the inversion process. It should also be remembered that the relative importance of surface generation compared to diffusion and depletion layer generation increases as the device size shrinks, making this source of generation very important in very small devices. Unfortunately the Zerbst plot does not permit a distinction to be drawn between the sources of carriers which are not dependent on depletion width. However, the Arrhenius plot provides conclusive evidence that the process is not a band to band excitation (this would have an activation energy > 1 eV) so eliminating the possibility of minority carrier diffusion being a major process in this case. Although such an Arrhenius plot could be obtained from conventional DLTS the Zerbst plot cannot be reconstructed and, again, ambiguity persists which only I-DLTS resolves.

The surface generation velocity S_g under the gate can be calculated from:

$$S_g(t) = \frac{\epsilon_s D_{ox}}{\epsilon_{ox}} \left[\frac{N_s Y_z(t)}{2n_i} - \frac{C_{ox} X_z(t)}{C_f \tau_g'} \right] \qquad (5)$$

where $X_z(t)$ and $Y_z(t)$ are the X and Y co-ordinates of the Zerbst plot; τ_g' is the effective generation lifetime calculated from the slope of the Zerbst plot and which includes any contribution by surface generation from the lateral depletion region C_{ox} is the oxide capacitance, D_{ox} is the oxide thickness, C_f is the quiescent capacitance in inversion, N_s is the substrate doping concentration, n_i is the intrinsic carrier concentration, and ϵ_s and ϵ_{ox} are the dielectric permittivities of the semiconductor and of the oxide, respectively. The value of surface generation velocity obtained from the data in Figure 3 is 90 cm/s over most of the time span.

4 - CONCLUSIONS

The I-DLTS method is an extremely useful addition to the assessment methods available for MOS devices. It complements the long-established Zerbst method and it is possible to construct a Zerbst plot from the I-DLTS measurement which can result in remarkably noise-free data. This allows the direct measurement of sub-micron devices. Very importantly, the combination of I-DLTS data with the Zerbst analysis permits a complete separation of the various mechanisms giving their dependence on depletion width and the activation energy of the process. The technique uses the high sensitivity of a conventional DLTS system for accurately measuring inversion layer formation on devices of any size, down to sub-micron dimensions. It is a routine matter to repeat the isothermal measurement at several temperatures and produce an Arrhenius plot of the DLTS peaks. This yields the activation energies of the dominant generation processes.

As it is possible to re-construct the capacitance transient which gave rise to the DLTS peak a complete Zerbst analysis can be carried out on the highly accurate capacitance-time data. Since the temperature of the sample is held constant throughout a scan, the time taken for a measurement is not limited by the time required to change and measure the temperature of the sample. Moreover, it is not necessary to stress the sample by continuously sweeping it through a wide range of temperatures.

5 - REFERENCES

/1/ ZERBST M. Z. *Angew Phys.* 22 (1966) 30
/2/ KANG J.S. and SCHRODER D.K., *Phys. Stat. Sol.* 89 (1985) 13
/3/ PEARCE N.O., HAMILTON B., PEAKER A.R., and CRAVEN R.A., *J. Appl. Phys.* 62 (1987) 576
/4/ SCHRODER D.K., and NATHANSON H.C., *Solid-State Electronics* 13 (1970) 577

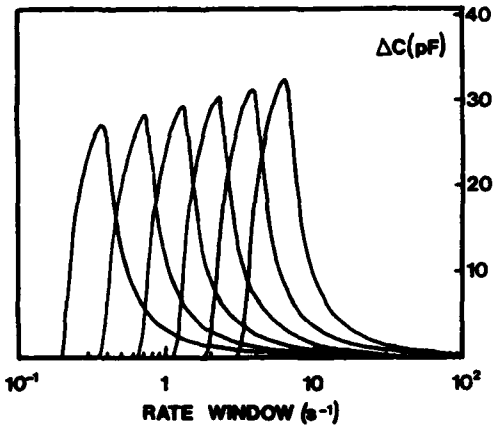


Figure 1

A family of DLTS processed inversion transients plotted as a function of rate window. The scans were taken at 342K, 332K, 322K, 312K, 302K and 292K. The highest temperature trace is on the right.

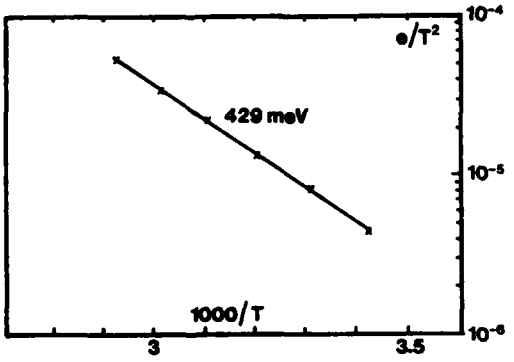


Figure 2

An Arrhenius plot of data derived from the position of the peaks in Figure 1.

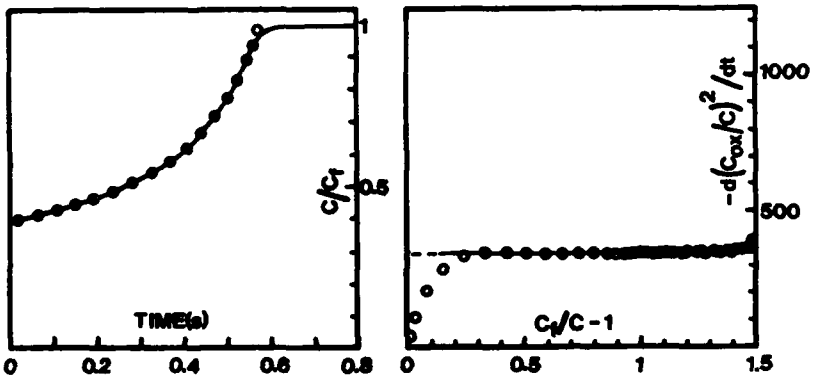


Figure 3

(a) A normalised capacitance time plot reconstructed from the I-DLTS measurement at 332K shown in Figure 1, and (b) The Zerbst plot derived from (a). The horizontal portion shows that the carrier generation process is not a function of depletion width and so the minority carriers are produced at the semiconductor surface or diffuse to the depletion region from the bulk.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CMOS TECHNOLOGY FOR CCD VIDEO MEMORIES

G.J.T. DAVIDS, P.B. HARTOG, J.W. SLOTBOOM, G. STREUTKER,
A.G. van der SIJDE and W. WIERTSEMA

Philips Research Laboratories, PO Box 80.000, NL-5600 JA Eindhoven, The Netherlands

ABSTRACT - A new double polysilicon gate technology for an 835 Kbit CCD video memory with a cell of 4x4 μm^2 [1] is presented. The spacer technology for the LDD MOSFET's is integrated in the isolation of the double poly CCD structure. This makes the CCD fully compatible with standard CMOS processing and relaxes the anisotropic plasma etching of second poly electrodes. The charge transfer efficiency is high for a SCCD without fat zero ($\epsilon \approx 2 \cdot 10^{-4}$). Measurements and calculations on the charge transfer show no degradation for doping concentrations below $8 \cdot 10^{15}$ cm^{-3}. The leakage current density is measured on the 835 Kbit memory and agrees with earlier measurements on a 308 Kbit CCD video memory [2,6].

INTRODUCTION

CCD memories are particularly suited for video and audio applications because of the serial data stream, and have a cell size which is 2 to 3 times smaller than a DRAM cell. Double polysilicon gate technology is well known for its application in CCD's. However, downscaling of the conventional double poly technology for high densisty CCD memories is not straightforward. The conventional thermal inter poly oxide, based on selective oxidation of phosphorus doped poly, is not compatible with short channel LDD MOSFET's. An attractive alternative is the integration of the well known oxide spacer technology for the LDD MOSFET's in the inter poly isolation of the SCCD. This makes the double poly CCD structure fully compatible with standard CMOS processing, and relaxes anisotropic etching of the second poly as will be shown. With the new double poly technology an 835 Kbit CCD video memory [1] with a cell of 4x4 μm^2 is realized in a 1.3 μm single metal CMOS process.
From measurements and calculations on the charge transfer in the new double poly CCD structure the substrate doping concentration is optimized. Leakage current densities at higher temperatures are measured in the 835 Kbit memory fabricated in the new double poly technology. They are compared with earlier measurements on a 308 Kbit CCD video memory [2,6].

TECHNOLOGY

The process flow for the double poly CCD structure and its integration with p-channel and LDD n-channel MOSFET's in a 1.3 μm N-well CMOS process are shown in fig. 1. A cross section of a CCD memory cell with a pitch of 4 μm is shown in fig. 2, and the gap between first and second poly is shown in fig. 3. The double poly structure with anisotropically etched second poly can be seen in a detail of the 835 Kbit memory in fig. 4. In contrast with a conventional thermal inter poly oxide based on selective oxidation, the spacer isolation gives the option to choose independently the thickness of the inter poly oxide and of the second gate oxide [3]. In our case (fig. 1) the thickness of the second gate oxide is 400 Å with an inter poly oxide thickness of 0.25 μm. The first gate oxide is 250 Å thick. The smooth interfaces of the inter poly oxide yield a high (> 25 V) inter poly breakdown voltage. The spacer isolation also relaxes the anisotropic plasma etching of the second poly, since the spacers of the first poly gates have positive slopes (fig. 2 and 3). Negative slopes that occur with a conventional thermal inter poly oxide can yield residual poly ridges after plasma etching causing shorts between adjacent second poly gates.
NMOS CCD test structures with relaxed dimensions without LDD's were processed with both the conventional and the new double poly technology for a comparison of the charge transfer. A detail and the process flow of a test structure with the conventional double poly isolation is shown in fig. 5. Silicon nitride spacers are applied (see fig. 5) with extra processing steps in order to prevent for the residual second poly ridges. The thickness of the first gate oxide is kept again at 250 Å, and with the selective oxidation 800 Å of second gate oxide is grown on monosilicon simultaneously with 2300 Å of oxide on first poly.

CHARACTERIZATION

The characterization and optimization of the new CCD memory structure will be discussed. With the combination of a 5 V four phase push clocking scheme (fig. 7) and thin gate oxides, a high charge density of $3 \cdot 10^{12}$ el./cm^2 is realized. This results in a high charge transfer efficiency as can be seen in fig. 6a. It shows that the signal loss after 200 transfers is relatively small for a SCCD without fat zero ($\epsilon \approx 2 \cdot 10^{-4}$). Although the monosilicon interface under the second poly electrodes has been exposed to the plasma etching during the formation of the spacer, see fig. 1b, careful processing gives a reliable second gate oxide. This also appears from a comparison of the charge transfer with measurements on the CCD structures that were processed with the conventional inter poly isolation. The measurements yield the same charge transfer efficiency and not a higher, as would have been expected in the case of a damaged monosilicon interface for the new technology.
Care has to be taken with the doping underneath the CCD electrodes. The result of 2D numerical simulations of the charge transport in fig. 7 demonstrates that, at high doping concentrations, unwanted potential barriers are introduced in the gap region between successive gates (see fig. 7 positions A and B). Below the line "inter-electrode potential barrier" in fig. 8 complete charge transfer is possible. However, if the gap becomes too small, high lateral electric fields will give avalanche charge multiplication [4,5]. Threshold voltage measurements on dual gate transistor structures agree with these simulations and show (fig. 9) that for an oxide spacer of 0.3 μm the doping concentration in the CCD should be kept below about $8 \cdot 10^{15}$ cm^{-3} in order to prevent for the unwanted potential barriers. Fig. 6b shows the strong degradation of the charge transfer at too high a doping concentration.
At higher temperatures the leakage current density increases strongly particularly due to diffusion of minority carriers from the bulk. The measured CCD leakage current densities in the 835 Kbit memory made in the new double poly technology agree (see fig. 10) with earlier measurements on a 308 Kbit memory which was processed in a poly/metal gate 2 μm NMOS technology [2,6]. The typical leakage current density at T = 90°C is 0.2 μA/cm^2. In the 835 Kbit CCD video memory the influence of the leakage current is much less than in the 308 Kbit memory not only because the cell area is 3 times smaller but also because the signal charge is 1.7 times higher. This is illustrated in fig. 11. The leakage charge compared to the signal charge is only 3% at 90°C and 20 msec delay time in the 835 Kbit memory, while for the 308 Kbit memory it is 15%.

CONCLUSIONS

With a new double poly technology an 835 Kbit CCD video memory [1] with a cell of 4x4 μm^2 is realized in a 1.3 μm single metal CMOS process. The integration of the oxide spacer technology for the LDD MOSFET's in the inter poly isolation of the SCCD makes the double poly CCD structure fully compatible with standard CMOS processing and relaxes the anisotropic etching of second poly. A high charge transfer efficiency without fat zero ($\epsilon \approx 2 \cdot 10^{-4}$) is obtained if the substrate doping concentration is kept below $8 \cdot 10^{15}$ cm^{-3}. CCD's fabricated with conventional inter poly isolation show a comparable charge transfer efficiency but are not compatible with short channel LDD MOSFET's. The leakage current density measured on the 835 Kbit CCD agrees with earlier measurements on a 308 Kbit memory, however in the 835 Kbit the influence of the leakage charge at higher temperatures is less due to the 3 times smaller cell size and 1.7 times higher stored signal charge.

References:

[1] H.J.M. Veendrick e.a., "An 835 Kbit video serial memory (VSM)", ISSCC 1988, p. 44.
[2] M.J.M. Pelgrom e.a., "A digital field memory for TV-receivers", Internat. Conf. on Consumer Electronics, June 1983, Chigago.
[3] R.M. Levin and T.T. Sheng, J. Electrochem. Soc. Vol. 130, no. 9, p. 1894 (1984).
[4] J.W. Slotboom et al., "High-density CCD video-memories: physics and technology", Conf. on Solid State Devices and Materials, Tokyo 1986, p. 315.
[5] J.W. Slotboom et al., "Surface impact ionization in silicon devices", IEDM 1987, p. 494.
[6] J.W. Slotboom et al., "Leakage current in high density CCD memory structures", IEDM 1981, p. 667.

1. Definition of N-well, channel stopper, field oxide, first gate-oxide and Vt implantation.
2. Deposition of poly-Si (0.4 μm).
3. Phosphorus doping.
4. Deposition of SiO_2 (0.25 μm TEOS).
5. Definition of poly-Si/SiO_2 stack.
6. N-MOST LDD lithography.
7. Phosphorus LDD implantation.
8. Resist strip, see fig. 1a.
9. Deposition of SiO_2 (0.25 μm TEOS).
10. RIE SiO_2, "spacer etching", (CF_4, CHF_3,Ar), see fig. 1b.
11. Oxide growth (steam, 900°C, 400 Å).
12. Deposition of poly-Si (0.4 μm).
13. Phosphorus doping.
14. Deposition of SiO_2 (0.15 μm TEOS).
15. Definition of poly-Si/SiO_2 stack.
16. Implantation of S/D of NMOST's and of PMOST's after the definition of implantation masks, resist strip, drive 30 min. 925°C N_2, see fig. 1c.
17. Standard backend of a single metal CMOS process.

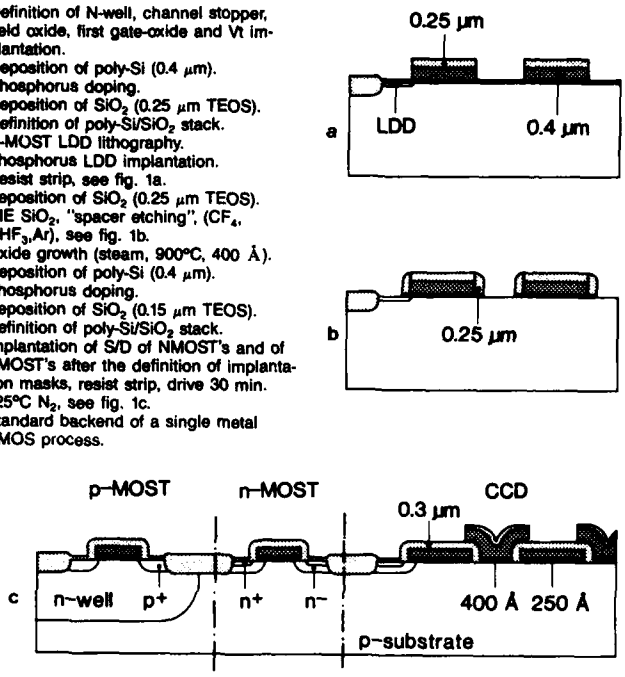


Figure 1.
Process flow for a double poly CCD structure and its integration in an N-well CMOS process.

Figure 2.
SEM photograph of the cross section of a CCD memory cell with a pitch of 4 μm.

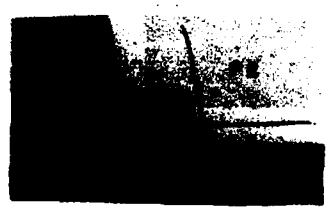


Figure 3.
TEM photograph of a cross section of the gap region between first and second poly (P_1 and P_2 respectively).

Figure 4.
SEM photograph of a detail of the 835 Kbit video memory showing a part of the double poly structure with anisotropically etched second poly.

1. Definition of channel stopper, field oxide, first gate-oxide and Vt implantation.
2. Deposition of poly-Si (0.5 μm).
3. Phosphorus doping.
4. Definition of poly-Si.
5. Grow oxide (steam, 950°C, 800 Å on mono-Si and 2300 Å on poly-Si).
6. Deposition of nitride (400 Å).
7. RIE of nitride.
8. Deposition of poly-Si (0.4 μm).
9. Phosphorus doping.
10. Deposition of SiO_2 (0.15 μm TEOS).
11. Definition of poly-Si/SiO_2 stack.
12. Implantation of S/D, anneal.
13. Standard backend of a single metal NMOS process.

Figure 5.
Conventional double poly CCD structure based on selective oxidation of phosphorus doped poly with nitride spacers; detail above and process flow below.

Figure 6.
Output pulses (without fat zero) of a small ($W = 2$ μm) serial CCD channel. At a high doping concentration of $2 \cdot 10^{16}$ cm^{-3}, the charge transport is degraded by potential barriers (see b), while at a low doping concentration of $2 \cdot 10^{15}$ cm^{-3} only a small signal loss due to interface trapping appears (see a).

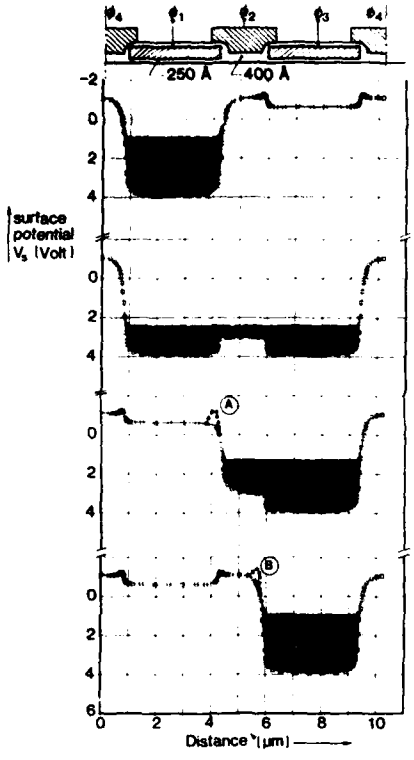


Figure 7.
Calculated surface potential distribution at four important time steps of the four phase push clocking scheme; A and B indicate unwanted inter-electrode potential barriers for high doping concentrations.

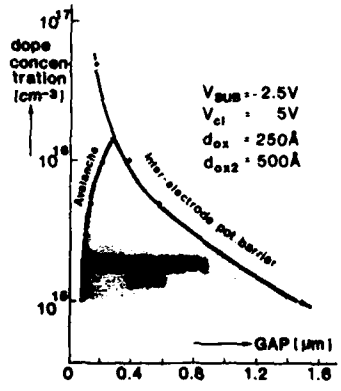


Figure 8.
For substrate impurity concentrations and gap sizes in the "safe-operation-area" no inter-electrode potential barrier is present and the maximum electric field is less than $2 \cdot 10^5$ V/cm, which is used as an indication of avalanche multiplication.

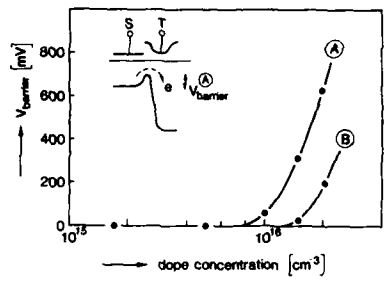


Figure 9.
Measured inter-electrode potential barriers as a function of the doping concentration for two situations A and B (see fig. 7).

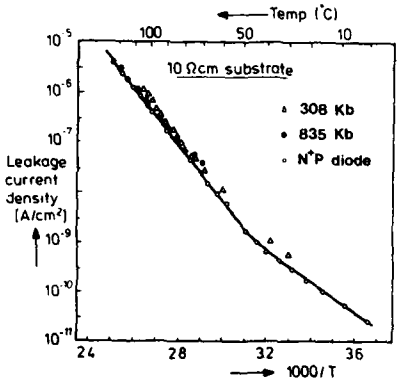


Figure 10.
Leakage current density for CCD and for gated diode structures as a function of temperature; (△) is measured on a 308 Kbit CCD, (●) on an 835 Kbit CCD and (○) on a gated diode.

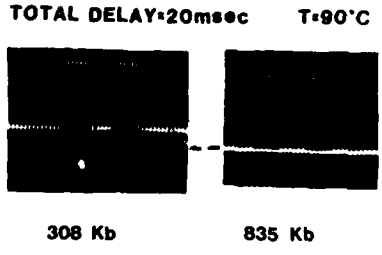


Figure 11.
Output pulses of 308 Kbit and 835 Kbit CCD memory structures for a comparison of the leakage charge.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

EFFECTS OF MECHANICAL STRESS ON MOS STRUCTURES WITH $TiSi_2$ GATES

P.J. REUTERS, J. GIESE, M. OFFENBERG, W. RICHTER, S. EWERT and P. BALK

Inst. of Semiconductor Electronics and Inst. of Physics, Technical University Aachen, D-5100 Aachen, F.R.G.

Résumé - Un piègeage d'électrons et une génération densité d'etats á l'interface par l'injection d'électrons ainsi qu'une deterioration du comportement de rupture ont été observé dans des capacités MOS ayant une électrode de grille en $TiSi_2$. Ces effets sont plus prononcés pour des épaisseurs d'électrode plus importantes. Ils semblent être causées par des contraintes méchaniques provoquées par l'électrode de grille de ce système MOS.

Abstract - MOS capacitors with $TiSi_2$ gate electrode show electron trapping, interface state generation upon electron injection and reduction of breakdown strength; these effects become more pronounced for electrodes with increasing thickness. Evidence is presented that this degradation of the insulator is due to mechanical stress in the MOS system caused by the gate.

1 - Introduction

$TiSi_2$ is an interesting material for interconnection, contact and MOSFET gate application because of its low resistivity and its ability to reduce native oxide films. For this reason this silicide finds application in the VLSI technology.

MOS capacitors with $TiSi_2$ gate were found to show properties typical for highly defective insulator films. They contain electron traps with capture cross-sections of 10^{-15} to $10^{-16}cm^2$. Similar traps were found in oxides implanted with Ti and electroded with Al. This suggests that the traps in the first case could also be due to Ti. The above explanation would fit in with the observed diffusion behavior of Ti implanted into SiO_2. However, the silicide gate leads to the occurrence of heavy compressive stress at the $Si-SiO_2$ interface. It will be shown that the compressive stress rather than indiffusion of Ti into the oxide is responsible for the degradation in MOS properties.

2 - Experimental

The samples used for electrical mesurements were MOS capacitors fabricated on p-type (100) Si substrates. After growing oxide films of 50 and 90 nm in O_2 silicide gates were prepared by e-gun coevaporation of Ti and Si. Polycide structures were fabricated by growing a 250 nm LPCVD poly-Si layer, doping it with As by ion implantation and finally depositing a Ti/Si film. Silicide formation was carried out at 800°C in Ar. Capacitors were defined by optical lithography and dry etching. Avalanche injection, Fowler-Nordheim injection, quasi static C-V and I-V measurements were performed to characterize the trapping behavior, the interface state generation and the breakdown of the structures.

The samples for stress determination were fully $TiSi_2$ coated oxidized Si wafers in the case of interferometeric measurement of the radius of curvature. For determining the shift of the Si TO phonon by Raman spectroscopy partially coated samples were used. Both measurement techniques yield information on the stress at the Si-SiO_2 interface.

3 - Results and discussion

Results on the electron trapping in a sample with 90 nm oxide and 200 nm $TiSi_2$ are shown in fig.1. This figure gives the shift in flatband voltage in dependence on the density of injected electrons. It follows from this plot that the oxide contains traps with cross-sections in the 10^{-15} to 10^{-16} cm^{-2} region. However, removal of the silicide and replacing it with an evaporated Al electrode leads to the disappearance of most of the traps. This suggests that mechanical stress is dominating the trapping behavior. In contrast to chemical interaction between the silicide and the SiO_2, stress effects should exhibit a dependence on the thickness of the silicide film. Fig. 2 shows electron trapping data for MOS capacitors with $TiSi_2$ gates of different thickness. The

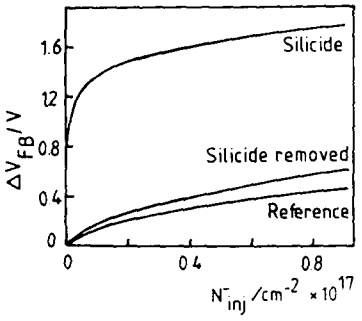


Fig. 1: Effect of substituting $TiSi_2$ by Al on electron trapping behavior

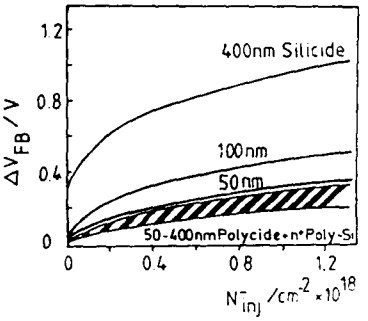


Fig. 2: Effect of $TiSi_2$ thickness on electron trapping behavior

SiO_2 film in these samples was 50 nm. For comparison, data on samples with polycide and with poly-Si gates are also shown. The low trapping density in the latter case is apparent. The trapping in the silicide gate capacitors indeed increases with the thickness of the gate electrode.

Similarly, the generation of the interface states upon Fowler-Nordheim injection of electrons is highest for the sample with the thickest $TiSi_2$ electrode (fig.3). The injecting field had to be lowered in order to avoid the breakdown of this sample. Already at a low density of injected electrons a large density of interface traps is observed. Before injection all samples exhibited interface trap densities around $10^{10}cm^{-2}$. In fig. 4 typical I-V curves showing the breakdown behavior of samples with different gates are presented. Polycide gate samples show a breakdown field comparable to that found with poly-Si gates. The values for the polycide gates are distinctly lower and decrease with electrode thickness. These observations indicate that the poly-Si interlayer in the polycide samples lowers the stress level near the Si-SiO_2 interface.

The optical interference method provides average stress values parallel to the interface over relatively large areas. In all silicide and polycide gate samples

this stress is compressive. Only in the poly-Si gate samples the stress was tensile but very small. Increasing $TiSi_2$ thickness leads to higher stress levels at the Si-SiO_2 interface (fig. 5). The stress in polycide gate samples was distinctly lower which indicates the role of the poly-Si film in reducing the overall stress.

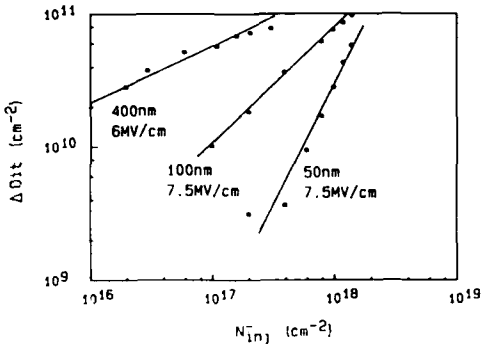


Fig. 3: Effect of $TiSi_2$ thickness on interface state generation

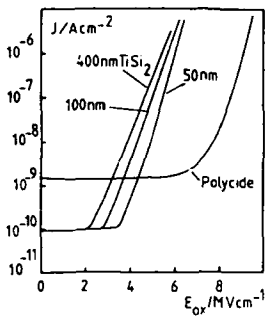


Fig. 4: Effect of $TiSi_2$ thickness on breakdown characteristics

On the other hand, Raman measurement of the phonon frequency presents information on the deformation of the lattice perpendicular to the interface/1/ In this case the observed reduction of the phonon frequency indicates a stretching of the Si-Si bonds, corresponding to tensile stress in this direction. In contrast to the interferometric approach the Raman measurement provides information on stresses localized in relatively small areas. These are determined by the beam diameter, which in our case amounted to approximately 20 µm. It was ascertained by working with different beam intensities that the observed effects were not affected by local sample heating. The derivation of absolute values for the stress is somewhat problematic for this method. The data in fig. 6 show the change of the stress in the vicinity of the

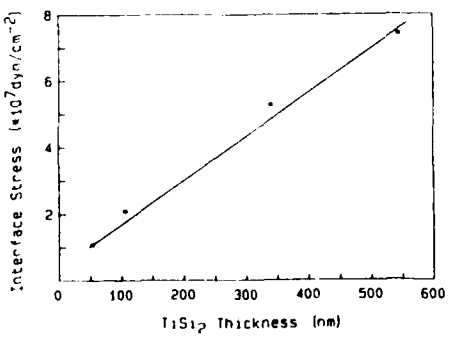


Fig. 5: Stress at the Si-SiO_2 interface form curvature measurements

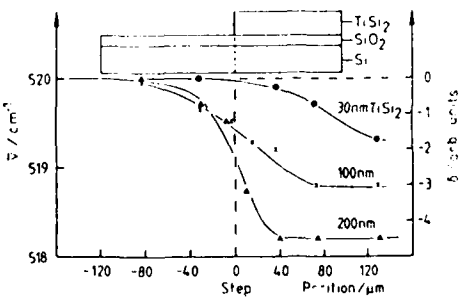


Fig. 6: Stress in Si at $TiSi_2$ edge from Raman study

edge of a silicide gate. Also in this technique an increase in the stress (which is tensile in this case) for thicker $TiSi_2$ films is observed. When passing the edge to the silicide the stress increases monotonically, i.e. without passing through a peak value. This finding may be contrasted with reports of stress concentration (i.e. a maximum value of this quantity) at such topographical features /2,3/. However, it should be remembered that our technique presents information on the stress perpendicular to the surface, whereas the data in /2,3/ refer to the parallel direction.

From the data presented in the present study it may be concluded that chemical effects do not play a determining role in the degradation of the MOS properties of silicide gate capacitors. Stress effects alone are responsible for these effects.

/1/ Richter, W.: in Springer Tracts in modern Physics Vol. 78, ed. G. Höhler, Z.A. Niekisch, Springer 1976

/2/ Zekeriya, V., Wong, A., and Ma, T.-P., Appl. Phys. Lett. 46 (1985) 80

/3/ Serebrinsky, J.H., Sol. St. Electr. 13 (1970) 1435

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

LATCH-UP FREE VLSI CMOS CIRCUITS CONSIDERING POWER-ON TRANSIENTS

W. RECZEK, J. WINNERL* and W. PRIBYL

Siemens AG, Components Group, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.
**Siemens AG, Corporate Research and Development, Otto-Hahn-Ring 6, D-8000 München, F.R.G.*

Résumé - Le présent article se propose de formuler les conditions d'apparition du latch-up à condition de power-on et d'en établir les origines physiques, d'indiquer les tendances actuelles permettant de réduire l'acuité du problème au niveau de la conception des circuits et de faire un bilan des mesures permettant une meilleure protection des systèmes à l'égard de ce phénomène.

Abstract - Power-on latch-up is depending on circuit design and technology. The use of an epilayer increases the latch-up hardness of conventional CMOS (LOGIC) while for n-well CMOS with VBB generator (DRAM) the latch-up suszeptibility is increased because of the capacitive voltage divider in the periphery and the high effective substrate shunt resistance. Therefore protection circuits provide latch-up free operation.

1 - INTRODUCTION

The success of a CMOS technology in the micron and submicron range is strongly related to latch-up hardness. So far, most latch-up investigation was done on simple four layer test structures, which gives an insight in the physics of the latch-up effect. For latch-up predictions the circuit environment is essential, especially for power-on latch-up. The usefullness of technology and circuit measures for latch-up prevention considering power-on are discussed.

2 - LATCH-UP TEST STRUCTURE AND EQUIVALENT CIRCUIT AT POWER-ON

The latch-up test structure Fig.1a is identical to a CMOS inverter provided with an additional field oxide polysilicon gate. The total n^+p^+ spacing is 6µm (X=2µm and Y=4µm). Each diffusion of the test structure can be connected separately. Samples have been fabricated in a 1µm n-well CMOS technology.
Besides the 'functional unit' latch-up path, the environment of the latch-up path also has to

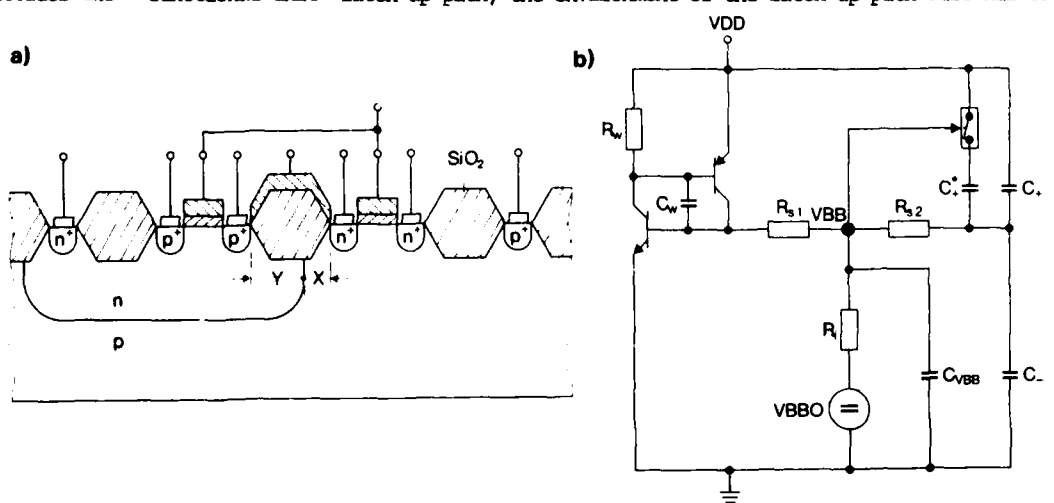


Fig.1: Latch-up test structure and equivalent circuit at power-on.
a.) Cross section of the test structure.
b.) Equivalent circuit of an n-well CMOS structure (DRAM) with the latch-up sensitive CMOS path, the cell area with cell plate generator and VBB generator /1/.

be considered. The conventional two transistor equivalent circuit is extended by elements describing the environment. As an example, Fig.1b shows the equivalent circuit of an n-well CMOS structure (DRAM, /1/) with the latch-up sensitive CMOS path, the cell area with the cell plate generator and the on chip VBB generator.

3 - LATCH-UP DURING POWER-ON

Fig.3 shows the power-on behaviour of an n-well CMOS structure with VBB generator. At the moment of power-on the substrate is floating because substrate bias can only be generated after a certain time delay. With floating substrate the effective substrate shunt resistance is nearly infinite. Therefore, the substrate potential VBB is determined through a capacitive voltage divider by the supply voltage VDD (see Fig.1b). VDD changes, especially the rise at power-on, are coupled into the substrate by C_+ (VBB-VDD) and buffered by C_- (VBB-VSS). The rising VDD leads to a positive peak on VBB. The level of this peak determines whether latch-up is triggered (Fig.2c) or not (Fig.2b).

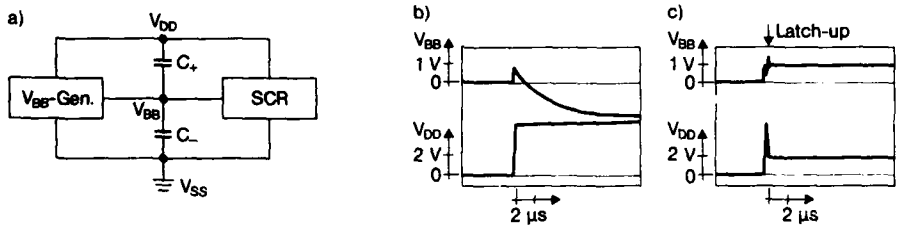


Fig.2: Power-on of an n-well CMOS circuit. a) measurement setup, b) VBB and VDD without latch-up, c) VBB and VDD with latch-up.

3.1 - CRITICAL CAPACITANCE RATIO FOR POWER-ON LATCH-UP

The necessary ramp rate to trigger power-on latch-up depends on the capacitances C_+ and C_-. Fig.3a shows the critical ramp rate (CRR) dV/dt vs. C_+ with VDD and C_- as parameters. For high C_+ (C_+>1nF) the CRR is independent of C_- ($dV/dt * C_+$ = const.). Decreasing C_+, a C_{+max} value is found, where the structure is latch-up free even for infinite ramp rate RR. C_{+max} decreases for increasing VDD, because for a given RR the charge built up by the displacement current increases with increasing final VDD level.

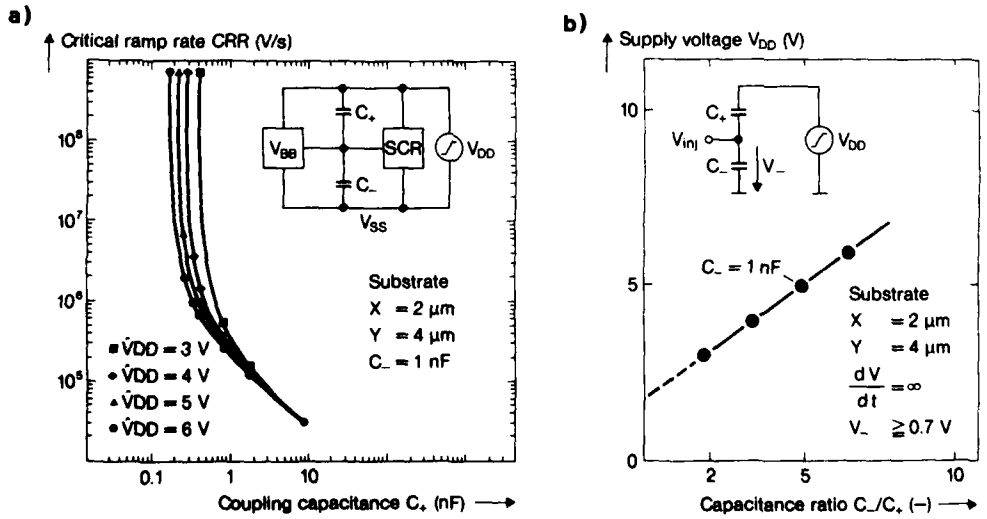


Fig.3: Power-on latch-up. a.) Critical ramp rate RR vs. C_+ with parameter VDD b.) minimum supply voltage VDD with infinite RR vs. ratio C_-/C_+.

Fig.3b shows the minimum VDD with infinite RR, which triggers latch-up, vs. the ratio C_-/C_+. Using the ratio C_-/C_+ as a variable, one single linear function describes the overall power-on latch-up behaviour.

3.2 - COMPARISON OF LOGIC AND DRAM

The maximum allowable RR and final VDD level for latch-up free power-on of circuits with VBB generator is determined by the ratio of C_-/C_+. In the case of LOGIC circuits a capacitance ratio C_+/C_- of 1/3 results at power-on. This ratio of 1/3 shows to be near the safe area. Additionally it has to be mentioned that in LOGIC circuits normally no VBB generator is used. The substrate is grounded to VSS.

In the case of DRAMs with cell plate bias (VDD or VDD/2) a typical capacitance ratio C_+/C_- of 20 is found. C_+ is mainly determined by the cell plate capacitance C_{cpl} (also called C_+^*).

4 - HOW TO AVOID POWER-ON LATCH-UP

The latch-up hardness of circuits can be improved by technology and circuit design (protection circuits).

4.1 - TECHNOLOGY

The use of an epilayer over highly doped substrate is well known to improve the static latch-up hardness. With grounded substrate the critical current to initiate latch-up is increased by orders of magnitude /3/. But for circuits with VBB generator (especially DRAMs) the power-on latch-up hardness is a main problem because of the increased effective substrate shunt resistance. The latch-up free regime for infinite RR in Fig.3b can not be improved. For finite RR and a given C_- the critical RR is even reduced by a factor of 2 with epi /3/. This is due to the higher current gain of the lateral parasitic bipolar transistor in the epi case. A comparison of the lateral bipolar transistor current gain for epi and non epi is shown in Fig.4.

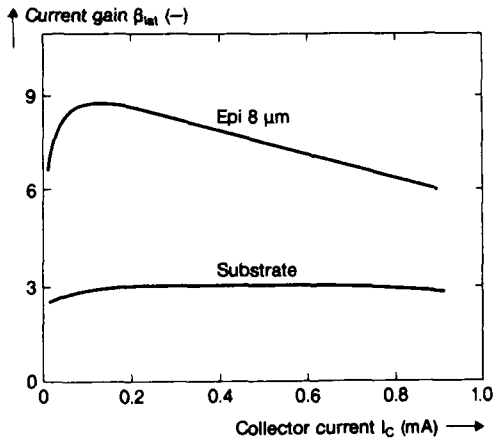


Fig.4: Current gain of the lateral bipolar transistor vs. the collector current I_C for epi and nonepi.

4.2 - PROTECTION CIRCUIT

In DRAMs the main part of C_+ is represented by C_{cpl}. A protection circuit reduces the effective C_+ by disconnecting C_{cpl} from VDD or the plate voltage generator during power-on (Fig.5). Fig.5b shows that the latch-up free regime is increased for a certain ratio C_-/C_+ so that no latch-up occurs.

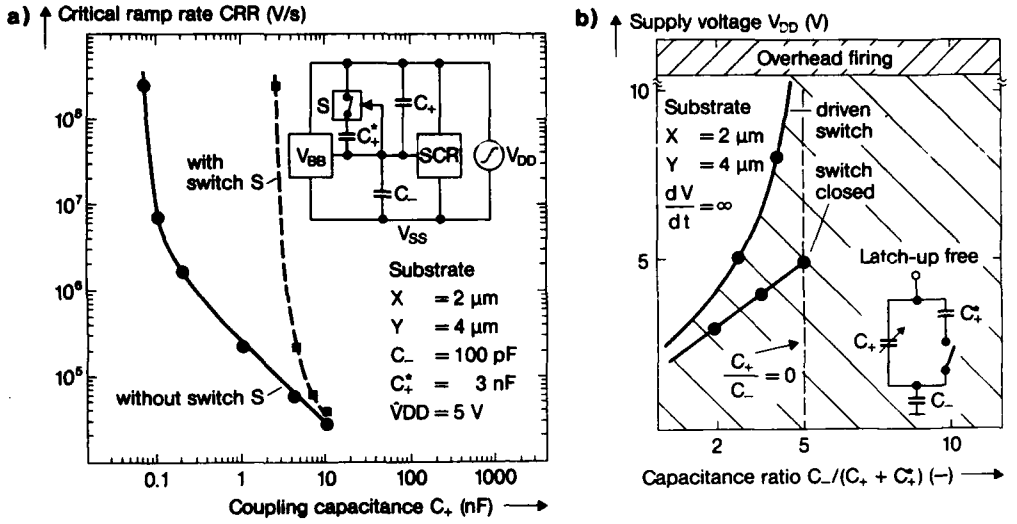


Fig.5: Power-on latch-up with and without protection circuit. a.) Critical RR vs. C_+ and b.) minimum supply voltage VDD with infinite RR vs. ratio C_-/C_+ .

5 - CONCLUSION

Besides the 'functional unit' latch-up path, which is represented by the latch-up test structure, the environment of this path also has to be considered. This is of special importance when the power-on latch-up hardness is investigated. Latch-up savety during power-on can be achieved by increasing C_-/C_+ or reducing VDD. For the case of a DRAM with VBB generator, the capacitance ratio C_+/C_- of 20 is leading to a high risk of latch-up during power-on. Epi, a unique measure for static latch-up suppression can even lead to a reduction of power-on latch-up hardness in circuits with on-chip generated substrate bias. In this case it is necessary to use protection circuits to achieve power-on latch-up free circuits.

ACKNOWLEDGEMENT

This work is based on a project which has been supported by the Minister of Research and Technology of the Federal Republic of Germany under the support-no. NT 2696. For the contents the authors alone are responsible.

REFERENCES

/1/ RECZEK W., PhD.thesis, Munich, FRG (1988).
/2/ MAZURE C., RECZEK W., TAKACS D., WINNERL J.,to be published in IEEE ED (1988).
/3/ TAKACS D., WINNERL J., RECZEK W., IEDM Tech. Dig., Washington, DC, December(1985) 504.

Session 1B Room B : Novel structures

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

NON-DESTRUCTIVE CHARACTERISATION OF DEVICE PROCESSING OF SILICON-ON-SAPPHIRE (SOS) WAFERS

C. PICKERING, S. SHARMA, S. COLLINS, A.G. MORPETH, G.R. TERRY and A.M. HODGE

Royal Signals and Radar Establishment, St Andrews Road, Malvern, GB-Worcs WR14 3PS, Great-Britain

Abstract - Spectrosopic ellipsometry has been used to obtain microstructural and layer thickness information from processed SOS wafers consisting of poly-Si/gate oxide/epi-Si/sapphire structures. Oxide thickness and interface roughness increased for 0.1 μm SOS wafers. Poly-Si formed by annealing α-Si deposited layers had the best dielectric function and similar structure and thicknesses were obtained on bulk and SOS wafers.

1 - INTRODUCTION

Processing of SOS wafers for CMOS devices involves the formation of a thin gate oxide (~400A) followed by deposition of a doped polycrystalline (poly-Si) layer (~4000A) for the gate contact. Thicknesses of these layers are usually estimated from similar layers grown on bulk monitor wafers. However, differences may occur due to the location of the wafers in the reactor or because of different oxidation and growth rates for bulk wafers. Standard measurement techniques used in silicon processing laboratories, eg single-wavelength ellipsometry or interferometry, cannot be used for SOS multilayer structures. In this work we report the first use of spectroscopic ellipsometry (SE) to measure directly the gate oxide and poly-Si thicknesses on SOS wafers. Comparison between bulk and SOS wafers indicates differences occurring as the Si epi-layer thickness is reduced. In addition, the depth profiling capability of SE has been used to obtain information on the rough surface overlayer, the degree of crystallinity of the poly-Si and its dependence on doping and annealing conditions.

2 - EXPERIMENTAL

A series of SOS wafers with Si epi-layer thicknesses of 0.1, 0.2, 0.3, 0.4 and 0.6 μm were thermally oxidised at 900°C in the same furnace to produce oxide thicknesses in the range 350-500A. Poly-Si gate contacts were produced on the oxidised 0.4 μm SOS wafers using two CVD deposition temperatures. Deposition at 580°C produced an amorphous Si (α-Si) layer while deposition at 617°C produced a small grain size (500-1000A) poly-Si layer. These layers were subsequently annealed at 900°C for 30 minutes in a $POCl_3$ atmosphere followed by 60 seconds HF etch. The high temperature doping step causes recrystallisation of the α-Si layer and an increase in grain size of the poly-Si layer, but the former has a more uniform distribution of large grains (1000A to 5000A) as assessed by TEM (A G Cullis et al, unpublished work). Similar processing steps were also performed on bulk Si wafers for comparison. Spectroscopic ellipsometry measurements were made at each step, over the wavelength range 230-850 nm with 75° angle of incidence, using a commercial SOPRA rotating-polariser instrument.

3 - RESULTS AND DISCUSSION

(i) Oxide Deposition Stage

Single-wavelength ellipsometry or Nanospec interferometry as normally employed at $\lambda = 633$ nm (He-Ne laser) cannot be used to measure the thickness of gate oxides on SOS, since the light penetration depth is greater than the epi-layer thickness (d). The wavelength-scanning facility of SE allows a wavelength region to be chosen where the penetration depth is less than d. For $\lambda = 230$-400 nm the light only penetrates a few hundred A into the underlying Si layer. This region has been used to make a direct comparison of oxides grown in the same furnace on bulk and SOS wafers. In addition to the oxide thickness, the theoretical fits to the spectra can be used to obtain an estimate of any non-stoichiometric or rough interface

layer (eg SiO_x).

Typical raw spectra are shown in Figure 1 which compares results from oxidised bulk and 0.1 μm SOS wafers, together with fitted curves. At about λ = 0.38μm the beginning of interference fringes can be seen as the light penetrates through to the sapphire substrate. The peaks in tan ψ and cos Δ are clearly shifted to longer wavelengths for the SOS wafer, providing an immediate visual indication of increased oxide thickness. In addition, the fits indicate a much higher interface thickness (modelled as SiO, a physical mixture of 1:1 c-Si:SiO_2) for the 0.1 μm SOS wafer. The variation of SiO_2 and interface thickness with SOS epi-layer thickness is shown in Figure 2. For d > 0.2 μm the total oxide thickness is the same for bulk and SOS wafers, although the interface is a higher proportion of the total thickness in the SOS case. The SiO interface observed here is presumably due to microroughness at the initial growth interface and the thicker interface for SOS compared to bulk wafers probably reflects the rougher initial surface of SOS /1/. However, for d ⩽ 0.2 μm, both the SiO_2 and interface thicknesses increase, with a dramatic increase for 0.1 μm SOS. For this epi thickness, the total thickness has increased by ~ 6% with the interface roughness almost doubled. The increased oxide thickness might be expected due to the greater density of (221) microtwin defects occurring within ~ 0.1 μm of the sapphire interface /2/. Any deviations of orientation from <100>, the slowest oxidation plane, would result in enhanced oxidation rates and the high density of microtwins may also produce a rougher surface and hence a thicker interface layer when oxidised.

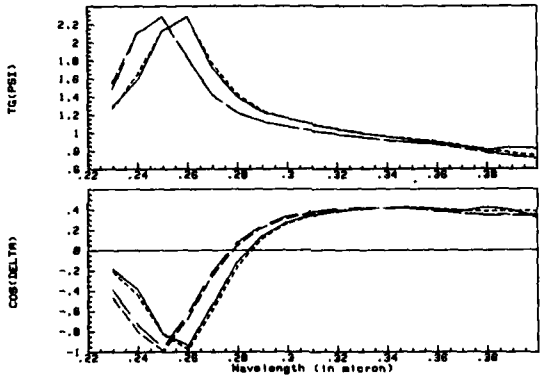


Fig 1 - Tan ψ, cos Δ spectra for oxidised bulk (—— ——) and 0.1 μm SOS wafers (——). Fitted curves:- 447A SiO_2/7A SiO/c-Si (— — —); 456A SiO_2/28A SiO/c-Si (- - - -).

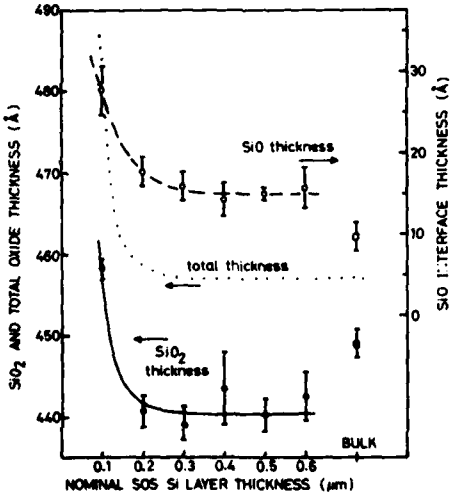


Fig 2 - SiO_2 (•) and SiO (o) interface layer thicknesses for oxidised SOS wafers with various epi-layer thicknesses. Error bars indicate standard deviation of cross-wafer variation.

(ii) Poly-Si Deposition and Annealing Stage

SE measurements in the short wavelength region for the poly-Si deposited wafers can be used to obtain information on the microscopically rough surface overlayer and the crystallinity of the poly-Si. The dielectric function thus calculated for the poly-Si is then used for the complete spectrum to obtain thicknesses of the multi-layer structure, ie the poly-Si and gate oxide as well as the Si epi-layer thickness. Figure 3 shows pseudo-dielectric function, $\langle\varepsilon\rangle = \langle\varepsilon_r\rangle + i\langle\varepsilon_i\rangle$, spectra for poly-Si layers produced on oxidised 0.4 μm SOS wafers. Identical results were obtained on the bulk Si comparison wafers. This series of curves provides a visual indication of the change in poly-Si structure as a function of deposition and annealing conditions. The best, ie highest magnitude $\langle\varepsilon_i\rangle$ result is obtained for the annealed layer deposited in amorphous form at 580°C. This spectrum agrees well with that for a c-Si sample with 25A native oxide except in the 0.36 μm region, where effects of the high P doping will be important. The 580°C deposited layer is clearly amorphous since no structure is observed at 0.295 or 0.36 μm, while the 617°C layer is small-grained poly-Si (a combination of α-Si, c-Si and voids) in agreement with TEM results. The $\langle\varepsilon_i\rangle$ values for this layer after annealing are improved but are not as good as the annealed 580°C layer, indicating a higher concentration of voids and/or a rougher surface. This sequence of poly-Si quality as a function of deposition and annealing conditions is in good agreement with TEM results and confirms the data obtained previously by Harbeke et al /2/. Detailed fits to the spectra indicate 40% c-Si, 40% α-Si, 20% voids for the 617°C as-deposited poly-Si layer, improving to 87% c-Si, 6% α-Si, 7% voids on annealing. However this compares with 98% c-Si, 1% α-Si and 1% voids for the annealed 580°C layer (see Fig 4), confirming that the lower temperature deposition produces a better quality layer on annealing.

The spectra for the poly-Si layers deposited on oxidised bulk wafers have been analysed in detail for layer thicknesses for comparison with Nanospec results. Figure 4 shows the result for an annealed 580°C layer and the quality of fit can be seen to be very good. The poly-Si layer thicknesses agreed with Nanospec readings within ±10%. Best agreement (within 4%) was obtained for the layer with poly-Si optical constants closest to c-Si. However for the other layers the Nanospec appears to underestimate the thickness when voids are present and overestimate when α-Si is present. This would be expected due to the assumption of a constant refractive index by the Nanospec analysis program, and indicates that inaccurate results will be obtained if the poly-Si structure is variable or significantly different to c-Si.

Individual layer thicknesses can also be obtained from the SOS multi-layer structures. Figure 5 shows a spectrum in the region beyond 0.4 μm where the light penetrates the complete structure. Fits to the spectra give poly-Si and epi-Si layer thicknesses with 90% confidence limits ~ 5%. Thicknesses of poly-Si layers produced under the same conditions on bulk and SOS wafers agreed within ~80A. The Si epi-layer thickness of SOS wafers was found to vary considerably but was within the specified tolerance of ±10% when allowance was made for the material converted to oxide. The two theoretical curves in Figure 5 show the sensitivity to the gate oxide thickness. This mainly affects the peaks at 0.47 and 0.62 μm and can be determined with a 90% confidence limit of ±20%. To determine this thickenss, measurements must be made at the 75° angle of incidence used here, since the 0.62 μm peak does not occur at other angles.

4 - CONCLUSIONS

Spectroscopic ellipsometry has been used to obtain microstructural and thickness information non-destructively from processed SOS wafers which cannot be obtained by standard methods. Thus the technique has potential as an on-line process monitor.

REFERENCES

/1/ Pickering C, Dixon S, Gasson D B, Robbins D J and Hodge A M, J Crystal Growth 84 (1987) 180.
/2/ Harbeke G, Krausbauer L, Steigmeier E F, Widner A E, Kappert H F and Neugerbauer G, Appl Phys Lett 42 (1983) 249.

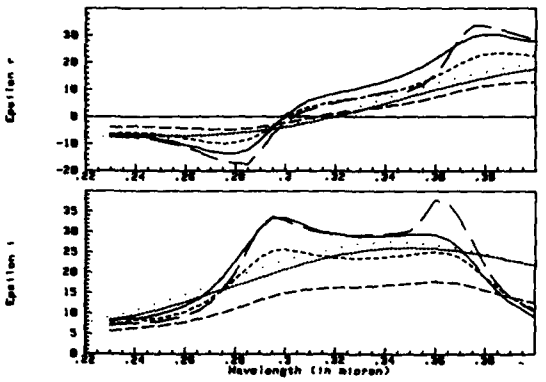


Fig 3 - $\langle\varepsilon\rangle$ spectra for poly-Si layers deposited on oxidised SOS wafers (580°C: as-deposited (- - - -), annealed (——); 617°C: as-deposited (—— ——), annealed (— — —); c-Si + 25A SiO_2 (—— ——); -Si (•••••))

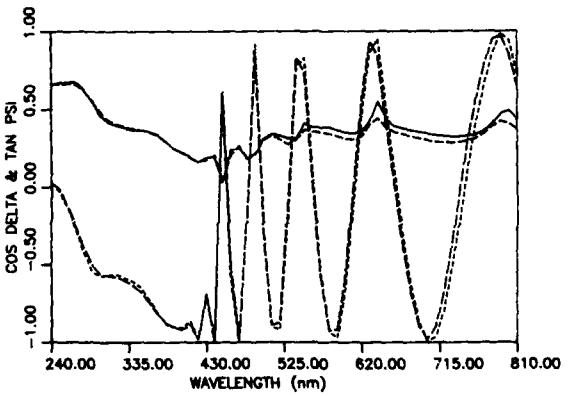


Fig 4 - Tan ψ, cos Δ spectra for poly-Si layer deposited on oxidised bulk wafer (— — —, —— —— experiment; ——, - - - - theoretical:- 82A overlayer with 11% voids, 3450A poly-Si (98% c-Si, 1% α-Si, 1% void), 538A SiO_2, c-Si).

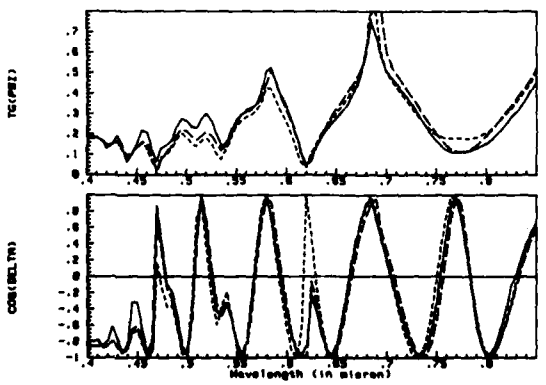


Fig 5 - Tan ψ, cos Δ spectra for poly-Si layer deposited on oxidised SOS wafer (experiment (——); calculated curves using 300A (- - - -), 400A (— —) oxide layer with 4120A poly-Si and 3420A c-Si epi-layer on sapphire substrate).

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

OSIRIS II, A TWO-DIMENSIONAL PROCESS SIMULATOR FOR SIMOX STRUCTURES

I. SWEID, N. GUILLEMOT and G. KAMARINOS

INPG/ENSERG, Laboratoire de Physique des Composants à Semiconducteurs, CNRS-UA 840, 23, Av. des Martyrs, F-38031 Grenoble Cedex, France

Résumé - La simulation des processus technologiques est très importante pour la fabrication des circuits intégrés. Dans cet article, les auteurs présentent OSIRIS II simulateur technologique sur SIMOX. Le programme prend en compte les conditions limites propres à cette structure. Les modèles utilisés sont brièvement rappelés. Un exemple de réalisation d'un dispositif N-MOS sur SIMOX est simulé par le programme.

Abastract - The simulation of technological processes is very important for the fabrication of integrated circuits. In this paper, the authors present OSIRIS II for the simulation on SIMOX structures. The program takes into account the limiting conditions associated with this material. The models used are briefly recalled. Finally, a complete simulation of an N-MOS device on SIMOX is presented.

1 - INTRODUCTION

Although many two-dimensional process simulator for bulk silicon do actually exist, they are not general enough to be able to simulate these process on Silicon-On-Insulator (SOI) structures. In this paper, we present OSIRIS II, a simulator for SIMOX (Separation by IMplanted Oxygen) structure [1]; it allows the complete simulation of device fabrication and includes the following steps :
- resist deposition or removing,
- oxide and nitride deposition or etching,
- ion-implantation of boron, BF_2, arsenic, phosphorus and antimony,
- growth of thin oxides and growth of the field oxide in the case of a LOCOS isolation,
- diffusion of impurities in an inert or oxidizing ambient.

2 - ION-IMPLANTATION

In order to simulate ion-implantation in silicon with a mask, we used the conventional convolution of a vertical distribution with a Gaussian lateral function. The vertical function depends on the impurity implanted : two joint half-Gaussian's for arsenic, phosphorus and antimony, and a Gaussian or a Pearson IV function for boron and BF_2 channeling is taken into account with an exponential tail.
The edge of the masking layer (oxide, resist ...) can be of arbitrary shape : the mask is then considered as an equivalent silicon layer and the resulting profile is obtained numerically.

3 - OXIDE GROWTH

To simulate the one-dimensional oxide growth, the very well-known Deal and Grove model is used, where the influence of the ambient (dry or steam), temperature, pressure and chlorine percentage are included in the linear and parabolic parameters. The thin oxide regime is simulated precisely with the model developed by Fargeix et al. [2].
In the case of LOCOS isolation, the field oxide growth is simulated by the analytical model proposed by N. Guillemot et al. for bulk silicon [4]. When the field oxide rejoins the buried oxide it is assumed that the buried oxide acts as a thermal oxide, and its growth follows the Deal and Grove model.

4 - DIFFUSION OF IMPURITIES

The simulation of impurity redistribution during thermal annealing (inert or oxidizing) involved the resolution of the diffusion equation in a domain which is variable in time; it is rectangular at first and gradually becomes triangular at the end of the process (Fig.1).

To avoid numerical problems, a "mixed" solution is used :
- In the isolation region (region I), where the impurity concentration is low, the diffusion equation is solved analytically [4,5]. - In the source/drain and channel regions (region II) the diffusion equation is non-linear : electric field drift, concentration dependence and impurity clustering must all be taken into account. A numerical solutions is thus used : the physical domain is mapped into a fixed-time and invariant rectangular domain by means of coordinate transformation. Then, the so-obtained differential equations are discretized with the classical method of finite-differences and a Crank-Nicolson scheme. The solution is obtained using the Gauss-Seidel method [6].
"Natural" boundary conditions are applied between the analytical and the numerical regions (the concentration obtained both analytical and numerically must be the same).
As studies on impurity diffusion in SIMOX structures are still not fully completed [4], diffusion models in OSIRIS II are the same as those used for bulk silicon in the one-dimensional program, SUPREM III.

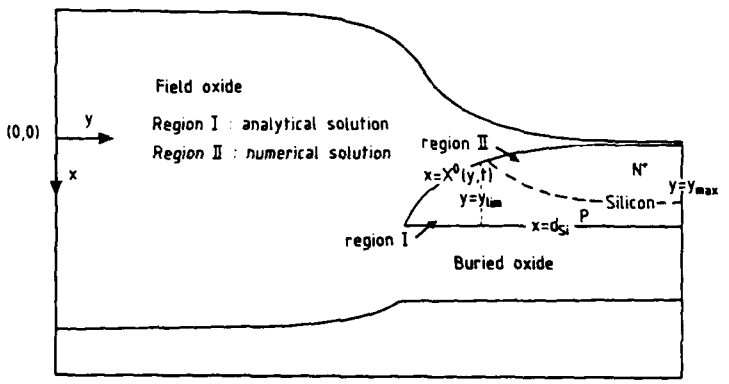


Fig.1 - Fig.1 - Geometries of silicon domain Ω when a mixed solution is used to solve the diffusion equation.

5 - SIMULATION OF A MOS DEVICE ON A SIMOX STRUCTURE

As a conclusion, OSIRIS II is used to simulate the processing sequence for a MOS device in Fig.2. The process schedule is the following :
- Field implant with boron (100kev, 10^{13} cm^{-2}).
- Field oxidation (270mn in steam ambient at 1000°C).
- Etching of 0.05µm of oxide.
- Dry oxidation (100mn, 950°C, dry, 3% HCL).
- Implantation of boron (10^{12} cm^{-2}, 30Kev) thus adjusting the threshold voltage.
- Implantation of source and drain with arsenic (10^{15} cm^{-2}, 120Kev).
- Aneal process (950°C, 30mn).

The surface topographies and corresponding equidensity contours for impurity concentrations at the end of the process are given in figure 2.

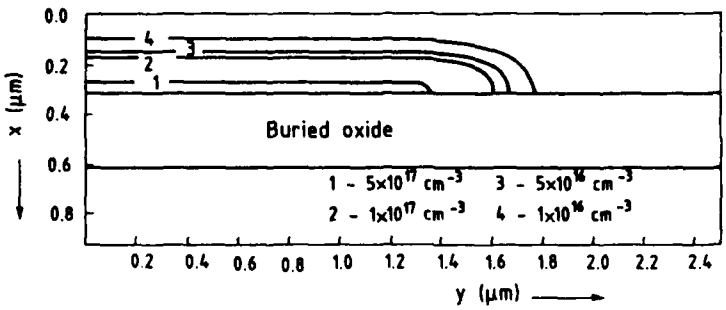


Fig. 2(a) - The implantation of a low dose of boron (10^{13} cm^{-2}, 100Kev) in the isolation region.

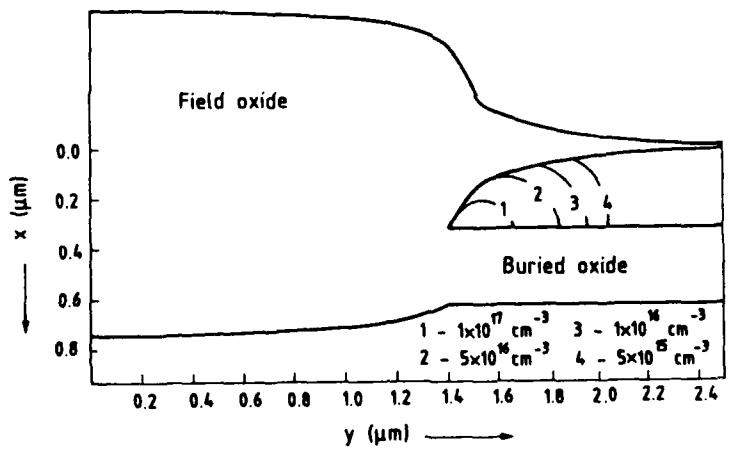


Fig. 2(b) - Equidensity contours after the field oxide formation (270mn in steam ambient at 1000°C).

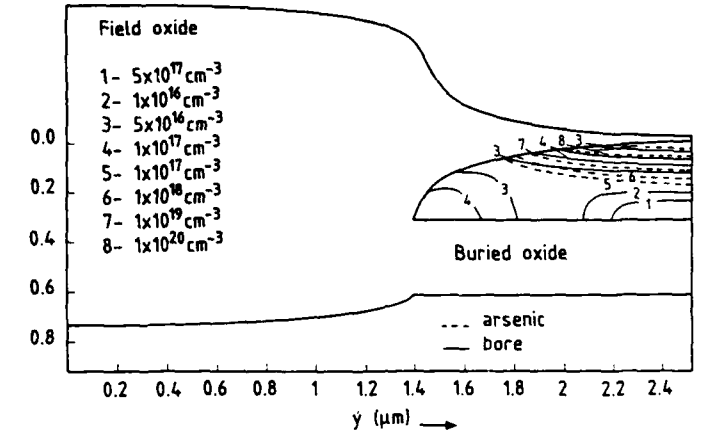


Fig. 2(c) - Equidensity contours after the implantation of boron ($10^{12}cm^{-2}$, 30Kev) in the channel region and arsenic ($10^{15}cm^{-2}$, 120Kev) in the source/drain region.

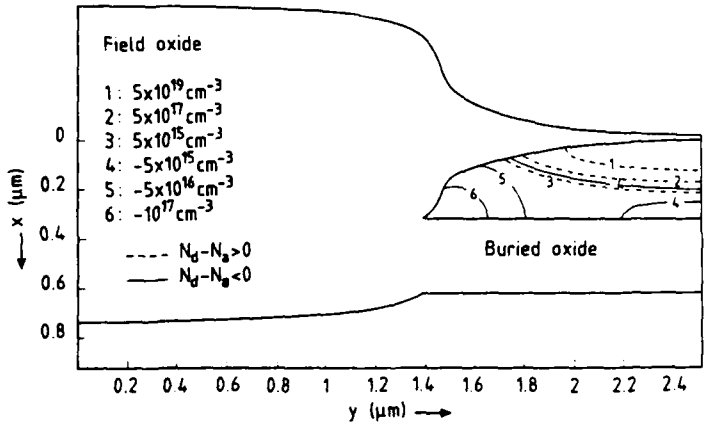


Fig. 2(d) - Equidensity contours of the net concentration after activation annealing (950°C, 30mn).

REFERENCES

/1/ Sweid, I., Guillemot, N., Kamarinos, G.,Eur. Silicon On Insulator Workshop (Meylan, France), 1988
/2/ Fargeix, A., Ghibaudo, G., Kamarinos, G., J. Appl. Phys. 54, (1983)2878.
/3/ Guillemot, N., Pananakakis, G., Chenevier, P., IEEE Trans. Electron Devices ED-34 (1987) 1033.
/4/ Sweid, I., Guillemot, N., Kamarinos, G., J. Appl. Phys. (to be published).
/5/ Sweid, I., Guillemot, N., Kamarinos, G., Rev. Phys. Appl. (to be published)
/6/ Guillemot, N., Chenevier, P., Deroux-Dauphin, P., Gonchond, J.P., Rev. Phys. Appl. (1984) 987

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

ELECTRONIC PROPERTIES OF SILICON INTERFACES PREPARED BY DIRECT BONDING

S. BENGTSSON and O. ENGSTRÖM

Chalmers University of Technology, Department of Solid State Electronics, S-412 96 Göteborg, Sweden

Abstract - The influence of interface charges on the properties of Si-Si and Si-SiO_2 interfaces prepared by direct bonding has been investigated. Surface potentials of N-N and P-P interfaces and recombination currents in P-N junctions depend on surface and heat treatments. In both cases lower magnitudes were measured in samples pre-treated in HF compared to samples pre-treated in HNO_3. Bonded Si-SiO_2 interfaces with interface state densities of about 10^{11} cm^{-2} eV^{-1} and low flatband voltages have been achieved.

1 - INTRODUCTION

The technique of directly bonding surfaces of materials together included in the silicon technology [1,2] has given rise to new possibilities for device geometries. Interfaces obtained by this kind of preparation have been included in novel insulating structures [2,3,4] and in discrete power devices [5,6]. In the works performed up to this date, the main efforts have been to demonstrate the mechanical properties of these interfaces, while the electronic behavior has been investigated only to limited extent. An interesting question, therefore, is whether this technique can be used in cases where device data critically depend on the electronic interface properties. The presence of interface charges accompanied by potential barriers and charge carrier recombination is the main factor limiting this possibility. In this work we have investigated the electronic properties of interfaces prepared by direct bonding of silicon wafer surfaces of both N- and P-type including materials with different bulk resistivities. Interfaces of the combinations N-N, P-P and P-N have been studied together with the bonded interface of the important combination SiO_2-Si.

2 - SAMPLE PREPARATION

Polished silicon wafers with flatness and warp similar to that of wafers used in integrated circuit production were cleaned by standard procedure in hot H_2O_2-NH_3 and H_2O_2-HCl solutions. This was followed by a treatment in HNO_3 or 2% HF:H_2O, water rinse and drying. The polished surfaces of wafer pairs were brought together using vacuum in a specially designed fixture. After releasing from the fixture, the wafers held together by adhesion forces, were loaded into an annealing furnace at room temperature. The furnace temperature was slowly ramped to the region 1000 - 1100 °C, where the samples were annealed for 1-2 hours, in an atmosphere of oxygen or nitrogen.

3 - EXPERIMENTAL RESULTS

Impurities and crystalline defects present at the bonded interface are expected to create charges and potential barriers. A charge Q at an interface between two semiconductors with a shallow doping concentration N creates an interface potential $V \sim Q^2 / N$. This means that the sensitivity to interface charges is less pronounced at higher doping levels. For resistivities higher than about 1 Ωcm, interface charge densities higher than about 10^{11} cm^{-2} are expected to cause considerable energy barriers (S. Bengtsson and O. Engström, unpublished). Further, small lateral variations in interface charge may cause considerable lateral variations in potential.

3.1 N-N AND P-P INTERFACES

Using spreading resistance measurements across an N-N interface lapped to a small angel for 10 Ωcm silicon an interface potential barrier was found. The result, shown in fig. 1, exhibit a clear peak in spreading resistance due to the interface potential barrier.

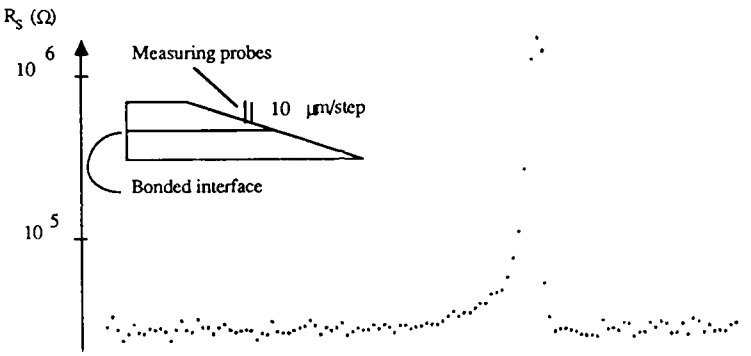


Fig. 1 - Spreading resistance measured across a bonded N-N-interface prepared of two 10 Ωcm silicon wafers and with HNO_3 used as prebond treatment.

The lateral distribution of the interface potential was investigated by scanning the wafer surface with a He-Ne laser beam and the photocurrent at the interface was measured as a function of the position of the laser spot [7]. No bias was applied during the measurement. The variation in photocurrent, displayed in fig. 2 demonstrates the lateral variations in interface potential. The influence of surface effects was negligible, as found from measurements after repeated surface treatments. This variation in interface potential gives current-voltage characteristics with a much smaller curvature than expected from a structure with a simple energy barrier.

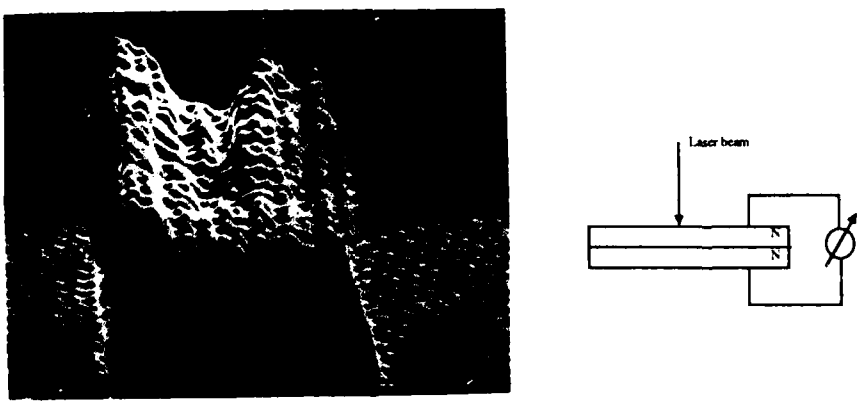


Fig. 2 - The lateral variation of photocurrent of a bonded N-N-interface.

The differential resistance at low voltage was used to estimate the influence of different surface and heat treatments on the interface charge. Fig. 3 displays the differential resistances before and after a post-bonding anneal at 1000 °C during about 90 minutes for two different pre-bond treatments and for different silicon resistivities. For samples with low resistivities, a considerable decrease in differential resistance after the heat

treatment indicates a decreasing interface charge for the samples pre-treated in HNO_3. The samples pre-treated in HF exhibit a considerably smaller potential barrier, already after the wafer bonding procedure, and it was not necessary to give these samples a post-bonding anneal. These results shows that the native oxide layer present on silicon surfaces after treatment in HNO_3 is important in the generation of defects influencing the interface resistance.

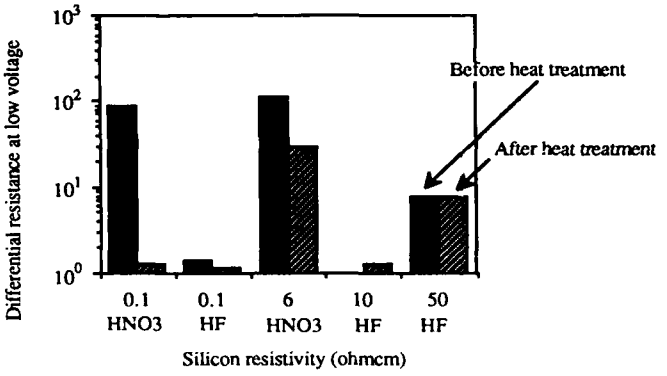


Fig. 3 - Differential resistance at low voltage for some samples with bonded N-N-interface as a function of silicon resistivity for different surface and heat treatments.

3.2 P-N INTERFACES

PN-junctions were prepared by bonding low ohmic P^+-type wafers of resistivities < 0,1 Ωcm to N-type wafers of 1-5 Ωcm resistivity. Current voltage characteristics in the forward direction are shown in fig. 4 for two samples with surface treatment before bonding using HF and HNO_3 respectively. The n-factors in the expression $I=I_0\exp(qV/nkT)$ were found to be in the region of 1,5 - 7, with the lowest values for samples treated in HF. This reflects a higher recombination current in samples treated in HNO_3, again indicating that surface treatment in HF before bonding of two silicon surfaces gives a lower density of electron states at the bonded interface than does a corresponding treatment in HNO_3. Breakdown voltages in the reverse direction of about 75 V were found for bonded P^+N-interfaces using 2-3 Ωcm N-type wafers.

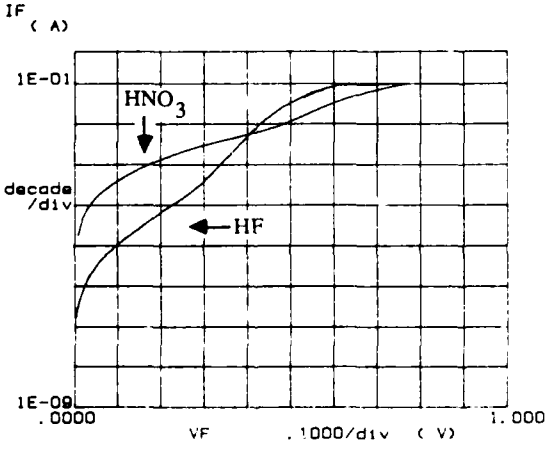
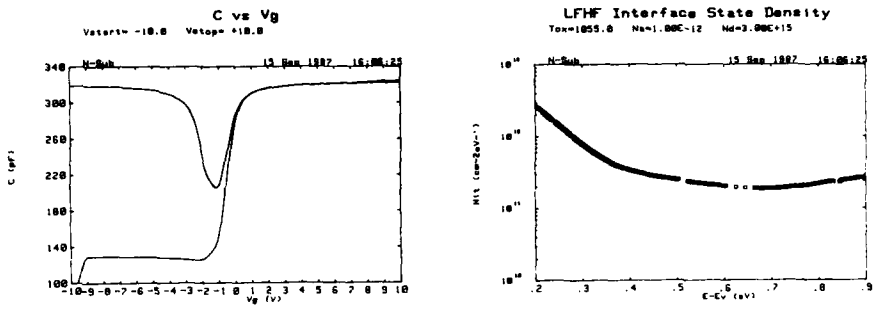


Fig. 4 - Current voltage characteristics of two bonded P^+N-junctions treated in HF and HNO_3, respectively, before wafer bonding.

3.3 MOS-STRUCTURES

MOS-structures were prepared by bonding a thermally oxidized silicon wafer to a bare silicon wafer. After this, the bulk silicon of the oxidized wafer was etched off using a selective silicon etch, leaving a

silicondioxide-silicon structure with a bonded interface. Aluminium contacts were evaporated onto both sides of the structure to make MOS-capacitors and the interface properties were characterised using CV-technique. Usually low flatband voltages and low midgap interface state densities at and above 10^{11} cm^{-2} eV^{-1} were found. Fig. 5 shows measured CV-curves and calculated interface state density for a MOS-capacitor with bonded interface. Reference capacitors with thermally grown interfaces oxidized in the same furnace as the wafers used in the bonding procedure, had interface state densities in the range 10^{10} to 10^{11} cm^{-2} eV^{-1}.



a) b)

Fig. 5 - a) HF and LF CV-curves for a MOS-capacitor with bonded SiO_2-Si interface.
b) Interface state density obtained from the curves in a).

4 - DISCUSSION

For N-N and P-P interfaces the interface charges were reduced by heat treatment or by using HF as surface treatment before wafer bonding. In fact, the lower differential resistances obtained after these treatments were of the same magnitudes as the contact resistances of the samples. Estimations of the interface charges based on these results give concentrations in the region below 10^{11} cm^{-2}. In a recently published work [1], low interface potential barriers were found for bonded P-P interfaces of samples with bulk resistivities of 0,004 Ωcm. However, for such low resistivities only interface charges of about 10^{13} cm^{-2} are expected to give a measurable interface barrier. It was also possible to obtain P^+N-junctions with low recombination currents. For P^+N-junctions prepared by direct bonding technique, the P^+-region may be expected to diffuse into the N-region, creating a P^+N-junction at a certain distance from the mechanical interface. As seen from fig. 4, this interface still has a certain influence on the properties of the PN-junction, but can be controlled by the chemical surface treatment before bonding. Finally, we have shown that SiO_2-Si interfaces can be prepared with low interface state densities and low flatband voltages. This demonstrates that silicon interfaces obtained by direct bonding can be utilized in cases where the electronic properties are crucial for the behaviour of a device.

5 - ACKNOWLEDGEMENTS

This project is financed and supported by the Swedish National Board for Technical Development. We also wish to thank Dr. M.S. Shivaraman and ABB Drives for assistance.

6 - REFERENCES

[1] M.Shimbo, K.Furukawa, K.Fukuda, K.Tanzawa, J. Appl. Phys. 60 (1986) 2987.
[2] J.B.Lasky, Appl. Phys. Lett. 48 (1986) 78.
[3] J.Ohura, T.Tsukakoshi, K.Fukuda, M.Shimbo, H.Ohashi, IEEE Electr. Dev. Lett. 8 (1987) 454.
[4] Xu Xiao-Li, Tong Qin-Yi, Xong He-Ming, Proc. ESSDERC 1987, 403 (Bologna).
[5] A.Nakagawa, K.Watanabe, Y.Yamaguchi, H.Ohashi, K.Furukawa, 1986 IEEE IEDM Tech. Dig. 122
[6] H.Ohashi, K.Furukawa, M.Atsuta, A.Nakagawa, K.Imamura, 1987 IEEE IEDM Tech. Dig. 678.
[7] O.Engström, B.Drugge, P.A.Tove, Physica Scripta 18 (1978) 357.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

THEORETICAL ANALYSIS OF THE TWO-TERMINAL MOS CAPACITOR ON SOI SUBSTRATE

P. PAELINCK[1], D. FLANDRE[1], A. TERAO[2] and F. VAN de WIELE

Laboratoire de Microélectronique, Université Catholique de Louvain, Place du Levant 3, B-1348 Louvain-la-Neuve, Belgium

Résumé - Des modélisations analytiques et numériques de la capacité métal-oxyde-semiconducteur sont proposées pour les substrats sur isolant comportant un film flottant. Le traitement théorique repose sur des considérations physiques et met en évidence l'influence de la distribution des diverses charges à l'intérieur de la structure globale. Les caractéristiques capacité-tension sont directement interprétables à partir des mécanismes physiques internes du dispositif.

Abstract - The present paper deals with the analytical and numerical study of MOS capacitors on SOI substrates. The major concern is the theoretical treatment of such devices with floating substrate. Special care has been taken to keep the mathematical modeling consistent with physical considerations. The importance and the role of the underlying silicon substrate have been highlighted. The different operating conditions of the capacitance-voltage characteristic have been interpreted in terms of device intrinsic physical mechanisms.

1 - INTRODUCTION

Owing to its simplicity of fabrication and analysis, the MOS capacitor is now intensively used for the control and the study of the electrical properties of MOS circuits integrated in the bulk.On the contrary, in SOI technology, alternative techniques to standard C(V) measurements have been proposed for characterizing the MOS system [1,2]. However, the structures that are considered in these experiments are 3 or 4 terminals devices, allowing to get rid of problems related to the floating substrate. Unfortunately, the need for minimizing parasitic capacitance effects (w.r.t. pure capacitance measurements) results in an increased complexity in the method for extracting the material parameters. Our purpose, guided by physical reasoning, is to propose accurate analytical and numerical modelings of the two-terminal MOS/SOI capacitor (Fig. 1). We will show that the capacitance characteristic is somewhat different from that of the classical bulk MOS capacitor. The discrepancies will be explained from the charges distribution within the whole structure. The influence of several technological and electrical parameters will be discussed.

The study is based on the one-dimensional solution of the Poisson equation. Since no current flows through the capacitor, the Fermi potentials are constant but different in the film and in the substrate. This implicitly assumes that a gradient of the Fermi potential exists within the buried oxide. Nevertheless, since the carriers densities vanish in the insulator, there is no current in the oxide. However, the Fermi potential in the SOI film can not be a priori known since there is no physical contact on the latter. Moreover, owing to the fact that the film is sandwiched between two insulators, carriers are impedded from leaving the film. These physical considerations lead to state that, whatever the voltage applied to the gate, the overall charge in the floating SOI layer remains constant and equal to the charge stored in the film at equilibrium. This constant charge depends on the doping profile, the interface charges and the geometrical characteristics of the structure. This is to say that, for any gate voltage, any variational charge appearing underneath the front gate must be compensated by an other charge (of opposite sign) above the back interface so as to keep constant the integrated space charge density in the film.The low-frequency capacitance-voltage characteristic is then easily derived by calculating the variation of the potential drop across the gate oxide with respect to the gate voltage.

---

(1) P. Paelinck and D. Flandre are granted by I.R.S.I.A.
(2) A. Terao is a research assistant of the National Fund for Scientific Research

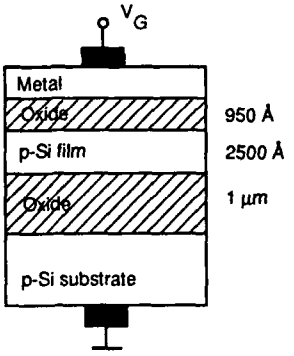


Fig. 1 - The MOS capacitor on SOI substrate

2 - MODELING

2.1. Analytical modeling

The analytical model is limited to fixed oxide charges and uniform doping levels. An analytical model for an ideal SOI MOS capacitor (zero interface charges, zero work function differences) has recently been published [3]. It points out that, in contradiction with classical approximations [4], the SOI MOS capacitor analysis requires to take into account the non zero width of the accumulation layer in the SOI film as well as the space charge region in the underlying non metallic silicon substrate.
Our model has now been extended to non zero interface charges and work function differences.
In order to solve the one-dimensional Poisson equation in the SOI film, we need an additional boundary condition, which depends on the following situations :
- as long as front and back space charge regions are not in interaction, we assume that they are separated by a quasi-neutral region
- when they interact, the sum of the widths of the front and back space charge regions is a constant equal to the film thickness; in that case, we have to approximate the different space charge region widths[5] and these constitute the single analytical approximations.
Finally, the analytical model consists in a set of nonlinear equations which can be solved using a classical iterative method.

2.2. Numerical modeling

Numerical modeling takes all the capacitor characteristics into account : film thicknesses, nonuniform impurities concentrations, fixed oxide charges and fast interface states densities.The strategy relies on a two-step procedure [6]. In a first phase, the structure under equilibrium conditions is simulated in order to determine the integrated charge in the SOI film. Then, for other gate biases, the Poisson equation is solved at each node of the discretization grid. The related set of equations is to be augmented with an integral conservation constraint, allowing to determine the unknown Fermi potential. This additional relationship requires that the overall variational charge must vanish in the floating substrate for any gate voltage. The latter entails a reasonable increase in the system matrix complexity, provided that the equations are reordered. The resulting system of equations is easily solved using a gaussian elimination.

3 - RESULTS

The first structure that has been investigated includes interface states densities at the three silicon dioxide/silicon interfaces. The impurities concentration in the film is one order of magnitude higher than in the substrate. The C(V) curve is quite dissimilar to the typical characteristics exhibited by classical bulk capacitors (Fig. 2).

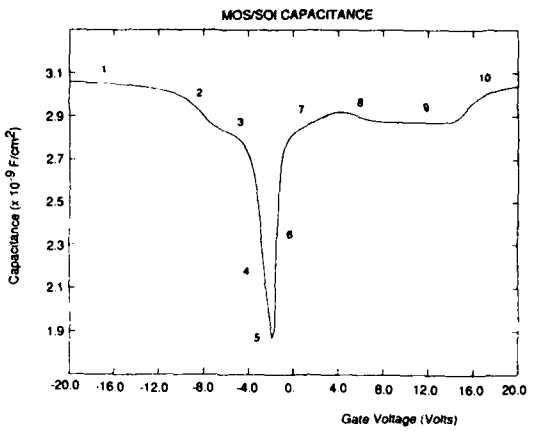


Fig. 2 - Capacitance-Voltage curve ($N_{A,film}=10^{15}$ cm^{-3}, $N_{A,subs}=10^{14}$ cm^{-3}) ($N_{ssf}=5.10^{10}$,$N_{ssb}=10^{11}$ and $N_{sss}=5.10^{10}$ cm^{-2}) (t_{film}=2500 Å)

This is due to the various charges appearing in the structure. The basic keypoint lies in the fact that the different interfaces reach accumulation, depletion and inversion for different gate voltages. In Table 1 are listed the charges controlled by the three interfaces for different operating conditions of the capacitor as referred in Fig. 2 (a=accumulation, d=depletion, w.i.=weak inversion, i=inversion).

Case	1	2	3	4	5	6	7	8	9	10
Front interface	a	a	d	d	w.i.	w.i.	w.i.	i	i	i
Back interface	i	i	i	i	i	i	i	w.i.	a	a
Substrate	a	a	a	d	w.i.	i	i	i	i	i

Table 1 - Modes of operation of the MOS/SOI capacitor

The charge in the substrate depends on the gate voltage in the same manner as in the classical capacitor . However, charges in the SOI film may interact or not, depending on the film thickness, the doping level and the interface charges densities. Under non equilibrium conditions, it can be observed that the overall variational charge in the floating film is zero. This is responsible for the modulation of the electrical potential at both front and back interfaces. Fig. 3 shows the influence of the substrate doping level on the capacitance. For substrate doping levels lower than the SOI film doping level, the capacitance depends on the substrate behaviour only over a small range of voltages around the minimum value. When the impurities concentration increases, the influence of the charge stored in the underlying substrate tends to modulate the C(V) curve over a wider gate voltage range. This explains why the silicon substrate must be included in the model instead of replacing it by a metallic gate.

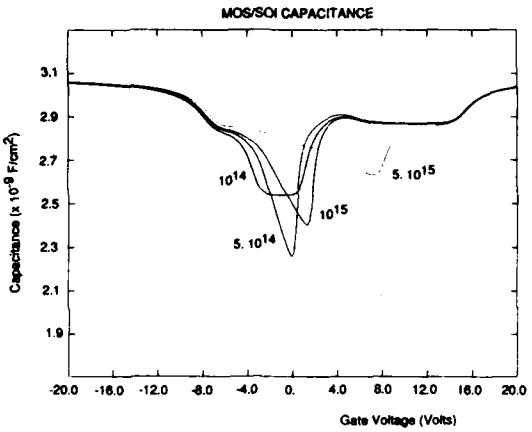


Fig. 3 - Capacitance-Voltage curve ($N_{A,film}=10^{15}$ cm^{-3}, $t_{film}=2500$ Å)
($N_{ssf}=5.10^{10}$, $N_{ssb}=10^{11}$ and $N_{sss}=5.10^{10}$ cm^{-2})

4 - CONCLUSION

A consistent theoretical treatment of MOS capacitor on SOI substrate has been proposed. We have shown that, under non equilibrium conditions, the Fermi potential within the floating substrate has to adapt itself so as to maintain the charge stored in the film at equilibrium. At present time, no experimental results concerning the MOS/SOI capacitor as a two-terminal device have been yet reported. This could be attributed to the quite unusual shape exhibited by the capacitance-voltage characteristic of this structure. From this point of view, the above-mentioned approach gives some insight into the major physical mechanisms and enables to state clearly the correspondence between the various modes of operation of the capacitor and its electrical behaviour. However, it is still to be determined whether the MOS/SOI capacitor can be used as a characterization device as in bulk technology.

REFERENCES

[1] Lee, J.-H. and Cristoloveanu, S., IEEE Electron Device Letters Vol. EDL-7 (1986) 537-539
[2] Vu, D.P. and Pfister, J.C., Applied Physics Letters 48 1 (1986) 50-52
[3] Flandre, D. and Van de Wiele F., IEEE Electron Device Letters Vol. EDL-8 (1988)
[4] Lim, H.-K. and Fossum, J.G., IEEE Trans. Electron Devices Vol. ED-30 (1983) 1244-1251
[5] Hauser, J.R. and Littlejohn, M.A., Solid-State Electr. 11 (1968) 667-674
[6] Van de Wiele, F., to be published

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

SILICON SELECTIVE AND LATERAL OVERGROWTH EPITAXY : GROWTH AND ELECTRICAL EVALUATION FOR DEVICES

J. FRIEDRICH, G.W. NEUDECK and S.T. LIU*

School of Electrical Engineering, Purdue University, West Lafayette, IN 47906, U.S.A.
**Honeywell SSED, Plymouth, MN 55441, U.S.A.*

Résumé

La croissance épitaxiale sélective du silicium monocristallin ainsi que la surcroissance épitaxiale latérale sont devenus des étapes importantes pour les technologies de traitement avancées. Les problèmes associés à ces étapes, tels que l'uniformité du dépôt, la qualité de la matière et la détérioration de l'oxyde masqueur ont été recherchés. Des éléments ont été construits pour évaluer la matière formée épitaxiallement en fonction de son adéquation pour la fabrication des circuits électroniques. Les résultats procurent des critères fondamentaux qui peuvent être utilisés dans n'importe quel type de réacteur pour atteindre une épitaxie sélective de qualité.

Abstract

Selective epitaxial growth of monocrystalline silicon and epitaxial lateral overgrowth have emerged as important processing steps for advanced processing technologies. The problems associated with these steps, such as deposition non-uniformity, material quality, and masking oxide deterioration have been investigated. Devices have been built to evaluate the epitaxially grown material with respect to its suitability for circuit fabrication. The results give fundamental criteria that can be used in any kind of reactor system to achieve good quality selective epitaxy.

Introduction

In selective epitaxial growth of monocrystalline silicon on a silicon substrate (SEG) the wafer is masked with SiO_2 into which seed windows are cut. Under appropriate conditions one can achieve epitaxial growth in these windows without nucleation on the oxide [1]. When deposition is continued, even after the epitaxial silicon has reached the oxide surface level, lateral growth occurs. Epitaxial lateral overgrowth (ELO) can be used to fabricate silicon-on-insulator structures [2]. Most recently, ELO has been employed as part of a 16 MBit DRAM technology [3]. SEG and ELO offer potential for increased device density, device performance, and design flexibility, while keeping process complexity comparable to that of conventional technologies.

The problems associated with low-temperature selective epitaxial growth are deposition uniformity, material quality, non-selectivity by nucleation of silicon on the masking oxide, planarity and faceting, integrity of the oxide mask, and, for ELO, the aspect ratio of vertical to horizontal growth. We have studied these effects in two different types of commercially available reactors, a radiant heated barrel reactor, and an induction heated pancake type vertical reactor. In this paper we report on some of the results obtained from our experiments and discuss implications for device fabrication.

Growth Mechanism and Uniformity

Most selective epitaxial processes employ $SiCl_2H_2$ (DCS) as the silicon source gas, H_2 as the carrier gas, and HCl to control selectivity and to prevent the deposition of polysilicon on the oxide. This system has been studied for the case of selective as well as non-selective deposition [4]. SEG and ELO in the $SiCl_2H_2$-HCl-H_2 system has been found to be sensitive to temperature, partial pressures of the reactant species, total gas flow, reactor type, ratio of masked to unmasked area on a wafer, and to crystal type and pattern orientation on the growth surface.

SEG poses a peculiar problem in CVD, as the deposition reaction is highly reversible, due to Cl-etching. Furthermore, the existence of two regions with different gas compositions close to each other, one over the masking oxide and the other over the seed area, challenges the ability of reactor systems to achieve good uniformity.

The key to good uniformity is operation at low temperature, where the effect of the Cl/Si ratio in the kinetic rate expression is subdued, by employing high gas velocities to reduce depletion effects, and by having low oxide coverage on the wafers. This way we have achieved 2 % non-uniformity across a wafer for 10 μm of growth.

In an inductively heated, pancake type reactor we found that growth rates were a function of masking oxide thickness (Fig. 1). This dependence is probably temperature related. The thermal conductivity of SiO_2 is two orders of magnitude lower than that of silicon. However, if the decrease in growth rate with increasing oxide thickness was primarily due to smaller thermal conductance, an almost linear drop would be expected. This was not observed, instead the growth rates remain essentially constant after the initial decrease somewhere between 80 nm and 300 nm. Furthermore, wafers with backside oxides did not exhibit lower growth rates, and thin wafers did not have higher growth rates than thicker ones. Thus our results point towards enhanced radiative heat transfer from wafers with oxide layers thicker than 300 nm. From experiments with locally varying masking oxide thickness it was concluded that the observed growth rate dependence is a bulk effect and not due to local variations in temperature and concentration of reactants.

Material Quality

One of the most important issues in low-temperature selective epitaxial growth is the required low defect density of the grown material. We have studied some of the parameters that influence the amount of defects and have

explored the limitations in growing low defect-density material at low temperatures in commercial reactors.

Meyerson *et al.* [5] have demonstrated that device quality material can be obtained at temperatures as low as 750°C in an ultra-high vacuum epitaxy system. In commercially available reactor systems, Borland and Drowley [6] and Drowley and Turner [7] were able to grow device quality epitaxial layers in the pressure range of 10 to 100 torr and at temperatures of about 850°C.

We have conducted experiments that elucidate the importance of the background water vapor and oxygen levels in the epitaxial growth environment, as previously demonstrated by Meyerson for an ultra-high vacuum reactor, and we have applied the results from Ghidini and Smith [8],[9] on $Si/O_2/H_2O/SiO_2$ equilibrium data to predict an operating range for commercially available epitaxy reactor systems where good crystalline quality material can be grown.

In their investigation of the interaction of H_2O and O_2 with silicon surfaces, Ghidini and Smith showed that there exist critical pressures for water vapor and oxygen below which etching of silicon by the reactions

$$Si(s) + O_2 \rightarrow 2SiO(g) \qquad (1)$$

$$Si + H_2O \rightarrow SiO\,(g) + H_2\,(g) \qquad (2)$$

takes place, where T_s denotes the substrate surface temperature. For partial pressures of water vapor or oxygen higher than these critical pressures, oxidation of silicon occurs. For temperatures greater than about 1000°C, the critical partial pressures for water, P_c (H_2O, T_s), are one order of magnitude higher than those for oxygen, P_c (O_2, T_s). However, for temperatures less than 950°C, they are about equal. P_c (O_2, T_s) can be expressed as

$$P_c(O_2,T_s)=4.4 \times 10^{12} \exp(-3.93\ eV / k_B T_s) \text{ torr} \qquad (3)$$

where k_B is Boltzmann's constant and T_s is the silicon surface temperature in degrees Kelvin [8]. For a given deposition pressure, one can thus compute a critical preclean and growth temperature as a function of moisture and oxygen content in the H_2 carrier gas. This is illustrated in Fig. 2. For example, if one has a known moisture plus oxygen content of 1 ppm in the H_2 carrier gas, and the reactor can be operated at a minimum pressure of 25 torr, the minimum deposition temperature to obtain good crystal quality material would be 875°C.

Figure 3 and Fig. 4 show growth above and below the critical temperature, respectively, for a deposition pressure of 150 torr and the moisture content in that particular run. It is observed that good crystal quality epitaxy was also grown at 890°C and 150 torr. At 860°C a high number of stacking faults due to local oxidation during the growth is visible.

The experiments reported here strongly support the assumption that the limitations of growth in commercial reactor systems at low temperature may indeed be governed by the water vapor and oxygen background pressure during preclean and growth. Low partial pressures of oxygen and water vapor can be achieved by low pressure deposition systems, point-of-use purification of the gases, or both. For the system studied, the lower temperature limit at 0.1 ppm residual moisture and oxygen, and a minimum deposition pressure of 150 torr, was about 890°C.

Oxide Integrity

In silicon-on-insulator (SOI) by ELO, the masking oxide functions at the same time as the insulating layer. In many of the new device technologies silicon is grown over a thin layer of SiO_2, either along a sidewall or a gate-type oxide. The integrity of these thin oxides are essential for the fabrication of complex devices and circuits by such processes.

Investigations of the effect of post-oxidation anneals on the breakdown characteristics of thin, thermally grown oxides showed that a considerable number of oxide defects are induced when the anneal is carried out in an oxygen deficient ambient [10]. Such conditions exist in epitaxial reactor systems, where the amount of residual water vapor and oxygen has to be minimized to obtain good crystalline quality material.

We have investigated the effect of H_2, H_2-HCl, and $SiCl_2H_2$-HCl-H_2 ambients at 950°C and 150 torr, conditions typical for SEG, on the electrical breakdown characteristics of thin, thermally grown oxides. We found defect formation in thermally grown oxide films during selective epitaxial depositions to take place, accelerated by the presence of a silicon source gas [11].

Oxidized, unpatterned wafers were divided into four groups, each group containing oxides of nominal 60 nm, 80 nm, 100 nm, and 120 nm thickness. To investigate the effect of the various gas components of the ambient used in SEG, these groups were subjected to different processing steps.

The first set of four (group I) was metallized with a layer of less than 300 nm of Al-Si in a sputtering system, patterned, and annealed at 450°C in dry N_2 for 20 minutes to form MOS capacitors. These wafers served as a standard against which the other wafers were compared.

The second set of wafers (group II) was placed into an inductively heated, pancake type epitaxy reactor, and an epitaxy cycle was carried out. However, no silicon source gas or HCl was added to the ambient. Hydrogen was the only gas introduced into the reaction chamber during the 20 minute bake at 950°C and 150 torr.

The third group of wafers (group III) was subjected to the same procedure as the second group, with the exception that HCl was introduced into the chamber in addition to the H_2 during the 20 minute bake at 950°C at 150 torr. The HCl partial pressure was about 3 torr.

The fourth set of wafers (group IV) was placed into the epitaxy reactor and a complete selective epitaxy run was carried out at 950°C and 150 torr. HCl and $SiCl_2H_2$ were added as reactant gases. The partial pressures of these gases were in the order of 1 torr. All runs were carried out for 20 minutes such that no polysilicon nucleated on the oxide surface. After the epitaxy the wafers of groups II to IV were metallized, patterned, and annealed as described for group I.

Figure 5 plots the yield for the four different groups of wafers. As expected, the control wafers (group I) yield almost 100% good devices. The yield decreases noticeably for wafers baked in H_2 and HCl at 950°C and 150 torr for 20 minutes (group III), and even more when $SiCl_2H_2$ was added to the ambient (group IV).

On a clean silicon surface, etching occurs in the presence of oxygen or water vapor when the O_2 or H_2O partial pressures are smaller than the equilibrium partial pressure of SiO [8,9], which is a strong function of temperature. At 950°C, partial pressures of O_2 or H_2O of less than about 0.5 mtorr lead to etching of exposed silicon by formation of volatile SiO. In the Si-SiO_2 system, formation of SiO can occur via the disproportionation reaction

$$Si + SiO_2 \rightarrow 2SiO(v) \qquad (4)$$

for sufficiently low O_2 and H_2O partial pressures at the reaction site. Such conditions have been achieved in the investigation of post-oxidation anneals of thermally grown SiO_2 by Hofmann *et al.* [10]. They reported that between 750°C and 1100°C oxide decomposition was non-uniform, initiated at nucleation sites and resulting in voids in the oxide.

At 950°C, the following conditions must be met for equation (4) to occur at measurable rates: the O_2 or H_2O partial pressures have to be lower than 0.5 mtorr, silicon must be supplied to the reaction site, and the reaction product SiO must be able to leave the reaction site. In our epitaxy system, all three conditions were met. Water content in the feed gas was always less then 1 ppm (typically 0.4 ppm). For a deposition pressure of 150 torr, that would translate into a water vapor partial pressure of about 0.2 mtorr, which at 950°C is less than the equilibrium partial pressure of SiO. The oxides that were treated in the epitaxy reactor with just H_2 or H_2 and HCl can be compared to those that went through the POA procedures as described by others. Here oxide decomposition would begin at weak points in the oxide, that may be caused by pre-oxidation substrate defects as well as post-oxidation contamination.

The oxides that went through a regular selective epitaxy, with DCS as the silicon source gas, exhibit a large number of defects. Here additional silicon is supplied to the oxide surface from the gas phase, by decomposition of the silicon source gas DCS at the hot surface. The abundance of Si can accelerate reaction (4), leading to higher defect densities. In addition, the reaction product SiO is easily transported away from the reaction site at the oxide surface.

Taking the equilibrium partial pressure of SiO as the dividing line between the region of oxide decomposition and oxidation, Fig. 6 shows the critical reactor pressure as a function of deposition temperature, with the carrier gas moisture content as a parameter. The shaded regions denote the conditions under which defects can be induced by oxide decomposition. One can see that the critical conditions fall right into a range where selective epitaxial depositions are commonly carried out. For depositions at 150 torr and a moisture content of 1 ppm, the maximum allowable deposition temperature would be about 920°C.

Devices

We have built bipolar transistors as shown in Fig. 7 into material grown above the critical minimum temperature, at 950°C and 150 torr in a vertical pancake reactor, and simultaneously in the silicon substrate in the same die. The devices show almost identical junction properties, indicating the quality of the SEG/ELO epitaxial material is of device quality. Peak betas of greater than 450 were measured that were still greater than 100 at 0.1 nA. The junctions had ideality factors from 1.0 to 1.03 with reverse currents less than 30 nA/cm^2.

For SOI by ELO, we have grown silicon selectively from seed windows, spaced 10 μm apart, such that the growth fronts merged, resulting in a planar overall surface everywhere. This was done in a radiant heated barrel reactor at 900°C and 50 torr and an HCl/DCS ratio of 2. The thickness of the buried oxide stripes was 400 nm. About 8 μm of vertical growth was required to obtain a planar surface, which was then etched back to about 1 μm by plasma etching. A cross-sectional view of the resulting SOI by ELO is shown in Fig. 8. After the plasma etch the film was further thinned to about 0.6 μm by wet etching to remove plasma damage. A row of discrete NMOS transistors of various sizes were built on the SEG seed region, and comparable transistors were built on the ELO region above the oxide stripes. The SOI wafers were doped uniformly to approximately $1.5x10^{17}$ cm^{-3} by multiple boron implantation. The gate oxide thickness was 20 nm. The gate polysilicon was doped N+ and was about 400 nm thick. The source and drain regions were formed by arsenic implantation and annealed at 925°C. After the contact and metal processing, the I-V characteristics were measured. Fig. 9 shows a typical result for an NMOS transistor made on SOI by ELO. The drawn features of this transistor were W/L=3.5μm/3.5μm. In this measurement the source was connected to the "ELO-substrate" to show the typical kink behavior of SOI devices. Evaluated at the low drain voltage, the channel mobility was determined as 410 cm^2/V-sec. The same size transistor in the SEG seed region exhibited the same mobility. The subthreshold slope and breakdown voltages were 120 mV/dec and 8.1-8.3 Volts, respectively, for both types of transistors. From these measurements we conclude that the electrical properties of SOI by ELO are the same as the ones of the SEG bulk region.

Acknowledgment

This work was supported by SRC N0. 87-SJ-108 and ONR No. N 00014-85-0438.

References

[1] K Tanno, N. Endo, H. Kitajima, Y. Kurogi and H. Tsuya, Jap. Jour. of Appl. Physics **21** L564 (1982)

[2] L Jastrzebski, A.C. Ipri, and J.F. Corboy, IEEE Electron Dev. Letters **EDL-4**, 32 (1983)

[3] G.B. Bronner et al., Symposium on VLSI Technology Digest, San Diego, 21 (1988)

[4] C.I. Drowley, Chemical Vapor Deposition 1987, Electrochem. Soc. **PV 87-7**, 418 (1987)

[5] B.S. Meyerson, E. Ganin, D.A. Smith, and T.N. Nguyen, J. Electrochem. Soc. **133**, 1232 (1986)

[6] J.O. Borland and C.I. Drowley, Solid State Technology **28 (8)**, 141 (1985)

[7] C.I. Drowley and J.E. Turner, Chemical Vapor Deposition 1987, Electrochem. Soc. **PV 87-7**, 243 (1987)

[8] F.W. Smith and G. Ghidini, J. Electrochem. Soc. **129**, 1300 (1982)

[9] G. Ghidini and F.W. Smith, J. Electrochem. Soc. **131**, 2924 (1984)

[10] K. Hofmann, G.W. Rubloff, and D.R. Young, J. Appl. Phys. **61**, 4584 (1987)

[11] J. Friedrich and G.W. Neudeck, to be published in the J.Appl. Phys.

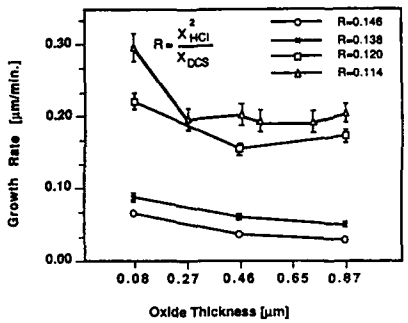


Fig. 1: Growth rate as a function of masking oxide thickness

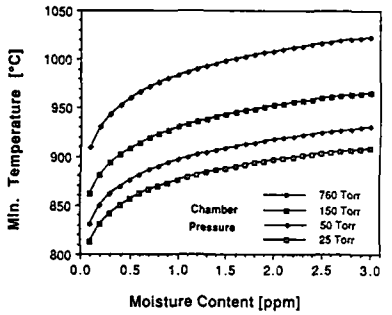


Fig. 2: Critical preclean and growth temperature as a function of moisture content in the growth ambient

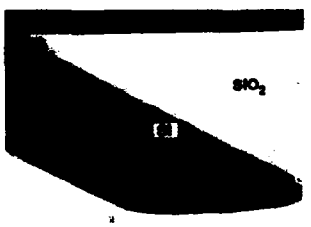


Fig. 3: SEM of SEG above the minimum required temperature

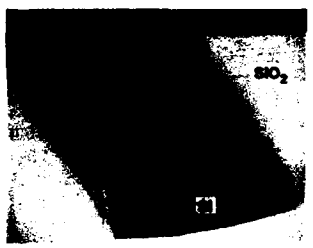


Fig. 4: SEM of SEG below the minimum required temperature

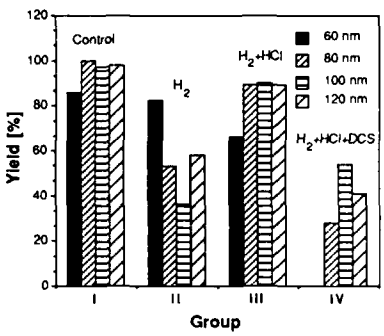


Fig. 5: Yield for capacitors build on oxides exposed to different ambients

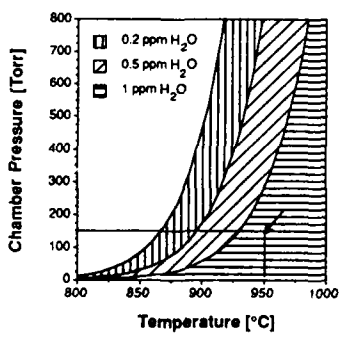


Fig. 6: Critical reactor pressure as a function of deposition temperature

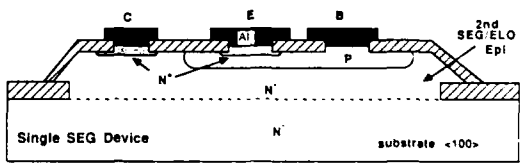


Fig. 7: SEG/ELO bipolar transistor built for material evaluation

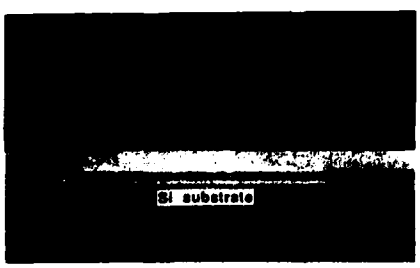


Fig. 8: SOI by ELO after merging and plasma etch back

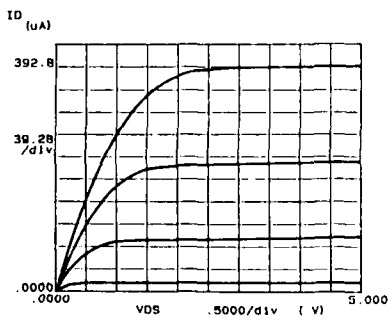


Fig. 9: SOI NMOS transistor I-V characteristics

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

HIGHER EFFICIENCY OF CMOS-PROCESS-COMPATIBLE PHOTODIODES IN SOI-TECHNIQUE BY REFLECTING FILMS

K. KNOSPE and K. GOSER

Bauelemente der Elektrotechnik, Universität Dortmund, D-4600 Dortmund, F.R.G.

Résumé: Une tension de plusieurs volts etait realisée sur circuit integré par integration de photodiodes en couches de silicium polycristalin en technique SOI. Pour accroire le rendement quantique une couche appropriée etait deposée au joint Si - poly-Si. Cette couche ne pose aucune restriction pour les étapes suivantes et les traitements thermiques. Une comparison des rendements quantiques mesurés avec les valeurs théoriques a été faite.

Abstract: A voltage of several volts was achieved on chip by integration of photodiodes in polycristalline silicon layers in SOI-technique. For increased quantum efficiency suitable layers were deposited at the Si - poly-Si boundary. These layers caused no restrictions for successing process steps and temperature treatments. A comparison of measured quantum efficiencies with theoretical values has been made.

Introduction

With the development of optical communication techniques the importance of opto-electrical transformers increases. Because the CMOS-process is well advanced and gives the highest integration level, the opto-electrical transformers should be compatible to this process. Parts of such transformers are photodiodes(PD). To switch a MOS-transistor a single PD is not sufficient, because the voltage of a single photodiode (caused by incident light) is too low. So we realized a series connection of PDs in SOI-technique to obtain the necessary photovoltage. The integration of all devices in a poly-Si-layer with a thickness of about 1μm causes a bad quantum efficiency of the PD, especially for radiation of wavelengths above 0,6μm (penetration depth >2μm, increases rapidly at higher wavelengths). To increase the quantum efficiency we deposited both an anti-reflecting-layer (ARL) (to minimize reflection at the top of the poly-Si-layer) and a back-

reflecting-layer (BRL) (to minimize the reflection at the bottom of the poly-Si-layer) /1/.

Manufacturing

The PDs (and the other devices) were manufactured by a SOI-technique, which is compatible to the CMOS-process used in our institute of the university of Dortmund. First different layers of SiO_2 and poly-Si were deposited on the Si-substrate. They form both the BRL of the photodiode and the insulating layer necessary for SOI-technique. Their thicknesses were optimized for radiation of a wavelength of 800nm, and they caused no restrictions to the following process-steps (e.g. annealing at 800°C or recristallisation at 1250°C). Then a 1,2µm thick layer of poly-Si was deposited on. This layer was recristallised 5 hours at 1250°C /2/. The medium size of the crystallites was 0,42µm. In this layer the photodiodes (and other devices) were integrated by different steps of oxidation, masking, etching and implantation of boron and phosphorous respectively. The separation of the photodiodes (for series connection) was realized by etching trenches into the poly-Si-layer down to the surface of the BRL. After metallisation an ARL was deposited on the top of the photodiodes areas. This layer consists of different SiO_2 layers (deposited from a plasma, thermally grown); the ARL was also optimized for radiation of a wavelength of 800 nm.

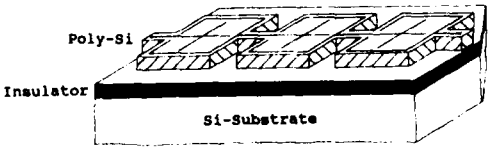


Fig. 1 Scematic view of the series connection

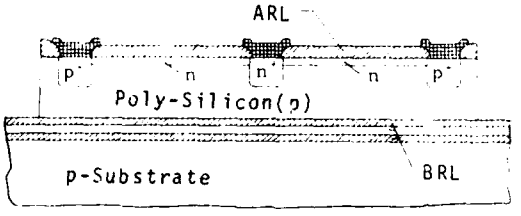


Fig. 2 Scematic cross section of a photodiode

Fig.3 is a REM-photo; it shows an etched trench between two PDs (dark, interupted, horizontal line). The connection between the PDs is realized by an aluminium layer (8µm width, raised areas).

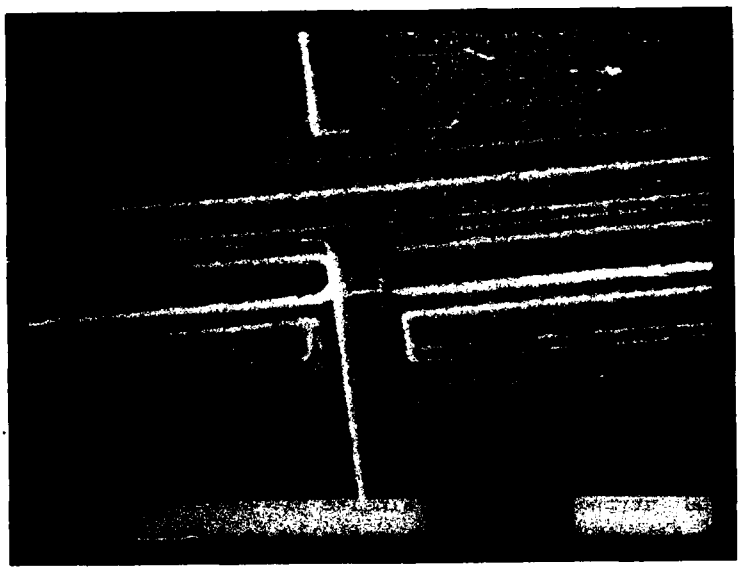


Fig.3 REM-photograph of the connection between 2 PDs

Experimental and theoretical results

A series connection of 10 PDs was measured. Under monochromatic light with low intensity we obtained the following photovoltages:

λ [μm]	U [mV]	Radiation power [μW/mm^2]
0,5	1170 ± 90	25
0,8	340 ± 20	5

Under white microscope light more than 3V were generated. The quantum efficiency (fig. 4) reaches up to 30% (response 120mA/W at this maximum); at higher wavelengths an oscillation of the curve can be seen. One maximum of the oscillation is at a wavelength of 0,8μm. At this wavelength we have got a quantum efficiency of 9%, the response was 60mA/W. The theoretical values were obtained by computing the absorbed part of the incident radiation within the poly-Si-layer, multiplied with a recombination factor (RCF). This RCF takes into account the leakage current and the recombination of the generated charge carriers inside the poly-Si-layer. At longer wavelengths the experimental and theoretical data agree well.

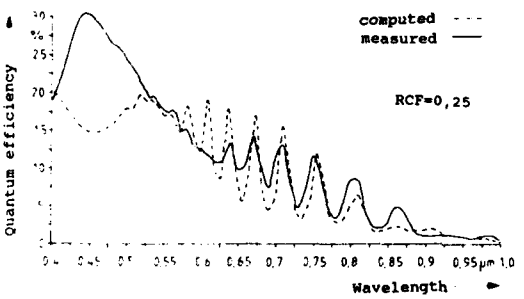


Fig. 4 Average quantum efficiency of 8 samples of the same lot

The deviations are due to the absorption coefficient and the differences between the calculated and constructed layer thickness. Because the exact data for poly-Si were not known (they depend strongly on the grainsize of poly-Si), those of crystalline silicon were used. Especially at lower wavelength, which are in the range of the grain sizes, the deviations are big.

Conclusions

By integrating PDs in poly-Si in SOI-technique we get a voltage of several volts on a chip. The use of reflecting films (which cause no restrictions to the manufacturing process) increased the quantum efficiency (and so the photocurrent) at specific wavelengths. The combination of all measures leads to an increase of the usable power of CMOS-process-compatible PDs. Laser recristallisation was used to increase the medium size of the crystallites in the poly-Si-layer (and thereby the quantum efficiency)/3/.

References

/1/ J.Müller, IEEE Trans.on El. Dev., Vol.25 (1978) 247
/2/ G.Harbecke,H.F.Kappert et al, J.Electrochem.Soc. 131(1984)675
/3/ U.Hilleringmann,K.Goser, Kleinheubacher Berichte 31(1987)65

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

COOL PLASMA ACTIVATED SURFACE IN SILICON WAFER DIRECT BONDING TECHNOLOGY

G.-L. SUN, J. ZHAN, Q.-Y. TONG, S.-J. XIE, Y.-M. CAI and S.-J. LU

Microelectronics Center, Nanjing Institute of Technology, Nanjing 210018, China

Abstract--A novel cool plasma surface activation method has been developed for high quality SOI/SDB (Silicon wafer Direct Bonding) preparation. The activation effectiveness of different plasma gases, espetially of O_2 plsma gases were investigated. The measurements of H.V-PMOS and L.V.-NMOS devices made on the SOI/SDB and on a bulk Silicon indicate that ratios of electron and hole mobility of SOI to those of bulk Silicon are 0.92 and 0.88,respectively.It was proved that our SOI substrate produced by SDB is of device level quality.

1.INTRODUCTION

SOI material has been the subject of an intense and challenging research over past few decads, because it has the potential to provide the ideal substrate for ultra-high speed,high voltage,high temperature and high radiation-hardness VLSI systems.Recently,SOI produced by Silicon wafer Direct Bonding (SDB) technology has attracted much attention because of its high quality,high flexibility and low cost(1),(2),(3).However,sucessful bonding depends in large extend on the surface treatment.In this paper we report a novel cool plasma surface activation method for SOI/SDB preparation.

2.EXPERIMENTS AND DISCUSSION

The bonding process adopted in this work is shown in Fig.1.Tow polished wafers were oxidized thermally, and the surface activation of tow wafers were performed before the prebonding process which was carried out in clean air condition.The bonding was then processed in N_2 or O_2 ambient at about 1000 °C for a few hours.

It has been found that the surface activation and subsequent absorption of OH groups are key process steps for suceessful prebonding.The surface treatment in this work is as follows:(a) Radical atoms/moleculars (O_2,N_2,NH_3 , etc.)produced by R.F.plasma activate the surface for 6 minutes.(b)Silicon wafers are immersed in a molecular type surface activation solvent for about 1 hours in order that the surface absorbs large amounts of OH groups.(c)Wafers are cleanned by standard IC wafer cleanning procedures.

The infrared transmittance curves (O, N, A and V) in Fig.2 show broad and split dips between 3100--3700 cm-1 owing to absorption of OH groups on the surface .The symbols O,N,A and V refer to samples treated at 850 °C for 6 minutes in O_2,N_2,NH_3 and vacuum plasma respectively,and θ is original sample.

It was found that the sample O in Fig.3 is deactivated after being heated at 1100 °C for 0.5 hour and then being passed steps (b) and (c).The broad and split dips no longer apper in the curve.It furthur confirms that plasma treatment activates the surface and enhances the absorption of OH groups.

The investigation of the influence of temperature and time on O_2 plasma activation process was carried out.Fig.4 show that temp. less than 850°C is better than 900 °C in terms of OH groups absorption.But there is almost no difference between IR transmittance spectra curves of O_2 plasma treatment for 6 minutes and 60 minutes at 850°C in Fig.5.It is assumed that there is a critical temperature value above which deactivation dominates in O_2 plasma treatment.

The fracture strength of the bonded silicon wafer is 110--145 kg/cm^2, which is close to the values of silicon wafers used. SEM microphotograph,in Fig.6 shows the cross-section of resulted bonding wafer.

The high voltage PMOS device, as shown in Fig.7 was designed.It was fabricated with standard voltage NMOS device both on the SOI/SDB substrate and bulk Si using Al-gate CMOS process to assess the electrical properties of the SOI layer.The SOI/SDB substrate has 4 μm thickness with resistivity of 20 Ω*cm.Fig.8 is I-V characteristic of 100V.H.V.PMOS device.The measurements indicate that ratios of electron and hole mobility on SOI layer to those of bulk Silicon are 0.92 and 0.88 respectively.It shows that our SOI/SDB is of device level quality.

3. CONCLUTION

A cool plasma can activate the SiO_2 of a Si wafer and helps the bonding process.Activation effectiveness of different plasma was compared and we found no sigificant difference among various plasma gases.The temperature during activation of O_2 plasma below 850°C is nessary to avoid deactivation.The devices made on SOI layer show that SOI/SDB subsrtrate is suitable to IC application.

This work was supported by Chinese Natural Science Foundation (6866011).

REFERENCE

(1) J.B. Laskey, S.R. Stiffler. F.R.White and J.R.Abernathey,Tech. Dig. IEDM pp 684--687.1985.
(2) Semicondutor World ,6,1986. (Japan)
(3) Li Hui, Sun Guo-Liang,Zhan Juan, and Tong Qin-Yi, Applied Surface Science 30, PP 397-401,1987.

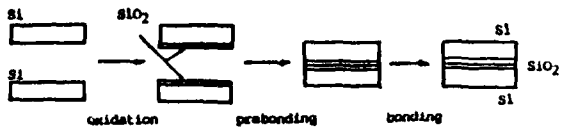


Fig.1 SDB process sequence

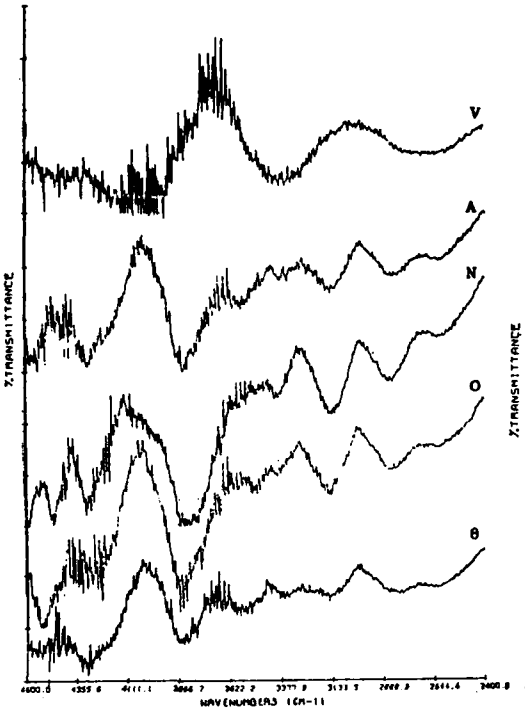


Fig.2 IR transmittance spectras

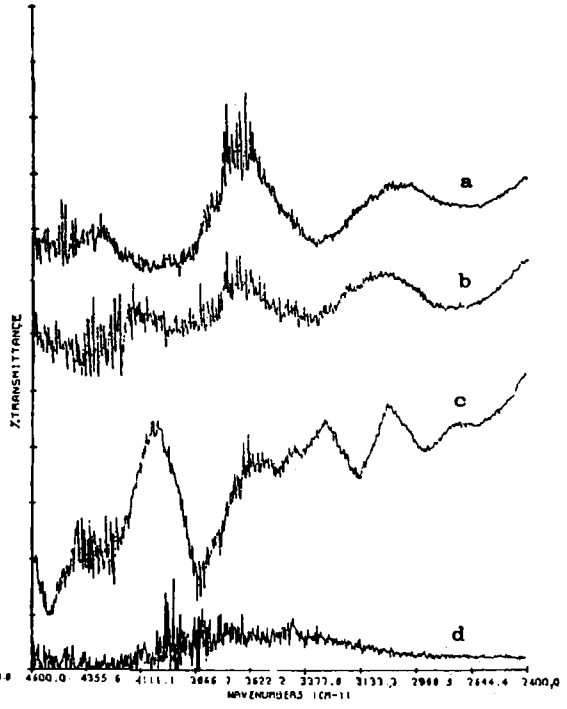


Fig.4 Temp.influence on IR spectras
a: at 300°C, b: at 700°C
c: at 850°C, d: at 900°C

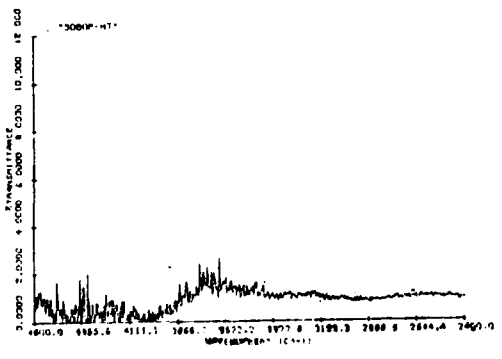


Fig.3 IR spectra of heated sample

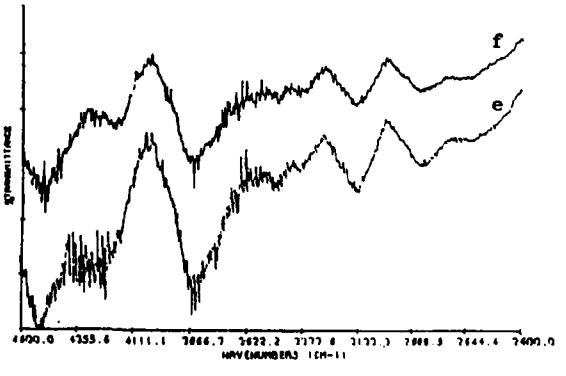


Fig.5 Time influencee on IR spectras
e: 6 minutes
f: 60 minutes

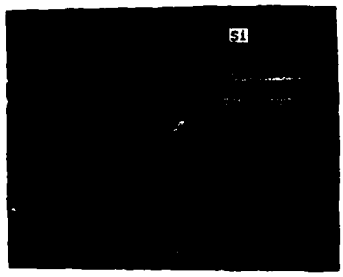


Fig.6 SEM cross-section microphotograph of SOI/SDB

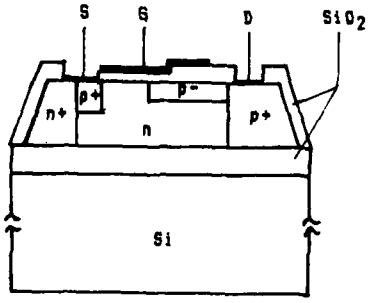


Fig.7 H.V.-PMOS structure

Fig.8 I-V characteristic of H.V.PMOS
Horizantal axis, V_{ce}, 20V/div; Vertical axis, I_d, 500uA/div; Gate volt., 2V/div

Session 1C Room C : Bipolar, BICMOS : technology

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CIT 1 AND CIT 2, ADVANCED NON EPITAXIAL BIPOLAR/CMOS PROCESSES FOR ANALOG-DIGITAL VLSI

C. VOLZ and L. BLOSSFELD

ITT Intermetall, Hans-Bunte-Str. 19, D-7800 Freiburg, F.R.G.

ABSTRACT

Deux processus Bipolaire/CMOS de haute performance ont été développés; les processus CIT 1 (2.0 µm) et CIT 2 (1.5 µm). La technologie CIT n'utilise ni l'epitaxie ni le buried layer. Des transistors bipolaires (npn, pnp) et des transistors MOS (canal n, canal p) sont implantes avec succes sur un meme chip sans diminuer les performances des deux technologies en présence.

Two N-well high performance Bipolar/CMOS processes have been developed; a 2.0 µm CIT 1 process and CIT 2, a 1.5 µm process. CIT technology uses neither epitaxie nor buried layer. Bipolar transistors (npn and pnp) and MOS transistors (n-channel and p-channel) have been sucessfully fabricated on the same chip without a decrease of the performance.

1 INTRODUCTION

For expanding markets like the consumer market, the automotive, or the telecommunication sectors, signal processors with working frequencies of 1 GHz and more are required. The integration of various kinds of functions on the same chip is advantegous. Today's system designs demand processes which allow integration on one chip [1 - 3]. For a high integration level the power consumption is also a very important issue. In the past bipolar devices have been used for these analog-digital applications.

In this paper two new high performance nonepitaxial Bipolar/CMOS processes are proposed; the 2.0 µm CIT 1 process and CIT 2 [4, 5], a 1.5 µm process. Bipolar transistors (npn and pnp) and MOS transistors (both n-channel and p-channel) have been succesfully fabricated on the same chip.

2 DEVICE STRUCTURE AND PROCESS SEQUENCE

In the present CIT processes p-type <111> substrates with a resistivity of 2 Ohmcm are used as a starting material. The n-wells for the CMOS transistors and the collectors of the bipolar transistors are formed by a phosphorous implantation. The n-layer has a sheet resistance of 1 kOhm per square and is 2.5 µm deep. Base and drain/source for the p-channel transistors as well as the emitter and n+ drain/source are made simultaneously.

The base is doped by a double boron implantation. The collector contact and the emitter area are defined by a nitride layer, which is used as a mask for the ion implantation of the extrinsic base. The intrinsic base is implanted through this nitride layer where the collector is protected by photoresist. The following oxidation creates a bird's beak. It allows a self aligned separation of the highly doped extrinsic base from the collector contact and the emitter area.

2.1 CIT 1

After an arsenic ion implantation for emitter and n+ drain/source a dielectric layer of PECVD oxide is deposited. In the MOS channel region this oxide is removed. The threshold voltage is adjusted by ion implantation through an 80 nm thick gate oxide. CIT 1 uses a metal gate. This process is very simple and needs only 6 photolithographic steps. Figure 1 shows the cross section of the 2.0 µm Bipolar/CMOS device structures.

2.2 CIT 2

After removing the nitride layer and gate oxidation, buried contacts for emitter, collector and drain/source are opened in the gate oxide, which has a thickness of 20 nm. Polysilicon is deposited and patterned. The polysilicon is used as an etch mask to remove the oxide on the external base. This step forms a vertical spacer. The 200 nm thick emitter polysilicon is doped by an arsenic ion implantation. To reduce the sheet resistance of the polysilicon (200 Ohms per square) and of the extrinsic base a 50 nm thick layer of platinum silicide with a resistivity of 8 Ohms per square is selectively deposited. The bipolar transistor surface is completely covered by this silicide. Both collector-base and base-emitter are electrically isolated by a vertical oxide spacer.

After gate patterning, the lightly doped drain region is defined by a shallow ion implantation. The threshold voltage is adjusted by a deep ion implantation through the polysilicon layer. The n-channel drain/source contacts are completely covered by polysilicon. The p-channel drain/source diffusion and the n-doped polysilicon are linked by platinum silicide. The silicided polysilicon can be used as a first metal level. It can interconnect all kind of MOS contacts, n- and p-channel.

The second metal is 1.2 µm thick aluminium with 2 % silicon. Between platinum silicide and aluminium a 170 nm thick barrier layer of titanium is inserted. A third metal level is available. Figure 2 shows the cross section of the 1.5 µm Bipolar/CMOS device structures.

3 DEVICE PERFORMANCE

Table 1 gives a summary of the performance.

3.1 NPN DEVICE PERFORMANCE

Bipolar transistors with a performance comparable to those produced by conventional collector implanted bipolar processes were fabricated. The current gain was 80 and independent of the collector current up to 10 nA. The breakdown voltages were BVEBO = 6.5 V, BVCBO = 18 V and BVCEO = 6.5 V. The transition frequency fT was 1.5 GHz for CIT 1 and 5.0 GHz for CIT 2.

3.2 CMOS DEVICE PERFORMANCE

The threshold voltages and the transistor conductance (W/L = 1) of the CIT 1 devices are Vth = 1.75 V, Beta = 8 µA/V2 for the n-channel transistor and Vth = -1.75 V, beta = 8 µA/V2 for the p-channel transistor. Using 1 additional mask the threshold voltages can be adjusted independently for the n- and

p-channel transistors. By using the n-well as a lighty doped drain for the n-channel MOS transistor a drain/source breakdown voltage of typically 40 V was measured.

The CIT 2 devices threshold voltages and conductances (W / L = 1) are Vth = 0.8 V, Beta = 50 µA/V2 for the n-channel transistor and Vth = - 0.8 V, Beta = 30 µA/V2 for the p-channel transistor.

3.3 CIT 2 BASIC GATES

To evaluate the basic gate delay, ringoscillators with 51 stages were fabricated. The minimum bipolar propagation delay time at 1.25 W is 180 ps for a fan out of 2. The CMOS oscillator is interconnected only by polysilicon lines. Its minimum propagation delay is 280 ps for a fan out of 1.

4 CONCLUSION

2.0 um and 1.5 um nonepitaxial Bipolar/CMOS processes with high performance ECL circuits have been developed. The very simple metal gate process CIT 1 with arsenic implanted emitter needs only 6 photolithographic steps. The high speed process CIT 2 uses a polysilicon emitter and silicided polysilicon lines as a first metal level. The process is optimized for ECL within the Bipolar/CMOS process. The minimum bipolar propagation delay is 180 ps. Both processes have the potential for monolithic multifunctional analog- digital VLSI.

5 ACKNOWLEDGEMENT

The authors thank the engineers of INTERMETALL Development Department for very helpful discussions and support. This work has been partly supported by the Federal Department of Research and Technology of the Federal Republic of Germany.

6 REFERENCES

[1] Hiroshi Momose, Hideki Shibata, Shinji Saitoh, Jun'ichi Miyamoto, Kohichi Kanzaki, Susumu Kohyama: 1.0 µm n-Well CMOS/Bipolar Technology, IEEE Journal of Solid-State Circuits, Vol.SC-20, No.1, February 1985

[2] Yutaka Okada, Kenji Kaneko, Satoshi Kudo, Kouichi Yamazaki, Takahiro Okabe: An Advanced Bipolar-MOS-I2L Technology with a Thin Epitaxial Layer for Analog-Digital VLSI, IEEE Journal of Solid-State Circuits, Vol.SC-20, No.1, February 1985

[3] Katsumi Ogiue, Masanori Okada, Shuuichi Miyaoka, Ikuro Masuda, Takahide Ikeda, Kenichi Tonomura: 13-ns, 500-mW, 64-kbit ECL RAM Using HI-BICMOS Technology, IEEE Journal of Solid-State Circuits, Vol.SC-21, No.5, October 1986

[4] C.Volz, L.Blossfeld: Collector Implanted Technology for Si Bipolar Devices, ESSDERC, paper, September 1986

[5] L.Blossfeld, C.Volz: CIT 2 - eine vollsilizierte neue bipolare VLSI-Technologie, IGT-Fachtagung Grossintegration, paper, March 1987

Table 1: summary of the BICMOS device performance

	parameter	CIT 1	CIT 2	unit
NPN	emitter size	2.5 * 2.5	1.5 * 5.0	µm2
	RE	18	15	Ohm
	RB ex	300	20	Ohm
	RC	500	170	Ohm
	Cje	65	28	fF
	Cjc	100	45	fF
	hFE	80	80	
	BVEB	6.5	6.5	V
	BVCE	6.5	6.5	V
	BVCB	18	18	V
	fT	1.5	5.0	GHz
N MOS	Lg	2.4	1.5	µm
	Vth	1.75	0.8	V
	BVDS	18(40)	12	V
P MOS	Lg	2.4	1.5	µm
	Vth	-1.75	-0.8	V
	BVDS	18	12	V

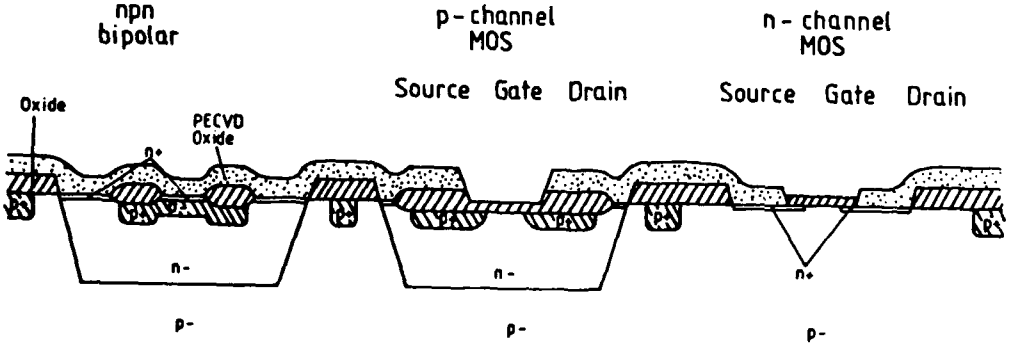


Fig.1: cross section of the CIT 1 BICMOS device structures

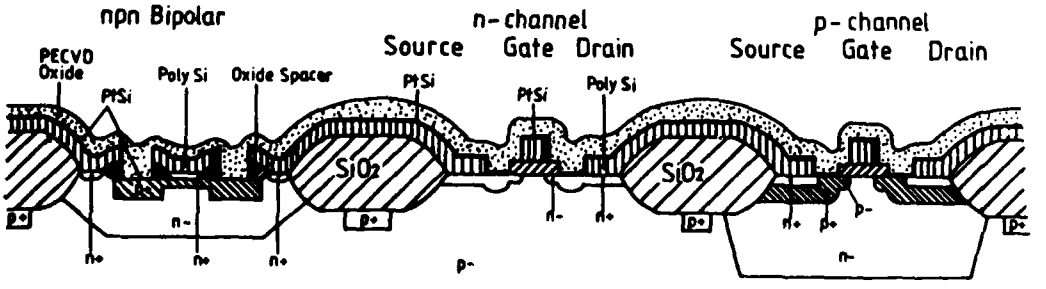


Fig.2: cross section of the CIT 2 BICMOS device structure

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

HIGH-SPEED OPTICAL DETECTION UP TO 2.5Gbit/s WITH A DOUBLE POLYSILICON SELF-ALIGNED SILICON BIPOLAR TRANSISTOR

W. BOCK, L. TREITINGER and W. PRETTL*

Siemens AG, Zentrale Forschung und Entwicklung, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.
** Institut für Angewandte Physik, Universität Regensburg, D-8400 Regensburg, F.R.G.*

Abstract – The photo response of a standard high-speed self-aligned silicon bipolar transistor has been investigated. The photosignal of the base collector diode is found to consist of at least two current components. Optical detection capabilities are demonstrated by excitation with modulated laser light at 830nm wavelength up to data rates of 2.5Gbit/s.

1 – INTRODUCTION

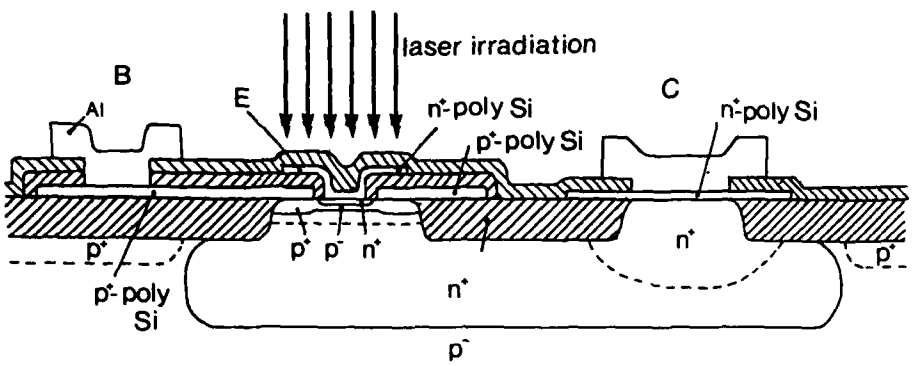
The development of double polysilicon (poly-Si) self aligning bipolar technology has provided a drastical increase of speed as well as reduced power consumption [1]. These goals are reached by cutting down the parasitics by using the self-aligned emitter base configuration as well as by scaling down the vertical dimensions enabled by the polysilicon emitter. In this work it is shown that these vertical npn standard bipolar transistors have additionally an optical detection capability with high sensitivity and remarkably good frequency response.

2 – EXPERIMENTAL

We report on the photo response of the base collector (BC) diode of a vertical npn transistor fabricated using a self-aligning silicon bipolar process. Fig.1 shows a cross section of the transistor and gives the most important dimensions and transistor parameters. Polycrystalline silicon layers are used as diffusion sources and as interconnections to the base and emitter contacts. To improve the optical sensitivity the emitter contact is located about 10µm away from the intrinsic transistor so that the emitter window enables interband light to penetrate into the BC junction. This special transistor does not demand for any additional process steps in comparison to the processing of the standard npn transistor.

A GaAlAs semiconductor laser beam with a wavelength $\lambda = 830$nm was used for optical excitation. Light intensity and laser pulse shape were monitored by use of a calibrated silicon photo diode.

To determine the sensitivity of the BC diode the laser was driven by a dc current source. The transient signal was investigated by application of a pulse generator and a bit pattern generator for modulation of the laser current. Fig.2 shows the measured dc photocurrent as a function of the light input power. The photosignal increases linearily within the range of 10^{-7}mW to 1mW of optical power. The sensitivity is about 0.046A/W corresponding to a quantum efficiency of 7%. This is in agreement with the limited vertical extensions of the device due to the shallow doping profile needed for transit frequency greater than 12GHz. The depth of the collector of 1.6µm has to be compared with the light penetration depth $1/\alpha \approx 20\mu m$, where α is the absorption coefficient of silicon at the used wavelength.



A_E = 2 x 8 µm²	emitter mask size	C_{BC} = 50 fF	BC junction capacitance
A_B = 6 x 12 µm²	base area	R_B = 150 Ω	base resistance
w_b = 0.2 µm	base width	R_C = 50 Ω	collector resistance
w_c = 1.6 µm	depth of the buried collector		

Fig.1 - Data and cross section of the self aligned transistor

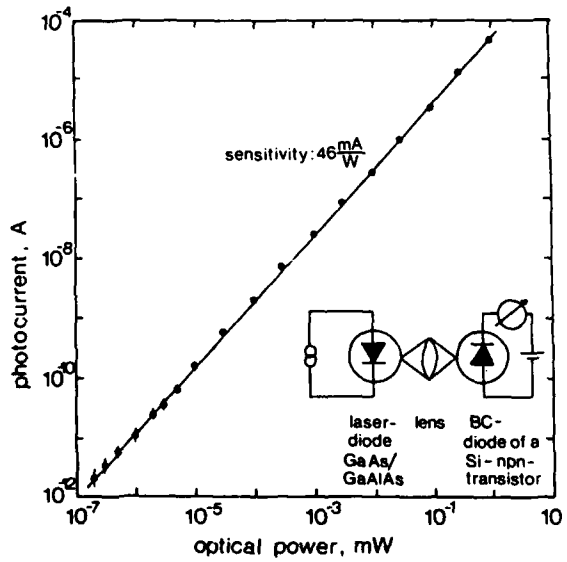


Fig.2 - Photocurrent versus incident optical power. The insert shows schematically the experimental arrangement.

The transient behaviour of the photosignal due to the excitation with a trapezoidally shaped light pulse is shown in Fig.3. The photo response S(t) plotted as $S_o - S(t)$ for the signal rise and as S(t) for the decay, where S_o is the steady state peak signal, shows a fast and a slow component. The fast component defines the first half of the total signal stroke and follows instantaniously the laser pulse with a time constant less than 250ps and 370ps for signal rise

and decay, respectively. For both transients the photocurrent approaches the final value with an exponential time constant τ = 0.55ns of the slow component.
Complementarily to the transient behaviour is the frequency dependence of the photosignal. Fig.4 shows a decay of about 20dB/decade of the photocurrent at modulation frequencies above 2GHz.

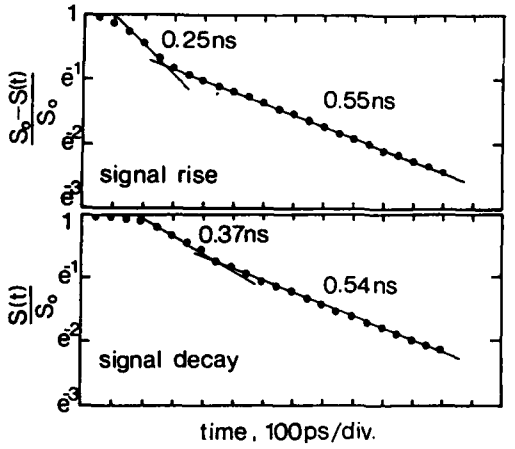


Fig.3 - Rise and decay of the photosignal S(t) normalized to the peak signal S_o due to pulsed excitation

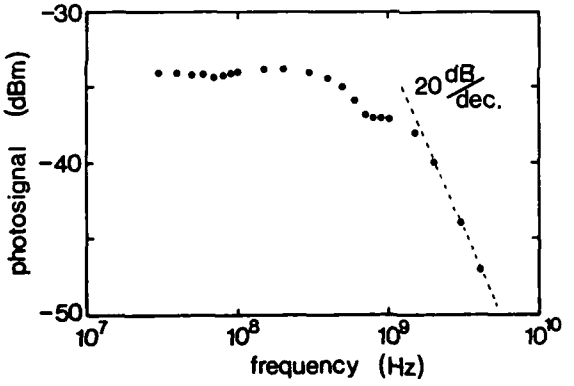


Fig.4 - Frequency response of the photosignal due to sinusoidal excitation

3 – DISCUSSION

To understand the time and frequency dependence of the photocurrent various time delays of the BC- diode have to be taken into account [3]. The time constant $\tau_j \leq$ 10ps given by the product of the base-collector junction capacitance C_{BC} with the sum of the external base and collector resistance, $R_B + R_C$ has not much influence on the transient signal. Similarily the drift time $\tau_d \approx$ 5ps of the carriers through the depletion layer cannot be responsible for the much greater delay measured. The dominant contribution to the transient behaviour originates from the photon absorption in the undepleted regions of the external base and of the collector. As the minority carrier lifetime in highly-doped polycrystalline silicon ($N_A \geq 10^{19}$ cm^{-3}) is shorter than 1ns [4], the main effect will be given by the relatively large vertical extension of the collector layer. The time delay τ_{diff} due to the diffusion of minority carriers towards the junction is approximately given by [5]

$$\tau_{diff} = \frac{d^2}{2.4 \cdot D}$$

where d is the width of the absorption region and D is the minority carrier diffusion constant. As the collector doping concentration increases towards the buried layer, the diffusion constant varies locally. For $d = 1.6\mu m$ and a reasonable average value of $D = 20 cm^2 s^{-1}$ we find $\tau_{diff} = 0.5 ns$ being of the same order as the time constant of the observed slow component. As a result we find that the total photosignal consist of at least two current components. A fast component originates from the absorption of photons in the depletion layer of the diode and a slower component with a time constant $\tau = 0.55 ns$ is due to a diffusion process of minority carriers to the junction.

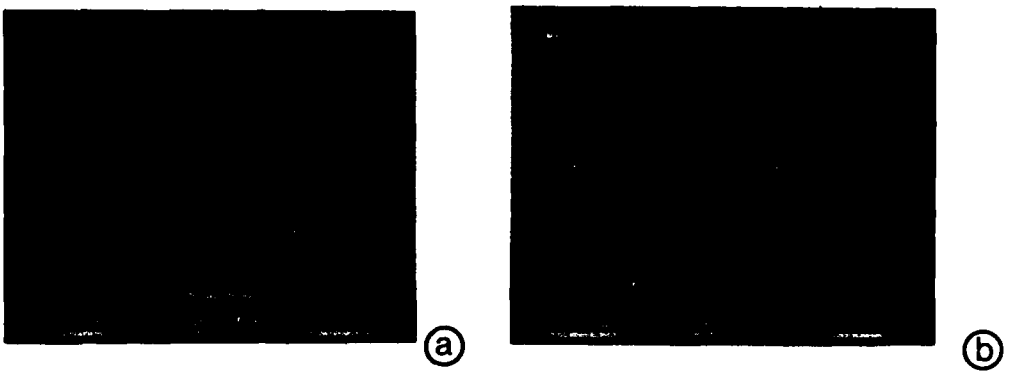


Fig.5 - Light pulse pattern at 700Mbit/s (a) and 2.5Gbit/s (b) detected by the transistor BC-diode (upper trace) and by a fast Si-photodiode (lower trace). Both signals are inverted.

Fig.5 shows the response of the photodiode in comparison to the incident laser light. The laser was driven by a bit pattern generator at the data rates of 700Mbit/s (a) and 2.5Gbit/s (b), respectively. The lower curve represents the photosignal of the reference photo diode, the upper one is the generated photocurrent measured at the load of 50Ω. In the photographs taken from the screen of a sample oscilloscope switching on the light corrsponds to the falling flanks and vice versa. The slow component limits the bandwidth and leads to a reduced signal stroke at high-low-high and low-high low sequences. The two photocurrent components are most obvious when looking at the transition from light off to light on in Fig.5(a). For data rates below the multi Gbit/s range, especially for 140Mbit/s and 560Mbit/s the transistor is a suitable detector without any restriction with respect to speed. By use of a limiting amplifier data rates up to 2.5Gbit/s can still satisfactorily be detected.

4 – SUMMARY

In summary modern self-aligned silicon bipolar transistors are promising devices for on-chip detection of light with a wavelength smaller than 900nm and data rates up to 2.5Gbit/s, the benchmark of future optical communication networks. The investigated device is the base-collector diode of an integrated vertical npn transistor, fabricated in a double poly Si self-aligning bipolar process.

REFERENCES

[1] Wieder A.W.; Siemens Forsch.– & Entwicklungsber., **13**, (1984), p.246
[2] Wieder A.W.; SSDM Conference on solid state devices and materials, Tokyo, (1986), p. 261
[3] Bock W., Prettl W.; to be published in *Electronics Letters*
[4] Poon E., Hwang W., Yang E.S., Evans H.L.; J. Appl. Phys. **57**(2), (1985), p. 338
[5] Sawyer D.E., Rediker R.H.; Proceedings of the IRE, (1958); p.1122

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

AN IMPROVED FULLY CMOS COMPATIBLE BIPOLAR STRUCTURE

A. GÉRODOLLE, G. GIROULT, S. MARTIN and A. NOUAILHAT

CNET/CNS, BP 98, Chemin du Vieux Chêne, F-38243 Meylan Cedex, France

Résumé - Une technologie bipolaire compatible CMOS étudiée au CNET a été simulée par le programme TITAN 5. Pour diminuer la résistance collecteur, une méthode incluant la formation de tranchées à fond perméable a été imaginée. La simulation process permet d'étudier l'intérêt d'une telle technique.

Abstract - A fully CMOS compatible bipolar technology is studied. 2-D process simulations of the device have been carried out before wafer processing ; comparisons of the results with experiments are presented. A new idea for making the collector, without standard epitaxy of a highly doped buried layer, is described. The 2-D simulation results enable the advantages to be checked with regard to the classical approach.

1- INTRODUCTION

To decrease the collector access resistance in BICMOS technology, the classical techniques are to form a buried collector by epitaxy, or to make a highly doped "retrograde well" with high energy and dose implant /1/. This latter technique could be used in CMOS technology, but it requires high dose and energy machines and is sensitive to certain mask problems.

The alternative idea presented here is to use trenches (filled with highly doped polysilicon) as the collector contact as well as additional dopant source for the collector, acting as a buried layer. This technique has the advantage of not needing any additional thermal or ionic implant step. The only additional steps, which come just after the formation of the locos and implantation of the retrograde well, are the formation, oxidation and polysilicon filling of the trenches. This technique is fully CMOS compatible, and avoids selective epitaxy for the buried layer of the collector. But, of course this is not simple to achieve. Process simulation of this structure has been performed, in order to validate the theoretical interest of this technology. Different values have been tested for the doping concentration and trench depth.

2- PRESENTATION OF BICMOS MAIN STEPS

The process flow is the same as in the 1 micron CMOS technology developed at the CNET, with only one additional mask level. The bipolar transistor process steps are those of the Nwell PMOS transistor, except for one additional ionic implant of boron to adjust the intrinsic base and the gate oxide etching before polysilicon deposition. These two process steps use the same mask.

The main features are the following : the gate arsenic-doped polycide acts as the emitter contact and as a doping source for the emitter. After polycide etching, a lightly doped extrinsic base region (LDEB) is implanted before deposition of the oxide-sidewall, to reduce the base resistance (Rbb'). SiO_2 spacers (0.3 microns wide) are used for the self-alignment of the extrinsic base. The drive in of the emitter and the dopants activation are obtained during the final RTA. Note that, without the additional implantation of the intrinsic base, a very shallow base can be obtained, resulting from the diffusion of the boron which has been implanted into the silicide at the extrinsic base implantation step. The N well acts as the collector. In order to improve the performances of the bipolar and the resistance to the latch-up of the CMOS devices, a retrograde well is used.

3- SIMULATION

The process simulations of the transistor have been performed using the TITAN-5 program, which allows multi-mayer ionic implantation and diffusion to be simulated. This helped to choose the energy and dose of the additional base implant. Profile calculations are made in polycide and silicon ; values of the segregation coefficient and diffusivity have been chosen in accordance to /2/ and the results are in good agreement with SIMS and electrical measurements (Figure 1). A complete study of arsenic and boron codiffusion in the three layers WSix, polysilicon, silicon is to be published in the near future.

Simulation of the complete structure, including the collector contact, can be provided, using a very sharp meshing in the active area (Figure 2).

4- SIMULATION OF AN IMPROVED BIPOLAR TRANSISTOR : A BURIED DIFFUSED COLLECTOR

The structure simulated is shown in Figure 3. The trenches are formed after the locos formation, in the same way as insulation trenches ; they are then slightly oxidized, and the oxide and the silicon are etched anisotropically, so that the trench becomes permeable to the dopant species. The trenches are then filled with highly doped polysilicon (Phosphorus, 10^{20} to 3.10^{20} atoms/cm^3). During the thermal annealing used for the retrograde well (150 mn, 1050°C), the dopant diffuses in the silicon, making a zone equivalent to the collector buried layer of the epitaxial technology. The simulations have been made on a 4 micron wide structure. Three simulations have been carried out with a 1.6 µm wide and 2.2 µm deep trench, with a phosphorus concentration from 10^{20} to 3.10^{20} atoms/cm^3. At each time, the square resistance of the collector has been calculated for several cross-sections (Figure 4).

With 10^{20}, the square resistance remains 1.2 kΩ from X = 3µm to 4µm. To obtain a resistance of 100 Ω the position should be at X = 2µm, which means that the dimensions have to be reduced by 2µm, which is not feasible.

With 2.10^{20}, the resistance is 1.2 kΩ at X = 4µm, 393 Ω at X = 3µm. This means that reducing the dimensions by only 1µ is sufficient to obtain a real improvement in the resistance.

With 3.10^{20}, the resistance is 973 Ω at X = 4µm, 80 Ω at X = 3µm. The collector access resistance value is low, but the highly doped collector area is too close to the base. To overcome this problem, another simulation with a deeper trench has been carried out. With a phosphorus concentration of 3.10^{20}, and a 3 µm deep trench (Figure 5), the transistor active area remains optimized, as expected. Moreover, the collector access resistance has been reduced : 43 Ω at X = 3µm, 134 Ω at X = 3.5 µm, 435 Ω at X = 4 µm. This is due to the fact that the total phosphorus dose in the polysilicon trench is higher, and therefore the phosphorus depletion of the trench is slower. At the end of the process, the remaining phosphorus concentration in the polysilicon is 2.10^{20} atoms/cm^3, instead of $1.5.10^{20}$ with the 2 µm deep trench.

5- CONCLUSION

A general 2-D simulator, which makes no assumption on the structure, can be used as a predictive tool, to validate new ideas, before starting the process itself. In this case a new technique for processing fully CMOS compatible bipolar transistors without epitaxy has been investigated. Simulations show that a very low collector access resistance can be obtained without disturbing the transistor behaviour, provided a 3 µm deep trench, which is dopant permeable at the bottom and insulated at the sides, can be achieved.

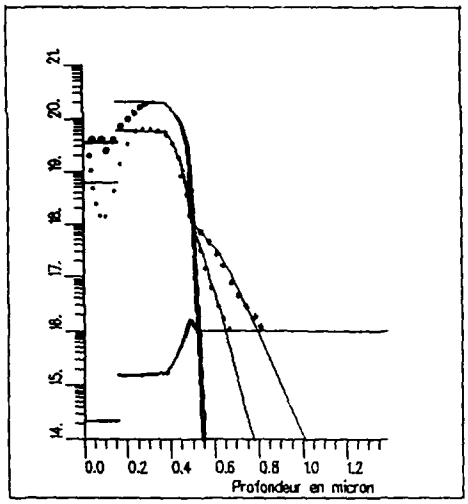


Fig. 1 - Intrinsic transistor area. Cross-section of As and B concentrations. SIMS measurements:
- * B with boron adjustement implantation
- • B without (the same as * in polycide)
- ● As in both cases
- ___ Simulation results

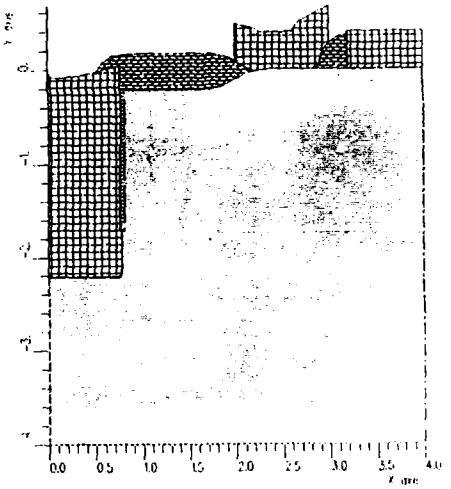


Fig. 3 -structure of the Bipolar T. The trench and emitter widths have to be divided by a factor 2 for the simulation. The real trench width is 1.6 microns.

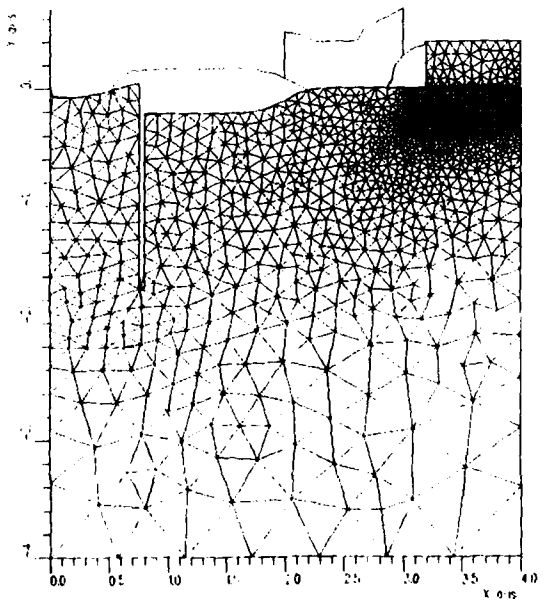


Fig. 2 -Exemple of meshing.

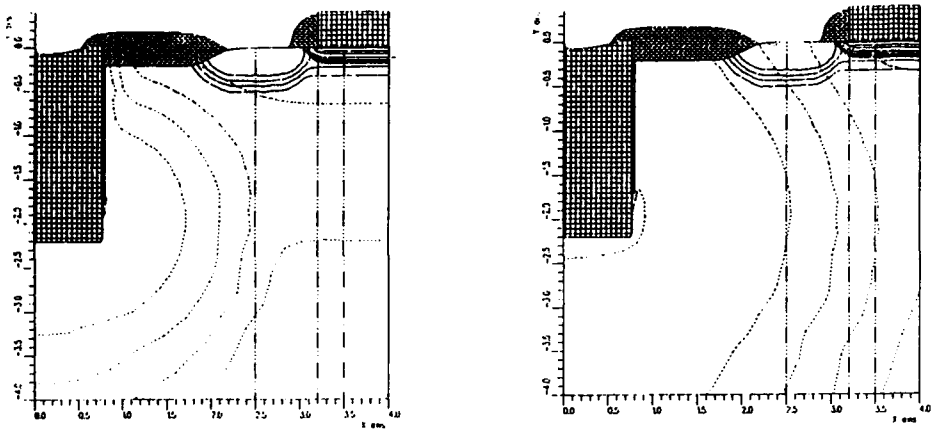


Fig. 4 -isoconcentrations of phosphorus corresponding to 10^{20} (a) and 3.10^{20} (b) atomes/cm3

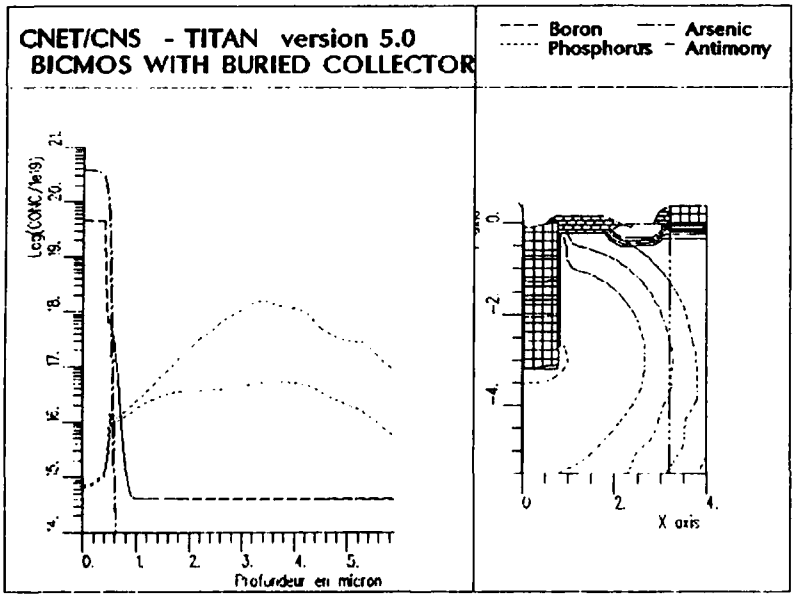


Fig. 5 -cross-sections of the doping concentration at X=3.2 um and X=4 um

The authors would like to thank M. Gauneau (C.N.E.T. Lannion) for SIMS analysis.

References

/1/ A.R. Alvarez, 'BICMOS technology',short course IEDM 1987
/2/ S.P.MURARKA, D.S. WILLIAMS, J.VAC.SCI.Technol.B,Vol5,N06,Nov/Dec 1987

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

WELL - OPTIMIZATION FOR HIGH SPEED BICMOS TECHNOLOGIES

H. KLOSE, T. MEISTER, B. HOFFMANN, J. WENG and B. PFÄFFEL

Siemens AG, Corporate Research, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

Abstract - A high speed BICMOS process with a polysilicon bipolar transistor is presented. Using this technology the well optimization is outlined. Different approaches to construct the well are compared to improve the high current behaviour of the bipolar transistor. Influences on the parasitics are discussed. Using the optimized process version bipolar transistors with a cut-off frequency of 9.5 GHz were fabricated.

1 — INTRODUCTION

Among the many tasks in the development of a BICMOS process the optimization of the well is the most specific one. The reason for this is that the well configuration necessary for appropriate performance of all BICMOS devices is not used in pure CMOS or in bipolar processes. Although there are many reports on BICMOS published so far, this is the first one describing a thorough optimization of this process block.

In this presentation we first give a brief description of the BICMOS process which we used as base-process for optimization. The main feature of this BICMOS process is the use of arsenic doped polysilicon for the formation of the emitter. Next the general optimization strategy is outlined. Three approaches are presented then to maximize the current drive capability of the bipolar transistor :

- 1 : single antimony buried layer (Sb - BL)
- 2 : combined antimony/phosphorus buried layer (Sb/P - BL)
- 3 : implantation of doubly ionized phosporus to increase the bipolar well doping (Sb — BL + P^{++})

Pros and cons of these approaches are discussed with respect to the most important characteristics of a BICMOS process : cutoff frequency and Kirk current of the bipolar transistor, drain junction capacitance of the PMOST, buried layer to buried layer spacing and the current gain of the parasitic vertical pnp bipolar transistor.

2 - PROCESS DESCRIPTION

The optimization approach is outlined using a BICMOS process which has the following features :

CMOS part : - 1.2 µm n - well CMOS process
- 2 layers of metal

bipolar part : - antimony buried layer
- collector sink
- polysilicon emitter bipolar transistor

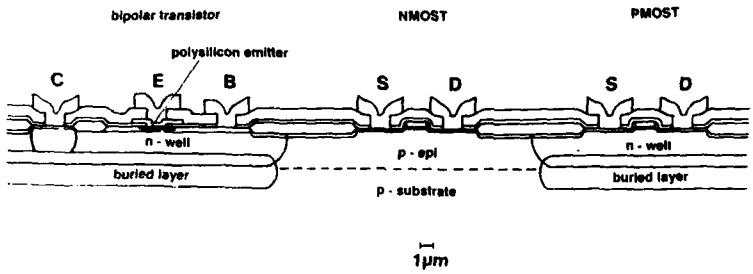


Fig. 1 - Cross-section of the BICMOS devices.

The cross-section of the resulting BICMOS devices is sketched in Fig. 1. As depicted in this figure the buried layer is used both beneath the bipolar transistor to reduce the collector series resistance as well as beneath the PMOST to avoid punch through between its drain and the p-type substrate. Reducing the punch on depth[1] of the buried layer improves the high current behaviour of the bipolar transistor significantly. At too small values of this quantity, however, the junction capacitance of the PMOST's drain is increased.

3 - OPTIMIZATION STRATEGY

The main goal for the optimization of the well process was both not to degrade the performance of the MOS transistors in comparison to the device performance of the underlying CMOS process as well as to have a bipolar transistor capable of driving high current densities without speed degradation. Thus for optimization a tradeoff has to be done between contrasting requirements.

First we checked how shallow the epi-layer can be designed without significant influence on the device performance of the PMOST. To monitore this the area-specific drain junction capacitance was used. The goal was to keep the increase of the value of the drain capacitance averaged over the 5V swing of CMOS logic below 10%.

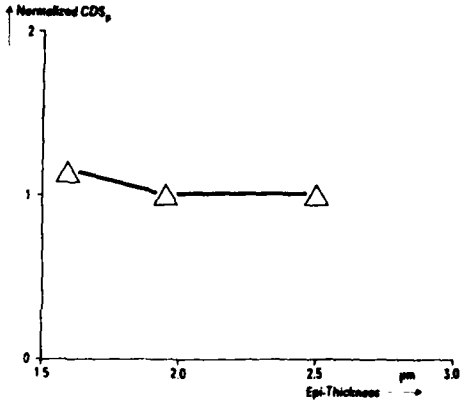


Fig. 2 - Dependence of the averaged drain junction capacitance of the PMOS transistor vs. epi -layer thickness

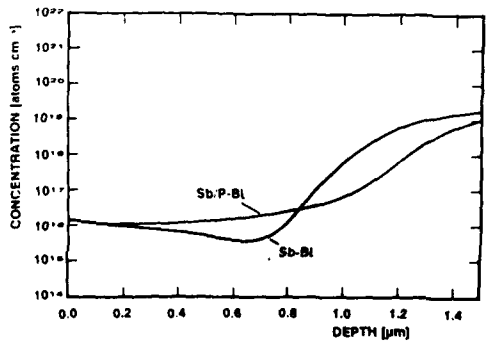


Fig. 3 - Doping profiles of the single antimony buried layer process and the combined antimony/ phosphorus buried layer process version (SUPREM-III simulation)

[1] Distance between the silicon surface and the 10^{17} cm^{-3} concentration point at the rising edge of the buried layer doping profile

The results of this experiment are shown in Fig. 2. Doing so we ended up with a minimal punch-on depth at the end of the whole process of approximately 0.85 to 0.9µm. The respective epi - layer thickness is 1.6 µm (see Fig. 2). With respect to this boundary condition the high current behaviour can only be improved further by increasing the collector doping concentration via a combined antimony/phosphorus buried layer or a high energy implant of phosphorus into the bipolar well. None of these two approaches require additional masking steps.

4. - RESULTS

4.1 - THE CUT-OFF FREQUENCY

SUPREM-III simulation of the process version with the single Sb-buried layer show that there is a considerable concentration dip right in front of the antimony buried layer (Fig. 3). Due to the low doping concentration within the dip that value of j_C which marks the 10% roll off point of f_T is as low as 0.16 mA/µm^2 (Fig. 4). This value of j_C is furtheron denoted as the Kirk-current density j_K.

One way to eliminate this concentration dip is the use of the combined Sb/P-buried layer. This kind of a conducting layer is formed by a double implantation using both antimony as well as phosphorus as dopant material. Thus the n-well above the buried layer is not only formed by 'down' - but also by 'up'-diffusion of phosphorus. As the n-well doping is enhanced not only within the bipolar but within the PMOST regions as well, the epi-layer thickness was somewhat increased to maintain unchanged drain junction capacitancies. Comparing the f_T vs. j_C plots both for the simple Sb-BL as well as for the Sb/P-BL a considerable improvement in the Kirk-current density was reached (j_K = 0.16 mA/µm^2 for Sb-BL and j_K = 0.37 mA/µm^2, Fig. 4). This, of course, leads to an improvement of the maximum f_T value too from 6.7 GHz to 9.2 GHz.

The second possibility to increase the current drive capability of the bipolar transistor is to increase the bipolar well doping by an implantation of doubly ionized phosporus. Doing this together with the base implantation no additional masking step is necessary. The doping profile which results from this approach is given in Fig. 5. The respective f_T vs. j_C dependence is sketched in Fig. 4. The maximum f_T for this process version is 9.5 GHz. The respective Kirk-current density is 0.4 mA/µm^2.

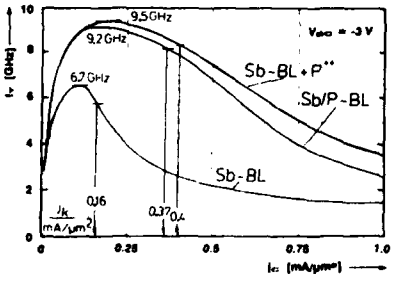


Fig. 4 - Measured cut - off frequencies vs. collector current densities for the process versions single antimony buried layer, combined antimony /phosporus buried layer and the enhanced bipolar n-well by doubly ionized phosporus implantation

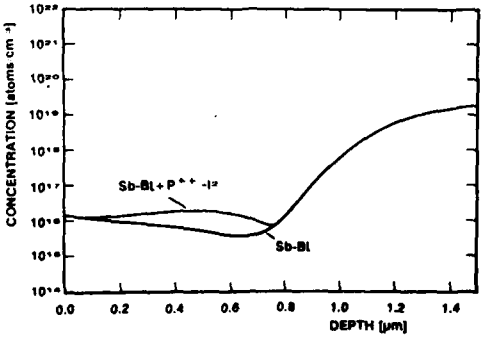


Fig. 5 - Doping profiles of the single antimony buried layer process version and the enhanced bipolar n-well version by implantation of doubly ionized phosphorus (SUPREM - III simulation)

4.2 - BURIED LAYER TO BURIED LAYER SPACING

One of the main quantities which determines the packing density of mixed CMOS/bipolar circuits is the BL-BL spacing. Due to the use of the shallow LOCOS isolation in BICMOS processes the channel stopper does not block the space charge regions of the buried layers as effectively as in pure bipolar processes.

	BL-BL spacing/µm	normalized β_{pnp}
Sb - BL	6	1
Sb/P - BL	8.5	$0.8 \cdot 10^{-1}$
Sb - BL + P^{++}	6	10^{-1}

Tab.1 - BL-BL spacing and the nomalized current gain β_{pnp} of the parasitic *pnp*

Because of the lateral out-diffusion of the phosphorus buried layer (Sb/P - BL) during the buried layer and the *n*-well drive- in the BL-BL spacing is quite large compared to those of the other apprcaches. The spacing of these process versions are equal of course because the P^{++} ions are implanted at a late stage of the process flow. Thus no significant diffusion occurs.

4.3 - CURRENT GAIN OF THE PARASITIC PNP

Again due to the shallow LOCOS isolation each bipolar transistor which is not entirely surrounded by a collector sink guard ring is a possible source of leakage currents in case of saturated bipolar action. The reason for this is the parasitic vertical pnp which is formed by the inactive base and the collector of the regular npn transistor together with the p-type substrate. To reduce possible leakage currents β_{pnp} has to be as low as possible. The results again are given in Tab.1. Due to the concentration dip which occurs in case of the single Sb-BL process (see Fig. 3) the current gain of this parasitic pnp device is far too high. Compared to that both the approach with the Sb/P buried layer as well as with the P^{++} implantation yield an improvement of one decade approximately.

5 - SUMMARY AND CONCLUSIONS

In this paper a high speed BICMOS process is presented. Different schemes to construct the well were compared. It was shown that the single antimony buried layer combined with a doubly ionized phosphorus implantation is the best approach both with respect to bipolar performance as well as with respect to reduced parasitic effects. The cut-off frequency obtained with this process version is 9.5 GHz. The current density which marks the 10% roll off point is 0.4 $mA/\mu m^2$. ECL ringoscillators with a voltage swing of 2•200mV designed with minimum size bipolar transistors and a one sided base contact show minimum gate delays of 140 ps. Thus the BICMOS process presented is a promising candidate for mixed ECL/CMOS circuits on the same chip.

ACKNOWLEDGEMENTS :

The work presented here is supported by the ESPRIT project 412.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

AN ULTRA HIGH SPEED TRENCH ISOLATED DOUBLE POLYSILICON BIPOLAR PROCESS

M.C. WILSON, S. DUNCAN and P.C. HUNT

Plessey Research (Caswell) Ltd., Caswell, Towcester, Northants, Great-Britain

Abstract - This paper describes a double polysilicon bipolar process incorporating a novel self-aligned emitter-base and deep trench isolation. The process has been designed primarily for ultra high speed by minimising parasitic capacitances, and also offers high packing densities. The performance of the technology is demonstrated by a 1/8 static divider operating at a frequency of 10.7GHz. Other representative SSI functions are also described.

1 - INTRODUCTION

Recent advances in silicon bipolar technology have been responsible for sub 50ps gate delays in ring oscillator circuits and demonstrated maximum operating frequencies of 10.38GHz in high speed divider circuits [1,2]. Generally, this performance is attributed to a device structure incorporating double overlapping polysilicon and emitter-base sidewall spacer technology, thus minimising parasitic capacitances due to a reduction in device size.

Several bipolar processes to date incorporate a modified LOCOS isolation [3,4]. This paper describes a 1µm process primarily aimed at ultra high speed circuits, and includes a two layer polysilicon structure with self-aligned emitter-base and utilises deep trench for isolation of the device. This potentially offers low buried n+ to channel stop capacitance, high packing density, and near planar topography for compatibility with multi-level metallisation systems. A 1/8 frequency divider operating at 10.7GHz is reported along with several other small scale demonstrator circuits.

2 - PROCESS TECHNOLOGY

Device processing begins with the growth of a 1µm epitaxial layer over a 20 ohms/□ buried n+ collector. Trenches 1µm wide by 5.5µm deep etched through a sacrificial CVD oxide mask provide device isolation. A channel stop implanted at zero angle controls the parasitic MOS threshold and suppresses the gain of the parasitic buried n+/substrate/buried n+ transistor. All masks are stripped and the trenches are refilled with a composite dielectric sidewall and polysilicon. The sidewall comprises thermal oxide, CVD nitride and a layer of low temperature oxide (LTO). Following polysilicon etchback, the LTO layer is stripped and the active area patterned. Field oxide is then grown resulting in the structure shown in figure 1(a). The continuous nitride layer lining the trench and active area prevents bird's beak encroachment.

The process sequence up to emitter-base formation includes an n+ implant to form the deep diffused collector, the removal of dielectrics over the active area, and the deposition and patterning of the p- poly extrinsic base contacts. This layer of polysilicon is protected by nitride over the active areas and oxidised in the field region. This provides thicker field oxide for minimised interconnect capacitance, and a planar topography for metallisation. The nitride is susequently removed (see figure 1(b)).

Emitter-base formation starts with the deposition of LTO over the p- polysilicon. A 1µm emitter window is anisotropically etched through the LTO and p- poly layers. Thermal oxide is grown and a layer of nitride deposited, during which p-type dopant is allowed to diffuse from the polysilicon into the single crystal to form the extrinsic base. The intrinsic base is then implanted through the window and a conformal layer of polysilicon deposited. This is then anisotropically plasma etched to form sidewall fillets 0.2µm wide as shown in figure 1(c).

The sacrificial polysilicon fillets protect the periphery of the nitride and oxide during subsequent dry and wet etching to leave the emitter window (figure 1(d)). The fillets are then removed by chemical etch. This method leaves the implanted base region free from plasma etch damage. A TEM cross-section of a fully formed sidewall is shown in figure 1(e).

A second polysilicon layer is deposited and implanted with arsenic. After appropriate patterning and etching, the emitter-base junction and collector contact are formed by rapid thermal annealing. The collector area in this process sequence is treated in the same way as the emitter. A cross section of this structure is shown in figure 1(f).

The completed device has a self-aligned PtSi/TiW contact structure and standard Al/Cu interconnect. Initial circuits were designed with a two level 5µm metal and via pitch. More advanced circuits were successfully manufactured with a 3.5µm contact pitch on a three level polyimide based metallisation system. A completed device is shown in figure 2.

3 - DEVICE PERFORMANCE

A gummel plot showing the DC characteristics of a typical device is shown in figure 3. The excellent low current gain behaviour is attributable to the use of a thermal oxide as part of the sidewall spacer rather than deposited oxide. The shallow emitter-base structure is a result of implanting the base through dielectric coupled with a polysilicon emitter and RTA processing. An initial version of the process produced Ft values of 14GHz. Subsequent development has resulted in a range of emitter-bases with Ft's of 17-22GHz. Ft curves of a standard and advanced device are shown in figure 4, where in both cases the effective emitter area was 0.6 x 14.4µm.

The use of trench isolation has reduced the collector-substrate capacitance to half that of LOCOS and improved packing density. This significantly reduces all parasitic cacpacitances and improves circuit performance. A list of basic electrical parameters for a standard 1 x 5µm device is shown in table 1.

4 - CIRCUIT PERFORMANCE

The high Ft of the process has enabled a substantial number of high performance circuits to be designed and fabricated. For example sub 40ps gate delays have been achieved on ring oscillators with a fanout of 1. However, ring oscillators are a poor example of high speed device technology. A more impressive demonstration of the process is a divide by 8 function operating at a frequency of 8.8GHz. This incorporated an emitter-base having an Ft of 17GHz, however, the more advanced emitter-base with an Ft of 22GHz enabled the same divide by 8 function to operate at 10.7GHz. The input and output waveforms of this static frequency divider are shown in figure 5.

Other SSI functions successfully processed and evaluated include a 1Kb RAM, a 1.6GHz universal shift register and an 8 bit DAC with a 1.2ns settling time. A micograph of the DAC is shown in figure 6. Several other pathfinder products demonstrating low power consumption and high toggle rates have been reported elsewhere [5],[6].

5 - CONCLUSION

A very high speed manufacturable bipolar transistor process has been described. Trench isolation and the new emitter-base sidewall technology provide high performance, low capacitances and high packing densities. Several SSI functions have been evaluated incorporating transistors operating with peak FT's of between 14GHz - 22GHz. The performance levels obtained from these circuits combined with the yield of a silicon based process make ultra high speed VLSI a reality.

ACKNOWLEDGEMENTS

The authors wish to thank the staff of the Silicon Technology Department at Caswell who were involved with the project. Also thanks go to Plessey Semiconductors for support and circuit design contributions, and to the Alvey Directorate for partial financial support.

REFERENCES

[1] S.Konaka, M.Suzuki, Y.Tammamoto and T.Sakai - Solid State Devices 1985, edited by P.Balk and O.G.Folberth, Elsevier Science Publishers B.V., 1986 - Printed in the Netherlands.

[2] K.Washio, T.Nakamura, K.Nakazato, and T.Hayashida - ISSCC Dig. Tech. Papers, p.48 1987.

[3] K.Y.Chui, J.L.Moll, and J.Manolui - IEEE Trans. Electron Devices, vol. ED-29, no. 4, April 1982.

[4] P.Zdebel, R.Balda, B.Y.Hwang,Y.d.l.Torre and A.Wagner - IEEE 1987 Bipolar Circuits and Technology Meeting, 1987, pp.172-175.

[5] P.C.Hunt and M.P.Cooke - Proceeding of IEEE 1988 Custom and Integrated Circuits Conference, Rochester, New York, May 16-19.

[6] S.Duncan, M.C.Wilson, P.C.Hunt and D.J.Bazley - 1988 Symposium on VLSI Technology, Digest of Tech. Papers, May 10-13.

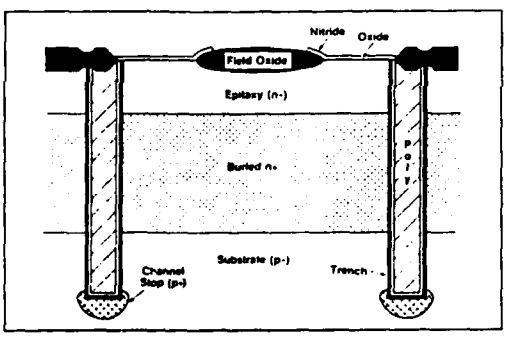


Fig.1(a) Device Isolation formation

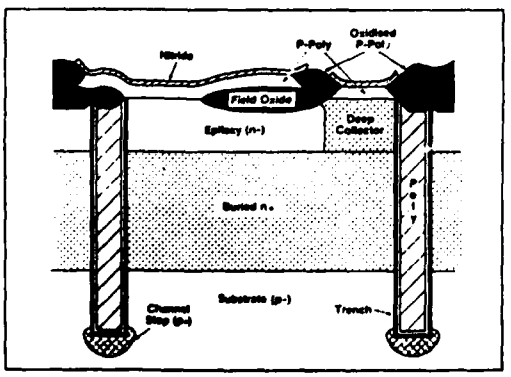


Fig.1(b) LOCOS of 1st layer of polysilicon

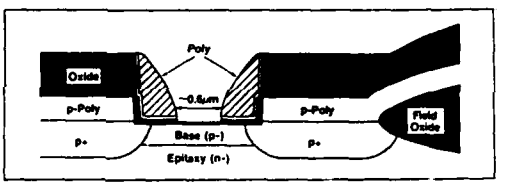


Fig.1(c) Emitter-base sidewall spacer formation

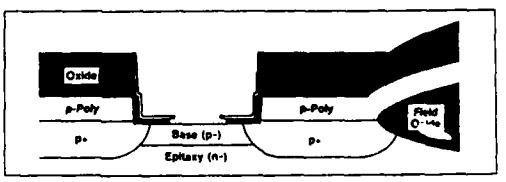


Fig.1(d) Spacer after polysilicon removal

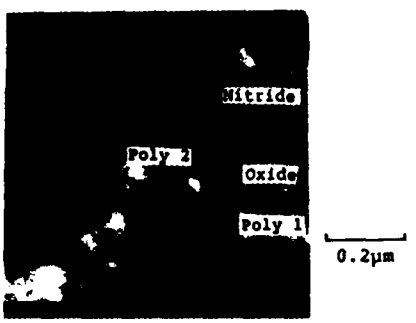


Fig.1(e) TEM cross-section of emitter-base spacer

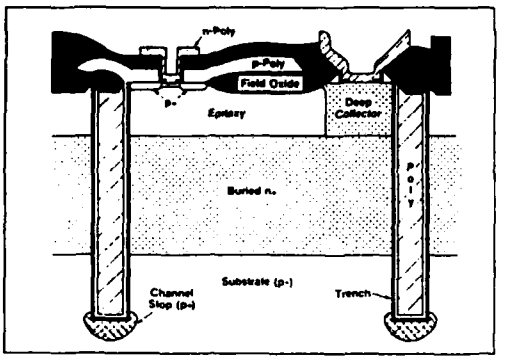


Fig.1(f) Device prior to metal deposition

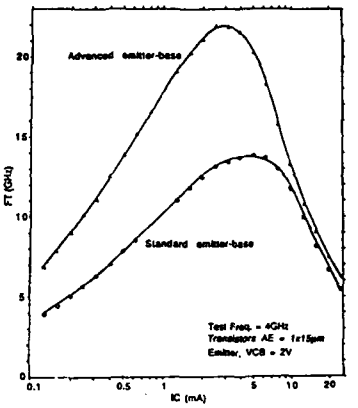


Fig.4 Ft curves of a standard and advanced emitter-base transistor

Fig.2 SEM micrograph of a completed transistor, the contact pitch is 5µm

Table 1 Typical electrical parameters of a 1x5µm^2 emitter device

Operating voltage	5V
Forward beta	100
Base resistance	150Ω
Collector resistance	75Ω
Emitter resistance	10Ω
Emitter-base capacitance	19fF
Collector-base capacitance	17fF
Base transit time	6.5pS
Maximum cut-off frequency	14GHz

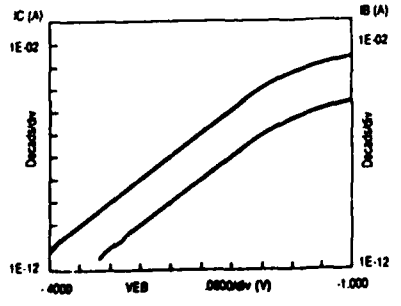


Fig.3 Gummel plot of a 1x5µm^2 emitter transistor (as drawn)

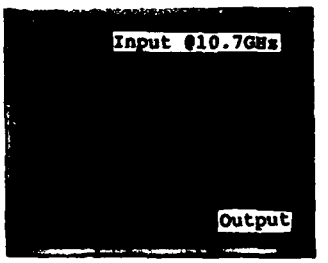


Fig.5 Oscilloscope trace of the input and output waveforms of a divide by 8 function running at 10.7GHz

Fig.6 A photograph of the 8 bit DAC

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A TRENCH ISOLATED HIGH SPEED BIPOLAR PROCESS FOR A 10K GATE, 950MHz, VLSI CIRCUIT

C. MALLARDEAU, M. ROCHE, M. DEPEY, F. DELL'OVA, D. THOMAS, D. CELI, P. HUNT* and A. HEFNER**

SGS-Thomson Microelectronics, BP 217, F-38019 Grenoble Cedex, France
**PLESSEY Research Ltd, Allen Clark Research Centre, Caswell Towcester, Northants, GB-NN12 8EQ, Great-Britain*
***Telefunken Electronic GmbH, Therensien Strasse, Postfach 1109, D-1700 Heilbronn, F.R.G.*

RESUME
Une technologie bipolaire rapide, de dimension critique micronique, utilisant l'isolement par sillons et un emetteur en polysilicium a été evaluée et caracterisée. Une démonstration du rendement de la technologie a été réalisée sur un circuit VLSI de 10000 portes.

ABSTRACT
A one micron bipolar process using trench isolation and polysilicon emitter has been evaluated and characterised. The demonstration of yield was achieved on a 10000 gate VLSI circuit.

1. INTRODUCTION

Very high speed circuits are required for super computers and communication systems. Bipolar technology has remained the technology of choice for applications where circuit speed is of primary importance, it has also demonstrated its VLSI capability as well. The key elements of the advanced bipolar technology are filled deep trench isolation, self aligned emitter base structure and polysilicon emitter contact (1),(2).

Both the self aligned structure and the trench isolation greatly reduce the device area and the associated parasitic capacitance, and hence significantly reduce the power delay product and increase the density of bipolar devices. However, yield is hard to assess with a trench isolation process, as it remains difficult to perfectly fill the trenches avoiding holes and crevices, and also to stop defect generation from the top corners of the trench. In order to characterise the trench isolation process in terms of defect generation and reproducibility, an advanced process has been set up (3), with one micron critical dimension, using only one polysilicon level for polyemitter. A 10K gate VLSI circuit was processed with this technology, in order to demonstrate the yield of a trench isolated bipolar process.

2. MAIN FEATURES OF THE PROCESS

The process is set up using 1 µm design rules .The photolitographic steps are implemented on CENSOR SRA 9535 , allowing submicron features . The lateral isolation between components is provided by deep trenches. The emitter-base structure shows improved performance by adding a polysilicon layer between the metallic contact and the monosilicon emitter region . Finally the interconnect system requires one polycide and two metal levels. Figure 1 shows a cross section of the basic transistor.

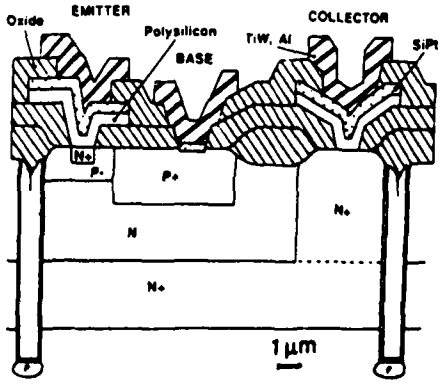


Fig.1- Cross section of the basic transistor

Deep trench isolation

After full sheet As buried layer implantation followed by 1.2 µm thick epitaxy deposition , trenches of 1.2µm wide and 5µm deep are used for lateral isolation .Special requirements for trench isolation have been described previously (4); and a conventionnal process of trench etching and filling was developed as follows :

_ reactive ion etching of monosilicon with CCl_4 and N_2 gases using an etching mask of stacked layers of thermal silicon dioxide, LPCVD nitride and PECVD oxide . A satisfactory trench profile was achieved with this process , as shown in figure 2..

_ oxidation of the trench sidewalls to provide dielectric isolation

_ channel stopper implantation at the bottom of the trench

_ conformal undoped polysilicon deposition in the trenches, and etch back planarisation with a sacrificial resist layer

_ oxidation for field oxide and polysilicon capping formation with a conventionnal locos technique

In order to demonstrate yield on a VLSI circuit, specific studies were carried out to achieve low defect generation and a uniform polysilicon planarisation.

-The trench bottom can be rounded by controlling the redeposition of oxide during RIE , to prevent stress from the bottom corners to appear. The oxide thickness on the trench sidewalls was also optimised. For field and capping oxidation , the temperature had to be increased up to 1000°C to prevent screw dislocations to propagate from the top corner of the trench . In these conditions no defects were revealed by Wright attack (fig.3).

-Amorphous silicon is conformaly deposited in the grooves . The thickness of the layer is ajusted according to the trench width . The etch back process is done with a sacrificial resist layer , with an end point detection on the nitride layer. Then an approriate oxidation followed by a wet etch is performed to remove the remaining polysilicon . With our process , the uniformity on wafer is satisfactory , and the integrity of the nitride is kept.

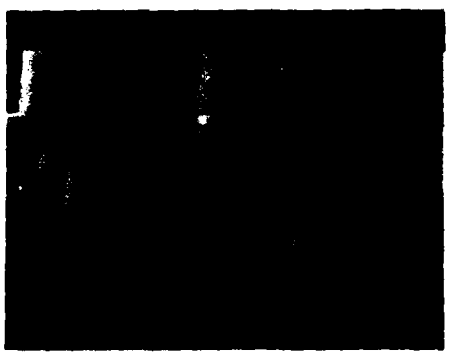


Fig.2- Trench profile

Fig.3- Defect analysis with Wright etch

Emitter base structure

The improvement in speed for bipolar transistors requires very shallow emitter base junction . However , reducing the distance between the metallic contact of the emitter and the emitter-base junction leads to an increase of the injection of minority carriers into the emitter , and thus creates a degradation of the current gain. The problem is solved by adding a polysilicon buffering layer between the metallic contact and the monosilicon emitter, to remove the recombination zone further, and to put an interface barrier to the diffusion of minority carriers.

After intrinsic base implantation with boron, a 200nm thick amorphous Si layer is deposited at 560°C under low pressure in the emitter contact. To minimize oxide contamination, an HF dip is performed prior to polysilicon deposition. Then the emitter is implanted with arsenic, and after silicon etch, a 10mn-975°C thermal treatment leads to a junction depth of 100nm with a 140nm wide base. SEM of this structure is shown in figure 4. The doping profile obtained from spreading resistance measurements is ploted in figure 5.

Fig.4- SEM cross section of the intrinsic transistor

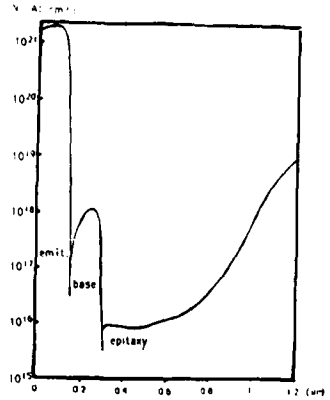


Fig.5- Spreading resistance measurements of the emitter base structure

Interconnect

The interconnect system consists of one polycide and two metal levels . The pitches are 3µm ,6µm and 9µm respectively .
A 80 nm thick PtSi is formed on monosilicon and polysilicon contact in order to realise one interconnect level inside the cells. A 500 nm thick plasma CVD SiO_2-P_2O_5 isolates metal 1 from polycide . Metal 1 consists of 150nm thick TiW barrier and 500nm thick AlSiTi The interlevel dielectric between metal 1 and metal 2 is made of 150 nm plasma CVD Si_3N_4 and a planarising layer of polyimide . The nitride layer is used to stop hillock formation. Metal 2 consists of 1.2µm thick AlSiTi. The circuit is finally passivated with polyimide.

3-PROCESS EVALUATION

The one micron demonstrator was processed on wafers, with the technology described above. The basic transistor is made of a 3-side trench walled base and a non walled emitter. The transistor area is 109µm2, and the size of the emitter-base junction is 2.4x1µm^2. A SEM cross sectionnal view of the NPN basic transistor is shown in figure 6.

Fig.6- SEM cross sectionnal view of the basic transistor

Table 1- Electrical characteristics of the basic transistor

Transistor area	109µm2
Base area	46µm2
Emitter area	2.4x1µm2
Intrinsic gain BF	120
IKF	5.1mA
RE	13.4Ω
BVEBo	6.21V
BVCBo	21V
Buried lay/subst.	46.4V
BVCEo	7.8V
CJE	11.3fF
CJC	16.4fF
CJS	7fF

Electrical characteristics of the basic transistor of the demonstrator are given in table1.
The Gummel plot (fig 7) shows a nearly ideal behaviour down to 100pA collector current, for the basic transistor. The current gain remains almost constant over five decades of collector current.
The cut-off frequency is 8GHz at VCE=4V and Ic=700µm. For a transistor with a longer emitter (hence smaller base resistance) than the one of the basic transistor, the measured Ft is over 12GHz (fig.8)

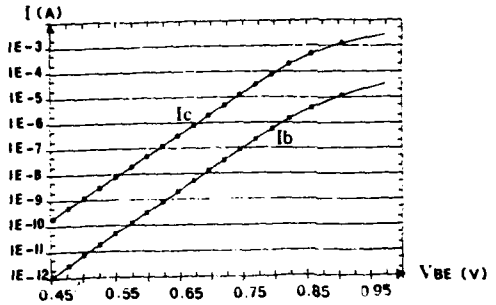


Fig.7- "Gummel plot" of the basic transistor Ic,Ib versus VBE at VCE=0

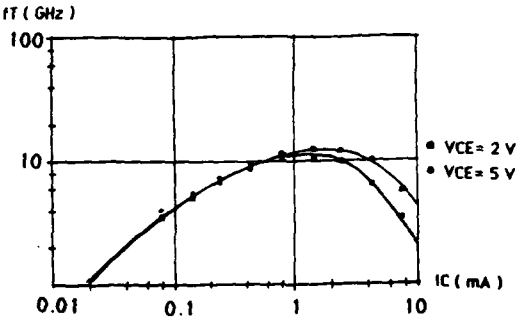


Fig.8- Ft versus Ic at VCE=5V (NPN TR3 - Em= 1x7.2 µm2)

4-YIELD ASSESSMENT ON A 10K SEA OF GATES

The functionnality of a 2048-stage shift registers was demonstrated. The aim of this demonstrator was twofold: first to illustrate the process capabilities in terms of speed, density, and power consumption, secondly to estimate the yield of the process The overall circuit uses 35216 transistors and is quite illustrative of the capabilities of the process technology. *Figure 9 shows a micrograph of* a completed chip. The chip size is 8.16x10.6mm². High frequency and speed were tested on packaged devices, demonstrating performances of 950MHz clock frequency for the shift register, with 1.8mW power consumption per stage, and 95ps propagation delay per stage for 21-stage ring oscillators processed with the same technology.

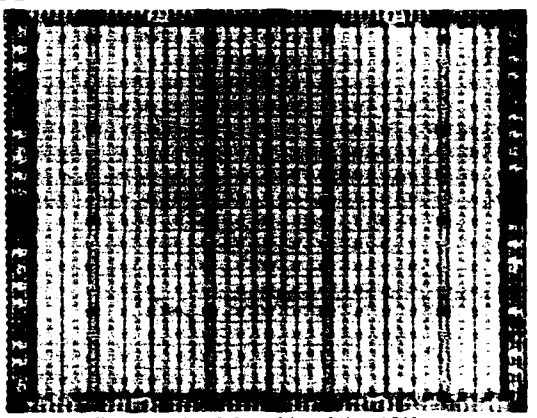


Fig.9- Micrograph of the chip of the 10K sea of gates (chip size: 8.2x10.1mm2)

5-CONCLUSION

The capability of a one-micron trench isolated bipolar technology to implement high speed circuits has been demonstrated on a 2048-stage shift register. Performances of 950MHz clock frequency were obtained. This circuit is very suitable for yield monitoring of a trench isolated bipolar process.

Acknowledgements

The authors would like to thank all the members of the project teams, especially B.Gros, C.Paillet, E.Blanchet, D.Neira and Y.Morand for process set up, B.Solignac, and F.Poilane for their assistance in characterizing the devices.

This work is supported by the EEC program ESPRIT under contract n°243.

REFERENCES

(1) M.VORA et al, IEDM Tech. Dig. (1985) p.34
(2) M.SUZUKI et al, IEEE J. Sol. Stat. Circuits Vol SC22, N°1,(1987) p.41
(3).M.ROCHE et al, ESSDERC 87.
(4) D.D.TANG et al, IEEE J. SOL. Stat. Circuits, Vol SC17, N°5, (1982) p.925

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

SHALLOW DOPING PROFILES FOR HIGH-SPEED BIPOLAR TRANSISTORS

K. EHINGER, H. KABZA, J. WENG, M. MIURA-MATTAUSCH, I. MAIER, H. SCHABER and J. BIEGER

Siemens AG, Corporate Research and Development, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

Abstract - Within the framework of a double polysilicon self-aligned bipolar technology shallow base widths W_B for npn transistors were obtained by low-energy implantation of boron into either crystalline Si (W_B = 150 nm), preamorphized Si (W_B = 125 nm) or by diffusing boron out of polysilicon (W_B = 85 nm). Rapid thermal processing was used for the emitter drive-in. A transit frequency of 22 GHz was achieved for W_B = 85 nm. Comparison with one-dimensional calculations indicates the potential of further enhancement of intrinsic device speed by optimization of the proposed technological approaches.

1 - INTRODUCTION

The combination of modern self-alignment schemes together with advanced lithographic capabilities resulted in a drastic reduction of lateral dimensions and parasitics and improved integration density and speed-power performance of bipolar transistors. For further enhancement of intrinsic device speed, vertical scaling of doping profiles becomes increasingly important. Heavily doped n^+-polysilicon used as emitter contact and diffusion source for arsenic led to shallow emitter junctions, especially if the emitter is driven-in by rapid thermal processing. To obtain very shallow base profiles we investigated three promising approaches:

A) A low-energy implantation (10 keV) into bulk Si was used to form the active base. However, further reduction of base width W_B by lowering the implantation energy seems to be restricted because of boron ion channelling.

B) To suppress channelling the Si substrate was preamorphized by a Ge implantation prior to the base implant. Great care has to be taken to obtain defect-free recrystallization.

C) A completely different approach inherently avoiding such problems is the diffusion of base and emitter out of the same polysilicon layer.

2 - DEVICE FABRICATION

In double polysilicon self-aligned npn bipolar transistors heavily doped n^+- and p^+-polysilicon layers serve as emitter and external base contacts, respectively. These layers are separated by sidewall oxide spacers remaining after an unmasked reactive ion etch-back (RIE) of a conformal SiO_2 layer. If the active base is implanted prior to the formation of these spacers,

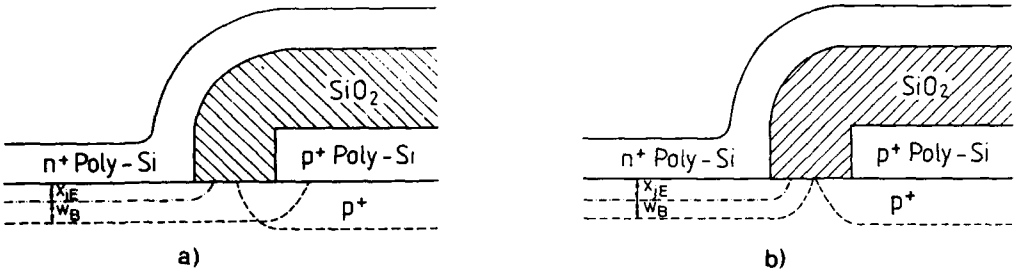


Fig. 1 - Schematic for dopant distribution of self-aligned bipolar with implanted base (a) and diffused base (b). In both cases the emitter is diffused out of the n^+-polysilicon.

perimeter emitter(E)-collector(C)-punchtrough due to incomplete overlap of active and inactive base (B) regions is easily avoided. Fig. 1a schematically shows the obtained dopant distribution in the vicinity of the oxide spacer. This advantage of a low resistive base link, however, has to be contrasted with the difficulty of having different reference planes for base implant and emitter diffusion due to the RIE step between base implant and emitter drive-in. With further vertical scaling this problem becomes more and more severe.

We fabricated devices with fairly shallow base-profiles by reducing the boron implantation energy to 10 keV. However, channelling of B^+ ions and their fast diffusion limits the formation of shallow B-profiles. Even worse, lowering the ion energy leads to an increased probability for channelling /1/. Therefore, further reduction of W_B merely by lowering the ion energy seems to be hard to obtain.

To reduce the influence of the channelling effect, we preamorphized the bulk Si by Ge implantation at 60 keV prior to the B^+ implant. To allow for a direct comparison of the benefits of preamorphization, the B^+ implantation was done at 10 keV as well. Of course, perfect recrystallization of the amorphous layer is of crucial importance. To remove the defects we employed a two step anneal. A first step (450°C, 30 min) is performed to smoothen the amorphous/crystalline interface /2/ and to create the proper conditions for defect-free recrystallization during the final high temperature step. In order to minimize the broadening of the base profile this anneal was done by rapid thermal processing (RTP) at 1075° (10 sec).

Difficulties with channelling and implantation defects are excluded if the base is not implanted but diffused. We used the same polysilicon layer as diffusion source for the active base and the emitter. Since B diffusion into the substrate is strongly suppressed by the presence of As, the drive-in for base and emitter is split. After boron diffusion, As was implanted and diffused by RTP at 1050°C (10 sec) like in the approaches described above. However, as the base is formed later than the oxide spacers, there is an inherent risk of insufficient base link with this scheme (fig. 1b). Therefore, careful tuning of diffusion of active and inactive base as well as oxide spacer thickness is mandatory in order to maintain low base resistance R_B.

3 - RESULTS AND DISCUSSION

The E/B dopant profiles as measured by SIMS are plotted in fig. 2. For the sake of clarity only the results of the 10 keV implanted (scheme A) and the diffused (scheme C) base are shown. Although the same emitter drive-in was used, a smaller emitter junction depth is obtained for the transistor with the diffused base. This is attributed to a retarded diffusion of As due to partial recrystallization of the polysilicon during the boron drive-in step. The resulting base widths W_B of the three approaches are listed in table 1.

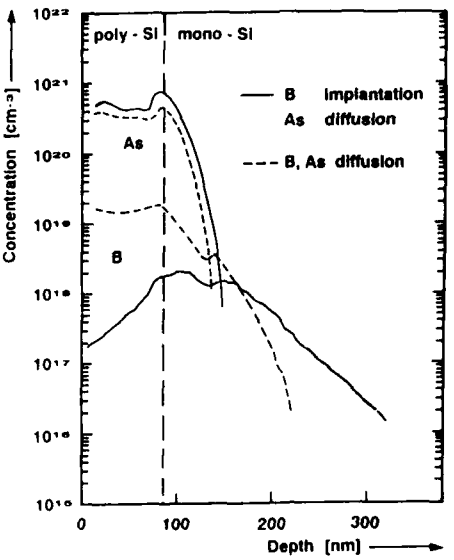


Fig. 2 - Dopant distributions measured by SIMS for transistors with implanted (10 keV) and diffused base (950°C, 10 min), respectively.

Table 1 - Device data for the different process versions

base formation process	W_B/nm	τ_F/ps	R_B/R_B(Process A)
A	150	8	1
B	125	7	1
C1 (950°C, 30 min)	105	5.5	3
C2 (950°C, 10 min)	85	5	7

Obviously, both novel base formation techniques result in an appreciable reduction of W_B. In the case of implanting B^+ into preamorphized Si (scheme B) a final W_B of 125 nm was achieved. The full benefits of preamorphization, however, will only be obtained for B^+ implantation energies below 10 keV. Thence a greater part of the B profile will be placed inside the amorphous layer and channelling will be more completely eliminated. By diffusing the base according to approach C, W_B of 105 nm or 85 nm, respectively, were obtained, depending on the B diffusion step (see table 1, scheme C1 or C2, respectively).

Progress in device performance by the novel base formation techniques becomes evident from fig. 3 where the transit frequency f_T vs. collector current density j_C is plotted as measured with HF-probes on wafer. In the case of process B the f_T-roll-off occurs at lower j_C values. This is due to an increased epi layer thickness (1.3 instead of 0.9 µm in the other cases). A maximum value f_T = 14.5 GHz is obtained, which compares to f_T = 13.0 GHz for the transistor with implanted base according to scheme A.

In order to demonstrate more clearly the influence of W_B we extracted the base forward transit time τ_F using the conventional ($1/2\pi f_T$ vs. $1/j_C$)-extraction scheme. In fig. 4 the τ_F values are compared with one dimensional calculations /3/. The boron concentration at the E/B junction N_A and the base width W_B were determined by SIMS, N_A also was independently confirmed by C(V)-measurement. From the fairly good agreement of experimental data and theoretical predictions, we expect that by further reducing W_B values of τ_F below 4 ps are attainable.

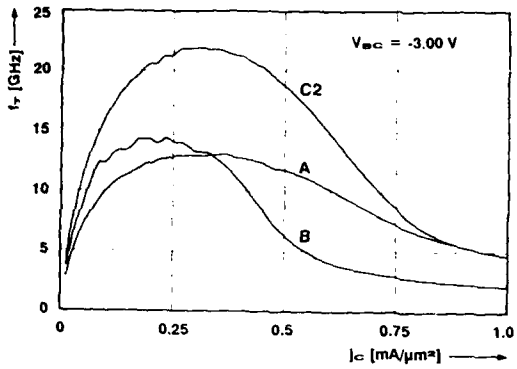


Fig. 3 - Transit frequency f_T for transistors with implanted base either into crystalline Si (curve A) or preamorphized Si (curve B) and diffused base (curve C2)

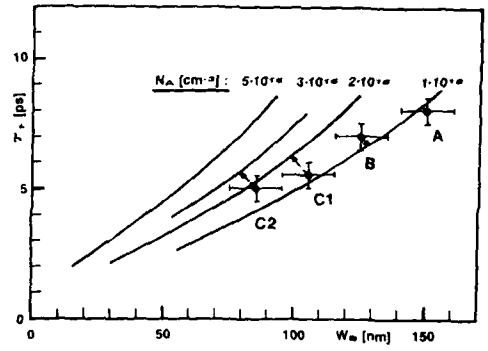


Fig. 4 - Comparison of experimentally determined forward transit times τ_F and base widths W_B with one-dimensional calculations (solid curves). The arrows indicate discrepancies in N_A between values used in the simulations and measurements

To check the link between external and internal base we measured base resistances R_B using a transistor with two isolated base contacts, one for feeding the regular base current and one for sensing the voltage drop across the base. In table 1 values of R_B, normalized to process version A are listed. Both devices with diffused bases suffer from an insufficient base link. In these cases adjustment of diffusion cycles and spacer thickness is necessary for optimum device operation.

4 - CONCLUSIONS

The key to further enhance the intrinsic speed of bipolar devices is the reduction of the base width W_B. This is demonstrated by a comparison of 1d simulations with experimental data. Simple reduction of base implant energy, however, has only limited chance of success because of the growing influence of the channelling effect. Two methods to circumvent this problem, i.e. preamorphization of the substrate prior to the base implant or diffusion of the active base out of the emitter polysilicon, have been demonstrated to open good perspectives for accomplishing very shallow base profiles. In first experiments basewidths down to 85 nm and forward transit times as low as 5 ps have been attained.

We would like to thank B. Hoffmann and D. Hartwig for supervision of processing, R. Kakoschke for RTP and T.F. Meister for valuable discussions.

REFERENCES

/1/ Cho, K., Allen, W.R., Pinstad, T.G., Chu, W.K., Liu, J. and Wortmann, JJ. Nucl. Instrum. Meth. in Phys. Res. B7/8 (1985) 265.

/2/ Myers, E, Rozgonyi, G.A., Sadana, D.K, Maszara, W., Wortmann, J.J. and Narayan, J, Mat. Res. Soc. Symp. Proc. Vol. 52 (1986) 107.

/3/ Schaber, H., Bieger, J., Meister, T.F., Ehinger, K. and Kakoschke, R., IEDM Tech. Digest (1987) 170.

SESSION 2

Session 2IP Room A : Invited papers

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

TRENDS IN INDIUM PHOSPHIDE MICROELECTRONICS

A. SCAVENNEC

Laboratoire de Bagneux, Centre National d'Etudes des Télécommunications, 196, Av. Henri Ravera, F-92220 Bagneux, France

RESUME - Parallèlement à l'optoélectronique 1,3 - 1,5 µm, et grâce à des propriétés de transport électroniques remarquables, une microélectronique sur les matériaux à base de InP se développe actuellement. Après quelques rappels sur les propriétés des matériaux InP et GaInAs, les diverses structures de transistors actuellement étudiées sont décrites, avec leurs mérites respectifs. Une analyse est ensuite présentée, portant sur les évolutions technologiques et les domaines d'application.

ABSTRACT - Together with 1,3 - 1,5 µm optoelectronics, a microelectronic technology is presently developing on InP-based materials, owing to their remarkable transport properties. After general considerations on InP and GaInAs properties, various transistor structures are presented and their comparative merits discussed. This is followed by an analysis of current trends in technology and fields of applications.

1 - INTRODUCTION

In the past 15 years, a large interest has concentrated on gallium arsenide (GaAs) devices for both microelectronic (GaAs MESFETs) and optoelectronic applications (AlGaAs/GaAs lasers and LEDs). This technology is now rather well-established, even if digital applications of GaAs MESFETs are unrestly facing progresses in silicon technology. More recently, another III-V material, InP, exhibiting similar potentialities in the fields of optoelectronics and microwave or high-speed applications, has experienced a renewed interest, mainly because of the shift in interest towards higher wavelengths for optical fiber transmissions. In spite of a higher susceptibility of the material to technological treatments than GaAs, InP-based optoelectronic devices are now fabricated and are commercially available. Conversely, no InP-based transistor has established itself as the plow-horse of an InP-based microelectronic technology, although a large research effort is building up worldwide in the investigation of various structures of field effect or bipolar transistors. This should soon result in reliable high-performance devices taking advantage of the attractive electronic transport properties of InP and of GaInAs, a ternary compound which, lattice-matched to InP, exhibits a mobility of more than 10000 cm^2/Vs.

2 - ELECTRONIC TRANSPORT IN InP AND GaInAs

Like in GaAs, electron mobility in InP and GaInAs is much higher than hole mobility, which favors n-channel FET or npn bipolar structures. With drift mobilities of 4000 and 10000 cm^2/Vs. respectively at room temperature, both materials exhibit a similar high drift velocity, of the order of 2.5 10^7 cm/s for 1 µm long gate devices (Fig1). This value, somewhat higher than in GaAs, mostly results from the larger separation between Γ and L valleys in the conduction band, confining electrons in the high mobility Γ valley till they gain an energy larger than about 0.5 eV from the electric field. Even for transistors with submicron gate length, this superiority of InP and GaInAs over GaAs is expected to remain valid /1/. Also related to the higher $\Delta E_{\Gamma L}$ value is the more prononced negative resistance behaviour used in Transferred Electron (Gunn) Devices (TEDs).

In table I, some parameters of interest regarding electronic properties of InP, GaInAs and AlInAs are given. This third compound is increasingly being used as a high bandgap material to confine electrons in the channel of GaInAs FETs, thanks to the large conduction band discontinuity with GaInAs. Unfortunately good quality AlInAs is difficult to grow by MBE or MO-VPE, with usual mobility in the 1000-2000 cm^2/Vs. range. Recently figures between 3000 and 4000 cm^2/Vs have been reported with both growth techniques; This illustrates the rapid improvement currently experienced in the epitaxial growth of InP-based material. It is also important to emphasize in table I, the small bandgap of GaInAs resulting in its high mobility but also in a low bulk resistivity, high temperature sensitivity and large impact ionization coefficients (fig. 2).

Table 1: Physical parameters of InP and lattice-matched GaInAs and AlInAs.

	InP		GaInAs		AlInAs
E_G (eV)	1.35		0.75		1.45
MOBILITY (cm^2/Vs)	4000		10000		3500
$\Delta E_{\Gamma L}$ (eV)	0.6		0.55		0.61
DRIFT VELOCITY (10^7 cm/s)	2.5		2.5		
BULK RESISTIVITY (Ωcm)	10^8		2.10^3		
BARRIER SCHOTTKY HEIGHT Φ_{Bn}(eV)	0.4		0.2		0.6/0.8
ΔE_C (eV)		0.2		0.55	

3 - KEY ASPECTS OF AN InP-BASED MICROELECTRONIC TECHNOLOGY

While n^+ doped substrates with a very low dislocation density are now available for optoelectronic devices, high quality semi-insulating (S.I.) substrates are not yet available. Because of an insufficient purity of starting materials, the resistivity of undoped crystals is usually only marginal (residual doping : a few 10^{15} cm^{-3} - n type). Semi-insulating material (10^7-10^8 Ωcm) is obtained by compensation of residual donors by Fe atoms with a concentration which increases from seed to queue of the ingot (for instance 10^{16} - 10^{17} cm^{-3}). Like in Cr doped GaAs, Fe atoms tend to redistribute under annealing, resulting in thermal unstability. Moreover a reproducible process for surface preparation (polishing) is not yet perfectly established. Obviously more work is necessary before S.I. substrates of adequate quality becomes available.

Actually most devices are fabricated from epitaxial layers, generally grown on a buffer layer preventing detrimental pollution from the substrate. VPE but more often MO-VPE or MBE are presently used for the growth of these structures. A good homogeneity of the grown layers (composition thickness, doping level) is then necessary not only to maintain the lattice matching condition in the case of heterostructure - which is not critical in the AlGaAs-GaAs system- but mainly to provide a small dispersion in threshold voltage. Very encouraging results have already been obtained in that respect /2/ : a deviation of only ± 1.5 % over 2 inch substrates of the sheet carrier concentration.

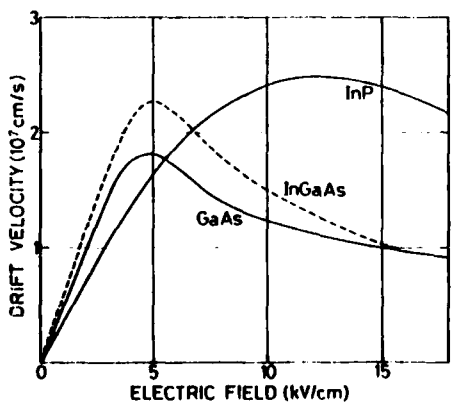


Figure 1:Drift velocity of InP and GaInAs under static electric field.

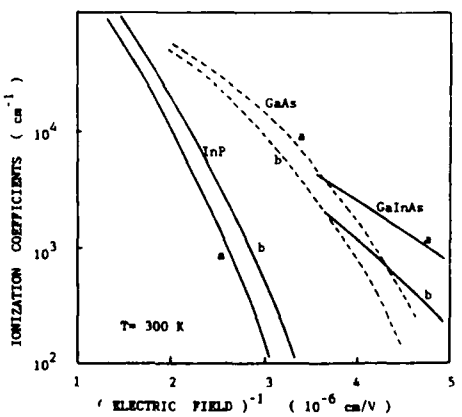


Figure 2: Ionization coefficients for holes(b) and electrons(a) in InP and GaInAs with reference to GaAs

Another important point which illustrates the difficulty faced in establishing an InP microelectronic technology is the material fragility : not only InP is more brittle than GaAs, making the handling of 2 or 3 inch wafers more critical, but also InP, and GaInAs as well, are quite sensitive to bombardment by energetic species as found in usual plasma etching or deposition systems. For instance Ar ion milling of InP surfaces even at low energy (300 V) degrades the material quality and results in the formation of an In rich surface layer unless special care such as cooling the slice at liquid nitrogen temperature is taken /3/. InP, and to a lesser extent GaInAs, also tends to decompose at relatively low temperature with preferential evaporation of the group V element (around 350°C under ultrahigh vacuum for InP /4/). Although this phenomenum can be fought by encapsulation, or phosphorus or arsenic overpressure, it obviously requires adjustements in the processing techniques previously developed for GaAs devices and circuits. Satisfactory techniques for surface protection are now available for diffusion and implantation annealing.

Related to its surface unstability, passivation is a very important issue for the reliability of InP-based devices : it is well-known that reverse biased GaInAs p-n junctions, when exposed to atmosphere, tend to degrade with conductive paths appearing at the surface, possibly associated with the formation of an In_2O_3 conductive layer. This problem has so far received limited answers. Only the development of suitable passivation schemes (PECVD SiN for instance in the case of GaInAs PIN or avalanche photodiodes /5/) allows the developemnt of stable devices. This passivation issue, which also hampered for quite a while the development of a viable GaAs MESFET technology, is now well addressed, partly due to the large interest in establishing a MISFET technology on InP and GaInAs, motivated by the attractive low surface state density observed on these materials.

To conclude this paragraph, a few considerations on metal-semiconductor contacts are now given : with most metals, InP and GaInAs surface tend to be pinned quite close to the conduction band. This has so far hampered the development of a MESFET technology on these materials. It has also partly motivated the use of AlInAs in various FET structure, owing to a larger Schottky barrier height, between 0.6 and 0.8 eV (see Table I). This Fermi level pinning at the surface is also put forward to explain the low and high resistivity of contact on n-type and p-type materials respectively, with some recent improvement in the latter case /6/,/7/. A technology for InP-based microelectronics is still far from being fully established ; but important progress is continuously being made. Such progress is best illustrated by the increase in performance of the various transistor structures presently investigated.

4 - InP FETs

First attempts to develop an InP MESFET technology have met limited success, due to the low Schottky barrier height. Since then, different schemes have been proposed to overcome this situation : for instance by adding a thin dielectric layer between the semiconductor surface and the gate metallization /8/ or more recently by use of a shallow diffusion converting a thin surface layer - completely depleted of carriers in actual operation - to p-type /9/. It has also been shown that by a suitable choice of the gate metal (Cd), a larger Schottky barrier appears /10/. None of these approaches has already established itself as a viable process so far, even though such a technology has been used in the fabrication of a monolithic laser-FET device /11/.

InP junction FETs have also been investigated, in particular with fully ion-implanted structures /12/. A major achievement has been the control of a shallow, abrupt p^+n gate junction, thanks to a co-implantation process. Submicron gate planar structures are difficult to fabricate, with the requirement of aligning the metal gate onto the p^+n region. This can be overcome by subsequent etching with the gate metal as a mask. More investigations are still needed to ascertain the viability of this structure (gate contact resistivity, mesa gate passivation) but promising results have already been obtained /13/. It is worth illustrating the potential of InP JFETs by noting that the monolithic receiver with the best sensitivity at 1.3 µm (-36.5 dBm at 200 Mbit/s) has been fabricated with InP JFETs /14/.

However most of the research devoted to InP FETs has so far focused on MISFETs : it was recognized quite early that the interface state density of insulator/InP could be very low, below $10^{12}/cm^2$ eV, in contrast to the situation with GaAs. Moreover in the prospect of a power transistor, a MISFET structure appears quite attractive. This was actually clearly demonstrated by the excellent results reported on a SiO_2 gate power MISFET : 3 W/mm /15/.

Unfortunately InP MISFETs are prone to large threshold voltage drift, in operation, due to unstability of the insulator-semiconductor interface. InP native oxydes, as usually grown exhibit poor dielectric properties. Good characteristics were reported however, with the oxidation of InP through an Al_2O_3 film /16/. This oxidation scheme leads to the formation

of a polyphosphate insulator, as recently identified. The direct formation of such an insulator by other means is obviously a promising technique /17/. As an alternative to oxidation, both nitridation and sulfidation have been investigated. In spite of an excellent effective mobility (3500 cm^2/Vs) characteristic of a low density of fast interface states, the dielectric quality of the thermally grown sulfide insulator doesn't allow the fabrication of transistors with low current drift /18/. Actually it is a second approach to the realization of MISFETs that has received most of the attention : that is the deposition of a dielectric film (SiO_2, SiN, Al_2O_3, PN, ...) using a large number of techniques, from e-beam evaporation to CVD. Most of them lead to a low density of interface states, and then acceptable mobility. However, drain current drift in CW operation, due to trapping and evidenced by hysteresis of C-V MIS diodes characteristics, remains a major obstacle to the development of an InP MISFET technology /19/.

Several criteria to be fulfilled in order to obtain an acceptable MIS structure are now well accepted : a good stoechiometry of the very first atomic layers of the InP surface, implying no high-temperature treatment or surface bombardment creating P vacancies ; a very thin interface layer, requiring in situ etching /20/, possibly As passivation /21/ ; a high quality dielectric. Simultaneous obtention of those characteristics is a real challenge : conventional CVD, indirect PECVD /22/, UV enhanced CVD /23/, flash CVD /24/, have in particular shown much promise in that respect, as evidenced by the reasonable drift behaviour reported with such techniques /25/. Obviously real progress has been made in recent years in the understanding and control of InP MISFETs drift behaviour. Although this problem has not yet received a fully satisfactory answer, continuous improvement is observed (fig.3), opening the door to a MISFET technology, which would be a major breakthrough in III-V microelectronics.

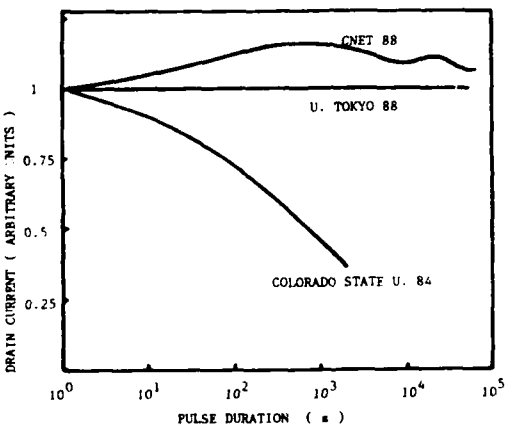


Figure 3: Drift behaviour of the drain current of enhancement-mode InP MISFETs after a 2V step is applied on the gate. V_{ds}= 0.1V.

5 - GaInAs FETs

$Ga_{0.47}In_{0.53}$ As lattice-matched to InP is a very promising material, with a high mobility and high saturation velocity. Moreover with a low-bangap, it is particularly suited for the absorption of light at wavelength of 1.3 and 1.55 µm. For these reasons, GaInAs transistors have received much attention in recent years. But which FET structure is the best is not clear yet, with continuous improvements being reported for the various structures (fig. 4). This progress can be clearly associated to the better quality of material grown by MBE or MOVPE including more efficient buffer layers at the substrate interface. As with InP, GaInAs MESFETs have not received much attention, due to the low Schottky barrier height. A few attempts have been reported recently aiming at artificially increasing the barrier height by inserting a thin insulating layer between the metal gate and the GaInAs channel or by the use of a Cd gate electrode .

The most popular GaInAs structure has long been the homojunction FET (JFET). Many different approaches have been reported for defining the gate : blanket p diffusion or implantation followed by gate layer etching, p-n junction epitaxial growth, selective diffusion or implantation,...with compromise to be found between gate length reduction and gate junction passivation. While selectively diffused or ion-implanted gate may be naturally passivated by the diffusion or implantation mask /26/ /27/, no efficient scheme has yet been reported for mesa etched gate. Better performance can be obtained with

heterojunction FETs (HJFETs) : selective etching of a p^+ InP gate layer down to a gate length of 0.5 µm has already been reported with a f_{max} of 60 GHz /28/, but the critical passivation problem has still to be solved.

GaInAs MIS structures, like InP ones, are characterized by a low interface state density (> 10^{12} cm^{-2} eV^{-1}) allowing the fabrication of MISFETs. Moreover, and this is probably attributable to the better surface stability of GaInAs, several Laboratories have already reported a quite acceptable drain current drift behaviour /29/. For instance with a self aligned structure with a SiO_2 UVCVD gate dielectric, a drain current drift of only 2 % in 70 hours has been mentionned /30/, allowing the fabrication of test circuits and microwave transistors with a 45 GHz f_{max}. Similar results (3 % in 30 hours) have since been reported on depletion mode GaInAs transistors with SiN gate dielectric deposited in a multipolar plasma after in situ native oxide removal /31/. The MISFET technology, only scarcely addressed till now, obviously offers much promise, in particular for logic applications and optoelectronics.

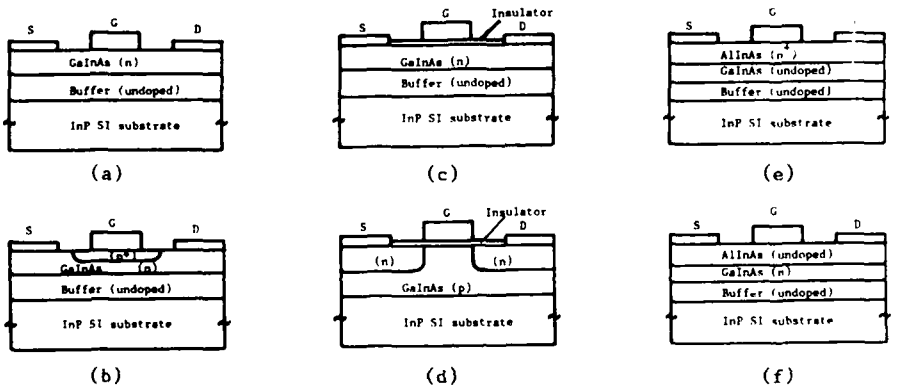


Figure 4: Schematic GaInAs FET structures
(a): MESFET (c): Depletion-mode MISFET (e): MODFET
(b): JFET (d): Inversion-mode MISFET (f): "MIS"FET

In the last 10 years AlGaAs/GaAs heterostructure FETs have been widely studied as an alternative to conventional GaAs MESFETs. Known in their original design as TEGFETs, MODFETs or HEMTs, they showed superior performances in terms of transconductance, cut-off frequency and noise figure. A parallel development has been experienced very recently with AlInAs/GaInAs heterostructure FETs, which offer a major advantage over their GaAs-based counterparts : the conduction band discontinuity, close to 0.55 eV (0.25 eV for usual AlGaAs/GaAs MODFETs) allows a much higher electron sheet concentration (about 3.10^{12} cm^{-3}) while preventing a parasitic channel to appear in the AlInAs barrier over a wider range of bias condition ; only drawbacks : a slightly lower Schottky barrier height (0.6-0.8 eV depending on AlInAs quality) and a material which has not yet reach the quality of AlGaAs. Several Laboratories have developed such heterostructure FETs which already outperform their GaAs counterparts of comparable gate length /32/. Since the report by NEC of a transistor with a transconductance of 440 mS/mm at room temperature, progress has been very fast, in particular with respect to high frequency performances (fig. 5). Recently devices with 1080 mS/mm transconductance and 175 GHz transition frequency for 0.1 µm gate length, with a minimum noise figure of 0.8 dB at 63 GHz (AG : 8 dB) for 0.2 µm gate length were reported /33/.

While the MODFET structure uses a doped barrier, which can act as a parasitic channel and be responsible for a high gate leakage current, other designs have been studied, directly derived from the AlGaAs/GaAs experience, with an undoped barrier, such as the inverted MODFET and the "MISFET". Such structures usually exhibit lower gate current, lower transconductance and gate capacitance than conventionnal MODFETs. Still their high frequency performance remains very attractive, as illustrated by the 100 GHz f_{max} reported by Bendix for a 0.7 µm long gate structure with a 200 Å thick channel doped to 4.10^{18} cm^{-3} /34/. With a more conventionnal channel (1000 Å with 2.10^{17} cm^{-3} doping) a 40 GHz f_{max} was obtained at CNET-Bagneux on 1 µm gate length devices /35/. Moreover a very good uniformity is expected for undoped barrier FETs with a conduction band discontinuity close to the Schottky barrier height ; this is an important point for circuit

applications. Worth mentioning in the prospect of circuits are the 280 mS/mm transconductance and near-zero threshold voltage reported by Sony on self-aligned refractory gate devices with 5 µm gate length /36/.

While improvement in AlInAs/GaInAs heterostructure FETs growth and processing is continuing, first promising results were reported recently with identical structures but for an InP channel : on 1 µm gate devices with a doped channel a transconductance of 155 mS/mm was measured and an f_{max} of 42 GHz inferred from S parameters measurements /37/. With pseudomorphic AlGaAs/GaInAs, heterostructure FETs as an example,pseudomorphic AlInAs/GaInAs, AlInAs/InP, AlGaAs/InP, GaInP/GaInAs have also been investigated; No clear improvement in performance as compared to lattice-matched devices has been demonstrated so far.

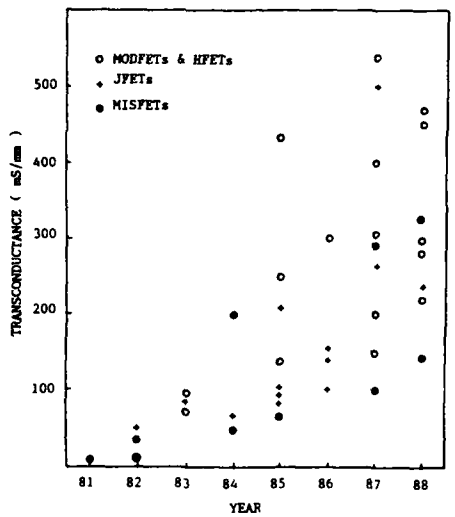


Figure 5: Increase in transconductance over the years illustrating the improvement in material quality. Only FETs with gate length >0.5 micron have been considered.

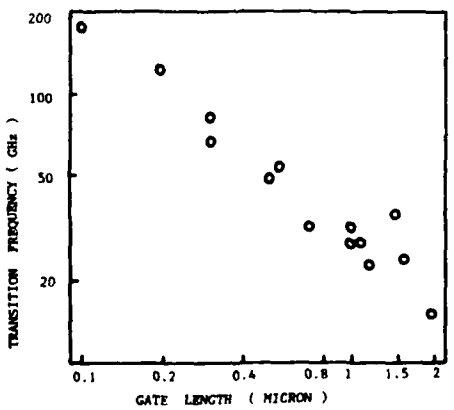


Figure 6: Variation of the transition frequency of GaInAs FETs with different structures versus gate length.

6 - GaInAs HETEROJUNCTION BIPOLAR TRANSISTORS

Bipolar transistors can also be fabricated from InP-based materials : as in Si or GaAs, they offer potential improvement over FETs in terms of drive capability, threshold voltage homogeneity, at the expense of a more complex processing technology. While the AlGaAs/GaAs HBT technology in now developing rapidly, the same HBT technology can also be envisionned on InP-based materials, with the associated advantages : low base sheet resistance, low emitter capacitance. With different material combinations available for the fabrication of HBTs, most investigations have rapidly concentrated on npn structures with a GaInAs base, and either GaInAsP or AlGaInAs emitter. These bipolar transistors have very interesting potentialities ; actually recent developments in the AlGaAs/GaAs HBT technology have shown that the device speed is now approaching a limit set by the transit time of electrons through the base and collector. The superior transport properties of InP-based materials could then allow an extension of the frequency range of operation of HBTs, with the additionnal advantage for circuit applications of a lower emitter-base voltage. Also of interest are the better behaviour of InP/GaInAs HBTs at low current density, resulting from the lower surface recombination velocity. This, as well as the invertibility of wide-bandgap emitter and collector, may be especially interesting in I^2L circuits.

However, the HBT technology is far more complex than the FET's one, and many technological problems have to be solved before a viable technology is established : abrupt p-n junctions, graded heterojunctions, isolation, passivation, non-diffusing contacts, ... Very promising results have already been obtained, in particular InP/GaInAs DHBTs grown by gaz source MBE with a transition frequency of 18 GHz for 5 µm emitter stripe width /38/, and devices grown by MOVPE with a current gain of 11000 /39/. In the AlGaInAs/GaInAs system, a current gain of 1200 and a transition frequency of 1.7 GHz for devices with 20 µm emitter stripe width have been measured /40/. Undoubtedly, these characteristics will soon be improved, taking into account the general improvement of the technology of InP-based materials.

7 - FROM DEVICES TO CIRCUITS

An important point has to be made : no InP-based microrelectronic device or circuit, with the exception of microwave Gunn diodes, has yet been developed, sold or used in systems, although a few companies intend to do so shortly. For the time being, most efforts are devoted to the demonstration of the feasibility of transistors or circuits. Low complexity circuits have been fabricated for demonstration of either logic or optoelectronic functions. Between 1980 and 1984 several InP MISFET ICs have been reported in the United States under Navy funding /41/. Such circuits were not able to operate DC, because of the drift in transistor characteristics. More recently GaInAs FET ICs have also been reported (GaInAs MISFET ring oscillator with a propagation delay time of 50 ps/gate /42/, GaInAs JFET base band amplifier). Most attention is presently focusing on using AlInAs/GaInAs heterostructure FET for digital ICs. In spite of a report by NEC on the dispersion of the sheet carrier concentration over a 2 inch wafer, almost no data is presently available on the threshold voltage dispersion of processed transistors. The quality of AlInAs material for instance is probably not yet sufficient to allow a reproducible threshold from wafer to wafer. A larger number of reports is now available on the use of InP-based transistors in optoelectronic circuits. Since the first demonstration in 1979 of an IC associating one GaInAs PIN and one FET /43/ a large interest has developed in the area of InP-based OEICs, particularly in Japan. A variety of InP and GaInAs FETs, and even bipolars, have been used in the fabrication of laser drivers or photodetector amplifiers. This large interest originates mainly from the expected improvement in performances, resulting from the reduction of interconnection reactances, as well as in cost and compacity. However the fabrication of high performance OEICs has so far been hampered by the necessary compromising for integrable structures of optoelectronic and microelectronic devices. Very significant achievements have been reported however, on either simple receiver circuits associating a photodiode and a transistor /44/ or even more complex circuits as recently reported by ATT Bell (GaInAs PIN plus 8 InP JFET amplifier /14/). Very informative is the analysis of the current trend in the increase in performance of monolithic photoreceivers : the sensitivity is steadily improving by about 2.5 dB/year since a few years and within 2 years from now, monolithic receivers should surpass their hybrid counterparts. Quite interestingly, a parallel evolution is presently observed with GaAs OEICs, with the recent report of a GaAs receiver exhibiting a sensitivity of -39 dBm at 250 Mbit/s, equivalent to the one of non-avalanching diode hybrid receivers at 0.85 µm wavelength /45/.

A similar evolution is observed with monolithic transmitters for which increasing bandwidths are being reported /11/. The variety of transistors, with InP or GaInAs channel used in these demonstration circuits still illustrates clearly the current uncertainty as to which transistor structures are the more promising.

8 - PROSPECTS FOR AN InP-BASED MICROELECTRONICS

The InP-based microelectronics is still in its infancy, but it is rapidly maturing as evidenced by the improvement in basic characteristics such as transconductance and bandwidth. Most recent reports refer to GaInAs channel FETs which benefit from simultaneous developments of high quality epitaxial growth techniques required for optoelectronic devices and of a mature AlGaAs/GaAs Schottky gate technology. Actually, for many years, and in spite of the known promising characteristics of InP, the motivation for developing an InP-based microelectronics was not strong enough to allow a sensible development. Only with the advent of 1.3 - 1.5 optoelectronics did InP-based microelectronics experience a real development. Also very important is the fact that the GaAs microelectronics expertise can quite directly be used for InP compounds /46/.

Thus supported by optoelectronics and GaAs technology, InP-based microelectronics is still facing specific problems : In the area of technology, several points still require particular attention, such as high quality (undoped ?) semi-insulating substrates, low temperature, low energy processing techniques, passivation ; but most importantly it will be necessary to concentrate future development on a limited number of transistor structures. In that respect more work has still to be done to ascertain the viability of structures with a GaInAs channel under operating conditions of real circuits; Beyond that point, the choice of the best structure is still an open question. For InP and GaInAs MISFETs, experiencing a real improvement of the stability of their threshold voltage, the prospects now look better, but much time is still necessary before a reproducible, industrial processing technology is established. Presently JFETs are quite appealing because of their simplicity. Good results have already been reported, and small complexity circuits have been fabricated. Heterostructure gates can possibly improve JFETs performances ; however heterostructure FETs of the MISFET type, or MODFETs, are more likely to become the prefered transistors either with InP or GaInAs channels. HBTs are also very attractive, but since they require an involved processing technology, it will take much time before they begin to challenge their AlGaAs/GaAs counterparts.

Optoelectronic integration is obviously a promising field of application for InP-based microelectronics. Fabrication of OEICs with acceptable yield will require adjustements in transistors structures and processing to ease integration with optoelectronic devices /47/. Competition might come from mis-matched epitaxy approaches, illustrated by the realization of GaAs/InP transmitters and receivers by NEC /48/. In the field of microwave applications, two areas of interest have to be considered : for low-noise amplication, AlInAs/GaInAs MODFETs already begin to challenge their AlGaAs/GaAs counterparts and will continue to do so, with industrial development likely to appear soon. For power microwave amplification, InP has already shown its potentialities. Unfortunately InP MISFETs used for such demonstration suffered from threshold voltage drift. While some progress is now observed on this point, other structures (JFETs and even MESFETs) are also investigated as alternatives, with the same potential for high power although probably with inferior performance in linearity. Finally one should not forget the excellent negative resistance properties of InP and GaInAs and of AlInAs/GaInAs barriers which may in the future allows new development in microwaves or digital applications /49/, /50/.

9 - CONCLUSION

Transistors made from InP compounds have already shown very attactive performances, mainly for microwave and optoelectronics applications. Still some basic work has to be carried on to establish suitable technologies and even to define device structures. Using the same tools as developed for GaAs ICs and for 1.3 - 1.5 µm wavelength optoelectronic devices, InP-based transistors are now able to enter the world of applications in both microwaves and optoelectronics. A few laboratories, in particular in Japan, are even already addressing the question of digital applications of GaInAs FETs. In Europe, several Laboratories are also engaged in research programmes, partly under EEC funding, aiming at establishing a viable InP-based microelectronics technology.

ACKNOWLEDGMENTS : The author thanks his colleagues of CNET-Laboratoire de Bagneux involved in InP-based microelectronics for fruitful discussions on the topic of this paper.

References :

/1/ A. Cappy, B. Carnez, R. Fauquembergues, G. Salmer, E. Constant : IEEE Trans. Electron. Dev. 27 (1980) 2158
/2/ H. Toyoshima, A. Okamoto, Y.Nakamura, K. Okata : Proc GaAs and Related Compounds Heraklion (1987)
/3/ N. Bouadma, J.F. Hogrel, J. Charil, M. Carré, IEEE J. Quantum Electron. 23 (1987) 909.
/4/ R.F.C. Farrow : J. Phys D. Appl. Phys 7 (1974) 2436.
/5/ S.R. Forrest, V.S. Ban, G. Gasparian, D. Gay, G.H. Olsen : IEEE Electron. Dev. Lett. 9 (1988) 217.
/6/ R. Kaumanns, N. Grote, H.G. Bach, F. Fidorra : Proc. GaAs and Related Compounds, Heraklion (1987).
/7/ P.E. Hallali, P. Blanconnier, L. Bricard, J.C. Renaud : Proc. ESSDERC Montpellier (1988) this issue.
/8/ S. Loualiche, A. Ginudi, H. L'Haridon, M. Salvi, A. Le Corre, D. Lecrosnier, P.N. Favennec : Proc. ESSDERC Montpellier (1988) this issue.
/9/ G.J. Gualtieri, G.P. Schwartz, G.J. Zydzik, L.G. Van Uitert : J. of Electrochem. Soc. 133 (1986) 1426.
/10/ L.G. Meiners, A.R. Clawson, R. Nguyen : Appl. Phys. Lett. 49 (1986) 340.
/11/ N. Suzuki, H. Furuyama, Y. Hirayama, M. Morinaga, K. Eguchi, M. Kushibe, M. Funamizu, M. Nakamura : Electron. lett. 24 (1988) 467.
/12/ S.J. Kim, K.W. Wang, G.P. Vella-Coleiro, J.W. Lutze, Y. Ota, G. Guth : IEEE Electron. Dev. lett. 8 (1987) 518.
/13/ W. Kruppa, J.B. Boos : IEEE Electron. Dev. lett. 8 (1987) 223.
/14/ S.J. Kim, G. Guth, G.P. Vella-Coleiro, C.W. Seabury : Proc. Dev. Conf. Boulder (1988).
/15/ M. Armand, D.V. Bui, J. Chevrier, N.T. Linh : Electron. lett. 19 (1983) 433.
/16/ T. Sawada, H. Hasegawa : Electron. lett. 18 (1982) 742.
/17/ E. Jolaguier, J. Joseph, Y. Robach, E. Bergignat, G. Hollinger : Rev. Phys. Appl. 22 (1987) 885.
/18/ G. Post, P. Dimitriou, A. Scavennec, N. Duhamel, A. Mircea : Electron. lett. 19 (1983) 459.
/19/ S.Krawczyk : l'Onde Electrique 67 (1987) 2/49.
/20/ N. Yoshida, K. Mizuguchi, T. Murotani, K. Fujikawa : Proc. 7th Dry Process Symposium (1985) 19.
/21/ R. Blanchet, P. Viktorovitch, J. Chave, C. Santinelli : Appl. Phys. lett. 46 (1985) 761.
/22/ L.G. Meiners : J. Vac. Sci. Technol. 21 (1982) 102.
/23/ G. Post, P. Dimitriou, A. Falcou, N. Duhamel, G. Mermant : Proc. ESSDERC Montpellier (1988) this issue.
/24/ Y.I. Nissim, C. Licoppe : Proc. ESSDERC Montpellier (1988) this issue.
/35/ L. Messick, D.A. Collins, R. Nguyen, A.R. Clawson, G.E. Williams : Proc. IEDM Los Angeles (1986) 767.
/26/ L. Nguyen, M. Allovon, P. Blanconnier, E. Caquot, A. Scavennec, B. Bourdon : Proc. ESSDERC Bologna (1986) 943.
/27/ H. Albrecht, Ch. Lauterbach : IEEE. Electron. Dev. lett. 8 (1987) 353.
/28/ J.Y. Raulin, E. Vassilakis, M.A. di Forte-Poison, M. Razeghi, C. Brylinski, G. Colomer : Proc. GaAs and Related Compounds Heraklion (1987).
/29/ J. Selders; H. Benehing : IEEE Electron. Dev. lett. 117 (1984) 173.
/30/ P.D. Gardner; S. Yegna Narayan, Y.H. Yun : Thin Solid Films 117 (1984) 173.
/31/ M. Renaud, P. Boher, J. Schneider, J. Barrier, M. Heyen, D. Schmitz : Electron. lett. (to be published).
/32/ K. Hikosaha, S. Susa, N. Harada, S. Kuroda : IEEE Electron. Dev. lett. 9 (1988) 241.
/33/ U.K. Mishra, A.S. Brown, S.E. Rosenbaum, M.J. Delaney, S. Vaughn, K. White : Proc. Dev. Res. Conf. Boulder (1988).
/34/ M.A. Fathimulla, H. Hier, J. Abrahams : Electron. Dev. lett 24 (1988) 93.
/35/ L. Giraudet, L. Nguyen, M. Allovon, M. Laporte, J.C. Renaud, A. Scavennec : Proc. 12e Workshop ou Compounds Semiconductor Dev. and ICS, Lugano (1988)
/36/ M. Kamada, H. Iskikawa, K. Kanedo, N. Watanabe : Electron. lett 24 (1988) 271.
/37/ M.A. Fathimulla, T. Loughran, L. Stecker, E. Hempfling, M. Mattingly, O. Aina : IEEE Electron. Dev. lett 9 (1988) 223.
/38/ R.N. Nottemburg, J.C. Biochoff, J.H. Abeles, M.B. Parish, H. Temkin : Proc. IEDM Los Angeles (1986).
/40/ J.L. Pelouard, J.P. Praseuth, P.E. Hallali, G. Tremblay : unpublished.
/41/ L.M. Messick : IEEE Trans. Electron. Dev. 31 (1984) 763.
/42/ L.C. Upadhyayula, P.D. Gardner, S.G. Liu, S. Yegna Naroyan : IEEE Electron. Dev. 7 (1986) 390.
/43/ R.F. Leheny, R.E. Nahory, M.A. Pollack, A.A. Ballman, E.D. Beebe, J.C. Dewinter, R.J. Martin : Electron. lett. 16 (1980) 353.
/44/ J.C. Renaud, L. Nguyen, M. Allovon, P. Blanconnier, S. Vuye, A. Scavennec: Journ. of Lightwave Techn. (to be published).
/45/ R.J.S. Bates, D.L. Rogers IOOC 1988.
/46/ M. Bon : l'Onde Electrique 67 (1987) 6/49.
/47/ J.P. Noblanc, J.C. Bouley, A. Scavennec : Optoelectronics-D and T-2 (1987) 177
/48/ A. Suzuki, T. Itoh, T. Terakado, K. Kasahara, K. Asano, Y. Inomoto, H. Ishihara, T. Torikai, S. Fujita : Electron. lett. 23 (1987) 954.
/49/ T. Inata, S. Muto, Y. Nakata, T. Fujii, H. Ohnishi, S. Hiyamizu : Jap. J. Appl. Phys. 25 (1986) 1983.
/50/ S.C. Binari, R.E. Neydert, G. Kelner, J.B. Boos : RCA review 45 (1984) 579.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988 C4-125

HETEROSTRUCTURE FIELD EFFECT TRANSISTOR, PHYSICAL ANALYSIS AND NEW STRUCTURES

G. SALMER

Centre Hyperfréquences et Semiconducteurs, CNRS-UA 287, Université des Sciences et Techniques de Lille-Flandres-Artois, F-59655 Villeneuve-d'Ascq Cedex, France

Résumé : Les principaux phénomènes physiques qui se produisent dans les MODFET conventionnels AlGaAs/GaAs sont décrits brièvement : structures quantiques, propriétés de transport - influence du champ électrique. La modélisation des composants permet d'obtenir dés informations très utiles sur leur fonctionnement et les principaux effets physiques qui interviennent tout en permettant d'en prévoir les performances. Les problèmes posés par les réalisations technologiques ou le fonctionnement dans des conditions particulières (haute température, grande puissance) sont examinés. On décrit les nouvelles structures telles que les MODFET inversés, les SISFET et les MODFET à couches pseudomorphiques ainsi que leurs performances potentielles.

Abstract : Physical phenomena that occur in conventional AlGaAs-GaAs MODFETs are briefly described. This covers quantum effects, carrier transport properties, influence of electric field... Progress in device modeling allow to obtain valuable information on device behaviour as well as various intervening physical effects and to predict the performance. Limitations resulting from both technological imperfections and specific device operations, e.g. low temperature, large power are investigated. New structures such as inverted MODFET, SISFET, and pseudomorphic MODFETs are described and their potential capabilities are discussed.

1. INTRODUCTION

The Modulation doped field effect transistor (MODFET), also called High Electron Mobility Transistor (HEMT), is one of the most promising devices for both microwave and high speed digital applications. It takes advantage of the enhanced mobility and velocity of electrons in the two dimensional electron gas formed at selectively doped GaAs-$Al_xGa_{1-x}As$ heterojunctions. For instance, MODFET ring oscillators have demonstrated switching delay times of 10.2 pS at 1.03mW/stage and 5.8 pS at 1.7m W/stage at 300 K and 77 K respectively using 0.35 µm gate length devices /1/. On the other hand, a 1.8 dB noise figure has been mesured at 60 GHz in a 0.25µm gate MODFET /2/. At 94 GHz, 7 dB small signal gain was also measured from a single stage amplifier. MODFET constitutesalso a good candidate for power amplification in the millimeter wave range as has been shown by several authors /3/ using multichannel structures. Several new device structures have been proposed and realized allowing a significant improvement of performance. Among these we should mention :

- the self aligned accumulation mode GaAs-MIS-like FET (SISFET) (4)
- the inverted HEMT /5/
- the AlAs/nGaAs superlattice HEMT for low temperature operations /6/
- the pseudomorphic GaAlAs MODFET /7/
- the multichannel HEMT /3/
- the n^+ AlInAs/GaInAs/InP MODFET /8/

In spite of the large number of theoretical and experimental studies that have allowed to progress in the understanding of the basic physical phenomena governing the device behaviour, several important aspects are not well known. First we will summarize the present situation of conventional GaAlAs/GaAs MODFETs. Then, we will describe some specific problems of device technology and behaviour under low temperature conditions and large signal operation. Finally, we will give a rapid survey of the various new HEMT structures and try to compare their capabilities.

2. PHYSICAL PHENOMENON IN CONVENTIONAL GaAlAs/GaAs MODFET

Conventional GaAlAs/GaAs MODFETs (Fig. 1) are multilayer structures in which the smaller band gap semiconductor layer (GaAs) is undoped and the larger band gap (AlGaAs) is n^+ doped. As the electron affinity of the former is higher than that of the latter, electrons are diffused towards GaAs, leaving behind their parent donor impurities. Electrons are therefore spatially separated from the ionized donors and can have impressive mobilities (1-2 000 000 cm^2/Vsec at 2°K). Fig. 1 represents a practical device in which an n^+ GaAs cap layer is used in order to

improve the access and contact resistances and a spacer layer in order to reduce the ionized-impurity scattering.

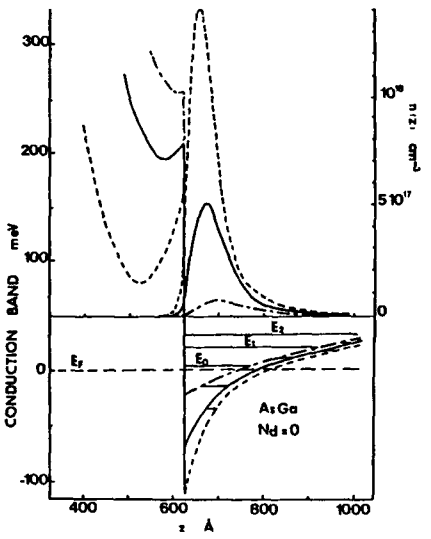


Fig.2 : Evolution of the conduction band minimum, subband structure and carriers concentration.
- - -Vgs =-1.6V ——Vgs = -2.3V
—·—Vgs = -2.7V

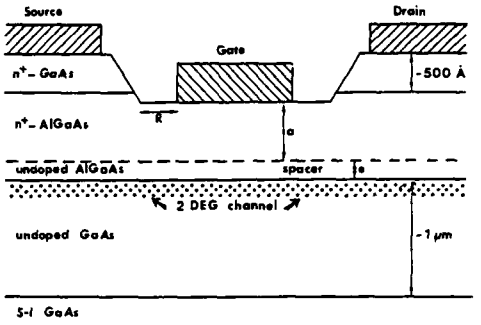


Fig.1 : Typical AlGaAs/GaAs MODFET

In figure 2, the spatial evolution of the conduction band minimum is shown for three values of the gate voltage. These results have been obtained by using a self consistent solution of the Schrödinger and Poisson equations /9/. It can be observed that as Vgs increases, the carrier concentration increases not only in the quantum well but also in the AlGaAs layer due to quantum effects.This occurs, of course, when the width of the potential well is narrow enough for quantum effects to occur along the direction (oz) perpendicular to the plane of the heterojunction. As the electrons keep two degrees of freedom, they form a two dimensional electron gas (2DEG) with a potential energy given by :

$$E_n(k//) = E_n + \frac{h^2}{2m^*}(k_x^2 + k_y^2)$$ where n is an integer.

The subband structure, that clearly appears on this figure is strongly dependent on the form of the quantum well and then on the gate voltage.

Figure 3 represents the relative positions of the subband energies /11/ with respect to the Fermi level E_{FP} as functions of the total carrier concentration in the well and then of gate voltage for a typical device at room temperature ($N_D = 10^{18}$ at/cm^3, a = 625Å, e = 30Å).

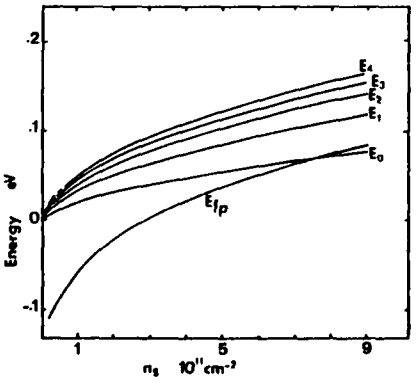


Fig.3 : Subband energy levels as functions of N_S

Looking for the energy levels of the subbands, we can easily conclude that we have to account for a minimum of five subbands if we want to give a precise description of the device. At equilibrium, about 98% of the carriers occupy the five lower subbands of which 90% on the lowest three. In the same way, as it has been shown by several authors, for instance YOSHIDA /12/, the calculation of the electron distribution must be based on Fermi Dirac statistics and using classical approximation can cause important errors.

It is well known that the total carrier concentration is an increasing function of the doping concentration in the AlGaAs layer and the conduction band offset ΔE_c and then of the Al mole fraction x.

Due to partial quantization, the 2D charge transport properties ought to be different from bulk free-charge transport. But due to the lack of information on this point, calculation of the matrix elements of the various scattering mechanisms (acoustic and polar optical phonons, intervalley transfert, impurities, etc...) is carried out by considering a wave function of the form .

$\psi(x, y, z) = \varphi_n(z) \exp i(k_x x + k_y Y)$, instead of the classical plane wave function, but however using a three dimensional potential (as in bulk). Of course, new scattering mechanisms such as intersubband scattering may occur. The main difficulty comes from the fact that the collision probabilities are strongly dependent on the wave function that depend on the gate potential and then vary along the channel.

Under low field conditions, electrons remain in the quantum well and their mobility is the most significant parameter. As the carriers are separated from the ionized impurities, the corresponding scattering mechanisms are greatly reduced and mobilities increase quickly when the temperature decreases. Values of 8.000 cm^2/Vsec, 200.000 cm^2/Vsec and 2.000,000 cm^2/Vsec have been measured respectively at 300°k, 77°k and 2°k /13/. The mobility improvement due to the existence of a spacer layer remains smaller than 15% at room temperature for a spacer layer thickness of 40A° /14/. As the carrier concentration strongly decreases /15/ when the spacer thickness increases, this last parameter must remain smaller than 25A° for devices used at room temperature. Of course, the situation may be quite different at low temperature : for instance, the mobility improvement may reach a factor of 20 at 4°k for a spacer thickness of 500A° /16/.

LEE and al /14/ have experimentally shown that the mobility of the two dimensional electron gas is nearly independent of the carrier concentration n_s, at room temperature. On the contrary, Monte Carlo simulations /9/ indicate that at 77°k, the carrier mobility increase by a factor of two when the carrier concentration increases from 10^{11} to $10^{12}cm^{-2}$. This enhancement is due to screening effect. This effect is more pronounced when, at low temperatures, nearly elastic scattering mechanisms dominate than at room temperature when polar optical phonon scatterings are predominant.

In biased MODFET, electrons in the 2DEG are submitted to a high electric field parallel to the interface. As a consequence, they are able to transfer first from the lowest subbands to the highest ones and when their average energy becomes comparable to the quantum well height, a lot of them becomes 3D. From a theoretical point of view, the treatment of the transition between a 2D system to a 3D one is a difficult task. Several authors /9/, /17/, /18/, /19/ have calculated the velocity field characteristics in such structures. Some of them have tried to account for several supplementary effects such as electron-electron interactions, degeneracy, screening and so on. On the other hand, MASSELINK et al /20/ have made some measurements of the velocity field characteristics in AlGaAs/GaAs modulation doped heterostructures at 300°K and 77°K. Similar conclusions can be drawn from both theoretical and experimental works : the values of the electric field that gives the peak velocity are smaller than those obtained for bulk GaAs material. This effect can be explained by observing that E_o, the energy of the lowest subband, is close to 50meV, and then the transfer to the L valley is facilitated. However, although theoretical peak velocities are the same (fig.4), experimental results indicate that the 2DEG peak velocity is lower than that of the lightly doped GaAs bulk. Real space transfer to the L and X valleys of the AlGaAs is probably responsible for this effect. For electric field values higher than 20KV/cm, the majority of electrons are in the L and X valleys and it is not surprising that theoretical calculations predict saturated velocity values similar to those observed in bulk GaAs. Finally, we can note that such kinds of theoretical simulations also allow to evaluate the perpendicular and parallel diffusion coefficients that are of primary importance for evaluating noise behaviour.

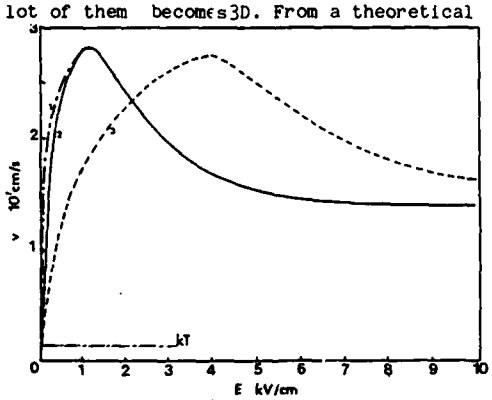


Fig.4 : Velocity field characteristic at 77°K. 2DEG :—·— N_S = 1,15 $10^{12}cm^{-2}$ —— N_S = $10^{11}cm^{-2}$ - - - bulk GaAs

3. TYPICAL BEHAVIOUR AND MAIN CHARACTERISTICS OF "INTRINSIC" GaAlAs/GaAs MODFET

A lot of Monte Carlo simulations /18, 21, 22, 23, 24/ and two dimensional solutions of the relaxation equations /25, 26, 27/ allow to improve our knowledge of physical behaviour of AlGaAs/GaAs MODFET. For instance, figure 5 represents the equiconcentration contours for a typical MODFET (N_D = 10^{18} at/cm^3, a = 360A° and e = 40A°) under usual d.c. bias conditions. We can easily observe that an important carrier injection in the bulk GaAs occurs at the exit of the gate ; this contributes noticeably to the drain current and is responsible for a large part of the output conductance. This result is consistent with the corresponding variations of the transverse component of the electric field that pushes the carriers into the bulk GaAs under the gate /24/. Equipotential and equienergy contours (fig. 6) obtained in similar conditions confirm the existence of a high field domain centered on the quantum well and located at the exit of the gate. Moreover, we can easily remark that the electron average energy is higher than the height of the quantum well during the main part of their drift under the gate, for usual d.c. bias conditions. As a consequence, the quantum character of their behaviour is not evident. It should be noted also that when electrons cross the interface from

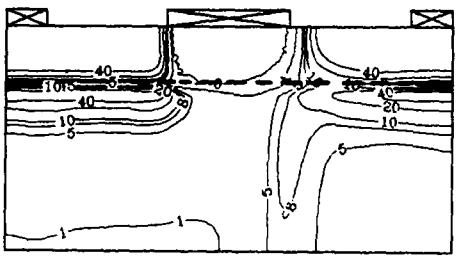


Fig.5 : Equiconcentration contours (in $10^{16}cm^{-3}$) : Vgs = -0.2V; Vds = 2.5V

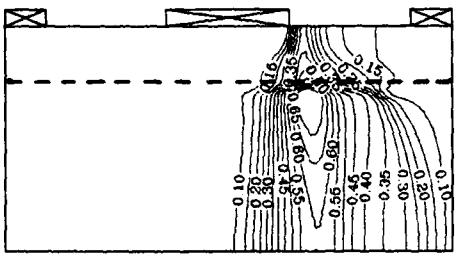


Fig.6 : Equienergy contours (in eV)

GaAs to GaAlAs, their average energy is reduced and they become colder. Consequently, in spite of the low GaAlAs mobility, the electron velocity in AlGaAs can be close to that reached in GaAs. As a consequence, real space transfer occurs at the exit of the gate, as it is clearly shown in the fig. 7, that represents the current density distribution in the same typical MODFET. This explains the decrease in the 2DEG carrier concentration in the gate-drain region so that it becomes smaller than under the source (fig.5). This effect demonstrates also that a large part of the drain current is drawn through the AlGaAs access layer.

Both Monte Carlo and two dimensional relaxation models confirm the importance of non stationary electron dynamics effects for submicron-gate MODFETs /24/. As it is shown in fig. 8, the electron velocities in the 2DEG are much higher than their saturated value during the main part of their drift under the gate and their peak value increases when the gate length decreases. Their average value increases also with the decrease in gate length. These theoretical predictions are very well confirmed by experimental results /28/ : the effective average velocity deduced from transconductance gm to gate capacitance Cg ratio increases from $1.3x10^7$ to $2.3x10^7$ when the gate length decreases from 2µm to 0.15µm. For smaller gate lengths, this quantity remains constant or slowly decreases : this variation is not completely explained and may be due to the combination of both two dimensional effects (small Lg/a ratios) and inertial effects.

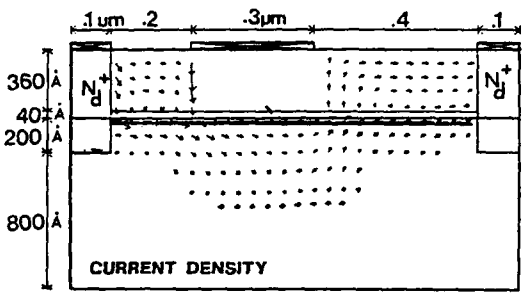


Fig.7 : Current density lines in a MODFET

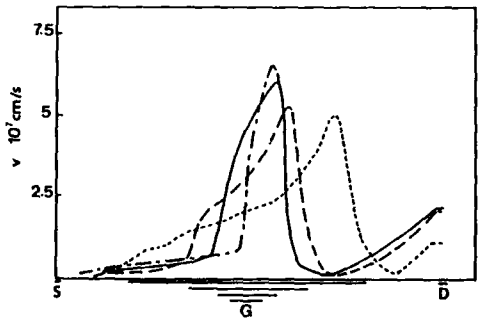


Fig.8 : Spatial evolutions of carriers velocities in the 2DEG/24/.
- - - - -Lg = 1.2µm — — —Lg = 0.6µm
——— Lg = 0.3µm —·—·—Lg = 0.15µm

Note that several authors use the concept of electron saturation velocity deduced either from d.c. current values by using a simple formula ($Jd = q N_s V_s$) or from the ratio gm/Cg. This concept is not consistent with velocity variations represented in fig. 8, where it appears that due to non stationnary electron dynamics effects, carrier velocities saturate only after the end of the gate and after the occurence of an important real space transfer.

Current gain cut-off frequency f_T is the most important parameter that allows to characterize the capabilities of FETs, whatever kind of application is envisaged : low noise or power microwave amplification, digital circuits ... It is close to $g_m/2\pi Cgs$ or $1/\tau$ where τ is the equivalent transit time. Both theoretical calculations and experimental studies indicate that f_T is mainly dependent on gate length. Figure 9 represents a typical dependence, deduced from Monte Carlo simulation, of this cut-off frequency on the gate length, for the intrinsic "HEMT previously considered /24/. It shows that this dependence remains ($1/L_g$) as far as the gate length (Lg) to GaAlAs layer thickness (a) ratio is higher than 5. For smaller ratios, edge

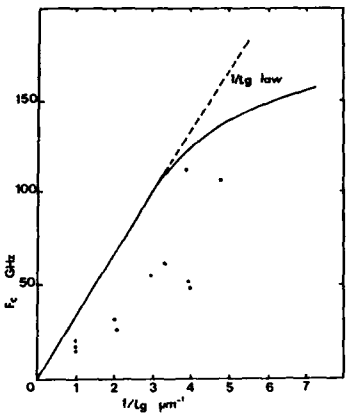


Fig.9 : Theoretical cut-off frequency dependence /24/ upon gate length : • experimental result

effects begin to take a prime importance and the improvement becomes smaller and smaller. As a consequence, we observe a significant improvement of the cut-off frequency due to a decrease of a, provided the ratio Lg/a is smaller than five.

On the same figure are given some typical experimental results reported in literature, and we can observe a significant difference comparing with theoretical predictions (that are quite similar for several authors). Of course, the differences can be partially due to the influence of access resistances (mainly source resistances) that is not always correctly taken into account in the experimental results. We think that surface potential effect that causes a partial dépletion of the GaAlAs layer and then a decrease of the 2DEG carrier concentration in the access zones and that is not taken into account in the simulations can explain this difference. As it appears from preliminary simulations, for high values of the drain current, carriers are hot before they enter under the gate and then the transit time and cut-off frequency can be degraded. As surface potential effect occurs mainly between the edge of the gate and those of the recessed zone, its influence is an increasing fonction of this distance R (Fig. 1). As the gate capacitances Cgs and Cgd of the equivalent scheme increase also when R decreases, this dimension is a critical parameter when designing such devices.

The maximum frequency of oscillation is proportional to the ratio transconductance gm over square root of the output conductance gd. As this last parameter is mainly related to carrier injection phenomenon into the GaAs layer, it is also degraded when the ratio Lg/a decreases and then two dimensional character of the carrier movement becomes predominant.

After we have described several aspects of the physical behaviour of intrinsic devices for classical d.c. bias, we will discuss briefly some problems that appear in practical devices used under various conditions.

4. PRACTICAL DEVICE BEHAVIOUR UNDER VARIOUS CONDITIONS

The equivalent scheme of one practical device includes three access parasitic resistances : source resistance Rs, gate resistance Rg and drain resistance Rd. Two of them Rg and Rs have a strong influence on device performance. Using T-shape electrodes and putting several gates in parallel, the gate resistance can be greatly reduced. For instance, for a 0.25μm gate length, the use of a T gate instead of a conventional one allows to reduce the gate resistance from 400Ω/mm to 80Ω/mm /2/. For smaller gate lengths, this technological improvement becomes essential.

Reduction of the source resistance by contacting, the most correctly as possible, the 2DEG is also a very difficult task. For this purpose, as it is shown in fig. 1, the majority of recent devices structures are recessed ones and an n^{++}GaAs cap layer is introduced between the contact and the AlGaAs layer. As it is heavily doped (several 10^{18}), it provides an additionnal current path nearer to the gate /30/. Ni GeAu alloyed contacts are generally used, but several authors estimate that due to the lateral diffusion, they destroy the 2DEG layer in front of the contact zone. Non alloying ohmic contacts using n^{+}InGaAs cap layer /31/ seem to be a promising solution. Another possibility lies in the use of self aligned ion implantation but it needs refractive gates ; it may increase the gate capacitances Cgs and Cgd and destroy the 2DEG. Its use seems to be limited to devices for digital applications. Physical analysis must account for the distributed character of the source resistance including the existence of various current paths through the cap layer, thin GaAlAs layer, the 2DEG (fig. 10) and through the different barriers by tunnelling effect. At room temperature, the current path through the cap layer is predominant. In microwave or transient regimes, it must be taken into account not only all these elements but also the displacement currents passing through the two depleted zone on both sides of the GaAlAs layer. It is possible by introducing parasitic capacitances (fig. 10). As it has been shown theoretically and experimentally by

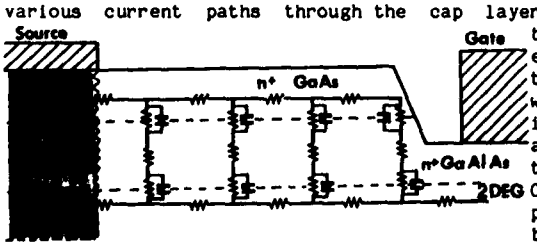


Fig.10 : Equivalent scheme of the source access zone/32/

VERSNAYEN /32/, this effect may contribute to a decrease of the equivalent Rs by a factor of 3 between 2 and 60 GHz.

For low noise amplification, it is almost established that the noise figure is an increasing function of both gate and source resistance. Moreover due to the better carrier-dynamics properties in the 2DEG, MODFET has better noise performance than MESFET for the same gate length /33/. However, some major differences appear between the various theoretical predictions and the agreement is always not very good. Very recently, A. CAPPY /34/ has proposed a new formulation for calculating the FET drain and gate noise sources, based on the concept of the local equivalent circuit at each abscissa in the channel and a new approach of the impedance field method. Typical results are resumed in figure 11, that shows the frequency dependence of noise figure for a quarter micron gate length MODFET by using the new method and the classical one. The influence of access resistance is clearly shown and a very good agreement with recent experimental results of P.M. SMITH /2/ et al can be observed up to 60 GHz.

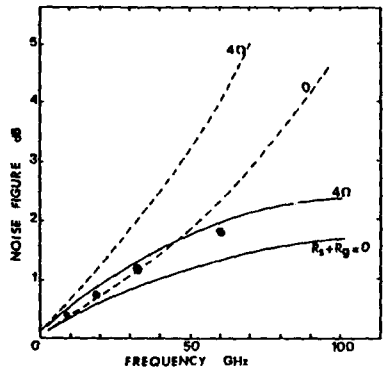


Fig.11 : MODFET noise figure versus frequency/34/. Lg = 0.25µm classical theory – – – ; New formulation ——— • Experimental points

The very large increase of electron mobility when the temperature decreases from room temperature to 77°K (factor 12 to 25) and to 4°K (factor 125 to 250) may let us imagine an important improvement of the performance. is partially true : for instance, late results obtained by K.M. DUH /35/ show that the cryogenic noise performance of MODFET (noise temperature of 5.3°K at 12.5K) approaches that of masers at frequencies below X band. But the improvement in current gain cut-off frequencies do not follow the ratio of carrier mobility but that of average velocity : Monte Carlo simulations /24/ predict only an improvement of 60% between room temperature and 77°K. Experimental results /36/ are in good agreement with these predictions and show also a corresponding decrease of the source resistance. However, these interesting properties cannot be always used due to deleterious effects such as collapse, persistant photoconductivity and variations of the threshold voltage with temperature. All these effects are due to the existence of deep-level trapping in AlGaAs by the so-called DX centers. Collapse is characterized by an important increase of drain resistance (for several ohms to several kΩ in certain cases) when positive strains are applied on the gate and the drain /36, 37/ ; it is due to the trapping of hot electrons in the AlGaAs layer after real space transfer at the gate end. It seems that it can be suppressed by choosing the distance between the gate end and the edge of the recessed zone smaller than 0.3 or 0.4µm /36, 39/. Our knowledge of deep trap level in AlGaAs has been significantly improved recently (40-41). As this deep level approximately follows the L valley minimum which changes with Al composition x and is taken to be located at 140 to 160 meV below the L valley, its influence can be suppressed only by using x values smaller than 0.15-0.17 for usual (10^{18}) and uniform doping levels. However in this case, the height of the potential well ΔEc and then N_s are reduced by a factor of two as compared to conventional x composition (0.25-0.3). Such devices may have a small drain current that constitutes a severe limitation for a lot of applications.

For other applications, several phenomenon limits strongly the capabilities of the device. First the maximum power is proportional to the product of the maximum drain current I_D and the breakdown voltage V_{BD}. For conventional structures, the drain current is smaller than 250mA-300mA/mm instead of 400mA/mm for a typical power MESFET. On the other hand, breakdown voltages are close to 7-10V that is also smaller than that of classical MESFET.
Secondly, one fundamental parameter for digital applications is the threshold voltage or the pinch-off voltage V_p and its uniformity along the wafer. As the pinch-off voltage is given by: $V_p = q.N_D a^2/2\varepsilon$, a very good uniformity of both doping level and epitaxial layer is needed.
In order to overcome the problems met with such conventional structures, a great number of device structures has been developed.

5. NEW MODFET STRUCTURES

We will describe briefly the main aspects and discuss the potential advantages and disadvantages of some of these new structures.

Inverted MODFET

In the first works on MODFET (for instance /29/, /42/, inverted structures with low-doped GaAs on the top were considered (Fig.12). As the channel in this case is much nearer to the gate, much higher transconductance can be expected for the same pinch-off voltage. But the first

realizations indicate small low field mobilities in the channel probably due to an out diffusion of Si impurities through the heterojunction. On the other hand, the conduction in the embedded AlGaAs N+ layer, if it is not completely depleted, may lead to an undesirable parasitic conductance. The optimization of growth conditions or the use of superlattice AlAs/GaAs structures instead of N+ GaAlAs layers has allowed to improve the carrier mobility in the channel (up to 20.000 cm^2/V sec at 77°K) and to obtain very promising performance. For instance CIRILLO et al /5/ have obtained the highest transconductance reported today for 1µm FET (1810 mS/mm at 77°K and 1180 mS/mm at 300°K) in spite of a high value of the source resistance, that may be due to surface effects.The main limitations seem to come from the gate current values. More recently FUJISHIRO /43/ has shown the very promising capabilities of improved structures using n and n+ layers on the top of the undoped GaAs for high speed integrated circuits : 11,8 pS/gate propagation delay and 0.33 mV threshold-voltage standard deviation by wafer. Mainly, they observe a large improvement in the output conductance for small gate lengths (2mS/mm for 0.3µm gate) that indicates a better confinement of carriers and the suppression of short channel effects.

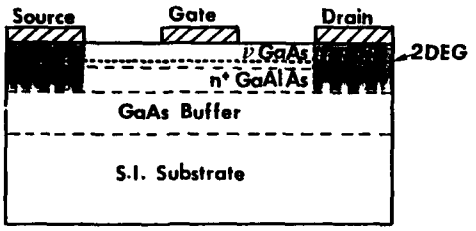


Fig.12 : Inverted MODFET

SISFET and MIS like FET

In order to improve the uniformity of threshold voltage, several authors /4, 44, 45/ have proposed to substitute by n+AlGaAs layer an undoped AlGaAs layer in a MODFET. In this case, the threshold voltage is independant of the characteristics of the AlGaAs layer and is only dependent on the materials used for the gate electrode. In the first realization of SISFETs (fig. 13), the gate was an n+GaAs /4/ with a threshold voltage close to 0V. More recently, the gate was a n+Ge covered by a metallization layer of WSi (Vth ~ 0.1 to 0.4V). In other structures called HIGHFETs or MIS-like FETs, a Schottky contact WSi/undoped GaAlAs gives a threshold voltage close to 0.8V. Very promising results have been obtained, for instance a threshold voltage standard deviation close to 11 mV over 2-inches wafer for a Ge gate SISFET, a transconductance as high as 470mS/mm for a 0.6µm gate and a 54 GHz cut-off frequency. Moreover, p channel MIS like FET's /46/ can be fabricated on the same wafer than n+ channel SISFET (fig. 14) and allow the realization of complementary structures. However, two main problems limit the capabilities of such structures, the gate current that may reach 10^2 to 10^3 A/cm^2 for a voltage of 0.8V above threshold (fig. 15) and the gate resistance that may by 100 times greater than a conventional MODFET. Note that in order to improve the confinement of the 2DEG and then the output conductance, new heterostructures such as the quantum well metal insulator inverted interface semiconductor FET (QWMI^{3}SFET) have been realized with n and p channel /47/.

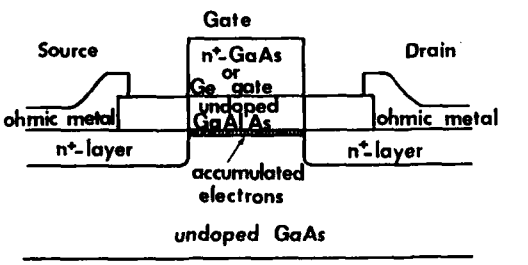


Fig.13 : Schematic structure of SISFET

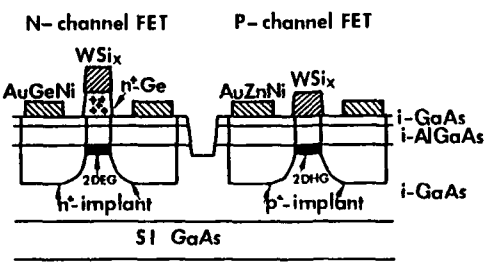


Fig.14 : Schematic view of the n channel and p channel MIS like FET

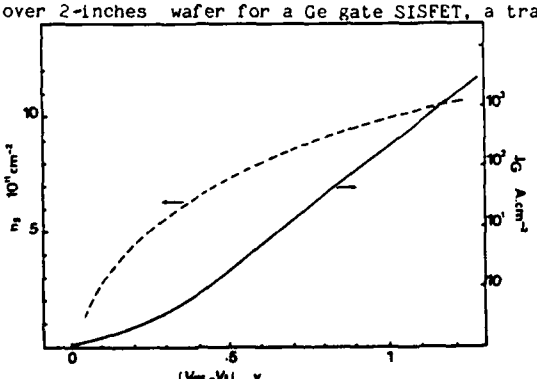


Fig.15 : Typical carriers concentration and gate current variations versus gate voltage for SISFET or MIS like FET

In order to suppress parasitic effects occuring under low temperature conditions, we can conceive the structure in order that the

majority of DX centers will be always depleted of electrons. It can be obtained by using a non uniform doping /36,49/ in the AlGaAs layer (fig. 16). However, such a structure may need gate self aligned implantation technique if we want to keep very small values of source access resistances. More complicated solutions have been proposed for instance by HIYAMAZU /48/. A delta planar doped AlAs/GaAs/AlAs quantum well is formed close to the 2 DEG. As the donor atoms are localized in the GaAs, no parasitic effects can be observed at 77°K.

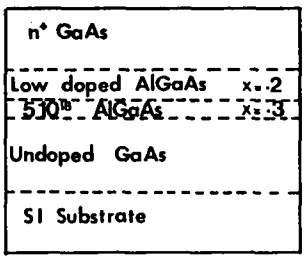


Fig.16 :Example of a specific structure allowing the suppression of low temperature parasitic effects/36/

Multilayer structures

For power applications, theoretical studies and device optimization are necessary to improve at the same time breakdown voltage, maximum current and linearity. Y. CROSNIER et al have shown that not only avalanche effects but also tunneling phenomena /50/ determine mainly the breakdown behaviour and the use of undoped layers under the gate (as in fig. 16) allow to improve the breakdown voltage by a factor of two. Recently A.I. AKINWANDE et al /51/proposed also to reduce locally the doping level in the AlGaAs layer between gate and drain and they obtained very promising results : a gate to drain reverse breakdown voltage close to 20 V. Using two, three, four or six DEG channels in parallel, multichannel HEMTs /3,52,53/ allow to obtain higher drain current and then higher output power : 600 mA/mm to 800 mA/mm instead of 250-300 mA/mm for conventional devices. However the structure becomes very complicated, with 16 different layers for instance, mainly on contacting all the 2DEG with low access resistance. The transconductance may become small when the deepest 2DEG becomes out of control of gate voltage. However, with a proper design and theoretical optimization /52/, a very good linearity can be reached as shown in fig. 17 for a triple channel MODFET. Promising results have been recently obtained in millimeter wave range by HIKOSAKA /54/ with a six channel MODFET (for instance 1,2 W at 30 GHz).

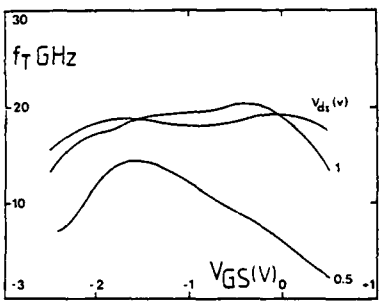


Fig.17 :Cut-off frequency variations versus gate voltage for a 0.8µm triple channel MODFET /52/

Use of InGaAs

The ternary compound $In_{0.53}Ga_{0.47}As$ matched on InP is appeared for a long time as a very attractive material for microwave applications, due to its high mobility (13.800 cm^2 V/sec) and high values of the average electron velocity in submicrometer devices /55/. With an $n^+In_{0.52}Al_{0.48}As$, it constitutes a very promising heterostructure /56/ because of the large conduction band discontinuity (0,53 eV instead of 0,23 eV for $Al_{0.3}Ga_{0.7}As$). As a consequence, 2DEG concentration N_s may reach 3.7 $10^{12}cm^{-2}$ instead of 10^{12} for a conventional AlGaAs/GaAs device. Due to good transport properties of n^+AlInAs (quite similar to GaAs),the deleterious influence of "parasitic MESFET" is less pronounced. Of course, some problems remain such as the quality of InP S.I. substrates, the epitaxial growth of InAlAs and the height of the Schottky barrier. But the device performance is very attractive /58-59/ ; for instance a 7.2pS/gate propagation delay (L_g = 0.2µm) and a current gain cut-off frequency of 130 GHz (L_g=0.2µm).

Pseudomorphic MODFET

As S.I. GaAs substrates are of better quality than InP ones, and having in mind the interesting properties of GaInAs, several authors have proposed to join these unmatched materials /7/. The lattice-mismatched layer is thin enough, so that the mismatch is accomodated as elastic strain : it can be free from dislocation /60/ and be considered as pseudomorphic (fig. 18). HENDERSON /61/ has shown that the optimum value of the In mole fraction y in $Ga_{1-y}In_yAs$ is close to 0.2 and that the critical thickness of the strained layer is a decreasing function of the In mole fraction (typically 200Å for y =0.15). Due to the higher value of the quantum well height (0.3eVfor $Al_{0.15}Ga_{0.85}As$), the carrier concentration is higher than in the conventional structure. Moreover carrier injection and then output conductance are reduced due to the potential barrier between the strained layer and the GaAs buffer. Although the transport properties for both electrons and holes under low and high fields remain not well established, a lot of various devices has been developed : n channel MODFET /60/, p channel quantum well MI^{3}SFET /62/, p channel simple /63/ and multiple /64/

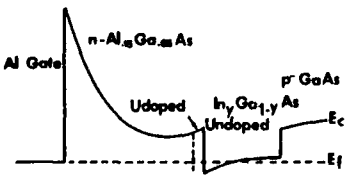


Fig.18 : Pseudomorphic AlGaAs/GaInAs MODFET

quantum well MODFET, planar n doped double quantum well MODFET /65/. At first we can note that due to the small value of the Al mole fraction in AlGaAs (typically x = 0.15), no parasitic effects such as PPC occur at low temperature. Secondly, p layer structures appear very attractive, due to the hole mobility improvement caused by strain, that lifts the zone center degeneracy at the top of the valence band. It has been demonstrated /64/ that this kind of heavy hole has a relatively light effective mass (m^*/me - 0.1 to 0.2).

Finally, a lot of works has been done on both the transport properties /66/ and device capabilities /67/ of pseudomorphic $In_yGa_{1-y}As/In_{0.52}Al_{0.48}As$ heterostructures, where y is the In mole fraction varying from 0.53 (matched on InP) to 0.65 for instance. The basic idea of this last family of devices is to try to profit from carrier velocity improvement resulting from the increase of In mole fraction in InGaAs. The first results are very promising : a drain saturation current of 1 A/mm and a cut-off frequency of 64 GHz /67/ for a 0.5µm gate mono-channel $In_{0.6}Ga_{0.4}As/In_{0.52}Al_{0.48}As$ MODFET.

In brief, pseudomorphic MODFETs have given the best performance known today : a current gain cut-off frequency of 165 GHz for a 0.18µm gate and a maximum output power of 0.67W/mm at 60GHz with 22% of efficiency /65/

6. CONCLUSION

Very important progress has been realized during the last years in our knowledge of phenomena that occur in conventional AlGaAs/GaAs MODFET and in device behaviour and modelisation. Of course, several basic research works must be followed concerning some specific device conditions, such as very low or high temperature operation for instance. Mainly due to the reduction of gate length up to 0.1µm, a significant improvement of performance has been recently obtained. However, due to parasitic effects (gate and source resistances) and carrier injection in the buffer layer, it has not been completely reapedprofit from this reduction. Technological efforts must be done and new kinds of structures can be used in order to overcome this difficulty. Among these new devices, heterostructures using undoped layers seem to offer a good answer to the problem of threshold voltage uniformity for digital applications. Carrier transport properties in pseudomorphic devices constitute a very large and interesting basic research field for the next future. Progress in our knowledge of such pseudomorphic devices and improvements in technological realization can lead to improve the present performance by a factor of two in the next years. For example, a 0.1µm gate device may possess current gain cut-off frequencies that reach 250 GHz and switching time smaller than 4pS at room temperature.

ACKNOWLEDGMENTS

I wish to thank my collegues A. CAPPY, E. CONSTANT, R. FAUQUEMBERGUE, A. VANOVERSCHELDE and J. ZIMMERMANN and several graduate students from Centre Hyperfréquences et Semiconducteurs A. BELACHE, D. DEPREEUW, P. GODTS, T. SHAWKI, F. TEMCAMANI and J.L. THOBEL for providing original results and for very helpful discussions.

REFERENCES

[1] N.J. SHAK and al, IEEE Trans. Elect. Devices, vol. ED-33, 1986.
[2] P.M. SMITH and al, 1988 IEEE MTT Symposium, New York.
[3] P. SAUNIER and al, IEEE El. Dev. Letters, vol. EDL7, n°9, p. 503, 1986
[4] K. MATSUMOTO and al, El. Letters, 24th May 1984, vol. 20, n° 11, p. 462.
[5] N.C. CIRILLO and al, IEEE Elect. Dev. Lett., vol. EDL7, n°2, Feb. 86,p.71.
[6] T. BABA and al, Jap. Journ. of Appl. Phys., vol. 23, n°8, pp. 654, 1984.
[7] A. KETTERSON and al., IEEE El. Dev. Letters, vol. EDL6, n° 12, 1985
[8] C.K. PENG and al., IEEE El. Dev. Letters, vol. EDL8, n°1, January 1987.
[9] J. ZIMMERMANN and WU YEN, Journal of Physique Appliquée, novembre 1987
[10] G. SALMER and al, IEEE Trans. on MTT, JULY 1988
[11] D. DEPREEUW, P. GODTS, Journal de Physique Appliquée, à paraftre, 1988.
[12] J. YOSHIDA, IEEE Trans. on El. Devices, vol. ED33, n°1, Jan. 1986.
[13] D. FRITSCHE, Solid State Electronics, vol. 30, n°11, 1987.
[14] K. LEE and al., Jap. Journ. of Appl. Physics, vol. 23, n°4, 1984.
[15] A. CAPPY, Thèse de Doctorat-ès-Sciences Physiques, Lille, 1986.
[16] J.J. HARRIS and al., Superlattices and Microstructures, vol. 2, n°6, 1986.
[17] K. YOKOYAMA and K. HESS, Phys. Rev. B, vol. 33, n°8, April 1986.
[18] M. MOUIS and al., to be published in Journal de Physique appliquée.
[19] M.A.R. AL MUDARES and B.K. RIDLEY, J. Phys. C. : Solid State Phys. 19, 1986, p. 3179.
[20] W.T. MASSELINK and al., 14th Symp. on GaAs and Related Compounds, 1987.
[21] U. RAVAIOLI and D.K. FERRY, IEEE Trans. on El. Dev., vol. ED 33, n°5,,May 1986.
[22] H. TOMIZAWA and al., IEEE El. Dev. Lett., vol. EDL6, n°7, July 1985.
[23] R. FAUQUEMBERGUE and al., Workshop on Compound Semiconductor, Visby, 1986.

[24] J.L. THOBEL, Thèse de Doctorat, Lille, 1988.
[25] D. WIDIGER and K. HESS, IEEE El. Dev. Let., vol. EDL5, n°7, 1984.
[26] D. LORET and al., IEEE Trans. on El. Dev., ED 33, 1986.
[27] T. SHAWKI and G. SALMER, 14th Symposium on GaAs and Related Compounds, Crête, 1987.
[28] P.R. DE LA HOUSSAYE, IEEE El. Dev. Let., vol. 9, n°3, March 1988.
[29] DELAGEBEAUDEUF and al., Electron. Letters, 16, 1980.
[30] FEUER, IEEE Trans. Electron Dev. ED32, n°7, 1985.
[31] S. KURODA and al., IEEE Trans. El. Dev. Let., vol. EDL8, n°9, sept. 1987.
[32] C. VERSNAEYEN and al., Electr. Letters, vol. 21, n°12, June 1985.
[33] A. CAPPY and al., IEEE Trans. Elec. Devices, vol. ED32, n°12, 1985.
[34] A. CAPPY and al., 4th Workshop of the European Simulation Group, Leuwen, 1988
[35] K.M. DUH and al., IEEE Trans. on Elec. Dev., vol. 35, n°3, March 1988.
[36] A. BELLACHE and al., this issue.
[37] A. KASTALSKY and al., IEEE Trans. on Elec. Dev., vol. ED33, n°3, March 1986.
[38] H. KINOSHITA and al., IEEE El. Dev. Let., vol. EDL6, n°9, sept. 1985
[39] R. FISCHER and al., IEEE Trans. on El. Dev., vol. ED31, 1028, 1984.
[40] T.N. THEIS, 14th Int. Symp. on GaAs and Related Compounds, Heraklion, 1988.
[41] P. GODTS and al., this issue.
[42] D. DELAGEBEAUDEUF and N.T. LINH, IEEE Trans. Elec. Dev. ED 28, P. 790, 1981.
[43] H. FUJISHIRO, 14th Int. Symp. on GaAs and Related Comounds, Heraklion, 1988.
[44] K. MAEZAWA, IEEE El. Dev. Letters, vol. EDL7, n°7, July 1986.
[45] K. ARAI, Jap. Journ. of Appl. Phys., vol. 24, n°8, Aug. 85
[46] M. HIRANO, IEEE Trans. on El. Dev., vol. ED34, n°12, Dec. 1987.
[47] A.KIEHL, IEEE El. Dev. Letters, vol. 9, n°6, June 1988.
[48] S. HIYAMIZU and al., Jap. Journ. of Appl. Physics, vol. 24, n°6, June 1985.
[49] P. GODTS, Thèse Université de Lille, 1988
[50] Y. CROSNIER, this issue.
[51] A.I. AKINWANDE and al., IEEE El. Dev. Letters, vol. 9, n°6, June 1988.
[52] F. TEMCAMANI, Journées Nationales Microélectroniques III-V, Oléron, Juin 1988.
[53] M. HUESCHEN and al., IEDM 1987
[54] K. HIKOSAKA and al., IEEE Elec. Dev. Lett., vol EDL8, n°11, nov. 1987.
[55] A. CAPPY and al., IEEE Trans. on El. Dev., vol. ED27, Nov. 1980.
[56] T. ITOH and al., Proc. of IEEE Cornel Conference, 1985.
[57] Y. NAKATA and al., Jap. Journ. of Appl. Phys., vol. 26, n°1, Jan. 1987.
[58] K. HIKOSAKA and al., IEEE Elect. Device Letters, vol. 9, n°5, May 1988.
[59] L. EASTMAN, Workshop on FET structures, MTT Symposium, New York, 1988.
[60] A. AYDINLI and al., 14th Int. Symp. GaAs and Related Compounds, Heraklion 1987.
[61] T.HENDERSON and al., IEDM, 1987.
[62] A. KIEHL and al., IEEE El. Dev. Letters, vol. 9, n°9, June 1988
[63] T.E. ZIPPERIAN and al., Appl. Phys. Let. 49 (8), August 1986.
[64] T.E. ZIPPERIAN and al., Appl. Phys. Let. 52 (12), March 1988.
[65] P.M. SMITH and al., 1988, IEEE MTT Symposium, New York.
[66] W.P. HONG and al., J.A.P., August 1988.
[67] A. FATHIMULLA and al., El. Letters 24, n° 11, May 1988.

Session 2A Room A : Characterization and noise

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CHARGE PUMPING IN SILICON ON INSULATOR STRUCTURES USING GATED P-I-N DIODES

T. ELEWA, H. HADDARA, S. CRISTOLOVEANU and M. BRUEL*

Laboratoire de Physique des Composants à Semiconducteurs (CNRS-UA), ENSERG, 23, rue des Martyrs, F-38031 Grenoble Cedex, France
**LETI, CENG, 85X, F-38041 Grenoble Cedex, France*

Résumé - Nous présentons une extension de la technique de pompage de charge sur des diodes P^+IN^+ à grille de contrôle, fabriquées sur silicium sur isolant. Cette méthode nous permet d'accéder aux propriétés des interfaces des structures SOS et SIMOX, tout en évitant l'utilisation de transistors MOS à 5 contacts. Les mesures effectuées sur SIMOX en pulsant la grille et/ou le substrat révèlent l'existence d'une forte densité d'états d'interface rapides et de pièges volumiques au voisinage de l'oxyde enterré.

Abstract - The extension of the charge pumping technique to gated P^+IN^+ diodes fabricated on silicon on insulator is analysed. This method allows us to evaluate the interface properties in SOS and SIMOX structures, without the need for 5-terminal MOS transistors. The experiment, performed on SIMOX films by pulsing both the gate and substrate, reveal the existence of a high density of fast interface states and bulk traps near the buried oxide.

1 - INTRODUCTION

The great interest in Silicon On Insulator (SOI) structures has generated an effort in the research of appropriate and reliable interface characterization techniques. Indeed, conventional MOS capacitance measurements suffer from the inherent existence of high series resistances and interface coupling in very thin films /1/. Current based techniques (dynamic transconductance /2/, low frequency noise /3/) are much more suitable for SOI.

It has been recently demonstrated that the charge pumping can be successfully applied to short-channel MOSFET's fabricated in SOI /4,5/. This technique is able to resolve the contributions of front and back interfaces while avoiding the influence of parasitic capacitances and series resistances. A major practical limitation is, however, due to the necessity of a film contact to measure the charge pumping current ; this requires the fabrication of special 5-terminal MOSFET's. In this paper we make use of the presence of the charge pumping phenomenon in standard gate-controlled diodes /6/ in order to demonstrate that charge pumping measurements can be used to characterize not only the two interfaces but also the bulk traps in SOI films.

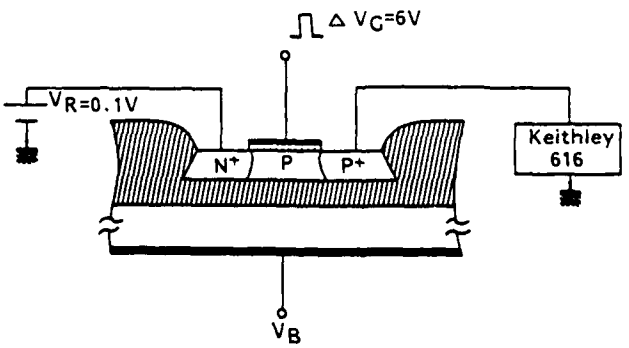


Fig.1. Experimental set-up of charge pumping technique

2 - EXPERIMENT

The schematic set-up of the experiment is shown in Fig. 1. The P^+PN^+ diode is reverse biased while the gate is continuously pulsed from strong inversion to accumulation. In strong inversion, minority carriers are rapidly supplied by the N^+ contact to form the inversion layer and fill the interface traps. As the device is pulsed into accumulation, the inversion layer electrons return almost instantaneously to the N^+ contact whereas trapped electrons recombine with majority carriers supplied by the P^+ contact.

This recombination and the underlying net flow of holes give rise to a charge pumping current I_{cp} in the film which is proportional to both the pulse frequency (Fig. 2) and the amount of interface traps /4-6/. It follows that the P^+ contact plays here the role of the film contact in the 5-terminal MOSFET. The method was validated by comparison with conventional charge pumping in 5-terminal MOSFET's and dynamic transconductance measurements.

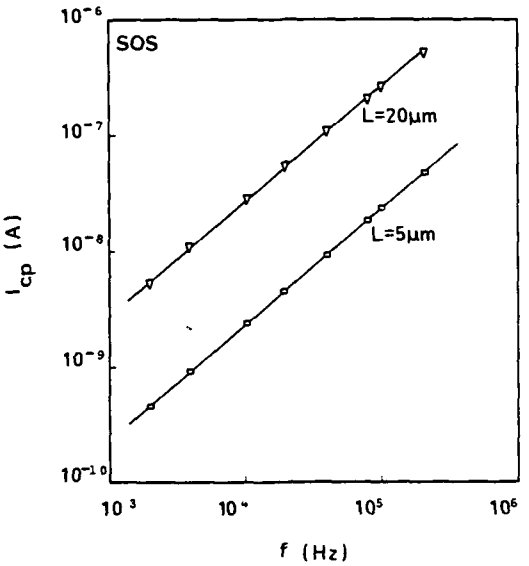


Fig.2. Charge pumping current versus frequency for SOS diodes

The influence of gate length, reverse bias and substrate bias are also investigated. A further development consists of taking into consideration the contributions of two emission processes. Firstly, some of the interface states can be directly filled by the emission of holes to the valence band. Secondly, a portion of the trapped minority carriers are emitted from the interface states in the silicon conduction band instead of recombining with substrate majority carriers. The model shows that by changing the 4 parameters of the gate pulse (top and bottom levels, rise and fall times) it becomes possible to determine the energy profile of interface states /7/.

The experiment was carried out in silicon on sapphire (SOS) films 0.5 µm thick and in SIMOX structures formed by deep oxygen implantation (200 keV, 1.7×10^{18} O/cm^2) and high temperature annealing. The thicknesses of the Si film and buried oxide were about 0.2 µm and 0.4 µm, respectively. The silicon overlayer forming the diode base was either P-type doped or natural with an unintentional residual doping of about 10^{15} cm^{-3} still subsisting as a consequence of the annealing process and oxygen activity.

3 - DISCUSSION

Figure 2 shows the variation of I_{cp} with the pulse frequency for two SOS diodes of different lengths (L = 5 µm, L = 20 µm). The linear behaviour of I_{cp} with the frequency is in agreement with the theoretical predictions [7], however the ratio between the two currents is greater than that between the two gate areas. This is attributed to a geometric component present in the longer device. Indeed, some of the inversion electrons may recombine with holes supplied by the P^+ contact and contribute to the total measured current (Fig. 3 for L = 20 µm). This effect is observed for transistors with relatively long channels : L ≥ 20 µm. Bearing in mind that the inversion layer formation delay in a P-I-N diode is twice as large as in a transistor, one should expect the geometric component to exist in diodes of gate lengths of the order of 10 µm.

Figure 3 shows I_{cp} versus the pulse base level for the two diodes. For L = 20 µm, I_{cp} does not go to zero when the whole pulse lies in strong inversion ; according to Elliot /8/, this confirms the presence of a geometric component. In contrast, for the shorter diode, I_{cp} decays rapidly as the pulse excursion is shifted towards accumulation or strong inversion.

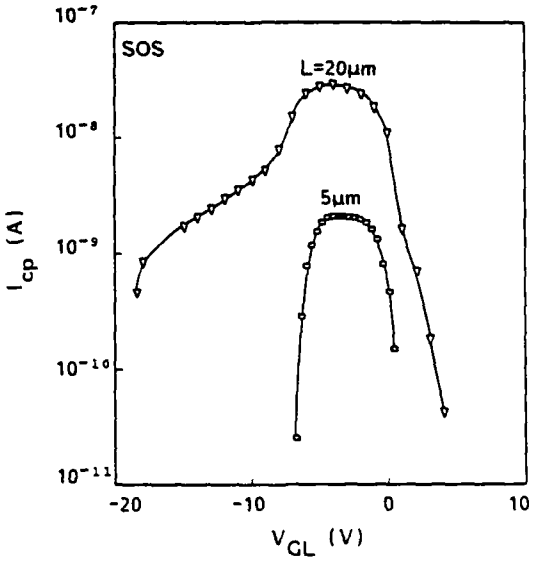


Fig.3. Influence of the geometric component on the charge pumping current measured as a function of the gate pulse base level

More information about bulk traps is obtained by pulsing simultaneously both gates. According to the new principle of *volume inversion* /8/, not only the two interfaces but also the whole film volume are now driven from accumulation into strong inversion. In our case (Fig. 4), the total value of I_{cp} is almost the same as that of the back interface. This implies that bulk traps are localized mainly near the buried oxide.

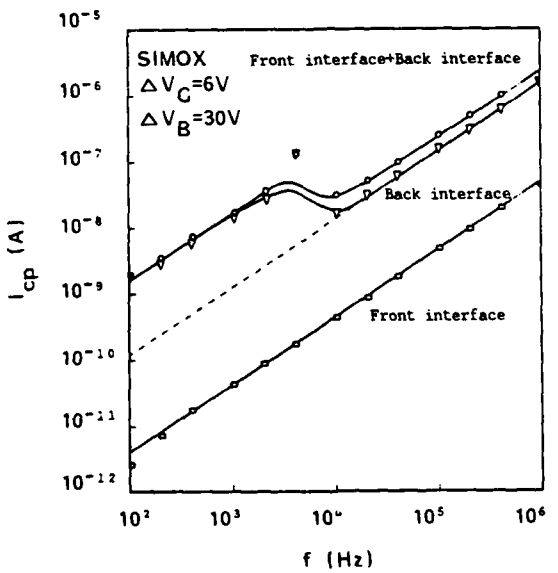


Fig.4. Charge pumping current versus frequency

Furthermore, Fig. 5 shows a maximum in I_{cp} at the front interface as the substrate voltage is varied. The behaviour of this curve can be explained by the fact that when the back interface is in strong inversion or in accumulation, the influence of the front gate bias on the back interface potential is negligible. On the other hand, when the back gate is biased in depletion or weak inversion a larger portion of the energy gap can be scanned near the buried oxide as a consequence of the front gate pulse. This leads to a partial pumping of those charges which are trapped near the back interface. The maximum in I_{cp} (Fig. 5) corresponds, therefore, to an additional contribution of the more defective region situated close to the buried interface.

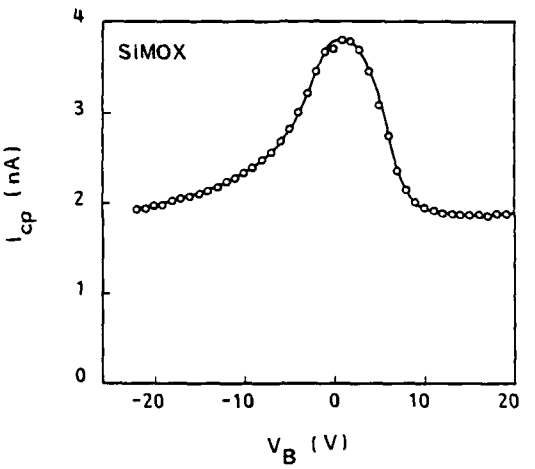


Fig.5. Front interface charge pumping current versus back gate bias

4 - CONCLUSION

In conclusion, the extension of the charge pumping method to gated diodes offers an elegant alternative to 5-terminal transistors for the assessment of bulk traps in the Si film and fast states at the interfaces of SOI structures. The density of fast states was found to be reasonably low (10^{11} $cm^{-2}eV^{-1}$) at the front interface in both SOS and SIMOX. A higher density of traps was determined for the buried interface and results probably from the oxygen implantation induced damage. The charge pumping in P-I-N diodes is a simple and accurate method and does not require any ad-hoc test device to be specially processed.

Acknowledgements. We thank Drs. C. Jaussaud, J. Margail, A. Auberton, B. Giffard (LETI), A. Chovet, G. Dimopoulos (LPCS) and Y. Gris (Thomson) for providing SIMOX and SOS structures.

REFERENCES

/1/ Lee, J.H. and Cristoloveanu, S., IEEE Electron Dev. Lett. EDL-7 (1986) 537.
/2/ Haddara, H. and Ghibaudo, G., Solid-State Electron. 31 (1986) 1077.
/3/ Chovet, A., Boukriss, B., Elewa, T. and Cristoloveanu, S., Appl. Surf. Science 30 (1987) 390.
/4/ Tack, M.R., Wouters, D.J., Groeseneken, G.V., Claeys, C.L. and Maes, H.E., E-SOI Workshop, Meylan, France (1988).
/5/ Elewa, T., Haddara, H., Cristoloveanu, S. and Davis, J., E-SOI Workshop, Meylan, France (1988).
/6/ Brugler, J.S. and Jespers, P.G.A., IEEE Trans. Electron Device ED-16 (1969) 297.
/7/ Haddara, H. and Cristoloveanu, S., Solid-State Electron. 29 (1986) 767.
/8/ Eliot, A.B.M., Solid-State Electron. 19 (1976) 241.
/9/ Balestra, F., Cristoloveanu, S., Benachir, M., Brini, J. and Elewa, T., IEEE Electron Device Lett. EDL-8, (1987) 410.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A CONTACTLESS TECHNIQUE FOR THE CHARACTERIZATION OF INTERNALLY GETTERED CZ SILICON

M.A. BRIERE[1]

Hahn-Meitner-Institut Berlin GmbH, Dept. Dataprocessing and Electronics, Glienicker Strasse 100, D-1000 Berlin 39, F.R.G.

Abstract - A characterization technique is described which is capable of measuring the salient electronic properties of the layered structure formed during the internal gettering process in Czochralski grown silicon. The method is contactless and involves the use of a two laser, pump-probe, system to measure the carrier lifetimes in the defect-free-zone (DFZ) and the bulk as well as to measure the width of the DFZ. Excellent agreement is found with the results of theory and those of the standard bevel and etch method.

INTRODUCTION

The use of oxygen rich Czochralski-grown silicon wafers, in the manufacture of electronic devices, has created a need for new methods to characterize the salient features of the resulting layered structure, which is shown schematically in Figure 1. High temperature processing, early in the manufacture of electronic devices, causes the oxygen, incorporated during crystal growth, to diffuse out of the near surface region. Subsequent hot processing allows the supersaturated solution of oxygen in silicon to form a mixture of silicon-oxygen phases. Due to the gross density mismatch between the silicon matrix and the SiO_x precipitates, silicon interstitials are generated at the precipitation site. These interstitials often condense to form dislocation loops and stacking faults. Therefore the resulting substrate is found to have a highly defective bulk, where precipitation is dense, and a near defect free zone (DFZ) near the wafer's surface where out-diffusion has denuded the region of sufficient oxygen for precipitation. This region is also known as a precipitate free zone (PFZ). This process of oxygen out-diffusion and precipitation is known as internal gettering, as the precipitate-dislocation-complexes are found to be effective sites for the trapping metallics away from the device active region [1]. Effective use of this gettering method requires the ability to evaluate the relationship of various process and material parameters on the efficiency of the gettering action and the effects on device performance and yield. Here, a contactless characterization technique is demonstrated, which yields excellent agreement with the standard bevel and etch method, for the determination of the DFZ/PFZ, as well as providing a measure of the near-surface and bulk carrier lifetimes.

The method described here is an extension of the pump-probe technique, pioneered by Harrick [2] and since advanced by several authors [3-5], which monitors the modulation of the transmision of a sub-bandgap probe by excess free carriers. In this case, a dye laser is used to generate free carriers at the wafer's surface with nanosecond pulses of highly absorbed blue light (480 nm). The infrared output of a CO_2 (11.7 μm) laser is used as a probe of the free carrier concentration. The two beams are made nearly collinear using a ZnS optic and are normally incident to the sample. The transmitted intensity of the infrared signal is measured using a fast (20ns), cooled, HgCdTe detector. A schematic of the experimental arrangement is shown in Figure 2. A very wide band-width amplifier is used so as not to elliminate any fourier components of the signal. The temporal evolution of the modulated infrared transmission is directly related to the free carrier dynamics. Sampling resolution on the order of tens of nanoseconds as well as signal to noise ratios of ≈100, allow the analysis of carrier dynamics, as the carriers diffuse from the surface into the bulk, using single transient measurements. Typical results for a homogeneous sample and a layered structure, are shown in Figures 3 and 4 respectively.

---

[1] Previous address : IBM East Fishkill Laboratories, Hopewell Jnct., NY 12533, U.S.A..Work was originated as part of an M.S. Thesis at the Physics Dept. of Worcester Polytechnic Institute, Worcester, MA 12603, U.S.A., completed 10.1987

THEORETICAL BACKGROUND

The absorption of light in silicon is determined by lattice and impurity vibrational coupling as well as electronic contributions,

$$\alpha_p = \alpha_l + \alpha_i + \alpha_e \qquad (1)$$

The electronic contribution to the absorption, for sub-bandgap light, is linearly related to the free carrier density,

$$\alpha_e = K_n n + K_p p \qquad (2)$$

where $K = \lambda_p^2 C$ and C is a material constant. Therefore, the transmission of sub-bandgap light is directly related to the free carrier density. Long probe wavelengths should be used for greater sensitivity.

Assuming an instantaneous pulse source of above bandgap light, the initial spatial distribution of excess free carriers is:

$$\rho(x,0) = \frac{I_{m0} \exp(-\alpha_m x)}{h\nu_m \alpha_m}(1 - R_m)^2 \qquad (3)$$

The modulated transmitted probe beam intensity is then:

$$\frac{\Delta I_T(t)}{I_{T_0}} = e^{-\int_0^s \Delta\alpha_p(x,t)\,dx} \qquad (4)$$

where the change in absorption is related to the change in free carriers as in eq (1). In the homogeneous case, the result is a single exponential transient, as shown in Figure 3. For the layered structure of Figures 1 and 3, the integrated free carrier concentration along the thickness of the sample is more dynamic {2},

$$\bar{\rho}(t) = \int_0^s \rho(x,t)\,dx = \bar{\rho}_{DFZ}(t)\exp(-t/\tau_D) + \bar{\rho}_B(t)\exp(-(t-t_d)/\tau_B) \qquad (5)$$

and

$$\bar{\rho}_{DFZ}(t) = \int_0^{x_{DFZ}} \rho(x,t)\,dx \qquad \bar{\rho}_B(t) = \int_{x_{DFZ}}^{s} \rho(x,t)\,dx \qquad (6)$$

where τ_D and τ_B are the carrier lifetimes in the DFZ and bulk, respectively and t_d is the time for the carriers to diffuse from the surface to the bulk. On average, this time is related to the width of the DFZ by:

$$t_d \approx \frac{x_{DFZ}^2}{D_A} \qquad (7)$$

For the high injection levels used here ($I_{m0} \approx$3-5 uJ), the carrier lifetimes are given by:

$$\tau \approx \tau_{n0} + \tau_{p0} \qquad (8)$$

as in the Shockley-Read-Hall (SRH) theory [6]. The effect of the gradient in carrier lifetimes at the boundary between the DFZ and bulk region prevents the straight forward application of the erf(x) in the solution of eqs. (6) and (7). In fact no closed-form solution is possible for this problem and numerical simulations must be used to completely evaluate experimental results. This work is presently being persued. However, some simple analysis is possible without analytical expressions for the free carrier dynamics. Using the time of flight concept, implicit in eq (5), it is possible to evaluate the width of the DFZ as well as to measure the effective carrier lifetimes within the DFZ and the bulk.

The analysis consists of evaluating the second natural log, (ln(ln(x)), of the measured modulated probe intensity which is described by eq. (4). The result is a measure of the temporal evolution of the integrated carrier concentration, on a logarithmic basis (ICCLB). An example is shown in Figure 5. The slope of this curve yields the effective carrier lifetimes as a function of time. This result can be converted to a depth profile of the lifetime if the carrier dynamics are adequately known in order to deconvolute the two factors in eq. (5). During the earliest portion of the transient, carrier dynamics are governed mainly by the properties of the DFZ, to which they are confined. The integrated carrier density changes little in this region, since the lifetime in the DFZ is much larger than the transit time for the carriers to reach the bulk. At latter times, the transient is governed increasingly by the properties of the bulk region. For long times, the slope of the ICCLB is a constant and yields the effective bulk lifetime. Though not done here, this lifetime must be corrected for the effect the DFZ in order to obtain the true bulk lifetime. The intercept of this line with the time axis yields the diffusion time of eqs. (5) and (7).

Even a brief description of the kinetics of oxygen precipitation in silicon is beyond the scope of this paper. The reader is referred to the wealth of information in the literature {3}. It will suffice to note that the width of the precipitate free zone can be adequately predicted on the basis of classical diffusion, nucleation and diffusion limited precipitation theories and is approximately given by [7]:

$$x_{DFZ} \approx 2\sqrt{D_{os} t_D}\left(\frac{C_p - C_s}{C_i - C_s}\right) \qquad (9)$$

{2} Of course this reasoning can be extended for any number of layers to describe an arbitrary lifetime profile.

{3} For example see, Mat. Res. Soc. Symp. Proc. 36(1984), 46(1985), 59 (1986).

SAMPLES

Silicon wafers, p-type (100) 125mm dia., from the seed section of a CZ crystal were used to evaluate the technique. The samples were etched and polished on both sides to an optically brilliant finish and had initial oxygen concentrations of 33 +/- .5 ppma, according to the 1978 ASTM standard [8] (for historical reasons), as measured with a Digilab FTIR. The samples were denuded at 1200 C for 4-20 hrs. and were then further processed so as to cause precipitation in the wafers bulk. Final processing was performed in an oxidizing ambient so as to passivate the wafer's surface. MOS capacitor measurements of the surface recombination velocity indicate a maximum value of .1 cm/s.

RESULTS and DISCUSSION

In Figure 6, the DFZ width as measured with this technique is compared with the results for the bevel and etch method (PFZ) and those of theory (DZ) (eq. 9). As can be seen, excellent agreement between the infrared method and theory exist. The results for the bevel and etch method are slightly higher, in agreement with previous results [7], and probably due to the sensitivity limit of 100 A diameter defects [9](the critical radius of oxygen precipitates in this case is about 6-10 A [7]). The error in determining the PFZ width with the bevel and etch method is high and the results are often subjective.

The measured effective bulk lifetimes are shown in Figure 7, plotted reciprocally vs. the corrected oxygen precipitation, which is related to the measured precipited oxygen by:

$$\Delta[O_i]_p(corr) = \Delta[O_i]_p\left(\frac{X_s}{(X_s - 2X_{PFZ})}\right) \qquad (10)$$

in order to yield a measure of the bulk precipitate density. The apparent linear relationship between the reciprocal lifetime and the bulk precipitate density is in agreement with SRH theory for recombination statistics [2], where:

$$\tau = \frac{1}{\sigma v_{th} N_T} \qquad (11)$$

This non-destructive technique therefore allows the characterization of the salient features of the layered structure formed during internal gettering processes. It can be extended to the analysis of other structures with layers of different lifetimes, ie. epitaxial systems or power device substrates. The practical (1 um) limiting factors in the determination of the DFZ width is the modulating pulse width and the measurement sampling rate. This technique may also be applicable for on-line monitoring of substrate properties during device processing.

REFERENCES

[1] T.Y. Tan, E.E. Gardner and W.K. Tice, Appl. Phys. Lett. 30, 175 (1977).
[2] N.J. Harrick, Phys. Rev. 103, 1173 (1956).
[3] M.A. Afromowitz and M. DiDomenico, J. Appl. Phys. 42, 3205 (1971).
[4] R.J. Schwartz and J.L. Gray, Proceedings of the 1978 IEEE Conferance of Photovoltaic Specialists, p.83 (1978).
[5] D.L. Polla, IEEE Trans. on Electro. Dev. Lett. EDL-4, 185 (1983).
[6] J.S. Blakemore, Semiconductor Statistics, Pergamon Press, New York (1962) Chapters 4,7-9.
[7] K. Yang, J.F. Carle and R. Kleinhenz, J. Appl. Physics, 62, 4890 (1987).
[8] ASTM Standard F 391-78.
[9] K.H. Yang, J. Electrochemical Soc. 131, 1140 (1984)

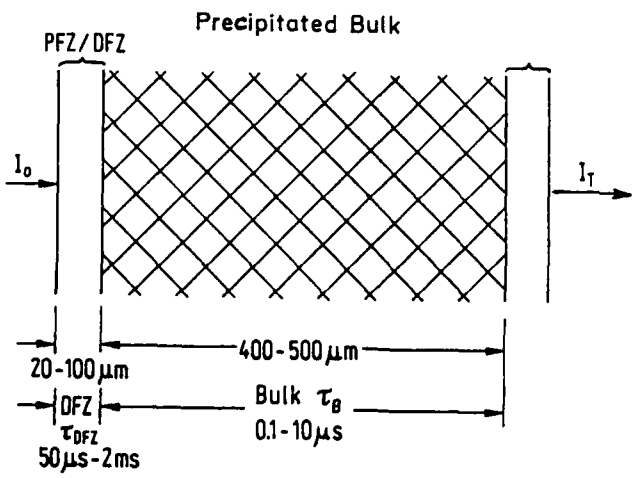


Figure 1
Schematic representation of an internally gettered CZ Si wafer. Transmission measurements are made through the thickness of the wafer. Excess carriers are generated near the wafer's surface.

Table 1
Definition of Symbols

Symbol	Current Usage
ρ	Excess Free Carrier Density [cm^{-3}]
n,p	Electron and Hole Densities,resp. [cm^{-3}]
τ	Free Carrier Recombination Lifetime [μs]
τ_{no}, τ_{po}	Low Level Injection Limit for Carrier Lifetimes [μs]
σ	Carrier Capture Cross Section [cm^2]
v_{th}	Carrier Thermal Velocity [cm s^{-1}]
N_T	Effective Density of Recombination Centers (SRH) [cm^{-3}]
t	Time [s]
t_d	Average Carrier Diffusion Time [μs]
t_D	Duration of Oxygen Denuding Process [s]
D_{Ox}	Oxygen Diffusivity in Silicon During Denuding Process [cm^{2}s^{-1}]
DA	Ambipolar Diffusivity of Free Carriers [cm^{2}s^{-1}]
α_m, α_p	Absorption Coefficient for Modulating and Probe Beams [cm^{-1}]
α_i,α_a,α_c	Electronic, Ampurity and Lattice Contributions to Absorption Coefficient [cm^{-1}]
K_n, K_p	Optical Absorption Cross Section for Electrons and Holes [cm^2]
$h\nu_m$	Energy of Pump/Modulating Photons [eV]
λ_p	Free Space Wavelength of Probe Beam [μm]
I_{m_o}, I_o	Incident Pump and Probe Beam Intensities, resp., [Wcm^{-2}]
I_T, ΔI_T	Quiescent and Modulated IR Transmission Intensities, resp.
R_m	Reflectivity of Modulating Beam
$[O_i]$	Interstitial Oxygen Concentration in Silicon [ppma]
$\Delta[O_i]_p$	FTIR Measured Precipitated Oxygen [ppma]
C_o, C_s	Initial and Solubility Limit for $[O_i]$ During Denuding Process [ppma]
C_p	$[O_i]$ at the Edge of the PFZ [ppma]
X_s	Sample Thickness [μm]
X_{DZ},	Width of the Defect Free Zone [μm]
PFZ	Precipitate Free Zone

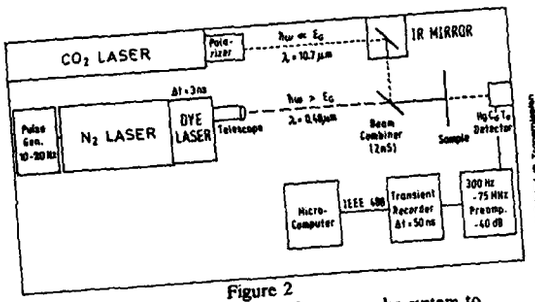


Figure 2
Experimental arrangement of pump-probe system to measure free carrier dynamics in semiconductors. The dimensions of the optical set are 2m by 1m.

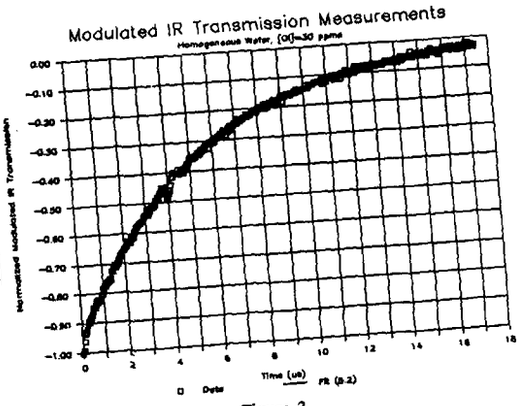


Figure 3
Modulated infrared transmission measurement for a homogeneous silicon wafer (ie. a wafer without a significant lifetime gradient along its thickness)

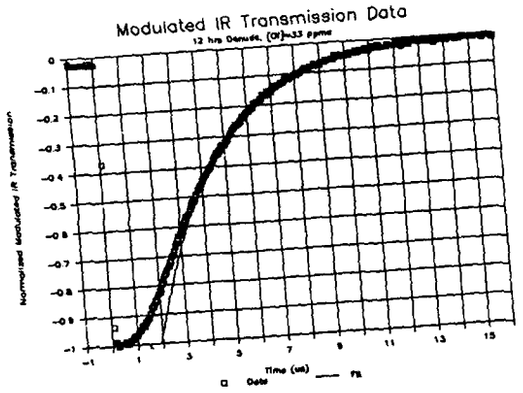


Figure 4
Modulated infrared transmission measurement for a denuded and precipitated silicon wafer. The duration of the denuding process was 12 hrs at 1200 C.

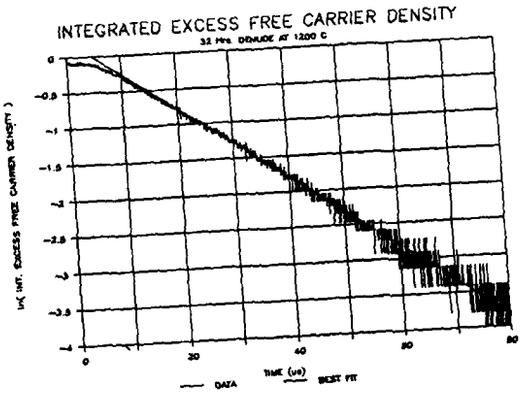


Figure 5
The temporal evolution of the integrated excess free carrier concentration as determined from the results of modulated IR measurements. This sample was denuded for 32 hrs at 1200 C.

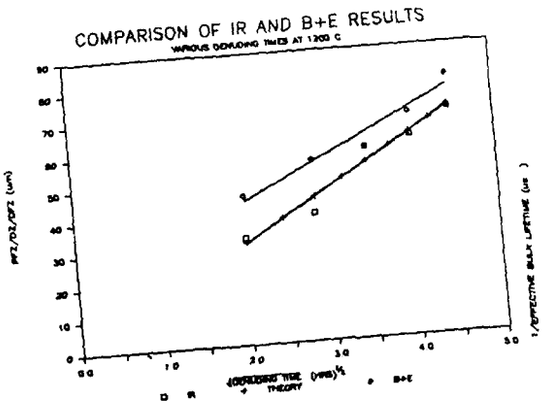


Figure 6
Comparison of results for the measurement of the width of the Defect Free (DFZ), Denuded (DZ) and Precipitate Free (PFZ) Zones as determined by the present IR technique, precipitation theory, and the standard bevel and etch method as a function of the square root of the denuding time. Excellent agreement amongst the results is seen. The solid lines are best fits to the data.

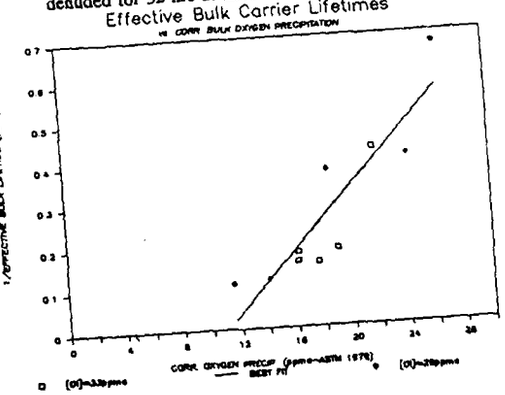


Figure 7
Reciprical of the measured free carrier lifetimes in the bulk vs. the corrected oxygen precipitation as measured by FTIR and corrected as described in the text.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A MICROWAVE METHOD FOR CONTACTLESS MEASUREMENT OF THE LIFETIME OF FREE CARRIERS IN SILICON WAFERS

T. OTAREDIAN, S. MIDDELHOEK and M.J.J. THEUNISSEN*

Delft University of Technology, Department of Electrical Engineering, Electronic Instrumentation Laboratory, PO Box 5031, NL-2600 GA Delft, The Netherlands
**Philips Research Laboratories, PO Box 80000, NL-5600 JA Eindhoven, The Netherlands*

Abstract - Following an optical excitation of the free carriers in silicon wafers, their lifetimes are determined by measuring the reflected-microwave power. The relations of the lifetime with the light intensity of the optical source and with the parameters concerning the deep recombination centers in the bulk and at the surface of the wafer are discussed.

1 - INTRODUCTION

To improve process control in modern IC plants, the ability to test and measure the important parameters of silicon wafers, both prior to and during processing, is highly desirable. The usual methods necessitate the formation of ohmic contacts and special test structures, which may not only change the parameters to be tested but may also cause damage, add contamination and make the interpretation of the results very difficult. This paper presents a convenient contactless method, which is capable of measuring the free carrier's lifetime in wafers. The theory is also addressed. The many advantages of the method include the freedom from damage caused by a probe and the avoidance of contamination as well as *in situ* and fast measurement. In addition, the sample needs no preparation, contact resistance-effects are eliminated and the measurement can be carried out readily after each processing step. The method is an improved version of that introduced by *Ogita* /1/ which is based on the measurement of reflected-microwave power from the wafer as a function of time after a pulsed-optical excitation. The change of reflected-microwave power is proportional to the excess concentration, if it does not exceed 3% of the majority concentration /2/. The lifetime of free carriers can be obtained by determining the reciprocal value of the slope of the logarithm of the signal, which is proportional to the reflected power. This measurement method can be also used for other semiconductor wafers.

2 - THE MEASUREMENT SYSTEM

The measurement system, which is used for the measurement of lifetime of the free carriers, has the functional diagram shown in Fig. 1.

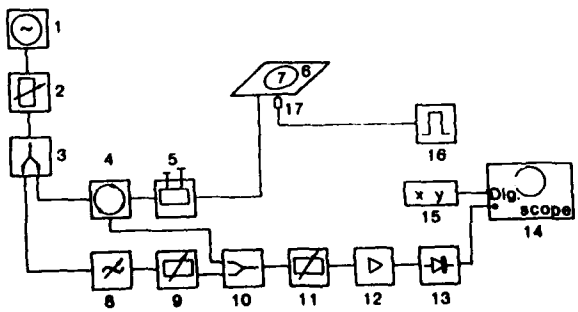


Fig. 1. The functional diagram of the measurement system

The microwave generator (1) generates microwaves with a power of 80 mW and a constant frequency of 2.8 GHz. The microwave power is directed via a waveguide to an attenuator (2) in order to adjust the emitting power level. The microwave power is then directed to a splitter (3), which divides the power into two equal parts. The first part is directed to an antenna (6) in order to irradiate the wafer (7) by the means of a

circulator (4) and a stub tuner (5). The stub tuner is used to match the impedance of the antenna and the wafer to the system. The reflected power from the sample is directed via the circulator to a combiner (10), through which this power is added to half of the original power. By adjusting the phase shifter (8) and attenuator (9) it is possible to eliminate the offset partially. The resulting power is then directed to an attenuator (11), a microwave power amplifier (12) and a microwave detector, which provides a voltage proportional to the microwave power. The output voltage can be visualized by a digital oscilloscope (14) and can be plotted by a plotter (15). The pulse generator (16) drives the optical source, which is a LED (17), with a wavelength of 875 um. An analogue-electronic circuit is also added to the measurement system, which provides the logarithm of the output signal.

3 - THE MEASUREMENT PRINCIPLE

The reflection coefficient of microwave power with a constant frequency, from a semiconductor sample, is dependent on the conductivity of the semiconductor, σ /2/. The change in the reflected-microwave power is a consequence of presence of excess-free carriers. In other words the change in the conductivity, (i.e. a small change in the free-carrier concentration) for small perturbations, almost 3%, can be approximated by Eq. 1,

$$\Delta P_{ref}(t) = [dR(\sigma)/d\sigma]\Delta\sigma(t)P_{in} \tag{1}$$

where P_{ref} is the reflected-microwave power from the wafer, P_{in} is the source-microwave power, $R(\sigma)$ is the reflection coefficient and $\Delta\sigma$ is the change in the conductivity, which is given by Eq. 2,

$$\Delta\sigma(t) = (\mu_n+\mu_p)q\Delta n(t) \tag{2}$$

where μ_n and μ_p are the mobilities of the electrons and holes respectively, q is the elementary charge and $\Delta n(t)$ is the excess concentration of the free carriers generated by the light pulse. From Eq. 1 and Eq. 2 it follows,

$$\Delta P_{ref}(t) = \Phi\Delta n(t) \tag{3}$$

where Φ is a value which remains constant during the measurement and is given by Eq. 4,

$$\Phi = [dR(\sigma)/d\sigma](\mu_n+\mu_p)qP_{in} \tag{4}$$

The effective lifetime can be obtained by calculating the reciprocal value of the slope of the logarithm of the excess-carrier concentration as a function of time as is given in Eq. 5,

$$\tau_{eff} = -dt/d\ln[\Delta n(t)] \tag{5}$$

4 - RECOMBINATION IN THE BULK OF SILICON WAFERS AT LOW AND HIGH INJECTION LEVELS

The decay process is theoretically modelled by using the Shockley, Read and Hall model in its general form. Information about these recombination centers can be obtained by matching a measured decay process on the model. A general expression is derived from this model, which can describe the behavior of lifetime at different temperatures and at different injection levels. The general form of the Shockley, Read and Hall model is given in Eq. 6,

$$\frac{d\Delta n(t)}{dt} = \frac{v_n v_p \delta_n \delta_p N_t (pn - n_i^2)}{v_n\delta_n n + v_n\delta_n n_i \exp[(E_t-E_i)/kT] + v_p\delta_p p + v_p\delta_p n_i \exp[(E_i-E_t)/kT]} \tag{6}$$

where,

$$n = n_0 + \Delta n \tag{7}$$
$$p = p_0 + \Delta n \tag{8}$$

v_n and v_p are the thermal velocities of electrons and holes, σ_n and σ_p are the capture cross-sections of electrons and holes, N_t is the trap concentration, n_i is the intrinsic concentration, n_0 and p_0 are the electron and hole concentrations, E_t is the trap energy level and E_i is the intrinsic Fermi level, Eq. 6 can be rewritten by substituting Eq. 7 and Eq. 8 as follows,

$$\frac{d\Delta n(t)}{dt} = \frac{-(A\Delta n(t) + B\Delta n^2(t))}{C + D\Delta n(t)} \tag{9}$$

with;

$$A = v_n v_p \delta_n \delta_p N_t (p_0 + n_0) \quad (10)$$
$$B = v_n v_p \delta_n \delta_p N_t \quad (11)$$
$$C = v_n \delta_n n_0 + v_n \delta_n n_i \exp[(E_t - E_i)/kT] + v_p \delta_p p_0 + v_p \delta_p n_i \exp[(E_i - E_t)/kT] \quad (12)$$
$$D = v_n \delta_n + v_p \delta_p \quad (13)$$

The effective lifetime can be determined by using Eq. 5 and Eq. 9 as follows,

$$\tau_{eff}(t) = \frac{C + D\Delta n(t)}{A + B\Delta n(t)} \quad (14)$$

For low injections, that is excess concentrations of less than 0.1% of the majority concentration, Eq. 14 can be approximated by C/A as follows,

$$\tau_{eff}(low) = \frac{v_n \delta_n n_0 + v_n \delta_n n_i \exp[(E_t - E_i)/kT] + v_p \delta_p p_0 + v_p \delta_p n_i \exp[(E_i - E_t)/kT]}{v_n v_p \delta_n \delta_p N_t (p_0 + n_0)} \quad (15)$$

For high injections, that is excess concentrations of more than 1000 times the majority concentration, Eq. 14 can be approximated by D/B as follows,

$$\tau_{eff}(high) = \frac{1}{v_p \delta_p N_t} + \frac{1}{v_n \delta_n N_t} \quad (16)$$

Thus, the lifetime changes for excess concentrations more than 0.1% and less than 1000 times the majority concentration. For this reason the lifetime in a high-ohmic wafer will be more sensitive to the light intensity than a low-ohmic wafer. In a high-ohmic wafer the majority concentration is low, therefore it is easy to generate an excess concentration of more than 0.1% of the majority concentration. A high-injection level for a high-ohmic wafer is low for a low-ohmic wafer.

5 - THE INFLUENCE OF RECOMBINATION AT THE SURFACE OF SILICON WAFERS

The general-differential equation, which describes the decay of the excess carriers at an intensity at which the lifetime has a constant value due to the bulk recombination and diffusion currents towards the surfaces is given by Eq. 17,

$$\frac{dn}{dt} = D\frac{d^2n}{dz^2} - \frac{n}{\tau_b} \quad (17)$$

where D is the diffusion coefficient. A solution for this equation for a light impulse is given by Eq. 18, /3/

$$n(z,t) = \exp(-t/\tau_b)\sum_0^\infty [A_m \exp(-\alpha_m^2 Dt)\cos(\alpha_m z) + B_m \exp(-\alpha_m^2 Dt)\sin(\alpha_m z)] \quad (18)$$

The average density over the whole thickness of the semiconductor as a function of time is given by Eq. 19,

$$n_i(t) = 1/d \int n_i(z,t)dz = \sum_0^\infty n_{0m} \exp[-(1/\tau_b + \alpha_m^2 D)t] \quad (19)$$

The response to the other pulse forms can be easily obtained by the convolution of the pulse form and this expression. The factor $\exp[-(1/\tau_b + \alpha_n^2 D)t]$ shows up in all the expressions and the term $1/\tau_b + \alpha_n^2 D$ can be defined as the total lifetime of the excess carriers. The term $\alpha_n^2 D$ is the contribution of the surface effects. The first term of the summation given by Eq. 19 is a good approximation for the recombination process. The boundary conditions due to the diffusion currents towards the surfaces and recombination with a velocity of S at the surfaces, give the following equation, from which the value of α_1 can be determined,

$$\tan(\alpha_n d/2) = S/D\alpha_n \quad (20)$$

The value of α_1 reaches the value of π/d for very high surface-recombination velocities. Thus the contribution of the surface effects can be mathematically approximated and subtracted from the results. However, it is experimentally verified that these surface effects do not play an important role in oxidized wafers and can be ignored.

6 - THE EXPERIMENTAL RESULTS

The results of the measurements, which are reported in this section, are all carried out with the measurement system described in section 2. In Fig. 2 two examples are shown. Fig. 2a shows the result of the measurement

on a low-ohmic wafer and Fig. 2b shows the result of the same measurement on a high-ohmic wafer. Both wafers were oxidized in order to reduce the surface recombination to a very low level. Both the original decay process and the logarithm of it, which is provided by an electronic circuit, are shown. The measurement is carried out at three different intensities, indicated by low, medium and high. The reciprocal value of the slope of the logarithm signal is proportional to the value of the lifetime, therefore a constant slope indicates a constant lifetime or an ideal-exponential decay process. It is clear from these figures that the lifetime has a constant value at low intensities and that it increases as the intensity increases. The change in the lifetime for low-ohmic wafers is less than that for high-ohmic wafers. These results are in agreement with the theoretical model, which has been discussed.

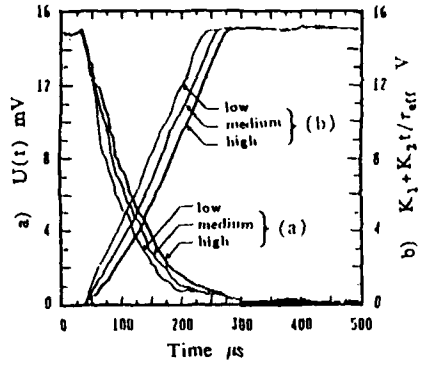


Fig. 2a. The result of the measurement carried out on an n-type silicon wafer with a resistivity of 10-15 Ω.cm. The measured lifetime at a low optical source intensity equals 48 μs.

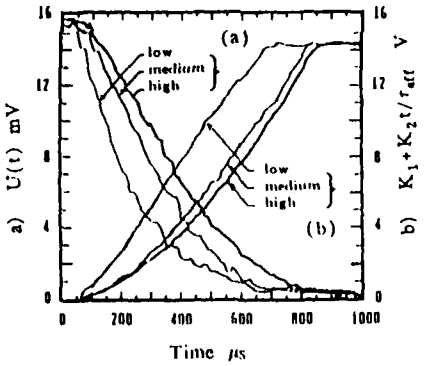


Fig. 2b. The result of the measurement carried out on an n-type silicon wafer with a resistivity of 700-1300 Ω.cm. The measured lifetime at a low optical source intensity equals 157 μs.

7. CONCLUSIONS

It has been shown that the bulk recombination lifetime of an oxidized silicon wafer can be easily measured, by using a low intensity. At higher intensities, the bulk lifetime changes from the value belonging to high injection levels to the normal bulk value.

REFERENCES

/1/ Y. Ogita, Jpn.J.Appl.Phys. 19 (1980), 2309.
/2/ M. Kunst, G. Beck, J.Appl.Phys. 60 (1986), 3558.
/3/ K.L. Luke, L.J. Cheng, J.Appl.Phys. 61 (1987), 2282.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

NUMERICAL SIMULATIONS TO IMPROVE THE ACCURACY OF ELECTRON-BEAM TESTING ON PASSIVATED INTEGRATED CIRCUITS

H. FREMONT, A. TOUBOUL*, D. GOBLED and Y. DANTO*

Texas Instruments-France, BP 05, F-06270 Villeneuve Loubet, France
**IXL, CNRS-UA 846, Université de Bordeaux I, F-33405 Talence Cedex, France*

RESUME : L'analyse sans contact au MEB des circuits intégrés par contraste de potentiel est non destructive. De plus, cette méthode permet par couplage capacitif l'évaluation du potentiel de pistes enfouies sous une couche isolante.
Pour augmenter la précision des mesures par cette technique, une simulation numérique en deux dimensions est utilisée. Nous avons en particulier étudié l'influence du potentiel des pistes voisines et de la grille d'extraction, ainsi que le rôle du positionnement du faisceau sur la précision des mesures.

ABSTRACT : The Scanning Electron Microscope associated with a Voltage Contrast system is a contactless and non-destructive tool for the test of VLSI Circuits. Moreover, the potential of passivated tracks can be evaluated by Capacitive Coupling Voltage Contrast (CCVC). This paper presents a numerical simulation applied to improve the accuracy of CCVC measurements. Particularly, the role of the location of the beam, and the influence of the potential of the neighbouring tracks and of the extraction grid is studied.

1 - INTRODUCTION

The increasing density of Integrated Circuits makes more and more necessary the development of new testing techniques for failure analysis and design validation. They must bring informations related to the internal behaviour of circuits and avoid contact problems or risks of mechanical destruction. The Electron Beam Voltage Contrast in the Scanning Electron Microscope (SEM) provides these features as it is contactless and non-destructive. Moreover, the use of the Capacitive Coupling Voltage Contrast (CCVC) /1/ makes possible the measurement of the potential of tracks buried in insulating layers. Because of the multiplication of metallisation levels and because of the circuits sensitivity to a chemical depassivation, this appears of increasing interest. However, experimental tests show that some difficulties have to be overcome in order to improve the accuracy of quantitative measurements. The two main problems are the oxide charging under electron irradiation and the voltage spreading.

Oxide charging mechanisms have been detailed elsewhere /2/. The purpose of this work is to provide a better understanding of the voltage spreading phenomenon, with the aid of 2-D numerical simulations.

2 - THE NUMERICAL APPROACH

The voltage spreading has not to be mistaken for "Local Field Effects". Following Kohler /3/, we think that the definition, as well as the effects of the voltage spreading are different. In the "image mode" (when the beam scans the IC surface), the voltage spreading creates a potential contrast between two metallic tracks. In the "waveform mode" (when the beam is focused on a specific point), a signal can be measured in the insulator between these tracks (see photo 1). Furthermore, the extraction grid located at 6 mm over the IC modifies the potential of the surface.

To solve this problem, we use a two-dimensional numerical simulation of the potential distribution in test layouts which are similar to existing configurations of IC's interconnections. The extraction grid has also been taken into account. Figures 1 and 2 depict the studied structures. The differential partial equation which describes the two-dimensional potential $V(x,y)$ distribution in the structures is a Laplace's equation : $\Delta V(x,y) = 0$. The potential of the different grids and the continuity of the electric induction at the oxide-vacuum interface represent the boundary conditions. A finite-difference technique has been used. Hence the structure is divided up into a number

of discrete meshes. Along the horizontal axis, the spatial discretisation fits the diameter of the electron-beam. The calculation is iterative, and is stopped as soon as a desired precision is obtained. Typical results of these simulations are depicted in figures 1 and 2.

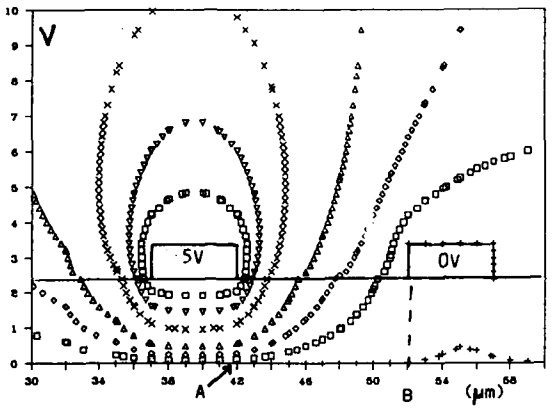


Fig. 1 : 2-grid layout simulating the configuration of a circuit under test
□0.2V, ◇0.5V, △1V, x 2V, ▽3V, □4V
Extraction grid potential Ve=0V

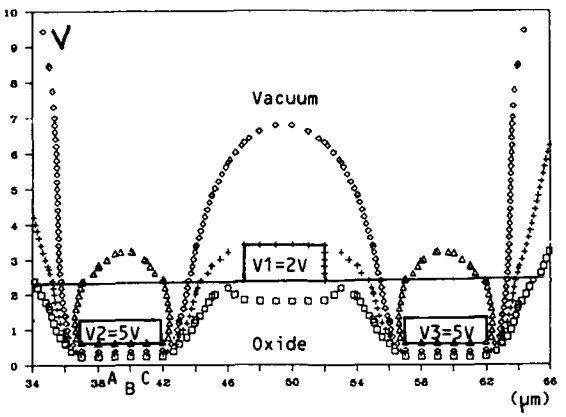


Fig. 2 : Typical dimensions and equi-potential lines of a 3-grid layout
□ 1.5 V, + 2 V, ◇ 2.5 V, △ 4 V
Extraction grid potential Ve=600V

From the potential maps, the interface potential can be deduced. In the absence of any buried conductor (see fig. 1) there is a non-zero interface potential between the two superficial tracks. This potential can be computed and experimentally measured by the contrast potential technique.

Figure 3 shows a fair agreement between the calculated and the measured values. The slight discrepancies which can be noticed stem from the noise superimposed to the experimental data.

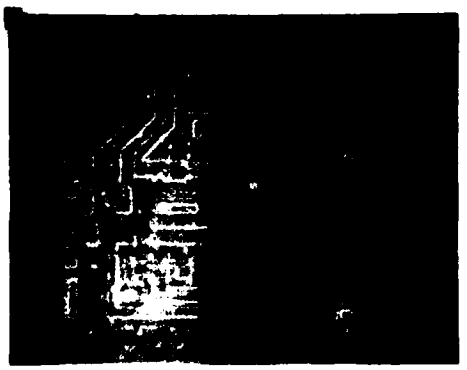


Photo 1 : Experimental evidence of the voltage spreading in the "image mode"

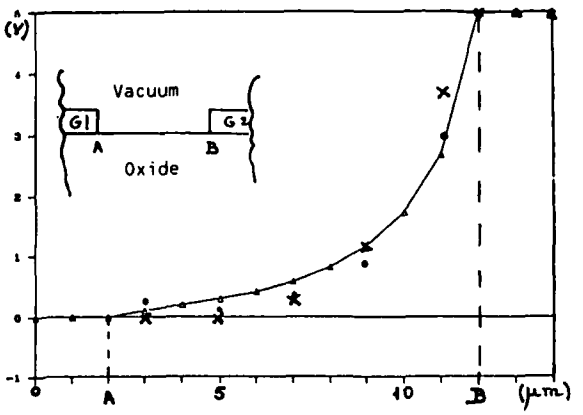


Fig. 3 : Surface potential variations for V1 = 0 V, V2 varying from 0 to 5 V
Δ: computed values
X : 2 MHz • 600KHz : experimental values

When the potential of a buried track is measured with the aid of the CCVC technique (see fig. 2) the non-zero potential created at the interface by the neighbouring tracks modifies the surface potential, which causes measurement errors. Both the CCVC experimental data and the numerical simulations give an information on the interface potential, as a function of the buried, the superficial, and the extraction grids potentials. But the capacitive coupling between the buried track and the virtual electrode created by the electron beam has not been taken into account for the potential calculation. However, it will be shown that this numerical approach can still be used to improve the measurements.

3 - ANALYSIS OF THE RESULTS - APPLICATION TO THE EXPERIMENTAL PROCEDURE

This approach has been applied to the structure presented on figure 2, in order to evaluate the influence of V1, and of the extraction grid polarisation Ve on the interface potential Vs on the top of the buried electrode. The calculated values of Vs depend linearly on V1 (see fig 4). To perfom a measurement, the beam must be positionned at the desired point, but the beam shifts during a measurement. These faults in the beam location create measurement errors as shown by the calculations.

Figure 4 presents the variations of Vs as a function of V1 for 3 different points over the buried track #2. The values, as well as the slopes are different from one point to another. Potential measurements by the CCVC technique give only access to potential variations : it is possible to detect ΔVS as a function of $\Delta V1$. Different slopes in the straight lines mean different $\Delta VS / \Delta V1$. Hence the positionning of the beam must be as precise as possible when the error due to a $\Delta V1$ on the measured value has to be calculated. The beam shift during a measurement also introduces errors, which can then be corrected with the help of the numerical solution.

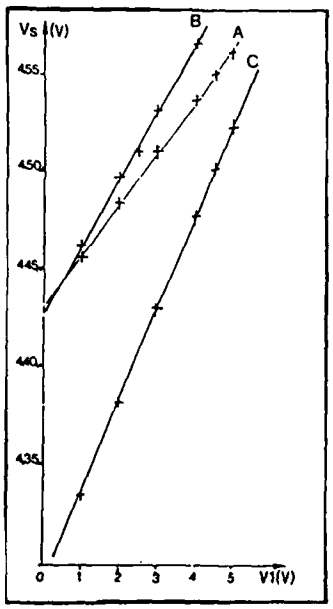


Fig. 4 : Computed values of Vs vs.V1 (see fig.2) at the the points A,B,C

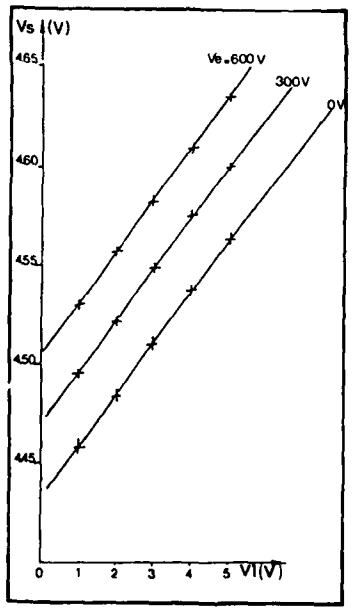


Fig. 5 : Computed values of Vs vs. V1 for different values of Ve

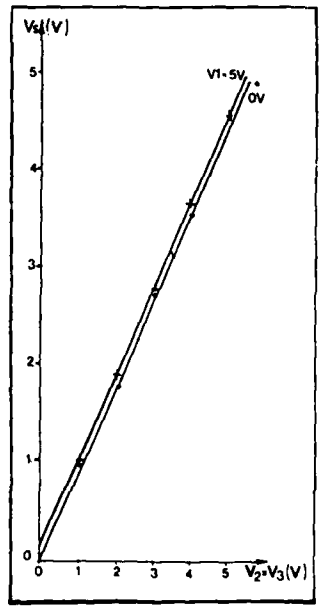


Fig. 6 : Computed values of Vs vs. V2 and V3 (buried at tracks) for two values of V1

In order to get a good voltage resolution, Ve has to be adjusted before every measurement by plotting the secondary emitted current vs. the potential of the filter grid ("S-Curve"). The calculations show that the setting of Ve to different values does not modify the slope of the plot (see Fig.5). Hence, the choice of Ve does not act on the measured potential variations. However, the problem of the oxide charging by the extraction grid /4/ has not been taken into account in the numerical simulation.

Figure 6 gives the variations of the surface potential as a function of the potential as a function of the buried tracks, for the two extremal values of V1 (0-5 V). The constant value of the slopes of the two curves proves that the variations of the potential of the buried tracks, have no influence on the potential surface variations, created by the neighbouring tracks potential V1.

In fact, the numerical solution cannot give access to the potential actually measured, because of the capacitive coupling existing, between the buried track and the virtual electrode created at the interface by the electron beam, which could not be simulated. But the results shown on fig.6 prove that the influence of neighbouring and measured tracks are uncoupled. That is to say that the numerical solution can be used to determine the influence of the neighbouring potential on the measured value, and hence to yield the actual potential of the buried track.

4 - CONCLUSION

A 2-D numercial potential simulation has been presented. Its application to improve the accuracy of electron beam testing of passivated IC's has been studied. We show that the position and the shift of the electron beam can widely create measurement errors, and that the potential of the extraction grid has no influence on the potential variation which can be measured by the CCVC technique. Though this model does not simulate the specific CCVC effect, the numerical results suggest that it can be used to correct the measured value and contribute to determine the actual potential of a buried track.

REFERENCES

/1/ GOERLICH S. - Duisburg University Thesis - VSI Reihe 9 Nr 61 (1986)
/2/ FREMONT H., TOUBOUL A., GOBLED D., DANTO Y.
Proc. Colloque International "Reliability and Maintainability"
Strasbourg - (Oct. 1988) - to be published.
/3/ KOHLER R. : Private communication
/4/ NYE P., DINNIS A. - Scanning 7, pp 117-124 (1985)

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

ACCELERATION 1/F NOISE IN SILICON MOSFETs

A.N. BIRBAS, Q. PENG, A. VAN DER ZIEL and A.D. VAN RHEENEN

Electrical Engineering Department, University of Minnesota, Minneapolis, MN 55455, U.S.A.

ABSTRACT - It is usually assumed that the 1/f noise in Si-MOSFETs is limited by collision 1/f noise. We found this to be the case for devices with relatively short channel lengths (L<10μm) but for channels of intermediate length (10μm<L<194μm) we found that the Hooge parameter varies as L^2. We attributed this to acceleration of the electrons by the applied field, accompanied by Bremsstrahlung emission and current 1/f noise generation. This is a new noise source.

INTRODUCTION

In short low noise Si-MOSFETs the magnitude of the 1/f noise is set by collisions processes /1/. Each collision process (normal, Umklapp, intervalley, etc.) has its own characteristic value of the Hooge parameter α_H and can be identified by it. The contributions of the various collisions must be added quadratically.

In long devices (L>10 μm) a new 1/f noise source has been identified in which the Hooge parameter varies as the square of the device length /2/, /3/. It is attributed to acceleration of the electrons by the applied electric field, accompanied by Bremsstrahlung emission and current 1/f noise generation. Since the process is in operation as long as the electron is travelling from source to drain, the contribution must be added linearly. This leads to an L^2- dependence; it represents a new noise source.

ACCELERATION 1/f NOISE

The general expression for the Bremsstrahlung power emitted by a single electron (c.g.s. units) is

$$P(t) = \frac{2e^2}{3c^3} a(t)^2 \quad \text{for} \quad 0<t<\tau \qquad (1)$$

where τ is the duration of a single radiation pulse and a(t) is the acceleration of the electron. We can obtain a linear pulse by writing:

$$\sqrt{P(t)} = \left(\frac{2e^2}{3c^3}\right)^{1/2} a(t) \qquad (2)$$

and now we make a Fourier transform

$$\int_0^\tau \sqrt{P(t)}\ \exp(-j\omega t)\,dt = \frac{2e^2}{3c^3}\,F(j\omega) \qquad (3)$$

where

$$F(j\omega) = \int_0^\tau a(t)\ dt = F(0)\ \frac{\sin(\omega\tau/2)}{\omega\tau/2}\exp(-j\omega t/2) \qquad (4)$$

For frequencies ω much smaller than the reciprocal of τ , $F(j\omega) = F(0)$. Applying Carson's theorem we find for the energy spectrum of the Bremsstrahlung

$$S_{\sqrt{P}}(f) = 2\left(\frac{2e^2}{3c^3}\right) F^2(0)\,\lambda \quad \text{in erg} \qquad (5)$$

where λ is the number of power pulses per second. Dividing by the quantum energy hf we can obtain the number spectrum of the photons and furthermore dividing by the duration of the pulse τ we have the spectum Sq(f) of the rate of the photon emission.
Hence

$$S_q(f) = \frac{4e^2}{3c^3}\,\frac{F^2(0)\,\lambda}{h\,f\,\tau} \qquad (6)$$

with an 1/f type spectrum. We now define the Hooge parameter from the Hooge equation by writing:

$$S_I(f) = \frac{\alpha_H I^2}{f\,N} = \frac{\alpha_H I}{f\,N}\,\frac{N e}{\tau} = \frac{\alpha_H e\,I}{f\,\tau} \qquad (7)$$

where $N=I\tau/e$ and τ the transit time. Since $\lambda=N/\tau$ we can write equation (6) as:

$$S_q(f) = \frac{4e^2}{3c^3}\,\frac{F^2(0)\,N}{h\,f\,\tau^2} = \frac{\alpha_H N}{2\,f\,\tau^2} \qquad (8)$$

so that :

$$\alpha_H = \frac{8\,e^2\,2\,\pi}{3\,h\,c}\,\frac{F^2(0)}{c^2} = \frac{4\,a_0}{3\,\pi}\,\frac{F^2(0)}{c^2}$$

using

$$a_0 = \frac{2\pi\,e^2}{h\,c} = \frac{1}{137} \quad \text{(the fine structure constant)}$$

We now descriminate between two processes:
a) noise produced by different scattering events. Here $m^*\,a(t) = \hbar\,dk/dt$ for a single collision process and:

$$F^2(0) = \left(\int_0^\tau a(t)\,dt\right)^2 \frac{\left(\Delta\left(\frac{h}{2\pi}k\right)\right)^2}{m^{*2}} \qquad (9)$$

where $\Delta(\hbar k)$ is the change in momentum during an individual collision process and so:

$$\alpha_H = \frac{4\,a_0}{3\,\pi}\,\frac{\left[\Delta\ (\hbar k)\right]^2}{m^*\,c^2}$$

b) acceleration type of noise. We consider a FET operating in the linear regime ($V_d \ll V_{ds}$). Then the electric field is constant and the acceleration $a(t) = eE/m^*$ is constant too. Hence:

$$F(0) = \int_0^\tau \frac{e\,E}{m^*}\,dt = \frac{e\,E}{m^*}\,\tau = \frac{e\,E\,\tau}{m^*}\,\frac{m^*\,\mu}{e\,\tau_0}$$

where t_0 is the intercollision time.
Since $\mu=e\tau_0/m^*$, $m^*\mu/et_0=1$

$$F(0) = \frac{\tau\,\mu E}{\tau_0} = \frac{\tau}{\tau_0}\,u_d = \frac{\tau}{\tau_0}\,\frac{L}{\tau} = \frac{L}{\tau_0}$$

where L is the effective channel length of the device.
Furthermore for semiconductors with a single effective mass $\tau_0=m^*\mu/e$ where τ_0 is the time

constant of the collision process (intercollision time). For p-channel Si MOSFETs there are three different effective masses $m1^*$,$m2^*$,$m3^*$, a more complicated expression for the time constant might be:

$$\frac{1}{\tau_0} = \sum_i \frac{g(\tau_i)}{\tau_i}$$

where $g(\tau_i)$ is the probability that the hole has the mass m_i^*.

We thus have for the Hooge parameter:

$$\alpha_H = \frac{4\,a_0}{3\,\pi} \left(\frac{L}{c\,\tau_0}\right)^2 \qquad (10)$$

The dimension of L/τ_0 is velocity and increases indefinetely with increasing L. At first it was thought that Δv could not become larger than c so that α_H might reach Handel's limit $2a_0/\pi$ by relativistic saturation. But as it is pointed out L/τ_0 is not a true velocity.

The above effect remains correct even for non linear operation. To show that we split the length L into sections Δx, having a length comparable to the free path length of the carriers and then :

$$\alpha_H = \frac{4\,a_0}{3\,\pi} \left(\frac{L}{c} \left\langle \frac{1}{\tau_0(E)} \right\rangle\right)^2 \qquad (11)$$

where < > denotes averaging.

EXPERIMENTAL EVIDENCE FOR THE ACCELERATION NOISE

In Figure 1 the variation of the Hooge parameter with the effective channel length of p-type MOSFETs made on the same chip is shown. The α_H varies as L^2 for intermediate channel lengths. The saturation of the curve is observed as we approach the smaller channel lengths (~ 2μm). There the collision generated 1/f noise dominates and no channel length dependence is expected for the Hooge parameter. Furthermore for long channel lengths no saturation (relativistic) effects are detected and so no indication exists for the approach of Handel's limit $2a_0/\pi$ even though the measured α_H for L=194 μm is approaching this value. Another question which may be raised is the exact extraction of the experimental Hooge parameter /4/ from the measured current spectrum if the shape of the spectrum is not exact 1/f. Even though the experimental data presented here are coming from current noise spectrums with $1/f^\gamma$ shape with $1<\gamma<1.1$ it can be shown that α_H has always meaning even if the noise spectrum is not exact 1/f (A. van der Ziel and A.D. van Rheenen, to be published).

In Figure 2 the experimentally obtained time constant is plotted as a function of the channel length and it is found to be constant and of the order of 10^{-12} sec. This value is resonable for the intercollision time of p-type silicon which is an evidence of the validity of the above approach.

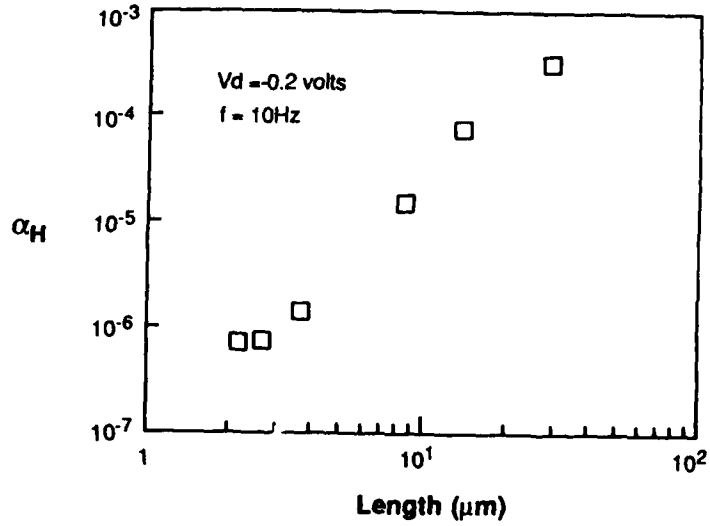


Fig.1- The Hooge parameter as a function of the channel length for p-type MOSFET's

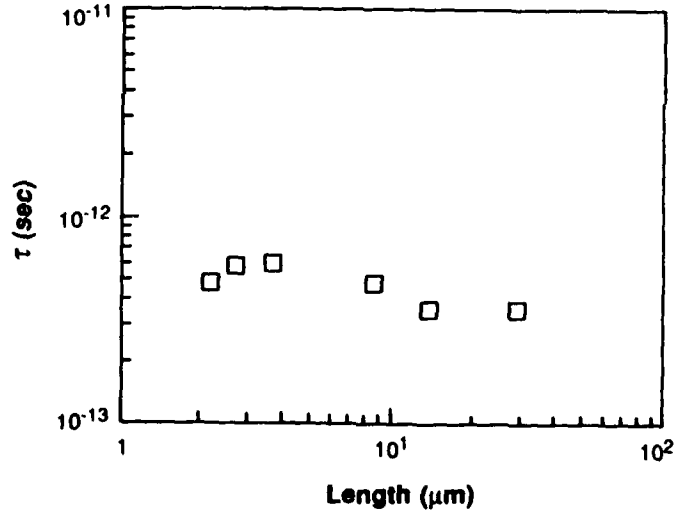


Fig. 2-The experimentally obtained effective time constant for the p-type MOSFET's

REFERENCES
/1/ A. van der Ziel, Proceedings of the IEEE, vol. 76, No. 3,(1988) 233
/2/ A.N. Birbas, Q. Peng, A.van der Ziel, A. van Rheenen and K.Amberiadis,accepted for publication, Journal of Applied physics, 65 , July 1988
(1988)
/3/ A. van der Ziel, Journal of applied Physics, accepted for publication, July 1988.
/4/ F.N. Hooge, Phys. Lett., A-29,(1969) 139.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

SPATIAL DISTRIBUTION OF 1/f NOISE SOURCE

L.K.J. VANDAMME

Faculty of Electrical Engineering, Eindhoven University of Technology, NL-5600 M.B. Eindhoven, The Netherlands

Résumé - Le bruit en 1/f a été étudié dans des résistances silicium type n dans des conditions d'accumulation et de déplétion. Les résultats sont interprétés en termes de phénomènes de volume et ils sont caractérisés par le paramètre α de Hooge avec des valeurs comprises entre 10^{-7} et 10^{-5}. La valeur de α pour la conduction en surface en régime d'accumulation peut être au moins un ordre de grandeur plus grande que la valeur pour une conduction loin de la surface.

Abstract - The 1/f noise of an n-type silicon resistor has been studied under conditions ranging from accumulation to depletion at 300K. The experimental results are interpreted in terms of a bulk phenomenon and are characterized by Hooge's empirical 1/f noise parameter α with values between 10^{-7} and 10^{-5}. The α-value for surface conduction at strong accumulation can be at least one order of magnitude larger than the value for bulk conduction.

1 - INTRODUCTION

There is a broad experimental evidence for the bulk origin of 1/f noise in metals and semiconductors /1/. The experimental results are then described by Hooge's empirical relation $S_V/V^2 = \alpha/fN$, where α is the 1/f noise parameter and N is the total number of carriers in the homogeneous sample subjected to homogeneous fields. As some experimental results point to a strong surface influence on 1/f noise, the controversy over the bulk or surface origin of 1/f noise has a long history.
Invoking the role of the interface became very popular in describing the 1/f noise in MOSTs. Experimental evidence has been put forward in favour of the Mc Whorter model /2-5/. The proportionality between 1/f noise and oxide-trap density in MOSTs was clearly taken as illustrating the interface or surface origin of 1/f noise /5/. Vandamme /6/ proposed a bulk model to explain the 1/f noise in MOSTs based on mobility fluctuations. However, some MOSTs are straightforward to interpret in terms of number fluctuations /5/.

Recent experiments indicate that α-values are associated with the number of crystal defects in the bulk /7/. Here we investigate the spatial distribution of the 1/f noise source regarding the influence of a gate voltage on the 1/f noise of an n-type resistor surrounded by a p-well. By depleting the Si - $Si O_2$ interface region with a negative gate voltage the bulk 1/f noise source is observed and characterized by α_b. Under a positive gate voltage accumulation is reached and the 1/f noise source close to the Si - $Si O_2$ interface is characterized by α_s, which can be greater than α_b, especially when α_b is as low as 10^{-7}.

2 - SAMPLE DESCRIPTION

The samples have been manufactured by Landis and Gyr (Switzerland) to serve as vertical Hall sensors. The geometry is presented in cross-section view in Figure 1 and the dimensions and parameters are summarized in Table 1. The inner dimension of the p-well and hence length and widths of the n-type caisson are ℓ = 150μm and t = 6μm. The contacts D_1, D_2, D_3, Q_1 and Q_2 all have a width $2x_\ell$ = 9μm. The current-carrying contacts are D_2 = D_3 and D_1, the sensor contacts are Q_1 and Q_2. With the MOS structure on top of the caisson between the contact regions the region close to the interface can be brought into accumulation at $V_G > 0$ or into depletion at $V_G < 0$. The current I is passed through D_1 and D_2, which is connected with D_3 and are used as a reference electrode in the sample. Two and four-probe tests have been carried out on this sample, which is important in estimating contact contributions. The

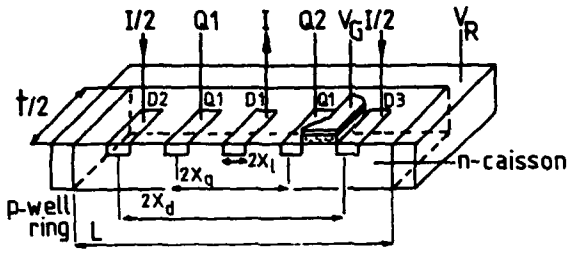


Table 1 Dimensions and semiconductor parameters

contact width $2x_\ell$(µm)	sensor distance $2x_q$(µm)	driver distance $2x_d$(µm)	caisson width t(µm)	donor concentration $n(cm^{-3})$
9	36	140	t = 12	10^{15}

Fig. 1 - Cross-section of the device. The drivers (current carrying contacts) are indicated by D_1, D_2 and D_3 and the sensors by Q_1 and Q_2. The values of distances and contact width are presented in Table 1.

so-called 2-probe noise test is obtained by passing a constant current between D_1 and $D_2 = D_3$ and observing the voltage fluctuations at the same contact pair. The 4-probe noise test, which enables us to suppress possible noise contributions of the contact interfaces at D_1, D_2 and D_3 measures the voltage fluctuations across the sensor contacts Q_1 and Q_2 while the current is passed through D_1 and $D_2 = D_3$.

3 - CALCULATED AND OBSERVED RESISTANCE

The resistance between the central contact D_1 and the connected outer electrodes $D_2 = D_3$ consists of two contributions in parallel : (i) a bulk resistance always present (ii) a surface resistance. At $V_G > 0$ the surface resistance will be the lowest and thus dominant. At $V_G < 0$ the current density close to the Si - $Si O_2$ interface can be neglected and the bulk resistance becomes dominant. To calculate the bulk contribution, we assume hemicylindrical equipotentials in the n-caisson all coaxial with the middle line of contact D_1. From the numerical calculations of equipotentials we see that hemicylinders are a good approximation. The resistance between two equipotentials with radii x and x + dx becomes $dR_b = \rho_b\, dx/\pi \times t$. Integrating from $x = x_\ell$ to $x = x_d$ results in

$$R_b = \int_{x_\ell}^{x_d} (\rho_b / \pi \times t)\, dx = (\rho_b / \pi\, t)\, \ell n\, x_d/x_\ell \qquad (1)$$

The contribution of a surface layer with resistivity ρ_s, thickness δ and width t is given by

$$R_s = \rho_s\, (x_d - 2\, x_\ell) / 2\, t\, \delta \qquad (2)$$

Fig. 2 shows the dependence of the resistance between D_1 and $D_2 = D_3$ as a function of the gate voltage. The observed resistance values between D_1 and $D_2 = D_3$ and between Q_1 and Q_2 are equal at $V_G > 7V$ because $2x_q \approx x_d/2$. The calculated value of R_b with eq. (1) and the values presented in Table 1 equals 3.6kΩ. The experimentally observed value was 3.8 kΩ at $V_G = 0V$. This means that the influence of the contacts are slightly under estimated.

4 - CALCULATED AND OBSERVED NOISE

The noise contribution of the bulk is calculated by assuming hemicylindrical equipotentials around the central electrode D_1. Hooge's empirical relation is applied on shell resistances of thickness dx. By integration of all contributions between x_ℓ and x_d we obtain for the two-probe noise a spectral density in the resistance of

$$S_{R_b} = \frac{\alpha_b\, \rho_b^2}{2\pi^3\, n\, t^3\, f} \left(\frac{1}{x_\ell^2} - \frac{1}{x_d^2} \right) \qquad (3)$$

Considering $1/x_d^2 \ll 1/x_\ell^2$ and applying eq. (1) we find for the normalized noise spectral density in R_b

$$\frac{S_{R_b}}{R_b^2} = \frac{\alpha_s}{2\pi (\ell n\, x_d/x_\ell)^2\, t\, x_\ell^2\, n\, f} \qquad (4)$$

The relative spectral density in the surface resistance R_s follows from Hooge's expression

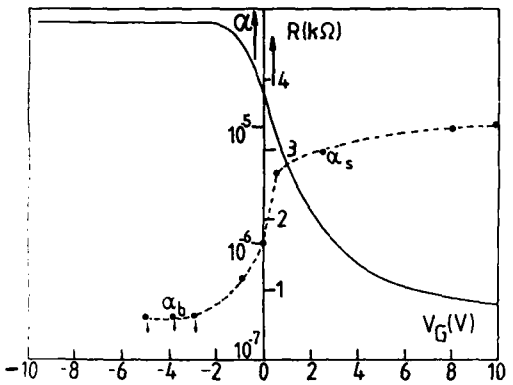


Fig. 2 - The resistance between D_1 and $D_2 = D_3$ versus the gate voltage at $V_{NP} = 0$ (solid line). Dotted line, the bulk α_b and surface α_s values of the 1/f noise parameter.

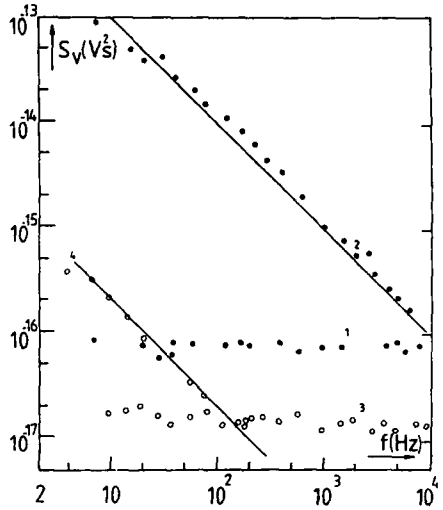


Fig. 3 - Spectra S_{VQ} between sensors at : (1) I = 0 (2) I = 0.41 mA at V_G = 0V and (3) I = 0 (4) I = 0.44 mA at V_G = 10.8 V.

$$\frac{S_{R_s}}{R_s^2} = \frac{\alpha_s}{2\,\delta t(x_d - 2x_\ell)\, n_s\, f} \qquad (5)$$

When the resistance and the noise are dominated by the bulk contribution, the 1/f noise parameter α_b can be calculated from the experimentally observed $f\,S_v/V^2$ value by using eqs(4) and (1) with the elimination of t.

$$\alpha_b = \left(\frac{f\,S_v}{R\,V^2}\right) \frac{2\,x_\ell^2\; \ell n^3(x_d/x_\ell)}{q\;\mu_b} \qquad (6)$$

The value of the 1/f noise parameter close to the interface α_s is calculated from eqs(2) and (5) and the experimentally observed $f\,S_v/V^2$ at $V_G > 0$, giving

$$\alpha_s = \left(\frac{f\,S_v}{R\,V^2}\right) \frac{(x_d - 2x_\ell)^2}{q\mu} \qquad (7)$$

To calculate the noise S_{VQ} at the sensors Q_1, Q_2 we use the general expressions for 4-probe arrangement /8/. Assuming only bulk contributions the ratio between sensor and driver noise $S_{VQ} / S_{VD} = 0.3$ for the dimensions presented in Table 1.

At strong accumulation $S_{VQ} / S_{VD} = 4x_q/x_d$. With the values in Table 1 we find $S_{VQ}/S_{VD} \simeq 1$. Whether or not it's a bulk or surface condition can be seen from the ratio experimentally observed S_{VQ}/S_{VD}.

Fig. 3 shows some spectra of the voltage fluctuations between the sensors Q1 and Q_2 at V_{PN}=0 for V_G= 0V and V_G= 10.8V at I=0 (thermal noise) and I = 0.42 mA (1/f noise). When reverse bias between p-well and n-type caisson is applied, the 1/f noise often is contamined with generation-recombination (G-R) noise with a characteristic frequency of 200 Hz. The observed ratio $S_{VQ}/S_{VD} \simeq 0.1$ for the G-R and 0.3 for the 1/f part of the noise spectrum which suggests that G-R contribution stems neither from the bulk nor from the Si - SiO_2 interface of the resistance, while the 1/f part are bulk contributions. Samples with G-R noise also show large leakage currents between p-well and n-type caisson of the order of 100μA per volt reverse bias. Therefore we suggest that the g-r component stems perhaps from the rim of the p-well around the caisson and the surface of the sample. The experimentally observed S_{VD} and S_{VQ} are proportional to I^2 as expected when I < 1mA. At V_G = 0, S_{VQ}/S_{VD} = 0.3 and at V_G = 10.8V, S_{VQ}/S_{VD} = 1 which is in agreement with bulk and surface noise respectively.

Concerning the noise results at $V_{PN} = 0$ and $V_G > 0$ we find : (i) the thermal noise at I = 0 is proportional to $V_G^{-\gamma}$ with $0.5 < \gamma < 0.75$, (ii) the 1/f noise is proportional to $V_G^{-\delta}$ with $2 < \delta < 3$. The value $0.5 < \gamma < 1$ cab be explained by $N \propto V_G$ and a reduction in surface mobility at increasing V_G. The value of δ can be explained by the empirical relation together with a mobility reduction and $N \propto V_G$.

Applying a reverse bias between p-well and n-type caisson reduces the effective resistance

width and $R \propto 1/t$. The 1/noise S_{V_D} and hence S_{V_Q} in terms of Hooge's empirical relation should be $S_V \propto V^2/N \propto R^2/t \propto R^3$. The observed proportionality is in agreement with the expected proportionality for bulk 1/f noise.

The experimentally obtained 1/f noise values α are summarized in Fig. 2. Most samples show α_b calculated from eq. (6) of at most 2.5×10^{-7} for $V_G < 0$. At $V_G > 0$, surface conduction becomes dominant due to accumulation and α_s values of 10^{-5} have been calculated from experimentally obtained fS_V/RV^2 and eq.(7). The spatial distribution of the 1/f noise source does not seem to be uniform over the whole sample.

5 - DISCUSSION AND CONCLUSIONS

By applying positive or negative gate voltages we have changed the resistance and the conduction path from the surface to the bulk. The ratio between the 1/f noise at the sensors and the drivers has changed from 0.3 typical for bulk conduction to the value 1, typical for surface conduction. Conduction noise S_V in general is proportional to $\alpha \int_\Omega J^4 \, d\Omega$ /8/ with J the current density and Ω the sample volume. The experimentally observed trends and the dependence of S_V on J^4 are arguments interpreting the noise results, by either eq(6) in terms of α_b or by eq.(7) in terms of α_s. The numerical factor in eqs (7) for our samples is 4.5 times as large as that factor in eq.(6). All observed α_b values (eq.6) at $V_G = 0$ appear to be of the order of 10^{-6} with trends towards $\alpha_s = 10^{-5}$ (from eq.7) at $V_G \gg 0$ and towards $\alpha_b \approx 2 \times 10^{-7}$ (from eq(6)) for $V_G \ll 0$. Hence, the spatial distribution of the 1/f noise source is not uniform over the whole sample. Similar results to ours obtained on quite different samples were observed by Clevers /9/. Carruthers and Mavor/10/ have investigated the noise dependence on bias conditions of burried n-channel MOSTs. From their results we have calculated α-values of approximately 10^{-5} at $V_G > 0$ and 3×10^{-7} for $V_G < 0$. For positive gate voltages, the conducting channel forms such that it is close to or touching the silicon/silicon dioxide interface.

At $V_G < 0$ or $V_{PN} < 0$ the 1/f noise contribution in spectra is often in competition with G-R noise. In this case only upper values of α_b could be calculated as indicated by the arrows in Fig. 2. From the different ratios S_Q/S_D for the G-R and the 1/f part of the spectrum we conclude that both noise sources stem from different regions in the sample, viz the junction between the p-well and n-caisson and the n-caisson.

In view of the dependence of α on the quality of the crystal lattice after different anneal treatments /7/, we suggest, without invoking a trapping model, that the 1/f noise source close to the interface can be higher due to higher crystal defect density. Experimental findings on the relation between oxidation and 1/f noise can be understood in terms of Hooge's empirical bulk relation for 1/f noise. The 1/f noise can decrease owing to oxidation of a defect rich region with a high α-value that is removed from the conduction path. The relation between surface states (oxide trap density) and 1/f noise in MOSTs /3-5/ often found experimentally, should be understood in terms of bulk 1/f noise as due to mobility fluctuations, if an increased number of surface states goes hand-in-hand with an increased defect density in the region just below the SiO_2. Then the results for MOSTs that are easily explained by the Mc Whorter model can also be explained by the empirical relation with different α_s values.

ACKNOWLEDGEMENT

The author would like to thank Dr R.S. Popovic from Landis and Gyr, Zug, Switzeland for providing the samples.

REFERENCES

/1/ Hooge, F.N., Kleinpenning, T.G.M. and Vandamme, L.K.J., Rep. Prog. Phys. 44 (1981) 479.
/2/ Mc Whorter, A.L., M.I.T. Lincoln Lab. Rep. vol 80 1955.
/3/ Klaassen, F.M., IEEE Trans. Electron Devices ED-18 (1971) 887.
/4/ Maes, H.E. and Usmani, S.H., J. Appl. Phys. 54 (1983) 1937.
/5/ Vandamme, L.K.J. and Penning de Vries, R.G.M., Solid-State Electronics 28 (1985) 1049.
/6/ Vandamme, L.K.J., Solid-State Electronics 23 (1980) 317.
/7/ Vandamme, L.K.J. and Oosterhoff, S., J. Appl. Phys. 59 (1986) 3169.
/8/ Vandamme, L.K.J. and van Bokhoven, W.M.G., Appl. Phys. 14 (1977) 205.
/9/ Clevers, R.H.M., Physica 147B (1987) 305.
/10/ Carruthers, C. and Mavor, J., Electronics Lett. 23 (1987) 1173.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

ON THE ORIGIN OF 1/F NOISE IN MOS TRANSISTORS

E.J.P. MAY

University of Exeter, School of Engineering, North Park Road, Exeter, BG-Devon EX4 4QF, Great-Britain

Resumé - Il est proposé que le bruit 1/f est dû aux états rapides de surface qui causent la redistribution des transporteurs par diffusion. C'est pour quoi la densité du transporteur est modulée. It est demontré que la racine carrée du module de la transformation Fourier de la densité est inversement proportionelle à la fréquence.

Abstract - It is proposed that 1/f noise is due to 'fast surface states' which give rise to a redistribution of carriers by diffusion. This leads to the modulation of the carrier density. The square of the modulus of the Fourier transform of the carrier density time function is shown to be inversely proportional to the frequency.

1 - INTRODUCTION

The origin of 1/f noise has puzzled workers for several decades. It is now generally recognised that 1/f noise is due to fluctuations in conductance described empirically by the expression /1/

$$< \frac{\Delta\sigma}{\sigma}^2 > = C\frac{\Delta f}{f} \qquad \ldots\ldots\ldots(1)$$

where C is a constant with a wide range of values.

The noise is strongly associated with fast surface states /2/ and yet there is evidence that it occurs in the bulk /3/. It has been observed in a wide range of materials and devices; any theory purporting to explain the mechanism of 1/f noise must therefore have its basis in universal phenomena.

In this model two phenomena are envoked: the 'fast surface states' and the resulting diffusion of carriers in the bulk. The 'fast surface states' give rise to charge impulses which diffuse into the bulk and produce a time varying carrier density. In the mathematical analysis which follows, it will be shown that the square of the modulus of the Fourier transform of the time varying component of the carrier density is inversely proportional to the frequency.

2 - ANALYSIS

The analysis is based on the continuity equation:

$$.J = -\frac{\partial\rho}{\partial t} \qquad \ldots\ldots\ldots(2)$$

where J is the diffusion current density and ρ the charge density. If one considers an MOS transitor as a thin strip of n-type silicon having one main face (outside) with an even distribution of 'fast surface states' and if J is the diffusion current density in the x-direction perpendicular to the main faces, the carrier density, n, is given by:

$$\frac{\partial n}{\partial t} = D_n \frac{\partial^2 n}{\partial x^2} \qquad \ldots\ldots\ldots (3)$$

The solution n = n (t,x) for a unit impulse is

$$n = \frac{1}{\sqrt{(D_n t)}} \; e^{-(\frac{x^2}{4D_n t})} \qquad \ldots\ldots\ldots (4)$$

The square of the modulus of the Fourier transform of n (t,x) is given by:

$$|N(\omega, x)|^2 = \frac{1}{D_n} \; \frac{1}{\omega} \; e^{-\sqrt{(\frac{2}{D_n}\omega)}x} \qquad \ldots\ldots\ldots (5)$$

For $x < \sqrt{(\frac{D_n}{2\omega})}$

$$|N(\omega, x)^2 \doteq \frac{1}{D_n \omega} \qquad \ldots\ldots\ldots (6)$$

If there are S impulses per unit time of mean square value m per unit surface area

then $$|N(\omega, x)|^2 = \frac{Sm}{D_n \omega} \qquad \ldots\ldots\ldots (7)$$

If N_o is the total number of carriers

$$\langle (\frac{\Delta\sigma}{\sigma})^2 \rangle = \frac{Sm}{D_n N_o^2 \omega} \; \Delta\omega$$

$$= \frac{Sm}{D_n N_o^2 f} \; \Delta f \qquad \ldots\ldots\ldots (8)$$

For $Sm = K N_o$

$$\langle (\frac{\Delta\sigma}{\sigma})^2 \rangle = \frac{K}{D_n N_o f} \; \Delta f$$

$$= \frac{C\Delta f}{N_o f} \qquad \ldots\ldots\ldots (9)$$

3 - COMMENTS

It should be noted that the condition imposed on equation (5) to obtain the approximate relation (6) limits the thickness of the MOS transistor channel to 10μm in order to maintain the 1/f noise law to high frequencies.

Although this 'surface state diffusion' theory has only been applied to MOS transistors, a more general approach has been accepted for publication elsewhere /3/.

Direct experimental support for the 'surface state diffusion' theory is described in work on the 'generation of augmented 1/f and 1/Δf noise'/4/. It is proposed that the augmented 1/f noise is produced by stimulating the 'surface states' using an external stimulus.

The main conclusion is that this model is in accord with the McWhorter-van der Ziel model /5/ insofar as it acknowledges the surface states as the origin of 1/f noise. It does not, however, require an oxide layer to produce convenient time constants; the carrier diffusion mechanism produces the 1/f fluctuations in conductance.

REFERENCES

/1/ Hooge, F.N., '1/f noise is no surface effect', Phys. Lett., 29A, pp. 139-140, 1969.

/2/ Sah, C.T. and Hielscher, F.H., 'Evidence of the surface origin of 1/f noise', Phys. Rev. Lett. Vol.17, pp. 956-958, 1966.

/3/ May, E.J.P., 'On the origin of 1/f noise' to be published in International J. of Electronics.

/4/ May, E.J.P. and Al-Charchafchi, S.H., 'The generation of augmented 1/f and 1/Δf noise in resistors by the injection of short duration pulses', J. IERE Vol.57, No.3, pp. 125-128, 1987.

/5/ van der Ziel, A., 'Proof of basic semiconductor flicker noise formula', Solid-St. Electron., 17, pp. 110-111, 1974.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

NOISE AND DIFFUSION IN p-TYPE SILICON

J.P. NOUGIER, A. MOATADID and J.C. VAISSIERE

Centre d'Electronique de Montpellier[1], Université des Sciences et Techniques du Languedoc, F-34060 Montpellier Cedex, France

RESUME: La théorie des fluctuations de la fonction d'occupation d'états, développée tres récemment, et la notion de temps de relaxation différentiel, permettent de donner ici l'expression du coefficient de diffusion transversal, qui est calculé numériquement dans si-p en fonction du champ électrique a 300 K entre 0 et 50 kV/cm.

ABSTRACT: The theory of the fluctuations of the state occupancy function, recently developped, together with the differential relaxation time, allow getting the transverse diffusion coefficient, which is computed in p-type Silicon, as a function of the electric field, at 300 K, from 0 to 50 kV/cm.

1. INTRODUCTION:

Except Monte Carlo simulations, which are much time consuming and thus very expensive, all the device modeling methods use transport equations, which require the knowledge of the diffusion coefficient D(E) as a function of the electric field E. D(E) is difficult to obtain. One way to get it is to study the diffusion noise, experimentally or theoretically. The purpose of this paper is to propose a new theoretical method for determining the diffusion coefficients, based on the differential relaxation time and on the fluctuation of the state occupancy function.

2. STATE OCCUPANCY FUNCTION:

In a previous paper [1], we settle a theory, derived from a series of nice papers from Gansevitch, Gurevitch and Katilius, they gathered in a review article [2]. In that paper, we studied the fluctuations of the State Occupancy Function $f(\mathbf{k},t)$ of hot carriers (labelled SOF in the following for convenience). $f(\mathbf{k},t)$ is a random function which takes the value 1 when the state $\mathbf{k}$ is occupied by a carrier, and the value 0 when the state $\mathbf{k}$ is empty. The SOF depends on the external electric field E, but for convenience we shall omit E, wherever it is possible. $f(\mathbf{k},t)$ fluctuates, due to the random motion of the carriers, which undergo scattering events. If N(t) is the total number of carriers inside the sample at time t, one has :

$$N(t) = \sum_{\mathbf{k}} f(\mathbf{k},t) \qquad (1)$$

The ensemble average $\langle f(\mathbf{k},t)\rangle$ of $f(\mathbf{k},t)$ is then the distribution function at time t, governed by the Boltzmann equation (labelled BE in the following). As $t\to\infty$, $\langle f(\mathbf{k},t)\rangle \to f_S(\mathbf{k},E)$, solution of the steady state BE in the uniform electric field E.

3. THE DIFFERENTIAL RELAXATION TIME:

Since we intend to study small fluctuations around the steady state, we are interested in small deviations of the distribution function $\langle f(\mathbf{k},t)\rangle$ around its steady state value $f_S(\mathbf{k},E)$. This can be done

[1] Laboratoire Associé au CNRS, UA 391.

This study was partially supported by GCIS (Groupement Circuits Intégrés Silicium), France.

using the differential relaxation time concept, introduced some years ago [3], defined as follows: When a small perturbation δE is applied, the new steady state distribution is $f_S(\mathbf{k},E+\delta E)$. When the perturbation δE is removed, $\langle f(\mathbf{k},t)\rangle$ evolves from $f_S(\mathbf{k},E+\delta E)$ to $f_S(\mathbf{k},E)$. The differential relaxation time $\tau(\mathbf{k},E)$ is then defined as:

$$\lim_{|\delta E|\to\infty}\left\{\left(\frac{\partial\langle f(\mathbf{k},t)\rangle}{\partial t}\right)_{/t=0}\right\} = -\lim_{|\delta E|\to\infty}\left\{\frac{f_S(\mathbf{k},E+\delta E) - f_S(\mathbf{k},E)}{\tau(\mathbf{k},E)}\right\} \qquad (2)$$

Indeed, as shown in [3], $\tau(\mathbf{k},E)$ can be expressed, using the steady state BE for $f_S(\mathbf{k},E+\delta E)$ and $f_S(\mathbf{k},E)$, and the transient BE for $\langle f(\mathbf{k},t)\rangle$. One then finds that $\tau(\mathbf{k},E)$ depends on the direction of the perturbation $\delta\mathbf{E}$ with respect to the electric field $\mathbf{E}$:

$$\tau(\mathbf{k},E) = -\frac{\hbar}{q}\,\frac{\delta\mathbf{E}\cdot\nabla_E f_S(\mathbf{k},E)}{\delta\mathbf{E}\cdot\nabla_k f_S(\mathbf{k},E)} \qquad (3)$$

When $\delta\mathbf{E}$ is perpendicular to $\mathbf{E}$, one gets $\tau_\perp(\mathbf{k},E)$. In polar coordinates with the polar axis along $\mathbf{E}$, $\mathbf{k}=(k,\theta)$ with $k=|\mathbf{k}|$ and $\theta=(\mathbf{E},\mathbf{k})$. $\tau_\perp(\mathbf{k},E)$ is then easily expressed as [3]:

$$\tau_\perp(\mathbf{k},E) = \frac{\hbar k}{qE}\,\frac{\partial f_S(\mathbf{k},E)/\partial\theta}{(\partial f_S(\mathbf{k},E)/\partial\theta)\cos\theta + k(\partial f_S(\mathbf{k},E)/\partial k)\sin\theta} \qquad (4)$$

Within the above hypothesis, the evolution of the distribution function $\langle f(\mathbf{k},t)\rangle$, around its steady state value $f_S(\mathbf{k},E)$, is governed by the equation:

$$\frac{\partial\langle f(\mathbf{k},t)\rangle}{\partial t} = -\frac{\langle f(\mathbf{k},t)\rangle - f_S(\mathbf{k},E)}{\tau(\mathbf{k},E)} \qquad (5)$$

As an example, figure 1 shows the variations of $\tau_\perp(\mathbf{k},E)$ versus $k=|\mathbf{k}|$, along three directions with respect to the electric field $\mathbf{E}$ ($\theta=0$, $\theta=\pi/2$, and $\theta=\pi$), in undoped p-type Silicon at 300 K, E=2 kV/cm (fig. 1a) and E=20 kV/cm (fig.1b). $f_S(\mathbf{k},E)$ was computed numerically, using a matrix method, so as $\tau_\perp(\mathbf{k},E)$ through eq.(4). We used a one spherical non parabolic band model. As can be seen fig.1, $\tau_\perp(\mathbf{k},E)$ is in the range 0.01 to 0.2 ps at 20 kV/cm, and increases as the field strength decreases.

4. FLUCTUATION OF THE STATE OCCUPANCY FUNCTION:

In order to study the fluctuations of the SOF, we apply the general theory developped in [1], to the particular case studied here, where the BE is replaced by eq. (5). According to [1], one sets: $\langle f(\mathbf{k},t)\rangle = f_S(\mathbf{k},E) + \delta f(\mathbf{k})\exp(i\omega t)$. This expression, carried into eq. (5), gives for the first order terms, $\delta f(\mathbf{k})$ as a solution of:

$$\left[i\omega + \frac{1}{\tau(\mathbf{k},E)}\right]\delta f(\mathbf{k}) = 0 \qquad (6)$$

The Langevin theory then relies the fluctuation $\hat{f}(\mathbf{k},\omega)$ to the white noise source term $\xi(\mathbf{k})$ by:

$$\left[i\omega + \frac{1}{\tau(\mathbf{k},E)}\right]\hat{f}(\mathbf{k},\omega) = \xi(\mathbf{k}) \qquad (7)$$

This gives, for the spectral density $S_f(\mathbf{k},\mathbf{k}',\omega)$ of the fluctuations of the SOF:

$$S_f(\mathbf{k},\mathbf{k}',\omega) = \frac{\tau^2(\mathbf{k},E)}{1 + \omega^2 \tau^2(\mathbf{k},E)} S_\xi(\mathbf{k}) \delta_{\mathbf{k}\mathbf{k}'} \qquad (8)$$

where $\delta_{\mathbf{k}\mathbf{k}'}=1$ if $\mathbf{k}=\mathbf{k}'$, and $\delta_{\mathbf{k}\mathbf{k}'}=0$ if $\mathbf{k}\neq\mathbf{k}'$. $S_\xi(\mathbf{k})$ can be evaluated using the variance $\langle\Delta f^2(\mathbf{k},t)\rangle$ which writes [1]:

$$\langle\Delta f^2(\mathbf{k},t)\rangle = f_S(\mathbf{k},E)\,[1-f_S(\mathbf{k},E)] = \int_0^\infty S_f(\mathbf{k},\mathbf{k},\omega)\ d\nu \qquad (9)$$

Carrying eq.(8) in eq.(9) readily gives $S_\xi(\mathbf{k})$, so that finally, with $F_S(\mathbf{k},E) = f_S(\mathbf{k},E)[1-f_S(\mathbf{k},E)]$:

$$S_f(\mathbf{k},\mathbf{k}',\omega) = \frac{4\tau(\mathbf{k},E)}{1 + \omega^2 \tau^2(\mathbf{k},E)} F_S(\mathbf{k},E) \delta_{\mathbf{k}\mathbf{k}'} \qquad (10)$$

5. DIFFUSION COEFFICIENT:

The spectral density $S_{j\alpha\beta}(\omega)$ of the fluctuations of the current density writes [2][1]:

$$S_{j\alpha\beta}(\omega) = \frac{q^2}{\Omega^2} \sum_{\mathbf{k}} \sum_{\mathbf{k}'} v_\alpha(\mathbf{k}) v_\beta(\mathbf{k}')\, S_f(\mathbf{k},\mathbf{k}',\omega) \qquad (11)$$

where Ω is the volume of the sample, and α and β mean directions with respect to the electric field. The local noise source term $K_{\alpha\beta}(\omega)$ is:

$$K_{\alpha\beta}(\omega) = \Omega\, S_{j\alpha\beta}(\omega) = 4\, q^2\, n\, D_{\alpha\beta}(\omega) \qquad (12)$$

which gives, using eqs.(11) and (10), replacing summations over $\mathbf{k}$ by integrations over $(\mathbf{k})$ space, and taking into account the density of states $1/8\pi^3$:

$$D_{\alpha\beta}(\omega) = \frac{\int \frac{\tau(\mathbf{k},E) v_\alpha(\mathbf{k}) v_\beta(\mathbf{k})}{1 + \omega^2 \tau^2(\mathbf{k},E)} F_S(\mathbf{k},E)\ d^3\mathbf{k}}{\int f_S(\mathbf{k},E)\ d^3\mathbf{k}} \qquad (13)$$

Using this method, we computed the transverse diffusion coefficient $D_\perp(E)$ in p-type Silicon, with doping impurity concentrations $N_A=0$ and $N_A=10^{17}$ cm^{-3}, at t=300 K, as a function of the electric field intensity E. Then, in eq.(13), $v_\alpha(\mathbf{k})=v_\beta(\mathbf{k})=v\sin\theta\cos\varphi$, $\tau(\mathbf{k},E)=\tau_\perp(\mathbf{k},E)$ and $F_S(\mathbf{k},E)\approx f_S(\mathbf{k},E)$ since in our case the semiconductor is non degenerate. $D_\perp(E)$ was also computed using a Monte Carlo simulation. The results are shown figure 2: taking into account the statistical dispersion of the Monte Carlo simulation, the results are in excellent agreement.

The result presented here is the first theoretical determination of the hot carrier diffusion coefficient using the distribution function, with non randomizing scatterring mechanisms Depending on the electric field strength, it is 10 to 50 CPU time less consuming than the Monte Carlo simulation.

REFERENCES:

[1] J.P. NOUGIER, J.C. VAISSIERE, Phys. Rev. B15, 8882 (1988)
[2] S.V. GANSEVITCH, V.L. GUREVITCH, R. KATILIUS, Rivista Nuovo Cimento, 2, 1 (1979)
[3] J.P. NOUGIER, M. ROLLAND, J. Appl. Phys. 48, 1683 (1977)

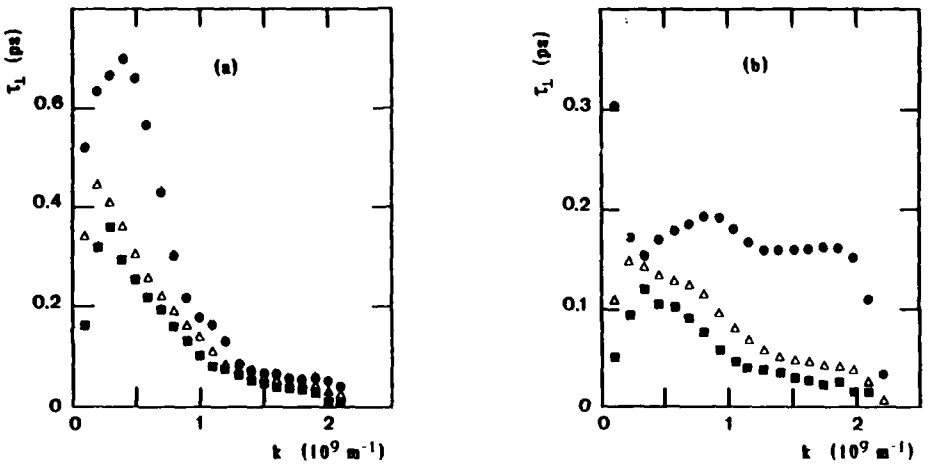


Figure 1: $\tau_\perp(\mathbf{k},E)$ for holes in undoped Silicon, as a function of $k=|\mathbf{k}|$, at 300 K, along three directions $\theta=0$ (●), $\theta=\pi/2$ (△), and $\theta=\pi$ (■), with respect to the electric field E. (a) E=2 kV/cm, (b) E=20 kV/cm.

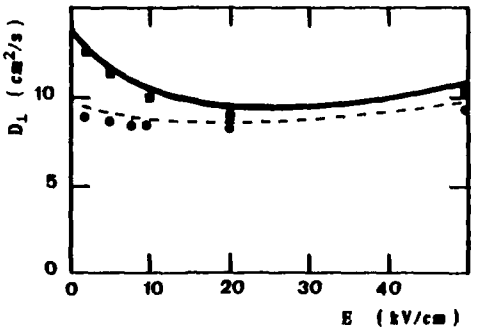


Figure 2: Transverse diffusion coefficient $D_\perp(E)$ of holes, as a function of the electric field intensity E, in pure ($N_A=0$) and doped ($N_A=10^{17}$cm^{-3}) p-type Si at 300 K.
Monte Carlo simulation: $N_A=0$ (——) and $N_A=10^{17}$ cm^{-3} (- - - -)
Computed using eq.(13): $N_A=0$ (■) and $N_A=10^{17}$ cm^{-3} (●)

Session 2B Room B : Metallization, silicides

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

PROPERTIES OF WSi_2 : OHMIC CONTACT TO N^+ AND P^+ Si, BARRIER BETWEEN Al AND Si, AND FEASIBILITY AS FIRST METAL IN MULTILEVEL METALLIZATION PROCESSES

S.-L. ZHANG, M. HAMMAR, T. JOHANSSON and R. BUCHTA

Swedish Institute of Microelectronics, PO Box 1084, S-164 21 Kista, Sweden

Abstract: The electrical and chemical properties of WSi_2 as well as the silicide formation have been examined with respect to applications in VLSI Si-Technology.

1. INTRODUCTION

Due to the shrinkage of device dimensions in modern VLSI technology, a barrier material between Si and Al metallization is of increasing importance. Such a barrier must meet three important requirements at all temperatures of interest: 1) low specific contact resistivity to Si, 2) good barrier properties to prevent Si-Al interdiffusion and 3) chemical stability with the adjacent materials i.e. Si and Al. Several metals and their silicides have been investigated for this purpose. None of the silicides examined in the literature possesses all the desired properties. They fail either due to Al-containing compound formation in the case of transition metal silicides or due to Si-Al interdiffusion in the case of refractory metal silicides at temperatures around 400°C [1]. W and WSi_2 have recently been more concerned due to the selective deposition technique as well as the material characteristics, i.e. thermal stability, low resistivity, etc. Although a number of reports on WSi_2 has appeared [2], very limited information is available about the Al/WSi_2/Si system.

SALICIDE (Self-Aligned siLICIDE) technology using WSi_2 could, as a first sight, be very attractive since a selective technique for W film deposition is available. However, a most recent study showed that stress induced large dislocation half-loops might be created along silicide edges when silicon substrate is used to react with deposited metal layer to form silicide [3]. Moreover, for the purpose of making a first level metal layer using WSi_2 for shallow junction MOS processing, it would be unlikely to chemically deposit W films directly onto heavily doped source/drain (S/D) Si areas because: 1) formation of encroachment or wormholes [4] during W deposition using WF_6; 2) dopant redistribution in S/D Si areas caused by the consumption of Si substrate during WSi_2 formation and thermal treatment at higher temperatures required for completing W/n^+- or p^+-Si reactions [5,6]; 3) a rough interface between WSi_2 layer and Si substrate resulting from rough interface between the as-deposited W film and Si substrate caused by a native oxide on Si surface prior to the W deposition and very uneven surface of LPCVD-W films [5,7-9].

Therefore, in order to prepare a WSi_2 layer contacting to the heavily doped S/D Si areas without affecting the shallow junction, a poly-Si buffer layer is required.

2. EXPERIMENTAL AND EVALUATION

3-inch <100>-oriented Si wafers with a resistivity of 16 to 24 Ωcm were used throughout this study. Diodes and Kelvin structures were made on test wafers

for electrical measurements. Heavily doped n-Si substrates with a surface carrier concentration of 7.5E19 cm^{-3} were made on boron doped p^--Si wafers by As-ion implantation and high temperature activation. P^+-Si substrates were prepared by BF_2-ion implantation with two different doses, on phosphorus doped n^--Si wafers. After activation, surface carrier concentrations of 1.6E19 and 2.6E19 cm^{-3} were respectively achieved. Undoped poly-Si layers of 2500 Å were then deposited followed by deposition of 1000 Å W films. Undoped poly-Si was chosen to complete WSi_2 formation at 800°C for 30 minutes because a substantial delaying in silicide formation rate was observed if Si is heavily doped [5,6].

Reference samples were prepared in the same procedure but on blanket Si wafers and some of them with different W and/or poly-Si thicknesses.

Leakage current and specific contact resistance (R_c) were measured on the test wafers. Rutherford backscattering spectrometry (RBS) and X-ray diffraction (XRD) analyses as well as sheet resistance (R_s) measurement were performed on corresponding structures on blanket wafers.

3. RESULTS AND DISCUSSION

1) As demonstrated in Fig. 1, a maximum R_s value appears around annealing temperature of 700°C, as a W/Si structure was heated. For achieving low sheet resistivity, a moderate annealing temperature of 800°C is sufficient for WSi_2 prepared in this way. At this temperature, a resistivity of 25 to 30 $\mu\Omega cm$ is obtained.

It is of great importance to be sure that there is no excess poly-Si left between WSi_2 and S/D Si areas after silicide formation so as to make a good contact. It is also of great significance that excess W on top of WSi_2 should be avoided for the study of WSi_2 as a diffusion barrier between Al and Si. XRD analyses with a diffractometer and Read camera showed that WSi_2 of tetragonal structure was formed and neither poly-Si nor W was left after annealing, within the resolutions of the techniques. The Al/WSi_2/Si structure was isochronally annealed for 1 hour in the temperature range of 350 to 575°C and the R_s values were then measured. Comparatively, the Al/W/Si structure was

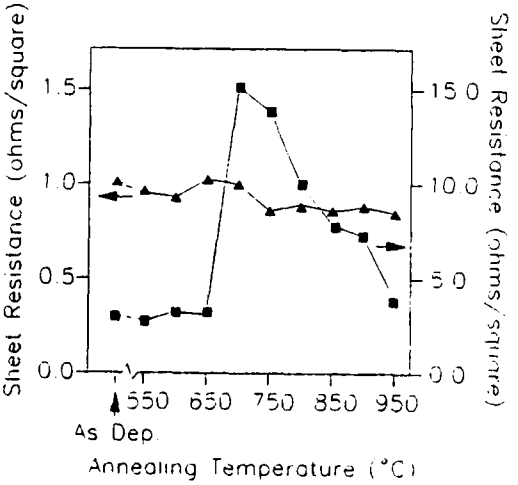


Fig. 1 R_s vs. anneal T for W/Si.

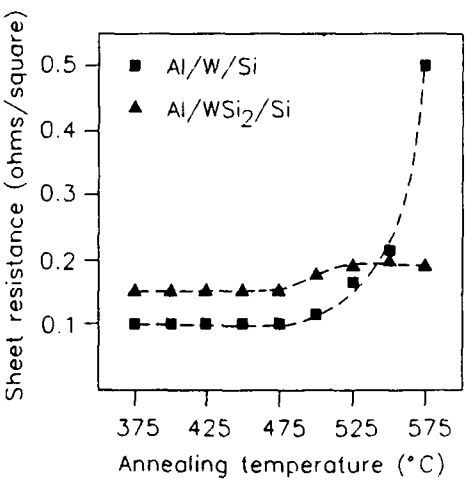


Fig. 2 R_S vs. anneal T for Al/W/Si and Al/WSi_2/Si.

also evaluated. Results from both structures are plotted together in figure 2. The substantial increase in R_s of the Al/W/Si system is due to a WAl_{12} formation [10]. For the Al/WSi_2/Si system, the minor increase of R_s observed at 475°C is probably due to the reaction between a small amount of W left after silicidation and the Al layer on top resulting in W-Al alloy, which was not detectable by XRD technique.

2) Undoped poly-Si layers of 4800 Å and W films of 3400 Å were deposited on both heavily and lightly doped Si substrates to show the retardation of silicidation process on the heavily doped substrates. Annealing of these structures was done in an argon flow furnace. Annealing temperatures used were 750, 800 and 950°C. Different times were applied. At low annealing temperatures, the retarding effect is obviously exposed by the diffractometer if the change of diffractory intensity of W-(110) plane is examined. Diffractory intensities of annealed samples are normalized to that of the as-deposited sample (figure 3). On the other hand, identical results, i.e.

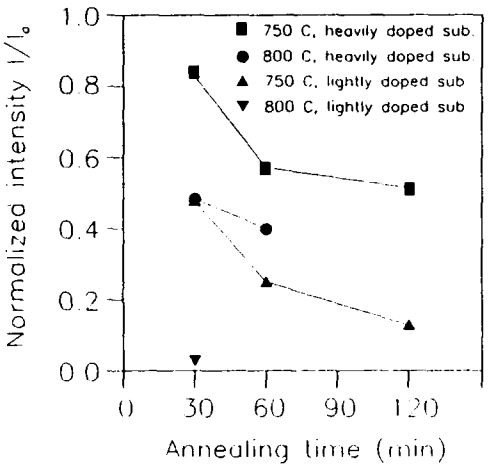


Fig.3 I/I_o vs. anneal T and time for W/poly-Si/mono-Si substrate.

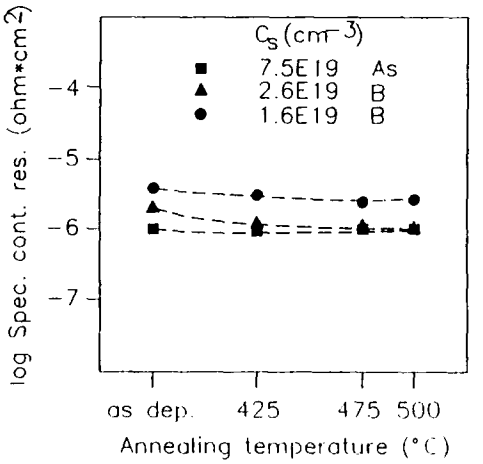


Fig.4 ρ_c vs. anneal T for Al/WSi_2/Si.

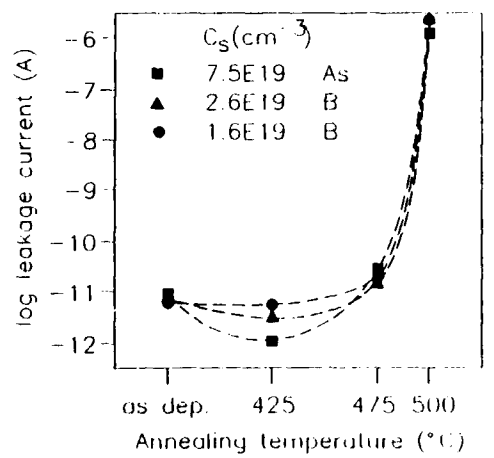


Fig. 5 Leakage current vs. anneal T for Al/WSi_2/Si.

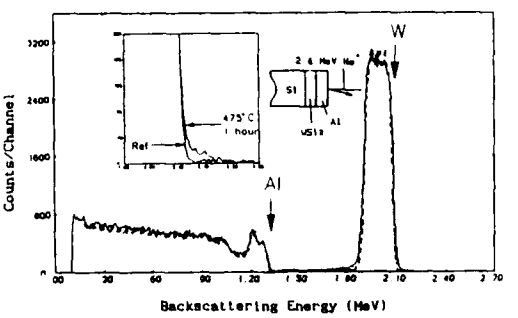


Fig.6 RBS spectra for Al/WSi_2/Si after 475°C, 1 hour anneal.

WSi_2 formation and no W left, were found for both heavily and lightly doped substrates when 950°C was used for silicidation.

3) The specific contact resistivity and diode leakage current of the system Al/WSi_2/Si were evaluated after heat treatment at different temperatures. RBS and XRD analyses and R_s measurement were performed to study the thermal stability of the system.

The contact resistivities of WSi_2 to an n^+-Si substrate with a surface carrier concentration of 7.5E19 cm^{-3} and a p^+-Si substrate with that of 2.6E19 cm^{-3}, were 9E-7 Ωcm^2 and 1E-6 Ωcm^2 (figure 4), respectively.

However, all of the diodes failed after annealing at temperatures higher than 475°C (figure 5), indicating interpenetration of Al and Si which was revealed by means of RBS analysis (figure 6). On the other hand, any kind of alloy formation has not been found even after the samples treated at a higher temperature of 600°C for 1 hour. This is probably due to Al atoms diffusing through WSi_2 grain boundaries within the silicide and perhaps relates to the absence of tungsten-aluminum reaction.

As a comprehension, the Al/LPCVD-W/Si system was also investigated. Here, the Al-Si interpenetration, the alloy formation of WAl_{12} and a substantial increase in sheet resistance and leakage current all occurred at 500°C [10].

4. SUMMARY

WSi_2 films were made by solid-phase reaction between LPCVD-W and undoped poly-Si on top of heavily doped Si substrates. Low and uniform specific contact resistances of WSi_2 to n^+ and p^+ Si substrates were achieved. Interpenetration of Al and Si was not found below 475°C when a WSi_2 layer of 2500 Å was used as a contact barrier. No W-Si-Al alloy formatiom was detected. The good and reliable ohmic contact of WSi_2, made in such a way described above, to both n^+ and p^+ Si makes WSi_2 a promising candidate as first metal in a multilayer metallization process, for which the planarization at high temperatures is desired.

REFERENCES

1. C. Y. Ting and M. Wittmer, J. Appl. Phys., 54, 937 (1983).
2. See, for example, K. C. Saraswat, D. L. Brors, J. A. Fair, K. A. Monnig and R. Beyers, IEEE Trans. Electron Devices, ED-30, 1497 (1983).
3. L. Van den hove, J. Vanhellemont, R. Wolters, W. Claassen, R. De Keersmaecker and G. Declerck, Proceedings of ECS Spring Meeting, Atlanta, USA, pp. 312-313, May 1988.
4. See, for example, M.L. Green and R.A. Levy, J. Electrochem. Soc., 132, 1243(1985).
5. L.D. Locker and C.D. Capio, J. Appl. Phys., 44, 436ε(1973).
6. T.I. Kamins, S.S. Laderman, D.J. Coulman and J.E. Turner, J. Electrochem. Soc., 133, 1439(1986).
7. E.K. Broadbent and C.L. Ramiller, J. Electrochem. Soc., 131, 1427(1984).
8. K.Y. Tsao and H.H. Busta, J. Electrochem. Soc., 131, 2702(1984).
9. C.M. Melliar-Smith, A.C. Adams, R.H. Kaiser and R.A. Kushner, J. Electrochem. Soc., 121, 298(1974).
10. M. Hammar, S.-L. Zhang, R. Buchta and T. Johansson, Proceedings of 13th Nordic Semiconductor Meeting, Ed. by M. Östling, pp. 101-104, Saltsjöbaden, Sweden, June 5-8, 1988.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

EFFECTS OF COMPOUND FORMATION WITH DOPANTS IN $TaSi_2$

V. PROBST[1], H. KABZA and H. GOEBEL

Siemens AG, Corporate Research and Development, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

Abstract - TaB_2 compound formation in $TaSi_2$ is demonstrated by means of SEM, SIMS- and XRD analysis. This holds for B-doping of $TaSi_2$ by means of diffusion from poly-Si as well as for direct implantation. The formation of TaAs could not yet be proven. As the mobile boron concentration level in $TaSi_2$ is low due to the compound formation diffusion from $TaSi_2$ into Si is rather limited.

1 - INTRODUCTION

With shrinking dimensions both laterally and vertically in IC technology the requirements on sheet and contact resistivities are becoming more severe. Therefore silicides have become a focus of interest. They are used as silicide, salicide, and polycide layers and also applied as a diffusion source for junction formation /1/. One of the straightforward applications is the formation of a direct contact to highly doped regions, either by siliciding doped silicon or by dopant diffusion from silicide into Si. Heat treatments like BPSG reflow or diffusion steps redistribute the dopants within the silicide. However, from thermodynamic considerations, it can be deduced that upon annealing the dopants not only diffuse, but also may form compounds with the metal of the silicide /2/. This undesirable behaviour has been demonstrated in the case of boron doped $TiSi_2$ /3/ and is also likely to be found in $TaSi_2$, because the heats of formation for these silicides are much lower than for the respective metal-boron compounds (heats of formation from /4/ in kJ/mole: $TiSi_2$: - 134,3 TiB_2: -323,9; $TaSi_2$: -119,3 TaB_2: -209,3. For $TiSi_2$ it was demonstrated that the formation of titanium boride particles effectively limits the concentration of freely diffusing boron at the interface to the mono-Si which leads to unacceptably high contact resistances /3/.
In this work boron and arsenic doped $TaSi_2$ was investigated with respect to compound formation and its influence on the diffusion behaviour.

2 - EXPERIMENTAL

In order to exclude any possible process induced enhancement of compound formation (like radiation damage or high concentration effects) direct implantation into the silicide was avoided and B or As-implanted poly-Si was used as a doping source for the $TaSi_2$ instead. 160 nm of amorphous CVD-Si were deposited on <100> oriented Si and implanted with $8 \cdot 10^{15}$ B cm^{-2}, 15 keV or $5 \cdot 10^{15}$ As cm^{-2}, 50 keV. Subsequently 100 nm of tantalum silicide were sandwich-sputtered (1st layer Ta; sandwich thickness 10 nm), capped with 200 nm of CVD-SiO_2 and subjected to a heat treatment of 900°C, 30 min.

---

[1] also Deparment of Integrated Circuits, Technical University Munich

In a second experiment the influence of metal dopant compound formation on junction formation was investigated. 200 nm of $TaSi_2$ were sandwich sputtered on substrate wafers and annealed at 900°C, 30 min. Subsequently, B-implantation (25 keV, $5 \cdot 10^{15} cm^{-2}$ or $2 \cdot 10^{16} cm^{-2}$), oxide capping (100 nm SiO_2) and diffusion steps from 800°C, 30 min to 1000°C, 30 min were performed.
SEM, SIMS, TEM and XRD analyses in combination with specific selective etching procedures were used to characterize the samples.

3 - RESULTS

3.1 - Evidence of metal dopant compound formation:

a) Boron

After removing the capping oxide and $TaSi_2$-layer selectively to the B-doped poly-Si, an SEM-micrograph (fig. 1) of the remaining surface reveals the existence of bright particles ($\phi \leq 0,5$ µm) like in the case of $TiSi_2$/3/.

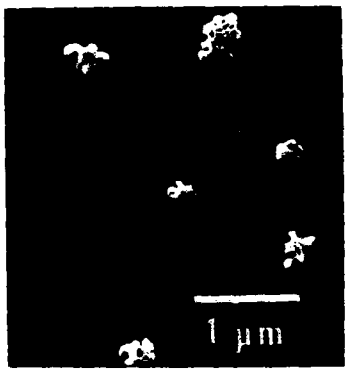


Fig. 1: SEM micrograph of the remaining poly-Si surface with particles after selectively etching the $TaSi_2$.

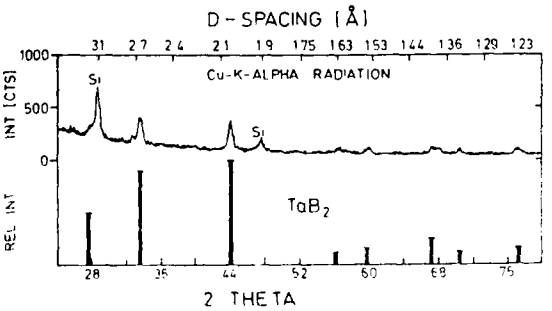


Fig. 2: X-ray diffraction patterns of the sample in fig. 1, obtained at grazing incidence and compared with the powder diffraction pattern for TaB_2.

An x-ray powder diffraction spectrum (fig. 2) at grazing incidence demonstrates that these particles consist of pure TaB_2. In a $NH_4OH:H_2O_2:H_2$ = 1:1:5 cleaning solution they can be completely removed. This is also demonstrated by SIMS analyses; fig. 3 shows the B and Ta signals as measured before and after this clean. For the latter case the apparent B surface concentration is drastically reduced and the Ta signal is at the detection limit.

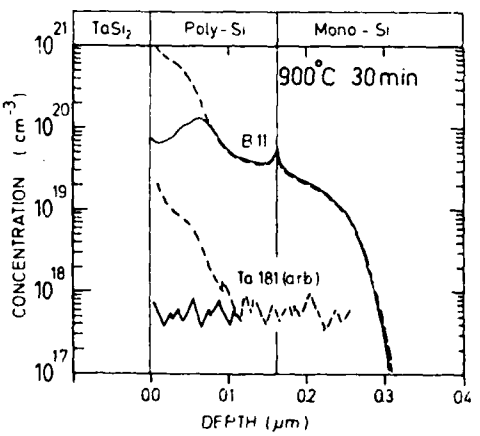


Fig. 3: SIMS-B-profile of the remaining poly-Si on mono-Si after selectively removing the $TaSi_2$ (---) and after an additional cleaning step (---) to remove the TaB_2 particles.

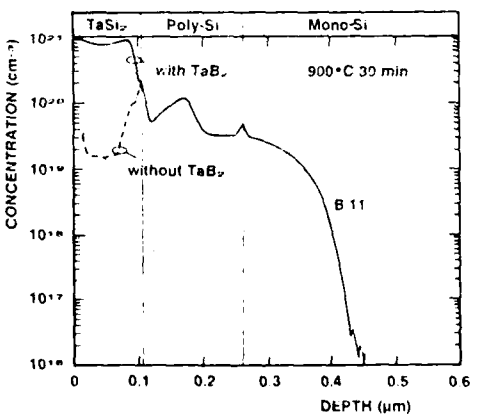


Fig. 4: SIMS profile of B before and after a selective etch of the TaB_2 particles.

Of course it is very important to separate the mechanism of TaB_2 formation from the boron diffusion processes. As nearly all TaB_2-particles are grown across the whole $TaSi_2$-layer, this separation can be done by selectively etching the TaB_2 crystallites out of the $TaSi_2$ bulk.On the remaining $TaSi_2$-layer surface holes are seen in the SEM at those locations, where TaB_2 crystallites had been. Distribution and size of these holes are exactly the same as those of the TaB_2 crystallites in fig. 1. SIMS profiles across the $TaSi_2$ -polycide structure before and after the etching procedure are given in fig. 4.
The B profiles within the polycrystalline and bulk Si are identical before and after the etch step. As expected they show a residual implantation peak due to slow B-diffusion inside the grains of the poly-Si, a segregation peak at the poly-Si/bulk-Si interface and a steep diffusion profile within the bulk-Si. Within the $TaSi_2$-layer, however, the B-profile changes drastically upon removal of the TaB_2 particles. The B concentration drops from $8{\cdot}10^{20}$ atms/cm^3 to appr. $1.7{\cdot}10^{19}$ atms/cm^3 which amounts to a factor of 50. The pile-up of boron at the $TaSi_2$/poly-Si interface for the etched sample is due to small TaB_2 particles ($\phi \approx 20$ nm) which are revealed by SEM analysis at the interface after the $TaSi_2$ layer has been removed. From these data an estimate of the average mobile concentration of B in $TaSi_2$ (from TEM: average grain size $d_g \approx 80$ nm) gives $2{\cdot}10^{19}$atms/cm^3 as an upper limit for 900°C, 30 min, as the existence of such small TaB_2 crystallites within the $TaSi_2$ layer, where they could not be selectively removed, cannot be excluded. The excess boron, supplied by fast grain boundary diffusion from the poly-Si, obviously nucleates in TaB_2 crystallites. Increasing the supply of boron would not increase the concentration of mobile boron in $TaSi_2$ but probably lead to further growth of the TaB_2 crystallites.

b) Arsenic
In order to detect any Ta_xAs_y compound, an identical set of analyses was applied to the As-doped polycide structure After etching off the $TaSi_2$-layer some kind of precipitates was found on the etched sample even after an overetch of 500%. X-ray fluorescence analysis indicated the presence of Ta and As. However, no clear evidence for a Ta_xAs_y compound could be found by XRD due to coincidence of the strongest lines for TaAs with those of $TaSi_2$. It cannot be excluded that during the anneal some trace of $TaSi_2$ migrates into the poly-Si layer like has been seen e.g. in the case of $CoSi_2$ /5/ and TiSi /6/. Such inclusions, of course, cannot be selectively removed from the poly-Si and therefore would provide the $TaSi_2$ signal in the XRD spectrum. Further analytical work is in progress.

3.2 - Impact of TaB_2-compound formation on the boron diffusion behaviour:

a) inside the $TaSi_2$ layer
As expected, direct implantation of boron into $TaSi_2$ (instead of poly-si) and additional diffusion steps from 800°C to 1000°C lead to TaB_2 formation as well. From SEM analyses it can be deduced, that the nucleation behaviour is quite different from the polycide case. A lot of tiny TaB_2 crystallites were formed in the regime of the implantation peak as well as at the interface to the monocrystalline Si. Their size (ϕ = 10 to 100 nm) and number increased with increasing annealing temperature and implantation dose. Upon B implantation into $TaSi_2$ ($2{\cdot}10^{16}$cm^{-2}, 25 keV) an annealing step of 900°C, 30 min leaves the implantation peak nearly unchanged (fig. 5) which is mainly due to quick TaB_2 formation. Also, quick grain boundary diffusion obviously leads to a transport of boron to the $TaSi_2$/Si interface. There again nucleation of TaB_2 crystallites occurs, possibly supported by catalytic effects of interface impurities like O_2. This causes a depletion of the boron concentration in the middle of the $TaSi_2$ layer.

b) Diffusion from B-implanted $TaSi_2$ into monocrystalline Si
After removal of the $TaSi_2$ layer and TaB_2-particles, the resulting B diffusion profile shows a shallow profile (x = 110 nm) and a very low interface concentration of $\approx 4{\cdot}10^{18}$ B cm^{-3} (fig. 6) in good agreement with /7/. As TaB_2-formation effectively limits the average concentration of freely diffusing boron within $TaSi_2$ to less than $2{\cdot}10^{19}$ B cm^{-3} B diffusion from $TaSi_2$ into Si is reduced accordingly. This is clearly demonstrated by comparison with a B profile obtained from $TiSi_2$ as a diffusion source /3/ (effective TiB_2 formation) and as opposed to that poly-Si as a diffusion source /7/ where there is no compound formation.

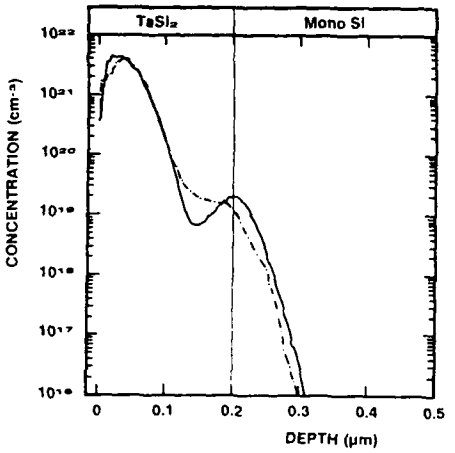


Fig. 5: SIMS profile of B implant ($2 \cdot 10^{15} cm^{-2}$, 25 keV) into $TaSi_2$ (---) and upon annealing at 900°C for 30 min (—).

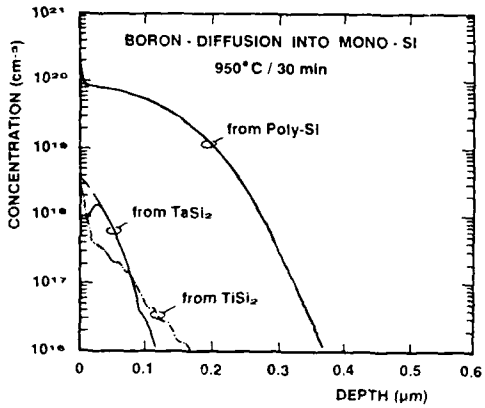


Fig. 6: SIMS profile of B-diffusion from $TaSi_2$ as compared with diffusion from $TiSi_2$ and poly-Si.

4 - CONCLUSIONS

As in the case of $TiSi_2$ also for $TaSi_2$ compound formation with boron has been demonstrated. So in order to obtain meaningful SIMS data special ways of sample preparation have to be followed. As most of the boron content is immobilized in the compounds the concentration level of freely moving boron within the $TaSi_2$ is severely limited and consequences for contact resistivities are to be expected.

ACKNOWLEDGEMENT

We gratefully acknowledge the contribution of several SIMS analyses by M. Metzger and P. Eichinger as well as process supervision by B. Hoffmann.

References

/1/ F. M. d'Heurle, Solid State Devices 1985, edited by P. Balk and O.G. Folberth, p. 213
/2/ S.P. Murarka, D. S. Williams, J. Vac.Sci Technol. B5(6), p. 1674 (1987)
/3/ V. Probst et al., Appl. Phys. Lett. 52 (21), p. 1803 (1988)
/4/ O. Kubaschewski, C.B. Alcoick, Metallurgical Thermochemistry (1979), p. 312
/5/ S.P. Muraka et al, J. Vac. Sci. Technol. B5 (4), (1987), p. 865
/6/ P. Lippens et al., to be published in ESSDERC 88, Montpellier
/7/ H. Gierisch et al, J. Vac. Sci. Technol. B5 (2), p. 508 (1987)
/8/ V. Probst et al, J. Electrochem. Soc. vol. 135, No. 3, p. 671 (1988)

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

NONSELECTIVE W/WSi_x-CVD TECHNOLOGY FOR LOW RESISTANCE VIA PLUGS ON ALUMINUM

E. BERTAGNOLLI, C. WIECZOREK, B. HOFFMANN and H. SCHABER

Siemens AG, Corporate Research and Development, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

Abstract: By reducing the thermal budget of a nonselective W/WSi_x-CVD technology, a via filling process completely compatible with a conventional aluminum track scheme has been obtained. A specific via resistance in the range of a few 10^{-8} Ωcm^2 was realized. Hillock formation was suppressed significantly. No fluorine pile-up was found at the WSi_x-/AlSiTi-interface so the formation of an AlF_3-interlayer has been avoided successfully.

Introduction

In multilevel metallization the formation of reliable and planar interlevel contacts is a key issue. Sputtered metals for via contacting fail due to self-shadowing problems as via dimensions are continuously shrunk. Blanket CVD-tungsten offers unique conformal step coverage. Therefore it is well suited for via filling purposes. Tungsten plugs in vias allow for the use of headless metal runners in a conventional aluminum track scheme. They fulfill the option of stackability of interlevel contacts. Tungsten on aluminum, however, suffers from a high interface resistance and an unfavourable thermal budget /1/.

The formation of a highly resistive AlF_3-layer is an inherent problem as long as tungsten fluorine (WF_6) is used for the deposition of tungsten and tungsten compounds /2/. A high deposition temperature prevents the formation of such compounds with low volatility, which influence the interfacial resistance. High temperature, however, enhances hillock formation. Also the reliability of an optional diffusion barrier at the aluminum-/silicon-interface in the contact area may be endangered. An important goal in every multilevel metallization scheme is the repeated application of the via filling process. This aggravates the problem of contact resistances in connection with low deposition temperature.

This paper describes a tungsten via plug process aiming at the minimization of via resistance and at hillock suppression. The main emphasis is put on optimization of the integrability in a multilevel wiring scheme based on AlSiTi-metallization.

In order to clarify the importance of interfacial phenomena and geometrical effects the measured resistances are compared with 2-D simulations. Calculations were done using the program VLSICAP /3/, which is based on a 2-D finite element method.

Experimental

A first metal layer of sputtered and patterned AlSiTi-alloy was planarized by an etch back of a plasma oxide/polyimide double layer /4/. 1000 nm plasma

oxide were deposited as intermetal dielectric. The different-sized viaholes were opened in a dry plasma etch process using CHF_3/O_2. The native oxide on the AlSiTi-surface in the vias was removed by a dip etch in dilute HF just before deposition of tungsten in a cold wall reactor.

Blanket CVD-tungsten was deposited using $WF_6/SiH_4/H_2$ as reactive gases. In order to avoid peel-off of tungsten due to bad adhesion on silicon oxide a 100 nm WSi_x-layer ($x \approx 2.4$) was deposited first. Then 500 nm of tungsten were grown on top of this interlayer. This double layer deposition was performed at various temperatures (T_D = 400, 425 and 450°C).

Vias with both lateral dimensions larger than 1 µm are not completely filled by this process; the inner part of such a via remains empty. A sacrificial layer of 300 nm polyimide was spun onto the wafers. This layer was etched prior to the tungsten etch back process, thus leaving protective polyimide plugs in the grooves between the vertical tungsten layers covering the sidewalls of the vias. Then the W/WSi_x on top of the dielectric was removed by an etch back process using a SF_6/O_2-plasma (etch-rate: 400 nm tungsten/min, uniformity: +/- 5%) /5/. Finally the polyimide plugs were removed with a resist stripping process. Thus the planar oxide surface was re-established.

A second metal layer of AlSiTi-alloy was sputtered and patterned. Finally the wafers were annealed for 15 minutes at 450°C in forming gas atmosphere.

Via resistances were measured using two level via chains with vias ranging down to an edge length of 1.0 µm. The via hole filling was controlled by SEM as well as by metallographic cross-sections. The composition of the layers was tested by electron microprobe mass analysis. Auger depth profiling was used to monitor the interface composition.

Results

Sheet resistance of the WSi_x-and W-layers are only weakly dependent on deposition temperature in the range investigated. Strong hillock suppression is, however, achieved by decreasing T_D to 400°C (fig. 1). Even for this low deposition temperature, the goal of via resistances in the $10^{-8}\Omega cm^2$ regime has been achieved. Wereas the high temperature ($T_D \approx 450$°C) via resistance is predicted quite well by the simulation based on measured sheet resistances of fig. 1, the rise of R_{via} with decreasing T_D is underestimated by the calculations. This discrepancy points at the existence of a resistive interfacial layer at low T_D.

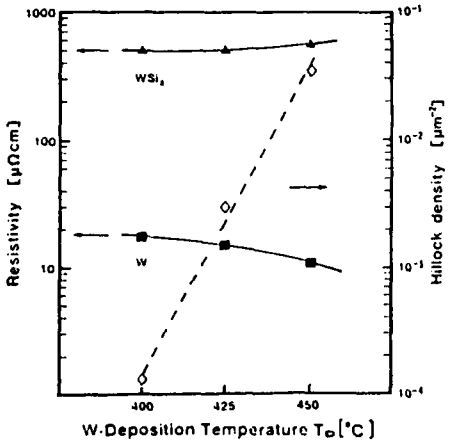


Fig 1: W- and WSi_x ($x = 2.4$) sheet resistances vs. deposition temperature T_D and hillock density on AlSiTi

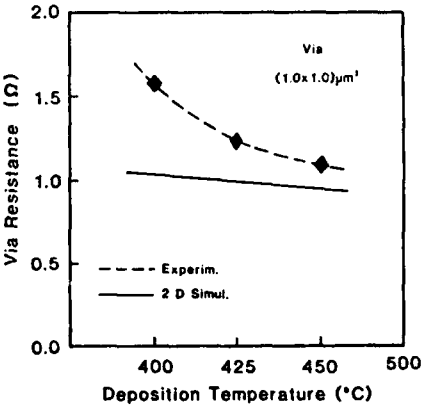


Fig.2: Resistance of a 1.0 x 1.0 µm² via vs. WSi_x/W deposition temperature. Simulations based on experimental sheet resistivities show only little increase of via resistance with decreasing deposition temperature

Vias ranging down to a nominal edge length of 1 µm have been sufficiently filled (fig.3a-c).

a) 1.0 x 1.0 µm² b) 1.1 x 1.1 µm² c) 1.2 x 1.2 µm²

Fig.3a-3c: SEM micrographs of via chains with CVD-tungsten filled vias of different size (T_D=400°C) after AlSiTi-2-patterning

A comparison of the measured via resistances for different via sizes with simulated values is given in fig. 4. The simulation was based on measured sheet resistances (fig. 1). Results are in good agreement with experimental data for T_D = 450°C. As the changes in sheet resistivity for various deposition temperatures are small (fig. 1), only a slight increase of via resistance with decreasing deposition temperature is predicted by simulation.
A disadvantage of the lower deposition temperature is the slightly reduced deposition rate (fig. 5). Below 400°C only insufficient growth rate could be obtained.

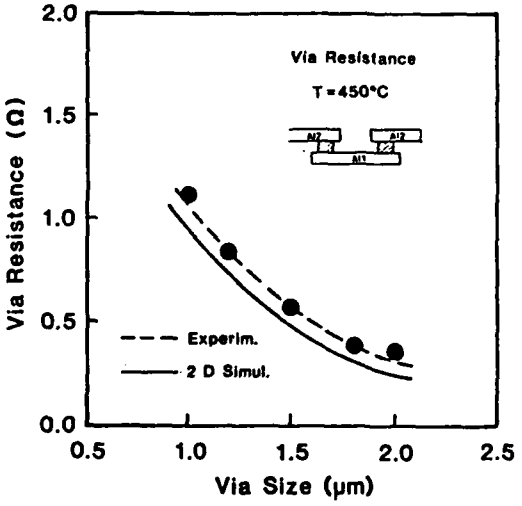


Fig.4: Via resistance vs. via size for T_D=450°C

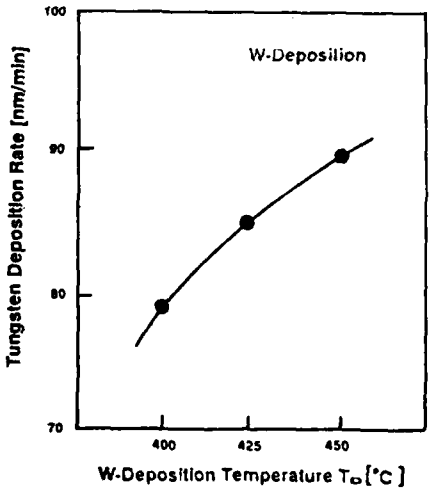


Fig.5: Dependence of deposition rate on deposition temperature

The investigation of the WSi_x-AlSiTi interface by AES-profiling did not show any pile up of fluorine (fig. 6). Fig. 7 shows a TEM-micrograph of this interface.

The formation of AlF_3 during the tungsten deposition process probably is prevented by the presence of silane. SiH_4 seems to enhance the formation of volatile fluorine compounds thus blocking the formation of AlF_3.

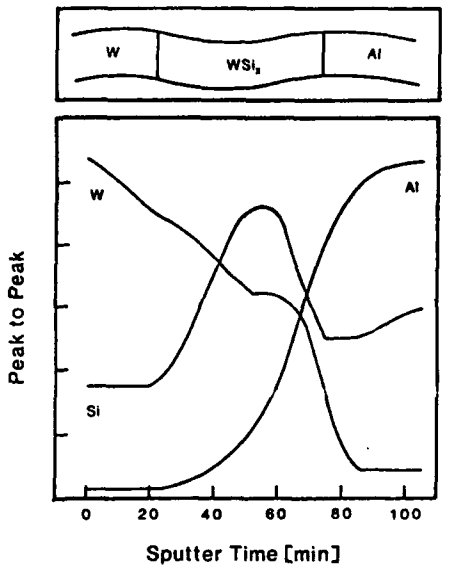


Fig.6: AES depth profile of W/WSi$_x$/AlSiTi-system after 400°C anneal. No indication of an interfacial layer (AlF_x) between AlSiTi and WSi$_x$ can be seen

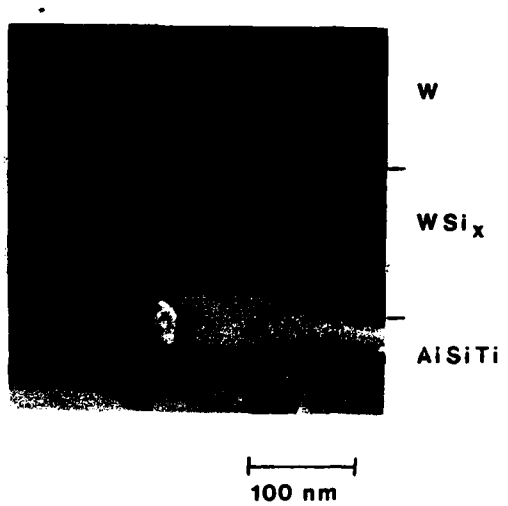


Fig.7: TEM micrograph of a cross-section of the W/WSix/AlSiTi-interface

Conclusions

With a low temperature W/WSi$_x$-CVD technology for via-filling a fully stackable multilayer interconnect system using conventional aluminum tracks comes into reach. In particular the well known trade-off between via resistances and hillock formation could be greatly alleviated.

Acknowledgement

The authors would like to thank M. Stolz for preparing the metallographic cross-sections, E. Renner for SEM analysis, A. Mitwalsky for TEM analysis, W. Pamler for AES profiling and I. Maier for electrical characterization.

References

/1/ R.S. Blewer, Solid State Technol., Nov. 1986, pp. 117

/2/ T. Moriya, S. Shinia, U. Hazuki, M. Chiba and J.M Kashiwagi, IEEE Techn. Dig. IEDM, N.Y. 1983, pp. 553

/3/ F. Straker, "VLSICAP", Thesis, Techn. Univ. Vienna 1985

/4/ V. Grewal, A Gschwandtner and G. Higelin, V-MIC 1986 Conf. Proc. IEEE, pp. 107

/5/ G. Higelin, C. Wieczorek, V. Grewal, V-MIC 1986 Conf. Proc. IEEE, pp. 443

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CONTROL OF A SELF-ALIGNED W SILICIDE PROCESS BY ANNEALING AMBIENCE

J. TORRES, J. PALLEAU, N. BOURHILA, J.C. OBERLIN, A. DENEUVILLE* and M. BENYAHIA*

CNET, PB 98, F-38243 Meylan Cedex, France
**CNRS, BP 166X, F-38042 Grenoble Cedex, France*

Résumé- Pour les technologies submicroniques,des étapes de metallisation autoalignées sont de plus en plus nécessaires . La formation autoalignée du siliciure de tungstene WSi2 par réaction entre W et Si a pu être maitrisée en réalisant des recuits sous oxygène et sous ammoniac . La formation latérale du siliciure est alors évitée même sur des espaceurs d'oxyde d'épaisseur 60 nm .

Abstract-With devices at submicronic dimensions , self-aligned processes are becoming of increasing interest . The self aligned W silicide formation by localised reaction between W and Si has been controlled using reactive atmosphere , oxygen and ammonia , during furnace annealing . No lateral silicide formation over SiO2 spacers,only 60 nm thick,has been observed .

1-INTRODUCTION

Refractory metal silicides are more and more widely used in semiconductor technology due to their resitivity as materials for gate interconnections, contacts on source and drain areas,and as a first buried metallic layer for local interconnections.For ultra large scale integration,with devices at submicronic dimensions, self aligned silicidation is becoming of increasing interest.The self aligned process can be achieved either by selective deposition of the silicide layers or by direct reaction, induced by thermal annealing ,between the refractory metal deposited over the whole wafer and the underlying silicon occurring only at localised points where these two materials come into contact .On the other hand the reaction between the refractory metal and oxidized silicon surface cannot take place.This latter approach is known as the salicide process.The critical step in such a scenario consists of the ability to well control the metal silicon reaction to avoid any lateral silicide formation over SiO_2 spacers, only a few thousand angstroms wide separating the silicide areas . Without this control bridging between the gates sources and drains could occur.

To make sure that the silicide process works properly the operating range for annealing time and temperature must be large enough to allow for control and reproducibility of the process. For this it appears necessary to slow down the silicide growth kinetics by inducing, at the surface of the metal layer, a second reaction between the metal and a reactive annealing atmosphere .Consequently a stable state of equilibrium will be reached between the metal silicide phase and the metal compound formed with the reactive atmosphere.Titanium is the refractory metal by far the most widely accepted for this technology /1/.The $TiSi_2$ growth kinetics has been found to be blocked by the formation of $TiNi_x$ compounds during thermal annealing under an N_2 flow.This paper presents recent decisive progress towards the control of the self aligned W silicide formation .

2-EXPERIMENTAL

Up to now silicide technology has not worked using W.It is difficult to achieve on the W/Si reaction due to the hight sensitivity of the reaction to small amounts of oxygen both at the W/Si interface and trapped in the W layer during metal deposition/2/ .Moreover tungsten does not react with N_2.On the contrary W reacts with oxygen,but in this case the reaction occurs so fast that tungsten is completely oxidized before the silicide reaction even begins .Recently W nitridation has been studied using ammonia .The WN_x compounds growth kinetics start at 550°C but seem to occur too fast compared with the W silicide formation which is generally blocked.

To overcome these difficuties encountered when using oxygen or NH_3 as the reactive annealing atmosphere the following solution has been envisaged.The WO_x and WN_x growth kinetics can be controlled by limiting the amount of reactive gas (oxygen or ammonia) in the annealing atmosphere.Consequently these reactions can be slowed down enough to allow W silicide formation at the W/Si interface .By varying the oxygen (or NH_3) partial pressure in the annealing atmosphere it is perfectly possible to control the state of equilibrium between the WO_x (or WN_x) phase and the WSi_2 phase at a fixed annealing temperature.On the other hand,when the oxygen (or ammonia) partial pressure has been fixed ,it is perfectly possible to control the reacted silicide thickness by varying the annealing temperature.Thus the W silicide process can be optimized by adjusting two free parameters :the partial pressure of the reactive atmosphere and the annealing temperature.For the experimental study silicon wafers with 600A SiO_2 on the top were used.By reactive ion etching (RIE) two reverse patterns were obtained:either 40um wide silicon lines with 4um wide SiO2 spacers or4um wide Si lines with 40um wide SiO_2 spacers.Over the whole patterned wafer a 50nm thick W layer was sputtered. Residual vacuum values of a few 10^{-8}T were reached in the sputtering system before starting the W deposition .The W silicon reaction was first studied on unpatterned Si wafers by using argon during thermal annealing . At 700°C the reaction is complete only for annealing times of about one hour ,but at 750°C the annealing time is reduced to less than a quarter of one hour. Annealings under reactive atmospheres of oxygen and ammonia,were then performed using patterned and unpatterned samples at annealing temperatures in the 700°C,1000°C range and for annealing times much greater than necessary for a complete reaction of the W layer with the silicon.The patterned samples were observed by SEM and X ray diffraction ,the unpatterned samples by RBS, nuclear reaction and X ray diffraction .

SEM observations

The photographs shown in Fig 1 and 2 correspond respectively to thermal annealing with oxygen and ammonia. As the SiO_2 spacers are only 60nm thick any lateral silicide formation of this order of magnitude over such spacers can easily be observed . Figure 1a shows the silicide grown at 850°C for 4mn, by using dilute oxygen and after chemical etching of WO_x.The silicide is perfectly aligned on the silicon line and no lateral silicide formation over the SiO_2 surfaces is observed. Fig 1b shows the edge of a SiO_2 line at the boundary with a silicide area.The salicide process has been well controlled. Fig 2a shows cross sections of patterned samples annealed at 850°C for 4mn under dilute NH_3.The view in Fig 2a has been taken before chemical etching of WN_x formed on SiO_2.The difference in the thickness of the deposited layer either over the SiO_2 surface or over the Si surface underlines the formation of two different phases over these two different surfaces. The photograph of Fig 2b shows the patterns after chemical etching of WN_x.Here also the W salicide process has been well controlled.

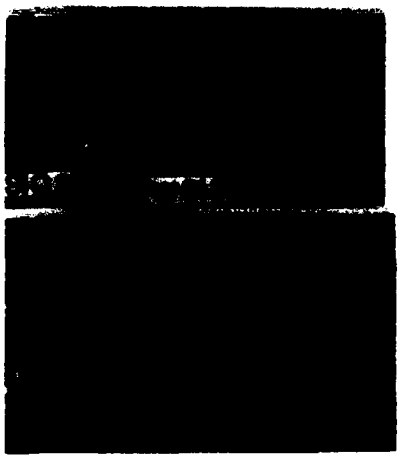


Fig.1-SEM observation of self aligned silicide grown at 850°C,4 mn, under oxygen
a-plane view
b-edge of a SiO2 spacer ,60 nm thick, at the boundary with a silicide area

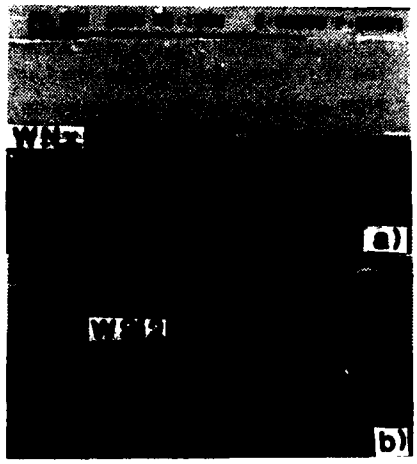


Fig.2-SEM cross sections of patterned samples annealed at 850°C, 4mn, under NH3
a-before chemical etching
b-after chemical etching

R.B.S and nuclear reaction

Annealing under NH_3 at atmospheric pressure (Fig 3):

In Fig 3a the squares correspond to thermal annealing at 700°C for one hour.Tungsten and silicon have not reacted .The corresponding nitrogen profile in Fig 3b obtained using the nuclear reaction $^{14}N(d,\alpha)^{12}C$ shows two peaks localised at the free surfaces of the sample (4E16 at/cm^2 dose) and the W/Si interface (3.1E16 at/cm^2 dose).

The crosses in Fig 3a are obtained when annealing at 750°C for one hour. A $WN_{0.7}$ compound has been found on the top of the layer, followed by a thick sublayer of unreacted W. The corresponding N profile shows a peak (4E16 at/cm2) localised at the W/Si interface.By increasing the temperatures for isochronal annealings the amount of N into W increases slowly.Finally a homogeneous WN stoichiometry is found when annealing at 900°C and above.From RBS and nuclear reaction, observations two important conclusions can be reached

1)N diffuses through the layer without precipitate WN_x compounds for annealing temperatures up to 750°C.As N solubility in W is very low, the N diffusion must occur at the W grain boundaries .The driving force could be the formation at the Si surface of SiN_x compounds thermodynamically favorable with regard to WN_x compounds.

2)The growth kinetics rate for the WN_x phases are low compared with the WSi_2 kinetics

3)As a result of the preceding remarks,it appears to be the rapid accumulation of N at the W/Si interface which blocks the silicon diffusion and impedes the WSi_2 formation.

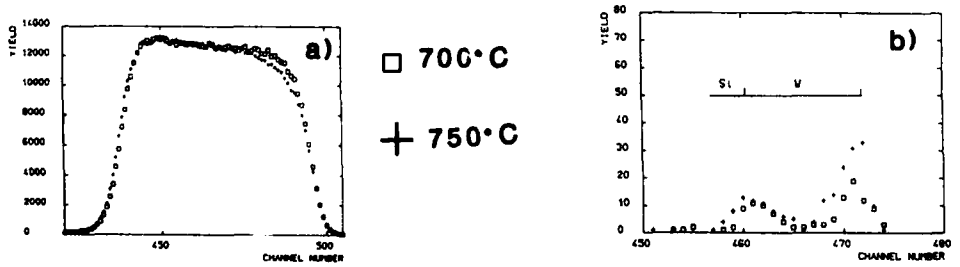


Fig.3 a-RBS spectra of W after annealing under atmospheric pressure
b-corresponding N profiles by nuclear reaction

Annealing under oxygen atmospheric pressure:

The W is completely oxidized and forms the WO_3 phase for annealing temperatures above 500°C .The RBS spectrum in Fig 4 corresponds to thermal annealing at 400°C for one hour.The corresponding oxygen profile has been obtained using the nuclear reaction ($^{16}O(d,\alpha)^{14}N$).A WO_3 layer is found at the free surface covering an oxygen free unreacted W layer .Moreover,not any oxygen peak at the W/Si interface has been observed.The W behaviour with oxygen is found to be very different from the behaviour found with N :

1)Free oxygen does not diffuse through the W layer to accumulate at the W/Si interface

2)The WO_3 phase forms at temperatures as low as 400°C,and its growth kinetics is very fast compared with the WSi_2 growth kinetics

3)As a result,it is the rapid formation of WO_3 which impedes the formation of WSi_2.

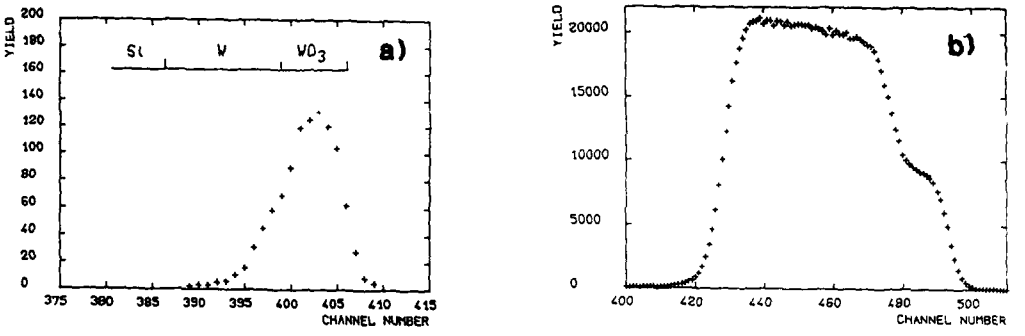


Fig.4 a-RBS spectrum of W after annealing under oxygen at atmospheric pressure
b-corresponding O profile obtained by nuclear reaction

X ray diffraction measurements and resistivity measurements have also been performed and the results confirm the RBS observations.

3-CONCLUSION

Self aligned W silicide formation has been controlled by using oxygen and ammonia as reactive annealing atmosphere . After furnace thermal annealing , W nitride and W oxyde can be selectively removed using a chemical etching which leaves only the WSi2 phase . Moreover , with respect to important processing aspects -dopants redristibution between WSi2 and poly or monocrystalline Si allowing to form very shallow junctions below silicided contacts , compatibility of the silicide with a first W C.V.D. interconnection level/3/- WSi2 presents a very interesting behaviour making it very attractive for self aligned silicide processes .

REFERENCES

/1/ M. E. Alperin, T. C. Hollaway, R.A.Haken, C.D.Gosmeyer, R.V.Karnaugh and W.D.Parmantie, IEEE Trans .Electron devices ED_32,141,1985

/2/G.Bomchil, G.Goltz and J.Torres, Thin Solid Films, 59, 1986

/3/F.C.Shone, K.C.Saraswat and J.D.Plummer, IEDM Technical Digest, 407, 1985

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

MODELLING DIFFUSION IN SILICIDES

P.B. MOYNAGH, A.A. BROWN and P.J. ROSSER

STC Technology Ltd., London Road, Harlow, GB-Essex CM17 9NA, Great-Britain

Abstract - This paper outlines models for the redistribution of dopant during the incorporation of silicides in VLSI processes.

1. INTRODUCTION

As device geometries are reduced the maximum speed of a circuit is governed increasingly by the RC constant of the interconnect. Refractory metal silicides are therefore receiving much attention as highly conducting additions to the polysilicon interconnects and source/drain regions of VLSI devices /1/. The silicon consumed during silicide formation is generally highly doped and it is important to understand and to model the resulting dopant redistribution. Silicide formation is also known to modify dopant diffusion by injection or depletion of point-defects in the underlying silicon /2/. The modelling of this silicide formation and the corresponding dopant redistribution within and below the silicide layer is discussed in section 3.

Silicides can also be used as a diffusion source to generate sub-100 nm junctions /3/. The silicide is formed and implanted and a subsequent anneal causes the implanted dopant to diffuse through the silicide and into the silicon substrate. Dopant redistribution, therefore, takes place with minimal movement of the silicide-silicon interface, and the junctions can be created with high surface concentrations and corresponding low contact resistivities. It is important to understand and model this process step if it is to be exploited in future VLSI processes. The modelling of dopant redistribution within and below these pre-formed silicide layers is discussed in section 2.

2. OUT DIFFUSION OF DOPANT FROM SILICIDE AND THE STATIC MODEL

Table 1 outlines an experiment to examine out-diffusion of the common dopants from cobalt silicide. Thick (0.55 μm) silicide layers were formed to allow dopant profile evolution within the silicide to be examined. Boron, phosphorus or arsenic were implanted to a range of about 75 nm and the films were then annealed for 30 seconds at 700°C or 1000°C. The electrical characterisation data from the resulting p+/n or n+/p junctions are given in Table 2. Good diode characteristics were not achieved with similar anneals using titanium silicide as the diffusion source. Scanning Electron Microscopy (SEM) analysis of the cobalt silicide films after formation and after subsequent anneals showed the silicide-silicon interface to undulate by 0.15 μm. Figure 1(a) shows a boron implanted cobalt silicide film after anneal at 1000°C for 30 seconds. The interface undulation is apparent, as is the silicide grain structure. Figure 1(b) shows the same sample after a junction delineation etch and the junction can be seen to follow the silicide-silicon interface. The dopant profiles were measured by Secondary Ion Mass Spectroscopy (SIMS), the layer conductivity by Spreading Resistance Analysis (SRA) and the interface position and junction depth were determined by Bevel and

Stain (B&S). The results are shown in Figure 2. The high SRA concentration within the silicide indicates high conductivity. The peak in the SRA profile corresponds to the point at which the analysis probes first reach the undulating silicide-silicon interface and the deeper inflection point is where the probes leave the interface. The flat boron concentration in the silicide bulk shown by the SIMS indicates very rapid diffusion to the silicon surface and a junction depth of 0.17 μm results.

Dopant diffusion from a pre-formed silicide was modelled by using the one-dimensional process simulator, SUPREM 3 /4/, with the silicide implemented as a user-defined materal. Figure 3 shows the simulated profile corresponding to the data of Figure 2. Simulation of the rapid diffusion in the silicide is achieved by using a diffusion coefficient for boron of 10^3 to 10^4 times the silicon intrinsic value. Similar values were required for arsenic and phosphorus diffusion in the silicide. This illustrates the magnitude and importance of grain boundary diffusion in cobalt silicide. A silicide-silicon segregation coefficient of unity gives good agreement between empirical and modelled values of silicon surface concentration, highlighting the suitability of cobalt silicide for the formation of juntions with low contact resistance. Also shown in Figure 3 is the predicted boron profile in the silicon substrate using the default diffusivity and using modified diffusion coefficients. An eight-fold enhancement of boron diffusivity is required for agreement with experiment. Arsenic diffusivity from the silicide into the substrate showed a two orders of magnitude enhancement. These diffusivity enhancements beneath the cobalt silicide are consistent with considerable vacancy injection by the silicide.

3. SILICIDE FORMATION ON DOPED SILICON AND THE DYNAMIC MODEL

Figure 4(a) shows an Auger Electron Spectroscopy (AES) analysis of a sputtered layer of titanium metal on arsenic implanted single crystal silicon. Figure 4(b) shows the same sample after titanium silicide formation. Layer evolution during the formation anneal is apparent, as is considerable dopant redistribution. Figure 5(a) shows an AES of a titanium silicide film formed over phosphorus doped polysilicon. A layered structure and considerable dopant movement is again observed.

The importance of modelling dopant redistribution during silicide formation is apparent from figures 4 and 5. For the static model, the movement of dopant to and across static boundaries was readily modelled. For the more general case, however, where dopant diffusion and segregation coincides with layer formation and boundary movement, a more complex modelling capability is required. A qualitative model for layer evolution during titanium silicide formation is shown in Figure 6. The deposited titanium layer readily absorbs oxygen when exposed to air. On annealing at 650°C, titanium and silicon react to commence silicide formation. Silicon diffuses through the existing silicide and reacts with the metal layer to continue the silicide formation, and the silicide-metal and silicide-silicon interfaces move. The oxygen present in the titanium layer is not highly soluble in the silicide and is 'snow-ploughed' into the titanium ahead of the advancing interface. Nitrogen is incorporated from the anneal ambient, creating a titanium nitride layer at the metal surface.

Figure 7 shows the layer evolution and the corresponding redistribution of arsenic or phosphorus during a self-aligned silicide process step. As the silicide forms, the dopant diffuses rapidly through it, segregates preferentially to the metal layer and some dopant segregates to the capping nitride layer where diffusivity is low. The capping nitride and unreacted metal are subsequently stripped. Also shown in Figure 7 is a self-aligned nitride formation at 1100°C. This results in the displacement of silicon and dopant by nitrogen at the silicide surface as the titanium nitride forms. Much of the displaced material diffuses through the silicide and is epitaxially redeposited on the substrate.

A preliminary investigation of the implications of this qualitative model was undertaken using a modification of the oxidation model in SUPREM 3. Silicide-silicon and silicide-ambient segregration coefficients were varied as was the diffusion coefficient and silicide-silicon boundary velocity. The low dopant diffusivity in the capping titanium nitride layer was found to limit the flux of dopant from silicon substrate to anneal ambient. The titanium nitride therefore plays an important role in minimising dopant loss during the silicidation of highly doped silicon.

A fully characterised moving boundary model for silicides must be implemented in a process modelling package like SUPREM 3 if the use of silicides in VLSI processes is to be optimised.

4. CONCLUSIONS

As silicides are incorporated into VLSI processes, it is becoming increasingly important to model their effect on dopant redistribution. Two initial models have been presented here which address this problem. The static model has been shown to give results which compare favourably with experimental data. A qualitative dynamic model is outlined which when implemented will provide a useful tool for determining the effect of silicides on VLSI processes.

REFERENCES

/1/ Rosser, P.J. and Tomkins, G., Vacuum, 35, (1985), 419-434.

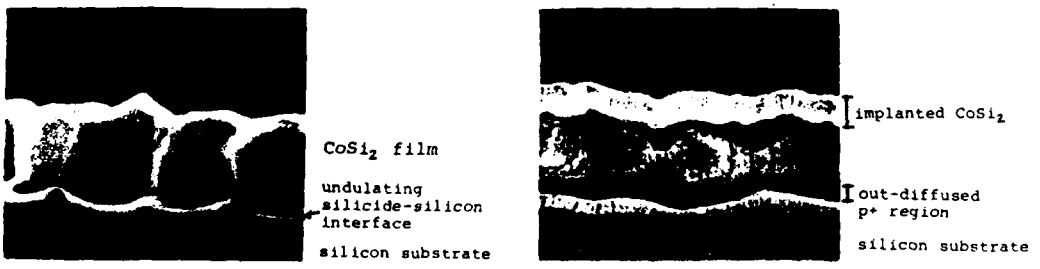
/2/ Wen, D.S., Smith, P.L., Osburn, C.M. and Rozgoyni, G.A., Appl. Phys. Lett., 51, (1987), 1184-1192.

/3/ Kowng, D.L., Ku, Y.H., Lee, S.K., Louis, E., Alvi, N.S. and Chu, P., J. Appl. Phys., 61, (1987), 5084-5088.

/4/ Stanford University/Technology Modelling Associates, SUPREM 3 process model.

ACKNOWLEDGEMENTS

This work was part sponsored by the Alvey Directorate under contracts ALV/APP/VLSI/005, 016, 066



(a) cleave & SEM

(b) cleave, delineation etch & SEM

Fig.1 Cobalt silicide film (0.55μm) implanted with boron ($1x10^{16}$ cm^{-2}, 25keV) and annealed (1000°C, 30secs) to form a silicided p+/n junction

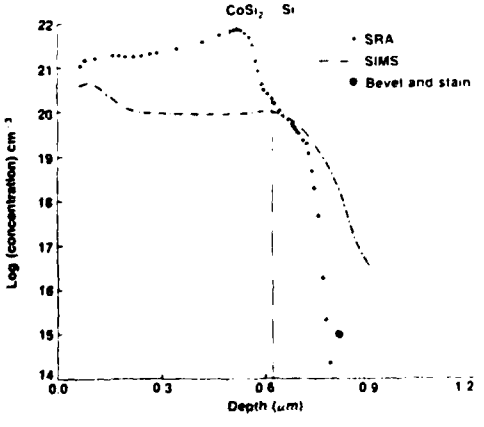


Fig.2 Experimental boron diffusion from a $CoSi_2$ source anneal: 1000 C, 30 secs

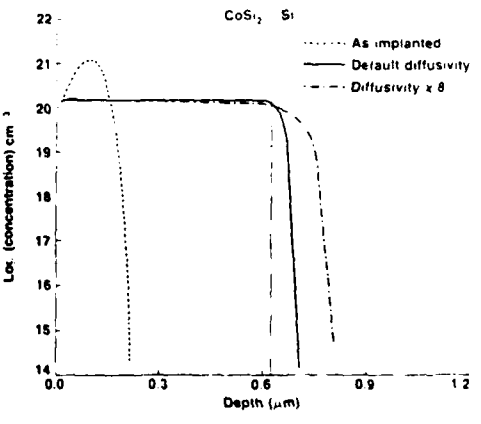


Fig. 3 Simulated boron diffusion from a $CoSi_2$ source anneal: 1000 C, 30 secs

Metal deposition:	Co
Silicide formed:	700C
Silicide implant:	B, P, As to 1E16cm^{-2}
Silicide anneal:	700C, 1000C, 30 SECS
Electrical char:	ideality, leakage
Material char:	SEM SIMS SRA B & S

Table 1. Experimental procedure for out-diffusion of dopant from silicide

$CoSi_2$

	Anneal temp	Ideality	Leakage
Boron	700C	< 1.1	good
	1000C	1.3	very good
Phosphorus	700C	1.5	OK
	1000C	2.3	good
Arsenic	700C	< 1.1	OK
	1000C	1.7	good

Table 2. Electrical characteristics of diodes formed by out-diffusion of dopant from silicide for 30 secs

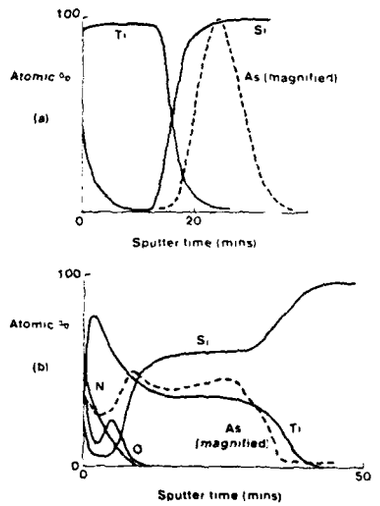


Fig. 4 AES of arsenic (a) before $TiSi_2$ formation (b) after $TiSi_2$ formation

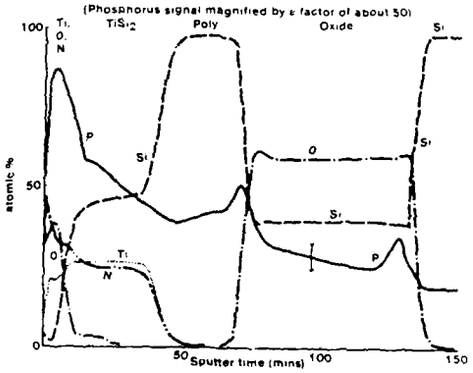


Fig. 5 AES analysis of a titanium film deposited over phosphorus doped polysilicon and annealed

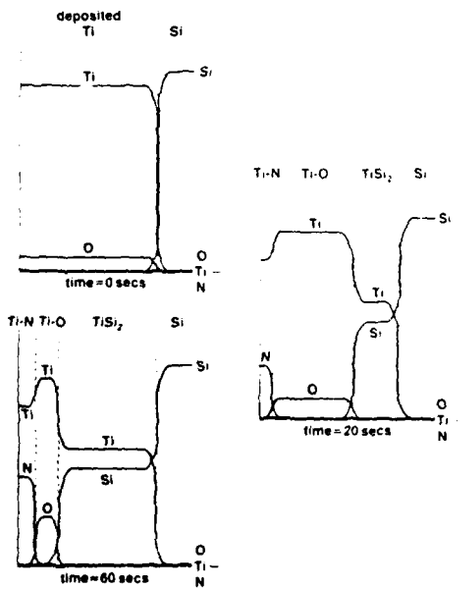


Fig. 6 Layer evolution during titanium silicide formation at 650 C

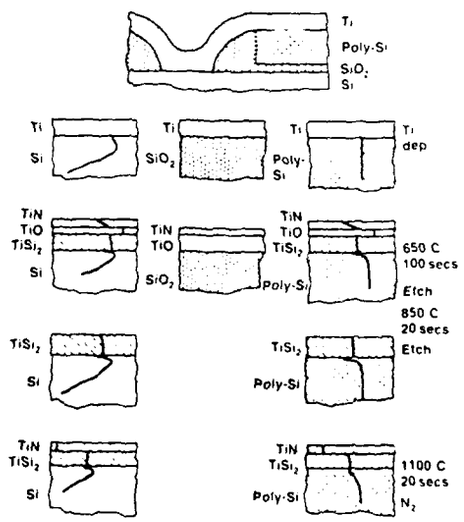


Fig 7 Structure of silicide during formation

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

DEGRADATION OF THE POLY-Si/SILICIDE STRUCTURE IN ADVANCED MOS-TECHNOLOGIES

P. LIPPENS, K. MAEX, L. VAN den HOVE, R. DE KEERSMAECKER, V. PROBST*, W. KOPPENOL** and W. van der WEG**

Interuniversity Microelectronics Center (IMEC vzw), Kapeldreef 75, B-3030 Heverlee, Belgium
**Siemens AG, Central Research and Development, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.*
***Department of Atomic and Interface Physics, Utrecht State University, PO Box 80000, NL-3508 TA Utrecht, The Netherlands*

Résumé. L'interaction dans le système poly-Si sur $TiSi_2$ sur un substrat de Si et dans le système polycide avec $CoSi_2$ ainsi qu'avec $TiSi_2$ est analysée quand ils sont soumis à une temperature élevée dans un four classique. On démontrera que le premier système souffre d'instabilité due à la croissance epitaxiale dans la phase solide qui peut seulement être évitée partiellement par le dopage de la couche poly-Si. Le système polycide montre une couche de $TiSi_2$ qui est devenue rude et qui pourrait poser des problèmes pour des couches poly-Si initiellement plus minces que 450 nm. Dans le cas de $CoSi_2$, la dégradation est désastreuse: le métal Co peut être détecté partout dans la couche poly-Si, ce qui est à peine évitable même par un dopage du $CoSi_2$ avec As ou B.

Abstract. The interaction in the system poly-Si on $TiSi_2$ on Si and in the polycide system both with $CoSi_2$ and $TiSi_2$ is investigated when subjected to furnace heat treatments at high temperature. It will be shown that the first system suffers from instability due to SPE-regrowth which is partially avoided by doping the poly-Si. The polycide system shows severe roughening of the $TiSi_2$. However, in the case of $CoSi_2$, the degradation is more pronounced, viz. the Co is detected all over the poly-Si layer and this can hardly be suppressed by doping of the $CoSi_2$-layer.

Introduction

In this work, the stability of layered silicide/poly-Si structures is investigated when subjected to high temperatures. These high temperatures are used in a full process to perform a glass reflow and, very recently, to diffuse the dopants into the Si using an implanted silicide as a diffusion source [1], [2], [3]. This investigation deals both with the system poly-Si on $TiSi_2$ on a Si-substrate and with the polycide-system (a $CoSi_2$ or $TiSi_2$ layer on a poly-Si-gate). The first structure has recently been proposed for application in an n-MOS process with $CoSi_2$ in which source and drain regions are formed by diffusion from a dopant gas ambient through a 'poly-plug' on the silicide in the contact holes [4]. The latter structure occurs in a self-aligned process, for instance when dopants are diffused from an ion implanted silicide for source and drain formation. In such a case, it would be very attractive to dope the poly-Si gate also by diffusion from the implanted silicide-layer.

Experimental

A $TiSi_2$-layer of approx. 90 nm thick was formed on Si-substrates of <100> orientation. The native oxide was removed by a 2% HF dip immediately before sputtering the metal. Then 100 nm of poly-Si were CVD-deposited at 600°C and were either left unimplanted or implanted with As or B to a dose of 5.E15 cm^{-2}. To prevent dopant evaporation during heat treatments, a cap layer of 100 nm CVD-SiO_2 was deposited. Finally, furnace heat treatments were performed at 800°C (2 h) or at 900°C (30 min and 2 h) in an N_2 ambient. This set of samples will further be referred to as the poly-plug-system. In another experiment a gate oxide of 20 nm was grown followed by deposition of 450 nm of poly-Si at 600°C. Then 90 nm of $CoSi_2$ or $TiSi_2$ was formed which was also either left unimplanted or implanted (B or As) and capped with 100 nm of CVD-SiO_2. Also here a native oxide removal in 2 % HF was carried out. Again heat treatments were performed at 700°C during 4 h, at 800°C during 2 h or at 900°C during 1 h in an N_2 ambient. This system will be referred to as the polycide structure. Both sets of samples were analysed by RBS, AES, cross-sectional SEM (XSEM) and TEM.

Results

Figure 1 shows 2 MeV $^4He^+$ RBS spectra of the unimplanted poly-plug structure with $TiSi_2$ that was heat treated at different temperatures. It is clear that after an 800°C, 2 h step, the RBS spectrum coincides completely with the one for the unheated sample which indicates that no interaction between the layers has taken place. But after a 900°C, 2 h treatment, the Ti-peak as well as the edge from the Si-substrate have shifted to higher backscattering energies indicating that the Si/silicide interface has moved towards the surface. The same behaviour was seen in a sample heated also at 900°C but only during 30 min: in this case, however, the Ti-signal is broader and much less 'block-shaped' and the Si-substrate edge has not moved that far towards higher backscattering energies. SEM-micrographs from cross-sections of the samples which were selectively etched show 3 layers on the Si-substrate after deposition. The sample which received a 900°C, 2h heat treatment on the other hand reveals the $TiSi_2$-layer directly under the cap-SiO_2 (fig. 2). This suggests a total transport of the poly-Si layer through the $TiSi_2$ and a solid phase epitaxial regrowth of this Si on the substrate. High resolution TEM-micrographs taken in the region of the interface silicide/Si-substrate revealed a perfect Si-lattice pattern.

It is also interesting to note that on the SEM-micrograph of fig. 2, the $TiSi_2$ is no longer a continuous layer, which could be expected after the heat treatment mentioned: an analogous behaviour was in fact detected for the system where $TiSi_2$ was grown on a Si-substrate (and capped with an SiO_2-layer) which was subjected to a high temperature for an extended time [5]. The Si from the substrate regrows epitaxially in between the silicide grains by the mechanism of thermal grooving, resulting in Si-mesas. In the case proposed here, this tendency seems to be enhanced because Si is supplied now from two directions (i. e. from the substrate and from the poly-Si layer). On the other hand, once a thermal groove has been formed the transport of the Si from the poly-Si layer towards the Si-substrate proceeds very quickly. This means that both epitaxial regrowth mechanisms strengthen each other.

The role of the dopants was investigated in the case of B and As. For As-implanted poly-Si and for various heat treatments at 900°C, the RBS-spectra indicate only a small broadening of the Ti-peak (fig. 3) due to increased $TiSi_2$ roughness (ball-up), which is considered to be the precursor stage of SPE-regrowth of the Si-substrate in between the $TiSi_2$-grains. Therefore, an attempt was made to show that even in this case, there is a transport of Si from the poly-Si layer into and through the $TiSi_2$. Samples were prepared with thicker silicide layers (300 nm) and a very fine-grained poly-Si layer to a total thickness of 200 nm was grown on top of them by a furnace crystallisation step during 10 min at 650°C of sputtered Si. The As-implantation was done in the sputtered Si film to a dose of 2.E16 cm^{-2}. The resulting RBS-profiles for heat treatments at 800°C and 900°C during 60 min, depicted in fig. 4, indicate that the backscattering yield of the Ti, resp. the Si in the $TiSi_2$ has decreased, resp. increased which points again to an interaction between the layers, though in an early stage. It can be concluded that there is a dopant-induced retardation effect and this was supported also by the results for B-doping of the poly-Si. Figures 5 a and b show TEM-micrographs of samples with B doping for the poly-Si layer, which were unheated or annealed at 900°C for 2 h, respectively. Figure 5 b shows that on some places the poly-Si has regrown almost completely on the Si-substrate, whereas on other places some Si is still left on top of the $TiSi_2$.

In the polycide-system, no drastic interaction of the poly-Si and the silicide layer is expected due to SPE-regrowth. This is indeed observed for $TiSi_2$ on poly-Si, where only a ball-up effect of the $TiSi_2$ is seen after a 900°C, 1h heat treatment (fig. 6). Nevertheless, this could be serious since it gives rise to a silicide layer with disconnected grains and to places where the remaining poly-Si layer between gate-oxide and $TiSi_2$-grains is very thin. For thinner poly-Si layers this could even lead to places where the $TiSi_2$ directly contacts the gate-oxide resulting in low oxide breakdown voltages and at least flat band voltage shifts.

However, in the case of $CoSi_2$ already an 800°C heat treatment for 2 h causes drastic effects, viz. the Co atoms are distributed over the entire poly-Si layer and no clear $CoSi_2$-phase can any longer be recognized in the RBS-profile of fig.7. A heat treatment at 700°C, even during 4 h did not show any degradation of the system except for a small roughening of the silicide layer. When the $CoSi_2$ is implanted with As (to a dose of 5.E15 cm^{-2}), it was seen that the Co is still moving over the poly-Si, although the As must have diffused partially into the poly-Si and stuffed the grain boundaries of both silicide and poly-Si layers in the first stage of the heat treatment (the RBS profile for As-doping is also included in fig. 7). A 900°C, 1 h heat treatment yields more or less the same result, which is also the case for B-doping of the silicide. These experiments indicate that the diffusion of dopants from a $CoSi_2$ layer in order to dope the underlying undoped poly-Si will not be possible for MOS-technologies. Further experiments are being processed to investigate the stability of the silicide on predoped poly-Si (either P solid source doping or implantation combined with a drive-in step). However, the first results do not indicate a drastic improvement, which was again to be expected from the results with As-diffusion from the $CoSi_2$.

Conclusion

It is concluded that the poly-plug system with $TiSi_2$ suffers from spatial instability due to SPE-regrowth when heated above 800°C which causes the poly-Si layer on top of the silicide to regrow on the substrate and moreover the $TiSi_2$ to ball-up while leaving Si-mesas in between the grains. The dopants can only avoid to

some extent this regrowth: even when large As-atoms are used to stuff the grain boundaries of the system the Si-content in the $TiSi_2$ increases considerably. This is in good agreement with the results from Murarka et al. [6] for the case of $CoSi_2$ and in-situ doped poly-Si and with the experiments of Lau and van der Weg for many other systems [7] . The polycide structure, subjected to a furnace heat treatment, shows a severe degradation both with $TiSi_2$ and $CoSi_2$. In the latter case, Co is detected all over the poly-Si layer, though X-Ray Diffraction spectra reveal that only poly-Si and $CoSi_2$ phases are present in the system. These results suggest that either $CoSi_2$ migrates as a whole all over the poly-Si layer or that $CoSi_2$ breaks up during the heat treatment, the metal diffuses all over the poly-Si layer and forms again $CoSi_2$-grains, which is in agreement with theories about the main moving species during silicide formation [8], [9]. In the former case an extensive epitaxial regrowth of the poly-Si grains in between the $TiSi_2$ grains is observed resulting in a poly-Si layer which is locally very thin under the balled-up silicide. In further experiments, the use of rapid thermal processing (RTP) will be investigated to avoid these poly-Si/silicide interactions.

Acknowledgements.
The authors would like to thank J. Vanhellemont for the TEM work and R. Wolters (Philips, Eindhoven) for access to X-Ray Diffraction analyses. P. Lippens is indebted to the Belgian Institute for Scientific Research in Industry and Agriculture (IWONL). K. Maex and L. Van den hove acknowledge the Belgian National Fund for Scientific Research (NFWO) for their fellowships as senior research assistants.

References.

[1] V. Probst et al., Proc. of the 17 th ESSDERC, Bologna, Sept. 14-17, 1987, p. 437
[2] R. Liu et al., Proc. Electrochem. Soc., Symposium on ULSI Science and Technology, May 10-15, 1987, p. 446
[3] P. Lippens et al., presented at the 6 th International Conference on Ion Beam Modification of Materials, Tokyo, Japan, june 12-17, 1988
[4] S. Vaidya et al., IEEE Trans. on El. Dev., Vol. ED-33, Nr.9 (1986), p. 1321
[5] L. Van den hove, Ph. D. Thesis Katholieke Universiteit Leuven, Belgium, p. 154
[6] S. Murarka et al., J. Vac Sci. Technol. B 5 (4), 1987, p. 865
[7] S. S. Lau and W. F. van der Weg, Thin Film Interdiffusion and Reactions, ch. 12, ed. by J. M. Poate, K. N. Tu and J. W. Mayer, J. Wiley, New York, 1978, p. 433
[8] G. J. Van Gurp et al., J. Appl. Phys., Vol. 49 (7), 1978, p. 4011
[9] F. M. d'Heurle et al., Thin Solid Films, Vol. 128, 1985, p. 283

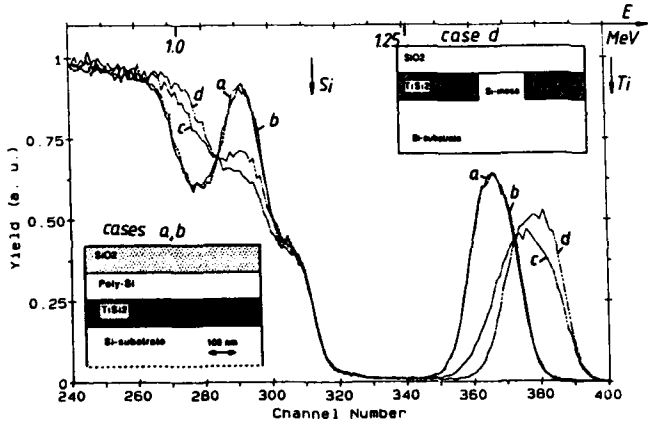


Fig. 1. RBS-spectra of the indicated layer structure (inset) for different furnace heat treatments in N_2. a.: Reference b.: 800°C, 2 h. c.: 900°C, 30 min. d.: 900°C, 2 h.

Fig. 2. Cross-sectional SEM-micrograph of the layer structure of fig. 1 subjected to a furnace heat treatment at 900°C for 2 h in N_2.

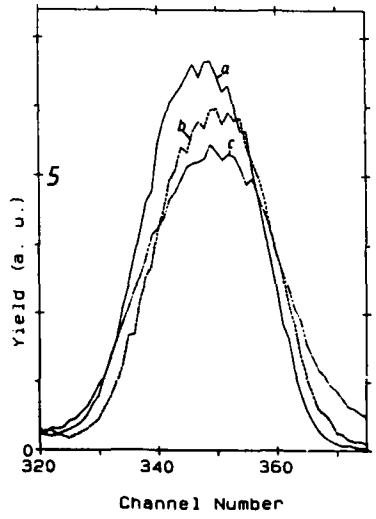


Fig. 3. RBS spectra of the Ti-peak in samples with As-doped poly-Si on $TiSi_2$ on Si (capped with 100 nm SiO_2) and treated at 900°C for 30 min (b), resp. 2 h (c). a is the reference sample.

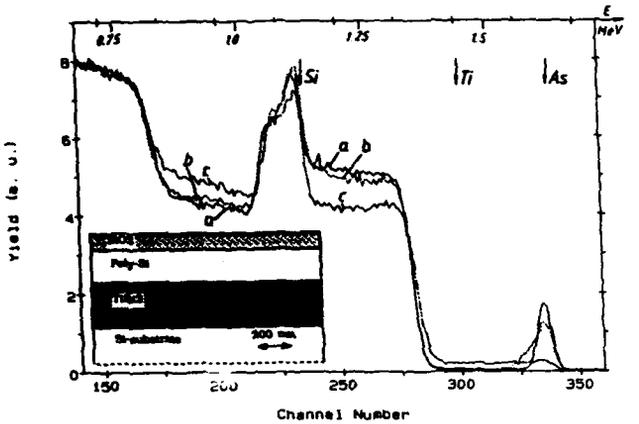


Fig. 4. RBS-spectra of the indicated layer structure (inset) in which the poly-Si was formed by crystallisation at 650°C during *10 min in N_2* followed by a second heat treatment during 60 min at 800°C (b), or at 900°C (c). Curve a is the reference.

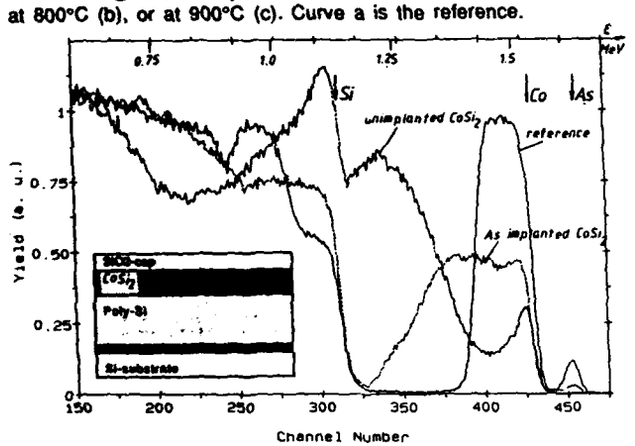


Fig. 7. RBS-spectra of the polycide-structure with $CoSi_2$ after an 800°C, 2 h heat treatment. The silicide is either undoped or As-doped. The spectrum for an unheated sample with As-implanted $CoSi_2$ is also included.

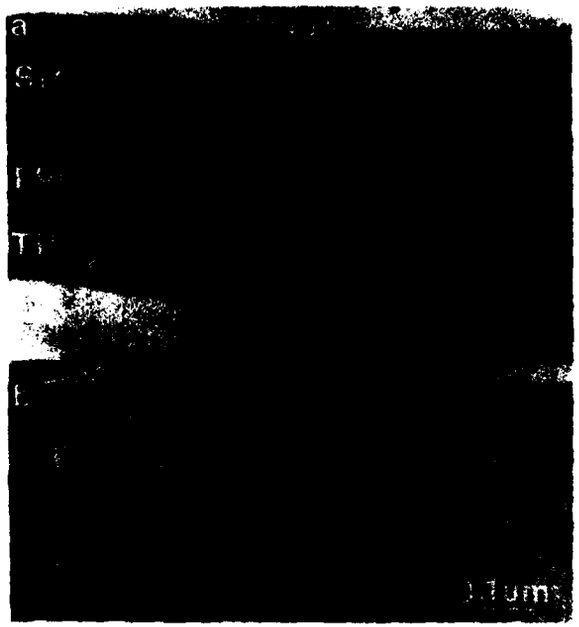


Fig. 5. TEM micrographs of B-doped poly-Si on $TiSi_2$ on Si (capped with SiO_2). a. reference. b. After a 900°C, 2 h heat treatment in N_2.

Fig. 6. Cross-sectional SEM-micrograph of the polycide structure with $TiSi_2$ and subjected to a 900°C, 1 h heat treatment.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A FULLY CHARACTERISED PROCESS FOR TITANIUM SILICIDE BY RTA FOR ONE MICRON CMOS

N.F. STOGDALE

Plessey Research Caswell Ltd., Caswell, Towcester, Northants., GB-NN12 8EQ, Great-Britain

Abstract A fully characterised process for self-aligned titanium silicide by RTA is described. Factors influencing formation; surface pre-clean, time-temperature schedules and ambient choice are all discussed. The sensitivity of the formation process to implanted arsenic is also described and a model presented for this effect. The technique is proven by its inclusion in a one-micron trench isolated CMOS process schedule and electrical results for devices and circuits thus fabricated are presented.

1. INTRODUCTION

Self-aligned titanium silicide formed on the gate and junction regions of small geometry CMOS devices gives enhanced device performance arising from the reduction of parasitic series resistances. The use of a Rapid Thermal Annealing (RTA) system allows tight control of ambient oxygen and greater process flexibility and control. A fully characterised process for self-aligned titanium silicidation by RTA is described, giving a silicide with good surface quality and low, uniform sheet resistivity. Factors influencing surface quality; surface pre-clean and ambient choice are described. Choice of thermal cycles, times and temperatures, for optimisation of the sheet resistivity and lateral encroachment are also discussed. The optimum film is analysed by SEM, SIMS and AES for surface quality, thickness and stochiometry. The sensitivity of the formation process to implanted dopant is described and a model presented for the sensitivity to implanted arsenic.

2. FACTORS INFLUENCING SILICIDE FORMATION

2.1 The Effect of Surface Pre-treatment

Silicide surface quality was found to be critically dependent on surface pre-clean. Several methods of pre-cleaning were evaluated: wet chemical, SF_6 plasma flash, HCL gas clean and 100:1 HF based spray clean. Sputter deposited titanium was used throughout and an in-situ RF (argon ion) sputter clean was also evaluated. The best silicide surface quality was achieved with the 100:1 HF spray clean, without a pre-deposition RF sputter clean.

2.2 The Effect of Ambient Choice

The choice of ambient; argon or nitrogen, was found to have a critical effect on silicide surface quality. Argon produced a poor quality 'broken' surface (sheet resistivity 3.3ohm/sq.), whilst nitrogen produced a good surface quality (sheet resistivity 4.5ohm/sq). Nitrogen forms a titanium nitride capping layer which protects the titanium from reaction with any ambient oxygen. The final silicide resistivity is increased by the consumption of titanium by this competitive TiN reaction.

2.3 RTA Time-Temperature Schedules

RTA time and temperature schedules over the range 600-1000°C, 10-60 seconds were evaluated for optimisation of sheet resistivity and lateral encroachment (see figure 1). The optimal process was found to be two-stage 675°C 30 seconds, followed by an unreacted titanium strip and a second stage anneal 1000°C, 10 seconds. Lateral encroachment of the silicide over surrounding oxide areas is inhibited by the use of a two stage process, no diffusion path exists at the high temperature anneal stage due to the unreacted titanium strip. An initial thickness of 500A titanium was used throughout.

3. ANALYSIS OF FILM PROPERTIES

Films were analysed for surface quality, sheet resistivity, thickness and stochiometry by SEM, 4-point probe, SIMS and AES. Surface quality was found to be of good, even texture. The optimal process gave a silicide with a sheet resistivity of 4.55ohm/sq (±-0.5). Thickness, determined from the silicon lattice transition in the SIMS profile (silicide to bulk), was found to be 680A (see figure 2). The film is found to be pure, stochiometric $TiSi_2$ (65% Si, 35% Ti) from AES. Trace elements carbon, oxgen and nitrogen were not detected at the 0.1% detection limit.

4. FORMATION SENSITIVITY TO IMPLANTED DOPANT

The silicide exhibited a formation sensitivity to implanted arsenic. Implant doses in the range $5E13cm^{-2}$ to $1E16cm^{-2}$ were evaluated, (implant energy 50KV). Silicide formation and sheet resistivity were seen to degrade with increasing arsenic dose. Below dose $1E14cm^{-2}$ the resistivity is a constant 4.55ohm/sq; above this dose the resistivity increases to 45ohm/sq. at dose $1E16cm^{-2}$ (see figure 3). No sensitivity to dopants boron, boron fluoride and phosphorus was indicated across the same dose range (see figure 4).

4.1 Model for Sensitivity to Implanted Arsenic

The sensitivity to implanted arsenic, for the activation conditions used, is found to be dependent upon excess dopant at the silicon surface. The excess dopant exists as interstitial or clustered arsenic and inhibits silicon diffusion across the interface. (The silicidation reaction is dominated by silicon diffusion across the silicon/titanium interface.) Excess dopant concentrations can be calculated from the SUMPREM3 model by subtracting active from chemical concentration at the surface. The thermal cycles assumed are two 1050°C 21 second dopant activation anneals in nitrogen plus the thermal cycles associated with the silicidation. A strong correlation is found between excess arsenic concentration and sheet resistivity (see figure 3). Excess dopant drops to zero at implanted dose $1.3E14cm^{-3}$, below this dose the sheet resistivity is a constant 4.55ohm/sq. Above this critical dose the excess dopant concentration increases to $3.3E20cm^{-3}$ at dose $1E16cm^{-2}$, with a consequent sheet resistivity of 45ohm/sq. Sensitivity to implanted arsenic can, hence, be controlled by adjustment of either the implant conditions or the activation anneals to assure zero excess dopant.

5. ELECTRICAL DATA

The optimised silicide process was fabricated on device structures on a one-micron trench-isolated CMOS process /1/. Polysilicon gates are fabricated with oxide sidewalls to allow fabrication of self-aligned silicided structures (see figure.5). A scanning electron micrograph of a fully silicided trench isolated CMOS transistor is shown in figure 7. The devices have good electrical performance, electrical parameters are shown in table 1. This technique has been used to fabricate working circuitry, in particular unloaded inverter ring oscillators with stage delays of 120ps.

6. CONCLUSION

A process for self aligned titanium silicide by RTA has been described. The silicide has good surface quality and is of low, uniform sheet resistivity. Sensitivity to implanted arsenic has been discussed and implant dose and activation conditions found to be critically important by its inclusion in a one-micron trench-isolated CMOS process schedule which has fabricated working devices and circuits.

7. ACKNOWLEDGEMENTS

This work has been carried out with the support of the Alvey Directorate and the Integrated Circuits Division, Plessey Research Caswell.

8. REFERENCES

/1/ Roberts, M.C. Bolbot, P.H. Foster, D.J. Proceedings this conf.

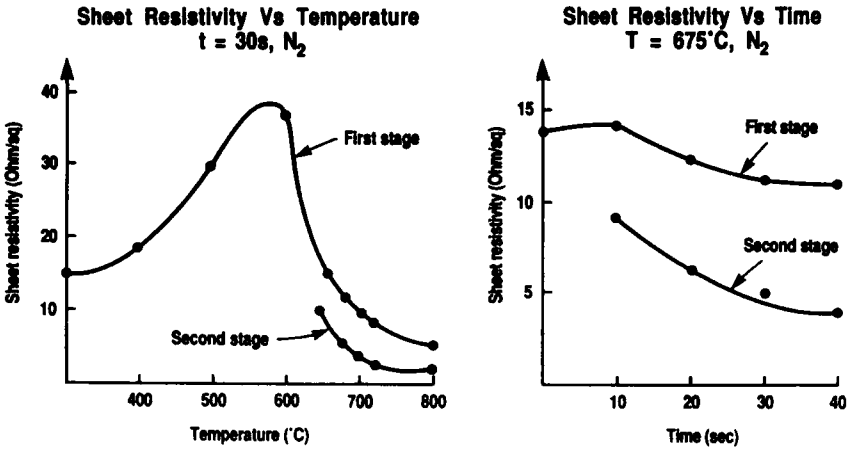


Figure 1: Sheet resistivity vs. RTA Time and Temperature.

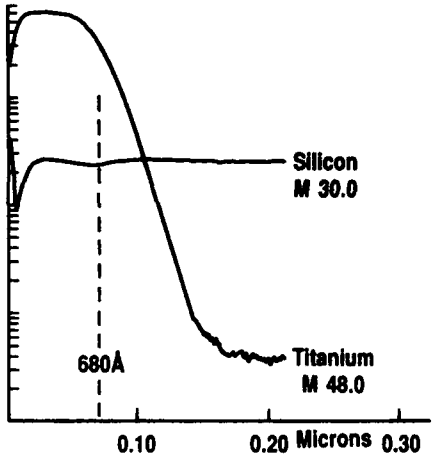


Figure 2: SIMS Profile showing lattice transition.

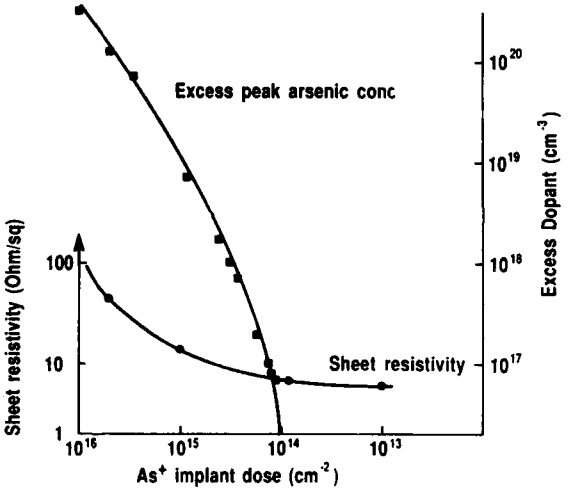


Figure 3: Sheet resistivity sensitivity to arsenic dose.

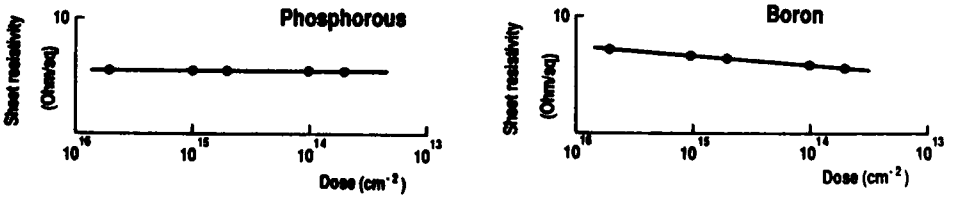


Figure 4: Sheet resistivity sensitivity to phosphorus or boron dose.

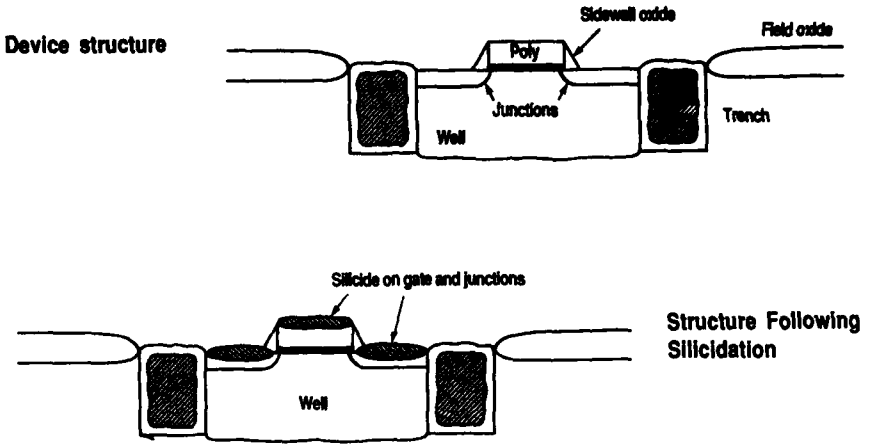


Figure 5: The Self-Aligned Silicide Process.

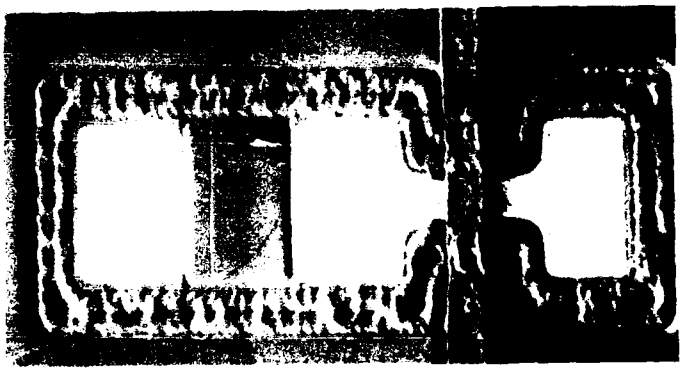


Figure 6: SEM micrograph of trench isolated CMOS transistor showing Self-Aligned Silicide.

	VT	SUB-SLOPE	DRIVE	GAIN
NMOS	0.7	95	5.5	85
PMOS	-0.7	-87	-2.7	-23
	VOLT	mV/DEC	mA	μA/V-2

RING OSCILLATOR (49 STAGE) SPEED:120ps.

Table 1: Electrical data for silicided transistors. (W=20 um, L=1um)

Session 2C Room C : Field effect transistors

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

OUTPUT IMPEDANCE FREQUENCY DISPERSION AND LOW FREQUENCY NOISE IN GaAs MESFETs

D. GITLIN, C.R. VISWANATHAN and A.A. ABIDI

Electrical Engineering Department, University of California, Los Angeles, CA 90024, U.S.A.

Abstract- **L'investigation traite de la relation entre bruit g-r et la dispersion de fréquence de l'impédance de sortie GaAs MESFET. On a employé une nouvelle technique pour l'identification des niveaux profonds qui sont responsables du bruit et de la dispersion. La dépendance du biais de la dispersion fut également investiguée et un modèle constant aux simulations en deux dimensions est presenté.**

The interrelation between g-r noise and output impedance frequency dispersion in GaAs MESFET's has been investigated. A new technique to identify the deep levels responsible for the dispersion is used. It is found that same traps are responsible for both noise and dispersion. The bias dependence of the dispersion was also investigated and a model is presented which is consistent with 2-D simulations.

1. Introduction

Low frequency anomalies in GaAs mesfets such as frequency dispersion in output conductance and transconductance, low frequency noise and backgating, have received considerable attention recently. Our experimental measurements are intended to clarify the origin and interrelation of these phenomena. This paper describes the results of experiments to relate the output impedance frequency dispersion and the low frequency noise in GaAs MESFET's. A new technique is used to identify the deep level responsible for output impedance dispersion. This technique consists of monitoring the phase of the output impedance vs frequency at various temperatures. We have also found that this same trap gives rise to g-r noise also.

2. Experiment

Depletion MESFET arrays, fabricated with a recessed gate structure on LEC grown substrates, and with nitride passivation were used in the experiments. Access regions between the gate and source/drain electrodes were 1μm long. The pinchoff voltage was -1.3 V.

The noise measurement was done by using a trans-impedance amplifier as described elsewhere.[1] The output impedance was obtained by using a lock-in amplifier to measure the differential voltage across a 50 ohm resistor in series with the drain, as well as the drain voltage. Both the in-phase and quadrature component were measured as a function of frequency, temperature and different bias conditions.

3. Results

The phase of the output impedance was plotted as a function of frequency under various bias conditions as shown in Fig. 1. The frequency at which the minimum in the phase occurs, corresponds to the inverse of the trap time constant, τ and is found to increase with temperature. G-r noise measurements were also carried out at several temperatures. The output noise power multiplied by frequency was plotted as a function of frequency (Fig 2).[2] The peak in this plot occurs at the same frequency as that corresponding to the minimum of the phase suggesting a common origin between the output impedance dispersion and the low frequency noise as illustrated in Fig 3 for the measurements at 89C. This leads to the conclusion that the dispersion is not surface state generated[3] since it is generally accepted that g-r noise is generated by bulk traps.[4]

An Arrhenius plot of τ yields an activation energy of 0.75 eV (Fig 5). which is comparable to the energy of the electron trap EL2 (E_a=.825 eV). The temperature gradient between the substrate and the actual device was taken into account following Hughes et. al.[2] Without any temperature correction an activation energy of 0.68 eV would have been obtained. The method employed here to identify the traps causing the z_{ds} dispersion has an advantage over the commonly used conductance DLTS method.[5] Only the traps involved in the a.c. phenomena are excited, whereas the DLTS involves also other traps associated with the surface and depletion

region. We have also studied the frequency dispersion in devices with different channel lengths at various bias voltages. The results show that the dispersion increases as the channel length is reduced and the drain bias is increased. The dispersion is larger at $V_{gs} \sim 0$, and is always negligibly small in the linear region of operation.

4. Discussion

A question arises whether temperature gradient alone can explain the bias dependence of the output conductance dispersion in analogy with the bias dependence of g-r noise as explained in Hughes et al.[2] Our data indicates that this is not the case. By looking at the phase vs frequency at different bias voltages, we see that the minimum of the phase is reduced strongly by increasing V_{ds} (Fig 4). On the other hand the frequency at which the minimum in phase occurs is seen to shift with temperature without any appreciable change in the minimum value. Thus, our results indicate that the output impedance is controlled by the electric field.

We can summarize the behavior of the device as follows. At V_{GS} close to the threshold voltage, the current saturates due to a true pinchoff of the conducting channel at the drain end. At $V_{GS} - 0$, however, the current limiting occurs due to velocity saturation, and an electrostatic dipole region forms close to the drain.[6] This modifies the flow of carriers by forcing them into the substrate.[7] as We have verified this by 2-D PISCES simulations (Fig 6). Carriers surmount the relatively shallow potential barrier (< 0.7V), and penetrate deeply into the SI substrate where traps are encountered. At low frequencies, a fraction of the carriers is captured and emitted by traps in step with changes in V_{DS}, leading to a large z_{ds} because only untrapped carriers appear at the drain, and the trapped carriers tend to reduce the channel thickness; whereas at high frequencies, the time constants of the traps are too long for this to occur, and the current flows as if it were in a trap-free material, decreasing z_{ds}. The amount of dispersion depends on the fraction of total carriers that are trapped. Increasing V_{DS} causes a larger voltage drop across the dipole, thereby strengthening it and causing more carriers to be deflected into the substrate; the amount of dispersion increases accordingly.

5. Conclusion

We have used a new method to measure the deep level responsible for z_{ds} frequency dispersion in GaAs MESFET's. This method uniquely identifies the origin of this phenomena. The results show that the same traps cause dispersion in output conductance and give rise to low frequency noise in these devices.

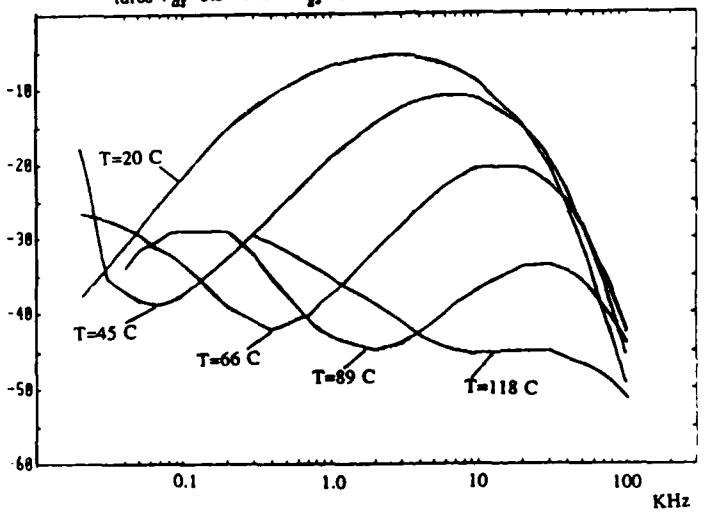
Our results indicate that the high electric fields that exist in modern MESFET's are responsible for the output impedance frequency dispersion. This result is consistent with the saturation mechanism that has been previously proposed.[7] Our experimental data are also consistent with the experimental observations that a p-type implant under the channel eliminates dispersion and reduces noise, by increasing the barrier experienced by carriers deflected towards the substrate.[8] At shorter channel lengths, however, the dipole gets stronger. Unless the operating voltages are also scaled down, such an implant may be unable to restrain carriers, and these low frequency parasitic phenomena, along with an undesirable bipolar action in the p-layer[9] will appear.

References

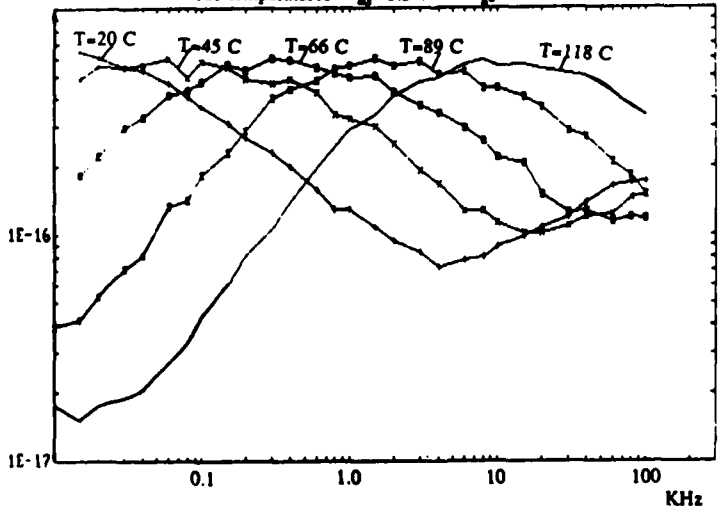
1. J. Chang, "Flicker noise in Silicon MOSFET's," *Master's Thesis UCLA*, 1988.
2. B. Hughes, N. G. Fernandez and J. M. Gladstone, "GaAs FET's with a Flicker-Noise Corner Below 1 MHz," *IEEE Transactions on Electron Devices*, vol. ED-34, pp. 733–741, April 1987.
3. P. Ladbrooke, and S. R. Blight, "Low-Field Low-Frequency Dispersion of Transconductance in GaAs MESFET's with Implications for other Rate-Dependent Anomalies," *IEEE Transactions on Electron Devices*, vol. ED-35, pp. 257–267, March 1988.
4. C. Su, H. Rohdin and C. Stolte, "1/f Noise in GaAs MESFET's," *IEDM Tech. Dig.*, pp. 601–604, 1983.
5. S. Sriram M. B. Das, "Characterization of Electron traps in Ion Implanted GaAs MESFET's on Undoped and Cr-Doped LEC Semi-Insulating Substrates," *IEEE Transactions on Electron Devices*, vol. ED-30, pp. 586–592, June 1983.

6. M. S. Shur, L. F. Eastman, "I-V Characteristics of GaAs Mesfet with Nonuniform Doping Profile," *IEEE Transactions on Electron Devices*, vol. ED-27, pp. 455–461, February 1980.
7. P. Bonjour, R. Castagne, J-F. Pone, J-P. Courat, G. Bert, G. Nuzillat and M. Peltier, "Saturation Mechanism in 1 μm Gate GaAs Fet with Channel-Substrate Interfacial Barrier," *IEEE Transactions on Electron Devices*, vol. ED-27, pp. 1019–1024, June 1980.
8. P. Canfield, J. Medinger, L. Forbes, "Buried-Channel GaAs MESFETs with Frequency Independent Output Conductance," *IEEE Electron Device Letters*, vol. EDL-8, pp. 88–89, March 1987.
9. B. J. Van Zeghbroeck, W. Patrick, H. Meier and P. Vettiger, "Parasitic Bipolar Effects in Submicrometer GaAs MESFETs," *IEEE Electron Device Letters*, vol. EDL-8, pp. 188–190, May 1987.

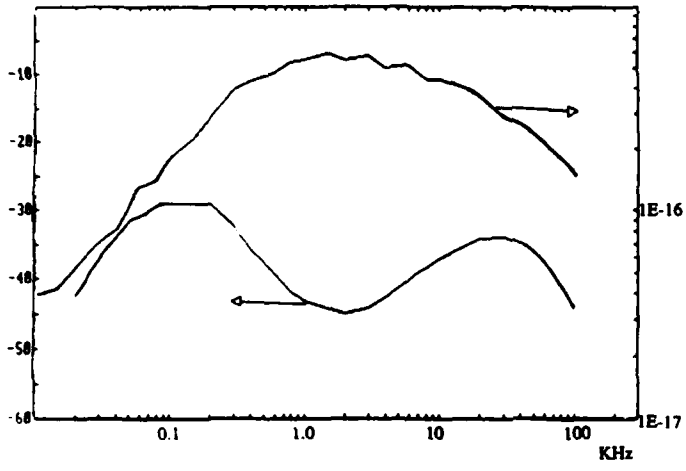
Figure#1
Phase of the output impedance at various temperatures V_{ds}=3.5 volts V_{gs}=0



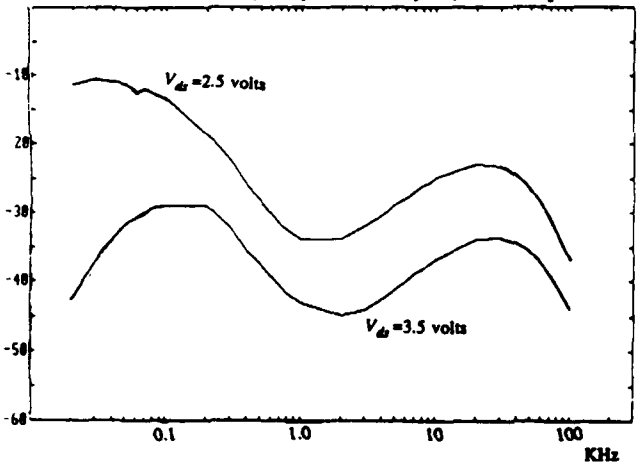
Figure#2
Output noise power times frequency vs frequency at various temperatures V_{ds}=3.5 v V_{gs}=0



Figure#3 Output noise power times frequency and phase of the output impedance vs frequency at 89 C



Figure#4
Phase of the output impedance vs frequency at 89 C V_{gs}=0



Figure#5

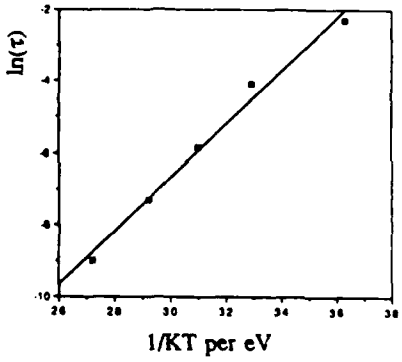
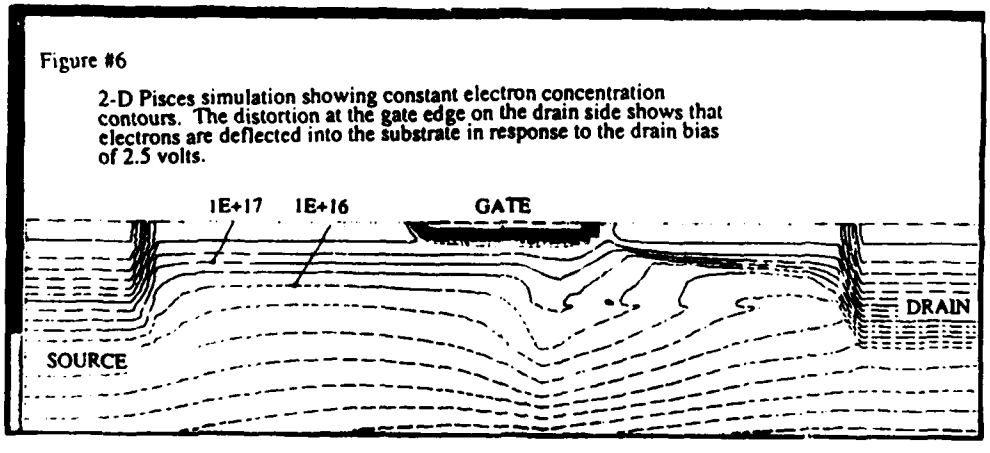


Figure #6

2-D Pisces simulation showing constant electron concentration contours. The distortion at the gate edge on the drain side shows that electrons are deflected into the substrate in response to the drain bias of 2.5 volts.



JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

InGaAs SINGLE- AND DUAL-GATE HIGH-SPEED FETs : PREPARATION AND PERFORMANCE

K. STEINER, K. NTIKBASANIS, U. SEILER, K. HEIME and E. KUPHAL*

Universität Duisburg, Halbleitertechnik/ Halbleitertechnologie, Sonderforschungsbereich 254, D-4100 Duisburg, F.R.G
**Forschungsinstitut der Deutschen Bundespost, D-6100 Darmstadt, F.R.G.*

Abstract

The preparation and performance of self-aligned single- and dual-gate InGaAs JFETs is discussed. Single-gate InGaAs JFETs exhibit maximum extrinsic transconductance of 350, 275 and 140 mS/mm at a gate length of 0.5, 1.5 and 3.5 μm, respectively. High overall potential barriers at the channel substrate heterointerface are necessary for control of threshold voltage uniformity over a wide range of gate lengths. For the first time device behaviour of self-aligned dual-gate InGaAs JFET is demonstrated.

1. Introduction

The compatibility with optoelectronic devices in longwavelength optical communication systems, the high electron mobility and velocity makes InGaAs lattice matched to InP attractive for microwave devices and amplifiers in OEICs. In the case of InGaAs FETs the low Schottky barrier of n-InGaAs /1/ prevents the fabrication of MESFETs. To overcome this problem high band gap materials such as GaAs /2,3/, InP /4/ or InAlAs /5/ were grown on top of the channel or InGaAs MISFETs were realized /6/. The best way to realize InGaAs FETs seems to be the JFET. This is due to the high potential barrier of the pn-junction which is solely determined by bulk properties. The pn-junction can be implanted /7,8/, diffused /9,11/ or grown by epitaxy /12-14/. State of the art are submicron high speed InGaAs-JFETs with self-aligned structures, which reduce parasitic resistance, noise and the mask-alignment effort. They achieve cut-off frequencies above 30 GHz at a gate length of 1.2 μm /15/ and above 70 GHz at a gate length of 0.5 μm /14/, extrinsic transconductances of 330 mS/mm for normally on /11/ and 550 mS/mm for normally-off devices /10/. In this work the fabrication and device performance of mesa type self-aligned single- and dual-gate InGaAs JFET with diffused gates is discussed.

2. Sample Preparation

All layers were grown by LPE. The growth started with a Zn doped p-InP buffer layer ($p<10^{17}$ cm^{-3}, ≈0,4 μm) on s.i. InP:Fe substrate (Sumitomo). It is followed by the n-InGaAs channel layer ($1\cdot10^{17}$ cm^{-3}, 0.40 μm). This layer is lattice matched at growth temperature to the InP buffer layer. The p^+-InGaAs layer is diffused from Zn doped spin-on film sources into the n-InGaAs. The pn-junction is diffused because thermal degradation effects during the fabrication process are lower as compared to ion-implanted and annealed or epitaxially grown layers /9/. The doped emulsions are spun on and dryed in air at about 100 °C for 30 min. During the open tube diffusion the samples are covered with a quartz plate to prevent thermal degradation effects at the surface of the sample. There is no mechanical damage due to the covering of the sample during the diffusion process. The sample is diffused at 535 °C within 10min. Hole concentrations well above 10^{19} cm^{-3} can be reached /9/. The junction depth is controlled by the diffusion time and obeyes a $\sqrt{t}$ law.

For the T-shaped gate Al is evaporated onto the whole sample (200nm). It is followed by a photolithography step and a second metallization (5nm Cr, 100nm Au). After lift-off the gate is defined by the CrAu patterns. The Al is selectively etched with a phosphoric acid (H_3PO_4 : CH_3COOH) using the CrAu cap layer as an etch mask. The T-shaped gate is completed by etching the p^+-mesa. The source and drain metal is directly evaporated onto the undercut structure and thus ensures the self-alignment. It is followed by an additional bond pad metallization. The fabrication process is finished by etching the n-mesa.

3. Experimental Results

After gate formation a Hall-mobility of 7300 cm^2/Vs was measured on an ungated Hall-bar. From magneto-transconductance measurements /16/ maximum drift mobilities between 5500 and 6500 cm^2/Vs were deduced. These values agree well with numerical Hall-factor calculations for InGaAs bulk material with a background carrier concentration of about $1 \cdot 10^{17}$ cm^{-3} /17,18/. Therefore it is concluded that the FET formation process does not degrade transport characteristics of the active InGaAs channel layer.

Fig. 1 displays DC-characteristics of InGaAs single-gate JFETs. Due to the T-shaped gate the gate lengths are 3.5, 1.5 and 0.5 μm. As it is expected /9/ the transconductance increases with decreasing the gate lengths. The maximum extrinsic transconductance comes up to 140, 275 and 350 mS/mm at a gate length of 3.5, 1.5 and 0.5 μm, respectively. These are the highest values ever reported for normally-on InGaAs-JFETs in the submicron and micron gate regime.

With decreasing gate length short-channel effects appear; i.e. the saturation and pinch-off behaviour become worse. The threshold voltage shifts from -1V to -1,6V with decreasing gate length. This is about the order which was predicted in /19/ by numerical simulations of InGaAs-FETs. In contrast to these results InGaAs-FETs with highly doped ($p_{INP} > 1 \cdot 10^{17}$ cm^{-3}) and thick (0.8 μm) p-InP buffer layers exhibit no threshold voltage shifts in the submicron gate regime /11/. InGaAs-FETs without buffer layer typically display tremendous short channel effects with decreasing gate length /6,9,13,18,20/. If a p-InP buffer layer is used it is assumed that the lack of short-channel effects is due to a better carrier confinement in the channel. The reason is a higher overall potential barrier at the channel-substrate heterointerface which is solely determined by the carrier concentration in the active region of the FET and the p-InP buffer layer. The high potential barrier suppresses carrier transfer from the channel into the substrate which may lead to the aforementioned short-channel effects /21,22/. FETs without a buffer layer - the channel is directly grown on the s.i. InP substrate - exhibit a channel-substrate potential barrier which is determined by the conduction band offset only ($\Delta E_c \approx 0.22$eV /23/). This is indicated by an accumulation layer in the InGaAs at the heterointerface /6,24/. In contrast to this accumulation layer a depletion region arises with increasing the distance between the Fermi-level and the conduction band edge in the InP; i.e. with doping the InP by acceptors. In the case of a small potential barrier at the channel-substrate heterointerface an electron transfer from the channel into the substrate is probable especially at higher electron velocities. Therefore short channel effects occure with decreasing the gate length. As a result potential barrier tuning at the channel-substrate heterointerface is necessary for the control of output characteristic at a certain gate length.

Fig. 2 displays device characteristics of self-aligned InGaAs dual-gate FETs. The device consists of two gate electrodes (L_G=1.5 μm) which allow a separate gain control by the second gate. The spacing between the two gates consists of a 4 μm wide floating metal layer. Fig. 2 clearly demonstrates the device behaviour. Biasing the second gate towards negative voltages strongly reduces the transconductance of the device. Such a dual-gate FET will find application in InP based high frequency mixers and oscillators.

In conclusion self-aligned single- and dual-gate InGaAs-JFET were fabricated on LPE layers. Single-gate InGaAs-JFET exhibit maximum extrinsic transconductances of 350, 275 and 140 mS/mm at a gate length of 0.5, 1.5 and 3.5 μm, respectively. For a better control of threshold voltage homogenity with decreasing gate length a high overall potential barrier at the channel-substrate heterointerface is necessary. For the first time device behaviour of self-aligned dual-gate InGaAs-JFET is demonstrated.

Acknowledgement

The authors are indepted to Dr. G.Schlamp and G.Ptaschek (Demetron, Hanau, FRG) for preparation of the spin-on films and M.Böhm for assistance in device preparation. The work was supported by Deutsche Forschungsgemeinschaft.

References:

/1/ Kräutle, H., Wölk, E., Selders, J., Beneking, H., IEEE Transactions on Electron Devices 32(6) (1985) 1119

/2/ Chen, C.Y., Cho, A.Y., Garbinski, P.A., IEEE Electron Device Letters 6(1) (1985) 20

/3/ Selders, J., Roentgen, P., Beneking, H., Electronic Letters 22(1) (1986)14

/4/ Bandy, S., Nishimoto, C., Hyder, S., Hooper, C., Appl. Phys. Lett. 38(10) (1981) 817

/5/ Ohno, H., Barnard, J., Wood, C.E.C., Eastman, L.F., IEEE Electron Device Letters 1(8) (1980) 154

/6/ Gardner, P.D., Bechtle, D., Narayan, S.Y. Colvin, S.D., Paczkowski, J., IEEE Electron Device Letters 8(9) (1987) 441

/7/ Chai, Y.G., Yuen, C., Zdasiuk, G.A., IEEE Transactions on Electron Devices 32(5) (1985) 972

/8/ Selders, J., Wachs, H.J., Jürgensen H., Electronic Letters 22(6) (1986) 313

/9/ Schmitt, R., Steiner, K., Kaufmann, L.M.F., Brockerhoff, W., Heime, K., Kuphal, E., Inst. Phys. Conf. Ser. No.79: Chapter 11, Paper presented at Int.Symp. GaAs and Rel. Comp., Karuizawa, Japan, (1985), 784

/10/ Albrecht, H., Lauterbach, Ch., IEEE Electron Device Letters 8(8) (1987) 353

/11/ Steiner, K., Seiler, U., Brockerhoff, W., Heime, K., Kuphal, E., Inst. Phys. Conf. Ser. No.91: Chapter 7, Paper presented at Int. Symp. GaAs and Rel. Comp., Heraklion, Creece, (1987), 721

/12/ Wake, D., Nelson, A.W., Cole, S., Wong, S., Henning, I., Scott, E.G., IEEE Electron Devices Letters 6(12) (1985) 626

/13/ Cheng, J., Forrest, S.R., Guth, G., Wunder, R., IEEE Transactions on Electron Devices 33 (6) (1986) 725

/14/ Raulin, J.Y., Varsilakis, E., Forte-Poisson, M.A.di, Brylinski, C., Razeghi, M., Inst. Phys. Conf. Ser. No.91: Chapter 7, Paper presented at Int. Symp. GaAs and Rel. Comp., Heraklion, Greece, (1987), 717

/15/ Schmitt, R., Heime, K., Electronics Letters 21(10) (1985)

/16/ Jay, P.R., Wallis, R.H., IEEE Electron Device Letters 2 (1981) 265

/17/ Takeda, Y. Littlejohn, M.A., Hutchby, J.A., Trew, R.J., Electronics Letters 17 (1981) 686

/18/ Takeda, Y., Littlejohn, M.A., Appl.Phys.Lett. 40 (1982) 251

/19/ Brockerhoff, W., Dämbkes, H., Heime, K., Proc. 3rd Int. Workshop on Phys. of Sem. Dev., S.C. Jain, S.Radhakrishna (Eds.), Madras, India (1985), 23

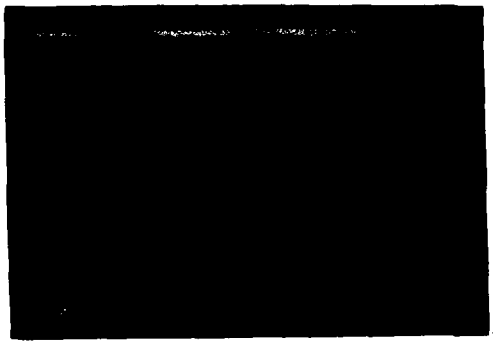
/20/ Wake, D., Livingstone, A.W., Andreas, D.A., Davies, G.J., IEEE Electron Device Letters 5,(7) (1984) 285

/21/ Reiser, M., Electronics Letters 6(16) (1970) 49

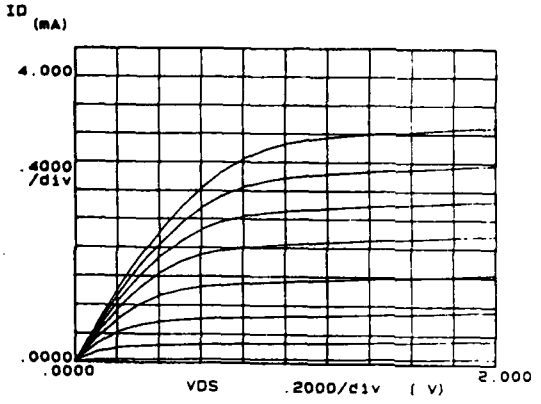
/22/ Matsumoto, K., Hashizume, N., Atoda, N., Awano, Y., Inst. Phys. Conf. Ser. No. 74: Chapter 7, Paper presented at Int. Symp. GaAs and Rel. Comp., Biarritz, France, (1984), 515

/23/ Forrest, S.R., Kim O.K., J.Appl.Phys. 53(8) (1982), 5738

/24/ Steiner, K., Schmitt, R., Zuleeg, R., Kaufmann, L.M.F., Heime, K., Kuphal, E., Wolter, J., Surface Science 174 (1986) 331

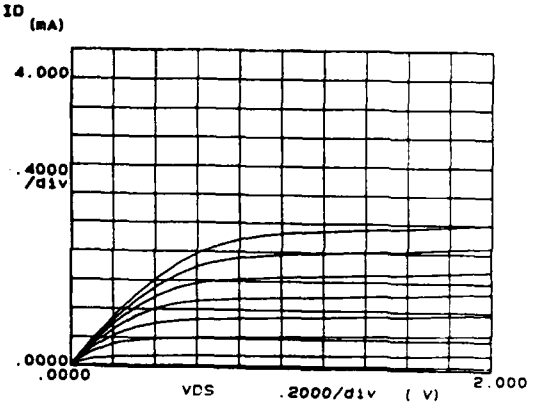


L_G = 4.5 μm

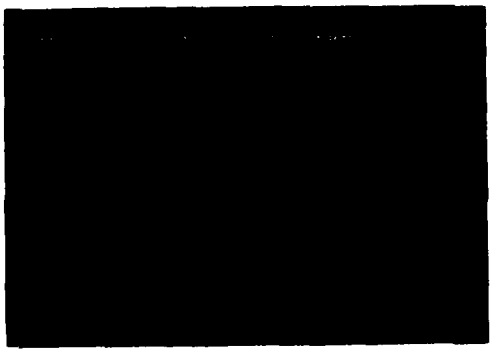


V_{G2S}=0 V, V_{G1S}=-0.2 V/Step

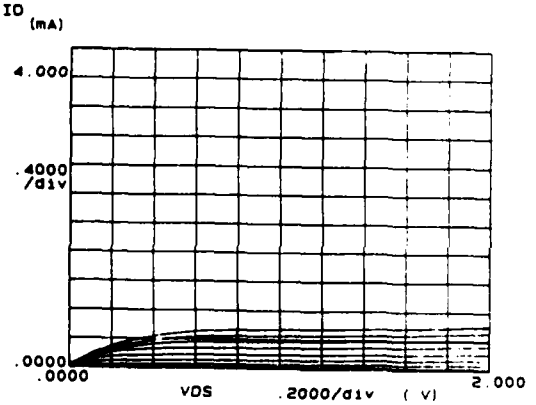
L_G = 1.0 μm



V_{G2S}=-0.5 V, V_{G1S}=-0.2 V/Step



L_G = 0.5 μm



V_{G2S}=-1.0 V, V_{G1S}=-0.2 V/Step

Fig. 1: Output characteristics of self-aligned single-gate InGaAs-JFETs in the same chip; $N_D = 1 \cdot 10^{17}$ cm^{-3}, channel thickness a = 0.18 μm, gate width W_G = 40 μm

Fig. 2: Output characteristics of self-aligned dual-gate InGaAs-JFET, W_G = 40 μm $N_D = 1 \cdot 10^{17}$ cm^{-3}, a = 0.18 μm

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

DEPLETION AND ENHANCEMENT MODE Si_3N_4/GaInAs MISFETs WITH NO CURRENT DRIFT

M. RENAUD, P. BOHER, J. SCHNEIDER, J. BARRIER, D. SCHMITZ*, M. HEYEN* and H. JÜRGENSEN*

Laboratoires d'Electronique et de Physique Appliquée (LEP)[1],
3, Av. Descartes, F-94451 Limeil-Brevannes Cedex, France
**Aixtron, Jülicher Str. 336-338, D-5100 Aachen, F.R.G.*

RESUME: Des transistors MISFETs à canal N, fonctionnant en mode désertion et enrichissement et ne présentant aucune dérive du courant drain source ont été fabriqués sur des couches épitaxiales de $Ga_{0.47}In_{0.53}As$ obtenues par la méthode MOVPE à pression réduite. Ces dispositifs utilisent un procédé de passivation qui inclut un retrait de l'oxyde natif de la surface de GaInAs suivi par un dépôt de Si_3N_4, les deux étapes étant réalisées en plasmas multipolaires. Les transistors MISFETs fonctionnant en mode désertion et enrichissement ($L_g = 2\,\mu m$, $W = 200\,\mu m$) ont une transconductance de 140 et 150 mS/mm, respectivement. Aucune dérive du courant drain source n'apparaît en 10^5 *s* pour les deux types de dispositifs grâce à ce procédé de passivation et en particulier à l' absence d' oxyde à l' interface.

ABSTRACT: N-channel depletion and enhancement mode MISFETs with no drain current drift have been fabricated on $Ga_{0.47}In_{0.53}As$ epitaxial structures grown by low pressure MOVPE. These devices use a passivation process which includes removal of native oxide from GaInAs surface followed by a Si_3N_4 film deposition, both steps performed using multipolar plasmas. Depletion and enhancement mode MISFETs ($L_g = 2\mu m$, $W = 200\mu m$) exhibit an extrinsic transconductance of 140 and 150 mS/mm, respectively. No drain current drift is observed in 10^5 *s* for both types of devices due to the efficiency of the passivation process and especially to the absence of oxide at the interface.

I. INTRODUCTION

Interest in $Ga_{0.47}In_{0.53}As$ alloy lattice matched to InP has been stimulated by its applications for both long wavelength optoelectronics and high speed field effect transistors (FETs). Its excellent electronic properties ($\mu_n(300K) \sim 12000\ cm^2V^{-1}s^{-1}$, $v_s > 2.5 \times 10^7\ cm/s$ and $\Delta E_{\Gamma L} = 0.55\ eV$) provide the opportunity to make FETs with high performances /1/. Due to the low Schottky barrier height of metals on n-GaInAs ($\sim 0.2\ eV$), a metal/semiconductor FET technology is not viable. So, GaInAs FETs with various alternative gate configurations such as Junction FETs (JFETs) /2,3/, Metal Insulator Semiconductor FETs (MISFETs) /4,5/ , Heterojunctions FET (HFETs) /6/ and Modulation Doped FETs (MODFETs) /7,8/ are now investigated.

The MISFET structure is very attractive but its development has been hampered until now by the difficulty to reproducibly achieve a good insulator/GaInAs interface. So, the key issue to develop a viable GaInAs MISFET technology is the control of the insulator/semiconductor interface. Nethertheless, demonstration of operating ring oscillators using inversion mode GaInAs MISFETs has already been done by P.D. Gardner et al /9/ and propagation delay as low as 50 ps was measured in a nine stage ring oscillator.

In this paper, we describe the technological process used to fabricate depletion and enhancement mode n-channel MISFETs. The passivation scheme includes removal of native oxide from GaInAs surface and deposition of a Si_3N_4 film, both steps performed in multipolar plasmas in a ultra high vacuum system. The major problem encountered on InP and GaInAs MISFETs relies on the long term stability of drain current. It will be shown that this passivation process is very efficient and particularly, due to the absence of oxide at the Si_3N_4 / GaInAs interface, highly stable devices are achieved. The drain current drift is shown to be less than 5 % over 10^5 *s*.

[1] LEP : a Member of the International Philips Research Organization

II. MATERIAL

Epitaxial layers have been grown by low pressure Metal Organic Vapour Phase Epitaxy (MOVPE) in a horizontal reactor, capable of handling two 2" wafers in one run /10/. The epitaxial structures are grown on iron doped semi-insulating InP substrates and consists of a nominally undoped InP buffer layer ($d_{InP} = 0.2\ \mu m$) and two n-doped GaInAs layers. Intentional n doping is obtained with SiH_4.The active GaInAs layer, doped to $N_D \sim 10^{17}\ cm^{-3}$ ($\mu_{300K} \sim 6000\ cm^2V^{-1}s^{-1}$) is about 0.1 μm and 0.05 μm thick for depletion and enhancement mode MISFETs, respectively. The top contact layer, doped to $N_D \sim 2 \times 10^{18}\ cm^{-3}$ is 0.1 μm thick (see also figure 1).

III. TECHNOLOGICAL PROCESS

III.A. Passivation scheme

Due to specific problems related to the III-V compounds surfaces such as their propension to loose their stoichiometry under high temperature or bombardment and also to their reaction with their native oxides /11/, a few conditions have to be satisfied to obtain a good insulator/III-V compound interface. Main requirements are a low temperature process with no energetic particles applied on a "stable" surface. The latter condition can be fulfilled either by achieving a stable oxide, this method is recommended by G. Hollinger et al /12/ on InP or by removing the native oxide at the semiconductor surface, that is our purpose on GaInAs. Moreover, we have shown in the past that the "native oxide" is the main contribution to drain current drift of MISFETs /13/.

Removal of native oxide is achieved by a hydrogen multipolar plasma. This cleaning step is done in a ultra high vacuum system and in-situ monitored by kinetic and spectroscopic ellipsometry /14,15/. Then, the sample is transferred in the deposition chamber where a Si_3N_4 film is deposited using a multipolar plasma of SiH_4 and N_2 at room temperature.

A detailed analysis of structural and electrical properties of these Si_3N_4 films has been reported in reference 16. Using optimized deposition conditions, the 1 MHz dielectric constant is 6.8, the breakdown voltage is $\sim 4 \times 10^6$ V/cm and the resistivity at a field strength of 1 MV/cm is $\sim 2 \times 10^{14}\ \Omega.cm$. Besides the oxygen content of the film is $< 5\%$.

This Si_3N_4 / GaInAs interface, characterized with MIS structures, presents a low interface states density, in the range of 3 to $5 \times 10^{11}\ cm^{-2}eV^{-1}$ over 0.4 eV of the band gap /14/ when annealed at 400 °C during one hour before Ti/Au metallization and for 15 min after the metal deposition.

III.B. Device fabrication

A schematic view of the MISFET structure is reported on figure 1. The main difference between the technologies for depletion and enhancement mode MISFETs lies in the thickness of the active GaInAs channel.

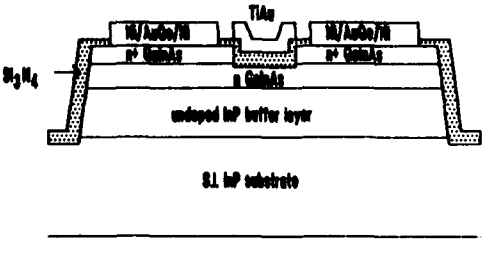


Figure 1: Schematic structure of described MISFETs.

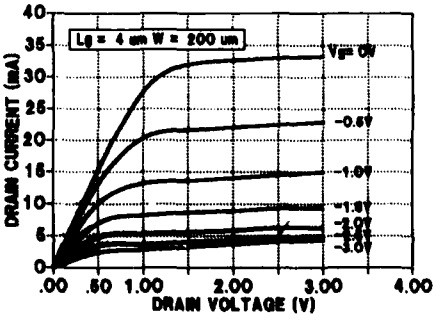


Figure 2: I(V) curve of a D-MISFET.

Device processing begins with the isolation of the active regions by wet chemical etching of n^+ and n GaInAs layers and InP buffer layer. Then, the channel is defined and the n^+ GaInAs layer is chemically etched. *Immediately after this*, the sample is introduced into the UHV system where the native oxide is removed and Si_3N_4 is deposited at room temperature. After annealing at 400°C during one hour, Ti/Au gate metallization is evaporated and defined by lift-off. Windows are opened in the dielectric film by reactive ion etching to define source and drain regions. Finally, ohmic contacts are formed by Ni/AuGe/Ni followed by an alloying.

IV. DEVICE CHARACTERISTICS

IV.A. I(V) characteristics

Typical values of transconductance of D-MISFETs with a 420 Å thick Si_3N_4 and with 200 μm gate width are 80, 115 and 140 mS/mm for 8, 4 and 2 μm gate lengths, respectively. Figure 2 shows the I(V) characteristics of a device with 4 μm gate length. Using a curve tracer, the hysteresis is small, indicating a low density of slow states but no pinch off is achieved.

E-MISFETs (L_g = 2 μm, W = 200 μm) exhibit a transconductance of 150 mS/mm. Transfer characteristics represented on figure 3 also shows no pinch off. This behaviour on both types of devices is suspected to be due to parallel conduction under the channel. Change of thickness as well as doping type of the buffer layer is assumed to improve the pinch off behaviour. From the transfer characteristic at low V_{DS}, the effective mobility of electrons in the channel has been estimated to 4800 $cm^2V^{-1}s^{-1}$.

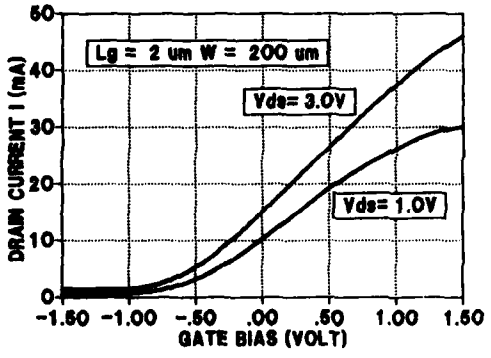


Figure 3: Transfer characteristics of an E-MISFET.

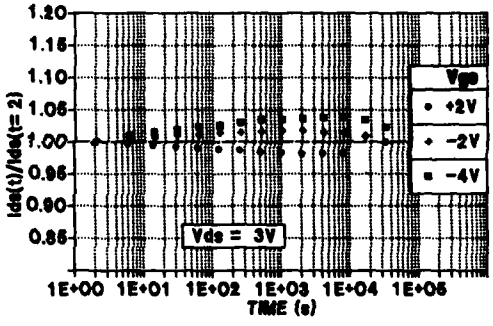


Figure 4: Drain current drift versus time.

IV.B. Stability measurements

In both cases of Depletion and Enhancement mode MISFETs, there is no drain current drift at all when operating at low V_{DS} and after a voltage step on the gate. When biased in the saturation region (V_{DS} = 3.0 V - figure 4), only a slight drift of drain current occurs (< 5% in 10^5 s) which could be due to hot electron injection in the insulator. These devices are remarkably stable compared to MISFETs usually reported in litterature for which the order of magnitude of the drain current drift is ~ 20%.

V. CONCLUSION

In summary, it has been shown that a Si_3N_4 film on a perfectly cleaned surface produces an efficient passivation of GaInAs. E and D-MISFETs exhibit an extrinsic transconductance of 140 and 150 mS/mm, respectively. Moreover, completely stable devices are obtained due to the absence of native oxide at the insulator/GaInAs interface. Same technology on devices with smaller gate lengths is expected to produce stable devices with higher performances.

ACKNOWLEDGEMENTS

The authors would like to thank Y. Hily and E. Boucherez for their contribution in the passivation and fabrication of GaInAs- MISFETs.

This work has been partly supported by the European Community within ESPRIT project 927.

REFERENCES:

1. *A. CAPPY, B. CARNEZ, R. FAUGUEMBERGUES, G. SALMER and E. CONSTANT* IEEE Trans. Electron. Devices, Vol. ED27, pp. 2158-2160, (1980)
2. *D. WAKE, A.W. NELSON, S. COLE, S. WONG, I.D. HENNING and E.G. SCOTT* IEEE Electron Device Letters, Vol. EDL 6, n° 12, p. 626, (1985)
3. *P. RIGLET, J.P. ANDRE, J.P. CHANE, P. DECONNINCK, J.J. VINGRIEF, J. BAELDE, P. GENTRIC* Proceedings of the 13th European Conference on Optical Communication ECOC 87, Helsinki, Finland, September 13-17, 1987, pp. 231-234
4. *P.D. GARDNER, S.G. LIU, S.Y. NARAYAN, S.D. COLVIN, J.P. PACZKOWSKI and D.R. CAPEWELL* IEEE Electron Device Letters, Vol. EDL 7, n° 6, pp. 363-364, (1986)
5. *J. SELDERS and H. BENEKING* IEEE Electron Device Letters, Vol. EDL 7, n° 7, p. 434, (1986)
6. *A. FATHIMULLA, H. HIER and J. ABRAHAMS* Electron. Lett. 24, n°2, p. 93 (1988)
7. *K. HIROSE, K. OHATA, T. MIZUTANI, T. ITOH and M. OGAWA* Proceedings of Int. Symp. on GaAs and related compounds, Karuizawa, Japan 1985, Inst. Phys. Conf. Ser. n°79, p. 529
8. *K. KAMADA, T. KOBAYASHI, H. ISHIKAWA, Y. MORI, K. KANEKO, C. KOJIMA* Electron. Lett. 23, p. 298 (1987)
9. *L.C. UPADHYAYULA, P.D. GARDNER, S.G. LIU and S.Y. NARAYAN* IEEE Electron Device Letters, Vol. EDL 7, n° 6, pp. 390-392, (1986)
10. *D. SCHMITZ, G. STRAUCH, J. KNAUF, H. JURGENSEN and M. HEYEN* To be published in Journal of Crystal Growth
11. *G.P. SCHWARTZ, B. SCHWARTZ, J.E. GRIFFITHS and T. SUGANO* J. Electrochem. Soc. 127, p. 2269, (1980)
12. *G. HOLLINGER, J. JOSEPH, E. BERGINAT, B. COMMERE and P. VIKTOROVICH* Journal of Vacuum Science and Technology, B5, p. 1108, 1987
13. *M. ERMAN, M. RENAUD and S. GOURRIER* Japenese Journal of Appl. Phys., Vol. 26, n° 11, pp. 1891-1897, 1987
14. *M. RENAUD, P. BOHER, J. BARRIER, J. SCHNEIDER and J.P. CHANE* Proceedings of the 17th European Solid State Device Research Conference ESSDERC 1987, Bologna (Italy) Sept. 14-17, (1987), pp. 129-133
15. *P. BOHER, M. RENAUD, J.M. LOPEZ-VILLEGAS, J. SCHNEIDER and J.P. CHANE* Appl. Surf. Science 30, pp. 100-107, (1987)
16. *P. BOHER, M. RENAUD, L.J. VAN IJZENDOORN, J. BARRIER and Y. HILY* J. Appl. Phys. 63 (5) pp. 1464-1472, (1988)

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

IN SITU PROCESSING OF InP BY FLASH LPCVD FOR SURFACE PREPARATION AND GATE OXIDE DEPOSITION

Y.I. NISSIM, C. LICOPPE, J.M. MOISON and C. MERIADEC

CNET, Laboratoire de Bagneux, 196, Av. H. Ravera, F-92220 Bagneux, France

Résumé - Des films de silice sont réalisés dans un réacteur CVD opérant à basse pression, refroidi par air et eau et commandé par chauffage optique rapide. Des vitesses de dépôt allant jusqu'à 100 Å/sec sont obtenues sous un flash thermique de 700°C. Ces dépôts sont effectués sur InP sans que sa surface soit endommagée. Les stuctures InP/SiO_2 ainsi formées ont d'excellentes propriétés électriques adaptées aux exigences du MISFET. L'amélioration des propriétés d'interface de cette structure est obtenue par exposition de la surface d'InP au silane avant le dépôt d'oxide. Une étude de surface indique que le silane réduit les oxydes natifs d'InP.

Abstract - Silicon dioxide films deposited on InP substrates are obtained in a reduced pressure, air and water cooled CVD reactor, with a rapid thermal heating. It is shown that a 700°C temperature flash results in SiO_2 deposition rates close to 100 Å/sec. High temperature deposition (700°C) is thus obtained in few seconds on InP substrates without any surface damage. These layers display excellent electrical properties well suited for MISFET applications. Improvement of the interface properties of this structure is obtained by flowing silane on the InP substrate prior to oxide deposition. Interface studies show that silane reduces the InP native oxides.

1 - INTRODUCTION

Passivation of InP has generated a lot of efforts for MIS applications. Since native oxides of InP exhibits electrical properties with are not suitable for gate dielectrics, there is a need to develop appropriate deposited insulators. The techniques proposed that can be sustained by an InP substrate are direct or indirect plasma /1,2/, low temperature pyrolitic CVD /3/ and photolytic CVD /4/. However in all cases, there is room for improvements in interface properties and bulk SiO_2 structures. When the thermal exposure of an InP substrate is reduced to a few seconds, the maximum temperature allowed before surface degradation can be considerably increased. In this work, the possibility of combining a conventional horizontal CVD reactor with a rapid thermal heating system using tungsten halogen lamps as a source of radiant heat was explored for InP processing. This flash CVD technique named also LRP after the work of ref. /5/ allows precise control of thermally driven surface reactions. In this case the substrate temperature rather than the flux of reactive gas, is used as a switch to turn the CVD reaction on and off. The thermal exposure of the substrate is thus minimized. This technique is a unique opportunity for III-V material to have access to high temperature processes such as CVD or LPCVD /6/. Deposition of SiO_2 was performed on InP at 700°C. Following an in-situ cleaning of the InP surface under a SiH_4 flow.

2 - EXPERIMENTAL

A flash CVD reactor was designed. It consists of an horizontal quartz chamber mounted above a row of halogen lamps. The reactor walls are constantly air and water cooled to insure fast temperature cycling on the substrate, and avoid undesirable contamination. The possibility to deposit SiO_2 on InP at temperatures above 600°C was explored. The reactant gase utilized are silane and oxygen diluted in flowing nitrogen. The total pressure of the system is kept constant at 5 Torr.

A full description of the deposition regimes, and of the kinetics of the reaction obtained on silicon substrates are described in ref. /7/. On InP substrates SiO_2 films are obtained up to 700°C of substrate temperature with no degradation of the InP surface. Deposition rates of 100 Å/sec are measured at this temperature with an oxygen to silane gas ratio of 10. The total cycle time for a 700 Å-thick gate oxide deposited on InP in less than 20 sec, with a temperature rise time of 3 sec and a natural cooling time of few seconds. The InP surface

which is usually thermally degraded at temperatures above 300°C is preserved here due to the fast temperature cycle and the auto-encapsulation of the substrate in the first stage of deposition.

III - CHARACTERIZATION OF THE SiO_2 FILMS ON InP

The index of refraction of this oxide was measured to be 1.46. Its infrared absorption spectrum shows that the $SiH_4 + O_2$ reaction is complete with no traces of Si-H absorption bands. The Si-O stretching and bending modes are located respectively at 1070 and 450 cm^{-1}. These figures are very close to the optical characteristics of thermally oxidized silicon films.

InP substrates oriented along the (100) direction with a residual n-type concentration of 7×10^{15} e/cm^3 were carefully prepared for electrical characterization. A 1000A-thick SiO_2 layer was then deposited at 700°C with the above specified parameters. MIS capacitors were then fabricated on this structure with the lift-off of a Ti-Au metallic film. A resistivity of 8×10^{15} Ω.cm and a breakdown strength of 7 MV/cm were measured on the oxide layer. Capacitance and conductance versus voltage characteristics were measured at 1 MHz. The resulting C(V) plot is displayed in Fig. 1a.

As can be seen from this curve accumulation and inversion regimes are easily reached. The residual hysteresis is 0.8 V which is a low figure for InP based structures. This value does not depend on scanning speed. Capacitance and conductance versus voltage characteristics were then taken between 10 and 100 Hz. At these frequencies the dielectric capacitance does not vary and no conductance peaks due to electron traps are observed in this frequency range.

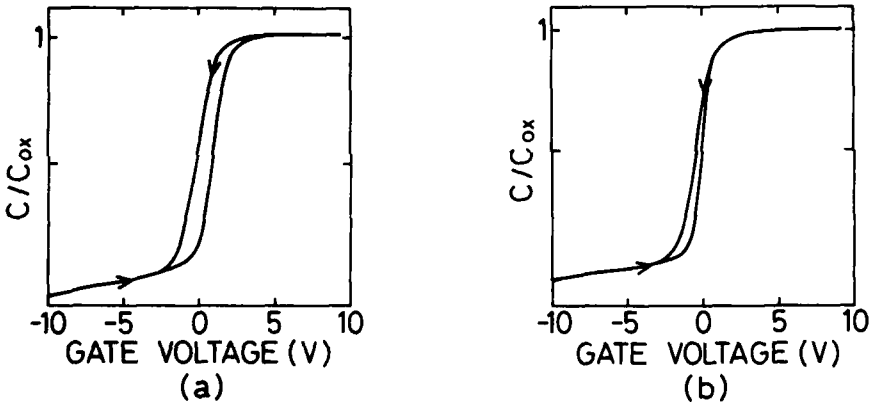


Fig. 1 : Capacitance voltage characteristic at 1 MHz of a MIS diode fabricated on InP by flash LPCVD.
a) deposition of SiO_2 on a virgin as polished substrate
b) deposition of SiO_2 after an in-situ cleaning in flowing silane.

IV - IN-SITU SURFACE CLEANING

One of the major advantages of the flash CVD techniques is that in-situ processing can be easily accomplished. The substrate is hot only during deposition or surface reaction and not during purging or gaz flow stabilization. Different processes can thus be made in-situ without cross contamination of one to the other. Surface preparation prior to oxide deposition can then be conceived.

The authors have shown that an oxidized InP surface exposed to a silane flow will react at 300°C /8/. Silane changes the bonding of surface InP oxides, restoring covalent bonding through a chemical reaction involving silicon bonding to InP and taking over the available oxygen to form Si-O bonds. This is illustrated by the change of AES signal of Indium and Phosphorus of Fig. 2. These spectra indicate that In and P recover their covalent bonding after silane exposure.

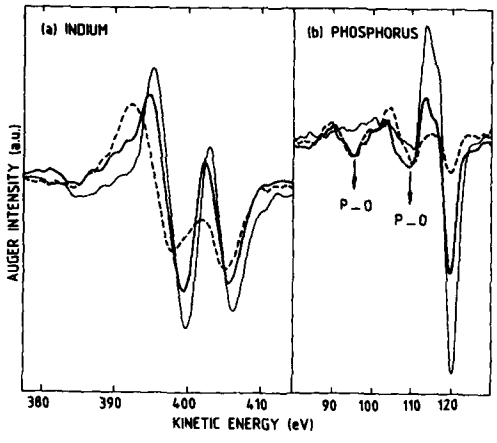


Fig. 2 : Indium and Phosphure Auger spectra for different treatments of the InP surface. Thin solid lines represent clean InP surface, dashed line represents oxidized InP surface and thick solid lines represent the restored InP surface silane exposure. Experiments were performed at a substrate temperature of 300°C.

The potential efficiency of silane preexposition of InP substrates on the improvement of semiconductor/insulator interfacial properties has been tested. MIS capacitors have been process with and without silane in situ cleaning. The cleaning procedure demonstrated in a UHV environment was transposed to the flash CVD reactor conditions. It results in a flash at 250°C for 10 sec in 0.1 Torr partial pressure of silane. The C(V) result of the in-situ cleaning is shown in Fig. 1-b. As can be seen, the hysteresis cycle is lowered by a factor of 3 to a final value of 0.25 eV. This improvement of the hysteresis width can be unambiguovsly assigned to a reduction of interface slow traps.

V - CONCLUSIONS

It has been shown that LPCVD can be controlled with radiant heat. This technique opens the access of high temperature processes to III-V materials such as InP. Furthermore multiple processing steps can be easily carried out in situ within such a reactor.

REFERENCES

/1/ Meiners, L.G., J. Vac. Sci. Tech., 21 (1982) 665

/2/ Pande K.P. and Nair, V.K.R., J. Appl. Phys., 55 (1984) 3109

/3/ Bennett, B.R., Lorenzo, J.P. and Vaccaro, K., Electron. Lett., 24, (1988) 172

/4/ Nissim, Y.I., Regolini, J.L., Bensahel, D. and Licoppe, C., Electron. Lett. 24 (1988) 488

/5/ Gibbons, J.F., Gronet C.M. and Williams, K.E., Appl. Phys. Lett 47 (1985) 721

/6/ Nissim, Y.I. and Bensoussan, M., CNET patent n°8614896

/7/ Nissim, Y.I., Bensoussan, M., Post G., Bensahel, D. and Regolini, J.L., in "Photon, Beam and Plasma Enhanced Processing, edited by Golanski, A., Nguyen, V.T. and Krimmel, E.F. (EMRS proc. XV, les éditions de Physique, Paris, 1987) 213

/8/ Licoppe C., Moison, J.M., Nissim, Y.I., Regolini, J.L., Bensahel, D., submitted to Appl. Phys. Lett.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

HIGH PERFORMANCE SCHOTTKY DIODE AND FET ON InP

S. LOUALICHE, A. GINOUDI, H. L'HARIDON, M. SALVI, A. LE CORRE, D. LECROSNIER and P.N. FAVENNEC

Centre National d'Etudes des Télécommunications LAB/OCM/TOH, BP 40, F-22301 Lannion Cedex, France

Résumé - Une diode Schottky de très haute qualité a été réalisée sur InP de type n en utilisant un traitement de surface original. Cette diode atteint des tensions de claquage de 60 V et le courant de fuite reste inférieur à 1 µA à 30 V. Le meilleur dispositif a un courant de fuite de 0.2 nA à - 1 V et une hauteur de barrière de 0.7 eV. Cette Schottky a été utilisée pour la fabrication de transistors à effet de champ (FET) par CBE et implantation ionique. Une transconductance de 140 mS/mm a été mesurée sur un transistor de 3 µm de grille.

Abstract - A high performance Schottky diode has been realized on InP by a special surface treatment. The diode reaches a breakdown voltage of 60 V and the reverse current remains at 0.6 µA for 30 V reverse voltage. The best device shows a reverse current of 0.2 nA at 1 V voltage with an ideality factor of 1.54. The Schottky has been used as a gate in the fabrication of FETs on InP by ion implantation and CBE. A transconductance of 140 mS/mn has been obtained on a 3 µm gate length FET in the frist experimental trial without any optimisaton.

The Schottky diode is widely used in high speed devices like FETs and optoelectronic devices like MSM (metal semiconductor metal). The InP and related ternary and quaternary materials are the most important semiconductors for the optoelectronic technology. These materials are used in the fabrication of emitter and receiver devices in the 1.3 µm and 1.5 µm wavelengths for lightwave communications in optical fibers. The InP and related compounds have also excellent transport properties (high mobilities and high saturation velocities) for high speed applications. But due to the absence of good quality Schottky diodes,

this material is not extensively used in high speed devices (FET, HEMT) like GaAs.

In the present work we have fabricated a good quality metal insulator semiconductor (MIS) tunnel Schottky device on InP and GaInAs. The fabrication has been done by a low temperature process (below 200°C) which leads to excellent devices.

The samples used to study Schottky diodes are undoped n type (5 x 10^{15} cm^{-3}) InP substrate or layers grown by gaz source molecular beam epitaxial layers (also called CBE : chemical beam epitaxy). The CBE layers are n type InP (undoped ~ 10^{16} cm^{-3} or n type GaInAs (~ 5 x 10^{15} cm^{-3}) of about 1 to 3 μ m thickness grown on semi-insulating InP : Fe substrates. The ohmic contact (Au-Ge annealed 360°C - 10s) is done first to avoid the annealing step to the Schottky contact. After the oxide fabrication, the Schottky metal (Au) is deposited through a metalic mask. The Schottky dots have two dimensions : 0.2 mm and 0.4 mm. The carrier concentrations are measured by C(V) technique using the Schottky device.

The current versus voltage (I(V)) electrical characteristic of a typical diode measured at room temperature is given in Figure 1. The breakdown voltage of this diode is about 10 V, the ideality factor is 1.54 and the apparent barrier height is 0.7 eV. The reverse leakage current at 1 V is below 0.2 nA. To the best of our knowledge these parameters seems to indicate that this Schottky diode is the highest quality device ever fabricated on n type InP and campares favorably to the Schottky diodes fabricated on GaAs. The study of the saturation current with temperature from 100 K to 300 K (Réf. 1) leads to evaluate the oxide thickness of the MIS tunnel diode of the figure 1 to 35 Å. By increasing the temperature of the diode fabrication by 50°C above that of figure 1, the breakdown voltage is increased from 10 V (Figure 1) to 60 V (Figure 2).

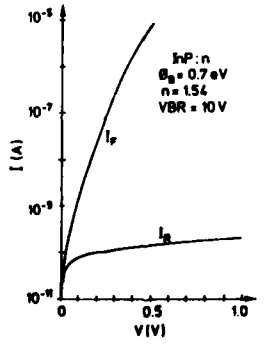


Figure 1 : Typical I(V) characteristic of 0.2 mm diameter diode on n : InP

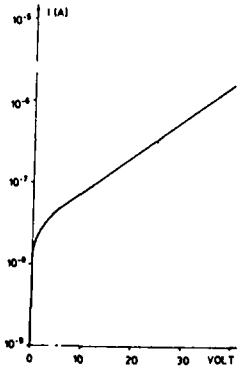


Figure 2 : Reverse I(V) characteristic of 0.2 mm diameter diode on n type InP. The leakage current is about 0.6 µA at 30 V.

The barrier height of this device is 0.65 eV and the ideality factor is 1.45 . The reverse leakage current of this diode is still below the microampere range at 30 V reverse voltage (Figure 2). The oxide thickness of this diode is estimated to 50 Å by I versus T measurement. An ellipsometry evaluation of the oxide thickness leads to 25 Å for the device of Figure 1 and 35 Å for that of figure 2. An other method has been used to evaluate the oxide thickness fabricated on InP by the present oxidation method. A part of on an n type InP sample has been protected by 100 Å Pt layer on its top, the other part has been oxidised and etched ten times to measure the thickness of the InP material used by the oxidation. These measure gives 200 Å for the InP removed by 10 successive oxidation and etch runs. This experiment leads to an oxide thickness of 40 Å if we suppose that $t_{oxide} = 2\ t_{InP}$. These three different methods give an order of magnitude of the optimum oxide thickness for the Schottky contact of 40 Å.

The oxide fabricated to enhance the surface barrier of the Schottky device seems to be stable. The Schottky device parameters does not change when the diode is stored at room temperature during a few months. The device has been polarised at 3 V reverse voltage during 48 hours and we do not observe any particular variation of the device characteristics. The annealing at 200°C for 10 minutes of the finished diode decreases the leakage current by a factor 2. The annealing of the oxide prior to the Schottky metal deposition at 200°C for 30 minutes improves the Schottky quality (better leakage current). At the temperature of 250°C and above the oxyde quality is degraded.

The gold (Au) has been chosen for the Schottky contact. Other metals (Al, Ag, Cr) are also tested, but the devices fabricated by te gold contacts give better performances.

A Deep Level Transient Spectrocospy (DLTS) measurement has been performed on the Schottky fabricated on undoped InP. The temperature range of the measurement lies between 90 K and 300 K and no trap has been observed in this temperature range. If we suppose a trap emission cross section of 10^{-15} cm^2, there is no trap, having its energy lying between 0.12 eV and 0.6 eV below the conduction band, introduced by the Schottky preparation technique. The experimental sensitivity, indicates that the trap concentration, if they exist, is below 10^{10} cm^{-2}.

The present oxidation method studied on n type InP has been tested on GaInAs layers grown by CBE on InP substrate. The GaInAs material in n type ($\sim 5\ 10^{15}$ cm^{-3}) undoped having about 1 μm thickness. The Schottky fabricated on this material has higher leakage current but the breakdown voltage is as high as 300 V (Figure 3). This is the highest breakdown voltage ever obtained on a Schottky device in GaInAs semiconductor.

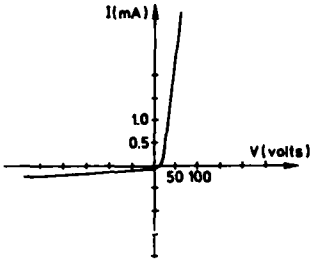


Figure 3 : Linear I(V) characteristic of a Schottky diode on n type GaInAs (1 μm thickness ; 5 10^{15} cm^{-3}) grown on a semi insulating InP substrate. Note the breakdown voltage (> 300 V).

The Schottky device fabricated on n type InP has good long term stability, does not introduce deep traps and the device parameters are suitable for electrical measurements (I(V), C(V), DLTS). This Schottky device has been also found suitable for the gate contact in field effect transistors (FET) fabricated on n type InP.

Two methods have been tried to dope the active layer of the transistor : Ion implantation and CBE. The silicon ion implantation on <100> semi insulating substrate at 2 doses and 2 energies (10^{12} cm^{-2} - 40 keV and 2 x 10^{12} cm^{-2} - 100 keV). To improve the ohmic contact a shallow n+ ion implantation (10^{13} cm^{-2} - 20 keV) has been done below the source and drain contact. The ion implantation annealing is performed at 750°C during 30 seconds.

The ohmic contact annealing occurs at 340°C for 15 seconds. Before the fabrication of the Schottky gate, we have done a 1200 Å etch below the gate by a chemical solution (Br_2 : HB_r: H_2O : 0.1 : 10 : 100) for the control of the saturation current. The gate Schottky contact has higher leakage current compared to the previous Schottky diode. For a 3 μm x 300 μm gate contact the leakage current is about 1 μA at - 1 V and 4 μA at - 2V. The best transconductance on a 3 μm gate implanted FET is 140 mS/mm and compares favorably to the best 1 μm gate MIS devices on InP (180 mS/mm : Ref. 2 - 200 mS/mm Ref. 3). But the device with the best transconductance does not reach the pinch off. The device reaching the pinch off has a transconductance of 110 mS/mm for a 3 μm gate. Other devices presents lower transconductances (Figure 4), but these devices can support Vds voltage above 11 V, and have excellent and constant output transconductance from Vgs = 0 to the pinch off. This property make them suitable for high power amplification with a good linearity.

The active layer grown by CBE consists of an undoped buffer layer (0.5 μm ; 10^{16} cm^{-3}) and a doped active layer (0.2 μm ; 2 x 10^{17} cm^{-3}) has been grown to improve the ohmic contact. The contact resistance is found to be better in CBE FET (3.5 Ω) than in the implanted one (30 Ω). It is necessary to etch under the gate to control the drain saturation current and to reach the pinch off of the transistor because of the high residual doping of the buffer layer. We have found an other methods to avoid the etch by growing a p type layer (500 Å ; 3 x 10^{17} cm^{-3}) between the buffer layer and the active layer. This method leads to transistors which reach the pinch without etching below the gate for a moderate gate voltage of 3 to 4 V. The gate leakage current is about 4 μA at - 3V.

The transconductance obtained in the case of CBE transistors is less than that obtained on implanted FET (50 mS/mm ; Figure 5). This is probably due to the poor quality of our buffer layer.

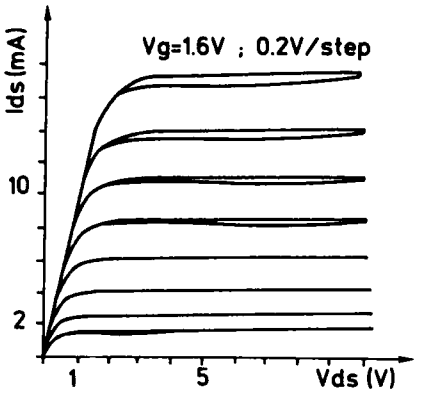


Figure 4 : Ids - Vds network of ion implanted FET on n type InP gate dimension 3 µm x 300 µm

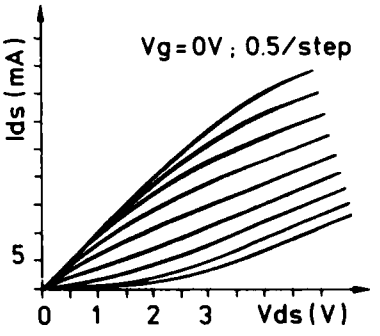


Figure 5 : Ids - Vds network of CBE FET gate dimension : 3 µm x 300 µm

In conclusion, a high quality MIS tunnel Schottky diode has been fabricated on n type InP. This device presents excellent long time stability, high breakdown voltage (60 V), and ultra low leakage current (0.2 nA). The fabrication process does not introduce deep traps. This Schottky diode is adequate when it is used as a gate in the FET. We have shown that the fabrication process of the diode is compatible with ion implanted technique or epitaxial technique for FET fabrication.

REFERENCES

[1] S.M. Sze, Physics of semiconductor devices, (Wiley, New York, 1981) p. 547

[2] A. Antreasyan, P.A. Garbinski, V.D. Mattera, H. Tenkin, J.H. Abeles, Appl. Phys. Lett. 51, 1097, 1987

[3] T. Itoh, K. Ohata, IEEE Transaction on Elect. Dev. ED 30, p. 811, 1983

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

ION IMPLANTED InP MISFET's WITH LOW DRAIN CURRENT DRIFT

G. POST, P. DIMITRIOU, A. FALCOU, N. DUHAMEL and G. MERMANT

CNET, Laboratoire de Bagneux, 196, Av. Henri Ravera, F-92220 Bagneux, France

Résumé

Des transistors à effet de champ M.I.S. sur phosphure d indium semi-isolant ont été élaborés. Les contacts et le canal sont dopés par implantation de silicium. Le diélectrique de grille est SiO_2 déposé par activation ultraviolette. Les dispositifs à désertion à canal de 2 microns ont une dérive du courant de saturation inférieure à 10 p. cent en 24 heures. Ils sont compatibles avec une technologie de lasers pour l'intégration opto-électronique.

Abstract

MIS field effect transistors on semi-insulating indium phosphide have been fabricated. The contacts and the channel are doped by silicon implantation. The gate dielectric is SiO_2 deposited under UV activation. Depletion-mode devices with a 2-micron channel length have a saturation current drift less than 10 per cent in 24 hours. They are compatible with a laser technology for integrated opto-electronics.

1. Introduction

Semi-insulating indium phosphide is the substrate material of choice for integrated opto-electronic emitters and receivers ([1],[2]). MIS and MIS-like field effect transistors can be made by ion implantation ([3],[4]), by epitaxy of n-type InP ([5] - [7]), or on the ternary arsenides lattice matched to InP ([8] - [10]). The implantation technology is appealing for integration purposes, due to its simplicity and the planar device geometry. However, the drift instability of implanted InP MISFET's has to be overcome if they are to compete against the InP JFET's [11], the stable GaInAs MISFET's ([12]), or the AlInAs/GaInAs HFET's ([13] - [16]).

We studied a MIS system prepared by UV-enhanced CVD of silicon dioxide at low temperature ([17]) and achieved a tolerably small amount of drift on 2-micron devices. Even if the absence of gate/contact self-alignment is a performance drawback, those transistors may be useful in a laser driver application.

2. Device Fabrication

A two-step implantation process is used to prepare the n^+ contact and the depletion-mode channel of the devices. The channel implantation consists of 10^{12} silicon atoms / cm^2 at 100 keV. A photoresist mask protects the substrate outside the channel regions. In order to keep the FETs entirely planar, the contact implantation is selective, too. A 300 to 400 nm layer of plasma-deposited phosphorous-doped silicon dioxide is used as the implantation mask for the high Si dose of 10^{14} at./cm^2. Well-defined channel masks of 1 µm have been prepared by reactive ion etching of the PSG material. The low-frequency plasma PSG deposition has been adapted so that the damaged layer on the fragile InP surface does not exceed 50 nm.

The implantation annealing has been carried out at 750 °C, after the stripping of the contact mask and the deposition of another 200 nm PSG capping layer. Following the removal of the annealing cap and the degraded InP surface layer, Au/Ge ohmic contacts have been evaporated and alloyed at 410 °C. The channels could then be etched chemically to bring the satured current down to a typical value of 100 mA/mm, which has been found to yield a FET threshold of -2 V on good substrates with low Fe background.

The gate insulator SiO_2 has been deposited at 180 °C from the UV-activated reaction of silane and nitrous oxide, using traces of Hg vapour as a catalyser. The kinetics of this process is critically dependent on the incident UV light intensity and the mercury concentration. The lack of control and the absence of in-situ thickness monitoring in our research-type reactor resulted in films scattering in a 20 nm interval about the intended thickness of 70 nm.

Titanium-gold gates have been electron-beam evaporated and delineated using a lift-off mask. Finally, the devices have been passivated under a double layer of UV-deposited silicon nitride and silicon dioxide, and Ti-Au contact pads have been added.

3. Properties of MIS Diodes

Control samples of unintentionally n-doped InP have been used to prepare MIS diodes along with each transistor process. Annealing of the diode structures between 200 and 250 °C has been successful in curing most of the evaporation-related damage and in improving the dielectric strength of the SiO_2. We obtained an insulator resistivity of 10^{14} Ohm cm, a breakdown electric field of 5 MV/cm and an accumulation capacitance dispersion below 3 % in the range from 100 Hz to

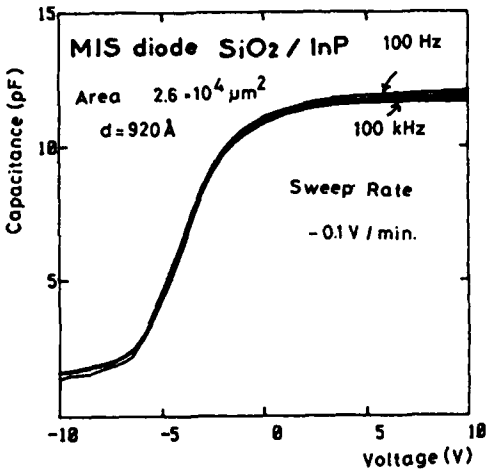


Fig. 2: Frequency dispersion of InP MIS capacitor

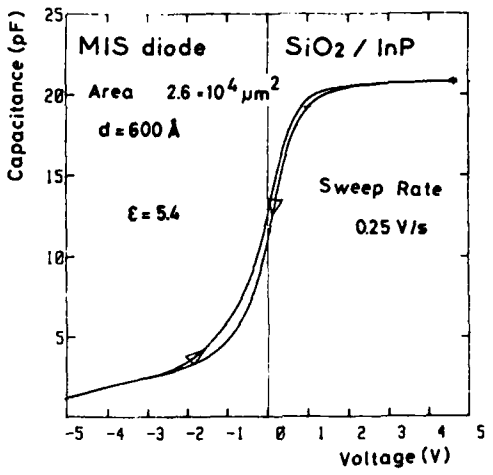


Fig. 1: 1 MHz C-V curve of UVCVD silicon dioxide / InP MIS system

100 kHz (Fig. 2). The dielectric constant is rather high (5 to 5.5), since the water and SiOH contents of the UVCVD layers are 7 atoms %, as estimated from the infrared absorption spectra. Rutherford backscattering analysis of the trace amount of Hg incorporated in the silicon dioxide gave a value below the detection limit (several 10^{16} atoms/cm^3 for thick control layers on graphite). An electron-injection type hysteresis of 0.5 V (Fig. 1) is characteristic of these MIS systems, for C-V cycles stressing the structure up to 1 MV/cm at a 0.25 V/cm sweep rate.

4. Transistor Characteristics

Three varieties of MISFET's have been studied on substrates with a low Fe doping level (< 10^{16} cm^{-3}): Enhancement-mode devices with 5-micron gates over 1 and 2-micron channels, depletion-mode transistors with 2-micron channels and 2...3 micron gate length. No self-alignment between the channel and the gate has been tried: our MIS systems do not withstand the implantation annealing treatment as do some GaInAs MIS systems ([6]).

The 2-micron enhancement and depletion mode devices have threshold voltages of 0 V and -2 V, respectively (Fig. 3), and transconductances from 70 to 110 mS/mm at 100 mA/mm saturated drain current. An implanted MISFET and its unimplanted neighbour exhibit very similar V_{gs} - I_{ds} curves (Fig. 4) and the transconductance peaks are almost the same. The dominant effect of the impantation is a threshold voltage shift; most of the active channel electrons are in the accumulation layer. The gate transfer characteristics in the linear region correspond to a peak effective mobility of 2200 cm^2/Vs. The gate-source capacitance of the depletion-mode FET's is about 2 pF/mm.

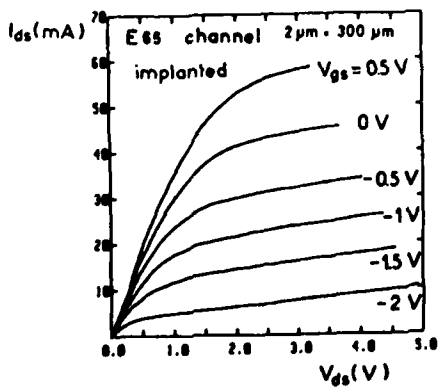


Fig. 3: Characteristics of implanted InP MISFET

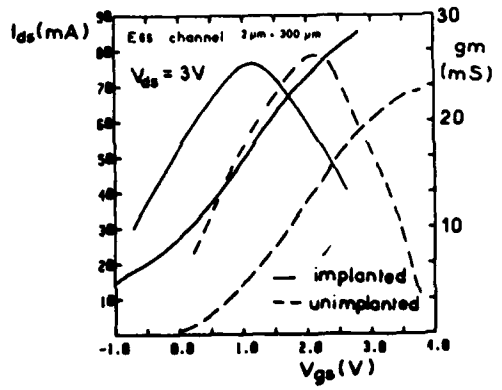


Fig. 4: Saturated drain current and transconductance of implanted vs. unimplanted MISFET

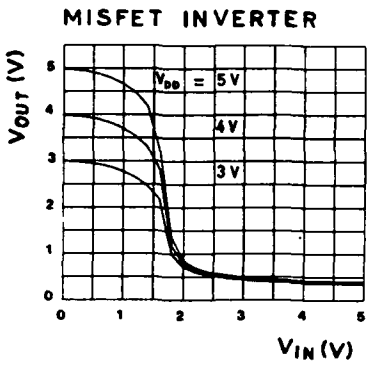


Fig. 5: InP MISFET inverter with 2-micron enhancement/depletion transistor pair

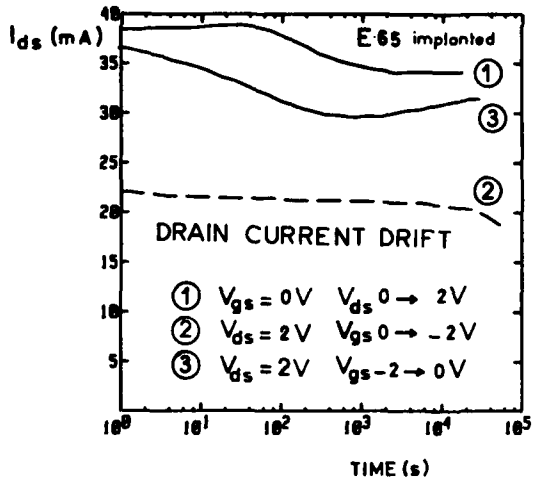


Fig. 6: Saturated drain current drift sequence (1,2,3) of implanted InP MISFET: the "t=0" voltage transitions are indicated.

Fig. 5 shows the static transfer curves of NMOS-style inverters composed of an enhancement-mode switch and a depletion-mode load, after repeated cycling of the device with a 0 to 5 Volt input. Before stress, the threshold was 0.5 Volts lower, which is consistent with the hysteresis observations on the C-V curves.

5. Drift Behaviour

The biggest problem of current InP MISFET technology is the long-term drift of drain current: a decrease is observed in general, corresponding to an upward drift of threshold voltage due to the injection and trapping of negative charges at the InP-SiO_2 interface ([18]). We used the common method of drift characterization which is the measurement of drain current evolution on a logarithmic time scale under small drain bias (V_{ds} = 0.1 V), after the application of a gate voltage step.

A sequence of 24-hour drift cycles on the same depletion-mode device in the linear region (V_g from 0 to -2 V, then back to 0 V) gave a reversible threshold voltage drift of less than 0.5 V. Three cycles in the saturated regime (step in V_d from 0 to 2 V, then V_g held at 0, -2 and 0 V) have been recorded next (Fig. 6). There are effects of positive and of negative charges, the positive ones being dominant at longer time constants. We cannot exclude alkali ion contamination as a possible parasitic drift effect since our processing environment and methods lack somewhat the stringent MOS discipline. The total amount of saturated drain current variation in 24 hours is equivalent to a threshold voltage drift of less than 0.5 Volts. That means that it can be compensated for in a typical laser driver application where a feedback stabilisation of the light output level is necessary in any case.

6. Conclusion

Wholly implanted InP MISFET's may be pushed to performance levels adequate for most applications in optoelectronics. The devices are entirely planar, have a negligible gate leakage current, their threshold may be adjusted between -2 and 0 Volts, and no epitaxial layers are required. On the other hand, the substrate preparation and the gate dielectric deposition are delicate processes. The residual drift of threshold voltage is still a matter of concern about device reliability, but important progress has been made on that subject. Better control of process steps is expected in the near future and we hope that the planar geometry of implanted FETs will make them fit easily into OEIC designs.

ACKNOWLEDGEMENTS

The authors are indebted to A. Scavennec for helpful discussions. P. Krauz, L. Bricard, D. Arquey, C. Joly and C. Mayeux contributed to device processing. This work has been supported by ESPRIT under contract No. 2638.

REFERENCES

[1] A. Suzuki, K. Kasahara, and M. Shikada: IEEE J. of Lightwave Technology 5 (1987), p.1479
[2] N. Suzuki, H. Furuyama, Y. Hirayama, M. Morinaga, K. Eguchi, M. Kushibe, M. Funamizu, and M. Nakamura: Electronics Letters 24 (1988), p.467
[3] K.P. Pande, M.A. Fathimulla, D. Gutierrez, and L. Messick: IEEE Electron Dev. Lett. 7 (1986), p.407
[4] T. Itoh and K. Ohata: IEEE Trans. Electron Dev. 30 (1983), p.811
[5] A. Antreasyan, P.A. Garbinski, V.D. Mattera, Jr., H. Temkin, and J.H. Abeles: Applied Phys. Lett. 51 (1987), p.1097
[6] P.D. Gardner, S. Yegna Narayan, S.G. Liu, D. Bechtle, T. Bibby, D.R. Capewell, and S.D. Colvin: IEEE Electron Dev. Lett. 8 (1987), p.45
[7] L. Messick, D.A. Collins, R. Nguyen, A.R. Clawson, and G.E. McWilliams: Proceedings IEDM-86 (1986), p.767
[8] P.D. Gardner, S.G. Liu, S. Yegna Narayan, S.D. Colvin, J.P. Paczkowski, and D.R. Capewell: IEEE Electron Dev. Lett. 7 (1986), p.363
[9] J. Selders and H. Beneking: IEEE Electron Dev. Lett. 7 (1986), p.493
[10] E. Frederic Schubert, W.T. Tsang, M.D. Feuer, and P.M. Mankiewich: IEEE Electron Dev. Lett. 9 (1988), p.145
[11] S.J. Kim, G. Guth, and G.P. Vella-Coleiro: IEEE Electron Dev. Lett. 9 (1988), p.306
[12] M. Renaud, P. Boher, J. Schneider, J. Barrier, D. Schmitz, M. Heyer, and H. Jurgensen: ESSDERC 1988, this issue
[13] M.A. Fathimulla, T. Loughran, L. Stecker, E. Hempfling, M. Mattingly, and O. Aina: IEEE Electron Dev. Lett. 9 (1988), p.145
[14] J.A. Del Alamo and T. Mizutani: IEEE Electron Dev. Lett. 8 (1987), p.534
[15] U.K. Mishra, A.S. Brown, L.M. Jelloian, L.H. Hackett, and M.J. Delaney: IEEE Electron Dev. Lett. 9 (1988), p.41
[16] M.D. Feuer, J.M. Kuo, S.C. Shunk, R.E. Behringer, and T.Y. Chang: IEEE Electron Dev. Lett. 9 (1988), p.162
[17] U. Sharma, RCA Review 47 (1986), p.551
[18] S.M. Goodnick, T. Hwang, and C.W. Wilmsen: Appl. Phys. Lett. 44 (1984), p.453

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

INTERFACE STATES PARAMETERS DEDUCED FROM DLTS, ICTS AND CONDUCTANCE METHODS ON TiAu/Si_3N_4/GaInAs MIS STRUCTURES

J. BARRIER, M. RENAUD, P. BOHER and J. SCHNEIDER

Laboratoires d'Electronique et de Physique Appliquée (LEP)[1],
3, Av. Descartes, F-94451 Limeil-Brevannes Cedex, France

RESUME: Plusieurs méthodes de caractérisation électrique utilisant les techniques de DLTS, d'ICTS, et de conductance ont été mises au point et associées. En prenant en compte les corrections dues à la variation de la section efficace de capture en fonction de l'énergie, ces méthodes permettent une détermination cohérente des caractéristiques des états d'interface et ont été utilisées sur des structures MIS ($TiAu/Si_3N_4/Ga_{0.47}In_{0.53}As$). L'application à l'optimisation d'un procédé de passivation efficace de l'interface $Si_3N_4/GaInAs$ utilisant un retrait de l'oxyde natif de la surface de GaInAs suivi d'un dépôt de Si_3N_4 par plasma multipolaire en ambiance ultra vide a permis la réalisation de transistors MISFET sur GaInAs possédant de bonnes performances.

ABSTRACT: A set of electrical characterization methods has been developed using DLTS, ICTS and Conductance techniques. Taking into account corrections due to the variation of capture cross section versus energy , this method allows for a coherent determination of interface states parameters and has been used on $TiAu/Si_3N_4/Ga_{0.47}In_{0.53}As$ MIS structures. It has been applied to perform an efficient passivation process of the $Si_3N_4/GaInAs$ interface consisting in an in situ native oxide removal and Si_3N_4 deposition by multipolar plasmas in a ultra high vacuum system. GaInAs MISFETs with good performances could be achieved using such an optimized process.

I. INTRODUCTION

$Ga_{0.47}In_{0.53}As$ is a very attractive material for the fabrication of high speed electronic components due to its excellent electronic properties. Several gate configurations of field effect transistors (JFET, HFET, MODFET) have been investigated with good results.

For insulated gate field effect transistors (MISFETs) technology, large interface states densities and drain current drift can appear with the insulator deposition, especially on III-V compounds MISFETS. Theses difficulties hampered the development of this technology but, recently, few works have demonstrated possibility to overcome them /1,2/. Such application requires an efficient passivation process to reduce the density of interface states while keeping a high stability with time. Many passivation methods have been applied with various results but main conditions are a low temperature process and a low energetic insulator deposition due to the III-V compounds surface instability.

Here, we have used a ultra high vacuum system and multipolar plasma treatments. Because "native oxide" seems to be responsible of the drain current drift /3/, we have chosen to remove it in a multipolar plasma treatment.

In order to improve this passivation process, we needed to develop a set of sensitive characterization methods. DLTS, ICTS, Conductance and Capacitance measurements have been performed on $TiAu/Si_3N_4/GaInAs$ MIS structures. Then, two kinds of treatments and there effects on the deduced interface state density were studied.

II. SAMPLE PREPARATION

Electrical measurements have been made on GaInAs layers lattice matched to n^+InP substrate grown by vapor phase epitaxy using the chloride method. The samples are chemically deoxidized in di-

[1] LEP : a Member of the International Philips Research Organization

luted HF before introduction into the UHV system. Then, the native oxide is completely removed in a H_2 multipolar plasma with a given hydrogen ion density that we have varied during this study. Finally, a Si_3N_4 film is deposited using a multipolar plasma of SiH_4 and N_2 at room temperature. Both steps can be monitored by in situ kinetic and spectroscopic ellipsometry /4/. After passivation, one hour annealing at 400°C is performed and TiAu dots are evaporated followed by a 15mn anneal at 400°C.

III. CHARACTERIZATION TECHNIQUES

III.A. DLTS-ICTS

We have recently presented a general method to extract the density distribution $N_{ss}(E)$ and the capture cross section $\sigma_n(E)$ from DLTS and ICTS measurements (J.M. Lopez Villegas, P. Boher, M. Renaud submitted to J. Appl. Phys. 1987). The coherence between the two spectroscopies is assumed only taking into account the variation of σ_n versus energy. This method is the most sensitive one but it assumes that the capture cross section is constant versus temperature.

III.B. CONDUCTANCE

Conductance method has also been applied on the same samples. Using a HP4192A impedance analyzer, Capacitance and Conductance can be measured from 500Hz to 10MHz at a constant bias on the metal gate of the sample which can be heated from room temperature to 110°C. The analytical expression of interface states admittance (Y_{ss}) has been calculated by Deuling /5/ only in two cases:

i) When the capture cross section and the density of states do not change with energy over a few KT:

$$Y_{ss} = \frac{eN_{ss}}{\tau} \left| \frac{\ln(1+\omega^2\tau^2)}{2} + j \tan^{-1}(\omega\tau) \right|$$

where τ is the interface time constant classically related to the capture cross section.

ii) For a variation of σ_n versus energy defined by $y = (KT/e) \times (d \ln(\sigma_n)/d\psi) = 0.5$, where the effect of the variation of σ_n is maximum. (The first case corresponds to y = 0):

$$Y_{ss} = \frac{eN_{ss}}{\tau} \{\Gamma | 1 - A | + jA \ln \left| \frac{1/A + \omega\tau/2}{1/A - \omega\tau/2} \right| \} \qquad \text{where } A = (1 + \omega^2\tau^2/4)^{-1/2}$$

Interface states parameters are obtained by fitting the experimental curves with both models. N_{ss} , σ_s (variance of surface potential fluctuation) and time constant are deduced from $Gp/\omega(F)$ curve and ψ_s(surface potential) from Cp(F) curve. σ_n is calculated fro τ and ψ_s . This method needs the precise knowledge of the impurity concentration. Measurements at various temperatures allows for the determination of any thermal activation of capture cross section.

IV. INTERFACE CHARACTERIZATIONS

Figure 1 reports the fit of $Gp/\omega(F)$ and Cp(F) curves with the two hypothesis (y = 0 and y = 0.5) for conductance measurements made at room temperature on one sample (gate bias: -1.4 V) .

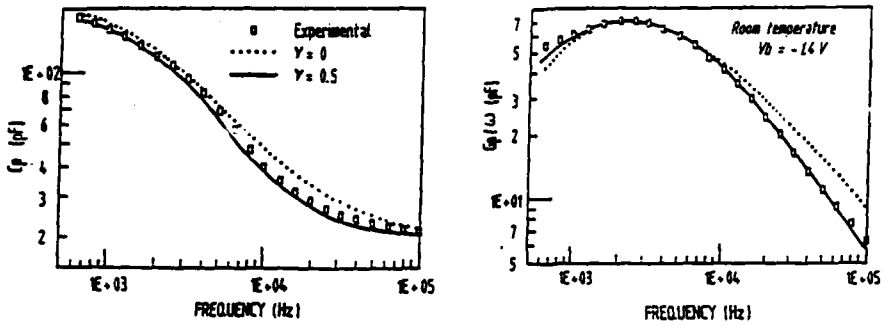


Figure 1: *Experimental and theoretical parallel capacitance Cp and conductance divided by the angular frequency Gp/ω versus frequency for the sample treated with the lowest hydrogen ion density (sample 1).*
$\gamma = 0$: $N_{ss} = 3.4 \times 10^{12} cm^{-2} eV^{-1}$, $\sigma_g = 0.03$, $\sigma_n = 5.6 \times 10^{-15} cm^2$, $\psi_s = 0.388\ eV$
$\gamma = 0.5$: $N_{ss} = 3.3 \times 10^{12} cm^{-2} eV^{-1}$, $\sigma_g = 0.67$, $\sigma_n = 1.03 \times 10^{-14} cm^2$, $\psi_s = 0.351\ eV$

The $Gp/\omega(\omega)$ curve cannot be well fitted by the constant cross section model even with no charge fluctuation ($\sigma_g = 0.03$). On the contrary, the variable cross section model improves both fits of $Gp/\omega(F)$ and Cp(F) curves and confirms DLTS and ICTS measurements which show a strong variation of σ_n (fig. 3a)

Effects of different hydrogen ionic densities during the UHV treatment on the interface states characteristics are shown on figure 2 for the density of states and on figure 3a and 3b for the capture cross section . It clearly appears that higher hydrogen ion density which is the more energetic one (sample 2) provide larger interface states density and also larger capture cross section and so faster interface states.

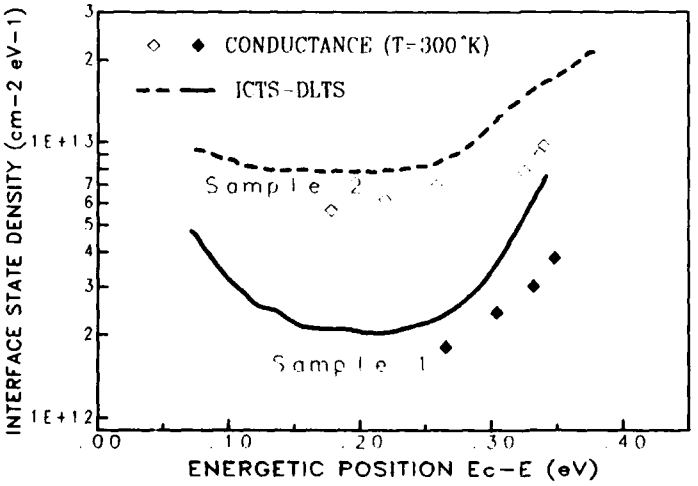
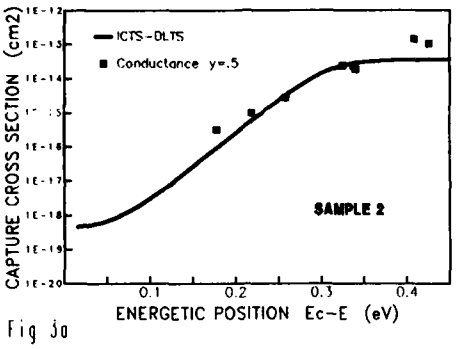


Figure 2: ***Interface states densities versus energy on GaInAs MIS structures fabricated either with a low (sample 1) or high (sample 2) hydrogen ion densities during the oxide removal treatment. The values extracted from conductance have been obtained using a variation of the capture cross section in the fits. The density of states measured on the same samples by DLTS and ICTS methods are also reported.***



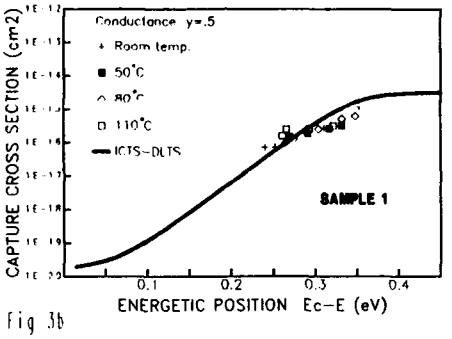


Figure 3a & 3b: *Capture cross sections versus energy on GaInAs MIS structures fabricated either with a low (sample 1) or high (sample 2) hydrogen ion densities during the oxide removal treatment. The values extracted from conductance have been obtained using a variation of the capture cross section in the fits. The capture cross sections measured on the same samples by DLTS and ICTS methods are also reported*

This result confirms that the whole process must be the less energetic as possible , especially for the oxide removal and for the Si_3N_4 first layers deposition. The excellent agreement between conductance method , DLTS and ICTS measurements can only be obtained taking into account the variation of σ_n versus energy. Conductance measurements have been made from room temperature to 110°C on the sample processed with the lowest hydrogen ion density . Capture cross sections calculated from the variable cross section model fit are reported on figure 3b and no thermal activation could be observed.

V. CONCLUSION

This work shows the necessity to take into account the capture cross section variation versus energy to correctly extract interface states parameters from conductance measurements on MIS structures. With these conditions, this method allows for the determination of a constant capture cross section versus temperature. Applied to the characterization of different multipolar plasmas treatments, this method shows the great damages produced by energetic ions in the plasma on the electrical properties of the interface. The optimized passivation process has permitted to the fabricate GaInAs MISFETS with good performances (140 mS/mm and no drain current drift, see other paper in the proceedings).

ACKNOWLEDGEMENTS

The authors would like to thanks J.P. CHANE for the growth of GaInAs samples. Thanks are also due to Y. HILY and E. BOUCHEREZ for the fabrication of MIS structures.

This work has been partly supported by the European Community within ESPRIT project 927.

REFERENCES:

1. *P.D. GARDNER, S.G. LIU, S.Y. NARAYAN, S.D. COLVIN, J.P. PACZKOWSKI and D.R. CAPEWELL* IEEE Electron Device Lett., Vol. EDL-7, n° 6, p. 363, 1986
2. *Y. TAKAHASHI, T. TAKAHASHI, T. SHITARA, Y. IWASE, Y-H. JEONG, S. TKAGI, F. ARAI and T. SUGANO* 14th Int Conf on GaAs and Rel. Compounds, Heraklion (Crête) 1987, Int. Phys. Conf. Ser. N° 91, pp 731-716
3. *M. ERMAN, M. RENAUD and S. GOURRIER* Japenese Journal of Appl. Phys., Vol. 26, n° 11, pp. 1891-1897, 1987
4. *P. BOHER, M. RENAUD, J.M. LOPEZ-VILLEGAS, J. SCHNEIDER and J.P. CHANE* Appl. Surf. Science 30, pp. 100-107, (1987)
5. *H. DEULING, E. KLAUSMANN and A. GOETZBERGER* Solid State Elec. n° 15, p. 559, (1972)

SESSION 3

Session 3IP Room A : Invited papers

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

MEGAPIXEL IMAGE SENSORS TECHNOLOGY

R.P. KHOSLA

Electronics Research Laboratories, Eastman Kodak Company, Rochester, NY 14650-2010, U.S.A.

Abstract-In this paper we will discuss the technology of megapixel sensor and megapixel imager status in Europe, Japan, and U.S.A. The discussion will also include the applications of these sensors for high-definition television (HDTV) and high-quality electronic photography.

Charge-coupled devices (CCDs) were invented by Boyle and Smith[1] in 1970. These devices were immediately proclaimed to have applications in the area of imaging, signal processing, and memories. Much work has gone into the development of CCDs for applications in all these areas. However, the greatest progress has been made in the area of solid state imaging. Since 1973, the number of picture elements has increased from a few thousand to the most recent announcement of a 2-million pixel imager[2] by Toshiba, thus giving a three orders of magnitude increase in pixel density in 15 years. Of course these advances have, in general, followed the advances in dynamic random access memory. During the same 15-year period, the D-RAM has also increased in memory by three orders of magnitude, from 4 kilobits of memory in 1973 to 4-megabit memory which were reported this year [Figure 1].

These advances have occurred as a result of the advances in VLSI technology which have given a push in improvements in materials, equipment, and processes. Sensors have been developed for consumer, commercial, and industrial applications. Consumer applications range from video to high-density television; industrial uses include robotic and machine vision, electronic imaging for advertising and communication, integrated text and images in offices, and publishing. In the areas of defense application, we are moving from general surveillance applications to imaging applications in the areas of robotic battlefields and intelligent robotics.

This presentation is specifically dedicated to advances in megapixel technology, and is a report on the status of megapixel sensor development in Europe, Japan, and U.S.A. In this paper, we will start with a brief background review of CCD technology and image sensors.

An image sensor consists of photosensitive elements that convert incoming light to an electric charge and readout structures that transfer the charge to the output. Three major classes of photosensitive elements are in use: the photodiode, the photocapacitor, and the photoconductor. The major classes of readout architectures for image sensors are the frame-transfer CCD, the interline CCD, and the x-y addressed metal oxide semiconductor array. Photodiodes have better sensitivity in the blue compared to photocapacitors. In frame-transfer architecture, the entire pixel area is photosensitive and, as such, the quantum efficiency is high compared to the pixel in interline or x-y addressed architecture.

The motivation for the development of high-density megapixel sensors is (i) to provide the consumer with high-quality pictures that are comparable to the pictures of traditional silver halide photography, and (ii) as sensors for high-definition television (HDTV). The equivalent resolution of a 35 mm film is of the order of 20 million pixels, and the requirement for high-definition television is approximately 2 million pixels. Until 1986, most of the work that has been done in the solid state imaging area based on CCD technology has been in the development of sensors that had pixel densities to satisfy NTSC television requirements and were usually in a 2/3-inch format. The pixel density for NTSC television is approximately 380,000. A large number of companies have pushed the technology for high-density sensors to reduce costs. Thus, the major driving force for continuing to push CCD technology for NTSC standard has been price/performance. Table I shows the list of 1/2-inch-format CCDs that are available today. As such, solid state sensors of moderate (200,000-400,000 pixels) resolution that are compatible with NTSC requirement now occupy an established and growing market.

The first large-density sensor was developed by TI and had a pixel density of 640,000 elements. Since then, there are at least eight companies that have announced products with megapixel resolution sensors: Ford Aerospace, Kodak, NEC, Sony, Tektronix, Texas Instruments, Thompson CSF, and Toshiba. Table II gives a comparison of the pixel density, pixel size, imager area, and other characteristics of the various megapixel sensors.

Although the advances in the megapixel technology have mostly resulted from the advances in semiconductor technology, there are specific sensor considerations such as sensor architecture, pixel size, signal, noise, blooming control, color filters, and others that must be addressed. Sensor considerations can vary with the specific application. For example, the Tektronix imager[3] was designed for astronomical applications, has a pixel size of 27 x 27 μm^2, and has a chip size of 55 x 55 mm^2. The Kodak megapixel sensor[4] has a higher pixel density of 1320 x 1035 and a pixel size of only 6.8 x 6.8 μm^2 and is compatible for 2/3-inch TV format.

To keep the total size of the device small, TI developed the virtual-phase CCD technology and Kodak developed the true-two-phase CCD technology. Reducing the size of the pixel to keep the chip size small has an added problem in that the charge handling capacity of the pixel or the amount of signal per pixel becomes low. Appropriate architecture such as the frame transfer device is generally employed for sensor design to take advantage of the large fill factor. Appropriate scaling of the pixel, from both design as well as processing, can immensely help the charge handling capacity of an image sensor, as has been well demonstrated in the development of megapixel sensor technology at Kodak.[4] The charge-handling capacity of imagers can also be increased by putting photoconductive overlayers on CCDs. With this scheme, the diode in the pixel that is generally used as a photosensitive element becomes a charge storage medium, and the total pixel containing the diode, the transfer gate, and the output structure now has a photoconductive overlayer and is nearly 100 percent photosensitive. This scheme has been utilized by Toshiba[2] to make a 2-million-pixel CCD. To maintain a high signal-to-noise ratio, the low-signal handling capacity of small pixels puts pressure on devising ways to reduce the noise levels in the sensors to a very low level. In addition to using ultra-high purity silicon wafers, intrinsic gettering techniques are commonly employed to precipitate impurities towards the center of the wafer, giving rise to a 50-75-um-deep denuded zone at the surface on which devices can be fabricated.

One can further reduce the noise by designing highly sensitive output amplifiers. One example of such a highly sensitive amplifier is the floating diffusion electrometer, which was designed at Kodak[4] to achieve an output structure noise level of only 15 RMS electrons. Other amplifier designs have been reported with RMS noise as low as ~1 electron, but at a cost of processing complexity. There are other noise sources in the CCDs that can be diminished by clever design and signal processing techniques. The CCD dark currents have been reduced down to .03 nanoamp per sq. cm. The signal handling capacity of many CCD imagers is such that a dynamic range as high as 70 dB has been reached.

As the size of the pixel becomes small and the number of pixels increases, signal overload (blooming), which is generally a problem with all solid state sensors, becomes even more significantly difficult. For larger pixel formats, it is possible to use simpler schemes such as lateral overflow drain for blooming control. However, with smaller pixel sizes, this is not possible. Therefore, antiblooming techniques using vertical overflow drain become extremely important. These require careful modeling and simulation of several ion implantations and diffusions in order to accomplish a higher blooming control factor.

For many consumer and commercial applications, it is necessary to have single-chip color sensors. This is required not only for keeping the cost of electronic image capture systems down, but to reduce the complexity of the cameras. Once again, there are a number of color filter schemes such as the checkerboard pattern and color stripes. These, too, require extensive modeling and simulation to ascertain the optimization of color filter design for a given system application. It is important that in the design of image sensors, the specifications are well defined so that the above-mentioned sensor considerations can be optimized.

CONCLUSIONS

To summarize, we have seen rapid development in the performance of solid state imaging devices which has paralleled the advances in DRAM technology. This has been accompanied by cost reductions and widespread availability of such sensors for a variety of applications. While extrapolations for future developments in DRAM development are commonly made, it seems unlikely that imaging devices, which are analog in nature, can continue at the same pace. This is due to fundamental limitations such as the absorption length of visible light and to photon counting statistics. Continued progress is, however, expected in device performance and image perfection as the technology for VLSI devices develops.

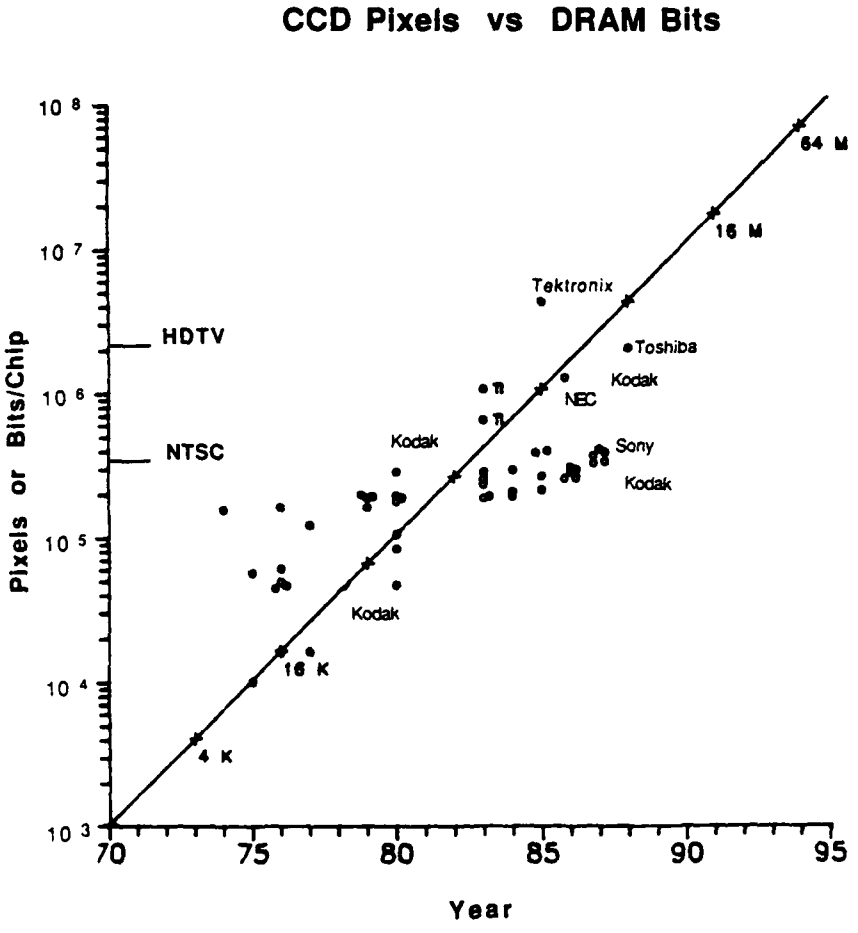


Figure 1 "Moore's Law" Curve Viewing DRAMS vs Year. Imagers are shown by open circles. Imager pixel density has progressed at the same rate as DRAMS.

TABLE I

1/2-INCH FORMAT CCD

COMPANY	DATE	ARCHITECTURE	PIXEL NO. H x V	PIXEL SIZE H x V um
Matsushita	1984	ILCCD	404x506	17.2x10
	1984	ILCCD	404x490	15.1x8.8
Mitsubishi	1985	CCPD	510x485	13x10
	1985	MOS	400x480	16x10
NEC	1986	ILCCD	768x490	8.4x9.8
Philips	1985	Accordian CCD	604x576	10x15.6
	1988	Accordian CCD	806x490	8.05x19.8
Sanyo	1984	FTCCD	400x542	17x10
Sony	1986	ILCCD	510x492	12.2x9.9
Texas Instruments	1986	FTCCD	774x488	8.5x19.8
Toshiba	1986	ILCCD	570x497	11.5x10

TABLE II

MEGAPIXEL SENSORS

COMPANY	DATE	ARCHITECTURE	SIZE	PIXEL NO. H x V	PIXEL SIZE H x V um
Ford Aerospace	1988	FFCCD	1"	1024x1024	12x12
Kodak	1986	FFCCD	2/3"	1320x1035	6.8x6.8
NEC	1986	ILCCD	1"	1280x970	9.9x9.8
Tektronix	1986	FFCCD	55x55mm	2048x2048	27x27
Texas Instruments	1987	FTCCD	2/3"	1024x1024	12x12
Thompson CSF	1987	FTCCD	20x20mm	1024x1024	19x19
Toshiba	1988	IL Si	1"	1420x1036	7.3x7.6

REFERENCES

/1/ Boyle, W. S. and Smith, G. E., "Charge Coupled Semiconductor Devices," 1970 Bell System Technical Journal, Vol. 49, pp. 587.

/2/ Manabe, S. et al., "A 2 Million Pixel CCD Imager Overlaid with an Amorphous Silicon Photoconversion Layer," 1988 ISSCC Digest of Technical Papers, pp. 50-51.

/3/ McCrath, R. and Freeman, J., "An 8 Megapixel/Second 800 x 800 Virtual Phase CCD Imager for Scientific Applications," 1983 IEDM Technical Digest, pp. 489-491.

/4a/Nichols, D. N., "A 1.4 Million Element, Full-Frame CCD Image Sensor with VOFD for Antiblooming and Low Color Crosstalk," 1987 IEDM Technical Digest, pp. 120-123.

/ b/Stevens, E. G., "A 1.4 Million Element CCD Image Sensor," 1987 ISSCC Digest of Technical Papers, pp. 114-115.

Session 3A Room A : Device modeling

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

TIME DEPENDENT BEHAVIOUR OF FIELD LIMITING RING PASSIVATION SYSTEMS

M.K. JOHNSON, D.J. COE, A.D. ANNIS and J.N. SANDOE

Philips Research Laboratories, Cross Oak Lane, Redhill, GB-Surrey RH1 5HA, Great-Britain

Abstract

Off-state modelling is used extensively within Philips to design field limiting ring passivation systems for switching devices. However, no work has been done to ensure that the quasistatic approach is valid in the transient case. This paper presents the results of fundamental simulations of the pulsed operation of a device with a field limiting ring. These are used to describe the ring's time dependent behaviour in detail and so to identify possible pitfalls in the conventional design approach.

Introduction

Field limiting rings represent a technique widely exploited for the edge passivation of discrete power devices. By reducing the electric field from junction curvature and surface effects, a suitable field limiting ring system can give devices a breakdown voltage approaching that of bulk silicon (see, for example, Kao and Wolley [1]).

Presently field limiting ring passivation schemes are designed within Philips using an off-state device simulator and it is assumed implicitly that a device whose electric field distribution is optimised for off-state operation will also work safely under transient conditions. Since many power devices are used in circuits where they are switched at high frequencies (eg switched mode power supplies), the validity of this assumption is critical to the design methodology. Previously no attempt has been made to verify this.

This paper presents fundamental modelling results from the pulsed operation of a device with a field limiting ring. The results are used to illustrate the mechanism by which the floating ring operates under switching conditions and indications are given as to whether such behaviour could ever become dangerous.

Simulation Results

The simulations were carried out at a fundamental level. In other words, the detailed distributions of potential and mobile charge carriers throughout the device were predicted by performing a consistent solution of Poisson's equation and the time dependent electron and hole continuity equations. The device simulated is a p^+n diode with one p^+ floating ring (see Figure 1) which was subjected to two reverse biassing voltage pulses as shown in Figure 2. Figure 3 depicts the resulting current flow, mainly displacement current, and the variation in the ring voltage (ie the voltage at which the ring floats). To explain the behaviour of the ring in this switching device, it is convenient to divide its operation (and hence Figure 3) into eight regions.

When the first pulse commences (region 1), the depletion region around the main junction begins to expand. The current waveform has the expected shape for a diode subjected to a constant voltage ramp. At $\sim 0.2\mu s$ a hump of current occurs. This coincides with the onset of punchthrough between the main junction and the ring and is caused by the increased capacitance of the system. Displacement current is now supplied by both the main junction and the ring junction. Figure 4(a) clearly displays the electric field pattern at the surface of the device during punchthrough. The plot of hole concentration in Figure 5 shows the expanding depletion region around the ring and the space charge produced by the punchthrough current.

Punchthrough current continues at a high level until the main junction voltage stops falling (at the end of region 2). During this time the voltage difference between the main junction and the ring, that is the punchthrough voltage, gradually increases. This "fan-out" effect is due to the two-dimensional nature of the device.

When the main junction voltage stops falling and remains steady at -200V (region 3), the displacement current quickly diminishes to leave only the leakage current of the reverse biassed diode. The ring voltage suffers a further slow fall due to the decay of displacement current between the ring and the main junction and hence the disappearance of the voltage drop required to support it.

As the main junction voltage rises (region 4), the ring voltage rises with it although at a markedly slower rate. Throughout this phase of operation there is no conductive path between the ring and the main junction as can be seen from the electric field distribution at the surface of the device in Figure 4(b) and the hole concentration in Figure 6. This change in the ring voltage is surprising because the only apparent mechanism by which its stored charge (and hence voltage) could rise is by generation of electron-hole pairs in the depletion region. It can be shown that current from such sources is far too small to explain the observed rate of change of the ring voltage. In fact, the cause of the ring's behaviour is simply the electrostatic coupling of two isolated charged bodies, ie Coulomb interaction. So it is not the internal charge on the ring but the electric fields external to the ring which are causing the ring voltage variation.

A change in the rate of rise of the ring voltage coincides with a second sudden jump in current (at the start of region 5). The cause is the reappearance of punchthrough and so a conductive path between the ring and the main junction once more exists. This is demonstrated by the plot of electric field in Figure 4(c) and the plot of hole concentration in Figure 7. This time holes flow onto the ring, since it is more negative than the main junction, and its depletion region slowly collapses. It is this displacement current that causes the sudden increase in the rate of rise of the ring voltage. The fan-out effect is again noticeable.

When the ring voltage reaches 0V (region 6), punchthrough ceases and the ring is left floating at some voltage below ground. The charge on it can now only decay by leakage current due to generation of carriers. So after 10μs when the second pulse is applied there is still a voltage remaining on the ring which has the effect of raising the contact voltage required to initiate punchthrough. This is noticeable in the current waveform as a shift in the time at which punchthrough occurs and the current hump appears (see Figure 8). Note that the effect of electrostatic coupling between ring and main junction reappears during region 7.

To summarise, the device's behaviour shows the following features:-

- the displacement current variation with time shows fine structure superposed on the normal waveform for a diode. This is due to the onset of punchthrough to the field limiting ring.
- due to electrostatic coupling effects, a change in the contact voltage causes a change in the ring voltage even when there is no conductive path connecting the two regions.
- at the end of the first pulse the ring is left charged and hence floating at a voltage below ground.
- fan-out is observed, ie the ring voltage does not change at the same rate as the contact voltage.

The first three effects are purely transient phenomena and cannot be modelled quasistatically by off-state simulations.

Experimental Verification

Experimental observations of displacement current show similar features to those seen in the simulations. Figure 9 shows the current waveform for a BU508 bipolar transistor passivated with field limiting rings, which is subjected to repeated voltage pulses where the contact voltage rises at a rate of 10V/μs. This waveform shows both:-

- fine structure in the form of several humps. These features are much smaller than the ones observed in the simulations because of the smaller capacitance of the rings in the experimental case.
- fine structure moving down the waveform as the pulse repetition rate is increased. Again this is because a short time between pulses means that the residual charge left on the rings at the end of one pulse does not decay significantly before the next.

Thus the experimental observations seem consistent with the simulation results.

Discussion of Results

From a device design point of view, the peak electric field within the device is the quantity of importance since this determines whether the device will breakdown. The variation of this with time is shown in Figure 10. Point B on this waveform gives the electric field predicted by off-state simulations. It can be seen that the electric field peaks twice during the first pulse, at A and C, and **not** at B. The difference in electric field between A and B arises from the voltage drop necessary to support the flow of punchthrough current between the main junction and the ring. Thus the size of this difference will depend on the resistance of the punchthrough path

(ie the device structure), the voltage ramp rate and the capacitance of the ring system. In the simulations shown here this difference is less than 2% and in most devices will be fairly small.

The peak electric field at C is dictated by the voltage required to punchthrough from the ring to the main junction. In most devices this is very close to the voltage needed for punchthrough from the main junction to the ring. However, point C occurs at the end of region 4 on Figure 3 when electrostatic coupling has already lowered the voltage on the ring and the punchthrough voltage has dropped due to the fan-out effect. So, it would require a device with very asymmetric punchthrough voltages for the field at C ever to exceed that at A.

Conclusions

The results from the fundamental simulations have revealed in detail the mechanism underlying the transient operation of devices with field limiting ring passivation systems. This turned out to be more complicated than initially envisaged. Experimental work has reproduced some of the effects seen in the simulation results and hence supports their validity.

The simulations suggest that the electric fields present under transient operation are very close to those in the steady state except when a significant voltage drop is required to drive punchthrough current between the ring and the main junction. This voltage drop is dependent on the resistance of the punchthrough path, the capacitance of the ring system and the voltage ramp rate.

Our improved understanding of the behaviour of field limiting rings in switching devices has lead to greater confidence in the current quasistatic design methodology as well as indicating its potential limitations.

[1] Y. Kao and E. Wolley, "High Voltage Planar p-n Junctions", Proc. IEEE, 55(8), August 1967.

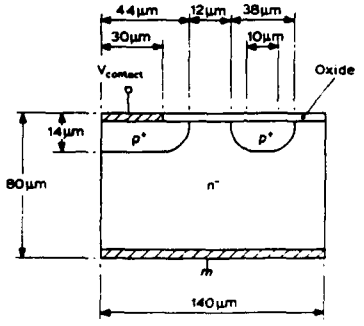


Figure 1. Diode edge passivated with one floating ring

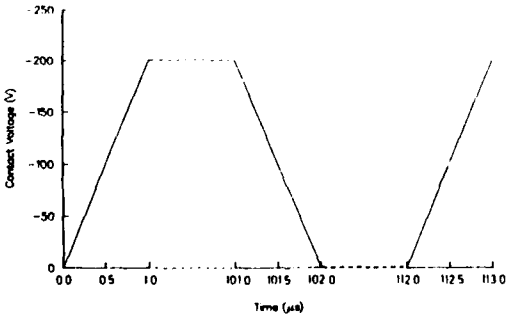


Figure 2. Pulses applied to the main junction contact

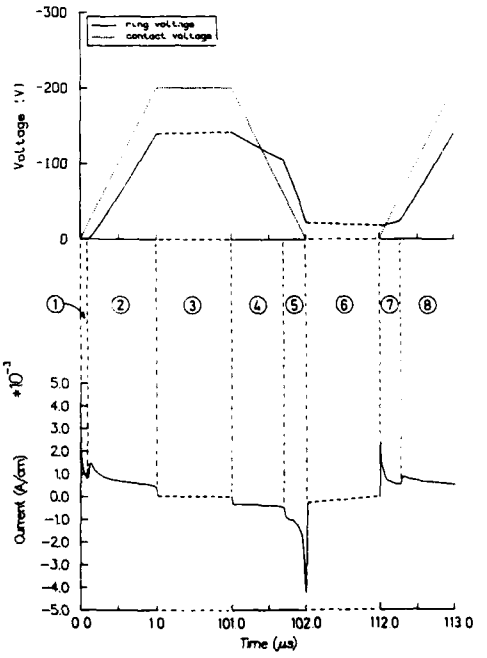


Figure 3. Devices behaviour divided into regions

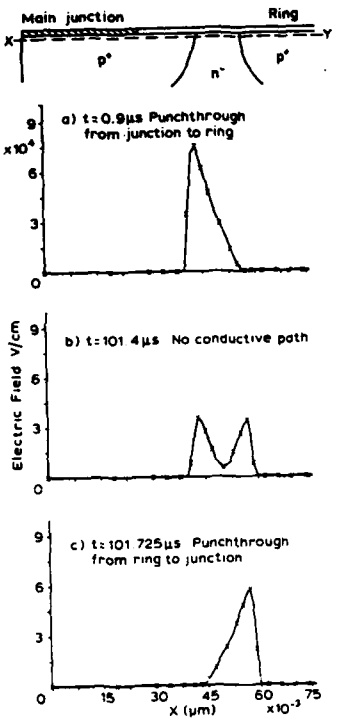


Figure 4. Plots of $|E|$ along the line X-Y.

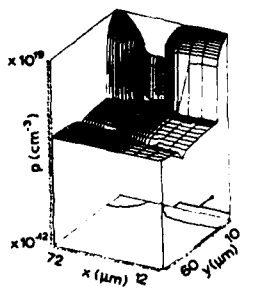


Figure 5. Hole concentration t = 0.7µs

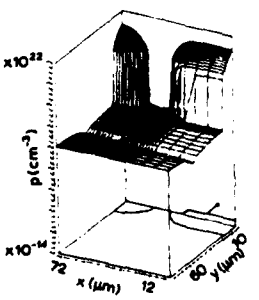


Figure 6. Hole concentration t = 101.4µs

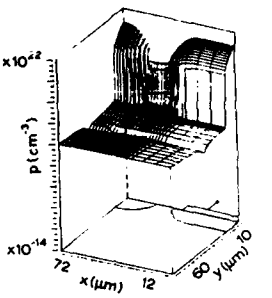


Figure 7. Hole concentration t = 101.725µs

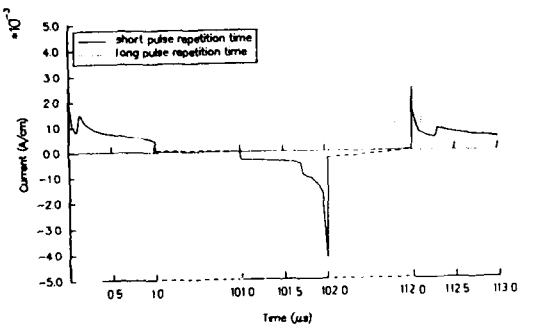


Figure 8. Current v. time.

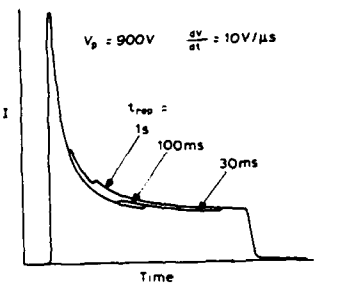


Figure 9. Current-time curves at different repetition times for BU508

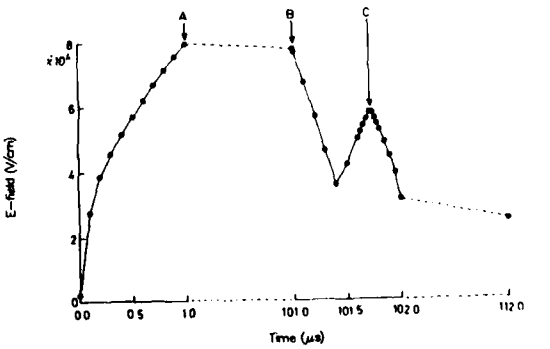


Figure 10. Peak electric field v time

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

3D MOSFET DEVICE EFFECTS DUE TO FIELD OXIDE

M. THURNER and S. SELBERHERR

Institut für Allgemeine Elektrotechnik und Elektronik, Technische Universität Wien, Gusshausstrasse 27-29, A-1040 Wien, Austria

Abstract – This paper presents 3D effects of MOSFET's due to the nonplanar nature of the field–oxide body. The investigations have been carried out by MINIMOS 5 our fully three–dimensional simulation program. Three–dimensional effects like threshold shift for small channel devices, channel narrowing and the enhanced conductivity at the channel edge have been successfully modeled.

1 Introduction

The shrinking dimensions of the elements of IC's require suitable device models in physics and mathematics for accurate simulation. The usual two–dimensional device simulations describe fairly well the electrical characteristics for wide channel transistors but the advanced VLSI technology led to serious problems in modeling very narrow channel devices and therefore a great demand appeared for 3D simulations. The three–dimensional effects in MOSFET's like the shift of the threshold voltage, enhanced conductivity or the large depletion region near drain at the channel edge caused by the finite channel width are not taken into account by the two–dimensional simulations [1]; the 2D programs are meanwhile state of the art. Accurate investigations of the previously stated effects and the knowledge of increased current densities under certain bias conditions at the channel edge are important not only for studying the electrical device characteristics but also for aging effects [2]–[3]. A realistic physical model and suitable mathematical algorithms have been developed to simulate the previously stated three–dimensional effects.
We shall report in Chapter 2 about the physics and the mathematics on which the simulations are based. In addition we shall present the device structure and a consideration of some aspects of the oxide body of the MOSFET.
The results of our simulations carried out by MINIMOS 5 are reported in Chapter 3 and will be

discussed there, too.

2 Physical and Mathematical Aspects

The basic equations which are implemented in our MINIMOS to describe current flow in silicon differ only slightly from the conventional equations. The Poisson equation and the continuity equations for electrons and holes are 'the' established basic equations which are commonly in use. A derivation of these equations can be found in e.g. [7]. The current relations for electrons and holes are somehow different from the conventional drift–diffusion relations. They include the hot electron transport effect. Detailed information on the 'hot–electron–transport' model can be found in [6,9]. In addition to the well known boundary conditions for the Poisson equation and the continuity equations we have to implement a boundary condition for the mobilities. We set the driving force F as in (1).

$$\vec{F}_n|_{int} = 0 \tag{1}$$

If we neglect this boundary condition we get unrealistic mobilities near the interfaces. For solving the set of differential equations we apply for discretisation the box integration method after Forsythe [10] to deal with the boundary conditions of the nonrectangular simulation region. The linearized equations are solved by the SOR method and the Gauß–elimination [4],[5],[8]. The geometry of the 3D MOSFET model is given in Fig. 1. We implemented an approximation to the complete oxide–volume which can be seen in Fig. 2. The upper plane denotes the interface of the oxide to the contacts and the surrounding volume, respectively. The lower plane denotes the interface of oxide and semiconductor. The definition of the interface is quite general and can be varied in a wide range. The gate contact is filled in the middle of the upper plane, the Source and Drain contacts are on the left and right. In the middle in the gate region the distance of the upper and lower plane equals the gate–oxide thickness. The drain currents calculated by 2D and 3D simulations cannot be compared in a straight forward manner if the field oxide is nonplanar. The oxide reduces gradually the channel width which is given by the mask specifications. Therefore we introduce an effective channel width w_{eff} for the 2D current calculation instead of the given mask width. The effective channel width is calculated by (2)

$$w_{eff} = w - \frac{1.}{D_{ch}} \int_0^w fox(z)|_{l/2}\, dz \tag{2}$$

w is the width specified by the mask, D_{ch} denotes the channel depth, $l/2$ denotes half of the channel length, $fox(z)$ is the function which describes the geometry of the channel in width direction.

3 Results and Discussion

The geometry of the investigated 3D MOSFET is given in Fig. 1: an n–channel MOSFET with an $1\mu m$ x $1\mu m$ channel and gate oxide of $15nm$. For demonstrating the effects at the channel edge we select two different bias points. The first is near threshold with $U_S = U_B = 0.0V$, $U_D = 3.0V$, $U_G = 1.0V$ (the threshold voltage for this device is $U_{th} \sim 0.75V$). The potential distribution in channel length and width direction at the semiconductor/gate–oxide interface is shown in Fig. 3. (This plane penetrates into the field oxide near the contact boundary of Source and Drain.) The corresponding minority carrier distribution is given in Fig. 4. A remarkable depletion region at the drain side causes the channel charge to be smaller (under certain bias conditions) than predicted by 2D simulations. The second bias point is far above threshold $U_S = U_B = 0.0V$, $U_D = 1.0V$, $U_G = 3.0V$. The corresponding potential distribution can be seen in Fig. 5. The location of the

plane for which the distribution is drawn, is the same as at the previous bias condition. The high increase of the potential distribution out of the channel is due to the gate contact overlapping the field oxide. Also interesting is the minority carrier distribution (Fig. 6), which shows the enhanced conductivity at the semiconductor field–oxide interface. Note that only one half of the channel width is shown in Fig. 3 – Fig. 6; $-0.5\mu m$ denotes the middle of the channel width and $0.0\mu m$ the boundary of Source and Drain contacts. The effect on the device characteristics of these effects depends on the gradient of the bird's beak and the channel width. A high gradient in the field oxide shape results in high parasitic current at the channel edge; this effect is less significant for low gradients. Narrow channel devices with high gradient have much higher currents than predicted by 2D calculations while the agreement with 2D simulations is good for wide channel devices in any case. Using a low gradient in bird's beak we get a very smooth potential distribution compared to a nearly rectangular shape.

Acknowledgement

This work was supported by the research laboratories of SIEMENS AG at Munich, FRG, by DIGITAL EQUIPMENT CORP. at Hudson, USA, and by the "Fonds zur Förderung der wissenschaftlichen Forschung", project S43/10. We are indebted to Prof. H. Pötzl for many helpful discussions.

References:

[1] S.M. Sze, Physics of semiconductor devices, ISBN 0-471-09837-X, John Wiley & sons, 1981

[2] S.M. Sze, VLSI–Technology, ISBN 0-07-062686-3, McGraw-Hill, 1983

[3] T. Iizuka,K.Y. Chiu, and J.L. Moll, Double threshold MOSFETs in bird's–beak free structures, IEEE Int. Electron Device Meet.,Wash.,D.C., 1981, p. 380

[4] L.A. Hageman,Franklin T. Luk,David M. Young, On the equivalence of certain iterative acceleration methods, SIAM J.NUMER.ANAL., pp 852-873, vol. 17 No. 6, Dec 1980

[5] R.G. Grimes,D.R. Kincaid,D.M. Young, ITPACK 2A – A fortran implementation of adaptive accelerated iterative methods for solving large sparse linear systems, Report CNA-164, Center for numerical analysis, University of texas at Austin, 1980

[6] W. Hänsch, M. Miura–Mattausch, A new current relation for hot electron transport, Proc. NASECODE IV Conf.,pp.311-314, Boole Press, Dublin, 1985.

[7] S. Selberherr, Analysis and simulation of semiconductor devices , ISBN 3-211-81800-6, Springer, WIEN NEW-YORK, 1984

[8] O. Axelsson, Solution of linear systems of equations;Iterative methods, Lecture notes in mathematics 574, SMT, 1976

[9] S. Selberherr, The status of MINIMOS, Proc. Simulation of semiconductor devices and processes, pp 2-15, Swansea, 1986

[10] G.E. Forsythe,W.R. Wasaw, Finite Difference Methods for Partial Differential Equations. Wiley, NEW-YORK, 1960

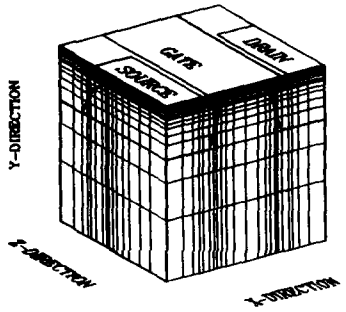


Fig.1: Perspective view of the three-dimensional MOSFET structure.

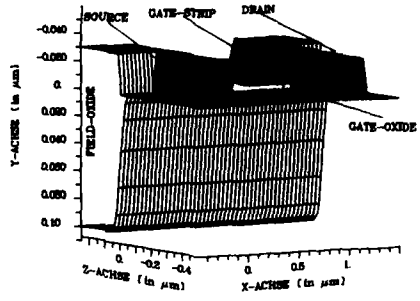


Fig.2: Oxide body of the MOSFET structure (the oxide is between the upper and lower plane).

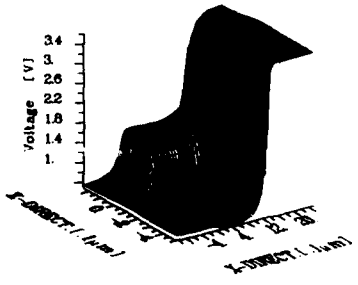


Fig.3: 3D-plot showing a detailed view of the surface potential at the channel edge along the channel length at bias $U_{DS} = 3.0V$, $U_{BS} = 0.0V$, $U_{GS} = 1.0V$.

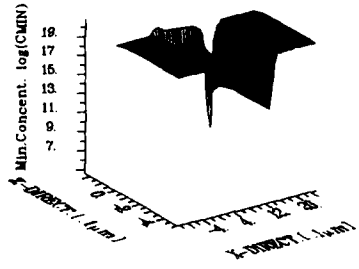


Fig.4: 3D-plot showing a detailed view of the minority carrier density at the channel edge along the channel length at bias $U_{DS} = 3.0V$, $U_{BS} = 0.0V$, $U_{GS} = 1.0V$.

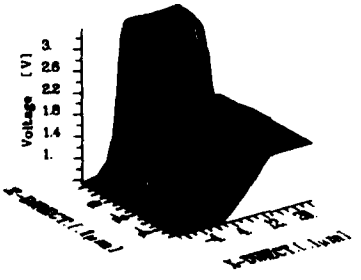


Fig.5: 3D-plot showing a detailed view of the surface potential at the channel edge along the channel length at bias $U_{DS} = 1.0V$, $U_{BS} = 0.0V$, $U_{GS} = 3.0V$.

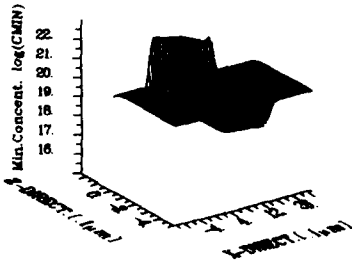


Fig.6: 3D-plot showing a detailed view of the minority carrier density at the channel edge along the channel length at bias $U_{DS} = 1.0V$, $U_{BS} = 0.0V$, $U_{GS} = 3.0V$.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

NOVEL CALCULATIONS IN THE FIELD OF ACCURATE ANALYTICAL MOS TRANSISTOR MODELLING

L. LAUWERS and K. DE MEYER[1]

IMEC, Kapeldreef 75, B-3030 Leuven, Belgium

Résumé: Des calculs analytiques détaillées, à partir de quelques modèles physiques de base, prouvent que le courant dans un transistor à canal court peut être précisément déterminé, sans que le nombre de paramètres empiriques augmente. En plus, il est démontré que quelques unes des approximations, sur lesquelles beaucoup de modèles actuels sont basés, sont incorrectes, et que les adaptations reliées manquent de base physique.

Abstract: Extended analytical calculations, starting from basic physical models with only few assumptions made, prove that the drain-source current for a wide range of MOS transistor geometry, can be accurately determined using one single model, without enlarging the number of empirical parameters. Furthermore, it is proven, that some of the assumptions, on which lots of actual models rely, are incorrect, and the related empirical improvements commonly in use lack physical background.

1. Situation

In most circuit simulators today relatively simple analytical transistor models are being used. These models do compromise accuracy with respect to minimizing computational effort mainly because of the large amount of transistors to simulate. However, the urge for reliable simulation of complex VLSI circuits emphasizes the need for more accurate analytical device models valid over a broader range of device dimensions and operating conditions. In current models too often one still relies on approximations with very limited applicability and which even already have become invalid for a long time.

In this work detailed calculations with only a limited amount of assumptions and simplifications were made in order to obtain accurate short channel MOS model equations, suitable for medium-size digital designs, as well as analog applications and device investigation (e.g. parameter sensitivity analysis). We propose new threshold and mobility models, valid for both long and short channel devices, based on sound physical principles.

2. Model improvements

We started from the basic device physics models, which are commonly used and at this moment still prove their validity, as long as one doesn't make wrong approximations. Carrier mobility and velocity are well described as a function of the lateral field E_{lc} and normal field E_{neff} at a certain point c in the channel [1][2] by:

$$v_c = \frac{\mu_c . E_{lc}}{1 + \frac{\mu_c . E_{lc}}{v_{max}}} \qquad \text{with } E_{lc} = \frac{\Delta V_c}{\Delta c}$$

and

$$\mu_c = \frac{\mu_0}{R(c)} = \frac{\mu_0}{1 + \Theta . E_{neff}} \qquad \text{with } E_{neff} = \frac{\zeta . Q_{inv}(c) + Q_{dep}(c)}{\varepsilon_{si}}$$

In each point between source and drain, an effective normal field, E_{neff}, is considered, which is typical for the mobility reduction at that point. Assuming E_{neff} to be the average of the surface field at the interface and the depletion field at the back of the inversion layer, ζ=0.5 [5]. However, we found ζ to be quite lower, which reduces the weight of the inversion charge in reducing the mobility.

The sheet current J_L, written as $J_L = Q_{inv}(V_c).v_c$, can now be integrated in Schockley-style[1][8], while the mobility reduction term R(c) is now considered to be dependent on the position in the channel. Hence, the current equation depends on expressions of both inversion and depletion charge integrated from source to drain.

---

[1] Professor at the *Katholieke Universiteit Leuven* and Research Associate of the *Belgian National Found for Scientific Research.*

$$J_L = \mu_0 \frac{1}{1+\frac{\mu_0}{v_{max}L}\int_0^{V_D}\frac{dV_c}{R(V_c)}} \int_0^{V_D} \frac{Q_{inv}(V_c)}{R(V_c)} dV_c \qquad \text{with } R(V_c) = 1+\theta.(\zeta.Q_{inv}+Q_{dep})$$

At this point, three basic ideas do fundamentally differ from classical calculations in this area, and enable us to achieve high model accuracy, without introducing any new parameter.

First, it is emphasized that both inversion and depletion charge, given by

$$Q_{inv}(c) = C_{ox}(V_{GS} - V_{TH}(c)) \qquad \text{with } V_{TH}(c) = V_{FB} + \psi(c) + \gamma(c)\sqrt{\psi(c)}$$

$$Q_{dep}(c) = C_{ox}\,\gamma(c)\sqrt{\psi(c)}$$

where $\psi(c) = 2\phi_F - V_{BS} + V_c$ is the local potential drop from the surface to the bulk, do vary from source to drain, as V_c, being the channel potential, raises from 0V at the source to V_{DS} at the drain.

Furthermore, it can be seen that, as γ is determined by the average doping level for the local depletion layer depth, a variable substrate doping concentration invalidates the use of a single bulkfactor γ, not only for different V_{BS} biases, but also for each applied drain-source voltage. Therefore, we developed a satisfactory relationship between γ and ψ, which causes the bulk-factor to be dependent on both V_{BS} and V_c:

$$\gamma(c) = \gamma_1 + \gamma_2.(\sqrt{\psi(c)}-\sqrt{2\phi_F}) + \gamma_3.(\psi(c) - 2\phi_F)$$

This three-parameter relationship offers an appropriate γ for all depths of the local depletion layer, within the range where γ_1, γ_2 and γ_3 are fitted on data. γ_1 keeps its physical meaning as the *surface bulk-factor at the source*.

Finally, classical threshold voltage models are linearized, using the Taylor series expansion for $\sqrt{\psi(c)}$ [3][4], where some correction factors have to account for the errors made in neglecting higher order terms, and using the expansion outside its convergence region. This doubtful technique can simply be avoided by taking $W_c = \sqrt{\psi(c)}$ as new integration variable. W_c then varies from $W_0 = \sqrt{2\phi_F - V_{BS}}$ at the source, to $W_D = \sqrt{2\phi_F - V_{BS} + V_{DS}}$ at the drain.

Taking above considerations rigourously into account, leads to the following formulation of the current:

$$I_{DS} = \mu_0 C_{ox}\frac{W}{L}.\frac{1}{1+\frac{\mu_0}{v_{max}L}\int_{W_0}^{W_D}\frac{2W_c}{R(W_c)}dW_c} \int_{W_0}^{W_D}\frac{T(W_c)\,2W_c}{R(W_c)}dW_c$$

where T(W) and R(W) are polynomials in W_c:

$$T(W_c) = T_A - T_B W_c - T_C W_c^2 - T_D W_c^3 \qquad W_c = \sqrt{\psi(c)}$$

$$R(W_c) = R_A + R_B W_c + R_C W_c^2 \qquad dV_c = 2W_c\,dW_c$$

In both numerator and denominator polynomial, only the first terms, T_A and R_A, are dominant, and depend on the applied gate- and bulk- voltage. The higher order coefficients only depend on the model parameters:

$$T_A = V_{GS} - V_{FB} - 2\phi_F - \gamma_1\sqrt{2\phi_F} + \sqrt{2\phi_F - V_{BS}} \qquad R_A = 1 + \theta\zeta T_A$$

$$T_B = \gamma_1 - \gamma_2\sqrt{2\phi_F} - \gamma_3\,2\phi_F \qquad R_B = \theta(1-\zeta)T_B$$

$$T_C = 1 + \gamma_2 \qquad R_C = \theta(1-\zeta)T_C - 1$$

$$T_D = \gamma_3$$

The analytical solving of this current integral, which is not explained here, leads to:

$$I_{DS} = C_{bas}.\ln\frac{P(W_0)}{P(W_D)} + C_{red}.\ln\frac{R(W_D)}{R(W_0)} + C_{W1}.(W_D - W_0) + C_{W2}.(W_D^2 - W_0^2) + C_{W3}.(W_D^3 - W_0^3)$$

with:

$$P(W_c) = 1 + \frac{2.\sqrt{R_B^2 - 4R_AR_C}}{2R_C\,W_c + R_B - \sqrt{R_B^2 - 4R_AR_C}}$$

$$C_{bas} = 2(-R_{BC}\,C_{Ta} + (R_{BC}^2 - 2R_{AC})\,C_{Tb})/\sqrt{R_B^2 - 4R_AR_C} \qquad R_{AC} = R_A/R_C$$

$$C_{red} = (C_{Ta} - R_{BC}\,C_{Tb})/R_C \qquad R_{BC} = R_B/R_C$$

$$C_{W1} = 2C_{Tb}/R_C$$

$$C_{W2} = (-T_C + R_{BC}\,T_D)/R_C \qquad C_{Ta} = T_A + R_{AC}\,T_C - R_{AC}R_{BC}\,T_D$$

$$C_{W3} = -2T_D/3/R_C \qquad C_{Tb} = -T_B + R_{BC}\,T_C + (R_{AC} - R_{BC}^2)\,T_D$$

The first term is the basic current carrying term which, due to mobility reduction, results in a logarithmic function of V_{DS}. In the first order zero-reduction case R(c)=1, this term limits to $(V_{GS}-V_{TH})V_{DS}$. The second logarithm, which is the basic mobility reduction term, then tends to 0. The third to fifth term are adjusting reduction terms, primarily dependent on respectively T_B, T_C and T_D.

3. Implementation in a CAD-model and remarks on data-fits

After introduction of some classically modelled small geometry effects, e.g. drain induced barrier lowering and the influence of the bird's beak [3], this equation is implemented in a full CAD-model, where, for device-investigation purposes, several options for the saturation region and -voltage are provided [3]. The saturation voltage is determined to obtain continuous slope in the transition region.

Fittings have been carried out on several I-V data sets by our parameter extraction program SIMPAR [6]. In the saturation region, also a slope fitting is performed, which is adequate for analog applications.

I_{DS}-V_{GS}, I_{DS}-V_{DS} fittings for the new model are shown on figs.1 to 6, where the simulated curves (dotted) are compared with the measured data (full lines). The relative RMS-residuals are never above 1%, for short channel as well as for long channel devices, both for nMOS and pMOS.

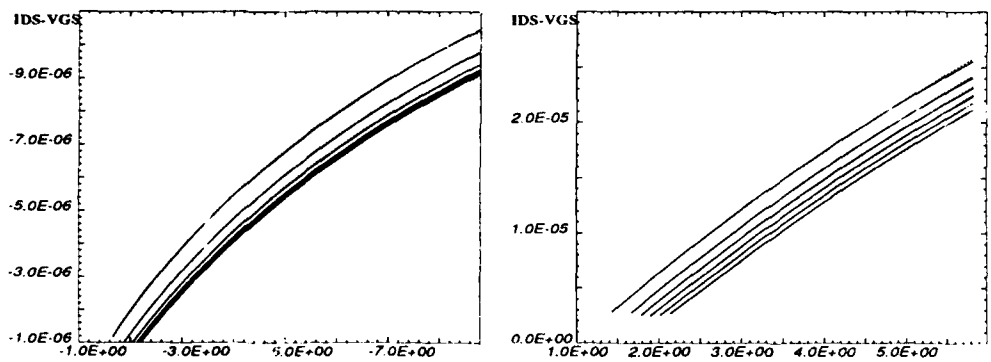


Fig. 1,2: I_{DS}-V_{GS} data fitting with the new model for long channel pMOS (left) and nMOS (right); t_{ox}=23n; W_m=20µ, L_m=20µ

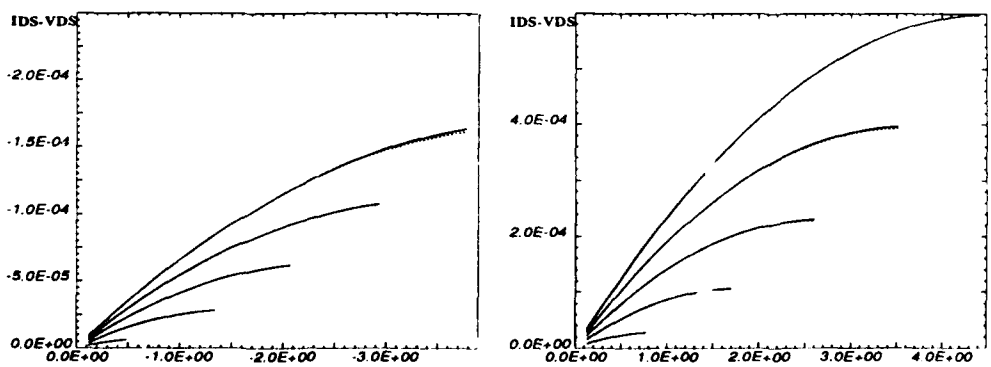


Fig. 3,4: I_{DS}-V_{DS} data fitting in the triode region with the new model for long channel pMOS (left) and nMOS (right); t_{ox}=23n; W_m=20µ, L_m=20µ

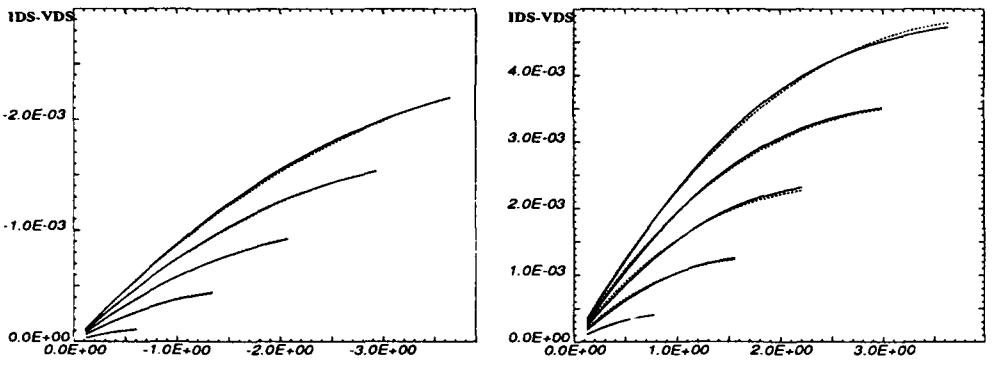


Fig. 5,6: I_{DS}-V_{DS} data fitting in the triode region with the new model for short channel pMOS (left) and nMOS (right); t_{ox}=23n; W_m=20μ, L_{eff}=1.3μ

As we compare this formula with other models, it can first be said that some physical parameters, e.g. mobility and threshold voltage, loose their unique device definition, as their values and effects have been locally taken into account. Furthermore, in view of the obtained accuracy in a wide range of data, applied voltages and device geometries. one can conclude that earlier commonly made approximations, were too rough, or restricted to a small validity range. Therefore, it is not surprising that parameters that resulted from these approximations, f.i. FB, α,... [3] kept causing problems either in the short channel or long channel model.

It is remarked that, besides by the series resistance and the saturation velocity, the current in the whole triode region is only influenced by two small geometry parameters, modeling drain induced barrier lowering and the bird's beak. No other parameters are involved.

4. Implementation in a circuit simulator

The full CAD model has been integrated for testing purposes in the circuit simulator ELDO [7]. Simulations are now being done, and some results will be presented at the conference.

5. Conclusions

The fitting results on several data sets are very accurate. The major importance of these detailed calculations is that we can now rely on a closed current formula, which performs unseen accuracy in analytical modelling. One can take this formula as a starting and reference point for making ones own approximations, possibly deriving a simplier expression the accuracy of which could be acceptable for certain operation conditions.

Acknowledgments

Our thanks are due to W.Maes, L.Dupas, E.Machiels and W.Magnus for various support to this work. We like to thank the CNET, France, for assistance with the implementation of the CAD model in ELDO. Part of this work has been supported by the Everest-Esprit project 962.

References

[1] B.T.Murphy: *A unified field-effect transistor theory including velocity saturation*, IEEE-JSSC, SC-15, jun1980, pp325-327
[2] S.C.Sun,J.D.Plummer: *Electron mobility in Inversion and Accumulation layers on thermally oxidized silicon surfaces*, IEEE-JSSC, SC-15, aug1980, pp562-573
[3] S.Liu: *A unified Mosfet CAD-model*, Berkeley, 1981
[4] G.Merckel, J.Borel, N.Z.Cupcea: *An accurate large-signal MOS-transistor model for use in CAD*, IEEE-TED, ED-19, 1972, pp681-690
[5] G.T.Wright: *Physical and CAD-models for the implanted channel VLSI Mosfet*, IEEE-TED, ED-34, apr1987, pp823-833
[6] W.Maes, K. De Meyer, L.Dupas: *SIMPAR: A versatile technology independent parameter extraction program using a new optimized fit strategy*, IEEE-TCAD, CAD-5, apr1986, pp320-325
[7] B.Hennion, P.Senn: *ELDO: A new third generation circuit simulator using the one step relaxation method*, proc.ISCAS, 1985, pp1065-1068
[8] S.M.Sze: *Physics of semiconductor devices*, Murray-Hill, Wiley

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A NEW ANALYTICAL AND STATISTICAL-ORIENTED APPROACH FOR THE TWO-DIMENSIONAL ANALYSIS OF SHORT-CHANNEL MOSFET's

M. CONTI, C. TURCHETTI and G. MASETTI*

Dept. of Electronics University of Ancona, v. Brecce Bianche, I-60131 Ancona, Italy
**DEIS University of Bologna, v. le Risorgimento 2, I-40136 Bologna, Italy*

ABSTRACT

An approximated analytical solution of Poisson's equation for the short-channel MOSFET operating in the subthreshold regime is presented. It is shown that the proposed approach predicts a dependence of the threshold voltage on process parameters and drain and substrate voltages in very good agreement with two-dimensional analysis and with available experimental data. Finally, the method of this work, which permits to gain a factor of about 10^3 in CPU time with respect to numerical modeling for threshold predictions, seems particularly suited for statistical modeling.

1. INTRODUCTION

In order to improve circuit performances and to reduce chip size, devices used in today MOS VLSI circuits have very small channel lengths and/or channel widths. Besides, when device dimensions are scaled down, second-order phenomena, such as drain induced barrier-lowering and influence of the curvature of the source and drain junctions [1-2], become essential in establishing the electrical behaviour of the MOSFET. As a consequence, the device performances should be determined, in principle, by the adoption of two- (or even three-) dimensional device simulators (see, f.i., [3-6]).

However, these simulators are very time consuming and, thus, very cost effective when multiple analyses, such as those required for statistical modeling, must be performed. On the other hand and as well known: i) in order to develop designs with the aid of circuit simulators, circuit designers require very simple (even if strongly approximated) analytical device models, while ii) in most cases, device designers, to get a general view of the way a process is going on or to evaluate the spread existing between the characteristics of devices processed with the same or with different runs, are interested in the knowledge of only a few device parameters of the MOSFET (such as threshold voltage and gain factor).

The purpose of this paper is to derive, through the weighted-residual method [8], an approximated analytical solution of the two dimensional Poisson's equation for a short-channel MOSFET operating in the subthreshold region. In particular, it is shown that the proposed approach is able to predict a distribution of the surface potential along the channel close to that achievable with the two-dimensional simulator HFIELDS [6]. Besides, this approach can also satisfactorily predict the dependence of the threshold voltage in short-channel devices for a wide range of values of both device parameters (such as oxide thickness, junction depth and substrate doping) and bias voltages. Finally,it is shown that with the analysis of this work threshold voltage predictions can be achieved with a reduction of a factor of about 10^3 in CPU time with respect to standard numerical analysers [6]; this result enables to qualify the approach as a cheap method to predict the sensitivity of threshold voltage on fabrication tolerances in short-channel MOSFETs.

2. MODEL FORMULATION

For a device operating in the subthreshold regime, by neglecting carrier concentrations Poisson's equation can simply be expressed as

$$\frac{\partial^2 \phi(x,y)}{\partial x^2} + \frac{\partial^2 \phi(x,y)}{\partial y^2} = \begin{cases} +\dfrac{q N_a}{\varepsilon_s} & \text{in the semiconductor} \\ 0 & \text{in the oxide} \end{cases} \qquad (1)$$

where ϕ is the electrostatic potential, N_a is the bulk doping and ε_s the permittivity of silicon.

Eqs. 1 , once defined the proper boundary conditions, can be solved by means of standard numerical procedures [3-6]. However, some approximate mathematical methods that reduce partial differential equations to a set of ordinary differential equations are available and more easily solved. One of the most powerful methods is the so-called weighted residual approach [8]. Following this technique we can seek for an approximate solution $\phi'(x,y)$

to (1) in the linear form

$$\phi(x,y) \approx \phi'(x,y) = \sum_{j=1}^{n} a_j \phi_j(x,y) \qquad (2)$$

where $\phi_j(x,y)$ are independent arbitrary functions to be properly chosen and a_j are coefficients to be determined.

The errors, or residuals (R_{Si} and R_{ox}), introduced by approximation (2) can thus be defined for eq. (1) as

$$R_{si} = \frac{\partial^2 \phi'}{\partial x^2} + \frac{\partial^2 \phi'}{\partial y^2} - \frac{q N_a}{\varepsilon_s} \qquad \text{in the silicon} \qquad (3a)$$

and

$$R_{ox} = \frac{\partial^2 \phi'}{\partial x^2} + \frac{\partial^2 \phi'}{\partial y^2} \qquad \text{in the silicon-dioxide} \qquad (3b)$$

Usually, to reduce the residuals, a suitable number of integrals of the error, weighted in different ways through some weighted parameters W_k, are required to be zero, that is

$$\int_{\Gamma} W_k R_{\Gamma} \, d\Gamma + \int_{\Omega} W_k R_{\Omega} \, d\Omega = 0 \qquad k = 1, 2, \ldots, n \qquad (4)$$

where Γ is the boundary of the domain Ω in which equations (1) have to be solved,while R_{Γ} and R_{Ω} are the residuals evaluated along Γ and inside Ω, respectively.

In this work, in order to get a relatively simple and low-time consuming approach, we choose to fit the electrostatic potential with a simple polynomial in x and y with independent coefficients, i.e.

$$\phi'(x,y) = \sum_{j=0}^{n} \sum_{k=0}^{m} a_{jk} x^j y^k \qquad \begin{matrix} 0 \le n \le 2 \\ 0 \le m \le 3 \end{matrix} \qquad (2')$$

Morever, always in the light to achieve a simple analytical formulation, we made a choice of the relationship (2') where, by imposing the boundary conditions, all the coefficients result as functions of only one of theme. This results in a strong simplification of the analysis which reduces the set of equations (4) for the residuals to one equation only. As a consequence, the evaluation of ϕ' results dependent on the evaluation of one parameter a_{jk} only.

Now, due to the very simple formulation (2') for ϕ' , the application of the weigthed-residual method gives a simple linear equation for the minimum value ϕ_{smin} of the surface potential:

$$\phi_{smin} = \alpha V_{GS} - \beta \qquad (5)$$

where α and β are numbers which depend on device geometry, device physical parameters (N_a, t_{ox}, V_{FB}) and on bias voltages.

3. SURFACE POTENTIAL AND THRESHOLD BEHAVIOUR

An example of the behaviour achievable for the surface potential along the channel of the device $\phi_s(y)$ by using the approach presented in this work is shown in Fig. 1. Device parameters were L=1.5 μm, $N_a = 10^{16}$ cm^{-3}, $N_d = 10^{18}$ cm^{-3}, t_{ox}= 25 nm, x_j= 0.25 μm, V_{BS}= -2 V, V_{DS}= 1 V, $V_{GS} - V_{FB}$= 1.2 V. In the same figure, for comparison, the behaviour of the surface potential achieved with the two-dimensional analyzer HFIELDS [6] is shown as dotted line. As can be seen, the agreement between the two behaviours is resonably good and, in particular, the values found for the minumum value ϕ_{smin} of the surface potential are pratically coincident. As far as the CPU time is concerned, for the results of Fig. 1 we found that our approach enables a reduction in computer time of factor 1: 1000 with respect to the calculations required from HFIELDS. This factor represents also an average value for the CPU time ratio for all the other comparisons we made by varying device parameters in the following ranges: L= 0.6 / 10 μm, $N_a = 4 / 50 * 10^{15}$ cm^{-3}, t_{ox}= 10 / 70 nm, x_j= 0.15 / 0.45 μm, V_{BS}= 0 / -4 V, V_{DS}= 0.1 / 5 V.

4. THRESHOLD VOLTAGE PREDICTIONS

In MOS technology a parameter extremely useful to "conventionally" define the conduction state of a single device and to identify process variations throughout a wafer and among chips built on different wafers with the same run (or different runs) is the threshold voltage. For the sake of simplicity and a) to take into account that the carrier injection in the channel of the MOSFET almost exponentially depends on the variations of the barrier-height at the source-channel junction with respect to its value at equilibrium and b) to account for the strong dependence of such a barrier-height on device parameters and bias voltages in short-channel devices, we defined the threshold voltage V_{TH} as the gate-source voltage which gives a certain value ϕ^* for the surface potential ϕ_{smin} . So doing we get:

$$V_{TH} = V_{FB} + \frac{\phi^*}{\alpha} + \frac{\beta}{\alpha} \qquad (6)$$

Fig.2 shows, as continuous lines, the threshold-voltage value, calculated from (6) and using $\phi^* = 2\,\phi_F$ (ϕ_F being the Fermi-voltage), as a function of channel length for several values of the body bias. Device parameters were: N_a= 10^{16} cm^{-3}, t_{ox}= 25 nm, x_j= 0.25 μm, V_{DS}= 5 V. The results achieved with CADDET [7] are shown as dots in the figure. As can be seen, the agreement between comparison is very good. Similar conclusions can be drawn as far as the dependence on oxide thickness, substrate doping and drain voltage; a comparison between the results predicted from the model of this work and the numerical simulations is shown for three values of the oxide thickness in Fig. 3.

Finally, a comparison between the results predicted from our model and experimental data [7] is shown in Fig. 4. As can be seen, for the examined channel-length and oxide-thickness range, the proposed model satisfactorily predicts the threshold behaviour.

5. STATISTICAL ANALYSIS OF THRESHOLD VOLTAGE

Due to the very low CPU time required, the approach presented in this work seems particuraly suited for statistical analysis of the threshold voltage in short-channel MOSFETs. In fact one can assume that threshold voltage given by eq.(6) is a random function of some process variables

$$V_{TH} = V_{TH} (L, t_{ox}, N_a, x_j) \qquad (7)$$

Then, by considering that the process variables can be characterized with a certain probability distribution, one can easily derive the distribution probability of threshold voltage. In particular we assumed a gaussian distribution for all the above variables and calculated the distribution $p(V_{TH})$ of the probability density associated to the threshold voltage, the mean value and the root-mean square error of the threshold.

Fig. 5a shows, for three values of channel length, the $p(V_{TH})$ distribution achieved by considering a constant mean-square error in channel length of 0.1 μm. The other device parameters were considered constant and equal to N_a= 10^{16} cm^{-3}, t_{ox}=25 nm, x_j= 0.25 μm, V_{DS}= 5 V, V_{BS}= -2 V. As can be seen, as the channel length decreases the $p(V_{TH})$ distribution becomes more broadened with a more pronounced asymmetry toward lower values of threshold. These conclusions are fully confirmed from the results shown in Fig. 5b, where the mean value and the value of the root-mean square error of the threshold voltage are reported against channel length. In particular, for the process considered, one can derive an increase from 0.07 % to 3.5 % in the ratio between the root-mean square error and the mean value of the threshold voltage when the channel length is decreased from 5 μm to 1 μm.

Similar conclusions can also be drawn by considering a gaussian distribution for the oxide thichness, the substrate doping and the junction depth.

6. SUMMARY AND CONCLUSIONS

In this work the weighted-residual method was utilized to find an approximated analytical solution of the two dimensional Poisson's equation in short-channel MOSFETs. In particular, it was shown that the proposed method:

i) is able to predict a distribution of the surface potential along the channel very close to that achievable with two-dimensional simulators;

ii) can satisfactorily predict the dependence of the threshold voltage in short channel devices for a wide range of variation of process parameters (such as channel length, oxide thickness, substrate doping and junction depth) and bias voltages;

iii) for threshold evaluation it requires a CPU time lower of about three order of magnitude with respect to two-dimensional simulators.

Finally, due to property iii) above, the analysis was extended to find the probability distribution of the threshold voltage induced by variation in process parameters. In particular, results were found for gaussian distributions of channel length, oxide thickness, substrate doping and junction depth.

ACKNOWLEDGEMENTS

This work has been supported by the Consiglio Nazionale delle Ricerche under the Project: "Materiali e Componenti per l'Elettronica a Stato Solido". Support from SGS-THOMSON Microelectronics is also gratefully acknowledged. Thanks are also due to the numerical analysis group of the University of Bologna for the support in the usage of HFIELDS code.

REFERENCES

[1] R.R.Troutman, *IEEE Trans. Electron Devices*, vol. ED-26, pp.461, 1979.

[2] B.Eitan and D.Frohman-Bentchkowsky,*IEEE Trans. Electron Devices*, vol. ED-29, pp.254, 1982.

[3] A.Husain and S.G.Chaimberlain,*IEEE Trans. Electron Devices*, vol. ED-29, pp.631, 1982.

[4] W.I.Engl et al, *Proc. IEEE*, vol. 71, pp. 10, 1983.

[5] S.Selberherr, "Analysis and simulation of semiconductor devices", Ed. Springer, Wien 1984.

[6] G.Baccarani et al, *Proc. of the Fourth International Conference on the Numerical Analysis of Semiconductor Devices and Integrated Circuirts*, J.J.H.Miller Ed., Dublin: Boole Press, pp. 3, 1985.

[7] T.Toyabe and S.Asai, *IEEE Trans. Electron Devices*, vol. ED-26, pp. 453,1979.

[8] B.A.Finlayson, "The Method of Weighted Residuals and Variation Principles", New York: Academic, 1972.

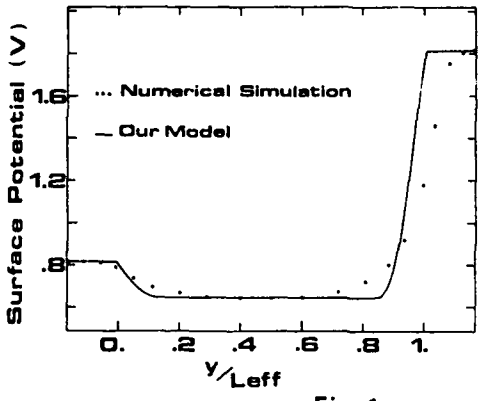


Fig. 1

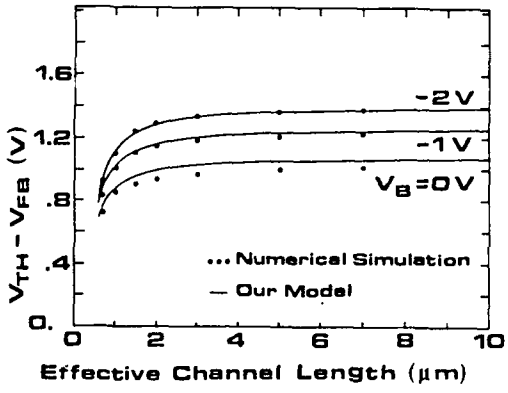


Fig. 2

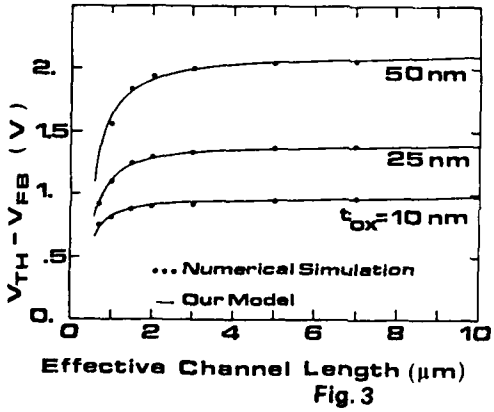


Fig. 3

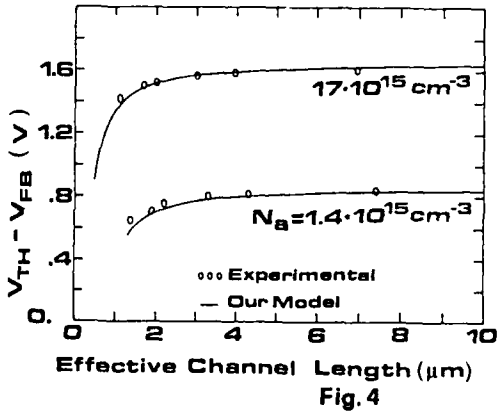


Fig. 4

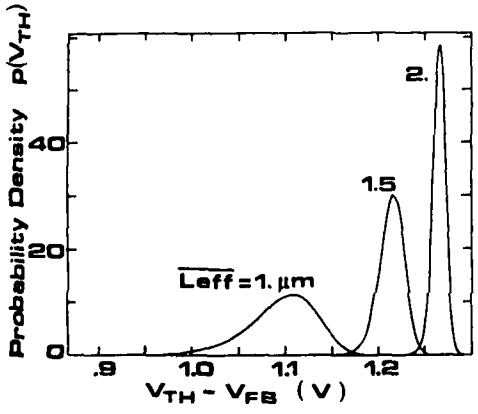


Fig. 5a

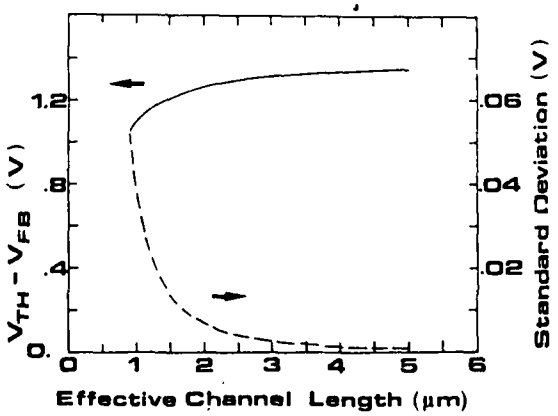


Fig. 5b

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

THE SERIES RESISTANCE OF SUBMICRON MOSFETs AND ITS EFFECT ON THEIR CHARACTERISTICS

F.M. KLAASSEN, P.T.J. BIERMANS and R.M.D. VELGHE

Philips Research Laboratories, PO Box 80000, NL-5600 JA Eindhoven, The Netherlands

Résumé - Une formule de modèle, liée au procès, pour les résistances série des MOSFETs est présentée, en supposant un profil de dopage linéaire pour des jonctions graduelles. La variable la plus importante du procès est le gradient de dopage latéral, tandis que la dose du drain légèrement dopé n'a pas beaucoup d'influence. Les valeurs calculées des resistances correspondent bien avec les valeurs mesurées, qui ont été extraites des paramètres d'un nouveau modèle de circuit pour des MOSFETs. Dans ce modèle le courant et la tension de saturation sont exprimés explicitement en termes des résistances série.

Abstract - A process-related design formula for the MOSFET series resistances is discussed, assuming a linear doping profile for graded junctions. The main process variable is the lateral doping gradient, whereas the LDD-dose has little effect. Calculated resistance values agree well with measured data, which have been extracted from the parameters of a new MOSFET circuit model. In the latter model the current and saturation voltage are expressed explicitly in terms of the series resistances.

1 – Introduction

In order to maintain a 5.5 Volt power supply in (sub)micron-size MOSFETs, graded source/drain junction profiles are widely used. Although the required reduction of hot carrier effects by this approach has been discussed extensively, little attention is paid to the associated effect of increased series resistance. In graded junctions the largest contribution to R_s arises from current crowding at the channel side (compare Fig. 1). However, when using implicit relations (derived for conventional junctions [1, 2]), to calculate this effect in LDD devices, large deviations from experimental data are found (compare Fig. 2). In this contribution a new process-related design formula for this resistance has been derived, which is based on the assumption of a linear lateral doping profile, in agreement with 2-D process simulation results. Furthermore its effect on the transistor characteristics and performance are discussed.

2 – Calculation of the LDD-MOSFET series resistance

According to the schematic representation of the current pattern of fig. 1, the source/drain series resistance of a MOSFET consists of four components: a contact resistance below the contact window, a sheet resistance, where the current flows along parallel lines, a spreading resistance due to current crowding at the vincinity of the channel end and an accumulation layer resistance owing to gate overlap. The contact resistance R_{co} has been calculated applying a transmission line model to the interface between the metal and the semiconductor [3]. Usually its value and that of the second component are much smaller than the spreading resistance. Using the assumption (based on numerical simulation) that the current converges within an angle of 1 radius from the full junction depth (y_j) to the thin accumulation sheet (thickness y_c), the latter resistance is given by

$$R_{sp} = 0.64 \int_{.64\, y_c}^{.64\, y_j} \frac{\rho \, d\, x'}{x'W} \quad , \tag{1}$$

where the constant of 0.64 comes from 1/tg1. The above integral has been evaluated in case of a uniform junction doping [1] and an exponential profile at the vincinity of the channel [2]. However even in the latter case, for LDD-type devices large deviations from experimental data are found (compare fig. 2). In fact, according to 2-D process simulation and experimental data [4], for the tail of graded junctions a far more linear doping profile applies.

Assuming an accumulation layer to be present between the undepleted junction end ($x = 0$) and ($x = -x_a$) (compare fig. 1), its resistance is simply given by

$$R_{ac} = \frac{x_a}{\mu_{ac} q N_{ac} W} \quad , \tag{2}$$

where qN_{ac} and μ_{ac} are the charge and mobility of this layer, respectively and W is the channel width. Defining a doping profile $N = N_0 - kx$, where k is the lateral doping gradient, according to (1) the spreading resistance is given by

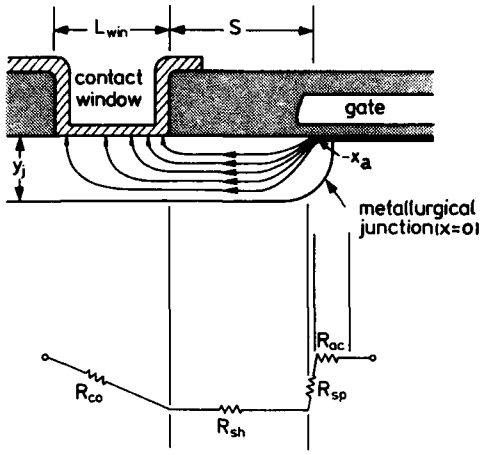


Fig. 1. Schematic diagram showing current pattern in the source region and its associated resistance components.

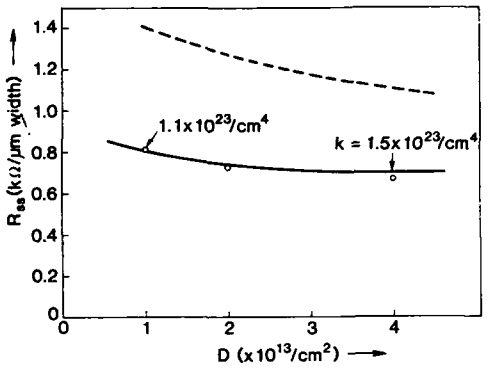


Fig. 2. Source series resistance vs LDD implantation dose. Dots are measured values, lines are calculated values (— = model, - - = ref. 2).

$$R_{sp} = \frac{0.64}{q\mu_s(N_0 + kx_a)} \ln\left[\frac{y_j(N_0 + kx_a)}{y_c(N_0 + kx_a + \cdot 64\ ky_j)}\right], \tag{3}$$

where μ_s is the average bulk mobility in the tail of the graded junction.
Considering that the total series resistance is determined by the minimum of the sum of R_{sp} and R_{ac}, x_a follows from the condition $d(R_{sp} + R_{ac})/dx_a = 0$. This procedure yields

$$x_a = k^{-1}\left[\{0.64(\mu_{ac}/\mu_s)\ N_{ac}kc\}^{1/2} - N_0\right], \tag{4}$$

where

$$c = \ln\left[\frac{y_j(N_0 + kx_a)}{y_c(N_0 + kx_a + 0.64\ ky_j)}\right].$$

Substituting the latter value in (3), we finally obtain

$$R_{sp} = \frac{0.80c^{1/2}}{qW(\mu_{ac}\mu_s)^{1/2}(N_{ac}k)^{1/2}}. \tag{5}$$

From this result we conclude that the main process variable, affecting the LDD series resistance, is the lateral doping gradient $k = d\,N_j/dx$, whereas the implant dose has only indirectly little effect. This is shown in fig. 2, where the calculated source series resistance has been plotted. The values of k have been obtained from process simulation. In addition, since μ_{ac} decreases with N_{ac} (V_{GS}), R_{sp} decreases slowly with gate bias [2]. For instance, when V_{GS} is decreased from 4 to 2 Volts, the resistance values of fig. 2 increase by 25%.

3 – Effect of series resistances on the transistor characteristics

Starting from the well-known model equation for a short-channel MOSFET [5], the effect of a source (R_{ss}) and drain resistance (R_{sd}) on the characteristics can be calculated by taking into account the Kirchhoff relations

$$\left.\begin{aligned} V'_{DS} &= V_{DS} - I_D(R_{ss} + R_{sd}) \\ V'_{GS} &= V_{GS} - I_D R_{ss} \end{aligned}\right\} \tag{6}$$

In the above relations V_{DS} and V_{GS} are the external drain-source and gate-source bias.
Using the above relations and assuming that for a practical device

$$\beta(R_{ss} + R_{sd}) < 0.50 ,$$

in which $\beta = \beta_o$ W/L is the MOSFET gain constant, the current equation can be rewritten into

$$I_D = \beta \frac{(V_{GS} - V_T) V_{DS} - \tfrac{1}{2}(1 + \delta) V_{DS}^2}{1 + \theta_A(V_{GS} - V_T) + \theta_B V_{SB} + \theta_c V_{DS}} . \tag{7}$$

In this equation V_T is the threshold voltage, δ is a factor representing the substrate backgate effect and θ_B is a mobility degradation parameter. However compared to the previous model [5] two parameters are present, which are generalisations of the mobility degradation (θ_0) and velocity saturation effect (θ_{co})

$$\theta_A = \theta_0 + \beta(R_{ss} + R_{sd}) , \tag{8a}$$

$$\theta_c = \theta_{co} - \tfrac{1}{2}\theta_0 - \beta R_{sd} . \tag{8b}$$

Furthermore from the internal saturation condition $d I_D/dV'_{DS} = 0$, the external saturation voltage can be calculated. In this way we obtain

$$V_{DSS} = \frac{2(V_{GS} - V_T)}{1 + \delta} \left[\left\{ 1 + \frac{2\theta_c(V_{GS} - V_T)}{(1 + \delta)[1 + \theta_A(V_{GS} - V_T) + \theta_B V_{SB}]} \right\}^{\frac{1}{2}} + 1 \right]^{-1} . \tag{9}$$

Generally, owing to the presence of R_{ss} and R_{sd}, under the same bias conditions the current is reduced, but the saturation voltage increases.

In principle the model parameters V_T, δ, β, θ_A, θ_B and θ_c can be determined from a comparison between the measured transistor characteristics and calculated model data. Usually V_T and δ are expressed in terms of the zero-bias threshold voltage V_{TO} and a threshold slope factor γ [5]. In addition for a short-channel transistor a drain feed-back effect $V_T = V_{TO} - fV_{DS}$ has to be taken into account. From measuring the characteristics of devices with different gate lengths, according to relations (8a) and (8b) R_{ss} and R_{sd} can be obtained from the slope of θ_A vs β and θ_c vs β plots. However since $\theta_{co} = (L E_c)^{-1}$, where E_c is the velocity saturation field, in the latter case the value of the parameter E_c is required.

4 – Comparison with experimental results

Generally the model equations (7) and (9) represent well the characteristics of MOSFETs with graded junctions. Fig. 3 shows a good fit to the measured characteristics of a device with a 0.5 μm gate length. Corresponding values of major parameters are given in the subscript. Next figs. 4 and 5 give the parameters θ_A and θ_c as a function of the gain constant β. The latter has been varied by taking devices with different gate length (compare the insets). In agreement with eqs. (8) a linear dependence is observed. From the slope value and the indicated value of E_c (obtained from non-LDD devices), a value R_{ss} = 680 Ω and R_{sd} = 560 Ω per μm gate width has been obtained. Unfortunately owing to additional effects, the accuracy of the above model is not sufficient to provide the gate bias dependence of the series resistances. Therefore the given data has to be considered as an average. However, if the resistance values are determined from the characteristics at low drain bias, a slightly larger value for R_{sd} is observed. This is due to the absence of an accumulation layer at large drain bias.

Following the above procedure, the resistance value has been measured as a function of the LDD-implantation dose. The results for the source resistance are given in fig. 2. Not only agree the measured data well with values calculated according to eqs. (3), (4) and (5), but the weak dependence on the dose D is confirmed too.

Finally fig. 6 gives the decrease in drive current of an optimized LDD-device compared to a conventional type. Owing to compensating effects (a decrease of θ_c and an increase of V_{DSS}) this decrease is less than the direct change in θ_A and θ_c.

References

1. G. Baccarani et al., IEEE Ltts **EDL-4**, 27 (1983).
2. K K Ng et al IEEE Trans ED-33 965 (1986).
3. D.B. Scott et al., IEEE Trans. **ED-29**, 651 (1982).
4. W. Rösner et al., VLSI Technology Symposium, 9 (1988).
5. F.M. Klaassen, chapter 12 in Advances in CAD for VLSI, Vol. 1, North-Holland, (1986).

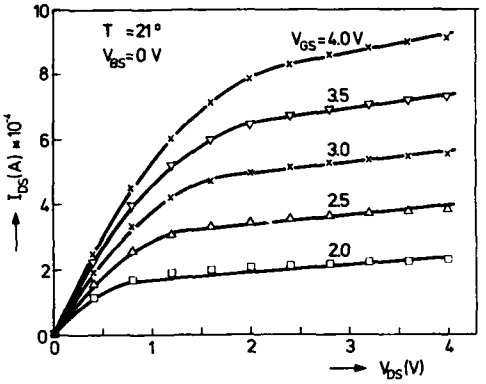


Fig. 3. Measured and calculated characteristics of a 0.5 μm MOSFET. V_{TO} = 0.70 V, γ = 0.10$V^{1/2}$, β = 1.48 mA/V^2, θ_A = 0.56 V^{-1}, θ_c = 0.22 V^{-1} f = 0.13.

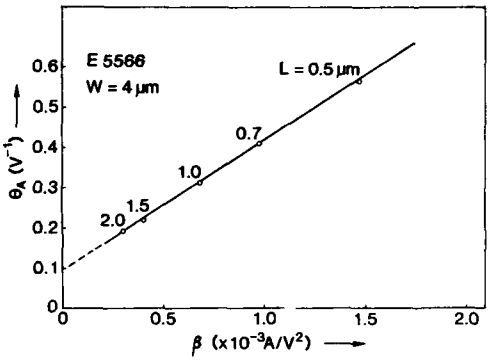


Fig. 4. Model parameter θ_A vs gain factor. The slope yields the value of (R_{ss} + R_{sd}).

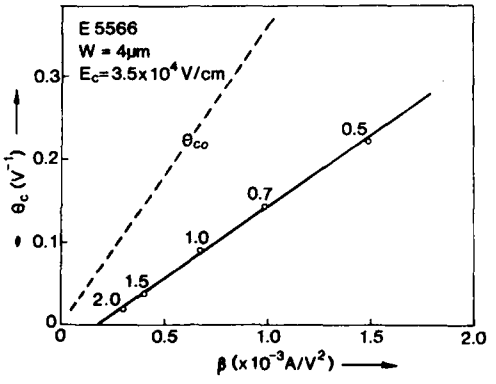


Fig. 5. Model parameter θ_c vs gain factor. From the slope and the velocity saturation field (E_c) the value of R_{sd} is obtained.

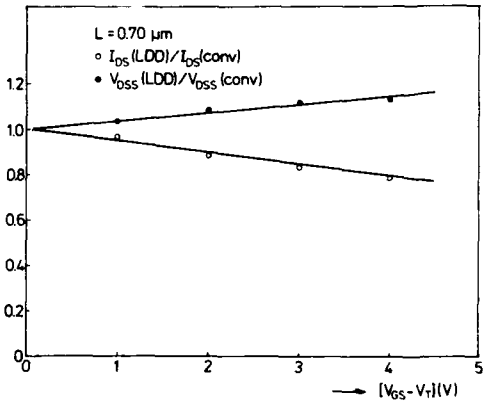


Fig. 6. Change in drain current and saturation voltage of a 0.7 μm - LDD MOSFET compared to a conventional device.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

ANALYTICAL ANALYSIS OF PUNCHTHROUGH IN BURRIED CHANNEL P-MOSFETs

T. SKOTNICKI, G. MERCKEL and T. PEDRON

CNET-CNS, BP 98, Chemin du Vieux Chêne, F-38243 Meylan Cedex, France

Résumé - Le phénomène de perçage dans les MOSFET déplétés à canal P (BC-P-MOSFET) est modélisé de manière analytique par l'utilisation de la transformation tension-dopage (VDT). On montre que le mécanisme du perçage dans les BC-P-MOSFETs est semblable à celui des MOSFET canaux N à conduction à surface, ceci à cause de l'abaissement de barrière provenant de la tension drain (DIBL). Cependant ce phénomène est renforcé, dans les canaux enterrés, par un phénomène d'écran électrostatique de la surface inversée. Ce nouveau modèle de courant de perçage, le premier du genre, traite correctement les effets d'écrantage et donne une grande précision sur une large plage de polarisation et de longeur de canal.

Abstract - The punchthrough phenomenon in burried-channel (BC) P-MOSFETs (depletion mode devices) is analysed analytically using the voltage-doping transformation (VDT) technique. The mechanism of punchthrough in a BC-P-MOSFET is shown to be, similarly as in a surface-channel (SC) N-MOSFET, due to the DIBL effect. However in a BC-MOSFET the DIBL effect is shown to be considerably enhanced by the effect of screening of the surface inversion layer of electrons. The new punchthrough current model, an unique up to date, deals correctly with the screening effect and shows high accuracy within a wide range of biases and channel lengths.

1 - INTRODUCTION

Punchthrough (pt) leakage currents give a significant rise to a stand-by power dissipation in MOS circuits. Thereby punchthrough becomes one of the severest limitations when shrinking MOSFETs to submicron dimensions. For this reason pt phenomenon has received a great deal of interest in the literature, but this uniquely in relation to surface-channel (SC) N-MOSFETs. To date there is no analytical model published, being able to predict punchthrough currents in burried-channel (BC) P-MOSFETs. Such a situation must become strange when taking into account that in CMOS circuits pt leakages of BC-P-MOSFETs contribute, at least, equally as those in SC-P-MOSFETs to a total stand-by power dissipation. The analysis presented in this paper is based on the voltage-doping transformation (VDT) which has already been successfully applied to the modeling of short-channel threshold voltage [1] and punchthrough [2] in SC-N-MOSFETs. Now an extension of the VDT to the punchthrough in BC-P-(or N-)MOSFETs is proposed.

2 - PHYSICS OF PUNCHTHROUGH IN BC-P-MOSFETs

According to the voltage-doping transformation [1] the 2-dim Poisson equation on the x-axis can be rewritten in the following 1-dim form

$$\left.\frac{\partial^2\Psi}{\partial x^2}\right|_{x\in x\text{-axis}} = \frac{q}{\varepsilon_s} N^*(x) \qquad (1)$$

where the x-axis is the loci of the potential minima along the channel and goes perpendicularly to the silicon surface. The effective doping $N^*(x)=N(x)-\varepsilon_s D(x)/q$ where $D(x)\triangleq\left.\partial^2\Psi/\partial y^2\right|_{x\in x\text{-axis}}$ is found analogically as in [2] from an approximation of the lateral potential distribution $\phi(y)=\Psi(x=\text{const},y)$ although now an exponential instead of a parabolic approximation is used. In a SC-N-MOSFET it is reasonable to approximate the lateral potential distribution by a family of parabolas because the potential across the p-n source-bulk and drain-bulk junctions varies quadratically with distance. However in a BC-P-MOSFET there is no n-p junction in the way from a source to a drain. Pfiester et.al. found in [3] a 2-dim analytical solution for a potential distribution in a BC-MOSFET. Their solution shows that the lateral potential distribution in a BC-MOSFET depends exponentially and not quadratically on a distance. This finding prompts the lateral potential distribution $\phi(y)$ to be approximated by a superposition of a constant and two exponents

$$\phi(y) = a + b\exp\left(-\frac{y + l_{vc}}{l_0}\right) + c\exp\left(\frac{y - L_{el} + l_{vc}}{l_0}\right) \qquad (2)$$

where $y=l_{vc}$ is the distance of the virtual cathode from the source boundary and l_0 will serve as a fitting parameter. The coefficients a, b and c are determined by the boundary conditions identical as in [2] and next $N^*(x)$ is found to be

$$\text{- for } 0\le x\le t_c \qquad N^*(x)\triangleq N^*_A = N_A - \frac{2\varepsilon_s}{q l^2_0} V^*_{DS}\exp\left(\frac{L_{el}}{2l_0}\right) \qquad (3a)$$

where

$$V^*_{DS} = -\frac{2}{d}\{[V_{DS} + (V_{bi}+V_{SB}-\eta\varphi_{cvc})(2-d) - 2\sqrt{(V_{bi}+V_{SB}-\eta\varphi_{cvc})(V_{bi}+V_{SB}+V_{DS}-\eta\varphi_{cvc})(1-d)}\,]$$

$$[(V_{bi}+V_{SB}-\eta\varphi_{cvc})(2-d) - 2\sqrt{(V_{bi}+V_{SB}-\eta\varphi_{cvc})(V_{bi}+V_{SB}+V_{DS}-\eta\varphi_{cvc})(1-d)}\,]\}^{1/2} \qquad (3b)$$

and $d=1-\exp(-L_{el}/l_0)$. t_c is the thickness of the implanted p-channel and η is the second fitting parameter.

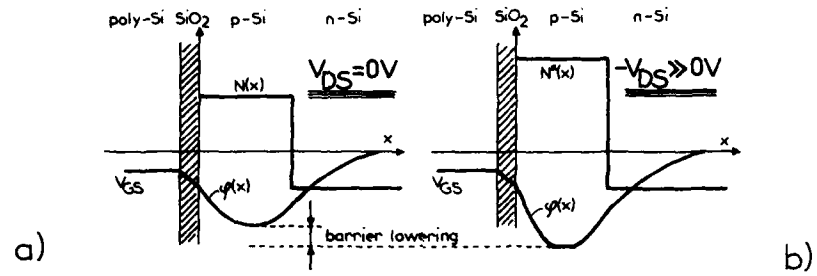


Fig. 1 - Physical mechanism of punchthrough in BC-P-MOSFET. Large negative drain bias increases the effective channel doping N^* pertaining to the gate field. As a result, potential within the channel becomes more negative, thus lowering the barrier height and thereby allowing more current to flow.

Additionally we assume that only the channel doping is modified by the drain field so that for $t_c \le x < \infty$ $N^*(x)=N_D$. Expression (3) describes explicitly the voltage-doping transformation. On this base the mechanism of punchthrough in a short- BC-P-MOSFET can be accounted for as follows: the drain bias increases the effective channel doping N^*_A pertaining to the gate field thereby reducing the penetration of gate and bulk electrical fields into the channel, see Fig. 1. As a result the potential minimum φ_{cvc} along the x-axis becomes more negative thus bringing the transistor towards the "on-state" as shown in Fig. 1b). In this way the drain current can become quite significant even for $V_{GS}=0V$ (when the channel should normally be "off"), which purely means punchthrough. It is to be noted that the above effect is suppressed when $L_{el} \to \infty$ because then $N^*_A \to N_A$ whatever the drain bias, compare eqn (3a).

3 - POTENTIAL BARRIER HEIGHT

The potential barrier encountered by holes on their way from a source to a drain equals $[\varphi_{vc}(x) - V_{bi} - V_{SB}]$, and as φ_{vc} varies with x thus also the barrier differs in height depending on depth x , where $\varphi_{vc}(x)=\Psi(x,y=t_{vc})$. The lowest barrier occurs at the depth t_{vc} (corresponding to the location of the virtual cathode) and equals $(\varphi_{cvc} - V_{bi} - V_{SB})$, where $\varphi_{cvc}=\Psi(t_{vc},t_{vc})$ is the potential of the virtual cathode. Thus the pt current flows within only a narrow strip at depth t_{vc} instead of a whole channel opening 0 to t_c. Consequently we will focus only on φ_{cvc} finding.

A - Entirely depleted channel

Equation (1) has a 1-dim form typical of the long-channel case so its solution (easily obtainable by double integrating) is well known from the long-channel MOSFET's theory. For this reason, as well as for the lack of space, it will not be quoted here. The solution for φ_{cvc} combined with a pt current expression (analogical to that used in [2]) allows BC-P-MOSFET punchthrough characteristics to be modeled. The resulting lines N in Fig. 2 show much less punchthrough than the measured ones denoted by M. A physical interpretation of this fact is that the surface inversion layer of electrons, neglected up to now, screens the gate electrical field lines thus giving completely the control over the channel up to the drain, and thereby enhancing punchthrough, compare [4].The simplest way of how to take into account the surface inversion layer screening (SILS) effect is to assume that the barrier heigth becomes independent on the gate bias once a strong inversion is induced, as prompted by the depletion theory. Lines P in Fig. 2 show, however, that such an approach (perfect SILS effect) must lead to a significant overestimation of punchthrough at large positive V_{GS} biases. From the above it becomes clear that the SILS effect is essential to punchthrough in BC-P-MOSFETs and thereby necessitates a more careful treatment.

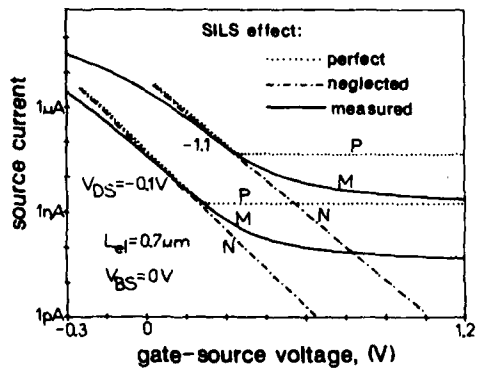


Fig. 2 - Comparison of the model of ideal SILS effect - lines P, and the neglect of the SILS effect - lines N with measurements - lines M.

B - Inversion at surface

In order to find the principle of the SILS effect the Poisson equation (1) must be completed with the term $N_D \exp(\Psi/U_T)$ corresponding to the concentration of electrons in the inversion layer

$$\frac{\partial^2 \Psi}{\partial x^2} = \frac{q}{\varepsilon_s}\left[N^*(x) + N_D \exp\frac{\Psi}{U_T}\right] \qquad (4)$$

A strict derivation of φ_{cvc} would necessitate in double integration of (4) which is analytically impossible. However it is to be noted that φ_{cvc} can be accurately described by a solution to (1) thus neglecting the inversion layer, if only the the actual depletion depth t_d, is derived from (4) which deals correctly with the inversion layer. Indeed, knowing the actual t_d value, which accounts correctly for the effect of the surface inversion layer of electrons we may expect good results for φ_{cvc} from the solution to (1) because whole the region $t_{vc} \le x \le (t_c + t_d)$ is free of electrons and contains only the space-charge of ionized impurities which is just correctly incorporated in (1). It will be shown in the next Section that such an approach ensures very good results even for gate biases much greater than that needed to produce the strong inversion at the surface. On the above basis φ_{cvc} is formally found by double integrating (1) to be

$$\varphi_{cvc} = -\frac{qN_D}{2\varepsilon_s} t^2_d \left(1 + \frac{N_D}{N^*_A}\right) \qquad (5)$$

and t_d by single integrating (4) to be

$$t_d = -t_c + \sqrt{t^2_c\left(1 + \frac{N^*_A}{N_D}\right) - \frac{2\varepsilon_s}{qN_D}\varphi_s} \qquad (6a)$$

where

$$\varphi_s = U_T \ln\left[\left(\frac{L_D \varepsilon_{ox} V_{gef}}{t_{ox} \varepsilon_s U_T}\right)^2 - \frac{q(N^*_A + N_D)}{2\varepsilon_s U_T} t^2_{dcrit} + 1\right] \qquad (6b)$$

t_{dcrit} corresponds to the onset of the strong inversion at the surface (φ_s=0) and is formally given by (6a) after substituting φ_s by 0. L_D is the Debye's length and reads $(\varepsilon_s U_T/2/q/N_D)^{1/2}$ and $V_{gef}=V_{GS}-V_{SB}-V_{FB}$.

4 - RESULTS

The presented model of nonideal SILS effect has fairly improved the accuracy. The mean error of the theory is made as small as 3±3% (currents in logarithm are compared). The comparison between measured and modeled BC-P-MOSFET subthreshold characteristics, as illustrated in Fig. 3, demonstrates an excellent agreement for long as well as short channels. The verified range of validity of the model includes gate voltages V_{GS} ranging from threshold voltage, in this technology $V_{th}\cong$-0.5V, up to large positive voltages much higher than those needed to induce a strong inversion at the surface. In the technology used for the verification the surface inversion is induced at $V_{GS}\cong$0.3V whereas the modeled characteristics shown in Fig. 3 extend up to 1.2V still being in very good agreement with measurements. As far as a drain bias is concerned the range of 0V down to -10V is treated very correctly by the model, however under a condition that the corresponding pt current does not exceed the level of about 1μA. Otherwise the current becomes space-charge limited which is not taken into account in the analysis. As regards the bulk bias, the range 0 to 5V poses no difficulty to the model as can be observed in Fig. 4. It is worth noting that the positive bulk bias suppresses punchthrough in BC-P-MOSFETs analogously as a negative bulk bias suppresses punchthrough in SC-N-MOSFETs.

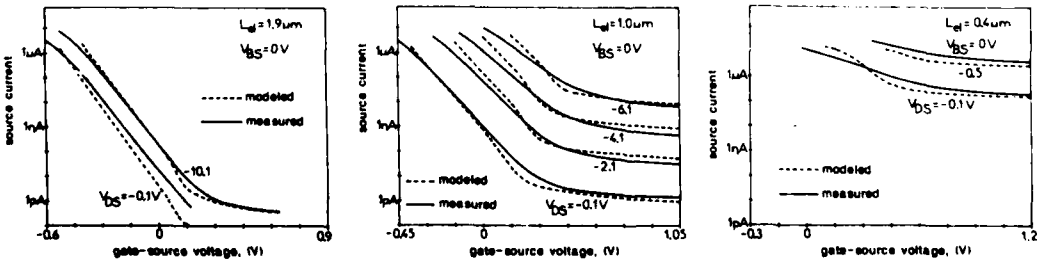


Fig. 3 - Measured (solid line) and analyticaly modeled (dashed line) subthreshold characteristics for a series of BC-P-MOSFETs with different electrical channel lengths and parametrically varied drain bias. Bulk bias is kept at 0V. The technological parameters read: N_A=1.44×10^{16}cm^{-3}, N_D=3.3×10^{15}cm^{-3}, t_c=0.21μm, t_{ox}=25ηm, μ_p=470cm^2/V/s and V_{FB}=-0.219V. The used values of the fitting parameters are: η=0.979 and l_0=0.304μm.

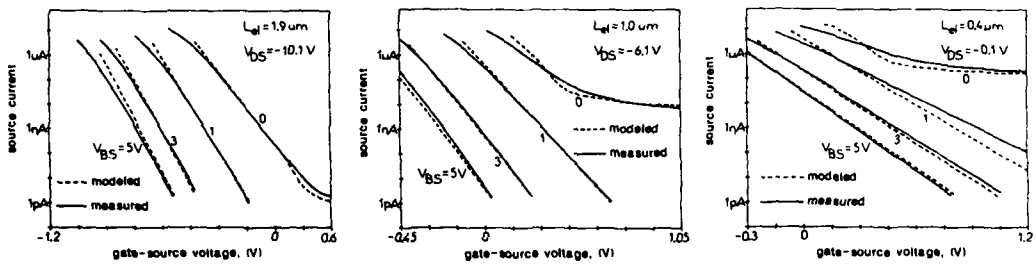


Fig. 4 - Measured (solid line) and analyticaly modeled (dashed line) subthreshold characteristics for a series of BC-P-MOSFETs with *different electrical channel lengths with parametrically varied bulk bias. Drain bias is kept constant.* The technological parameters read: N_A=1.44×10^{16}cm^{-3}, N_D=3.3×10^{15}cm^{-3}, t_c=0.21µm, t_{ox}=25ηm, μ_p=470cm^2/V/s and V_{FB}=-0.219V. The used values of the fitting parameters are: η=0.979 and l_0=0.304µm.

5 - CONCLUSION

In conclusion, our analysis for the first time enables the BC-P-MOSFET characteristics to be modeled analytically. The model exhibits an excellent accuracy and a wide range of validity. The mean error of the theory is of 3% (*currents in logarithm are compared*) whereas the *number of code lines needed to implement it on the HP9000 calculator is as small as 30 lines of BASIC code.* Thanks to the analytical analysis the *punchthrough phenomenon in BC-P-MOSFETs has for the first time be entirely accounted for. The impact of the surface inversion layer screening (SILS) effect on punchthrough is deduced and investigated. As a result the analysis is believed to find* applications in design, process optimization and simulation of advanced BC-P-MOSFETs.

REFERENCES

[1] Skotnicki T., Merckel G. and Pedron T., IEEE Electron Device Lett., EDL-9, (1988), 109.
[2] Skotnicki T., Merckel G. and Pedron T., IEEE Trans. Electron Dev., to be published in July 1988.
[3] *Pfiester J.R., Shot J.D. and Meindl J.D., IEEE Trans.* Electron Dev., ED-32, (1985), 333.
[4] Merckel G., Process and Device Modeling for Integrated Circuit Design, edited by F. Van De Wiele *et.al.*, Noordhoff, Leyden, (1977), 667.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A MOBILITY MODEL FOR MOSFET DEVICE SIMULATION

A.J. WALKER and P.H. WOERLEE

Philips Research Laboratories, NL-5600 JA Eindhoven, The Netherlands

Abstract The device characteristics of sub-micron MOS transistors depend strongly on the mobility of the charge carriers in the inversion layer. A new low lateral field mobility model for the normal field dependence will be presented. Good agreement was found between the predicted and measured mobilities for a large variety of samples. The model was successfully incorporated into a two - dimensional device simulation program which gave predictions agreeing well with experimental data.

1. INTRODUCTION

The importance of sub - micron silicon MOS transistors in future generations of memory and logic circuitry has led to intense interest in the accurate modelling of these devices with particular attention being paid to the mobility of the charge carriers in the inversion layer. Scaling the transistors into the sub - micron regime demands the use of thinner gate oxides and higher substrate doping densities. The influence that these changes have on the mobility must be built into any model if accurate predictions are to be made about device performance.

This paper presents a new physically - based model for the normal effective field dependence of the electron and hole mobilities in silicon inversion layers at low lateral fields. Recent results of an investigation into the application of this model to measurements carried out on MOS devices with gate oxide thicknesses and surface doping densities scaled for process generations down to 0.25μm are given. Furthermore, for the first time such a model was incorporated in a two - dimensional device simulation program and comparisons between predictions from this program and measurements on sub - micron transistors are presented.

2. MODEL

It has been shown elsewhere that the electron [1] and hole [2] mobilities follow universal curves when plotted as functions of an effective normal field, E_{eff}, given by

$$E_{eff} = \frac{1}{\epsilon_{si}}(Q_{dep} + \eta Q_{inv}) \qquad (1)$$

with $\eta = 1/2$ for electrons and $1/3$ for holes and where Q_{dep} and Q_{inv} are the depletion and inversion layer charges per unit surface area respectively.

In this new model a semi - empirical approach was taken to model the mobility - E_{eff} curves. It was assumed that the room temperature inversion layer mobility is dominated by three scattering mechanisms [3], namely Coulomb, carrier-phonon and surface roughness, each with its own contribution to the net mobility. These contributions are designated by μ_c, μ_{cp} and μ_{sr} respectively. Actual modelling of scattering processes in inversion layers is very complex due to the quantum mechanical nature of these processes and the fact that at temperatures above absolute zero more than one sub - band is filled. Therefore, a simplified semi - empirical approach was adopted.

To obtain a simple first order expression for carrier mobility due to screened Coulomb scattering, the Brooks - Herring formula was adopted [4]:

$$\mu_c = \frac{CT^{1.5}}{ln(1+b) - b/(1+b)} \cdot \frac{1}{N_I} \tag{2}$$

where $b = (24m^*\epsilon_{Si}k^2T^2)/(\hbar^2q^2N_o)$, N_I is the charged impurity density, N_o is the mobile carrier density, C is a constant, ϵ_{Si} is the dieiectric constant of silicon and the other symbols have their usual meanings.

The following modifications were made for the inversion layer :

$$N_o = n_{inv}/ < z > \qquad\qquad N_I \longrightarrow N_I + N_{oz}/ < z >$$

where n_{inv} is the number of mobile carriers per unit surface area, $< z >$ is the average distance of mobile carriers from the Si/SiO_2 interface and N_{oz} is the fixed oxide charge per unit surface area.

As a result of numerical calculations [5], $< z >$ can be given as a function of the mobile inversion charge number density per unit surface area, n_{inv}, and therefore as a funtion of E_{eff}.

As was given elsewhere [6], the mobility μ_{cp} due to carrier - phonon scattering and the mobility μ_{sr} due to surface roughness scattering can be approximately given by

$$\mu_{cp} = k_{cp}T^{-1}E_{eff}^{-1/3} \tag{3}$$

$$\mu_{sr} = k_{sr}E_{eff}^{-2} \tag{4}$$

where k_{cp} and k_{sr} are constants.

The net mobility μ was calculated using Matthiesen's rule. i.e.

$$\frac{1}{\mu} = \frac{1}{\mu_c} + \frac{1}{\mu_{cp}} + \frac{1}{\mu_{sr}} \tag{5}$$

with μ_c, μ_{cp} and μ_{sr} given in equations 2, 3 and 4 respectively with the constants C. k_{cp} and k_{sr} as fitting parameters.

The above equations were incorporated into a two - dimensional device simulation program, CURRY [7], in which the effective field at any point from source to drain was calculated by averaging the normal field in the depth direction over the inversion layer or surface accumulation layer at that point. The model is therefore a non - local one with the mobility at any point dependent on the values of the normal electric field at other points.

3. EXPERIMENTAL

To measure the mobility, a set of devices was fabricated which included both n- and p-type transistors and associated capacitors. The devices were all made on (100)-oriented p-type Si wafers with a resistivity of 20 Ω.cm. An n-well was formed using high energy phosphorus and arsenic implants while the p-substrate was implanted with high energy boron for the n-type devices. Gate oxides of thickness between 10nm and 23nm were thermally grown at 900^oC in an oxygen/nitrogen mixture (10% oxygen by volume). Polysilicon gate electrodes were used and doped with phosphorus. Capacitor gate electrodes had areas of up to $3.3 \times 10^{-4}cm^2$ with transistor gate areas up to $4 \times 10^{-4}cm^2$. Gate oxide fixed charge and interface state levels of below $1 \times 10^{11}cm^{-2}$ were measured in the devices.

The effective inversion layer mobility, μ, was extracted from conductance measurements in the linear region on 200μm x 200μm MOS transistors. The mobile inversion charge per unit surface area, Q_{inv}, was measured on the capacitor structures [8]. The sub - micron transistors were fabricated using electron beam lithography. The n-MOS device had a gate length of 0.53μm and width of 3.9μm. Its gate oxide was 12.5nm thick. The p-MOS device had a gate length and width of 0.60μm and 4.2μm respectively with a 10nm thick gate oxide.

4. RESULTS AND DISCUSSION

The results from mobility measurements on the large n- and p-MOS devices are given in Fig.1 and 2 respectively in which mobility is plotted as a function of normal effective field. The surface doping density, N_s, and gate oxide thickness, t_{ox}, are also shown in the figures. The points are the experimental data with the

solid lines being the fits of the model to these data. Fig.1 shows that the electron mobility in the inversion channel is almost independent of surface doping density at high effective fields but depends strongly on the doping at lower fields. In the case of the pMOS devices, different boron implants had been introduced into the originally n-type channel but it was not clear as to what the actual density of charged impurities would be since the impurities are not completely ionized.

It was apparent in the model for electrons that the same values for C (equ.2), k_{cp} (equ.3) and k_{sr} (equ.4) could be used for all samples provided that the surface doping density, N_I, and the oxide fixed charge, N_{oz}, were known. The values used in Fig.1 were $C = 2.63 \times 10^{17}$, $k_{cp} = 1.35 \times 10^{7}$ and $k_{sr} = 7.0 \times 10^{14}$ all expressed in cgs units.

In the model for holes, the value of N_I was unknown. Therefore a fit was carried out to find the optimum value of $C/(N_I + N_{oz}/ < z >)$ using fixed values of k_{cp} and k_{sr}. This was 0.225 for $k_{cp} = 3 \times 10^{6}$ and $k_{sr} = 2.22 \times 10^{14}$ all in cgs units. This is the solid line in Fig.2.

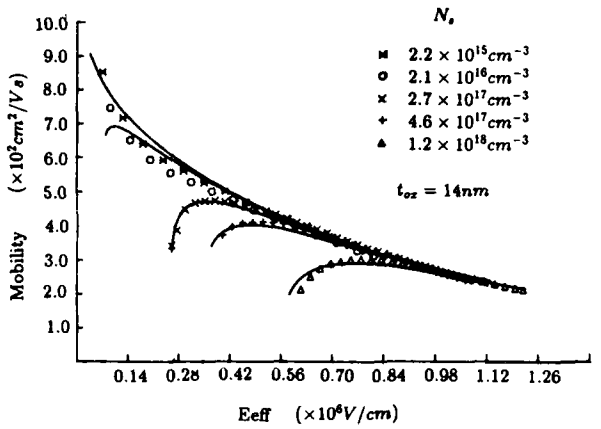


Fig.1 . The low lateral field electron mobility rsus the normal effective field for samples with gate oxide thickness and surface doping density as given in the inser The points are the experimental values while the solid lines are the fits of equation 5 with μ_c, μ_{cp} and μ_{sr} as given in equations 2, 3 and 4.

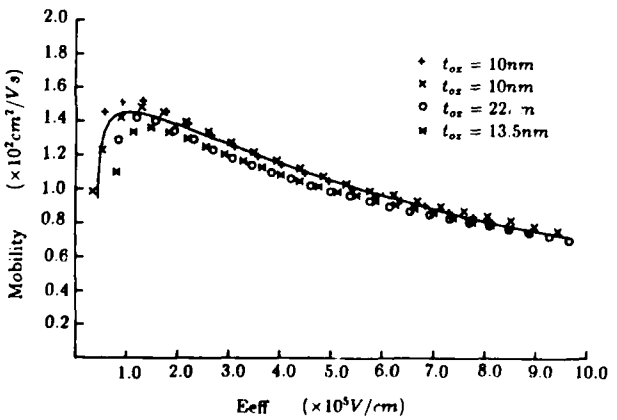


Fig.2 . The low lateral field hole mobility versus the normal effective field for samples with gate oxide thickness as given in the insert. The points are the experimental values while the solid line is the fit of equation 5 with μ_c, μ_{cp} and μ_{sr} as given in equations 2, 3 and 4.

In Fig.3, simulations of a half - micron n-MOS transistor together with experimental data are shown. Good agreement between the calculated and experimental data is observed. Similar data for a buried channel p-MOS transistor are shown in Fig.4. The $I_{ds} - V_{gs}$ characteristics at low V_{ds} are modelled well but systematic deviations in the $I_{ds} - V_{ds}$ characteristics at high drain bias suggest that the lateral - field model for holes is not accurate at high lateral fields.

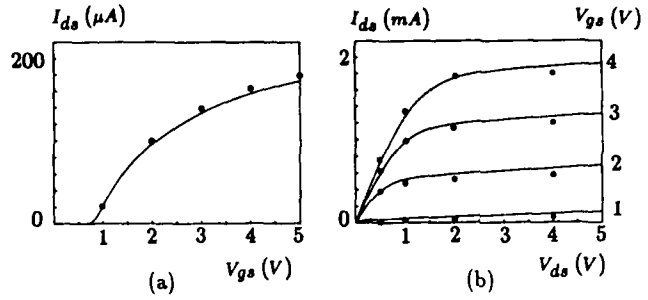


Fig.3. (a) I_{ds} - V_{gs} (V_{ds} = 0.1V) and (b) I_{ds} - V_{ds} characteristics of an n-MOS lightly-doped drain transistor with a gate length of 0.53μm and width of 3.9μm and gate oxide thickness of 12.5nm. The dots are calculated data points using the device simulation program.

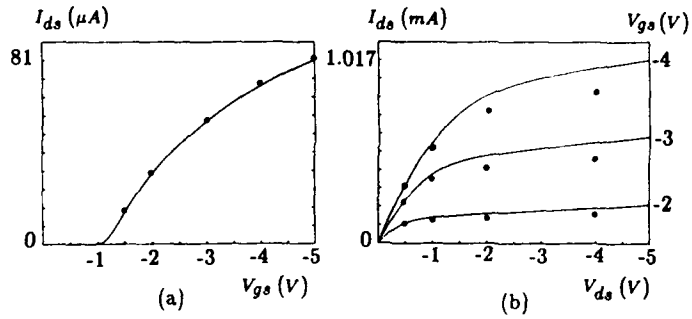


Fig.4. (a) I_{ds} - V_{gs} (V_{ds} = - 0.1V) and (b) I_{ds} - V_{ds} characteristics of a p-MOS buried channel transistor with a gate length of 0.60μm and width of 4.2μm and gate oxide thickness of 10nm. The dots are calculated data points using the device simulation program.

5. CONCLUSIONS

It appears that a simple mobility model with three parameters can describe a large amount of experimental data. The model was successfully incorporated into a two - dimensional device simulation program which subsequently predicted 0.5μm MOS transistor behaviour in good agreement with experiment

References

/1/ A.G. Sabnis and J.T. Clemens, IEDM Tech. Dig., pp. 18-21, Dec.1979.
/2/ J.T. Watt and J.D. Plummer, Symp. on VLSI tech., pp. 81-82, 1987.
/3/ Y.C. Cheng and E.A. Sullivan, J. Appl.Phys., Vol.44, No.8, pp.3619 - 3625, Aug. 1973.
/4/ T. Sugano, K. Hoh, H. Sakaki, T. Iizuka, K. Hirai, K. Kuroiwa and K. Kakemoto, in J. of Faculty of Engineering, Univ. of Tokyo (B) XXXII, 155, Appendix F, pp. 362-363 1973.
/5/ J.A. Pals, Ph.D Thesis, Technical Univ. of Eindhoven, Dec. 1972.
/6/ A.J. Walker and P.H. Woerlee, ESSDERC Tech. Dig. pp. 667 - 670 Sept. 1987
/7/ S.J.Polak, C. den Heijer, W.H.A. Schilders and P. Markowich, in Int. J. for Num. Meth. in Eng., Vol.24, pp. 763-838, April 1987.
/8/ C.G. Sodini, T.W. Ekstedt and J.L. Moll, Solid-State Electronics, Vol.25. pp. 833-841. Sept.1982.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

"WCAP" : WORST CASE ANALYSIS PROGRAM : A TOOL FOR STATISTICAL CIRCUIT SIMULATION

N. BALLAY and B. BAYLAC

SGS-Thomson Microelectronics, Central R and D/CIS/TDMC, BP 217, F-38019 Grenoble Cedex, France

Résumé - Le but du programme 'WCAP' est d'extraire automatiquement les dispersions des paramètres du premier ordre d'un modèle CAO, directement depuis plusieurs histogrammes d'indicateurs classiques (tension de seuil, courants statiques, ...) usuellement mesurés sur des testeurs industriels.

Abstract - The goal of 'WCAP' Program is to extract automatically first order CAD model parameter spreads directly from various histograms of classical indicators (threshold voltage, static currents, ...) usually measured on industrial parametric testers.

INTRODUCTION

Usually, these spreads of first order parameters were either directly measured on sophisticated R & D parametric testers for new processes, or "estimated" by process/design/model people for industrial processes. In both cases, the estimated spreads of first order model parameters let to overestimated and unrealistic spreads of electric quantities (current, access time, ...) mainly because 3-sigma variations on each first order parameter induce 5 to 8 sigma variations on current or access time which are complex functions of ALL the first order model parameters.

In 'WCAP' program, the indicators used to control the process are considered as linear functions of the first order model parameters whose sensitivity coefficients are calculated using multivariable regression procedure. Chapter 1 will show how at this point, their theoritical histograms may be easily calculated analytically. The fit with experimental distributions thus leading to REALISTIC minimum and maximum values for the first order model parameters is described in chapter 2. A real application is then developed.

1 - MATHEMATICAL MODEL OF PROBABILITIES

The fundamental basis of the program is the exact calculation of probabilities for linear functions of variables. This calculation is rather easy in the case of 2 or 3 variables when each of them varies uniformaly over a given range and figure 1 summarizes one particular case with 3 variables. Larger number of variables should induce triple, quadruple, ... integrals and complex distribution (Gaussian instead of Uniform) lead to non-analytical integrals. That is why we assumed the 2 following keypoints :

i) each indicator can be considered as a linear function of the first order model parameters.
ii) the distribution of each first order model parameters is uniform.

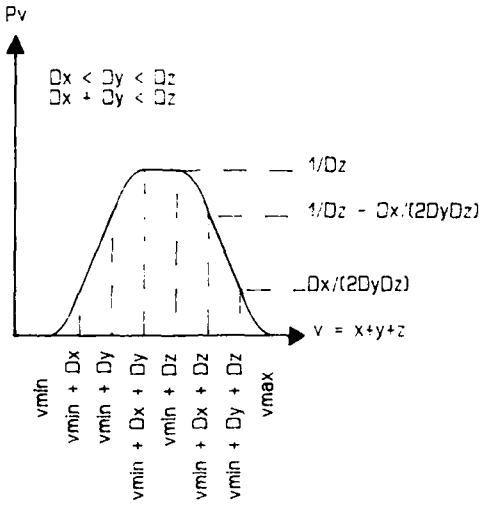


Fig. 1 - Theoritical histogram of a linear function of 3 variables uniformaly distributed

2 - ADJUSTMENTS OF MODEL PARAMETER SPREAD TO FIT HISTOGRAMS OF INDICATORS

The previous assumptions are valid when the indicators are properly chosen. In the case of MOS devices, the following indicators should be selected :

i) threshold voltage on a long and wide MOST (may be different substrate biases)
ii) static current on a long and wide MOST (may be different biases)
iii) static current on a short and wide MOST (may be different biases)
iv) static current on a long and narrow MOST (may be different biases)

A multivariable regression applied on each indicator versus the different model parameters (TOX, NB, UO, DL, DW) can be easily performed and the agreement is fair. Note that the other second order model parameters have been measured on typical wafers and are kept constant. This remains correct as soon as there is no correlation among parameters. That is why we used a specific home-made model for MOS devices whose parameters are independent /1/, /2/ and translation to other simple models are then performed.

The simple linear formula used to represent the indicators allow an easy and fast calculation of their distribution, assuming an uniform distribution of every model parameter. At this step, we use Marquartd's algorithm /3/ to fit actual distributions by modifying the window of first order model parameters.

3 - APPLICATION TO CMOS 1.2 U PROCESS

We validate this program on a 1.2 u CMOS process and we now present the case of the N-channel MOSFET family. The chosen indicators were in that case :

i) threshold voltage at VBS=0 for W/L=25/25
ii) static current for W/L=25/25 and VDS=0.1, VGS=5, VBS=0
iii) static current for W/L=25/25 and VDS=5.0, VGS=5, VBS=0
iv) statis current for W/L=25/1.2 and VDS=0.1, VGS=5, VBS=0
v) static current for W/L=25/1.2 and VDS=5.0, VGS=5, VBS=0
vi) static current for W/L=1.2/25 and VDS=0.1, VGS=5, VBS=0
vii) static current for W/L=1.2/25 and VDS=5.0, VGS=5, VBS=0

The model parameters whose spread were fitted are :

i) oxide thickness
ii) channel doping level
iii) mobility
iv) difference between drawn and electrical poly length
v) difference between drawn and electrical poly width

Figures 2 to 4 are output examples pointing out the accuracy of this analysis. The different assumptions are so realistic and, moreover, there is a strong relation between the final spread of each model parameter and its real one that could have been measured (see figure 5).

Dashed histograms : experiment
Dotted histograms : theory

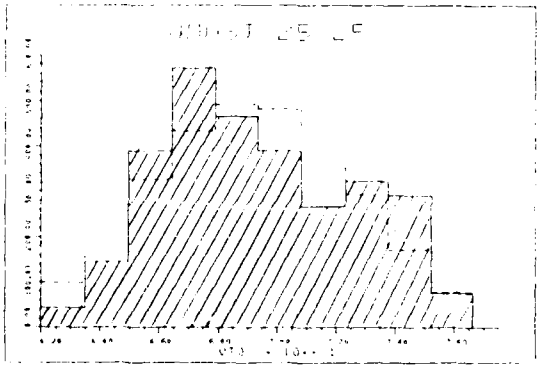


Fig. 2 - Distribution of threshold voltage for a long and wide MOST

Dashed histograms : experiment
Dotted histograms : theory

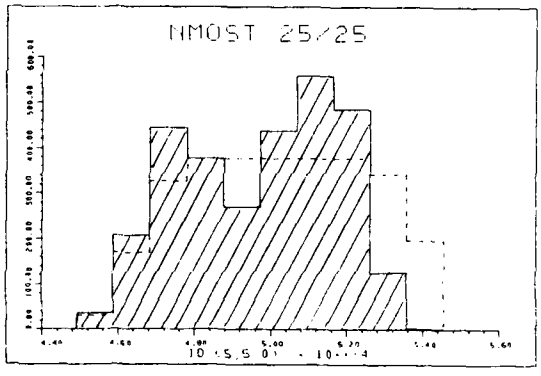


Fig. 3 - Distribution of saturation current for a long and wide MOST

Dashed histograms : experiment
Dotted histograms : theory

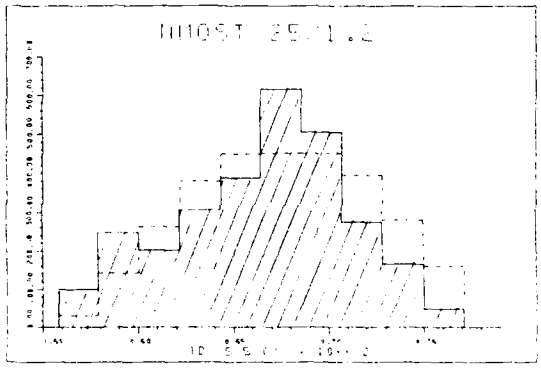


Fig. 4 - Disiribution of saturation for a short and wide MOST

Stars : experiment
Cont : theory

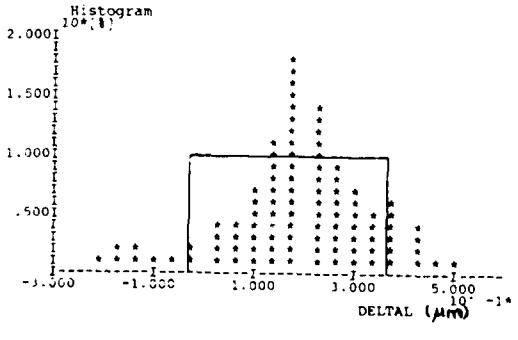


Fig. 5 - Actual and extracted by 'WCAP' histograms for DELTAL model parameter

4 - CONCLUSION

This program is an efficient tool to evaluate accurate spreads of model parameters needed for circuit simulation directly from histograms built with data from production line.

Once the spread of each model parameter is known, actual worst case parameter sets can be built. The same kind of analysis might be applied on results of SPICE simulations to know the theoritical distribution of desired variables (access time, gain, bandwidth, ...). Once this is done, the design yield of a circuit can be evaluated. To do so, some software must be written to generate SPICE input files and to extract automatically its results of simulation. This is the key to achieve proper simulations of analog circuits when worst cases cannot be easily defined.

The methods used in 'WCAP' program can be applied to every model, whose input parameters are independent, for MOS technologies as well as BIPOLAR ones.

REFERENCES

/1/ N. Ballay, B. Baylac, N. Rauber, "Threshold Voltage Model for Complex MOS Structures of VLSI Technologies", ESSDERC Tech. Dig., 1984

/2/ N. Ballay, B. Baylac, N. Rauber, "3-D Analytical Model of N and P MOSFET's for Enhancement and Depletion Mode Devices in the Submicron Range", ESSDERC Tech. Dig., 1984

/3/ D.W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters", J. Soc. Indust. Appli. Math., Vol. 11, no.2, Jun 1963

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

BAND-TAIL SHOCKLEY-READ-HALL RECOMBINATION IN HEAVILY DOPED SILICON

M.Y. GHANNAM*, R.P. MERTENS, S.C. JAIN, J.F. NIJS and
R. VAN OVERSTRAETEN

*Electronics and Communications Department, Cairo University, Guiza, Egypt
Interuniversity Microelectronics Center, Kapeldreef 75, B-3030 Heverlee, Belgium

Résumé - La densité des centres de recombinaison type Shockley-Read-Hall due aux états localisés situés dans le prolongement de la bande des porteurs minoritaires a été calculée pour le silicium fortement dopé. Nous prouvons que l'effet de ces états devient dominant pour des concentrations de dopage supérieur a 10^{17} cm^{-3}.

Abstract - Shockley-Read-Hall recombination center density due to the localized states in the minority carrier band tail has been calculated. It is shown that in heavily doped silicon, the effect of these band tail states is comparable to or more important than that due to deep states and modifies the lifetime of minority carriers significantly.

1 - INTRODUCTION

In the bulk of a semiconductor, Shockley-Read-Hall (SRH) recombination assumes the presence of all recombination centers at a single energy level in the energy band-gap. The recombination rate is given by

$$U_o = \frac{\sigma_n \sigma_p v_{th} N_t .(pn - n_i^2)}{\sigma_n[n + n_i exp(E_t - E_i)/kT] + \sigma_p[p + n_i exp(E_i - E_t)/kT]} \tag{1}$$

where σ_n and σ_p represent the electron and hole capture cross sections, N_t is the recombination center density per unit volume, v_{th} is the carrier average thermal velocity, n and p are the electron and hole concentration respectively, n_i is the intrinsic carrier concentration, E_i is the intrinsic energy level and E_t is the recombination center energy level, k is Boltzman constant and finally T is the absolute temperature.

When a continuum of localized states is inserted in the energy band-gap a more generalized formula has to be used. In this case the density of the recombination centers N_t has to be modified to consider the continuous nature of the localized states. A per unit energy per unit volume density of states $D(E)$ now replaces N_t in equation (1) and the recombination rate equals the integral of the per energy rate over all the continuum levels. The resulting expression has the form

$$U = \int_{Eg} \frac{\sigma_n \sigma_p v_{th} D(E).(pn - n_i^2) dE}{\sigma_n[n + n_i exp(E - E_i)/kT] + \sigma_p[p + n_i exp(E_i - E)/kT]} = \int_{Eg} u(E) dE \tag{2}$$

This procedure has been followed for instance to describe the surface recombination rate U_s. In this case, however, $D(E)$ has a per unit energy per unit area dimension. In a heavily doped semiconductor, potential energy fluctuations result in energy tails that appear at the conduction and valence band edges and extend deeper in the band-gap. The localized states in these band tails act as recombination centers. Though they are situated far away from midgap, the density of these states might be so high such that they contribute significantly to U. The total recombination rate in the bulk of a heavily doped semiconductor should therefore be expressed by equation (2). The integration should be carried out starting from the lowest localized state level in the tail of the valence band up to the highest localized state level in the tail of the conduction band. This integral can therefore be splitted into three components. The first component represents the recombination rate caused by the centers situated at the discrete level near midgap which is responsible for the recombination process taking place in unperturbed-band material. The second and third components represent the recombination rates caused by the centers in the tail of the valence band and in the tail of the conduction band. In heavily doped semiconductors, equation (2) can therefore be written as

$$U = U_o + \int_{VBtail} u(E).dE + \int_{CBtail} u(E).dE \tag{3}$$

In terms of lifetime, equation (3) can be rewritten

$$1/\tau_{eq} = 1/\tau_o + 1/\tau_{vbt} + 1/\tau_{cbt} \tag{4}$$

where τ_{eq} is the equivalent lifetime, τ_o represents the lifetime due to recombination at the discrete center near midgap, and τ_{vbt} and τ_{cbt} represent the lifetime due to recombination in the tail of the valence band and in the tail of the conduction band, respectively. Note that band tails are fundamental phenomena in heavily doped semiconductors and therefore the recombination taking place through the band tail centers is associated with a lifetime that represents a fundamental limit. We shall neglect here the recombination rate due to recombination via the localized states in the majority carrier band tail, if such states would exist. In such a situation equation (4) reduces to

$$1/\tau_{eq} = 1/\tau_o + 1/\tau_{bt} \tag{5}$$

where τ_{bt} represents the lifetime due to recombination via the centers located in the minority carrier band tail and is given by

$$1/\tau_{bt} = \sigma_{mino} v_{th} N_{tbt} \tag{6}$$

where N_{tbt} represents the effective density of recombination centers in the band tail. By combining equations (2), (3), (5) and (6) N_{tbt} for n-type material can be expressed as

$$N_{tbt} = \int_{Emin}^{Emax} \frac{D(E).dE}{1 + (n_i/n).(\sigma_p/\sigma_n).exp(E_i - E)/kT} \tag{7}$$

In order to calculate N_{tbt}, $D(E)$ should be known. In section 2 we introduce some functions that can be used to represent $D(E)$. Also, the limits of integration must be defined. Section 3 deals with the model followed to determine the localization edge in the minority carrier band tail. In section 4 the doping dependence of N_{tbt} (τ_{bt}) is discussed and in section 5 some numerical examples are presented and fitting of experimental data is carried out. Finally a short discussion followed by a conclusion is presented in section 6.

2 - DENSITY OF STATES IN THE BAND TAILS

In heavily doped semiconductors, energy band tails are created due to potential energy fluctuations as it is schematically represented in Fig.1. The parabolic representation of the energy band states does not hold anymore. A relatively simple model for the new density of state function was given by Kane [1]. In his model, the potential energy fluctuation is assumed gaussian with a standard deviation δ. The potential energy fluctuation arises from the Coulomb potential of the ionized impurities per unit volume that are randomly distributed on lattice sites. The rms potential energy fluctuation δ is therefore given by [2]

$$\delta = q^2/(4\pi\epsilon).[2\pi(N_D^+ + N_A^-)L_s]^{1/2} \tag{8}$$

where L_s is the screening length which depends on the carrier concentration and on temperature. The border between degenerate and non-degenerate material is controlled by a critical temperature T_c which for n-type material is given by [3]

$$T_c(K) = (3/\pi)^{2/3}(h^2/8km^*)n^{2/3} \tag{9}$$

where m^* represents the effective density of states electron mass. The material is non-degenerate when T is greater than T_c. In this case, the screening length L_s is quite accurately represented by the extrinsic Debye length [2]

$$L_s = (kT\epsilon/q^2n)^{1/2} \tag{10}$$

On the other hand, the material is degenerate when T is smaller than T_c. In this case, L_s can be satisfactorily represented by the Thomas-Fermi screening length [1,2]

$$L_s = (\pi\epsilon h^2/m^* q^2)^{1/2}.(\pi/3n)^{1/6} \tag{11}$$

For the valence band in an n-type material, Kane's density of states has the form

$$D(E) = (2\eta_v)^{1/2}.(m_p^{3/2}/\pi^2h^3).y(E/\eta_v) \tag{12}$$

where m_p is the effective density of states mass for the valence band and $\eta_v = \sqrt{2}\delta$. The dimensionless variable y is represented in integral form by

$$y(x) = \pi^{-1/2} \int_x^\infty (z - x)^{1/2}.exp(-z^2)dz \tag{13}$$

Substituting some values for n-type silicon in the set of equations (8) to (13), we obtain $D(E)$ for the valence band (minority carrier band) which is diplayed in Fig.2 together with the parabolic band.

Other models have been proposed to represent the band tail states. For comparison, the density of states

based on Halperin and Lax model [4] is plotted together with Kane's density of states in Fig.3 [5]. The model of Halperin and Lax takes into account the kinetic energy of the carriers when solving Schroedinger equation. This kinetic energy was neglected in Kane's model and therefore it results in a much larger density of states. The extent to which kinetic energy effects reduce band tailing depends on the kinetic energy $h^2/2m^*L_\delta^2$ in relation to δ. Localized tail states have also been described by mobility edge models (e.g.[6]). We use Kane's model here merely to illustrate the importance of band tails in determining U. The effect can be evaluated for any other model of band tails in a similar way.

3 - DETERMINATION OF THE LOCALIZED STATES IN THE BAND TAILS

In order to determine the recombination center levels in the band tail, one has to define sharply the localization edge which is the energy level that separates the localized states from the extended states in the tail. Using a carrier transport model based on effective medium theory [7] and percolation theory [8], one can determine accurately the edge of localization in the minority carrier band tail. This model assumes that, due to band edge distortion, the volume of a heavily doped silicon crystal is divided into potential wells and barriers of different height as shown in Fig.4. The barriers constitute non conductive regions and free carriers can move between the hills only. The fractional space f occupied by free holes of energy E is given by

$$f = 1/2 erfc(E/\eta_v) \tag{14}$$

It is known that in a binary mixture of conducting (with conductivity g_1) and insulating particles, the effective conductivity in three dimensions is given by [9]

$$g = 1/2.g_1.(3f - 1) \tag{15}$$

Therefore, if f is smaller than the percolation threshold $f_{pc} = 1/3$ the conductivity g becomes zero. Putting $f = 1/3$ in equation (14) we get the percolation threshold energy

$$E = E_{pc} = 0.4306\,\delta \tag{16}$$

which represents the localization edge and in the same time the lower integration limit E_{min} of equation (7).

All states with energy above E_{pc} are localized and act as recombination centers. Due to their high density, only those centers having energies within a few δ's from E_{pc} are effective. Numerical results show that centers having energies greater than 6 δ contribute to the band tail recombination by less than 0.01 %. Therefore, we substitute for the higher integration limit of equation (7) by $E_{max} = 6\,\delta$.

4 - DOPING DEPENDENCE OF THE BAND TAIL LIFETIME

As discussed in the introduction, the recombination taking place via band tail centers is characterized by a lifetime τ_{bt} and an equivalent center density N_{tbt}. At low doping N_{tbt} is extremely small such that the band tail recombination is negligible. In this case, the lifetime τ_{eq} is totally governed by recombination via the centers situated at the discrete energy level near midgap. Assuming that the density of these centers is independent of the doping level, τ_{eq} will also be independent of the doping level and equals τ_o. At high doping levels the density of the band tail centers N_{tbt} becomes large such that τ_{eq} becomes much smaller than τ_o and approximately equal to τ_{bt}. Since N_{tbt} increases rapidly with the doping level, τ_{eq} will follow the inverse behavior. In a certain doping range results of numerical calculations show that

$$N_{tbt} = C.N^a \tag{17}$$

where C and the exponent a are constants and N is the doping concentration. Combining equations (5) and (17), the equivalent lifetime τ_{eq} can be expressed as

$$\tau_{eq} = \frac{\tau_o}{1 + (N/N_o)^a} \tag{18}$$

Equation (18) states that at low doping concentration $(N/N_o)^a << 1$ and $\tau_{eq} = \tau_o$ while at high doping levels $(N/N_o)^a >> 1$ such that $\tau_{eq} = \tau_o(N_o/N)^a$.

5 - NUMERICAL RESULTS AND COMPARISON WITH EXPERIMENTAL DATA

The band tail recombination center density N_{tbt} as expressed in equation (7) is plotted in Fig.5 versus doping level taking the electron to hole capture cross section ratio R as an independent parameter. At each doping level $D(E)$ is replaced by its corresponding Kane's function. E_{min} and E_{max} are determined in the way described in section 3. The intrinsic carrier concentration is modified to include band-gap narrowing [10]. As expected, a continuous increase in N_{tbt} with doping level is obtained. In the doping range 10^{16} - 10^{19} cm^{-3} the exponent a of equation (17) is strongly dependent on R. In order to fit our theory to the reported experimental data representing the fundamental minority carrier lifetime in n-type and p-type material [11] we have to:

1) determine the constants τ_o, N_o and a of equation (18) from the experimental dependence. Indeed, τ_o is the constant lifetime of the low doping region, N_o is the doping level at which $\tau_{eq} = \tau_o/2$ and a is the slope of the high doping region.

2) use Fig.5 to relate the value of a to its corresponding R value (note that due to the difference in the effective mass of the minority carrier, $D(E)$ hence N_{tbt} are not equal for p-type and n-type silicon). Consequently the N_{tbt} curve that has to be considered is defined. Knowing N_{tbt} and τ_{bt} and using equation (6) we are able to extract the value of the minority capture cross section. Since R has been determined in a previous step, we can now obtain a value for the majority capture cross section.

The results of such procedure for n-type and p-type material are given in table 1 and the fit is displayed in Fig.6. The values of R seem reasonable and lie in the range reported in the literature. On the other hand, the values of σ_n and σ_p are very small compared to the values usually met. This result could have been expected since, as discussed earlier, Kane's density of states are overestimated and a more accurate model such as the one in [4] would reduce $D(E)$ by two orders of magnitude. Therefore, N_{tbt} would roughly be reduced also by the same amount which leads to an increase in the extracted capture cross section values by a factor of 100. On the other hand, the idea that the tail recombination centers have smaller capture cross sections than the deep level center is not totally rejected.

Table 1

	quantity	n-type	p-type
fitting param.	R	0.01	3
constants	τ_o (s)	$4x10^{-4}$	$7x10^{-4}$
	a	0.743	1.171
	$N_o(cm^{-3})$	$3x10^{15}$	$4x10^{16}$
results	$\sigma_p(cm^2)$	$1.334x10^{-20}$	$2.328x10^{-20}$
	$\sigma_n(cm^2)$	$1.334x10^{-18}$	$7.000x10^{-20}$

6 - DISCUSSION

More generally, the reported experimental values of the fundamental minority carrier lifetime suggest values for the quantity $\sigma.N_t$ between 10^{-4} and 10^{-1} /cm in the doping range 10^{16} - 10^{19} cm^{-3} respectively. Considering that reasonable values of σ would lie between 10^{-13} and 10^{-17} cm^2, the corresponding N_t values should be 10^9 -10^{13} cm^{-3} at a doping of 10^{16} cm^{-3} and 10^{12} - 10^{16} cm^{-3} at a doping of 10^{19} cm^{-3}. By consulting Fig.5 we clearly see that even if N_{tbt} is reduced by two order of magnitude its value could still easily lie in the range of values stated above. We therefore conclude that band tail SRH recombination is a phenomena that cannot be neglected in heavily doped silicon. One should be very carefull when trying to model the fundamental lifetime in silicon when the doping level exceeds 10^{17} cm^{-3}.

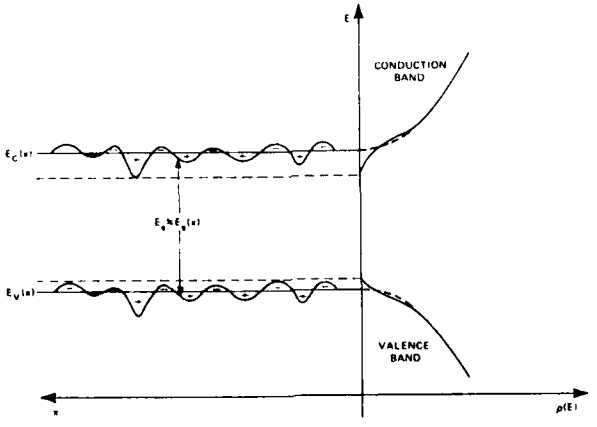


Fig. 1 - Schematical representation of the band tail formation resulting from random spatial variation of the electrostatic potential.

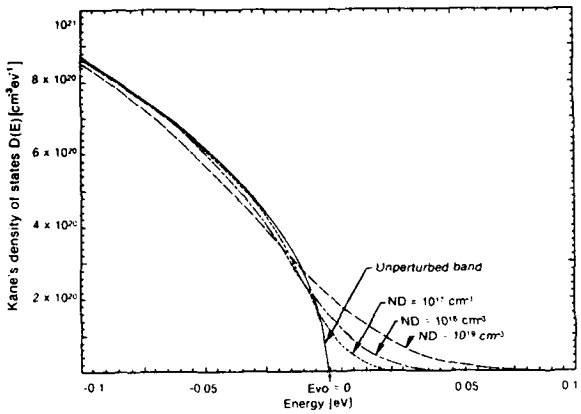


Fig. 2 - Density of states in the valence band of n-type silicon as calculated using Kane's model [1].

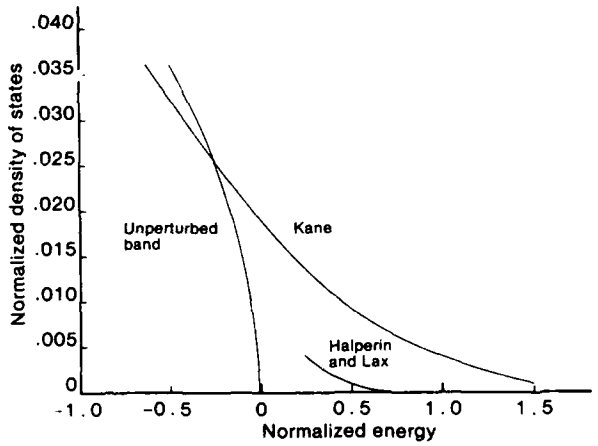


Fig. 3 - Comparison between the density of states calculated using Kane's model [1] and Halperin and Lax model [4].

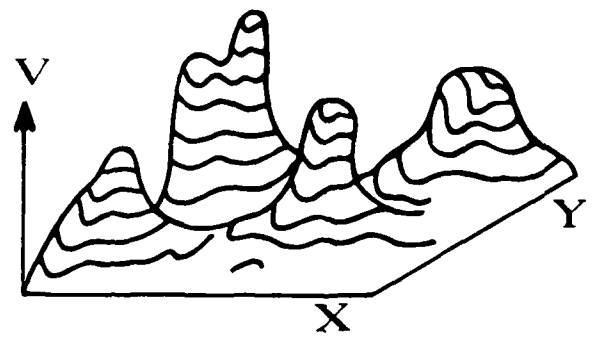


Fig. 4 - Schematic representation of the three dimension spatial variations of the valence band edge forming potential barriers and wells due to inhomogeneous impurity distribution in heavily doped n-type silicon.

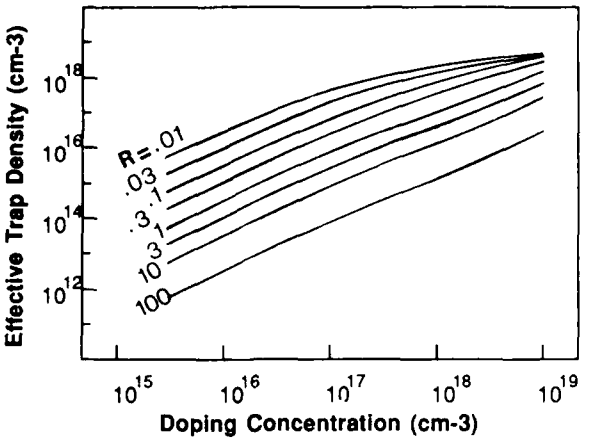


Fig. 5 - Equivalent recombination center density in the band tail of n-type silicon (D(E)=Kane's) versus doping concentration. R represents the hole to electron capture cross section ratio.

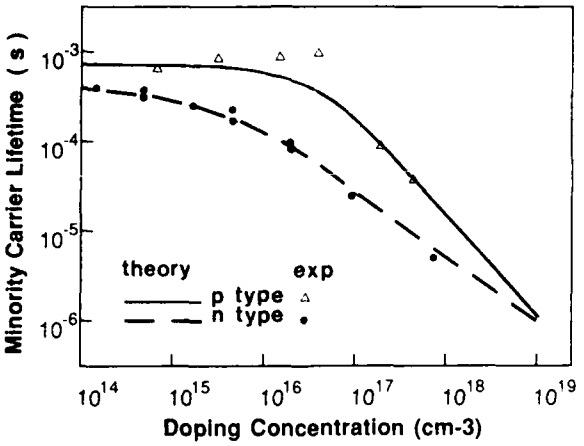


Fig. 6 - Fitting of the experimental data [11] concerning the fundamental lifetime versus doping for n-type and p-type silicon using a combined SRH lifetime that includes band tail recombination.

REFERENCES

[1] E. O. Kane, Phys. Rev. 131 (1963) 79.

[2] H. C. Casey, Jr. and M. B. Panish, "Heterostructure Lasers," Part A p. 135, Academic Press Inc., 1978.

[3] W. Shockley, "Electrons and holes in semiconductors," p. 242 Van Nostrand, Princeton, 1950.

[4] B. I. Halperin and M. Lax, Phys. Rev. 148 (1966) 722.

[5] R. A. Abram, R. J. Rees and B. L. H. Wilson, Advances in Physics 27 (1978) 799.

[6] J. G. Fossum, D. E. Burk and S. Y. Yung, IEEE Trans. Electron Devices ED-32 9 (1985) 1874.

[7] E. Arnold, Surf. Sc. 58 (1976) 60.

[8] S. C. Jain, M. Y. Ghannam, R. P. Mertens, J. F. Nijs and R. Van Overstraeten, IEEE Photovoltaic Specialist Conf., Las Vegas, Sept 26-30 (1988).

[9] S. Kirkpatrick, Phys. Rev. Lett. 27 (1971) 1722.

[10] J. W. Slotboom, Solid-St. Electr. 20 (1977) 279.

[11] J. G. Fossum and D. S. Lee, Solid-St. Electr. 25 (1982) 741.

Session 3B Room B : Reliability

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

GATE OXIDE QUALITY OF DRAM TRENCH CAPACITORS

S. RÖHL, M. ENGELHARDT, W.-U. KELLNER and A. SCHLEMM

Siemens AG, Microelectronic Technology Center, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

Abstract - The quality of thin trench capacitor oxide dielectric for 4M and 16M DRAM generations is investigated by leakage current and time dependent dielectric breakdown measurements. Geometric trench shape and surface smoothness influence leakage currents and thus dielectric lifetime, but etch chemistry, side wall redeposition and contamination are more important factors which reduce gate oxide quality. It is shown that silicon removing post treatments regain gate oxide quality comparable to planar capacitors.

1 - INTRODUCTION

Oxides as thin as 10 nm are currently used in trench structures as capacitor or gate dielectric of 4M and 16M DRAMs . Leakage currents and breakdown voltages have been discussed (ref.1) but few reliability and lifetime studies of large area trench structures have been published to date and the factors influencing gate oxide quality in trenches are not very well known. In this work we have investigated the influence of different trench etch processes and post treatments on gate dielectric quality.

2 - EXPERIMENTAL METHOD

Etch processes with $CBrF_3$ in a single wafer triode etcher and processes using BCl_3/Cl_2 and $CBrF_3$ in a batch reactor were optimized with respect to the trench profile. The aim was a round trench bottom, sidewalls perpendicular to the wafer surface without bowing or undercut. The trench profiles obtained ($1/\mu m^2$ cross section, 4-5 $/\mu m$ depth) are summarized in fig.1. For single wafer processing, no.1 to no.3, $CBrF_3$ gas flow is increased and for no.4 $CBrF_3$ gas is diluted by N_2, because processes were more stable with respect to homogeneity and trench profile reproducibility at higher gas flow. Batch processing no.5 and 6 showed less sidewall redeposition as compared to no. 1-4 and had slight trenching (no. 5) and severe trenching (no. 6), respectively. Process no. 6 was therefore only applicable after a wet anisotropic etch removing some 100 nm of Si. A typical example of sidewall film is shown in the SEM of fig.2. Strong sidewall films help to avoid trenching at the bottom, produce a round shape and prevent undercut. Auger depth profiles for O, C and N were taken at three trench depths to analyse the sidewalls. They proved that the redeposition (fig.2) consists mainly of SiO_2. No excess carbon was found inside the trench as compared to the cleaved surface outside the trench. Nitrogen was however found for the diluted process (no.4). This reduced to below the detection limit of 2 at.% only after O_2 plasma treatment (fig. 3). Different post treatments were applied to remove sidewall films, contamination or Si damage viz: buffered HF (BHF) oxide etching, O_2 barrel reactor plasma, CF_4/O_2 isotropic dry Si etch and aniso-

Fig.1: SEM micrographs of the trench process of six different etch processes after removal of sidewall redepositions.

a) Single wafer processes with $CBrF_3$
- # 1 low
- # 2 medium
- # 3 high
- # 4 N2 diluted

} gas flow

b) Batch processes
- # 5 $CBrF_3/BCl_3/Cl_2$
- # 6 BCl_3/Cl_2 (including wet anisotropic Si etch)

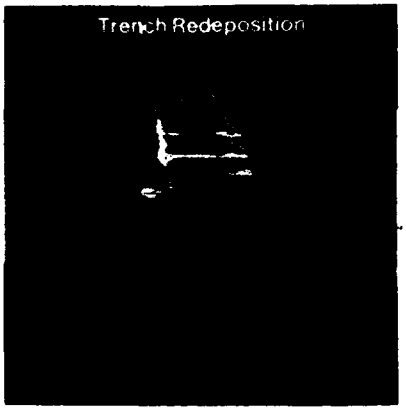


Fig.2: SEM micrograph (oblique view) of trench process # 2 showing typical sidewall redeposition

Table 1: Process flow

p-type <100> Si
deep p-well
LOCOS isolation
trench mask
mask etching
trench etching (4-5 um deep)
redeposition etching
post treatment
As trench doping
sacrificial oxide
gate oxidation 13.5 nm
poly-Si deposition and doping
trench refill
poly-Si electrode patterning

tropic wet Si etching using an organic base etchant. Trench and planar capacitor arrays of $1mm^2$ active area and an oxide dielectric of d=13.5 nm were then fabricated with the process flows given in table 1. A SEM cross section of the gate oxide at the upper trench corner is given in fig.4.

Tunneling currents of trench and planar capacitor arrays were measured (fig.5) revealing sharp corners at the upper trench edge (gate positive) or at the trench bottom (gate negative). Early failures were determined by loading the capacitors with a constant current of $60 \mu A/cm^2$ to the saturation voltage (70 msec) and accepting only those systems that reached 5-12 MV/cm. To test stability a test condition of $1 mA/cm^2$ for 0.5 sec was chosen as an accelerated stress to obtain reasonable measuring times. This condition corresponds to a charge density of 5.10^{-4} C/cm^2 and occuring failures are classified as defect related (ref.2). On a few batches long term measurements at the same current density were performed in order to determine the charge necessary for an intrinsic breakdown (ref.3).

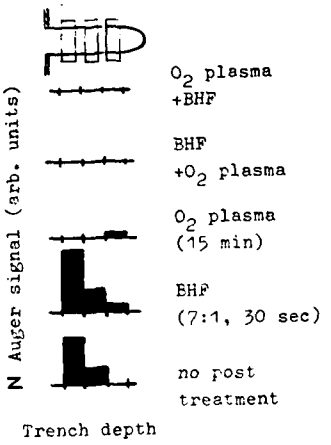


Fig.3: Auger signal of nitrogen on trench sidewall (process # 4) at three depths, normalized to the cleaved surface outside the trench

Fig.4: TEM cross section of the upper trench corner
a) as etched (surface roughness on trench sidewall is 3nm)
b) after sacrificial and gate oxidation

3 - RESULTS AND DISCUSSION

The trench tunneling curves of (fig.5) are due to the oxide thinning at the upper trench corner (fig.4). The close agreement of trench and planar curves for gate negative show that the trench bottom is round. Process no.5 and 6 show here reduced values and a much larger spread. The small spread of trench curves indicates a very homogeneous thin oxide at the upper trench corner across wafer and batch. The oxide thinning should therefore only reduce intrinsic oxide breakdown. If defect related breakdown occurs it should not be related to this oxide thinning.

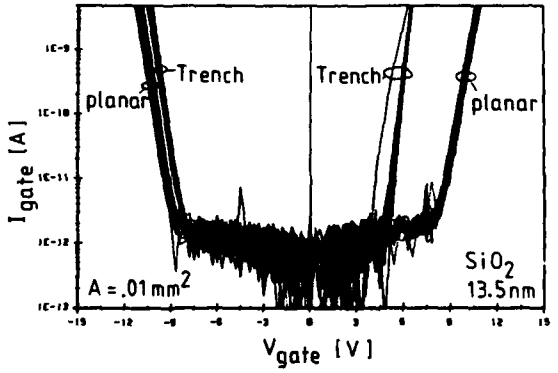


Fig.5: Tunneling currents for both gate polarities from 25 die on one wafer for planar and trench capacitor arrays (process # 2)

The gate oxide quality taken as the percentage of capacitors that withstood the above test criteria is shown in fig.6. A reference of planar capacitor arrays is given on the left and is compared to trench process no.1 to 6. Different post treatments are detailed and comparison with a triple trench dielectric (ONO) consisting of an oxide-nitride-oxide stack of 13.5 nm oxide equivalent thickness is shown on the right (ref.3).

Dielectric quality close to that of planar capacitors is found for process no.1 and 2 but these lack reproducibility in etch rate and trench profile from batch to batch. The higher gas flow leads to more stable geometric profiles and etch rates and to high repeatability (no.3, 4) but these processes need a complex post treatment to reach similar yields. The O_2 plasma removes the nitrogen (see fig.3) while HF-etching removes the SiO_2 redepositions. Etching off approximately 50 nm Si in CF_4/O_2 plasma or by wet che-

mical etching improves gate oxide quality. We assume that the strong redeposition formation of process no. 1 and 2 hinder contamination of Si sidewalls by etch products whereas reproducibility is degraded by a high degree of redeposition formation. This is corroborated by the fact that reduction of nitrogen (compare fig. 3) and etching of Si improve gate oxide quality substantially . Another possible cause for gate oxide degradation would be crystal damage at the trench bottom. This is however ruled out by experiments showing no improvements when process no.2 which had shown a good oxide quality is added to process no.4 as in situ dry post treatment

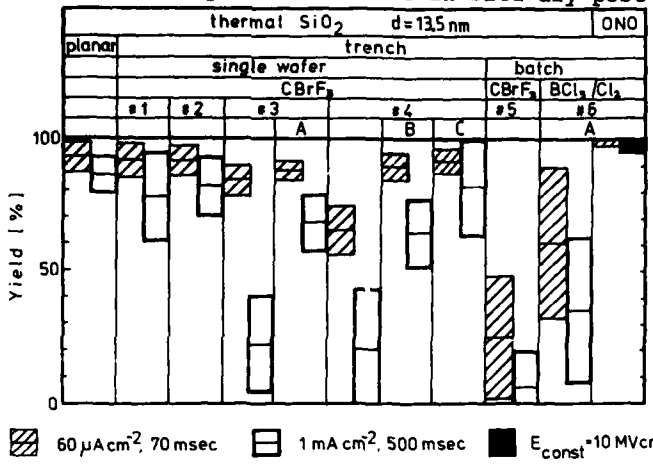


Fig.6: Gate oxide results for single wafer and batch etch processes
Hatched boxes indicate means and standard deviations.
For # 3 and # 4 results for the following post treatments are shown: A = wet anisotropic Si etch, B = O_2 plasma + BHF C = O_2 plasma + BHF + dry isotropic Si etch. Triple trench dielectric (ONO) and planar reference are shown for comparison.

The single wafer processes were superior to the batch processes investigated in this work, presumably because of enhanced sidewall redeposition.

The extremely high percentage of early failures of process no. 5 can be explained by trenching problems of varying degree at the trench bottom. Similar experiments using ONO dielectrics showed trench shape to have little influence, except in the case of process no.5, and also resulted in superior dielectric quality.

4 - CONCLUSION

Gate oxide quality in trench structures is sensitive to trenchwall contamination by etch byproducts. Sidewall films can protect against these contamination but lead to unstable process conditions. To increase the process window post treatments are necessary. Composite dielectrics (ONO) are much less sensitive to the trench etch process. Even if these take over in present and future DRAM generations trench oxide quality will remain an efficient tool for the optimization of trench etch processes in VLSI applications.

This work was supported by the Federal Department of Research and Technology (sign NT 2696)

References:

/1/ K. Yamabe, K. Imai, IEEE Trans. Electron Devices, vol.ED-34, no.8, 1987, p1681
/2/ D.R. Wolters and J.J. van der Schoot, Philips Journal of Research 40, 1985, p 115
/3/ P. Hiergeist, A. Spitzer, S. Röhl, to be published in IEEE Transactions on Electron Devices

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

IMPROVING RELIABILITY USING DESIGN CENTERING

W. HEIMSCH, R. KREBS, K. ZIEMANN and D. MOEBUS*

Siemens AG, Corporate Research and Development, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.
**Institut für Netzwerk und Systemtheorie, Universität Stuttgart, Stuttgart, F.R.G.*

RÉSUMÉ
Ici sont analysées les performances d'une porte NOR en technologie BiCMOS. Afin de réduire l'influence de la dispersion des paramétres de fabrication sur les fluctuations des différents courants internes, un logiciel d'aide á la conception (Design Centering Program) optimise la largeur des transistors MOS. Ceci permet, pour une fluctuation de 5% du courant de chaque transistor MOS, d'améliorer les performances (temps de propagation) du circuit d'environ 37%.

ABSTRACT
In this work the reliability of a BiCMOS NOR Gate is investigated. A design centering program is used to dimension the width of the MOS transistors in order to make the circuit utmost insensible to current efficiency fluctuations caused by process parameter deviations. Therewith the reliability could be increased by 37% while the current efficiency of the single MOS transistors fluctuates within its 5% value.

1. INTRODUCTION

For conventional circuit design it is almost impossible to take into account processing fluctuations and environmental influences (e.g. supply voltage, temperature) on circuit reliability. One way to get rid of this problem is the so-called "worst case design". But this is often an elaborate way. A better choice for improving circuit reliability is a design centering program, that will give a circuit design almost insensitive to parameter fluctuations. The design centering program applied here uses a Monte Carlo search algorithm for that case.

2. DESIGN CENTERING

The reliability of circuits depends on the fluctuation of design parameters. Mobility of the charge carriers in the channel, threshold voltage difference, oxide thickness for MOS transistors and current gain hfe for bipolar transistors are the most important ones. To increase the reliability of a BiCMOS NOR Gate (fig. 1), an essential part of a BiCMOS SRAM decoder, we used a design cen-

tering program /1/. The circuit has the task to switch a capacitive node of 5 pF faster than 2.5 ns.
The switching time of a circuit is directly coupled with the current efficiency of the single transistors which can be described by their channel width W. Mobility, threshold voltage, hfe, gate oxide thickness and channel length fluctuate within the 3σ value according to table 1.

Table 1: Parameter fluctuations

oxide thickness	±2.5 nm
threshold voltage	±250 mV
supply voltage	±10%
channel length	±10%
carrier mobility	±10%
current gain hfe	±10%

The design centering program generates a gaussian distributed cloud of points in the parameter space around the nominal value of the transistor width. All points which lead to a switching time shorter than 2.5 ns are acceptable points and the mean value of all these acceptable points becomes the new nominal value. Then a new region of acceptability is generated. This method leads finally to a circuit design which is utmost insensitive to process fluctuations.

3. RESULTS

The result can be seen in fig. 2. The design optimized with respect to switching time using ideal design parameters leads to a delay time of 2.1 ns. Its reliability for a switching time shorter than 2.5 ns reacts extremely sensitive to current efficiency fluctuations (▲). The design centered version (●) shows less dependence on this fluctuations even for extremely high fluctuations. For a standard deviation of 5% current efficiency fluctuation the improvement on reliability is 37%.
The design parameters (table 2) for this worst case analyses are based on a 1.2 µm BiCMOS process with a polysilicon emitter npn bipolar transistor (fig. 3).

Table 2: Process characterization.

bipolar transistor		MOS transistor	
emitter area	1.6 µm * 10 µm	gate oxide thickness	250 Å
hfe	100	N-channel L_{eff}	1.1 µm
ft	9 Ghz	V_{th}	0.87 V
Re	11 Ω	p-channel L_{eff}	1.1 µm
Rc	40 Ω	V_{th}	-0.9 V
Rb	325 Ω		
C_{je}	60 fF		
C_{jc}	78 fF		
C_{js}	70 fF		

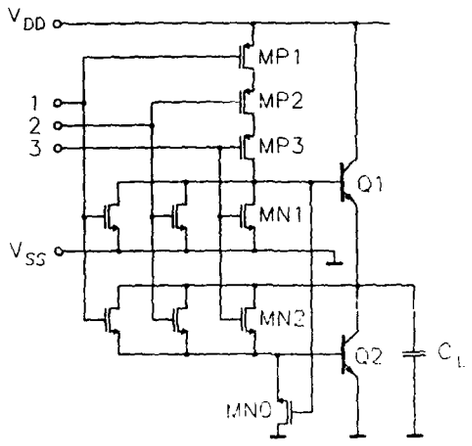


Fig. 1: BiCMOS NOR gate

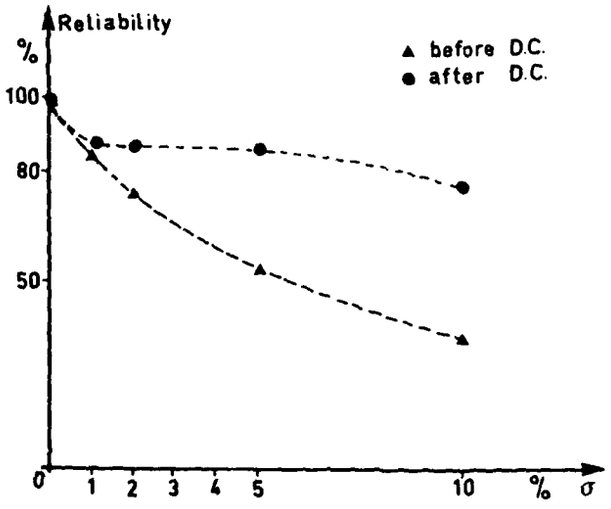


Fig. 2: Reliability versus standard deviation of the transistor current efficiency before and after design centering (D.C.).

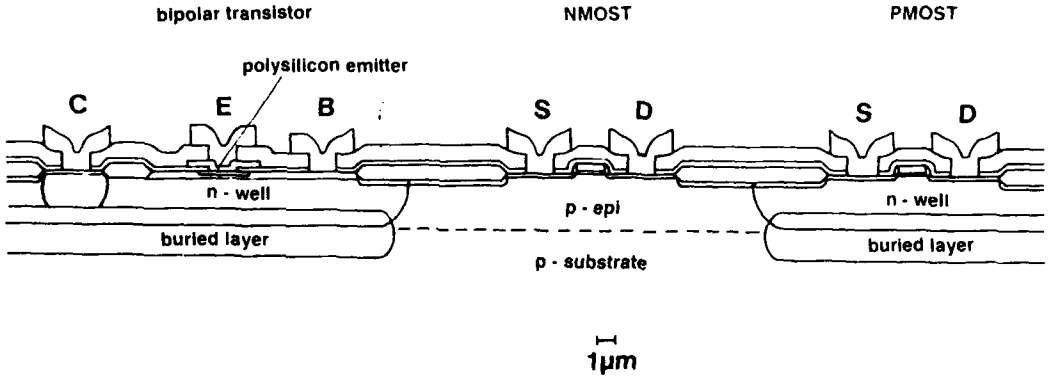


Fig. 3: Cross sectional view of bipolar and MOS transistors.

This work was carried out within ESPRIT project BiCMOS 412.

REFERENCES

/1/Eckstein T., Lüder E., 'Design Centering by improved Monte Carlo analyses of the region of acceptability', Proc.IEEE, Int. Symp. Circuits and Systems, San Jose, USA, 1986.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988 C4-291

ELECTRON BEAM DIRECT WRITE EFFECTS ON CMOS DEVICES

K. BARLOW

Plessey Research Caswell Ltd., Caswell, Towcester, Northants., GB-NN12 8EQ, Great-Britain

Abstract Electron beam exposure effects on metallised MOS devices have been studied. The results show that exposure leads to the generation of interface traps and positive oxide charge that can be almost completely removed using low temperature annealing (450°C). However, bulk oxide traps generated during exposure are only partly removed by the anneal and this leads to an increased trapping efficiency, which can seriously degrade MOS device integrity.

INTRODUCTION

E-beam direct wafer writing is especially useful in metal patterning for discretionary wiring and wafer scale integration. However exposure of MOS devices to high energy electrons can result in the generation of oxide bulk and interface trapping centres, which can impair device performance. Whilst low temperature annealing is effective in reducing interface traps, bulk oxide traps can only be completely removed by high temperature anneals([1],[2]) and this creates problems after metal patterns have been deposited. Thus e-beam exposure of metallised MOS devices can lead to degradation of device performance and lifetime because of the restrictions on post e-beam anneal temperatures and times. This paper shows the results of an investigation into the effects of e-beam exposure on MOS devices, exposed to doses in the range 0.15μC/cm^2 to 5μC/cm^2. The doses are applicable to those required for metal patterning and due to the limited range of anneal temperatures that can be used, this is an important area.

Experiment

An advanced CMOS process employing trench isolation, with a 20nm gate oxide, oxide sidewall spacers and silicided source, drain and gates was used. The devices were fully processed with their surfaces protected by thick dielectrics and patterned metal (0.5μm of Al/Cu on top of 0.15μm of TiW). Parts of the wafers used were kept as controls, whilst the rest were subjected to e-beam doses in the range 0.15μC/cm^2 to 5μC/cm^2 at 20kV. No resist was used on the wafers during exposure and the doses were chosen to simulate typical exposures required for metal patterning. Measurements were made before and after exposure, then the samples were annealed in H_2/N_2 at 425°C for 30 minutes.

Results and discussion

Fig.1 shows the threshold voltage shift, Vt as a function of dose for nmos transistors with drawn gate lengths in the range 5μm to 1μm. The threshold voltage was measured before and after exposure to obtain the Vt shift. The plot shows that a rapid shift is seen, for even the smallest dose (0.15μC/cm^2) followed by a tendency toward saturation of Vt as the dose increases further, also the Vt shift tends to increase with gate length. The shift, which corresponds to a reduction in Vt is due to the generation of positive charge throughout the gate oxide during exposure. Besides positive charge generation, e-beam exposure also produced a large increase in the interface trap density which led to transistor subthreshold and gain degradation. The subthreshold slope increased from a pre exposure level of 98mv/dec to around 300mv/dec after a 0.5μC/cm^2 exposure and up to 660mv/dec after a 5μC/cm^2 exposure, whilst the gain was reduced by around 30% after a 5μC/cm^2 exposure. Junction leakage also increased dramatically from a pre exposure level of a few pA to 1-10μA after a 0.5μC/cm^2 dose. This increased leakage was mainly attributed to degradation of the trench parasitic transistors. These parasitic transistors are normally off and suppression of them is paramount for good and reliable circuit operation([3]). The lateral and vertical

trench parasitic voltages before e-beam exposure were around 5.5V and 5V respectively. After exposure the potentials fell by 50% after a 0.15μC/cm² dose, and for higher doses the trench potentials could not be measured on the test system due to the very high leakage currents flowing at zero volts. The degraded trench potentials are due to the generation of interface traps and positive oxide charge in the trench sidewall oxide, this leads to increased depletion and higher leakage currents flowing.

Low temperature annealing (425°C, H_2/N_2, 30 mins) led to almost complete recovery of the trench potentials, with the worst case being 93% recovery after exposure to a 5μC/cm² dose. The anneal produced almost complete recovery of the nmos device characteristics by reducing interface trap levels to their original values and removing the majority of the generated positive charge, due to recombination with electrons tunnelling from the silicon. Fig.2 shows the Vt shift as a function of dose after annealing, comparison with Fig.1 shows a drastic recovery in Vt, the recovery improving with gate length. The gain is recovered to its pre exposure levels, whilst subthreshold and junction leakage improved due to the reduction in the interface trap density and increase in the trench potentials, Fig.3 shows the recovery in the subthreshold characteristics of a 1μm gate length nmos device exposed to a 0.15μC/cm² dose after annealing.

Reliability

In order to assess the effects of e-beam exposure on reliability, charge injection measurements were performed on gate oxides and electrical stress tests on nmos devices, biased to ensure hot carrier injection into the gate oxide were made.

Charge injection measurements were performed using a constant current density of 20μA/cm² and the injection mode was such that electrons were injected into the oxide over the oxide/si barrier. Prior to e-beam exposure injection measurements determined the oxide bulk trap density to be $\sim 10^{10} cm^{-2}$, after exposure injection measurements showed a large increase in electron trapping. This large increase in trapping sites is shown in Fig.4a where the maximum flatband voltage shift as a function of dose is plotted, for a 5μC/cm² dose the bulk trap density has increased to $\sim 5.10^{11} cm^{-2}$.

Low temperature annealing, successful in reducing the interface trap density and generated positive charge, only partly reduced the density of neutral traps as Fig.4b shows, with the bulk trap density now around $2.3.10^{11} cm^2$ for an oxide exposed to a 5μC/cm² dose. Thus exposure of the gate oxide leads to a large increase in the density of neutral traps that can only be partially removed using a low temperature anneal. This has important implications for device lifetime as Fig.5 and Fig.6 show. Fig.5 shows the Vt shift as a function of electrical stress time and Fig.6 shows gain degradation as a function of stress time, for exposed and control (unexposed) devices with 1μm gate lengths. The bias conditions were set to give the maximum substrate current and arsenic source and drains were employed to ensure any differences between the Vt and gain plots of the exposed and control devices could be clearly observed without the need for large biases. Fig.5 shows that the device exposed to a 0.15μC/cm² dose and then annealed shifts 55mv in 17 hours whilst the control device shifts 20mv in 20 hours, whilst Fig.6 shows that gain degradation for the e-beam exposed device is significantly larger than for the control.

Summary

Electron beam exposure leads to the generation of oxide bulk and interface trapping centres and positive oxide charge which leads to degradation of nmos device and trench parasitic characteristics. Low temperature annealing almost fully recovers subthreshold, junction leakage and gain characteristics, whilst the Vt is found to recover to within 25mv of its original value, depending upon gate length. However the large density of neutral traps introduced during exposure are only partly annealed out and this leads to increases in the oxide trapping efficiency during hot carrier injection which greatly enhances Vt shifts and gain degradation, which can seriously affect circuit performance. Neutral traps are only effectively removed by annealing at temperatures greater than 700°C; annealing at such temperatures after metallisation would require the use of non aluminium based alloys, otherwise reduced doses or energies during e-beam writing of metal patterns could be required.

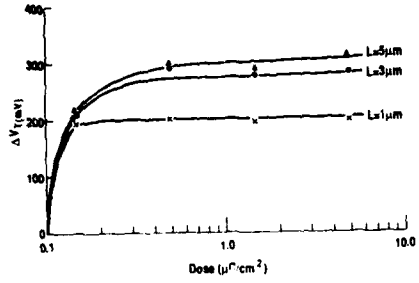


Figure.1 : Threshold voltage shift as a function of dose, with gate length as parameter.

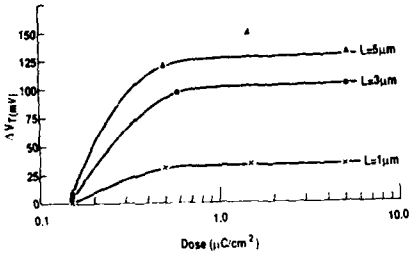


Figure.2 : Threshold voltage shift as a function of dose, with gate length as parameter, after annealing.

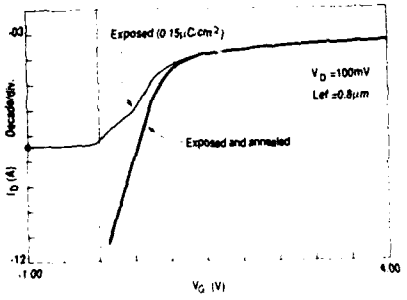


Figure.3 : Pre and post anneal subthreshold characteristics for a 20/1 nmos device exposed to a 0.15μC/cm² dose.

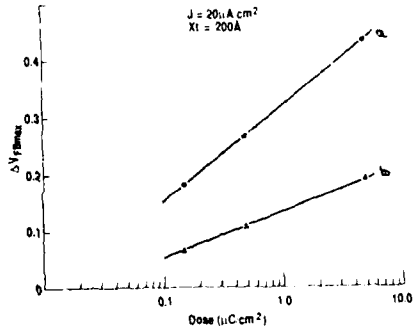


Figure.4a : Maximum flatband voltage shift as a function of dose.

Figure.4b : Maximum flatband voltage shift as a function of dose, after annealing.

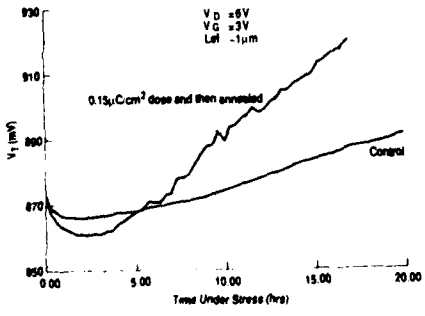


Figure.6 : Threshold voltage shift as a function of stress time.

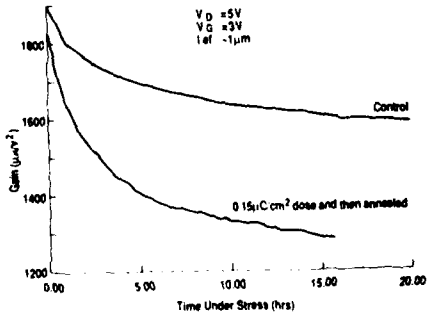


Figure.6 : Gain degradation as a function of stress time.

Acknowledgements

The author would like to thank the Alvey directorate for financial support and A. Marsh for useful discussions.

References

1. M. Peckerar, R. Fulton, D. Brown and R. Whitlock. 1979, J.Vac.Sci.Technol, 16(6), 1658-1661.

2. M. Shimaya, N. Shiono, O. Nakajima, C. Hashimoto and Y. Sakakibara. 1983, J.Electrochem.Soc, 130(4), 945-950.

3. M.C. Roberts, D.J. Foster, P.H. Bolbot and P.L. Medhurst. 1987, proc. 17th ESSDERC Conf, Bologna, Italy, 589-592.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

EFFECTS OF GAMMA RADIATION ON TRENCH ISOLATED CMOS

P.L. MEDHURST and D.J. FOSTER

Plessey Research Caswell Ltd., Caswell, Towcester, GB-Northants NN12 8EQ, Great-Britain

Abstract The gamma radiation hardness of trench isolated CMOS has been assessed with respect to the shift in threshold voltage of parasitic sidewall devices and minimum channel length NMOS and PMOS transistors. For parasitic devices, saturation of the threshold voltage shift is apparent after 20kRads. This is attributable to the low volume oxide films inherent to the device structure. Drifts in the parasitic threshold voltage over a 0 to 50kRads range are significantly lower than for similar LOCOS field devices. Off-state leakage current is generated in NMOS transistors and has been characterised as a function of gamma ray dose; such leakage current levels also saturate with irradiation dose.

Introduction

The radiation hardness tolerance of an MOS transistor is becoming an increasingly important criterion for space applications. This paper describes the effect of cumulative doses of gamma radiation on the electrical performance of NMOS and PMOS transistors fabricated by a one micron CMOS process. Trench technology was used for device isolation and associated parasitics were compared to field transistors in conventional LOCOS isolation. The positive charge generated in thick field oxides (approximately 0.5μm) of LOCOS devices can rapidly cause parasitic devices to link-up and short power supplies. With trench isolated devices, a thin sidewall oxide (0.15μm) offers a better degree of tolerance due to its lower volume. Devices were irradiated using a cobalt 60 gamma ray source. In both cases, for LOCOS and trench, 5V was applied to either polysilicon over field oxide or the trench polysilicon; this was especially severe for the trench case where the polysilicon remains around 0V during circuit operation.

Experiment

The trench device architecture uses 1.8μm deep n-wells in p-type substrate and 2μm deep trenches. There are two sources of oxide charge within the transistor structure: (i) the 1500A oxide passivation layer lining the trench sidewalls, (ii) the 200A gate oxide layer of the top transistor. The trench sidewall oxide is the dominant feature since it surrounds the entire device. On top of this is the polysilicon filling which would act as a parasitic gate for any leakage path to the substrate. A cross section of the trench structure is shown in Fig.1.

The devices were mounted on test boards and wired for the specific bias conditions of 5V for NMOS devices and -5V for PMOS devices, with the bias applied to either the trench polysilicon or to the silicided gate electrode. A high electric field would subsequently be exerted across the sidewall oxide for parasitic devices and the gate oxide for the top transistors.

Devices were irradiated under these bias conditions in accumulative doses up to 50kRads, being withdrawn at the required dose for d.c. measurement. Two dose rates were available: 5.3kRads/hr and 1.75kRads/min, the latter being selected for incremental dosages of 10kRads and above. The shift in threshold voltage was monitored as a function of gamma ray dose together with any associated increase in off-state leakage current. These were measured for both lateral and vertical trench parasitics, and for the top transistors. Fig.2 shows a schematic of the device cross-section and associated current flow from the parasitic transistors.

The effect of gamma radiation is to create electron-hole pairs uniformly in the oxide layer. Electrons lost to the substrate will allow a net positive charge build-up in the oxide. This increases with further irradiation, causing the threshold voltage to shift in the negative direction. The magnitude of this shift has been reported to follow a cubed relationship with oxide thickness (1). For PMOS devices, the negative shift enhnces the suppression of parasitic action. However, for NMOS devices, the thresholds are seen to fall thereby allowing additional leakage currents to flow.

The shift in threshold voltage as a consequence of gamma radiation was observed to saturate for both NMOS and PMOS parasitics and top transistors. This can be related to the small volume of oxide adjacent to the device channel which limits the amount of positive oxide charge build-up. Saturation in the threshold shift for the top transistors (Fig.3a) and PMOS parasitics was observed to be at 20kRads. The transition to saturation was softer for the NMOS parasitics with the onset occurring between 35 to 40kRads. Fig.3b shows the saturation effect for the sidewall parasitic devices.

The lateral and vertical PMOS parasitic devices showed drifts in the threshold voltages of 68mV/kRad and 104mV/kRad respectively. These results are significantly lower than those from similar LOCOS devices. Parasitic threshold drifts for LOCOS have been measured as high as 420mV/kRad for aluminium field and 930mV/kRad for polysilicon field devices (Fig.4), and therefore can only tolerate approximately 5 to 10kRads irradiation. Lateral and vertical NMOS parasitic devices have demonstrated drifts in threshold voltage of 340mV/kRad and 230mV/kRad respectively. Although higher than their PMOS counterparts, they are still three to four times lower than LOCOS. Off-state leakage was negligible for all PMOS devices over the range 0 to 50kRads. Some leakage was observed for the lateral NMOS parasitic, rising to a peak of ~9μA after 30kRads where it saturated. The vertical NMOS parasitic showed a smaller rise, peaking at ~18nA at the 30kRad saturation level. Fig.5a shows the variation in off-state leakage current with gamma-ray dose for NMOS parasitics. The observed saturation in threshold shift and leakage current of the trench devices as a result of the thin oxide, implies a tolerance to radiation doses much higher than 50kRads.

The threshold voltages associated with the top transistor showed lower drifts (because of the thinner gate oxide) of 2mV/kRad for NMOS and 3.2mV/kRad for PMOS. This again is a significant improvement over LOCOS. The leakage associated with the NMOS top transistor was negligible at low doses, only being detectable after 30kRads and rising to 2μA at 50kRads (Fig.5b). This is due to the charge inversion along the trench sidewall. A thinner oxide and higher well doping would reduce the inversion. In circuit operation, the field across the sidewall would be about 0V since the trench polysilicon would be zero. Therefore a lower off-state leakage would be expected.

Conclusions

The small volume of the trench sidewall oxide is significant in maintaining tolerance to radiation harsh conditions. It has been shown that the saturation in the parasitic threshold voltage shift through oxide charge build-up, occurs at 20 to 40KRads and that subsequent irradiation will have no further effect on threshold shift. The off-state leakage identified in NMOS devices could be reduced by higher p-well implants, thinner oxides and lower fields across the trench oxide during irradiation, with saturation levels kept low to effect the functionality of static logic circuitry such as used in gate arrays. The relatively low leakage and threshold drifts identified for trench isolation compared with LOCOS offers excellent potential for this technology for operation in radiation harsh environments.

Acknowledgements

The authors would like to thank the Plessey Company for permission to publish this work, and the Alvey Directorate for their financial support.

References

(1) G F Derbenwick, B L Gregory, IEEE Trans.Nucl.Sci., Vol.NS-22, No.6, Dec.1975.

Figure 1

Cross-section of a trench structure

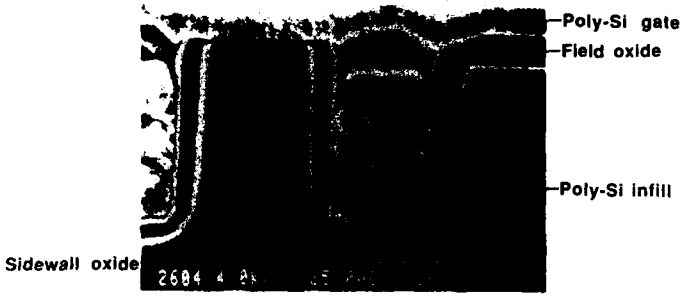
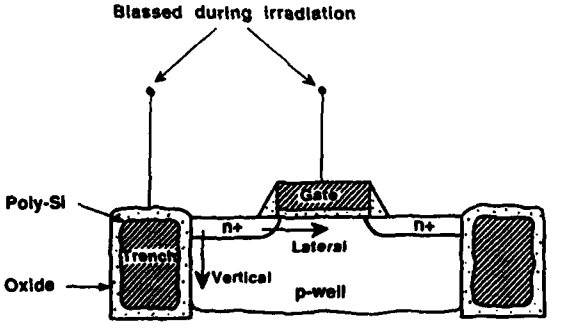


Figure 2

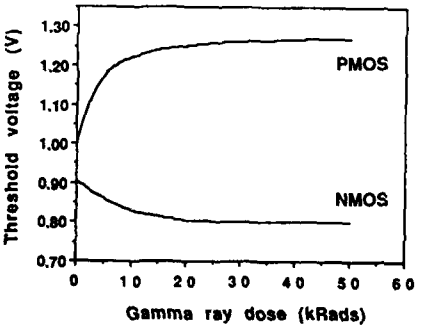
Schematic of Device and associated parasitic transistors



Current flow as a result of oxide charge

Figure 3a

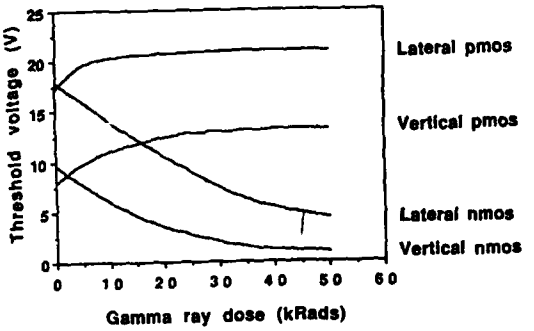
Threshold characteristics of the nmos and pmos top transistors as a function of gamma ray dose



Thresholds are negative for pmos and positive for nmos

Figure 3b

Parasitic threshold voltages as a function of gamma ray dose



Thresholds are negative for pmos and positive for nmos

Figure 4

Field threshold characteristics for LOCOS devices

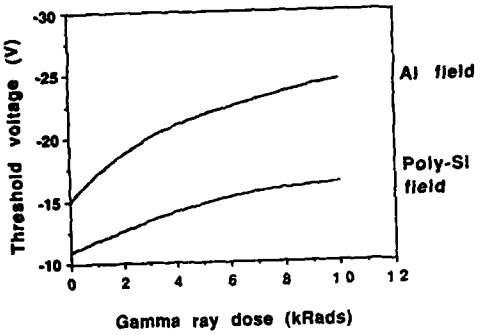


Figure 5a

Off-state leakage currents for nmos parasitic devices

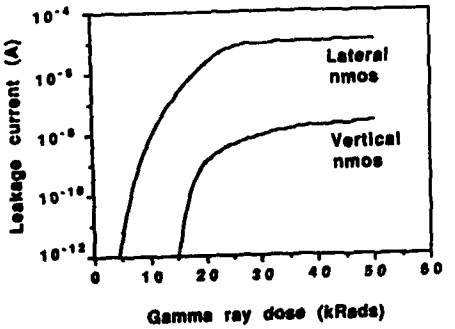
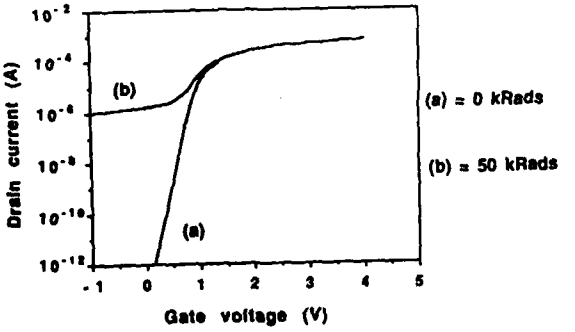


Figure 5b

Threshold characteristics of an nmos transistor as a consequence of irradiation



JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

THE INFLUENCE OF X-RAY DAMAGE ON ELECTRON-INDUCED INTERFACE DEGRADATION IN MOS CAPACITORS

U. SCHWALKE, E.P. JACOBS and B. BREITHAUPT*

Siemens AG, Corporate Research and Development, Microelectronics, D-8000 München 83, F.R.G.
**Sietec, Arbeitsgruppe Mikrostrukturtechnik, D-1000 Berlin 13, F.R.G.*

Abstract - The radiation response of MOS capacitors and their degradation resistance after annealing has been investigated. Compared to unexposed samples, irradiated and subsequently annealed MOS capacitors were found to be more prone to electron-induced interface degradation. The enhanced degradation correlates with the initial radiation damage and is lowest in $TaSi_2$ silicide gates. To which extent hydrogen contamination of the oxide or mechanical strain may account for the observed results will be discussed.

1 - INTRODUCTION

The realization of fine-line CMOS structures with dimensions < 0.5 µm will require advanced lithography and etching techniques. X-ray lithography has proven to be a promising technique for future submicron semiconductor fabrication /1/ and reactive ion etching is already in use for micron and submicron feature patterning. Since photon energies used in x-ray lithography and the energies of the intense radiation generated in plasma etching systems are well above the SiO_2 band gap, electron-hole pairs are generated in the oxide during irradiation. As a result of hole-trapping and interface-state generation, structural and chemical changes occur near the SiO_2/Si interface /2/ which will damage the oxide. The radiation response is known to be affected by hydrogen contamination /3/ introduced in the oxide during processing as well as mechanical strain /4/ induced by gate electrodes. Usually, radiation damage is removed by low temperature (300°C ≤ T ≤ 500°C) forming gas metallization anneals. However, little is known regarding the degradation behavior of annealed MOS devices. There is concern that the relaxation of the oxide damage is unstable, leading to enhanced device degradation during operation /5/. Especially in submicron devices with increased internal electrical fields and more severe hot-carrier problems this may become a serious limitation for device reliability.

In this work, the radiation response of MOS capacitors with different gate materials and gate oxide thicknesses as well as their degradation resistance after annealing have been examined.

2 - EXPERIMENTAL PROCEDURES

MOS capacitors used in this study were prepared by means of a standard CMOS process on 4-inch <100> Si wafers. Gate oxides of 11 to 45 nm were grown at 900°C in dry O_2+HCL ambient and as gate electrodes n^+-poly-Si, n^+-($TaSi_2$)-polycides and $TaSi_2$-silicides were used respectively.

All x-ray exposure experiments (E_{pk}≈1.6keV) were performed at the lithography beam line of the BESSY Synchrotron Facility in Berlin. Only specified areas of the wafers were irradiated so that several differently exposed as well as unexposed reference chips could be obtained on the same wafer. To be close to practical processing conditions, no bias was applied to the devices during

irradiation and the exposures were in the sensitivity range of the resists envisaged for x-ray lithography. The x-ray doses absorbed in the gate oxide with various gate electrodes were calculated by means of the XMAS program /6/. Compared to n^+-poly-Si gates, the total dose absorbed in SiO_2 was found to be reduced by ≈ 30% for $TaSi_2$ silicide and polycide gates. Annealing of the exposed wafers was either performed at 450°C in forming gas for 15 minutes or at 700°C in pure N_2 .

Electron-induced oxide degradation was accomplished by Fowler-Nordheim constant-current injection from the Si-substrate (positive gate bias = PGB) as well as from the gate electrode (negative gate bias = NGB). The interface state density (D_{it}) spectrum was calculated from the HF and quasi-static C(V) curves /7/ whereas flatband voltage shifts (ΔV_{FB}) were determined from the HF-C(V) curves.

3 - RESULTS AND DISCUSSION

3.1 - Radiation Response of MOS Capacitors

In Fig.1A radiation-induced negative flatband-voltage shifts as a function of the incident dose are shown for MOS capacitors with different gate electrodes but the same oxide thickness of t_{ox} = 25 nm. The V_{FB} shifts increase with dose and saturate at ≈ 0.3 J/cm² due to space charge buildup in the SiO_2 layer. For different gate materials large variations in the saturation ΔV_{FB} values were observed which are paralleled by the D_{it} spectra shown in Fig.1B. The radiation damage was lowest in $TaSi_2$ silicide gates which is mainly attributed to the corrective mechanical stress created by the silicide /8/. Consistent with the strained-bond model /4/ a corrective compressive strain in the SiO_2 induced by the $TaSi_2$ film would compensate for the intrinsic strain induced by oxidation. As a result of the increased Si-O bond strength at the interface, an increased radiation hardness (Fig.1) and an improved hot-carrier stability /9/ is observed. On the other hand, in n^+-polycide gates the mechanical stress of the silicide is expected to be efficiently buffered by the 300 nm poly-Si layer so that the radiation response should be comparable to n^+-poly-Si gates. However, due to the increased amount of hydrogen incorporated into the oxide during processing (900°C silicide formation anneal in forming gas) a largely increased radiation sensitivity is obtained, consistent with previous results /3/.

Since oxide thicknesses will be reduced in scaled CMOS devices, we have also studied the radiation response as a function of t_{ox}. For t_{ox} values of ≥ 20 nm, ΔV_{FB} is proportional to the oxide thickness, indicating a constant radiation-induced trapped oxide charge of ≈ 5.5 x 10^{-8} C/cm². However, as t_{ox}

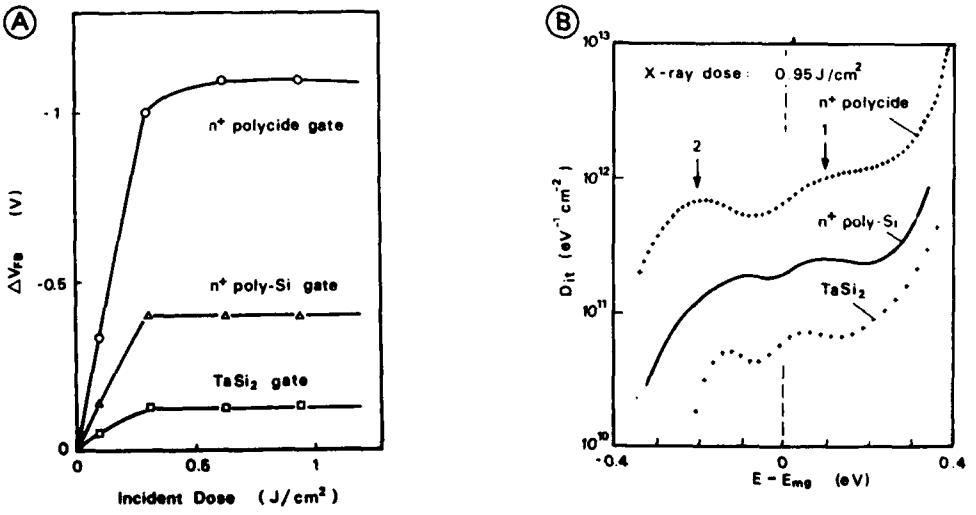


Fig.1: V_{FB} shifts (A) and D_{it} distributions (B) of MOS capacitors with different gate electrodes after x-ray exposure.

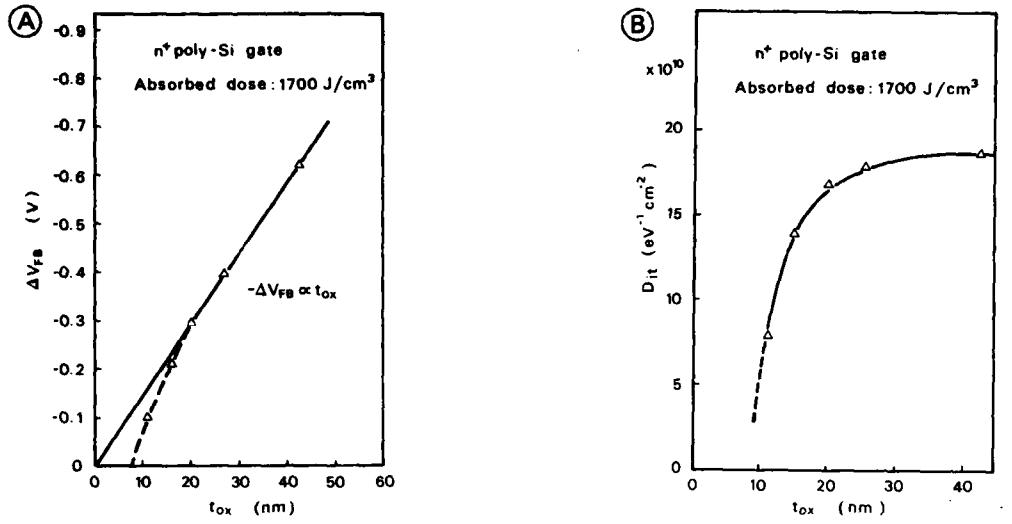


Fig.2: Dependence of ΔV_{FB} (A) and D_{it} (B) on oxide thickness.

is reduced below 20 nm, a departure from linear dependence is observed (Fig.2A), suggesting a reduced hole trapping in thin oxides. The beneficial effects of thin oxides with respect to radiation hardness are even more obvious in the D_{it} spectra shown in Fig.2B. Again, reduced interfacial stress present in thin oxides or different growth properties during initial oxidation may account for the increased radiation stability.

3.2 - Electron-Induced Interface Degradation of Annealed MOS Capacitors

After a brief 450°C forming gas anneal, a complete relaxation of the radiation-induced damage has occurred, as shown in Fig.3. However, annealing up to 700°C in H_2-free ambients did not result in a complete reduction of the

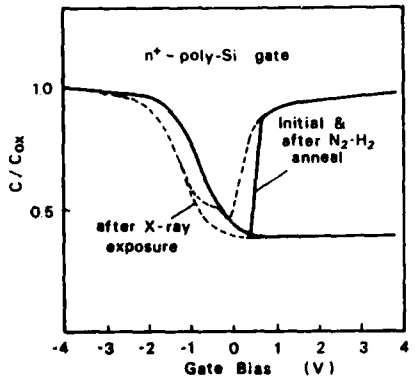


Fig.3: C(V) curves illustrating recovery from radiation-induced oxide damage after 450°C anneal in N_2-H_2.

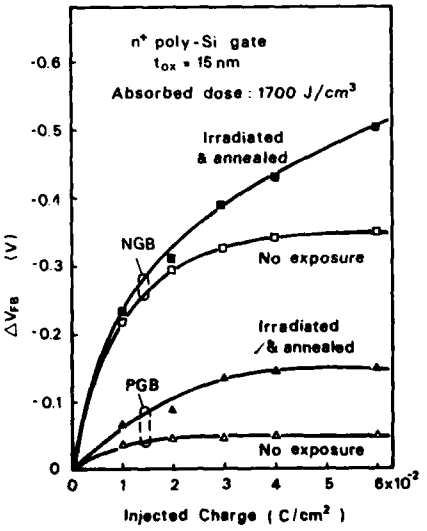


Fig.4: V_{FB} shifts as a function of injected charge.

radiation damage. This result suggests that H-atoms bonded at defect-sites in SiO_2 are responsible for the relaxation of the radiation-induced oxide damage.

During subsequent constant-current injection experiments the irradiated MOS capacitors were found to be more prone to interface degradation compared to the unexposed ones, as evident from the V_{FB} shifts (Fig.4) and the D_{it} spectra (Fig.5A). The enhanced degradation is explained by the reduced bond energy of the Si-H bonds ($\leq$ 3 eV) which are therefore more easily broken then Si-O bonds ($\approx$ 5 eV). Consistent with this interpretation, the susceptibility to electron-induced interface degradation was found to correlate with the initial radiation damage level and was lowest in $TaSi_2$ silicide gates (Fig.5B).

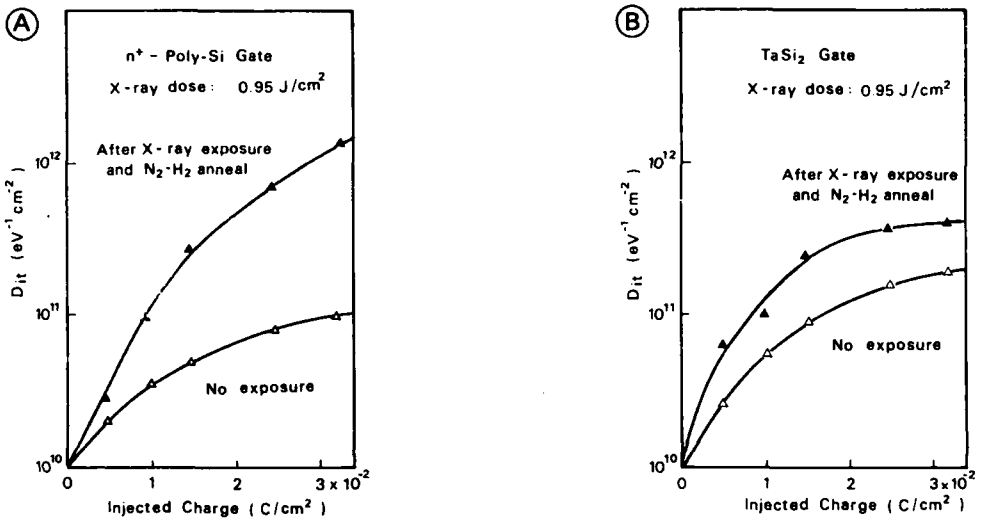


Fig.5: Electron-induced interface degradation in n^+-poly-Si (A) and $TaSi_2$ gates (B) with t_{ox} = 25 nm.

4 - CONCLUSION

We have shown that low temperature metallization anneals in forming gas will only conceal radiation-induced oxide damage. Compared to unexposed samples, an increased susceptibility to electron-induced interface degradation was observed which was found to correlate with the initial x-ray damage level. Nevertheless, radiation hardness and hence degradation resistance can be improved by the proper use of advanced gate materials and thinner oxides as well as avoiding H_2 contamination of the oxide during processing.

REFERENCES

/1/ A. Heuberger, in Microcircuit Engineering 1986, Ed. H.W. Lehmann North-Holland, Amsterdam.
/2/ E. F. da Silva,Y. Nishioka and T.P. Ma, Appl. Phys. Lett. 51, 270 (1987).
/3/ Y. Nishioka, E.F. da Silva and T.P. Ma, IEEE Trans. Nucl. Sci., NS-34 (1987) 1166.
/4/ V. Zekeriya and T.P. Ma, IEEE Trans. Nucl. Sci., NS-31 (1984) 1261.
/5/ F.C. Hsu, J. Hui and K.Y. Chiu, IEDM Techn. Dig. (1985) p.48.
/6/ H. Oertel, H. Betz and A. Heuberger, Microelectr. Eng. 3 (1985) 387
/7/ R. Castagné and A. Vapaille, Surface Sci. 28 (1971) 157.
/8/ F. Neppl and U. Schwabe, IEEE Trans. Electr. Devices, ED-29 (1982) 508.
/9/ A. Lill and M. Orlowski, ESSDERC '85 Europhys. Conf. Abstr., 9H (1985) 106.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

MOS-DEGRADATION IN INPUT AND OUTPUT STAGES OF VLSI-CMOS-CIRCUITS DUE TO ELECTROSTATIC DISCHARGE

X. GUGGENMOS

Siemens AG, Corporate Research and Development, Microelectronics, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

Abstract - MOS transistors have been used as sensors to study non-catastrophic effects of electrostatic discharges in input stages with protection circuits as well as output stages of VLSI circuits. Stress voltages far below the destructive level were found to cause both, severe threshold voltage shifts and transconductance degradation. As a result a reduction in circuit reliability is observed. To prevent degradation, selected and improved ESD protection circuits have to be used. It will be shown that the standard criterium for ESD-hardness needs to be extended in order to account for these requirements.

1 - INTRODUCTION

Micron and sub-micron CMOS-Devices are known to show an increased susceptibility to excessive voltages and currents caused by electrostatic discharge (ESD) or similar electrical overstress /1/. Consequently the ESD-hardness of the input and output stages has to be improved. The accepted characterization criterium for the ESD-hardness is given by the critical voltage V_{crit} of the ESD test model (e.g. Human Body Model - HBM, Machine Model - MM) at which significant leakage current or malfunction occurs /2/. This test will only consider catastrophic failures caused by single discharge pulses. Soft damage, such as threshold voltage and transconductance degradation in the CMOS devices will remain undetected /3/-/5/. However, this unrevealed ESD damage may be a serious limitation of device reliability. This issue is expected to become even more relevant in scaled down MOS devices with reduced gate oxide thicknesses and increased hot-carrier stress during operation.

In this work we have studied to which extent non-catastrophic effects of ESD-stress will cause degradation of the MOS devices of input and output stages. In CMOS input stages ESD protection circuits are integrated to limit the voltage at the susceptible MOS-gates. Their efficiency to prevent degradation has been investigated. To account for ESD effects on output stages, the driver performance of output transistors has been studied.

We will provide experimental evidence that ESD stress will cause significant device degradation and reduce device reliability. Protection circuits selected with a high ESD-hardness according to the standard tests were found to be insufficient in some cases.

2 - TEST STRUCTURES AND MEASUREMENT SETUP

For the electrical characterization of input stages special test structures consisting of an input protection circuit combined with a typical CMOS input inverter were used. The effects on output stages were studied at single driver transistors without additional protection circuits.

An especially developed ESD-tester enables a combination of a two pin ESD-stress and a full DC-characterization of the test inverter. In addition to the

standard characterization (leakage current, malfunction, etc.), MOS parameters of the individual transistors of the input inverter such as threshold voltage (V_{th}) and transconductance (g_m) are determined after each ESD-stress.

3 - RESULTS AND DISCUSSION

3.1 - MOS DEGRADATION DUE TO ELECTROSTATIC DISCHARGE

To determin the ESD-hardness of an input stage according to the standard test the voltage of the ESD-model is stepwise increased. Fig. 1 illustrates the V_{th} shift and the g_m degradation of the NMOS-transistor of the input inverter during a step stress test with the HBM. Note that only for stress voltages exceeding the critical voltage (V_{crit}) of -4600 V a significant leakage current occurs. However, much below this value starting at a voltage of about -2000 V, the MOS characteristics are largely degraded. Since the observed V_{th} shifts are negative in all cases, it is concluded that positive charge trapping in the SiO_2 has occured. The observed V_{th} shift and the changes in the transconductance indicate that voltage spikes pass the protection circuit and reach the gate which are not high enough to cause gate oxide rupture but will cause a current injection into the gate oxide.

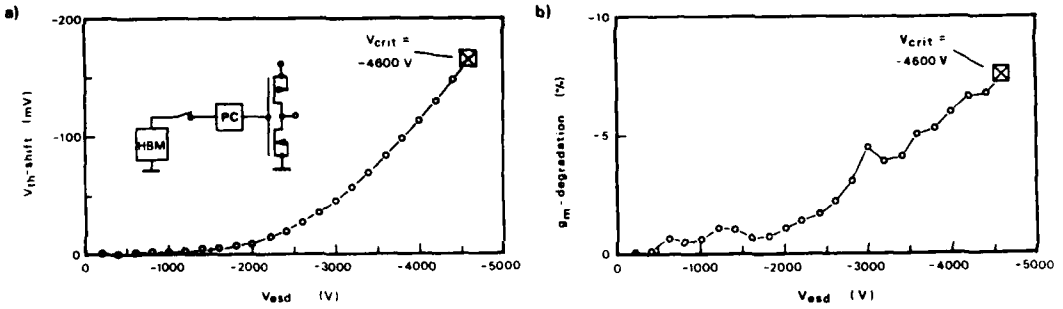


Fig. 1 Threshold voltage shift (a) and degradation of the transconductance (b) of the NMOS-transistor of an input inverter with protection circuit (PC) vs. stress voltage (Human Body Model 100 pF, 1.5 kOhm)

Protection circuits with different elements and layout variations have been compared. The protection circuit b) of Fig. 2 has a higher critical voltage V_{crit} of -3900 V compared to -2600 V of structure a), but already at a voltage of about -400 V a significant shift of the threshold voltage of the PMOS-transistor takes place. Up to a voltage of about -2300 V structure a) shows no significant degradation. Accordingly, structure a) would be the better choice. This result demonstrates that the standard ESD specification according V_{crit} can be misleading with respect to device degradation.

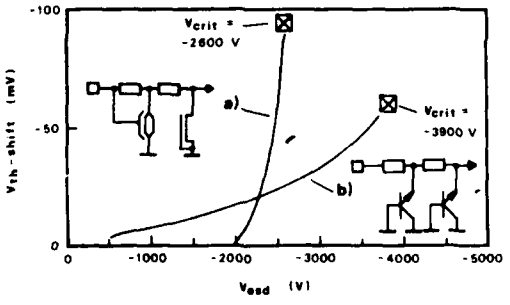


Fig. 2 Threshold voltage shift of the PMOS-transistor of an input inverter for two different protection circuits vs. stress voltage (Human Body Model 100 pF, 1.5 kOhm)

Next let us consider the influence of ESD stress on output stages without an additional protection circuit. Fig. 3 shows the resulting degradation of an output transitor. In this case the ESD stress voltage is applied between the drain and the source of the NMOS-transistor (gate floating). Compared to the degradation of input transistors, a similar degradation behaviour is observed. For this type of stress there are two possible mechanisms for the degradation: The high vertical electrical field between drain and gate may induce a transient gate oxide current (loading the floating gate) and/or the high lateral electrical field at the drain edge may generate hot-carriers. /3/

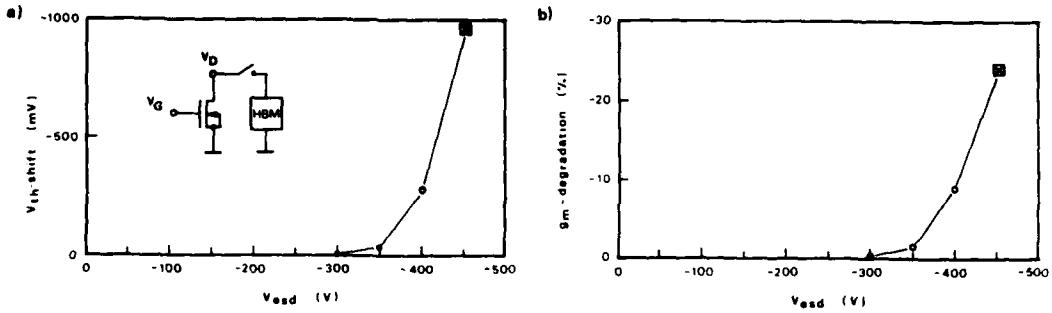


Fig. 3 Threshold voltage shift (a) and degradation of the transconductance (b) of the NMOS-transistor of an output inverter without protection circuit vs. stress voltage (Human Body Model 100 pF, 1.5 kOhm, V_D/V_s)

3.2 - EFFECTS ON CIRCUIT RELIABILITY

Now we will concentrate on ESD induced longterm effects. At first the influence of single ESD-events on the hot-carrier degradation behaviour of input transistors is studied. Fig. 6 shows the V_{th} shift and the g_m degradation of an NMOS-transistor due to hot-carrier stress (L_{eff}=1 µm, V_D=7 V, V_G=3.7 V). After 128 seconds a single ESD-stress (Machine Model, 50 pF/0 Ohm/-500 V) is performed. In this case, the hot carrier stress causes a negative V_{th} shift and the negative shift due to ESD stress accumulates. After a brief delay, the g_m degradation is shiftet to a higher level. In any case the additional ESD-stress reduces the device lifetime because a specified degradation level is exceeded earlier. These findings are paralleled by results reported recently in the case of output transistors /3/.

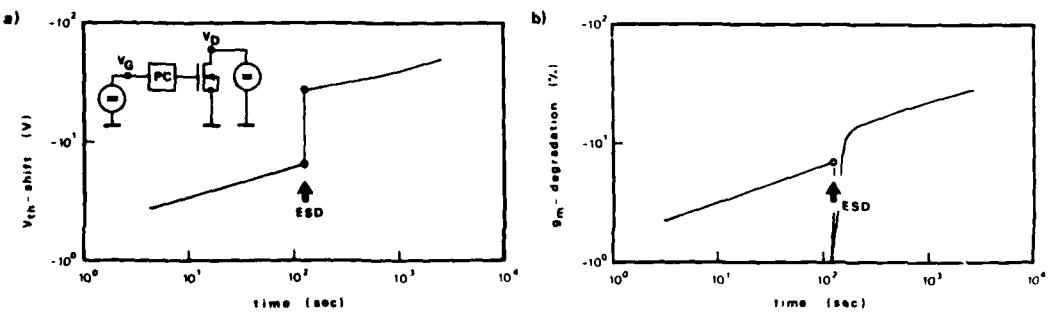


Fig. 4 Threshold voltage shift (a) and transconductance degradation (b) of an NMOS-transistor due to hot carrier stress (V_D=7 V, V_G=3.7 V) with an additional ESD-stress after 128 sec.

Next the effect of a multiple ESD-stress is discussed. Fig. 5 shows the V_{th} shift caused by an increasing number of discharges with an amplitude far below the critical voltage (compare Fig. 1). The degradation caused by each single ESD event accumulates resulting in a large total V_{th} shift. The number of discharges resulting in a functional failure can be estimated from this curve.

Both experiments show that the degradation caused by ESD-stress reduces the lifetime of the MOS devices due to cumulative effects. Even low level ESD events have to be avoided in order to maintain circuit reliability.

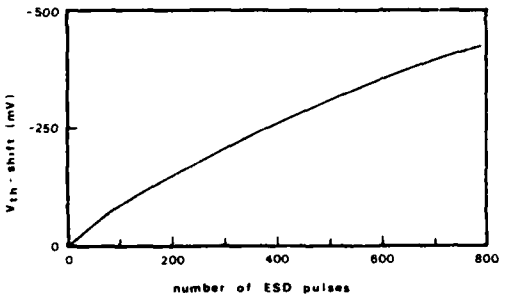


Fig. 5 Threshold voltage shift of the NMOS-transistor of an input inverter with protection circuit vs. number of discharges with a constant amplitude (Human Body Model 100 pF , -1500V, VIN/VSS)

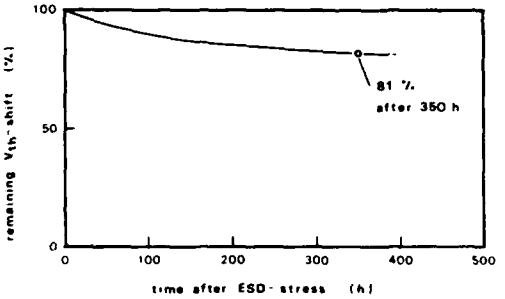


Fig. 6 Annealing behaviour of the threshold voltage shift vs. time after the initial ESD-stress at a elevated temperature of 140 °C

Further evidence of the importance of the results mentioned above with respect to longterm effects was found in the annealing behaviour of ESD induced degradations. Fig. 6 shows the annealing behaviour of the threshold voltage shift of an NMOS-transistor which was caused by a single ESD-event. Even after 350 hours at an elevated temperature of 140 °C about 81% of the initial shift of 430 mV remains. It this therefore reasonable to assume that the relaxation of ESD damage is not fast enough to relief from degradation effects under normal operating temperatures.

4 - CONCLUSIONS

It was shown that MOS devices of input/output stages degrade significantly at stress voltages far below the destructive level when unsuitable protection circuits are used. Since the observed V_{th} shifts and g_m degradation alter trigger voltages and switching times of the input/output stages, the performance is reduced. Multiple ESD events as well as single ESD events in combination with hot-carrier stress during operation was found to accumulate. As a result of this enhanced degradation, a reduction of circuit lifetime will occur.

As a consequence protection circuits have also to be selected and improved according to degradation effects. Standard characterization methods were found to be insufficient and in some cases, misleading. An extended characterization method has been developed. In test structures the MOS-transistors of input as well as output stages can be used efficiently as sensors for dangerous transient overvoltages. This method enables a proper selection and systematic improvement of protection circuits.

ACKNOLEDGEMENTS - The author wish to thank Dr. Neppl and Prof. Dr. Ruge for supporting this work, and Dr. Schwalke, Dr. Kerber and Dr. Mazuré for helpfull discussions.

REFERENCES:

/1/ Duvvury, C., et al., Int. Reliability Physics Symp., Proc., 24 (1986)199
/2/ MIL-STD 883, Method 3015
/3/ Aur, C., et al., Int. Rel. Physics Symp., 24 (1986)199
/4/ Holmes, G.C., Electr. Overstress Electrostatic Discharge Symp., (1985)27
/5/ Amerasekera, E.A., Cambell, D.S., Quality and Rel. Eng. Int., 2 (1986)107

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

MASSIVE HILLOCK GROWTH ON CATHODE SIDE OF TEST STRUCTURE DURING ELECTROMIGRATION EXPERIMENTS

B. BACCONNIER, G. LORMAND* and F. MARTIN

D. LETI/IRDI, CEA, CEN/G, 85X, F-38041 Grenoble Cedex, France
**Laboratoire d'Etude des Matériaux, Bât. 303, INSA, F-69621 Villeurbanne Cedex, France*

Résumé - Nous observons l'apparition d'excroissances sur la cathode des structures de test au cours d'essais d'électromigration. Ce phénomène est opposé à celui habituellement observé en présence de gradients thermiques élevés aux extrémités des lignes de test. Sur les films Al-Si-Ti déposés à 80°C, l'apparition des excroissances est accompagnée de celles de fortes dépressions du coté de l'anode. En l'absence d'un réseau continu de joints de grains, la quantité de matière traversant entièrement la ligne est négligeable. Il en résulte une divergence importante du flux de matière aux deux extrémités de cette ligne. Sur les films Al-Cu, l'apparition des excroissances et le déplacement de leur front vers l'anode est controlé par l'électromigration progressive du cuivre. La divergence du flux de matière est due, comme dans l'expérience des bandes en croix, à une différence entre les coefficients de diffusion intergranulaire des zones riches et pauvres en cuivre.

Abstract - We observed a massive hillock growth on the cathode side of test structures during electromigration experiments. This phenomenon is opposite to the one observed when high self heating thermal gradients are presents. On Al-Si-Ti films, a mass depletion occurs on the anode while hillocks grow on the cathode. The discontinuity of the grain boundary network is responsible for these effects, which are correlated to an enhanced lifetime. On Al-Cu films, hillocks appear at the negative pad - end contact segment junction and their front progagates towards the anode with a velocity proportionnal to current density. As in the cross stripe experiment, the phenomenon results from the difference between the grain boundaries diffusion coefficients in Cu depleted and Cu rich areas.

1 - INTRODUCTION

During electromigration experiments on thin Al films, a mass flow oriented towards the anode goes through test structures. Mass is primarily transported by grain boundary diffusion /1,2/ and mass flow presents local divergences at grain boundary triple points. However, more generalized flux divergences can be induced by differences in grain sizes or self heating thermal gradients. The latter effect leads to mass accumulation on the anode and mass depletion on the cathode, which has provided in the past an extensively used technique for measuring mass flow. Table I presents some of the experiments published on this technique. Not withstanding the fact that mass accumulation or depletion occurs in a quite heterogeneous way (hillocks or voids formation), the microstructure is assumed to be homogeneous in regard to stripe dimensions. This hypothesis does not seem to be questionnable when grain size is much smaller that line width.

We observed a massive hillock growth on the cathode side of test structure during the test of Al-Si-Ti and Al-Cu metallizations. We explain how this particular behavior is related to microstructural inhomogeneities.

TABLE I. Observation of mass accumulation or depletion.

Metal	Method	Line width (µm)	Grain size (µm)	Accumulation Side	Depletion Side	Reference
Al	T.E.M.	25	0.5 - 5	anode	cathode	1
Al	Electron beam	45	1	anode	cathode	3
Al-Cu-Cr	S.E.M.	7	1800	anode	cathode	4
Al-Mg	E.P.M.A.	7	----	anode	cathode	5
Al	S.E.M.	4	6	anode	cathode	6
Al-Si-Ti	S.E.M.	1.4	1	cathode	anode	present work
Al-Cu	S.E.M.	4	1.5	cathode	anode	present work

2 - EXPERIMENTS

Al-Si-Ti (1 wt.% Si - 1.8 wt.% Ti) and Al-Cu (0.5 wt.% Cu) films were sputtered on respectively Ti-W and Ti-N diffusion barriers both on 1000 Å of thermally grown SiO2. Film thickness was 0.6 µm for Al-Si-Ti and 0.7 µm for Al-Cu films. Al-Cu films were deposited with 100V bias voltage. Table II presents elaboration and electromigration test conditions. Failure criterion was an increase of 10% of the electrical resistance. Al-Si-Ti films deposited at 80 and 200°C were studied in T.E.M.. Cu profile on both tested and untested Al-Cu lines was measured by electron probe micro-analysis (E.P.M.A.). All lines were observed in S.E.M. after testing.

TABLE I I. Elaboration and electromigration test conditions .

Metal	Deposition temperature (°C)	Annealed 30mn (450°C)	Line width (µm)	Current density (GA/m2)	Temperature (°C)
Al -Si -Ti	80	yes	1.4 , 4	15	200 , 170
Al -Si -Ti	120, 160, 200	yes	1.4 , 4	15	200
Al-Cu	Room Temp	yes	4	10	200
Al-Cu	Room Temp	no	4	10	200

3 - RESULTS

We observe hillock growth and mass depletion respectively on the cathode and on the anode of Al-Si-Ti 1.4 µm lines deposited at 80°C (Fig. 1). When the deposition temperature increases, both effects disappear and the lifetime of 1.4 µm lines decreases while that of 4 µm lines increases (Fig. 2). Grain size is ranging from 0.5 to 1.5 µm both for 80°C and 200°C deposited films. Grain boudaries of 80°C deposited films look more like cell walls than like well defined grain boundaries. Mass depletions observed on the anode have lateral dimensions of about the grain size. We have noted that hillock growth can occur far from the end-contact segment and test line junction (farther than 0.1 mm) and is more heterogeneous that anode depletion. Hillocks volume, evaluated by S.E.M. observations, corresponds to a mass flow of 0.35 µm3 / s for 200°C and 0.15 µm3 / s for 170°C testing.

The behaviors of annealed or unannealed Al-Cu films are roughly the same : hillocks first appear at the negative pad and end-contact segment junction (Fig. 3) then hillock front propagates towards the anode with a velocity proportionnal to current density (Fig. 4). We detect no change in conductor shape on the anode. Cu profiles indicate a Cu depletion in the areas where hillocks are present. Figure 5 show the typical variations of electrical resistance of unannealed lines during test. We notice, after a an exponential type decrease in the first hours, a steep decrease of resistance occuring approximatively at the same time for all films. Failure occurs randomly during this decrease.

4 - DISCUSSION

Effects induced by electromigration can occur when mass flow divergence is not equal to zero everywhere in the conductor. However, it often seems necessary to define mass flux in a given section of the test structure in order to calculate diffusion coefficients from the observed effects /1/. This flux can be expressed as the sum of the fluxes in each grain boundary of the considered section and assuming in first approximation that grain boundaries are the same may be written :

$$J=Z^*.e.\rho.j \,/\, (kT).\delta.D0.\exp(-Q/kT).n.t \qquad (1)$$

where J is the mass flux, $Z^*.e$ the effective charge, ρ the resistivity at the temperature T, j the current density, n the number of grain boundaries crossing the section, δ the grain boundary width, t the film thickness, D0 and Q define the grain boundary diffusion coefficient D, and k and T have their usual meaning. We are going to see now how variations of n and D on our test structures allow to explain the observed effects.

Al-Si-Ti films deposited at 80°C have a grain size close to the width of 1.4 µm lines. Hence, there is always a discontinuity of the grain boundary network on the test-line not far from the end-contact segments, which means locally n=0. In this place, mass flux divergence is strictly proportionnal to mass flux, and leads to strong accumulations or depletions. Mean time to failure of these lines has been found higher than that of larger lines. This result can not be explained simply by a grain size dependence of electromigration resistance as reported for others materials /7,8/ : when deposition temperature increases, 4 µm lines have a larger lifetime than 1.4 µm lines, the contrary being obtained on

80°C deposited lines with the same grain size. The cell wall aspect of grain boundaries in the 80°C deposited films may explain a more homogeneous diffusion coefficient than well defined grain boundaries and thus divergences at triple points of these structures are less important. The shorter current stress duration (lifetime being shorter) associated with the high temperature deposited films is certainly the reason why we do not see any hillocks on the cathode of these films. The variation of the extruded volume with temperature gives an activation energy of 0.5eV (± 0.2eV) between 170°C and 200°C. Eq. 1 applied to the end contact segment gives a value of 1E-16 cm3 / s for dD at 200°C, which is close to the value of pure aluminium /5/. The fact that hillocks can grow far from the flux divergence section would require further considerations on stresses and defects concentration to be discussed.

It would be interesting to enlighten the reasons why such a hillock growth has not been reported for others alloys. A competition between thermal gradients and microstructure effects is an hypothesis, because our test structures tend to lower the thermal gradients more than structures in which line is in direct contact with pad. This hypothesis becomes questionnable when grain size is much lower than pad size (see Table I, /6/) because even if thermal gradients are high, bulk diffusion is always much smaller than grain boundary diffusion at usual test temperatures. However, the case of passivation cracking localized on the cathode of test structure during electromigration tests on layered structures has been reported /9/. This unexplained behavior was related to an increase of lifetime.

The apparition of hillocks on the cathode side of test structure is related to Cu depletion. According to the numerous studies of Cu influence on Al grain boundary diffusivity /10,11,12/, we may state that grain boundaries of Cu depleted areas present a much higher diffusion coefficient than those of non affected areas. Hence a higher mass flow in the depleted areas and a mass accumulation (hillock growth) arises at the edge of the depleted area. This explanation has already been given by Ho and Howard /11/ to account for the simultaneous apparition of hillocks and voids during the cross stripe experiment. When the edge of depletion area travels along the test line, an additionnal aluminium volume is pushed into the line, leading to the observed steep decrease of electrical resistance. It is difficult to determine quantitavely whether hillocks keep growing or not after areas have been completely depleted. However, the size of hillocks is relatively constant on the end contact segment. Futhermore, the steep decrease of resistance ends when hillocks have invaded the whole test line before failure. Threfore we conclude that hillock growth results mainly from the difference in diffusion coefficients.

ACKNOWLEDGMENTS

The authors would like to thank A. M. Papon for T.E.M. observations, A. Ermollieff et S. Marthon for E.P.M.A. measurements, M. Papapietro, J. P. Joly and C. Arena for encouraging discussions.

REFERENCES

/1/ I. A. Blech and E.S. Meieran, "Electromigration in thin Al films", J. Appl. Phys. 40 (1969) 487.

/2/ R. Rosenberg and L. Berenbaum, "Resistance monitoring and effects of nonadhesion during electromigration in aluminium films", Appl. Phys. Lett. 12 (1968) 201.

/3/ J.Weise, "Quantitative measurements of the mass distribution in thin films during electromigration experiments", Thin Solid Films 13 (1972) 169.

/4/ A. Gangulee and F.M. d'Heurle, "Anomalous large grains in alloyed aluminium thin films. II. Electromigration and diffusion in thin films with very large grains", Thin Solid Films 16 (1973) 227.

/5/ A. Gangulee and F.M. d'Heurle, "Mass transport during electromigration in Aluminium-Magnesium Thin films", Thin Solid Films 25 (1975) 317.

/6/ J. M. Pierce and M. E. Thomas, "Electromigration in aluminium conductors which are chains of single crystals", Appl. Phys. Lett. 39 (1981) 165.

/7/ S. Vaidya, T. T. Sheng, and A. K. Sinha, "Linewidth dependence of electromigration in evaporated Al-0.5%Cu", Appl. Phys. Lett. 36 (1980) 464.

/8/ H. U. Schreiber, "Electromigration mechanisms in aluminium lines", Solid-St. Electron. 28 (1985) 1153.

/9/ R. E. Jones, Jr. and L. D. Smith, "Contact spiking and electromigration passivation cracking observed for titanium layered aluminium metallization" , Proceedings VLSI Multilevel Interconnections Conf. (1985) 194.

/10/ P. S. Ho and J. K. Howard, "Nonlinear propagation of void front during electromigration in alloy films", Appl. Phys. Lett. 27 (1971) 261.

/11/ P. S. Ho and J. K. Howard, "Grain-boundary solute electromigration in polycrystalline films", J. Appl. Phys. 45 (1974) 3229.

/12/ F. M. d'Heurle, N. G. Ainslie, A. Gangulee, M. C. Shine, "Activation energy for electromigration failure in Aluminium films containing Copper", J. Vac. Sci. Technol. 9 (1971) 289.

Fig. 1. S.E.M. micrograph of a typical 80°C deposited Al-Si-Ti (w=1.4 µm) tested structure (J=15 GA/m2, T=200°C, t=200 h) : mass accumulation occured on the cathode and mass depletion occured on the cathode. Line lenght is 300 µm.

Fig.3. S.E.M. micrograph of a 4 µm Al-Cu tested line. Hillocks first appear at the negative pad / end contact segment junction and hillocks front propagates towards the anode.

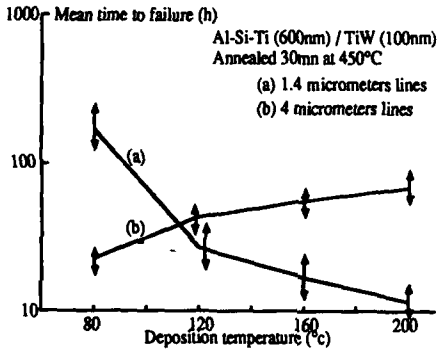


Fig. 2. M.T.F. of Al-Si-Ti 1.4 and 4 µm lines versus deposition temperature. Failure criterion is an increase of 10% of the electrical resistance.

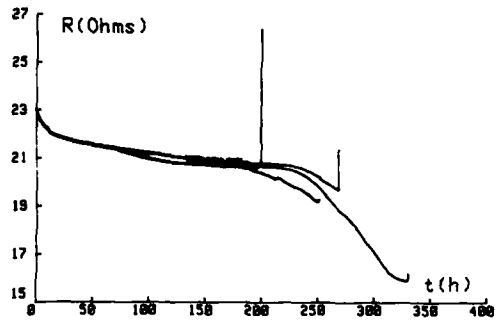


Fig. 5. Typical variation of unannealed Al-Cu lines resistance during test (w=4 µm, J=10 GA/m2, T=200 °C)

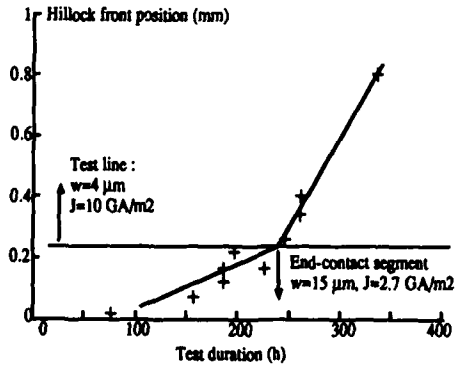


Fig. 4. Hillock front position versus test duration for unannealed Al-Cu lines. Slope is proportionnal to current density.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CHARACTERIZATION OF FIELD-OXIDE-TRANSISTOR-INSTABILITIES CAUSED BY SOG-PLANARIZATION

B. VOLLMER, G. RÖSKA and J. WINNERL

Siemens AG, Central Research and Development, Microelectronics, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

Abstract

A spin-on-glass planarization process in multilevel metallization can cause instabilities of metal-2-gate field oxid transistors. The threshold voltage shift of a p-channel transistor is reduced for negative gate bias and it can be accelerated by gate bias and temperature. The threshold voltage shift is caused by a thermally activated and field enhanced diffusion process. The planarization process leaves SOG inclusions, but the instabilities are also observed in strcutres where no inclusions are present.

1 - INTRODUCTION

In view of the decreasing feature size of CMOS devices, the metallization will become the dominant factor in determining chip size and speed performance, two or more metallization layers will be necessary. Planarization techniques have to be used to smooth topography and thus to guarantee a sufficient step coverage of the second metal layer. Planarization processes which are based on an etch-back-method using spin-on-glass (SOG) as a sacrificial layer meet the geometrical requirements of subµ technologies. This auxiliary layer is not totally removed but remains in narrow gaps between first metal lines /1/. The aim of this paper is to investigate instabilities of the field isolation with metal-2 gate under bias-temperature stress caused by a SOG planarization process.

2 - EXPERIMENTAL

The samples were fabricated in a 1µm CMOS technology with minimum fieldoxide width of 1.8µm and thickness of 600nm. The first metal is isolated from gate level by a sandwich of 150nm undoped (TEOS) and 650nm boron/phosphorus doped (BPSG) oxide. The planarization process starts with a PECVD oxide deposition. Undoped SOG is then spun-on and cured to remove the solvents. The following non selective etch back process provides a smooth and semi-planar surface. The SOG remains only in contact holes, at the edges of vertical side walls and within narrow gaps between first metal lines. The SOG inclusions are encapsulated by a further PECVD-oxide layer. The total intermetal dielectric thickness is 1µm.

Metal-2 lines over field oxide form parasitic field oxide transistors. The metal-2 gate can additionally be crossed by metal-1 lines. These two cases are represented by the test structures shown in Fig. 1a. The gate oxides of the metal-2 field oxide transistors consist of field oxide, TEOS, BPSG and PECVD oxide with a total thickness of about 2.4µm. The structures with metal-1 bars contain SOG inclusions, whereas without bars the sacrificial SOG layer is totally etched back. The SOG planarization process involves the risk of remaining impurities like mobile ions or polarisable structures. P-channel transistors with metal-2 gate were used to investigate field isolation instabilities.

The transistors were stressed on wafer level with an increased gate bias (40-100V) at temperatures from 70 to 140°C. They are always characterized at 90°C as a typical operating temperature, independent of the stress temperature.

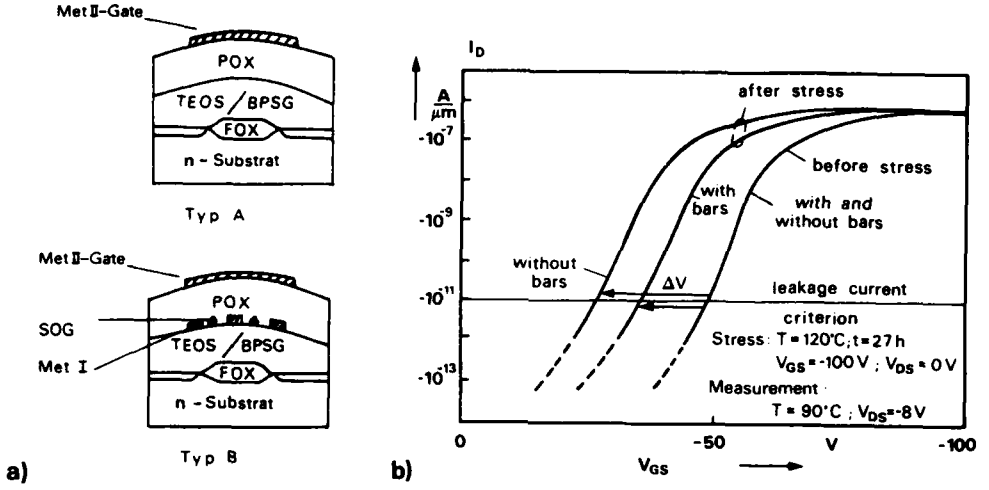


Fig. 1 a - Schematical cross section of a p-channel metal - 2 gate field oxide-transistor with and without metal-1 fingers.

Fig. 1 b - Subthreshold characteristic for 1.8µm field oxide transistor with metal-2-gate (with and without metal-1 bias before and after -100V gate bias stress at 120°C). The turn-on voltage shift V is defined as gate bias shift at 10pA/µm width leakage current.

3 - RESULTS AND DISCUSSION

Fig. 1b shows typical subthreshold characteristics for metal-2 p-channel transistors with and without metal-1 bars stressed with -100V gate bias at 120°C. The characteristics after stress are shifted to lower gate voltages. This shift is due to a threshold voltage reduction, while the substhreshold slope is nearly unchanged. The turn-on voltage V_{T_0} is defined as the gate voltage at 10pA/µm leakage current.

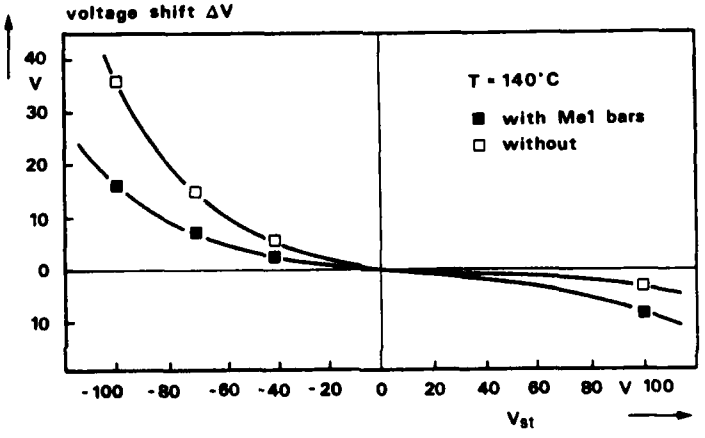


Fig. 2 - Threshold voltage shift for positive and negative gate bias.

The turn-on voltage shift ΔV after a defined stresstime increases with increasing gate bias (Fig. 2). Both structures show an enhanced degradation for negative bias. A negative gate voltage is the realistic bias condition for p-channel transistors.

Within the examined time period of 20h and stress voltages up to -100V the time dependence of the V_{T_0} shift follows a square root law (Fig. 3). A deviation from this law is only observed for high V_{T_0} shifts ($\geq$ 20V). At a stress temperature of 100°C, not only the structure with SOG inclusions shows a degradation of the turn-on voltage but also the structure without bars, V without bars is even higher at this temperature. The time dependence indicates that the voltage shift is caused by a diffusion process rather than by polarisation which should show a logarithmic time dependence. The voltage dependent time constant shows that the diffusion process can be enhanced by the gate field. This allows accelerated degradation experiments at increased gate bias.

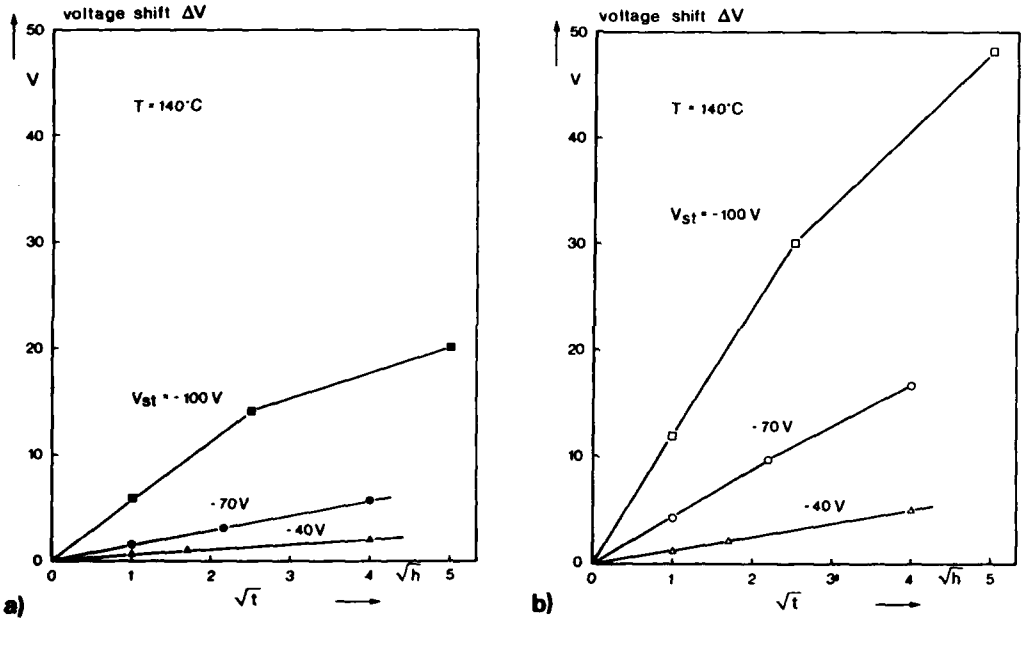


Fig. 3 - Time dependence of the threshold voltage shift of a field oxide transistor with (a) and without (b) metal-1-fingers.

Plotting the logarithm of ΔV versus the reciprocal temperature a linear dependence results for both structures. This exponential reverse temperature and the square root time dependences indicate a thermally activated diffusion process. The activation energy of samples without fingers stressed at -100V is approximately 0.6eV and higher than of samples with fingers (0.4eV). Therefore the degradation of samples with fingers dominates at low temperatures. In both cases the activation energies increase at reduced gate bias. With a voltage dependent activation energy the ΔV can be described by the following relationship.

$$\Delta V \propto t^{n} \exp\left(-\frac{E_a\,(V_{st})}{kT}\right) \qquad (1)$$

With the gate bias during stress V_{st}. In Fig. 4b the voltage shift is plotted versus gate bias for 140°C. The observed exponential dependence has a

comparable slope for structures with and without bars. The activation energy of (1) can be expressed by

$$E_a = {}^{0}E_a + {}^{F}E_a (V_{st}, T) \qquad (2)$$

as sum of a zero bias activation energy ${}^{0}E_a = E_a (V_{st} = 0)$ and a field enhancement part ${}^{F}E_a (V_{st}, T)$.

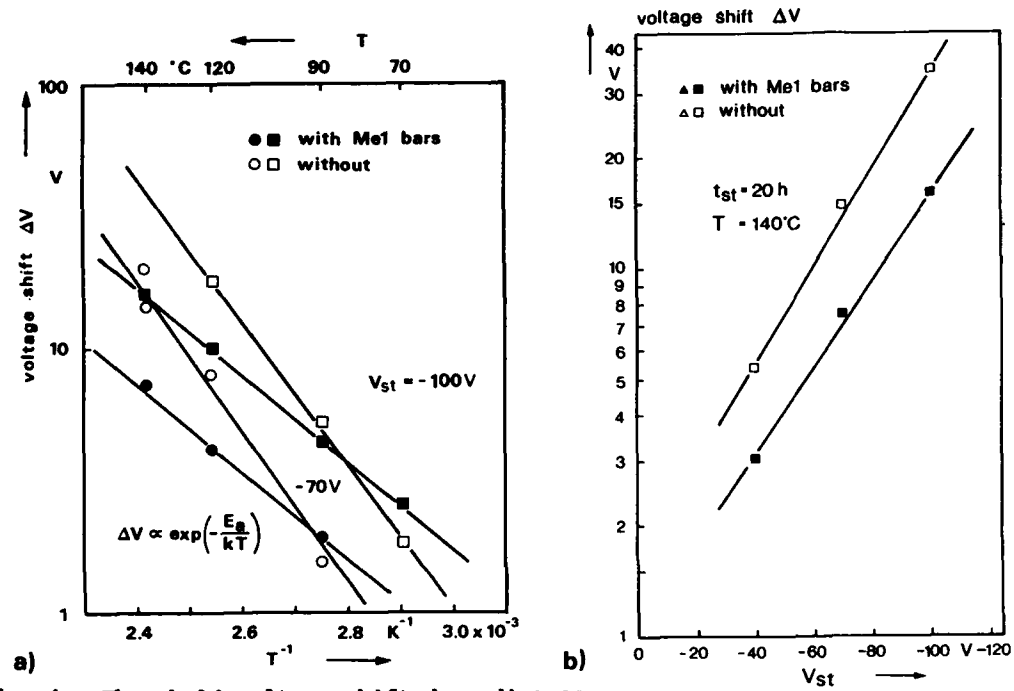


Fig. 4 - Threshold voltage shift depending on
(a) stress temperature
(b) gate bias

From literature activation energies are known for example for a Na^+-diffusion in SiO_2 of about 1eV and for hydrogen ions (protons) in SiO_2 of 0,36eV /2/. The above extraction values are within them, they are lower than the values for Na^+-diffusion in SiO_2. The activation energy with bars is about 0.4eV and would fit to a hydrogene-diffusion. The activation energies depend on the ion concentration and on the type of oxide (pure or doped SiO_2).

4 - CONCLUSIONS

For the described SOG planarisation process bias/temperature stress leads to degradation of metal-2 field oxide transistor isolation. The threshold voltage shift is caused by a thermally activated and gate field enhanced diffusion process. The instabilities are also observed in structures without metal-1 fingers and therefore the SOG inclusions themselves can not only be responsible for the instabilities.

This work has been supported by the Federal Department of Research and Technology of the Federal Republic of Germany. The authors alone are responsible for the contents.

REFERENCE

/1/ P. Elkins, K. Reinhardt, R. Tang, Proc. of V MIC Conf., p. 100 (1986).
/2/ Eldrige, Kerr, J. Electrochem. Soc., p. 986 (1971).

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

STATISTICAL ANALYSIS OF IMPLANT ANGLES EFFECTS ON ASYMMETRICAL NMOSFETs CHARACTERISTICS AND RELIABILITY

P. DARS, T. TERNISIEN d'OUVILLE, H. MINGAM and G. MERCKEL

CNET-CNS, BP 98, F-38243 Meylan Cedex, France

Résumé: - L'analyse statistique de la dissymétrie des caractéristiques électriques des transistors NMOS LDD traduit l'influence des angles d'implantation sur la variation du non recouvrement grille-drain (ou grille-source) observé sur les dispositifs réalisés sur une même plaque et sur les différentes plaques d'un même lot de fabrication. La conséquence de cette dispersion sur le vieillissement des structures montre l'importance du suivi de ce paramètre pour la fiabilité des transistors et la prudence nécessaire pour l'interprétation des résultats de stress électrique effectué sur les transistors.

Abstract: - Statistical analysis of asymmetry in LDD NMOSFETs electrical characteristics shows the influence of implantation angles on non-overlap variation observed on devices realized on a 100 mm wafer and within the wafers of a batch . The study of the consequence of this dispersion on the aging behaviour illustrates the importance of this parameter for reliability and the necessity to take it in account for accurate analysis of stress results .

INTRODUCTION:

The development of short channel NMOSFETs needed the design of new structures in order to decrease the peak value of the Electric field near the Drain edge. Several ways have been investigated (eg.LDD ,DDD architecture) which have given good results in reducing hot carriers degradation. On the other hand , those devices are known to be sensitive to process parameters like implantation angles (Tilt and electrostatic scanning), gate etching profiles ... , which have a strong influence on the asymmetry of NMOS devices , depending on the position of the device with respect to the ion beam [1,2] . The asymmetry of transistors (due to the shadowing effect of the gate edge on lightly doped implants) is shown to play an important role on the aging behaviour of the devices.

EXPERIMENTAL

In order to analyse the consequences of tilt and electrostatic scanning angles on devices , we have designed a test structure including two perpendicular typical NMOSFETs (50x1 μm^2) . LDD devices are fabricated through a 1 µm N-well CMOS process using a sidewall spacer technique,with a gate oxide thickness of 25 nm . The lightly doped junction implant is done through an arsenic implant dose of 10^{13} cm^{-2} , with 7° off-axis. Initial electrical test measurements are made on an automatic test line in order to analyse the differences in bulk and drain currents at a given voltage condition (Vd=7 V, Vg=3.5 V) for both perpendicular structures in the forward and reverse mode (inverting the role of drain and source). This work is carried out for all the wafers (Φ =100mm) in a batch , in several positions per wafer (20 or 60 depending on the process) (see fig.1)

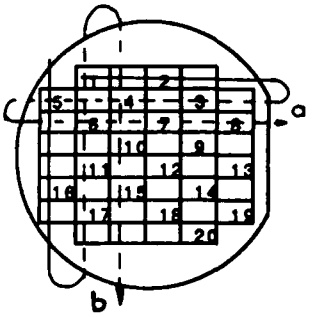


fig.1 Probing pattern
a. direction parallel to prim.flat ,b. direction perpendicular to primary flat

Electrical stresses are performed on those devices, in order to show the consequences of the asymmetrical behaviour of devices on the reliability, along a diameter of a 100 mm wafer.
Transistors are stressed, at room temperature, for 10 hours. The stress conditions are : Vd= 7 Volts, Vg= 3.5 Volts

RESULTS

1. INFLUENCE OF ELECTROSTATIC SCANNING

We analysed the differences in bulk or drain current for the two perpendicular devices with the test made in a given order (from chip 1 to 20). We defined the difference in drain currents by

$$Diffid = 2 \times \frac{(Idf-Idr)}{(Idf+Idr)}$$ Idf = Id in forward mode (Vd=7V , Vg=3.5 V)
Idr = Id reverse mode

The same work was made for bulk currents and gave similar results.

To illustrate the influence of the two electrostatic scanning angles , we plot the differences in two directions (see fig.1)

a.Along the direction parallel to the primary flat (see fig.2a) ,the gradient is evident for devices with the channel parallel to the flat, with a difference between the upper and the lower device of about 6 % .

b. Along the direction perpendicular to the primary flat (see fig.2b.) , the gradient is significant only for the devices with the channel perpendicular to the flat . The dispersion between the upper and the lower device is about 4 %.

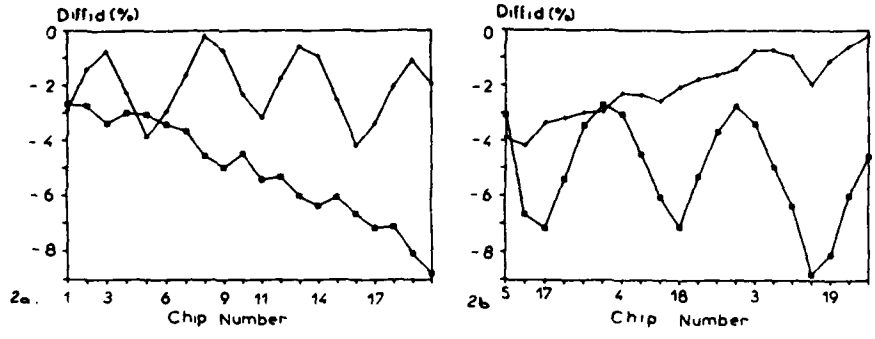


fig 2. Gradient of drain current asymmetry along the direction parallel (2a.) and perpendicular (2b.) to the primary flat (◘ for channel parallel to primary flat , + for channel perpendicular to primary flat)

The two different behaviours are explained by the influences of the electrostatic scanning angles of the implantor on a whole wafer .The differences in the two directions are due to the different scanning angles along the two directions in our implantor . It means that , even with a 0 off-axis implant angle, we find some devices with a non-overlap drain structure .

2-INFLUENCE OF THE TILT ANGLE

For avoiding the channeling effect during the implant , the wafers are implanted with a tilt angle of 7 degrees . That tilt angle gives different results on wafers depending on the position of the wafer during the operation (see fig 3). If we plot the differences in currents for the two perpendicular directions (see fig 4a.) , the points are distributed along a circle . That circle is due to the fact the tilt angle has a complementary effect on the perpendicular transistors . The values measured on a whole wafer (illustrated by the same number) are centered around a median position depending on the random position of the wafer during the low dose implant . If the analysis is done in only one direction ,it is possible to locate symmetrical devices that give the best results in stress measurements. However ,this case means the worst case for perpendicular devices .

The diameter of the circle shows the process sensitivity to the tilt angle.(see fig 4a. and fig 4b.)

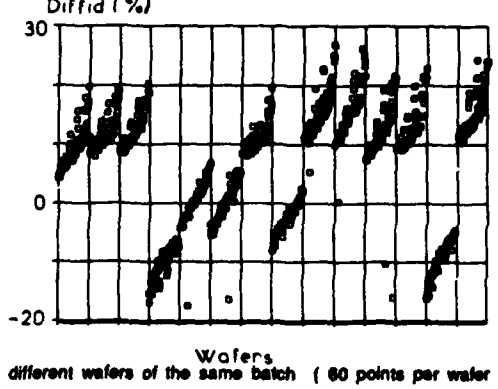


fig3. Drain current asymmetry on different wafers of the same batch (60 points per wafer)

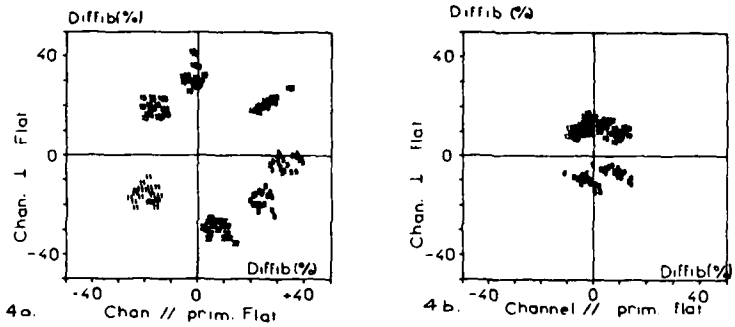


fig4. *Influence of the technology on bulk current asymmetry* a. Poly spacer technology b. Oxide spacer process

3-CONSEQUENCES ON DEVICES RELIABILITY

Asymmetrical devices are known to be very sensitive to electrical stresses performed in the non-overlap region [3] . In order to show the consequence of the dispersion of electrical characteristics over a wafer, we analysed one with a central value of Diffid =0. In that case, transistors structures change from a drain overlap to a non-overlap situation along a diameter . The results of the stress show only a slight degradation of the transconductance ($dGm/_{Gmo}$< 5 % *for 10 hours of* stress) until we reach the non-overlap condition, which gives the worst results (*see fig 6). A factor of 4 in Gm* degradation is observed between the two extreme cases, after 10 hours of stress .

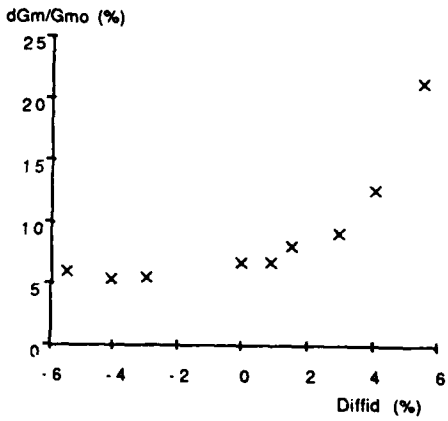


fig.6 Influence of asymmetry on Gm degradation (after 10 hours of stress)

CONCLUSION

Such an analysis ,with only a few complementary measurements , gives an accurate idea of the process sensitivity to implant angles and provides a predictive tool for aging behaviour ofadvanced CMOS structures .

AKNOWLEDGEMENTS

We wish to thank people from CNET pilot line for providing samples and helpful discussions, A. MAITRE and C.DENAT for technical support .

REFERENCES

[1] CHAN T.Y., WU A.T., HU C.H., RAZOUK R.R., IEEE Electron Device Lett., EDL-7, (1986), 16.
[2] PFIESTER J.R., BAKER F.K., IEDM Techn. Dig., (1987), 51.
[3] DOYLE B.S., BOURCERIE M., LECLAIRE P., BOUDOU A., DARS P., Electronics Lett. (1987), Vol 23, n°25,1356.

Session 3C Room C : Optoelectronic devices

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

PHOTODIODE FOR COHERENT DETECTION : MODELING AND EXPERIMENTAL RESULTS

J.E. VIALLET, S. MOTTET, L. LE FJEROU and C. BOISROBERT

Centre National d'Etudes des Télécommunications, F-22300 Lannion, France

Résumé - En détection cohérente, la puissance optique de l'oscillateur local peut conduire le photodétecteur en régime de forte injection. Des dégradations de performances du dispositif peuvent alors être constatées à partir de simulations numériques et de mesures.

Abstract - In coherent detection, the optical power of the local oscillator can lead the photodetector in the high injection regime. Degradation of the device performance can then be observed from numerical simulations and experiments.

1 - INTRODUCTION

Coherent detection can subtantially improves optical communication system performances [1] . Wavelength stabilized semiconductor lasers provide for the necessary local oscillator. Optical power of the local oscillator greater than the radiant power corresponding to the optical signal, yields better signal to noise ratio.

Photodetectors design is a compromise between different parameters such as sensitivity, response time, capacitance, gain, noise. The intrinsic material width of a PIN photodiode must be such that sufficient light is absorbed (sensitivity), that the free carriers thus generated are quickly collected (response time), driven at their top velocity by an uniform electric field, and so that the junction capacitance is low to limit RC time constant.
At high level of optical illumination, free carrier densities can be such that the electric field in PIN photodiodes is no longer uniform and exibit low and high values areas. This behaviour is expected to affect the response time of the device.

The goal of this study is to determine the influence of the optical power on the photodiode performances, in coherent detection, when a high radiant optical power shines on the photodetector, mixed along with the transmitted low level signal from the fiber output. The operation of Si, Ge, GaInAs/InP photodiodes under low and high optical power are described in terms of sensitivity, linearity, response time and diode capacitance as predicted by numerical simulation and confirmed by measurements.

2 - MODELING AND NUMERICAL METHODS

Unidimensional numerical simulation of PIN photodiodes are performed, using finite difference methods, solving the following set of equation that describes the behaviour of semiconductor devices. This set includes Poisson equation, electron and hole continuity equations and current formulation. Thermal generation recombination term as well as optical generation are taken into account [1].

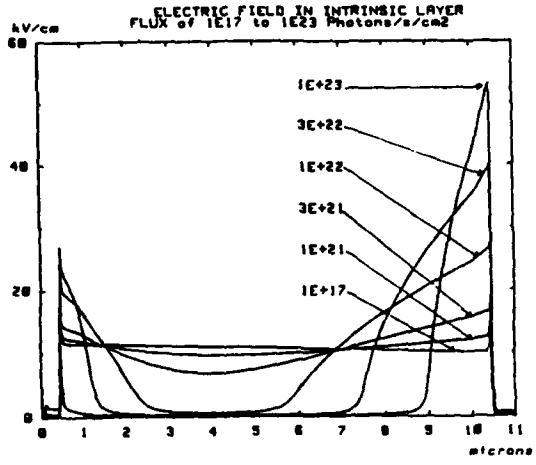


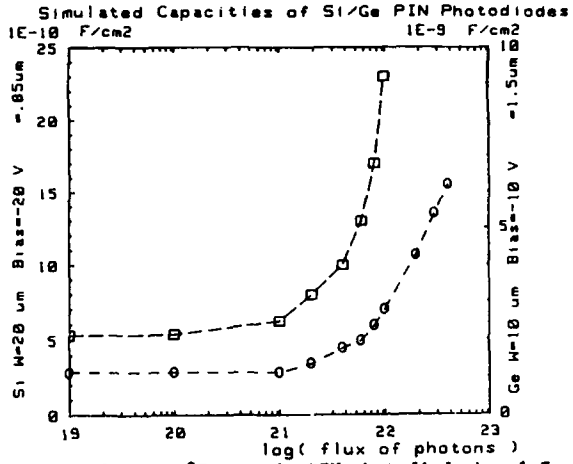
Fig.1 - Electric field distribution in Ge PIN photodiode, with high and low field value.

$$\begin{cases} \operatorname{div} \epsilon \overrightarrow{\operatorname{grad}} \varphi = q\cdot(n - p - C) \\ \dfrac{\partial n}{\partial t} = \dfrac{1}{q}\operatorname{div}(n\,\mu_n \overrightarrow{\operatorname{grad}} E_{Fn}) - U_{SHR} + G_{opt} \\ \dfrac{\partial p}{\partial t} = -\dfrac{1}{q}\operatorname{div}(p\,\mu_p \overrightarrow{\operatorname{grad}} E_{Fp}) - U_{SHR} + G_{opt} \end{cases} \quad \text{and} \quad \begin{cases} G_{opt}(x) = \displaystyle\int_\lambda \alpha(\lambda)\cdot\exp[-\alpha(\lambda)\cdot x]\, d\lambda \\ U_{SHR} = \dfrac{n\cdot p - n_1\cdot p_1}{\tau_n(p + p_1) + \tau_p(n + n_1)} \end{cases}$$

With n, p; τ_n, τ_p; μ_n, μ_p; E_{Fn}, E_{Fp} are respectively electron and hole carrier densities, carrier lifetimes, mobilities and Fermi levels. The electrostatic potential is φ. The dielectric constant is ϵ and the fixed charge density is C . Thermal generation recombination is described by the Schockley-Hall-Read formulation U_{SHR} and the external optical generation $G_{opt}(x)$ takes into account the α absorption value at position x, as a function of wavelength λ.
The set of equation is linearized and solved within an uncoupled scheme of resolution, using a variable implicit method to solve the transient case.
Fig.1 - Electric field distribution in Ge PIN photodiode, with high and low field value.

3 - NUMERICAL SIMULATION RESULTS

At high optical power, the electric field in the intrinsic layer of the PIN photodiodes is no longuer uniform (Fig 1). As the optical power increases, larger potential drops arise at the $P^+\nu$ and νN^+ junctions and increasingly wider areas of low electric field spread in the intrinsic material. The peak value of the electric field does not modify carrier velocities and is unsufficient to enhances avalanche. With electric field as low as a few kV/cm, carrier velocities can drastically decrease and lead to large transit time of the carriers across the intrinsic layer. With slow drift, in poor quality material, the carriers can suffers from recombination before being collected and thus sensitivity lowers. The obtained electric field distribution is equivalent to that of two junctions. The sum of the corresponding space charge regions is then smaller than the intrinsic layer width and it can be expected that the overall capacity increases.
Accurate computation of the capacity of the device is performed taking into account the local variation of the electric charge and of the displacement vector between two close bias [2]. Evolution with optical power is given in Fig 2.



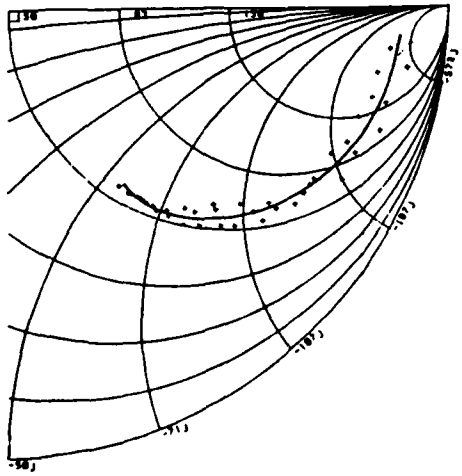
o Si PIN photodiode λ = .85 µm □ Ge PIN photodiode λ = 1.5 µm

Fig.2 - Increase of PIN photodiodes capacities, with increase of the local oscillator radiant power, as obtained by numerical simulation.

Numerical simulations show that for optical power larger than 10^{21} photons/s/ cm^2 the behaviour of the photodiodes can be far different than the low optical regime such photodiodes had been designed for. An optical power of 5 mW shining over 100 μm^2 (surface of the core of a monomode fiber) is equivalent to some $5 \cdot 10^{22}$ photons/s/cm^2.

4 - MEASUREMENTS AND IMPROVED EQUIVALENT CIRCUIT

The complex impedance of the photodiodes has been measured on a 50 Ω load, using a network analyser, through S_{11} parameter, after proper calibration, in the .1 - 2 GHz band width.



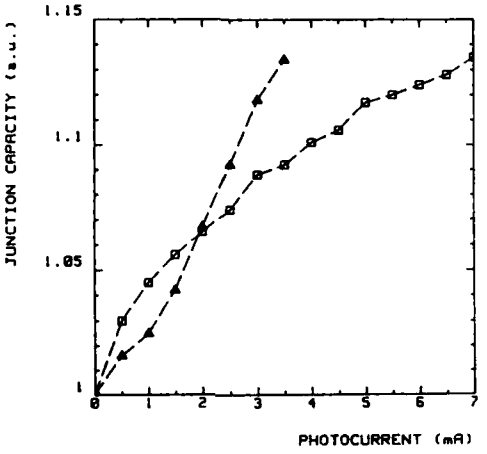
Experimental data (+) and impedance fit (continuous line).

Fig.3 - Improved equivalent circuit for Ge photodiode and value extraction method provide good impedance fit in the .1-2 GHz range.

The different elements of the equivalent circuit of the photodiode (junction and package capacitances, junction and serial resistances, wire inductances) are obtained through an

improved parameter extraction method. This method allows to determine the best fit values of the different elements of the equivalent circuit described by the operator without any prior knowledge of the values or ranges. The equivalent circuit of the photodiodes had to be completed by elements such as wire inductance or chip to package capacitance to obtain correct fit as shown in figure 3.

The variation of the junction capacitance and serial resistance of the photodiode have thus been extracted for optical power at 1.3 µm ranging from obscurity to 4.4 mW for the GaInAs/InP diode (3.5 mA photocurrent, 7.5 10^{20}photons/s/cm^2) and 9 mW for the Ge diode (7 mA, 1.5 10^{21}ph/s/cm^2).



△ GaInAs/InP PIN diode □ Ge PIN diode

Fig.4 - Relative junction capacitance increase versus diode photocurrent, as 1.3µm local oscillator radiant power increases. The junction capacities are obtained from microwave measurements and improved equivalent circuits of photodiodes.

As it can be seen in figure 4, the junction capacity of the Ge and GaInAs/InP photodiodes continuously increase with increasing optical power. The time constant (junction capacitance - serial resistance product) continuously increases but for such optical power is still below transit time.

5 - CONCLUSION

Numerical simulations of photodiodes predict that above an optical power corresponding to flux greater than 10^{21}photons/s/cm^2, the response time and the junction capacitance will continuously rise. Experimental measures confirm this increase of the capacitance and $R_s \cdot C$ time constant. For standard diode active diameter (70 µm), the optical power delivered by a conventionnal semiconductor laser (a few mW) is beneath the limit for which photodiode performances would shrink.

6 - REFERENCES

[1] Mears, C.L. and Batchman, T.E., IEEE vol LT 5, N°6, June 87, pp 827-837
[2] Viallet, J.E. and Mottet, S., Nasecode IV Conference, Dublin 1985 - pp 530-541
[3] Mottet, S. and Viallet, J.E., Nasecode V Conference, Dublin 1987, pp 289-294

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CHARACTERIZATION OF AN AlGaAs/GaAs METAL-SEMICONDUCTOR-METAL PHOTODETECTOR

M. ZIRNGIBL, R. SACHOT and M. ILEGEMS

Institut de Micro- and Optoélectronique, Swiss Federal Institute of Technology, CH-1015 Lausanne, Switzerland

Résumé: Nous reportons la caractérisation d'un photodétecteur métal-semiconducteur-métal en fonction de sa géometrie, de la tension appliquée et de la puissance optique incidente. De l'oxyde d'étain et d'indium (ITO) est déposé sur des substrats GaAs et des couches MBE afin de définir les contacts Schottky interdigités qui forment la surface active. Les détecteurs montrent une efficacité quantique externe élevée et une réponse rapide (temps de montée < 70 ps, largeur à mi-hauteur < 185 ps).

Abstract: The characteristics of metal-semiconductor-metal photodetectors as a function of layout geometry, applied voltage and optical input power are reported. The structures are fabricated using interdigitated indium tin oxide Schottky contacts deposited by sputtering on GaAs substrates and MBE-layers to form the active area. The detectors show high external quantum efficency and fast response (risetime < 70 ps, FWHM < 185 ps).

In an optical telecommunication or read-out system, sensitive and broadband photodetectors are necessary as interface elements between an optical waveguide and a electronic circuit. There is an increasing interest in metal-semiconductor-metal photodetectors (MSM) because of their high speed, their planar structure which is compatible with integration with electronic devices and their large active area which facilitates the coupling of the light into the detector. The capacitance of their structure is detemined by the lateral distance between the contact fingers at a given active area of the detector so that the RC product is small even for detectors with large areas (typical 200 fF for 200x200 μm^2 active area at punch through voltage). Finally, the fabrication of the MSM's is very simple compared to that of vertical detectors like p-i-n diodes. In spite of these promizing features, the explanation of the internal gain mechanism is still a subject of research[1-8].

We have fabricated MSM's with an indium tin oxide (ITO) metallization on undoped and Cr doped GaAs substrates (referred to as GaAs-MSM and GaAs(Cr)-MSM) and on undoped AlGaAs/GaAs MBE--layers (referred to as MBE-MSM) using a simple one-level lift-off technology. ITO, which acts as an anticoating layer at 820 nm, is sputter-deposited on the semiconductor surface to form the interdigitated Schottky contacts (Fig.1) so that the whole detector area can be illuminated. The size of the active area is 50x50 μm^2 and 200x200 μm^2. The geometry is adapted to be mounted into a 50 Ohm coplanar microstripline. The distance between the fingers, L, and the finger width, d, varies from 1 to 16 µm. The MBE-layer consists of an undoped 1 µm thick GaAs active layer and an underlying 0.5 µm thick $Al_{0.3}Ga_{0.7}As$ buffer layer to limit the light collection zone.

Measurements of sensitivity are made by focusing the beam of a laser diode which emits at 820 nm on the detector so as to achieve an uniform illumination over the entire active area. The optical input power is measured with a calibrated Si photodiode. In Fig. 2 we plot the photo- and dark current against bias voltage for a MBE-MSM with finger width and finger spacing of 3 μm. The dark current is probably governed by structural defects rather than by the Schottky barrier heigth as evidenced by the asymmetry of the I-V characteristics. Contrary to photoconductors, the photocurrent is not proportional to the bias voltage but determined by the reverse Schottky diode caracteristics. The sensitivity is 0.46 A/W (Fig. 3) independent of the optical input power. Fig. 4 shows the detector response to a short laser pulse of 60 ps risetime and 100 ps full width at half maximum (FWHM). The detector risetime of 72 ps is limited by the optical pulse while the FWHM of 184 ps is probably limited by the hole transit time (120 ps at $5x10^6$ cm/s saturation velocity taking into account the carriers generated under the fingers)

The detectors fabricated directly on the substrates show the same risetime but a higher FWHM (about 250 ps). Their photocurrent as a function of the finger spacing L for different values of the average electric field is represented in Fig. 5. The data are taken on diodes of 200x200 μm^2 active area with 4 um and 8 μm nominal finger width. For the detectors on undoped GaAs substrates, the dependence on geometry and electrical field is weak, contrary to those on GaAs(Cr) which show an increase of the photocurrent with L. The sensitivity of the GaAs(Cr)-MSM seems to depend only on L and not on the finger width. Unity external quantum efficiency is situated at 290 μA. Thus, the GaAs(Cr)-MSM's have external efficiencies up to 500% for large L. To our knowledge, these are the highest values reported so far for this kind of detector. Taking into account reflexion losses the GaAs-MSM's show also a small internal gain. The measured dark currents at 2 V/μm average field are 10 nA and 100 nA for the MSM's on GaAs(Cr) and on the undoped GaAs, respectively and do not show any clear dependence on the device geometry. The GaAs(Cr)-MSM with a small finger spacing show a linear dependence of the photocurrent with the input power (Fig. 6) while the sensitivity of those with a large L varies at first sublinearly with input power and then increases sharply above 100 μW for the structures operated at 29 V bias applied (Fig. 7). At lower optical input power, the photocurrent is independent of the finger spacing L. Focusing the laser spot alternatively on the two electrodes, we have observed that the photocurrent is maximum on the negative electrode and minimum on the positive electrode. The ratio of these currents is 10 for the GaAs(Cr)-MSM and 2 for the GaAs-MSM.

The interpretation of the measurements of the GaAs- and GaAs(Cr)-MSM's is not clear. The observed gain is not believed to be due to a photoconductive effect because the contacts are not ohmic. Avalanche multiplication may be involved in areas of locally enhanced electric field, although the average lateral field is as low as 2 V/μm. There may be an optical induced lowering of the Schottky barrier heigth, but no clear dependence of the photocurrent on the contact area is observed. Measurements of the barriers under illumination and comparisons between MSM's with different metallizations are being carried out to further elucidate the observed behavior.

In conclusion, we have fabricated and characterized a very simple MSM-photodetector with interdigitated ITO Schottky contact. This type of detector could find widespread applications in optical telecommunication- and data systems because of its simplicity and low cost, its high speed (risetime < 70

ps, FWHM < 185 ps) and sensitivity (up to 500% external quantum efficiency), and its planar structure.

Acknowledgements: We thank A. Nerini for her excellent technical assistance and J. M. Breguet and P. Maréchal for measurements. This work was supported by the Swiss National Science Fondation, research program 13.

[1]C. W. Slayman, L. Figueroa, IEEE Electron Dev. Lett. EDL-2, 112 (1981)

[2]C. J. Wei, H.-J. Klein, H. Beneking, Electron. Lett. 17, 689 (1981)

[3]M. Ito, O. Wada, K. Nakai, T. Sakurai, IEEE Electron Dev. Lett. EDL-5, 531 (1984)

[4]H. Schumacher, P. Narozny, Ch. Werres, H. Beneking, IEEE Electron Dev. Lett. EDL-7, 26 (1986)

[5]M. Ito, O. Wada, IEEE J. Quant. Electron. QE-7 , 1073 (1986)

[6]R. L. Robertson, IEEE J. Quant. Electron. QE-7 , 1185 (1987)

[7]D. L. Rogers, IEEE Electron Dev. Lett. EDL-11, 600 (1986)

[8]D. K. W. Lam, R. I. MacDonald, J. P. Noad, B. A. Syrett, IEEE Trans. Electron. Dev. ED-5, 1057 (1987)

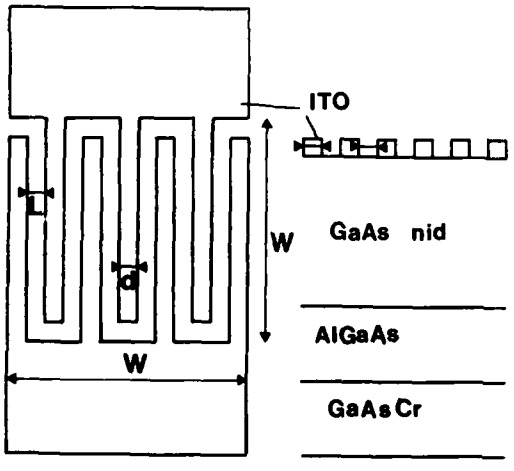


Fig.1: Top view and cross section of a MSM on a MBE layer: The dimensions are W = 50 µm and 200 um, L = 1,2,3,4,8,12 and 16 µm and d = 1,2,3,4 and 8 µm (nominal values on the mask, exact values are measured by SEM). The total area of the Schottky contacts measures between 9000 and 30000 μm^2 for the 200x200 um^2 MSM's and 1250 μm^2 for the 50x50 μm^2 MSM's. The MSM is designed to be integrated into a 50 Ohm microstripline.

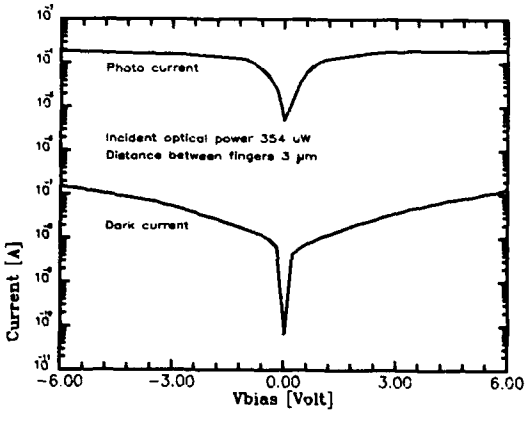


Fig. 2: I-V characteristics in the dark and under illumination of a MBE-MSM with an active area of 50x50 um^2 and 3 µm finger width and spacing.

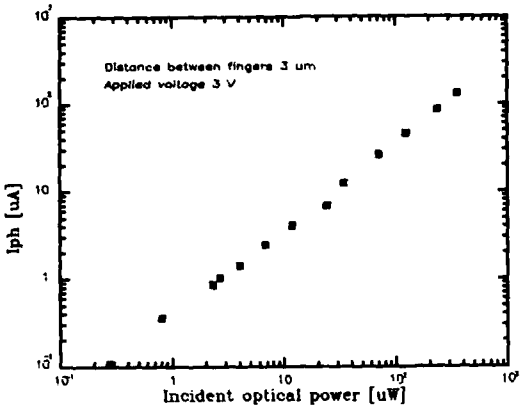


Fig. 3: Photocurrent versus incident optical power for the same MSM as shown in Fig. 2.

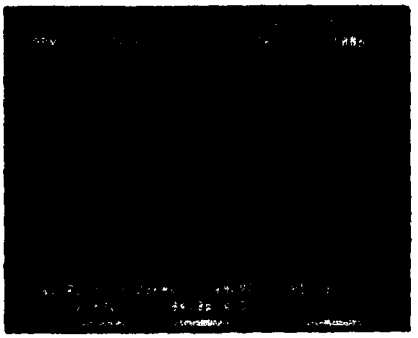


Fig. 4: Response of a MBE-MSM detector to a laser pulse of 60 ps risetime and 100 ps FWHM. The structure tested is the same as shown in Fig. 2 and 3 . The data are taken under 16 V applied bias to the detector. The MSM is mounted on a coplanar microstripline to assure 50 Ohm matching and the output signal is measured with a S4 Tektronix sampling head.

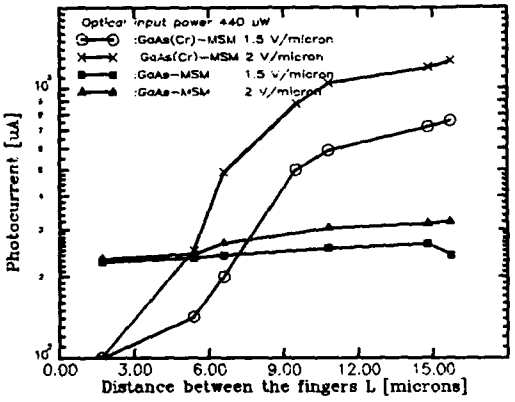


Fig. 5: Photocurrent as a function of finger spacing L for different values of the average electric field (V_{bias}/L) between the fingers. The Schottky contact area, which is different for the 7 structures, does not influence the sensitivity.

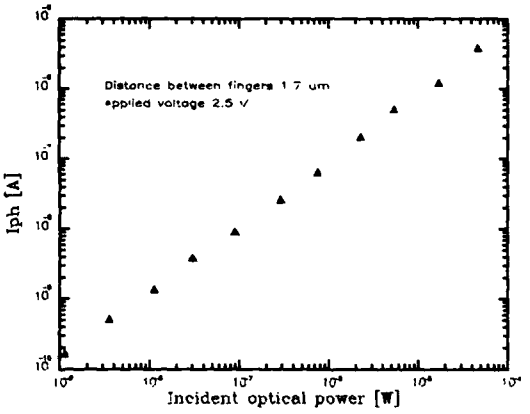


Fig. 6: Photocurrent versus incident optical power for a GaAs(Cr)-MSM with 1.7 μm finger spacing measured at 2.5 V bias.

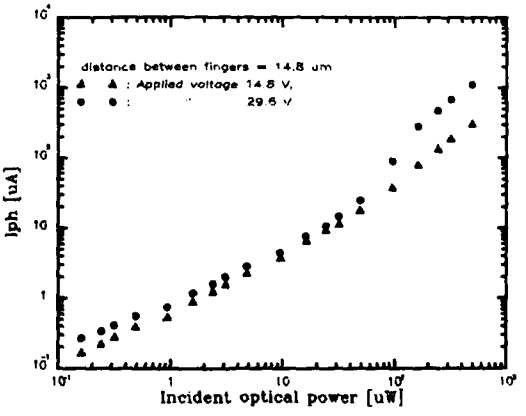


Fig. 7: Photocurrent versus incident power of a GaAs(Cr)-MSM with 14.8 μm fingerspacing measured at 14.8 and 29.6 V bias.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

HIGH SPEED InAlAs/InGaAs DOUBLE HETEROSTRUCTURE p-i-n's

J.-C. BISCHOFF[1], T.H. HOLLENBECK*,[2], R.N. NOTTENBURG*,[3], M.C. TAMARGO*, J.L. DE MIGUEL*, C.F. MOORE* and H. SCHUMACHER*

Institute for Micro and Optoelectronics, Swiss Federal Institute of Technology, CH-1015 Lausanne, Switzerland
**Bell Communications Research Inc., 331, Newman Spring Rd., Red Bank, NJ 07701, U.S.A.*

Résumé - Des photodiodes p-i-n InAlAs/InGaAs à double hétérostructure ont été produites par MBE et montées sur des guides d'onde coplanaires. Deux méthodes: (i) l'incorporation de couches InAlAs non dopées entre la zone d'absorption InGaAs et les couches InAlAs dopées et (ii) l'utilisation d'hétérojonctions graduelles ont été utilisées pour réduire l'accumulation de porteurs aux interfaces. Bien que les deux méthodes améliorent les performances des diodes p-i-n à double hétérojonction, les meilleurs résultats ont été obtenus avec les hétérojonctions graduelles: une efficacité quantique de ≈38 % à 1.3 µm, un temps de montée de 21 ps et une largeur à mi-hauteur de 40 ps ont été obtenus pour une diode ayant une surface de 24x24 µm² et une couche d'absoption d'une épaisseur de 0.5 µm.

Abstract - MBE grown double heterostructure InAlAs/InGaAs p-i-n photodiodes have been flip-chip mounted on coplanar waveguides. Two methods: (i) incorporation of doping setback InAlAs layers between the InGaAs absorption region and the doped InAlAs layers and (ii) compositional grading have been used to reduce carrier pile-up at the heterointerfaces. Although both methods improve the diode performances, the best results were obtained with compositional grading: a quantum efficiency of ≈38 % at 1.3 µm, a rise time of ≈21 ps and a FWHM of ≈40 ps have been obtained for a device with a 24x24 µm² area and a 0.5 µm thick absorption region.

One of the most promising approaches for detection in the 1.0 - 1.7 µm wavelength range is the use of a p-i-n photodiode coupled to a field effect transistor amplifier. $In_{0.53}Ga_{0.47}As$ based p-i-n photodiodes have already shown low dark current, good quantum efficiency and high speed operation /1//2//3/. Most devices investigated to date are of the homojunction type with a p-type layer formed by diffusion of Zn in the InGaAs layer. In contrast, the devices investigated here are MBE grown $In_{0.52}Al_{0.48}As/In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ double heterostructure p-i-n's.

Immediate advantages of a double heterostructure are an excellent control of the absorption region thickness and the elimination, in the photoresponse, of the diffusion tail due to absorption in low field regions. The pulse response of these devices is however, typically characterized by long tails due to carrier pile-up (trapping) at the heterojunction interfaces /4/. We present results of an investigation of two possible methods to reduce the deleterious effects of carrier trapping: (i) incorporation of doping setback layers at the heterointerfaces and (ii) compositional grading of the heterointerfaces.

---

[1] Work done while at Bell Communications Research Inc., Red Bank, NJ 07701, U.S.A., as a resident visitor
[2] Present address : Princeton University, Princeton, NJ 08544, U.S.A.
[3] Present address : AT&T Bell Laboratories, Murray Hill, NJ 07974, U.S.A.

The samples were grown by Molecular Beam Epitaxy (MBE) on InP (Fe-doped) semi-insulating substrates. Three types of structures were realized. The first one is a double heterojunction p-i-n structure with abrupt interfaces which consists of an undoped $In_{0.53}Ga_{0.47}As$ absorption layer sandwiched by two $In_{0.52}Al_{0.48}As$ layers, one doped with Si (n-type) and the other with Be (p-type) The two others are slightly modified structures. One contains 300 Å setback layers of undoped $In_{0.52}Al_{0.48}As$ separating the doped layers from the intrinsic $In_{0.53}Ga_{0.47}As$ and the other one, compositionally graded (In, Ga, Al) As layers gradually going from the $In_{0.52}Al_{0.48}As$ composition to the $In_{0.53}Ga_{0.47}As$. These graded layers are 300 Å thick and undoped.

Mesa diodes, designed for back illumination through the semi-insulating InP substrate, were fabricated from the as-grown wafers and mounted, upside down, on coplanar waveguides. With this technique, the losses due to the coupling of the p-i-n diode to the coaxial connector (Weltron K connector) can be maintained below 3 dB at 26 GHz.

Dark currents of respectively 328 pA and 537 pA at 5 V were measured for devices with compositional grading and areas of respectively 14x14 μm^2 and 24x24 μm^2. For these devices, whose absorption layer thickness is 0.5 µm, the capacitances at 0 V bias are respectively 50 fF and 145 fF.

We present in fig. 1 the reverse I-V characteristics of double heterostructure p-i-n diodes with abrupt heterointerfaces, in the dark (lower curve) and front illuminated with white light through the prober microscope (upper curve). The photoresponse shows a turn-on voltage at ≈ 0.4 V. Dispersion in the characteristics of diodes issued from different processing runs indicates that this turn-on voltage is sensitive to the processing and is related to the quality of the Schottky type Ti-Au contacts. Above the turn-on voltage, the sensitivity of the photodiodes is fairly independent of the applied bias voltage.

In fig. 2, we have reported the peak amplitude of the response of double heterostructure p-i-n diodes to optical pulses generated by active mode locking of a laser diode as a function of the reverse bias voltage (characteristics of pulses generated by laser #1: rise time ≈ 44 ps, FWHM ≈ 45ps, repetition rate ≈ 350 MHz and peak power ≈ 14.6 mW.). It appears that these devices have, under pulsed excitation, a behavior which differs significantly from that under continuous illumination. Indeed the characteristics reported in fig. 2 show that our devices do not respond to fast optical pulses when their reverse bias voltage is below a threshold value which depends on the type of heterointerface: ≈1 V for devices with compositionally graded interfaces, ≈2.8 V for devices with doping setback layers and ≈5 V for devices with abrupt interfaces. Above threshold, the peak current of the three types of devices increases linearly with bias in the voltage range shown in fig. 2. The best sensitivity, obtained at 6 V bias for a double heterostructure p-i-n with a 0.5 µm thick absorption layer, is 0.4 A/W (peak amplitude = 291 mV). At 1.3 µm, this corresponds to an external quantum efficiency of 38%.

At higher bias voltages, the pulse response peak amplitude levels off (fig. 3, characteristics of pulses generated by laser #2: rise time ≈ 5.3 ps, FWHM ≈ 7.4 ps, repetition rate ≈ 440 MHz and peak power ≈ 6 mW.). Only the devices with compositional grading were measured in the higher voltage range. Indeed for devices with abrupt interfaces or doping setback, the breakdown voltage, ≈ 8 V, was reached before any leveling of the response could be observed.

Devices with abrupt interfaces show a 80 ps rise time and a 1.4 ns fall time at 5.5 V bias. This long fall time is strongly temperature dependent and can be reduced to 180 ps by elevating the detector temperature by directing a hot air flow to the diode. Devices which employ doping setback and compositional grading show threshold response at bias levels of 3 V and 1.5 V respectively. Operating the devices, with doping setback, at 5.5 V bias, we observe over an order of magnitude improvement in fall time: 120 ps versus 1.34 ns for the diode with abrupt interfaces. The same observation can be made at 3.3 V bias by comparing the devices having compositional grading with the ones having doping setback: 110 ps versus 1.02 ns.

Although the speed of all the devices is increasing with bias voltage the best results are achieved with compositional grading; we obtain a 55 ps rise time, a 70 ps FWHM and a 102 ps fall time at 6 V. This response is limited primarily by our laser diode (laser #1). Indeed measurement of devices with compositional grading on an identical set-up, except for the laser diode (laser #2), have resulted in a 33 ps rise time, a 54 ps FWHM and a 82 ps fall time at 8 V bias for an active area of 24x24 μm^2 (fig. 4). Devices of smaller size have not shown any speed improvement over the largest ones. This is due to their smaller top contact to mesa area ratio resulting in an increased series resistance.

To evaluate the intrinsic rise time and FWHM of our devices, we have made the following assumptions: (i) the pulses have a gaussian waveform, (ii) the S4 sampling head rise time is 25 ps and the corresponding FWHM is 35 ps and (iii) the light pulse rise time and FWHM are 5.3 ps and 7.4 ps /5/. Under these assumptions, we obtain for the device whose pulse response appears in fig. 4 an intrinsic rise time of 21 ps and a FWHM of 40 ps.

During this work, we made the following observations. The dependence of the diode responsivity on bias voltage is apparent only under pulsed excitation. The responsivity of the diode is larger when the areal density of carriers trapped at the heterointerface under dark conditions is low, for example, when the band discontinuity is small (compositional grading of the heterointerface) or when the band potential minimum is above (for electrons) the quasi-Fermi level (structure with doping setback layers under reverse bias). These observations lead us to the assumption that the time spent by a carrier at the heterointerface, and thus the probability to be lost by recombination, decreases when the photogenerated carrier density in the well is large in comparison with the carrier density in the well under dark conditions. However, further work will be necessary to determine exactly the mechanisms involved in the transport of carriers through band discontinuities in double heterostructure p-i-n diodes.

In conclusion, we have presented two methods to reduce carrier trapping at the heterointerfaces in InAlAs/InGaAs double heterostructure p-i-n photodiodes: incorporation of doping setback InAlAs layers and of compositionally graded InAlGaAs layers. The best results have been obtained with devices with compositional grading and are very promising: external quantum efficiency $\approx$ 38 %, rise time $\approx$ 21 ps and FWHM $\approx$ 40 ps.

The authors wish to acknowledge R. F. Leheny and M. Ilegems for their support and M.-H. Meynadier, C. Caneau, B. W. Meagher, D. A. Humphrey, R. J. Martin and B. Zimmermann for their assistance during the course of this work. We are particularly grateful to H. Temkin for his advise and for providing the laser diodes.

References

/1/ J. E. Bowers, C. A. Burrus, "Optoelectronic Components and Systems with Bandwidths in Excess of 26 Ghz", RCA Review, Vol. 46, pp. 497-509, Dec. 1985.

/2/ J. E. Bowers, C. A. Burrus, "InGaAs PIN Photodetectors with Modulation Response to Millimetre Wavelengths", Electronics Letters, Vol. 21, pp. 812-814, 29th August 1985.

/3/ H. Temkin, R. E. Frahm, N. A. Olsson, C. A. Burrus, R. J. Mccoy, "Very High Speed Operation of Planar InGaAs/InP Photodiode Detectors", Electronics Letters, Vol. 22, No. 23, pp. 267-269, 6th Nov. 1986.

/4/ Y. Zebda, P. Bhattacharya, M. S. Tobin and T. B. Simpson, "Design and Performance of Very High-Speed $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ p-i-n Photodiodes Grown by Molecular Beam Epitaxy", IEEE Electron Device Letters, Vol. EDL-8, No. 12, December 1987, pp. 579-581.

/5/ J. P. van der Ziel, "Active mode locking of double heterostructure lasers in an external cavity", J. Appl. Phys., Vol. 52, No. 7, pp. 4435-4446, July 1981.

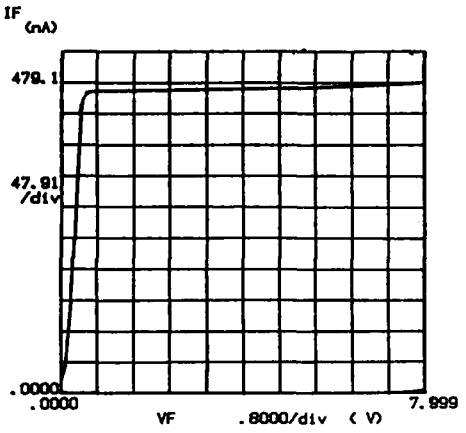


Fig. 1 I-V characteristics of a 24x24 μm² double heterostructure p-i-n diode with abrupt interfaces in the dark (lower curve, I = 586 pA for V = 5.0 V) and illuminated with white light through the prober microscope (upper curve).

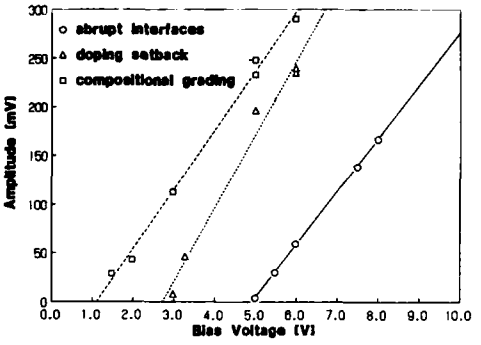


Fig. 2 Dependence of peak amplitude on bias voltage of 24x24 μm² diodes with abrupt interfaces, doping setback layers and compositional grading. Optical pulse characteristics: rise time ≈ 44 ps, FWHM ≈ 45 ps, repetition rate ≈ 350 MHz and peak power ≈ 14.6 mW.

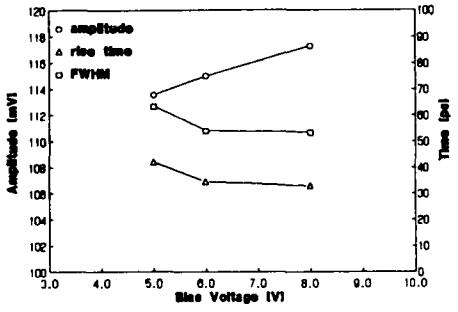


Fig. 3 Dependence of peak amplitude, rise time and FWHM on bias voltage of a 24x24 μm² with compositional grading. Optical pulse characteristics: rise time ≈ 5.3 ps, FWHM ≈ 7.4 ps, repetition rate ≈ 440 MHz and peak power ≈ 6 mW.

Fig. 4 Pulse response of a 24x24 μm² diode with compositional grading at 8 V bias. Optical pulse characteristics: rise time ≈ 44 ps, FWHM ≈ 45 ps, repetition rate ≈ 350 MHz and peak power ≈ 14.6 mW.
Scale x: 50 ps/div.
y: 20 mV/div.,

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

GaSb AND GaInAsSb PHOTODETECTORS FOR $\lambda > 1.55$ µm PREPARED BY METAL ORGANIC CHEMICAL VAPOR DEPOSITION

G. BOUGNOT, F. PASCAL, F. ROUMANILLE, J. BOUGNOT, L. GOUSKOV, F. DELANNOY, P. GROSSE and H. LUQUET

Centre d'Electronique de Montpellier (CNRS-UA 391), Université des Sciences et Techniques du Languedoc, Place Eugène Bataillon, F-34060 Montpellier Cedex, France

Résumé - Des couches de GaSb de type p et n ont été élaborées par dépôt en phase vapeur à partir d'organométalliques (MOCVD) sur des substrats de GaSb et de GaAs semi-isolant ; les plus faibles niveaux de dopage obtenus sont p = 2 x $10^{16}cm^{-3}$, n = 8 x $10^{15}cm^{-3}$. Des couches de $Ga_{1-x}In_xAs_ySb_{1-y}$ de type p ont aussi été fabriquées par cette méthode avec une composition assurant la photodétection à 2.5µm. Les propriétés des couches et des jonctions obtenues sont décrites.

Abstract - p and n GaSb layers have been grown on GaSb and semi-insulating GaAs substrates by Metal Organic Chemical Vapor Deposition (MOCVD) method ; the lowest doping levels obtained are : p = 2 x $10^{16}cm^{-3}$, n = 8 x $10^{15}cm^{-3}$. P type $Ga_{1-x}In_xAs_ySb_{1-y}$ layers have also been fabricated by MOCVD with a composition insuring the photodetection at a wavelength of 2.5µm. The properties of the layers and of the as grown junctions are described.

With the aim of realizing a Separated Absorption and Multiplication (SAM) photodetector, at a wavelength of 2.5µm preparation and characterization of GaSb and $Ga_{1-x}In_xAs_ySb_{1-y}$ layers and devices have been undertaken.

1 - FABRICATION /1/

The different layers are elaborated by Metal Organic Chemical Vapor Deposition (MOCVD) using trimethyl-gallium (TMG), trimethyl-indium (TMI) and trimethyl-antimony (TMSb) in a vertical reactor at atmospheric pressure. The n doping is obtained using dimethyl tellurium (DMTe) diluted in H_2. The results presented here are relative to :

- n and p type GaSb layers deposited on semi-insulating GaAs ; GaSb mesa diodes : p GaSb/n GaSb substrate.
- mesa diodes : p $Ga_{1-x}In_xAs_ySb_{1-y}$/n GaSb substrate, this device has allowed to test the infrared photodetection in the MOCVD ternary and quaternary alloys.

2 - LAYERS CHARACTERIZATIONS

p GaSb

The lowest p_H value (Hall p density) obtained in our native GaSb is 2 x $10^{16}cm^{-3}$ at 300K with an associated mobility μ_{pH}= 900cm²/V.s. The variations of p_H and μ_{pH} with temperature (77 < T < 300K) are shown on figure 1a and b. A model involving two acceptor levels and a compensating donor level has been used to account for the p variation and leads to the determination of the energies E_{A_1} = 23meV, E_{A_2} = 76 meV with respective densities N_{A_1} = 1.3 x $10^{16}cm^{-3}$, N_{A_2} = 1.03 x $10^{16}cm^{-3}$. The donor density is negligible, the hole effective mass has been choosen equal to 0.23 m_o/2/, the spin degeneracy factor g_{A_1} = 4 ; the factor g_{A_2} = 2 has been derived from the calculation. This result indicates that this p layer is weakly doped and weakly compensated and that the involved elaboration technique does not introduce any new electrically active defect. The existence of two native acceptor levels has already been derived from optical and electrical measurements in various GaSb samples ; our N_{A_1} and N_{A_2} values are nearly equivalent indicating that E_{A_2} may correspond to the 2nd ionization of A_1. Our less doped sample is comparable to the best publicated results of H. Miki et al./3/ ; p_H= 1.1 x $10^{16}cm^{-3}$, μ_{pH}= 620cm²/V.s and M. Lee et al/4/ : p_H= 7.8 x $10^{15}cm^{-3}$, μ_{p_H}= 950cm²/V.s.

n GaSb

The n doping level has been controlled by varying the DMTe pressure in the reactor /5/. The lowest n_H value has been obtained with the following growth parameters : T = 560°C, P(DMTe) = 1 x 10^{-6} Atm ; V/III ratio = 1 : n_H = 8 x 10^{15} cm^{-3}, μ_{nH} = 5000cm^2/V.s at 300K. The figures 2a and b show the variations of n_H and μ_{nH} as a function of T for three representative samples. The increase of n_H with decreasing temperature can be explained by the contribution of two types of carriers (Γ and L bands) coming from a degenerated Te donor level.

In conclusion, despite of the huge lattice mismatch ($\Delta a/a = 7.8.10^{-2}$) the electrical properties of these p and n layers are very similar to those of single crystal materials.

3 - OPTOELECTRONIC PROPERTIES OF THE DEVICES

GaSb diodes

They are made of natural p MOCVD GaSb layer on n GaSb substrate. The mesa diode diameter is 190µm.

- Capacitance - Voltage (C-V) characteristics.

Some devices exhibit a linear $1/C^2 = f(V)$ variation, however, generally such a variation is not observed and two limiting effective doping levels N_{B_1} at low voltage and $N_{B_2} > N_{B_1}$ at high voltage can be deduced. This fact could be related to an interaction layer - substrate during the growth which needs further investigations. The experimental N_B values does not exactly coïncide with the calculated ones deduced from the hole density in the layer and the electron density in the substrate ; M. Yano et al /6/ have already observed such a discrepancy and attributed it to interface defects.

- Current voltage I-V characteristics.

A typical example of room temperature I-V characteristic is shown on figure 3. A generation-recombination current I_{G-R} is dominant in the low voltage range with a τ value of 10^{-10}s. For higher polarization values an excedentary current is present greater than the expected current. This current can be due to the extension of the space charge in a perturbed region with lower values, this region could be the metallographic interface ; this current may also be due to surface leakages in these unpassivated devices. For V > 7V a more or less soft breakdown occurs which can be attributed to multiplication. The multiplication is confirmed by the variation of the I-V characteristic with temperature which exhibits a shift of the breakdown to lower V values and a more abrupt avalanche ; it is also confirmed by the observed photocurrent multiplication measured at 1.3µm (figure 4).

- Spectral responses

The figure 5a represents a typical spectral response of a GaSb diode with a junction depth of 4 µm. The external quantum efficiency at 1.55µm is equal to 0.22. Based on a single model of homojunction, the diffusion length in the MOCVD GaSb layer can be estimated : Ln ≃ 4µm.

$Ga_{1-x}In_xAs_ySb_{1-y}$ diodes

In this first step of investigation, the non optimized, and generally mismatched heterojunctions made of a p $Ga_{1-x}In_xAs_ySb_{1-y}$ deposited on n GaSb substrate have been elaborated in order to determine the growth parameters corresponding to x and y values allowing the photodetection at a wavelength of 2.5µm.

The spectral response of figure 5b is relative to the diode P20 in which the quaternary layer was $Ga_{0.8}In_{0.2}As_{0.18}Sb_{0.82}$ is matched to GaSb, the figure 5c is relative to the diode I9 on which the ternary layer was $Ga_{0.63}In_{0.37}Sb$. A backward illumination (through the substrate) has been used to do the measurement.

For the matched heterojunction, the spectral response is classical, the cut-off wavelength in the near infrared range is about 2.3µm. For the mismatched heterojunction, the cut-off appears beyond 2.5µm (the evaluated gap for x = 0.37 is 430 meV) but the photoresponse cannot be explained by the presence of only two materials : GaSb (Eg = 730 mev), $Ga_{0.63}In_{0.37}Sb$ (Eg = 430 meV), the increase of the response at about 1.9µm is probably due to a material of intermediate composition.

4 - CONCLUSION

MOCVD method has allowed to elaborate p and n type GaSb layers with electrical properties very similar to the best ones measured on material grown by other techniques.

The growth method has also allowed to deposite $Ga_{1-x}In_xAs_ySb_{1-y}$ layers with a gap value adapted to the photodetection at λ = 2.5µm.

This work must be extended with three main purposes : the decrease of the doping level in GaSb, the control of the growth conditions in order to fabricate $Ga_{0.74}In_{0.26}As_{0.24}Sb_{0.76}$ layers matched to GaSb with Eg = 490 meV and the evaluation of the phototransport parameters in this absorbing layer.

REFERENCES

/1/ Bougnot, G., Foucaran, A.F., Marjan, M., Etienne, D., Bougnot, J., Delannoy, F.M.H., Roumanille, F.M., J. of Crystal Growth 77-400 (1986).

/2/ Nakashima, K., Jap. J. of Appl. Phys. 20-1085 (1981).

/3/ Miki, H., Segawa, K., Fujibayashi, K., Jap. J. of Appl. Phys. 13-203 (1974).

/4/ Lee, M., Nicholas, D.J., Singer, K.E., Hamilton, B., J. of Appl. Phys. 59-2895 (1986).

/5/ Bougnot, G., Pascal, F., Roumanille, F., Grosse, P., Bougnot, J., Gouskov, L., Foucaran, A.,2nd Eur. Workshop on MOVPE - Fife (Scotland) 19-22 june 1988.

/6/ Yano, M., Suzuki, Y., Ishii, T., Matsushima, Y., Kimata, M., Jap. J. of Appl. Phys. 17-2091 (1978).

This work has been supported by DAII.

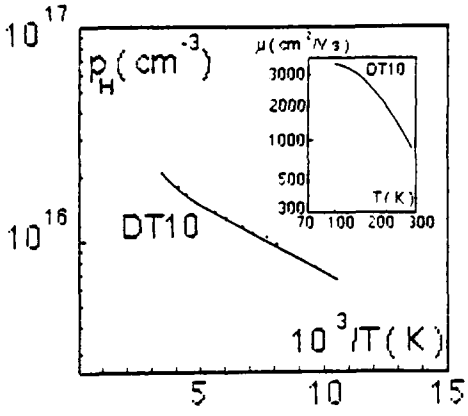


Fig.1 - Hole density and mobility in p type MOCVD GaSb layer as a function of temperature

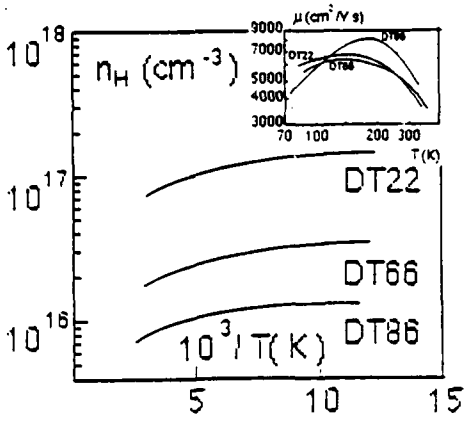


Fig.2 - Electron density and mobility in n type MOCVD Te doped layers as a junction of temperature

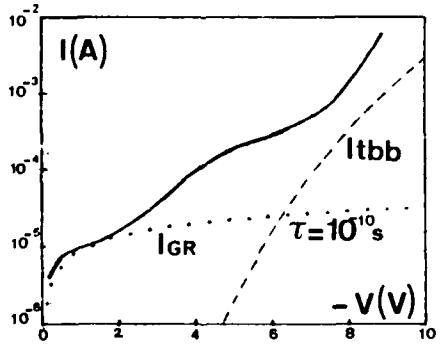


Fig.3 - Reverse dark I-V characteristic at 300K for a typical GaSb diode.
...Calculated generation-recombination I_{GR} and band to band tunnel I_{tbb} currents

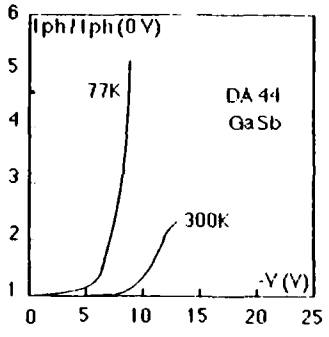


Fig.4 - Photocurrent multiplication at 300 and 77K measured on a GaSb diode (λ = 1.3µm)

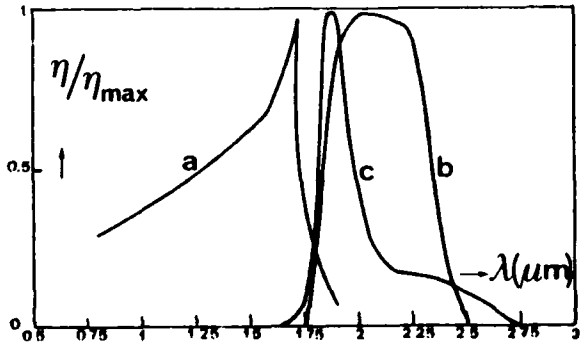


Fig. 5 - Spectral responses measured on :
a) GaSb photodiode DT29 η_{max} = 0.32
b) matched $Ga_{0.2}In_{0.8}As_{0.18}Sb_{0.82}$/GaSb heterojunction P20
c) mismatched $Ga_{0.63}In_{0.37}Sb$/GaSb heterojunction I9

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

1.5µm PHASE SHIFTED DISTRIBUTED FEEDBACK LASER

M. GILLERON, J. CHARIL, D. LESTERLIN, P. CORREC and J.C. BOULEY

Centre National d'Etudes des Télécommunications, Laboratoire de Bagneux, 196, Av. Henri Ravera, F-92220 Bagneux, France

Résumé - Nous présentons la conception, la technologie de fabrication et les propriétés d'émission de lasers à contre-réaction distribuée déphasée. L'utilisation d'un ruban de largeur non uniforme conduit à un pourcentage plus élevé de fonctionnement monomodal.

Abstract - We present the design, technology and lasing properties of phase-shifted distributed feedback lasers. The use of a non-uniform width stripe leads to a higher percentage of single-longitudinal-mode operation.

1-INTRODUCTION

GaInAsP/InP distributed feedback (DFB) lasers are attractive light sources for long-haul, high bit-rate optical fibre transmission systems. In order to improve the single-mode operation yield, many structures have been proposed /1,2,3/ to introduce an optical *phase-shift inside the cavity to remove* the wavelength degeneracy and suppress the wavelength dependence on the facet positions usually observed in the conventional DFB lasers. In this paper, we present the technology and properties of a 1.55µm Phase-Shifted Buried Ridge Stripe (PS-BRS) laser easily obtained by adjusting the optical wave propagation constant with a longitudinally non-uniform stripe width.

2-STRUCTURE

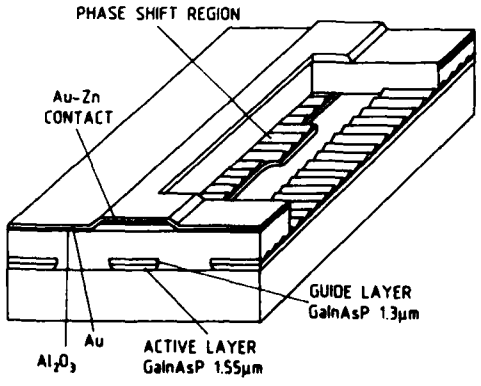


Fig. 1: Schematic view of the 1.5µm Phase-Shifted Buried Ridge Stripe laser.

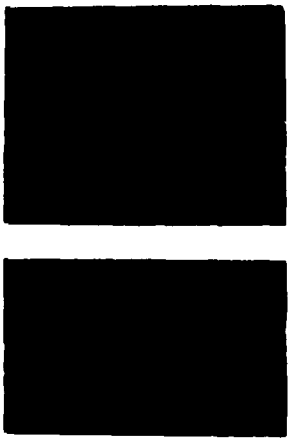


Fig. 2: SEM photograph of the phase shifter before the second epitaxy.

The PS-BRS laser, derived from the BRS structure /4/, is schematically shown in Fig. 1. The laser reported here has been entirely grown by Liquid Phase Epitaxy (LPE). First, a n-InP buffer layer, a 1.5µm GaInAsP active layer and a 1.37µm GaInAsP guide layer were successively grown by two-phase LPE. Then, a 2nd order grating was chemically etched into the guide layer and a non-uniform width stripe selectively etched into the active and guide layers (Fig. 2). Finally, single-phase, low-temperature LPE was used to overgrow the

burying p-InP and the contact p⁺-quaternary layers. GaAs protection cover was used to reduce the grating deformation likely to occur during the heating period prior to overgrowth. Classical technologies were used to obtain 300µm-long laser chips.

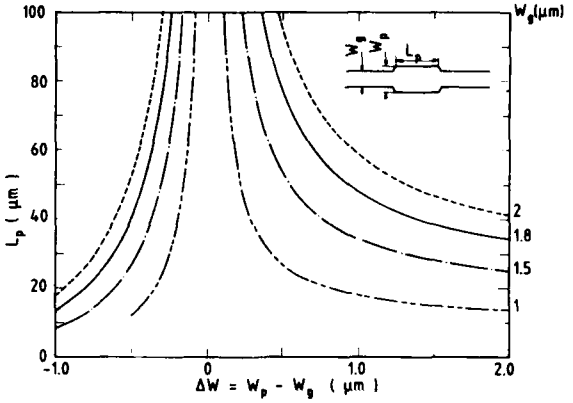


Fig. 3: Calculated relationship between the phase shifter length L_p and the stripe width widening (narrowing) of the stripe ΔW.

The wider stripe region of width W_p and length L_p=50µm acts as a phase shifter and has been designed to produce a $\pi/2$-phase shift /1,3/. The nominal stripe width of the mask is 2µm but, because of underetching, the actual uniform stripe width W_g is 1.8µm. To avoid parasitic reflections at the ends of the phase shifter, tapered regions were introduced to connect it smoothly to the uniform regions and the widening of the stripe $\Delta W = W_p - W_g$ was set to 1µm. Fig. 3 shows the calculated relationship between the phase shifter length L_p and the widening (narrowing) of the stripe ΔW. We notice that L_p is rather sensitive to W_g. For example, if, because of excessive etching, W_g decreases from 1.8 to 1.5µm, the optimum phase shifter length decreases from 50 to 35 µm and a 50µm-long phase shifter designed for W_g=1.8µm and $\Delta W = 1\mu m$ will produce a total phase-shift of 0.7π. Hence, the lasing properties of the resulting lasers will be severely degraded.

3-LASING CHARACTERISTICS

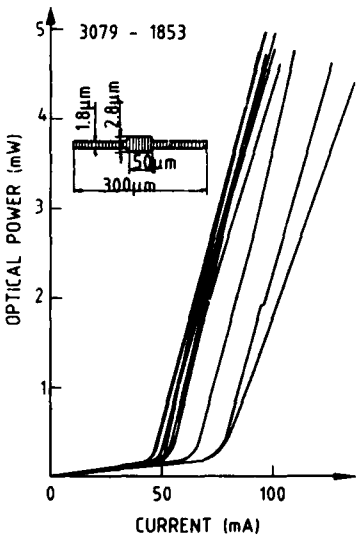


Fig. 4: Typical light-current characteristics of a PS-BRS.

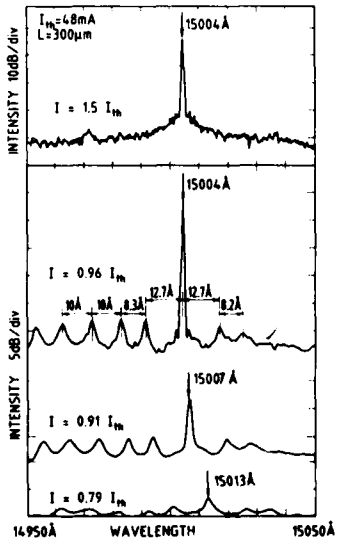


Fig. 5: Sub- and above-threshold spectra of a PS-BRS laser with an anti-reflection coated front facet

Fig. 4 shows the light-current characteristics of 12 unselected lasers measured from the same cleaved bar. Threshold currents range between 40 and 50 mA. Almost 90% of the mounted lasers exhibited CW single longitudinal mode (SLM) operation with a side mode suppression ratio greater than 30 dB.

Fig. 5 shows typical sub- and above-threshold spectra of a PS-BRS laser with an anti-reflection coated front facet. Above threshold, the laser exhibits mode hopping free SLM operation up to the optical power saturation due to the current leakage of the BRS structure. SLM operation is also observed with a main mode and side modes located symmetrically to the main mode. This symmetry demonstrates the effect of the phase shifter on mode selection and validates the stripe design.

4-CONCLUSION

We have reported on the design, technology and lasing characteristics of PS-BRS lasers. We have experimentally observed that the percentage of SLM operation is higher for such lasers that for conventional lasers with uniform width stripes.

5-REFERENCES

/1/ H. Soda, Y. Kotaki, H. Sudo, H. Ishikawa, S. Yamakoshi and H. Imai, IEEE J. Quantum Electron., QE-23 (1987) 804.

/2/ K. Utaka, S. Akiba, K. Sakai and Y. Matsushima, IEEE J. Quantum Electron., QE-22 (1986) 1042.

/3/ B. Broberg, S. Koentjoro, K. Furuya and Y. Suematsu, Appl. Phys. Lett. 47 (1985) 4.

/4/ D. Lesterlin, J. Charil, B. Rose, M. Gilleron, P. Correc and J.C. Bouley, Proc. ESSDERC'87 (Sept. 1987) 803.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

FABRICATION AND LASING PROPERTIES OF A NOVEL SINGLE LONGITUDINAL MODE 1.5 μm GaInAsP/InP DISTRIBUTED REFLECTOR (DR) LASER

K. KOMORI, H. SUZUKI, K.S. LEE, S. ARAI, Y. SUEMATSU, M. AOKI and S. PELLEGRINO*

Department of Physical Electronics, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-Ku, Tokyo 152, Japan
**C.S.EL.T., Centro Studi e Laboratori Telecomunicazioni SPA, Via Reiss Romoli 274, I-10148 Torino, Italy*

Résumé - Un nouveau laser, du type à single mode dynamique, avec réflecteur distribué dans les regiones active et passive a été realisé. On a identifié des intérèssantes proprietes comme un elevé rapport de soppression des modes satellites, un petit courrant de seuil et des proprietés d'émission asymmetrique.

Abstract - A new type of dynamic single mode laser, which has distributed reflector (DR) both in active and passive regions has been realized; attractive properties such as high sub-mode suppression ratio, low threshold and asymmetric properties have been detected.

Introduction

We report the fabrication process and preliminary results concerning a new dynamic single longitudinal mode (DSM) laser with the aim of high power operation and high single mode fabrication yield. Asymmetric DFB lasers in High/Low reflection configuration reported up to now and obtained by coating front and rear facets, have a DSM expectation value around 80% (1). Hence we developed a structure capable of high power, high yield DSM operation as well as suitable for monolithic integration, by joining a Distributed Bragg reflector to a DFB structure, in a Bundle Integrated Guide (BIG) (2) configuration to maximize the coupling between DFB and DRB regions. The theoretical analysis of the structure has been published elsewhere (3).

Structure and Fabrication

A schematic drawing of the structure is shown in fig.1, and in fig.2 the fabrication sequence is outlined. The island-type mesa process (4) was employed, during the second within the four total LPE growth, in order to obtain flat growth conditions of the GaInAsP (λ_g=1.3 μm) passive waveguide. High quality first order grating on AZ photoresist has been obtained employing an Ashing process at low oxygen pressure and RF power;the Ashing rate is shown in fig.3. The grating was transcripted using (HBr:HNO3:H2O=1:1:10) solution: the chemical characterization of the system and its etching properties will be described elsewhere. Finally, usual Buried Heterostructure was formed during the fourth LPE growth.

Results

Minimum CW threshold current of 25 mA was obtained for a 400 /um DFB, 200 /um DBR regions length, respectively. Since the stripe width was 3 /um, hte threshold current density was as low as 2 KA/cm2. In pulse conditions asymmetric emission properties were detected, as shown in fig.4, with a front to rear power ratio of 2.1, demonstrating good match between DFB emission and DBR reflectivity spectrum. Single longitudinal mode emission at 1.5 /um up to 3 times threshold (6 mW output power) were observed, as shown in fig. 5. The lasing wavelength temperature coefficient and characteristic temperature To were 0.1 nm/deg at 1.2 times threshold and 53 K,respectively (fig. 6). The large fixed single mode operation range shows that the laser is operating in DR rather than in DBR mode.

Conclusion

A new kind of DSM laser was developed; threshold current as low as 25 mA, asymmetric emission properties and sub-mode suppression ratio of 38 dB at 3 times threshold were recorded.

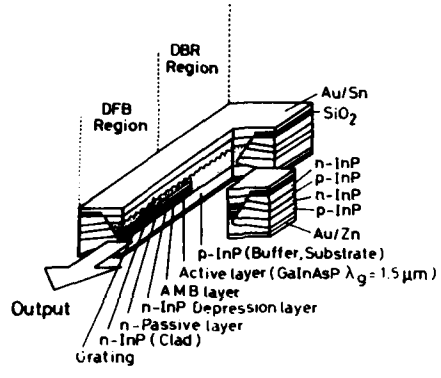


Fig. 1 - Distributed Reflector (DR) laser.

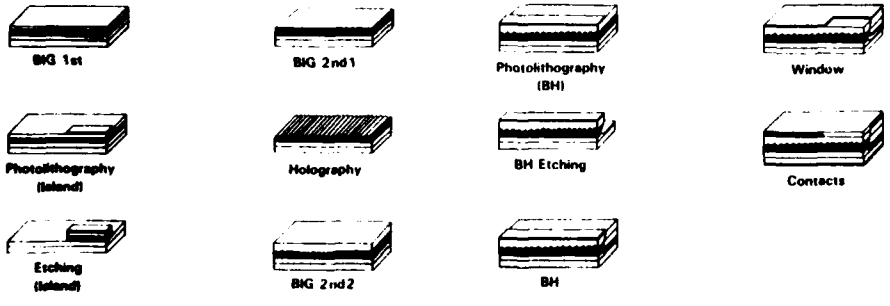


Fig. 2 - Fabrication Flow-Chart (read by columns).

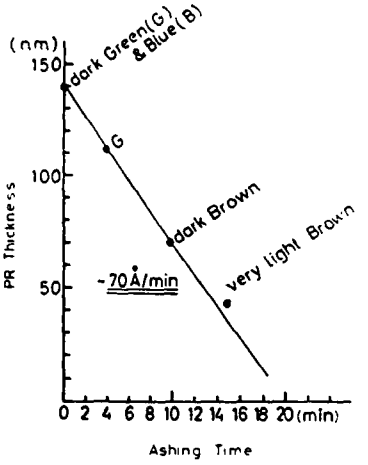


Fig. 3 - Ashing rate of AZ photoresist
oxygen pressure = 4*10-2 mbar
RF power = 30 W

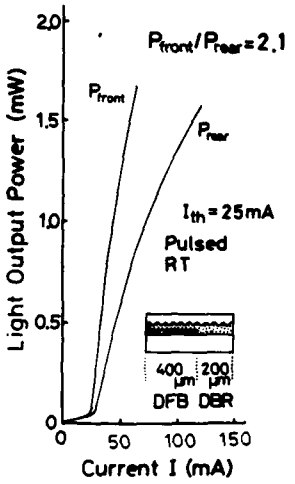


Fig. 4 - L-I curve for front-rear emission.

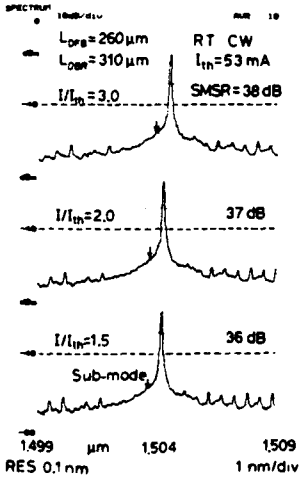


Fig. 5 - Typical lasing spectrum.

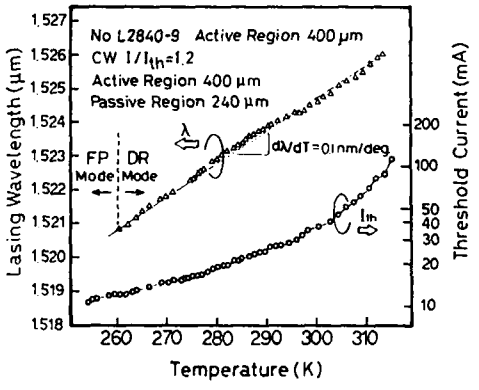


Fig. 6 - Temperature dipendence of lasing spectrum and threshold current.

References

(1) Henry, C.H. : IEEE J. Quantum Electron. QE-21,(1985), 1913
(2) Tohmori, Y.,et al. : Electron. Lett. 21,(1985), 743
(3) Komori, K., et al.:Trans. IEICE of JAPAN E71,(1988), 318
(4) Choi, I.H., et al. : Jpn. J. Appl. Phys., 26,(1987), L1593

SESSION 4

Session 4IP Room A : Invited papers

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988 C4-347

SILICON MOLECULAR BEAM EPITAXY : STATUS ; DEVICES ; TRENDS

E. KASPER

AEG Research Center, D-7900 Ulm, F.R.G.

Abstract - The single crystal growth, by molecular beam epitaxy, of materials compatible with silicon is described, and device applications are given. Process temperatures are decreased to 240°C, heterojunction devices are fabricated, strained layer Si/Ge superlattices are investigated, and silicon based monolithic integration of heterojunction devices and conventional devices is suggested.

1 - INTRODUCTION

Todays microelectronics development is governed by the shrinkage of lateral dimensions. As a consequence the technology is forced also to reduce vertical dimensions, and to decrease process temperatures. Other material systems, like GaAs, are competing with Si, especially for high speed devices and for optoelectronic applications. The electronic functions are confined to a rather thin surface layer on a several hundred micron thick substrate. With increasing tendency this surface layer is fabricated by epitaxy - a single crystal deposition process at temperatures well below the melting point. The substrate serves as stable handling support which gives orientation information to the epitaxy layer grown on it. In a traditional way a uniform epitaxy layer is grown, which serves as a thin substrate to the following device definition processes. The well proven chemical vapour deposition (CVD) process fulfills nearly perfectly the requirements of this simple layer structure as long as layer thickness is well above the micron.
The introduction of a new method as silicon molecular beam epitaxy (Si-MBE) is justified if besides better solutions to remaining problems new opportunities for device physics and fabrication will be opened. Si-MBE offers submicron layers, extremely abrupt junctions, complex dopant profiles for complete realization of the vertical device structure, processing at temperatures low enough to suppress totally solid state diffusion of group III/V elements, and heteroepitaxy of Si-semiconductor, Si-metal and Si-insulator structures. Indeed, one of the objectives of this article is to stimulate device designers and manufacturors not to be satisfied with a simple layer on a substrate but to demand realization of vertical device structures with high degree of design freedom.
The reader is referred to three books /1 - 3/ which allow an in depth sight of the main topics treated here. Individual references are only given for very recent results. A good opportunity for more information is given to European scientists with the 3rd Int. Symposium on Si-MBE (E-MRS Spring Meeting, May, 29th - June, 2nd, Strasbourg, 1989). Call for papers can be obtained from the author.

2 - STATUS

A rather small community (roughly fifty groups around the world) from universities and industrial research laboratories define the technical status of SI-MBE.

2.1 - The method

MBE as physical vapour deposition (PVD) process refines the evaporation method by condensation of the molecular beams on heated surfaces within an ultrahigh vacuum (UHV) environment (Fig. 1). Of essential importance are the cleanliness of the surface and an understanding of the adsorption of atoms on the growing surface. The cleanliness of the surface is determined by an in situ cleaning step before epitaxy and by the clean ultrahigh vacuum surrounding. Atoms from the molecular beam impinging on the surface are bonded to the surface with a desorption energy lower than the binding energy of the crystal. The adatoms diffuse rather easily on the surface creating a twodimensional

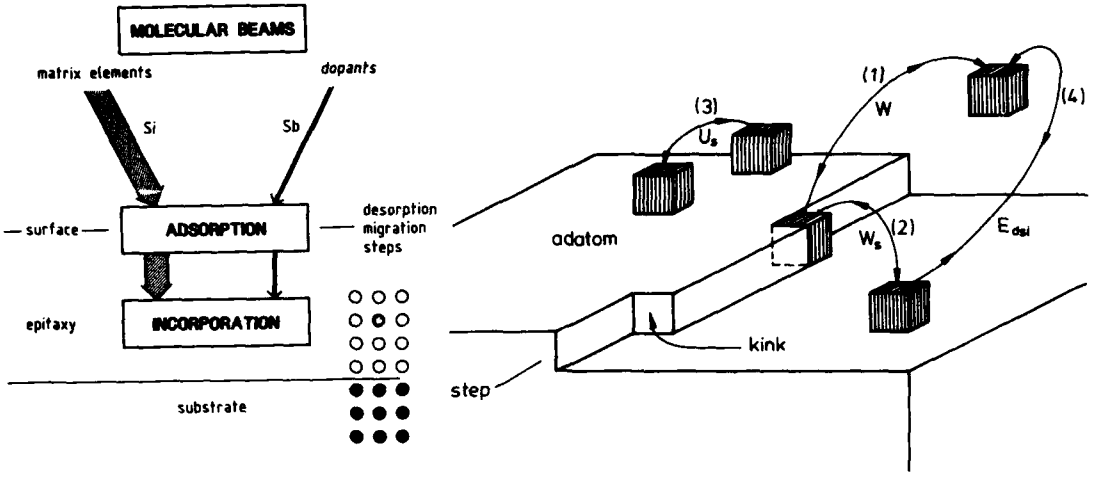


Fig. 1: Molecular Beam Epitaxy (MBE). Adsorption and incorporation of beam atoms (left side). Surface model (right side). Desorption energy $E_{dsi}(4)$, binding energy W(1) and diffusion barrier $U_s(3)$.

adatom gas. Adatoms are incorporated at surface steps which move forward by this adatom catching. Vertical growth proceeds by lateral movement of atomic steps.

2.2 - The apparatus

Within the ultrahigh vacuum pumped growth chamber are three subsystems (table 1). Matrix and

Table 1: Subsystems of an Si-MBE growth chamber

Subsystem	Sources	Wafer Heater	In situ Monitoring
Typ. examples	Effusion cell, E-gun evaporator, low energy ion implanter	graphite meander (heated by electrical current)	Quadrupol mass spectrometer (QMS), Reflection high energy electron diffraction (RHEED)

dopant elements are evaporated from the sources equipped with beam shutters. Fig. 2 shows an effusion cell used for dopant evaporation. As dopants from effusion cells Sb for n-doping and Ga,B for p-doping are used. Low energy ion implanters use preferably As for n-doping and B for p-doping.

The growth chamber is or can be connected with further UHV-chambers, e.g. load lock, preparation chamber, analyses chamber. Many of the equipments are home made with differing design. Within the last years equiment manufacturers (table 2) improved their apparatus considerably.

Table 2: Si-MBE equipment manufacturers

Anelva	Atomika	Perkin Elmer	Riber	Varian	VG
(Japan)	(FRG)	(USA)	(France)	(USA)	(England)

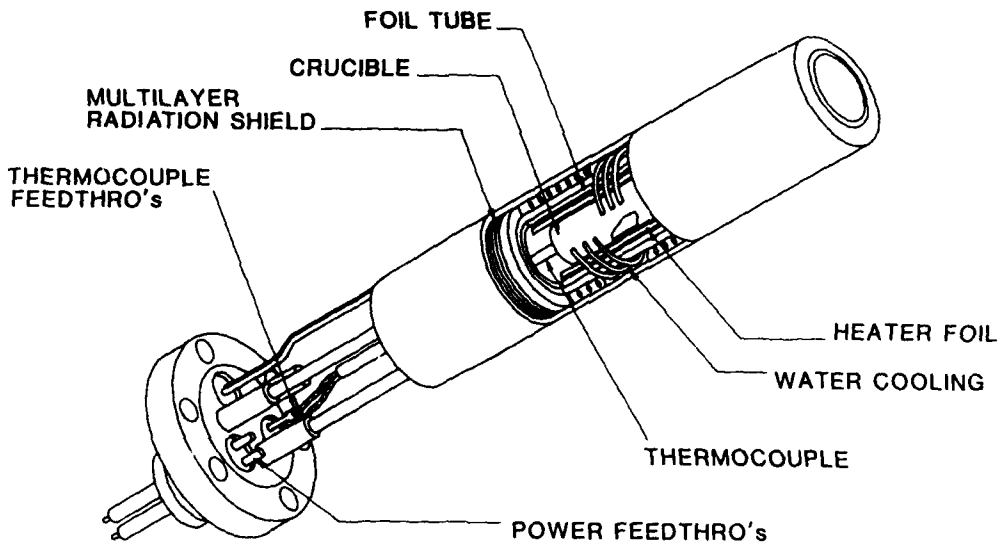


Fig. 2: MBE-equipment: Typical effusion cell design. [1,2]

2.3 - Doping techniques

Coevaporation of the dopant material is of limited use in Si-MBE because of desorption and segregation phenomena. Therefore ion assisted doping techniques were developed which may be interesting also for outside MBE applications. Low energy ion implantation operates with ion beams which are deaccelerated to 1 to 3 keV energy and focussed to an entrance port to the MBE chamber. With the help of an electrostatic deflection plate the beam is scanned over the wafer.
Doping by secondary implantation (DSI) utilizes the Si-ions always present in electron gun evaporation (typ. 0.1% of the atomic beam density). The Si-ions are accelerated by an applied substrate

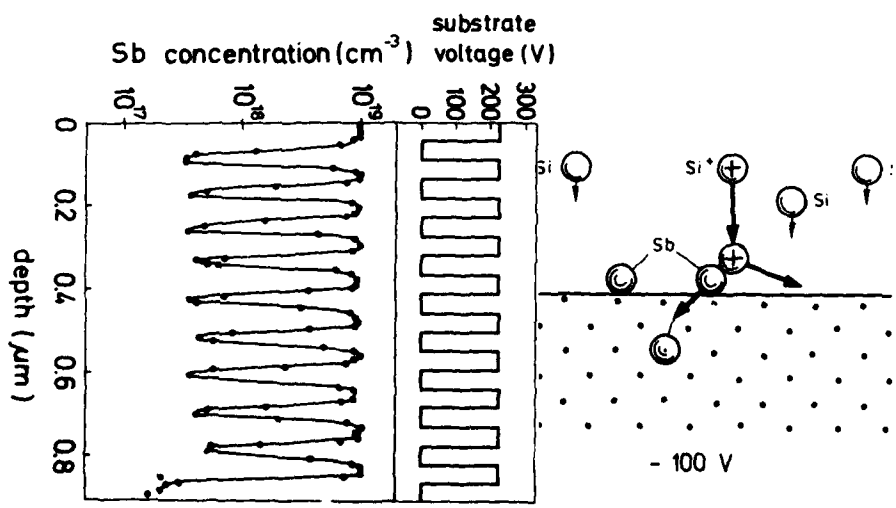


Fig. 3: Doping by secondary implantation (DSI). Dopant profile (Sb) and corresponding substrate voltage (left side). Scheme of process (right side).

voltage (typ. 100 V - 1000 V) toward an preadjusted Sb-adatom layer (typ. 0.1 monolayer (ML)) on the wafersurface (Fig. 3). The Si-ions knock on the Sb-adatoms and implant them a short distance (~1 nm). Very sharp transitions (Fig. 3) and δ-doping spikes can be created by this technique.

2.4 - Heteroepitaxy

Semiconductor-silicon, but also metal-silicon and insulator-silicon couples can be grown by Si-MI (table 3). Silicides, calcium fluoride and silicon germanium are the most prominent partners of

Table 3: Si-based heteropitaxy partners

material class	semiconductor	metal	insulator
partner	SiGe, Ge	$NiSi_2$,	CaF_2
	GaAs	$CoSi_2$	

silicon based heterosystems. GaAs on Si is usually grown in an GaAs-MBE equipment although the in situ cleaning is Si-MBE typical. A connection between Si-MBE and III/V-MBE equipments will allow growth of a wide variety of materials.

3 - DEVICES

Device classes are selected which are not only DC characterized.

3.1 - Integrated bipolar transistors

In one approach the simple uniformly doped layer of roughly one micron thickness was varied systematically in thickness. With the given layout and process sequence a maximum transit frequency of 7 GHz was obtained with - 1.2 µm layer thickness. The transit frequency decreased to 6 GHz and 4 GHz for layer thicknesses of 0.9 µm and 1.7 µm respectively. 2.5 GHz-Frequency dividers were fabricated with high yield from Si-MBE layers.

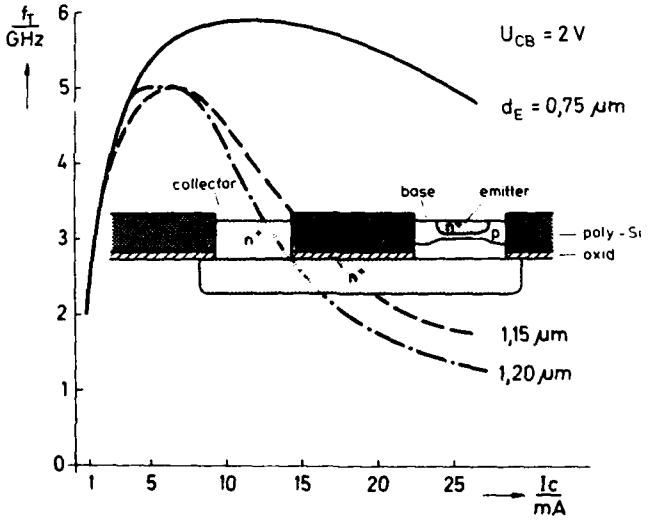


Fig. 4: Integrated bipolar transistor from differential epitaxy layers. Frequency f_T versus current f_c. Inset: transistor structure.

A more interesting approach demonstrated lateral definition of device structures by growth into oxide windows. This growth mode is called differential epitaxy (Fig. 4). The first results were encouraging resulting in transit frequencies comparable to the above mentioned data. There was, to my knowledge, no real trial to fabricate the whole structure of a high frequency transistor by Si-MBE.

3.2 - Silicon mm-wave integrated circuits (SIMM WIC)

Up to now the brightest success with Si-MBE structures was obtained for discrete mm-wave (> 30 GHz) devices. The first commercial Si-MBE device /4/ - a 90 GHz-Impact Avalanche Transit Time (IMPATT) diode - is the existing most powerful semiconductor device in this frequency range, and has the potential to occupy oscillators from 60 GHz to 300 GHz.
A real break through was the monolithic integration of the IMPATT into a 75 GHz-oscillator /5/. In the meanwhile mm-wave receivers and transmitters with integrated antennae were realized (Fig. 5) pushing Si-technology to the 100 GHz regime and beyond this border.

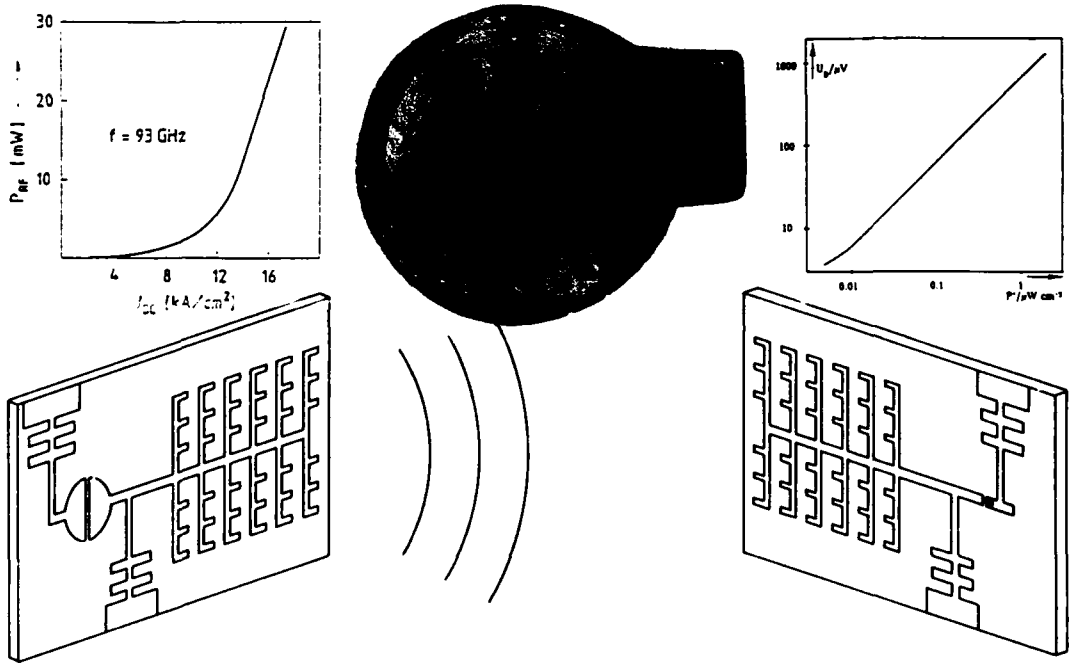


Fig. 5: Si-based integration of 100 GHz devices with microstrip oscillators, receivers and antennae (SIMMWIC). Scheme (below), outputpower (above right), receiver voltage (above left) and a receiver chip (above, middle) are shown.

3.3 - SiGe heterostructure devices

Heterostructures offer the designer a new freedom because of the unique properties not available in homogeneous materials. A nice example is J.F. Luy's /6/ work on a MITATT diode with improved noise properties. But the classical examples are the n-channel MODFET and the waveguide APD optical detector (Fig. 6) fabricated by H. Dämbkes and T.P. Pearsall from E. Kaspers and J. Beans material, respectively. A review of the recent work on SiGe heterostructure devices is given in /7/.

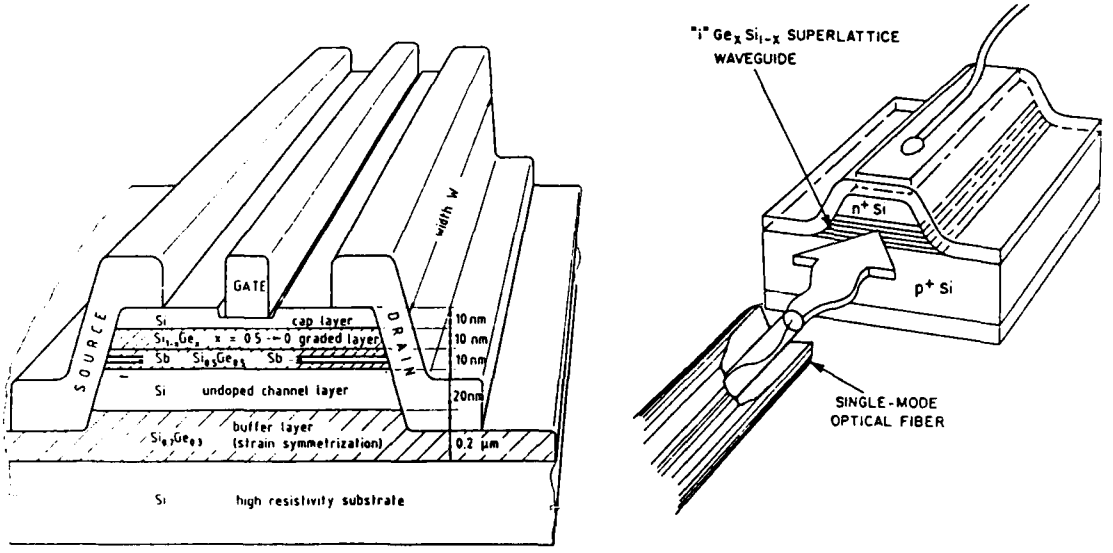


Fig. 6: N-channel MODFET (left side) and optoelectronic (1.3 µm) APD-receiver (right side) based on SiGe heterostructures on Si substrates.

4 - TRENDS

MBE fits very well into the general trend of semiconductor processing toward low process temperatures and submicron structure dimensions. Silicon on insulators (SOI) are gaining increasing importance. This paper will discuss MBE-typical solutions to SOI structures. A very large contribution to the performance advance of III/V-devices is given by the systematic exploitation of heterojunction properties. The exploitation of heterojunction effects in silicon based devices attracts attention for extension of the frequency range and for novel concepts. As examples are treated for the former the heterobipolar transistor (HBT) and for the later the modification of band gap character by strained layer superlettices (SLS). Monolithic integration of such heterostructure and superlattice devices with silicon based conventional IC's is proposed to be one of the innovative technology directions of the 90's.

4.1 - Low process temperatures

Si-MBE has reduced the growth temperatures of epitaxy layers far below the temperatures usually applied by the CVD-process. Standard growth temperatures for the Si-MBE process are between 550°C and 750°C. This is low enough to suppress completely the solid state diffusion of the common dopant elements from group III and group V of the periodic system of elements. Recent investigations are directed toward a further reduction of the process temperatures. A lower limit of single crystal growth of silicon seems to be reached between 140°C and 240°C. The most important results of such investigations are in contradiction to earlier results obtained from CVD experiments.

- On a single crystalline substrate epitaxy is obtained at all temperatures above temperature T_{Epi}, which is slightly dependant on growth rate. For very thin layers single crystal growth is also possible at temperatures below T_{Epi} but with decreasing crystal perfection as thickness increases.
- The transition at T_{Epi} is between the single crystalline phase and the amorphous phase. Poly crystalline layers are created only if ordering information is failing either by an amorphous layer, e.g. SiO_2, or by a contamination of the surface.
- The condensation coefficient is unity at the Si-MBE growth temperatures. That means all material from incoming molecular beams condenses at the surface with negligible reflection or desorption events.

4.2 - Silicon on insulator (SOI)

The creation of silicon on insulator (SOI) structures does not belong to the original strengths of

Si-MBE. But now, at least three ways are offered to create SOI-structures:
- Epitaxy on porous silicon. The low growth temperatures avoid a change of the properties of porous Si, which is a problem at process temperatures above 900°C. The unchanged porous oxide can be oxidized with high lateral oxidation rates from windows etched into the epitaxy layer. Single crystal islands or stripes are lying on an oxide.
- Lateral solid phase epitaxy (L-SPE). High quality solid phase epitaxy is obtained by a slight variation of the MBE-process. At room temperature the impinging molecular beams condens to an amorphous layer which can be eiter n- or p-doped up to levels of $10^{20}/cm^3$. This amorphous layer recrystallizes from the interface at temperatures around 600°C. Lateral solid phase epitaxy (L-SPE) is obtained if the deposition takes place on a prepatterned SiO_2 layer with windows (seed area) to the substrate. This elegant method has the disadvantage of overgrowth of only a few microns. Maximum overgrowth width of 50 µm was obtained with highly doped layers.
- CaF_2/Si heteroepitaxy. This most promising technique is based on the single crystal growth of different material stackings (heteroepitaxy). Generally for heteroepitaxy, problems with lattice mismatch, with different thermal expansions, with surface morphology (threedimensional nucleation) and with quality of interface structures (dislocations, stacking faults, twins, antiphase boundaries) have to be overcome. From several groups encouraging results /8/ were obtained with the CaF_2/Si heterosystem. CaF_2 is a good insulator with a small lattice mismatch to silicon.

4.3 Heterobipolar transistor (HBT)

It is predicted, and in the III/V material system already proven, that bipolar transistors would benefit from heterotransitions between the base and emitter regions. For silicon based transistors two concepts are now under actual consideration.

(i) The low bandgap base HBT (Fig. 7) with a thin, strained SiGe base layer between Si collector and Si emitter. Since end of 1987 at least five companies (AEG, AT&T, IBM, NEC, British Telecom) are developing this transistor typ /9, 10/ with a predicted over 40 GHz operation of a silicon based device.

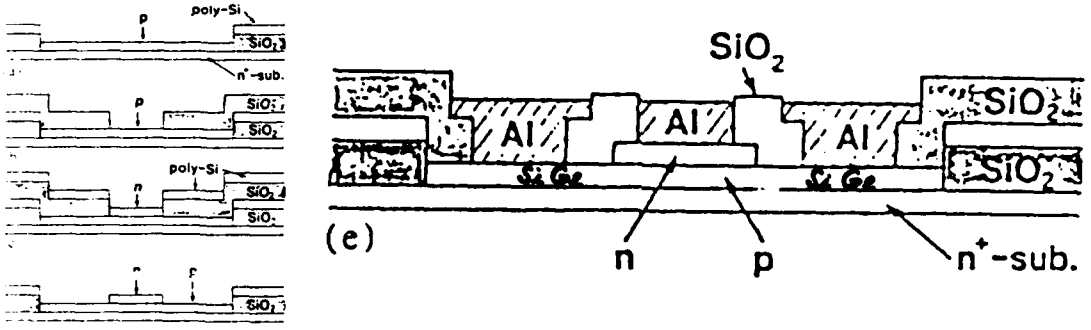


Fig. 7: SiGe base HBT. Structure and fabrication steps of the NEC transistor /9/

(ii) The metal base transistor utilizes ultrathin metal layers cladded by Si layers as vertical transistor structures. As metal layers $NiSi_2$ and $CoSi_2$ are used. Both silicides exhibit low lattice mismatch to silicon.

4.4 - Industrial equipment

The main body of Si-MBE work is now done under research conditions. Therefore, the equipment manufacturers are mainly oriented to this nonindustrial market. The increasing number of industrial customers will certainly strengthen efforts toward industrial equipment. Cooperation between equipment manufacturers and semiconductor/electronics industry is necessary. Such cooperation leads to the development of a batch process for Si-MBE /11/, to process control based on optical in situ monitoring /12/, and to an MBE machine /13/ based on the design principle of highest simplicity.

The design principle of highest simplicity /2/ requires that an industrial equipment contains only subsystems which have been proven to be absolutely necessary. Such an equipment will allow efficient, cost effective, reliable operation. Fig. 8 shows a two chamber equipment with a storage chamber containing a cassette type 25 wafer magazin for up to 150 mm wafers, an automated wafer transfer to

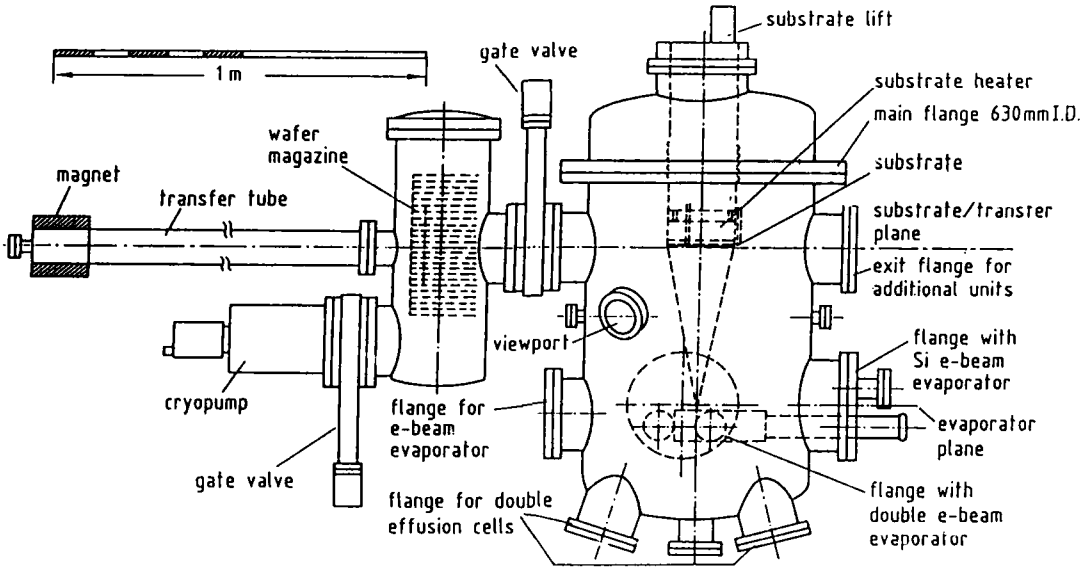


Fig. 8: Si-MBE equipment for single wafer processing of up to 150 mm diameter wafers /13/. Design pricriple of highest simplicity /2/.

the growth chamber, a growth chamber with mass spectrometer controlled e-gun evaporation of the matrix beams (Si, Ge).

4.5 - Strained layer superlattices

Periodic stackings of heterostructures (superlattices) open the possibility to tailor material properties of semiconductors (artificial semiconductor). Material partners with lattice mismatch can be combined without crystal perfection degradation, if the superlattice period is thin (strained layer superlattices - SLS). The most prominent Si-based SLS system is the SiGe alloy with up to 4 % lattice mismatch. Adjustment of strain, mobility enhancement of holes and electrons, strain induced band ordering and phonon zone folding effects were already demonstrated in this system /14/.
Most spectacular predictions promise for ultrathin Si/Ge superlattices a strong enhancement of direct band gap transitions. This would allow novel optoelectronic functions in a silicon based material system.
A monolithic integration of heterostructure or superlattice devices with conventional IC's on a silicon substrate has the potential to influence future electronics in a dramatic way /Fig.9/. The silicon substrate would give the stable, perfect substrate, the conventional IC would define the complexity level, and the heterostructure devices would define the ultimate performance (speed, combination of microelectronic and optoelectronic functions). The heterostructures will occupy only a small part of the chip area, and they will be defined by low temperature processes after the conventional IC part. For the solution of the problems connected with, research on mismatched structures, low temperature processing and heterostructure stacking with high perfection is needed.

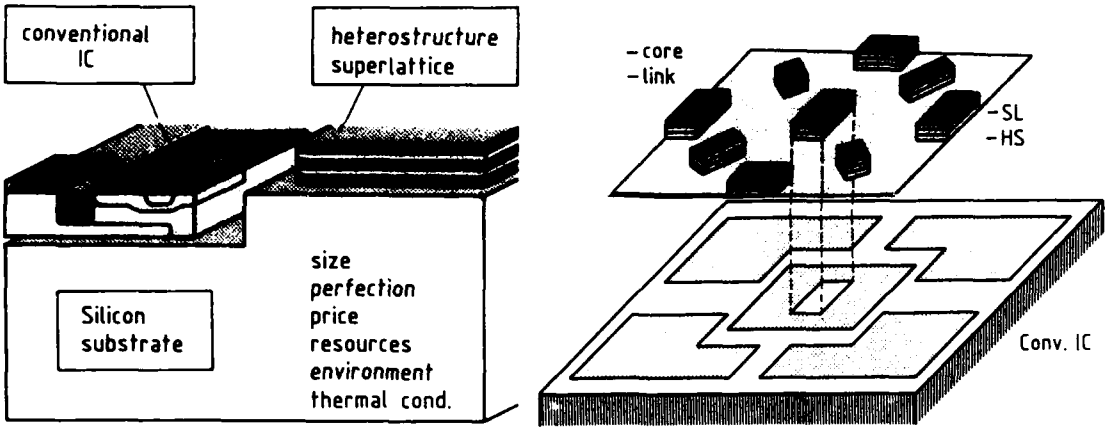


Fig. 9: Si-based monolithic integration of conventional IC with superlattices / heterostructure devices (left). A small part of the chip area (core regions, intrachip links) is occupied by the heterostructures (right).

/1/ The practice and technology of MBE, ed. by E.H.C. Parker, Plenum Press, New York, 1987
/2/ Silicon molecular beam epitaxy, ed. by E. Kasper and J.C. Bean, CRC Press, Boca Raton, Florida, 1988
/3/ Proc. 2nd Int. Symp. SI-MBE, ed. by J.C. Bean and L. Schowalter, Electro-Chem. Soc., Pennington, USA 1988
/4/ Data sheet. W-Band IMPATT, AEG, P.O. Box 3580, D-7900 Ulm, FRG
/5/ J.F. Luy, F.M. Strohm, J. Büchler, Microwave and Optical Technology Letters (MOTL) 1,117 1988
/6/ J.F. Luy, will be published
/7/ E. Kasper, Heterostructures on Silicon: A Step Further with Silicon, ed. by Y. Nissim, K. Ploog, and E. Rosencher, will be published
/8/ J. Phillips, J.L. Batstone and J.C. Hensel, MRS Spring Meeting 1988, Symposium A
/9/ T. Tatsumi, H. Hirayama, N. Aizaki, Appl. Phys. Lett. 52, 895 (1988)
/10/ S. Iyer, H. Tamkin, H. Dämbkes in /3/,
P. Narozny, will be published
/11/ VG Semicon / AT&T (Bell, J.C. Bean)
/12/ Riber / AEG / GEC, ESPRIT project no. 305
/13/ AEG / MPI f. Festkörperforschung / Atomika, German ministry of research and development contract
/14/ E. Kasper, H. Kibbel H. Jorke, H. Brugger, E. Friess and G. Abstreiter, Phys. Rev. B, will be published

Session 4A Room A : Bipolar transistors : characterization

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

MEASUREMENT OF MINORITY-CARRIER LIFETIME AND INTERFACE RECOMBINATION VELOCITIES IN P-I-N DIODES, FROM HIGH FREQUENCY RESPONSE OF A BIPOLAR JFET STRUCTURE[1]

G. VITALE and P. SPIRITO

Dept. of Electronic Engineering, University of Napoli, Via Claudio 21, I-80125 Napoli, Italy

Abstract - A new method to measure the minority-carrier recombination lifetime in the low-doped layer of a p-i-n diode, and its emitter recombination current, is presented.
The method is based on the measurement of the cutoff frequency of a three terminal device structure, similar to a vertical JFET, that incorporates the diode under test. The paper displays the basic theory of the measurement and some experimental results.

1 - INTRODUCTION

This paper presents a new method to measure the minority-carrier recombination lifetime in the low-doped layer of a p-i-n diode, as well as the recombination current in the p^+ layer of the diode.

The method is based on a three terminal device structure, that incorporates the p-i-n diode under test, as reported in Fig.1. The test structure is similar to a vertical Junction Field Effect Transistor, in which the p-i-n diode under test can be either the Gate - Drain or the Source -Gate diode. A similar device structure was first proposed in [1] as a means to measure the recombination velocity at the high-low transitions that appear in the device structure, and subsequently used in [2] to characterize the BSF of a Solar Cell. The method, which is being presented here, extends the characterization of p-i-n diode (or a Back Surface Field Solar Cell), by using the same test structure to measure also the lifetime of the low-doped layer and the recombination into the p region.

For the purposes of the present analysis, the test structure must be operated with the gate forward biased, with respect to the source, thus inducing an electron-hole plasma within the low-doped layer of the device. If a small-signal current is added between the gate and the source, an amplified current is detected at the drain. A measurement of the cutoff frequency of the device, as a function of the d.c. bias current, allows the simultaneous determination of the epilayer lifetime and of the recombination velocity associated with the p^+ layer of the p-i-n diode.

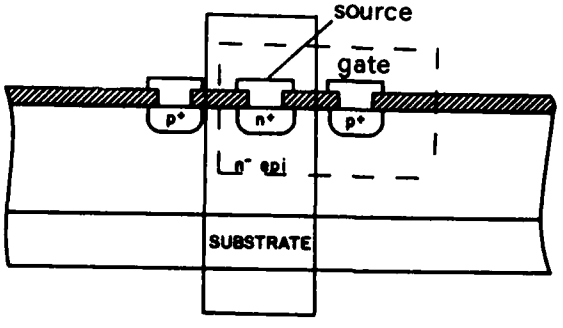


Fig.1 - The structure of the test device. The boxes indicate the subdivision of the test structure into the $n^+ n\, n^+$ and the $p^+ n\, n^+$ structures.

[1] This Work was Supported by ENEA under CNR-PFE2 Contract no. 170/87

2 - THEORY OF THE MEASUREMENT

A small-signal analysis of the test structure, has been carried out by solving the fundamental transport equations, linearized for small-signal. To that purpose, the device has been splitted into two, one-dimensional structures, as shown in Fig.2 and 3. These are: (1) The n^+-n-n^+, source-drain structure and (2) the surface p^+-n-n^+ (gate-source) structure. When the latter is forward biased, an electron-hole plasma is induced into the epilayer, within a thickness, X_{1o} which is controlled by the source-drain bias. From the d.c. analysis [3], it turns out that X_{1o} is given by:

$$X_{1o} = \frac{2\, q\, D_n\, (p_{so} - N_D)}{I_{do}\, A_d} \qquad (1)$$

where: p_s is the hole density at the source, N_D is the epilayer doping, A_d is the device area, I_d is the train current, the subscript o indicates d.c. quantities.

In the analysis, the source - drain structure is further subdivided into two parts, namely the conductivity modulated region, close to the source transition, and a drift region. The gate - source diode has been solved by summing the impedances of both the p^+-n and the n-n^+ transitions and by accounting for the recombination within the epilayer, the latter being a function of the thickness of the conductivity modulated region, X_1 The two structures are bound together by setting a common boundary condition both for the static and the dynamic equations. Finally the small - signal impedances, Z_{21} and Z_{22} have been calculated by integrating the electric field over the various regions to obtain the voltage $\tilde{V}_d$ and by recalling that:

$$Z_{21} = \left.\frac{\tilde{V}_d}{\tilde{I}_g}\right|_{\tilde{I}_d=0} \qquad (2a); \qquad Z_{22} = \left.\frac{\tilde{V}_d}{\tilde{I}_d}\right|_{\tilde{I}_g=0} \qquad (2b)$$

The small-signal, common source current amplification, h_{fs} has been obtained as: $h_{fs}=-Z_{21}/Z_{22}$.

If it is further assumed that the transport within the drift region is Ohmic, the general expression for the small-signal drain voltage is:

$$\tilde{V}_d = V_T \frac{\tilde{p}_s}{p_{so}} - j\omega\tau_1 V_T\left(\frac{\tilde{p}_s}{N_D} - \frac{\tilde{I}_d}{3I_{do}}\frac{p_{so}}{N_D}\right) + \frac{I_{do}\,\tilde{p}_{so}}{\sigma\, A_d\, s_o} + \frac{\tilde{I}_d W}{\sigma\, A_d} \qquad (3)$$

where: $\tau_1 = X^2_{1o}/4D_n$ is the transit time into the conductivity modulated region, I_d is the drain current, p_s is the hole density at the source, W is the epilayer thickness σ is the epilayer conductivity, the tilde indicates small-signal quantities.

It is worth to point out that the transport regime of the drift region does affect the low-frequency value of h_{fs} but it doesn't change its cutoff frequency.

The current at the input source-gate diode, has been derived by taking into account the effects of transit time between source and gate, the recombination within the epilayer, including the effects of the floating boundary, X_1. The recombination currents at the source and gate layers have been modeled by means of the recombination factors:

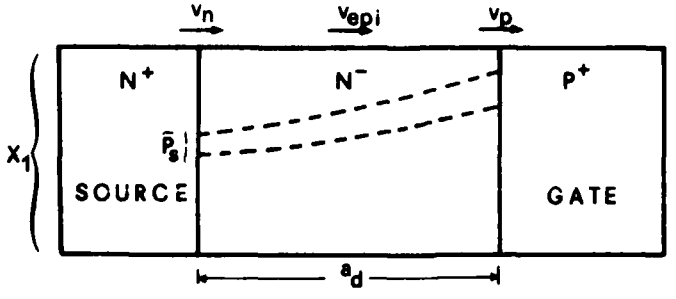


Fig.2 - The $p^+ n\, n^+$ structure and the minority-carrier distribution across the epilayer

$$\alpha_s = S_s A_s / N_D \; ; \; \alpha_g = S_g A_g / N_D \qquad (4)$$

where: A_s (A_g) are the source (gate) areas, S_s (S_g) are the source (gate) recombination velocities, that are functions of the minority-carrier transport properties of these high-doped layers. It is assumed that the impedance associated with the transport within these layers enters into the model at frequencies much higher than those of interest in the present analysis.

With these assumptions ,the gate current at the source interface is:

$$\tilde{I}_g = q\,\tilde{p}_s \frac{\dfrac{A_d X_{1o}}{\tau}(1 + j\omega\tau) + 2\,\alpha_g p_{so} + 2\,\alpha_s\, p_{so}\,(1 + j\omega\tau_D)}{1 + \left[1 - \dfrac{A_d\, X_{1o}}{2\, a_d\, A_g} - \dfrac{2\,\alpha_g\, p_{so}\,\tau}{A_g\, a_d}\right] j\omega\tau_D} \qquad (5)$$

where: τ is the epilayer minority carrier lifetime, τ_D is the diode transit time

From Eq.3, Z_{22} is obtained as the ratio $\tilde{V}_d/\tilde{I}_d$ with $\tilde{p}_s=0$. The latter condition is in fact equivalent to: $\tilde{I}_g=0$. Similarly, if $\tilde{I}_d$ is set to zero, in Eq.3, and the ratio of Eq.1 to Eq.3 is taken, the transfer impedance, Z_{21} is obtained.

The general expression for h_{fs} can be written as:

$$h_{fs} = h_{fo} \frac{(1 + j\omega/\omega_z)\,(1 + j\omega/\omega_{zd})}{(1 + j\omega/\omega_o)\,(1 + j\omega/\omega_p)\,(1 + j\omega/\omega_{pd})} \qquad (6)$$

where: $$h_{fo} = \frac{2\, V_T\, \sigma\, A_d}{q\, N_D\, W \left[\dfrac{A_d\, X_{1o}}{\tau} + 2\, p_{so}(\alpha_s + \alpha_g)\right]} \qquad (7)$$

$$\omega_o = \frac{1}{\tau} + \frac{I_{do}(\alpha_s + \alpha_g)}{q\, D_n\, A^2_d} \qquad (8)$$

The other zeroes and poles lie far above ω_o over a wide range of device parameters,thus allowing to simplify h_{fs} as:

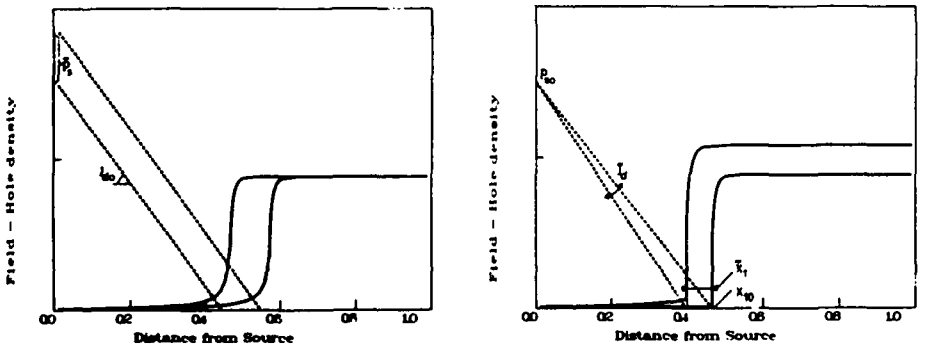


Fig.3 - Field (solid lines) and hole (dashed) distributions for constant drain current (left) and for constant gate current (right) between source and drain.

$$h_{fs} = \frac{h_{fo}}{(1 + j\omega/\omega_o)} \qquad (9)$$

From Eq.8, it turns out that a plot of ω_o vs. I_{do} is a straight line whose intercept with the ω axis equal to $1/\tau$ and a slope:

$$\frac{d\omega_o}{dI_{do}} = \frac{\alpha_s + \alpha_g}{q\, D_n\, A^2 d} \qquad (10)$$

is proportional to the recombination velocities of the Source and Gate layers. Fig.4 displays the experimental determination of these quantities, by using the method outlined above. The Base layer is n-type epitaxial, 2 10^{14} cm^{-3} doped. The measured lifetime is 2.2 μsec.

The separation of the two components of the recombination term, can be carried out by measuring independently the Source recombination velocity, by using the method reported in Ref.1. From Fig.4, the recombination velocities are: S_S = 3 cm/sec, S_G = 2.2 cm/sec.

By a proper design of the test structure, namely by making either the source or the gate area much larger than the other, one of the recombination terms can be made to dominate over the other, thus improving the accuracy of the measurement.

REFERENCES

[1] S.Bellone et al. "Measurement of the Effective Recombination Velocity of Semiconductor High-Low Transitions" IEEE Trans. Electron Devices, Vol.ED-32, pp.1771-1775 (1985)
[2] T.Daud and F.A.Lindholm "Direct experimental determination of voltage across high-low junctions" J.Appl.Phys., Vol.59(1), pp.285-287 (1986)
[3] S.Bellone et Al.:" A quasi-one-dimensional analysis of vertical FET devices operated in the bipolar mode" Solid-St. Electron. vol.26, pp.403-413 (1983)

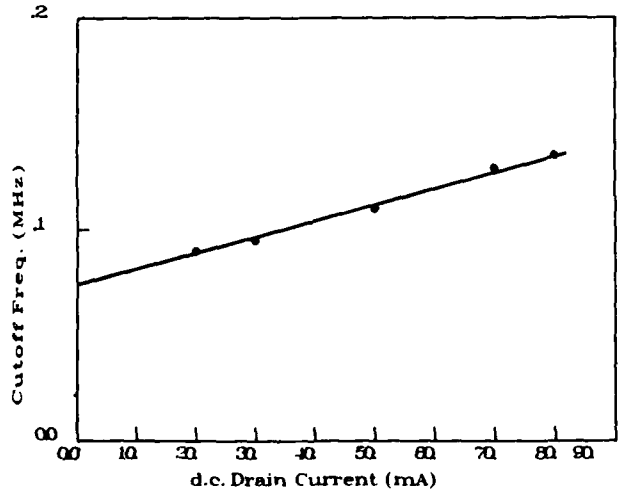


Fig.4 - The experimental determination of minority carrier lifetime and of gate recombination velocity.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

JFET FOR COMPLETELY DEPLETED HIGH RESISTIVITY SILICON

V. RADEKA, P. REHAK, S. RESCIA, E. GATTI*, A. LONGONI*, M. SAMPIETRO*, G. BERTUCCIO*, P. HOLL**, L. STRUDER** and J. KEMMER***

Brookhaven Nat. Lab., Upton, NY 11973, U.S.A.
**Politecnico di Milano, Dipartimento di Elettronica e Centro di Elettronica Quantistica e Strumentazione Elettronica CNR, P.za Leonardo da Vinci 32, I-20133 Milano, Italy*
***Max-Planck-Institut, Fohringer Ring 6, D-8000 München, F.R.G.*
****TU München, 8048 Garching and MBB GmbH, Postfach 801149, D-8000 München 80, F.R.G.*

Résumé - On présente le projet d'une nouvelle classe de transistors à effet de champ à jonction JFETs, conçue pour l'intégration sur le wafer à haute resistivité d'un détecteur de radiation complètement en dépletion, et les résultats expérimenteaux obtenus. Ces dispositifs vont rendre possible l'intégration d'un préamplificateur à faible bruit directement sur le détecteur et par conséquent une amélioration de la résolution. Ces nouveaux JFETs ont été projetés pour obtenir un g_m/C_G de 500 MHz et on a mesuré pour eux un g_m/I_D de 1/850 mV^{-1} et un pinch-off de 1.5 V.

Abstract - The design and the experimental results of a new class of JFETs, suitable for integration on the high resistivity wafer of a fully depleted radiation detector, are presented. These devices will make possible the integration of a low noise preamplifier directly on the detector, improving the achievable resolution. The new JFET has a measured g_m/I_D of 1/850 mV^{-1}, a pinch-off of 1.5V and a designed g_m/C_G of 500 MHz.

1 - INTRODUCTION

During the last few years, there has been an increasing interest in the integration of front end electronic circuits onto some new type of fully depleted large area detectors for position and energy measurements of ionizing radiation (Semiconductor Drift Chambers, fully depleted CCD, etc.) /1,4/.
As it is well known, the signal-to-noise ratio for amplitude measurements, as well as the timing resolution, increases by lowering the output capacitance of the detector and matching to this capacitance (which includes stray capacitances of connection to the preamplifier) the capacitance of the input device of the preamplifier /5,6/.
These detectors can be built with a very low output capacitance of the order of tens of femtofarad practically independent of their active area /7,9/.
The energy and position resolution achievable with these detectors have been limited, so far, by the excessively high input capacitance of even the best commercially available descrete field effect transistor.
To fully exploit the unique property of very small capacitance of these detectors, it is compulsory to build the input preamplifier on the same chip of the detector. In this way it is possible to reduce to negligible values the strays of the connections and to design an input FET with a capacitance matched to the detector one.

2 - DESIGN SPECIFICATIONS OF THE INTEGRATED JFET

The peculiar characteristics of the detectors impose some design constraints to integrated JFETs /10/ :

- Implementation on the detector wafer (high resistivity, detector grade, n-type silicon, 2Kohm.cm, <111> surface orientation).
- Implementation with the limited production steps used for the detector. In effect in order to keep the leakage current low, only relatively low temperature (800 C) processes in the detector production are allowed, with no diffusions and no polysilicon process presently available.
- requirement of operation on a fully depleted n-type substrate.
- surface connection of the p+ gate with the n+ output anode of the detector.

3 - STRUCTURE OF THE JFET

A schematic cross section of the device is shown in Fig.1. The doping is obtained exclusively by ion implantation. The Source and Drain electrodes are obtained with a high dose, low energy Phosphorous implant ($5x10^{15}cm^{-2}$, 30keV), while the Gate is obtained with a high dose, low energy Boron implant ($5x10^{14}cm^{-2}$, 12.5keV). A deep Phosphorous implant ($8.5x10^{11}cm^{-2}$, 520keV), centered at 0.65 um depth in the bulk, forms the channel of the FET. A deeper Boron implant ($1.8x10^{11}cm^{-2}$, 480keV), centered at 1.03um depth in the silicon bulk, separates the n^+ channel from the underlying completely depleted bulk.

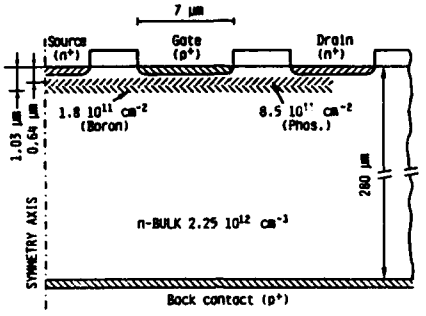
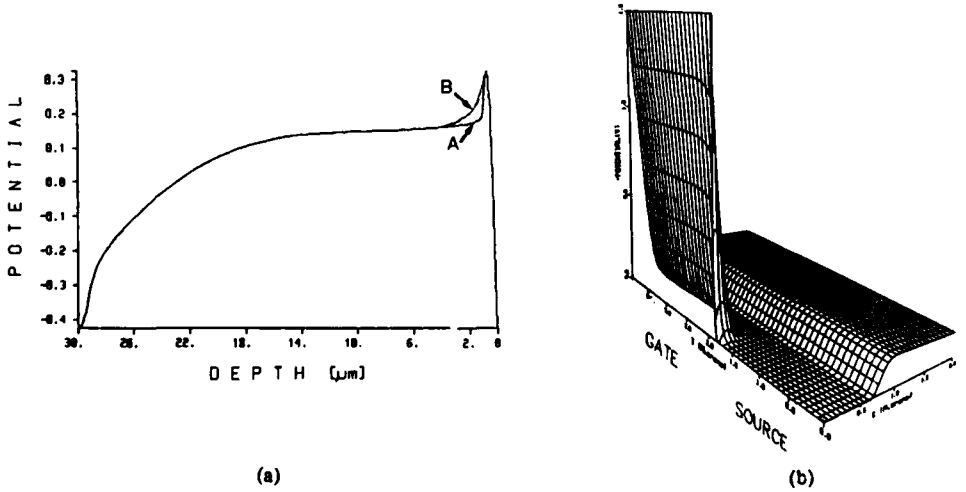


Fig.1 - Schematic cross section of the JFET built on high resistivity silicon.

The fixed negative charges of the depleted Boron deep implant create an electric field which enhances the confinement of the electrons injected from the source within the n^+ channel. Fig.2a shows the effectiveness of the deep p^+ implant in confining the electrons in the channel. Curve A represents the case in which the deep p^+ implant is present, curve B the one in which the deep p^+ implant has been omitted. A schematic 3D plot of the electron potential energy in the Source and Gate region of the transistor is shown in Fig.2b.

In order to isolate each transistor from the others in the integrated amplifying circuit and from the detector sensitive area, the device is surrounded by a thin p^+ electrode reverse biased with respect to any present n^+ electrode. The bulk is completely depleted of all mobile charges by a suitable high negative bias voltage at the p^+ implant on the back side of the wafer. Silicon nitride has been deposited for passivation purposes and to allow the deposition of the required aluminium connections.



(a) (b)

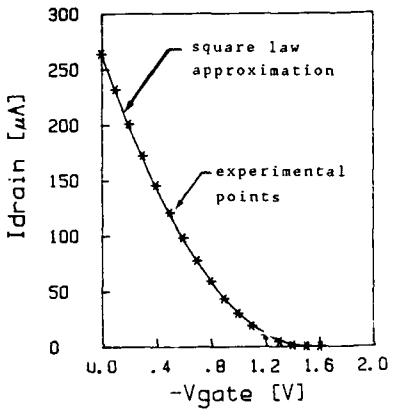
Fig.2 -a) Potential in the channel region with (curve A) and without (curve B) deep p+ implant. Only 30um depth is shown. b) Schematic 3D plot of the potential energy for electrons in the Source and Gate region of the JFET, with a negative bias applied to the Gate and 0V to the Drain.

The described structure has been used for JFETs of different geometries. Some of the geometrical parameters are the same for nearly all the transistors designed, in particular the Gate length is 7um, the spacing between the Gate implant and the Source and Drain implants are respectively 4um and 5um. The Gate width ranges between 100um and 2000um. Small transistors have a central square Source completely surrounded by concentric gate and Drain electrodes. Larger transistors have a comb structure.
A test wafer with JFETs of different geometries has been produced. Each transistor has , as mentioned in section 2, an n^+ pad connected to the gate simulating the anode of the detector and used as a connection pad for bonding and testing, and n+ source and drain pads of suitable dimensions for a wedge bonding, all electrically insulated by p^+ guards. Some n^+ additional pads have been integrated near the transistor in order to collect the reverse current from the depleted bulk.

4 - CHARACTERIZATION OF THE JFET

All the produced test transistors have been successfully tested. The I_D-V_{DS} characteristics of a transistor having a gate width of 100 um are shown in Fig.3a and the corresponding transcharacteristic I_D-V_{GS} for V_{DS}=2.5V is shown in Fig.3b.

(a)



(b)

Fig.3 - a) Characteristics of the JFET (Hor. 0.5V/div., Vert. 50uA/div, Vgate 200mV/step). Vback=-140V, Vguard=-3.5V. b) Related transcharacteristic curve for V_{DS}=2.5V

The resulting pinch-off voltage is V_p=-1.5V. The maximum transconductance g_m, at I_d=260 uA, is 0.3 mA/V corresponding to a g_m/I_D=1/850 mV^{-1}. The gate capacitance C_G is extimated to be of the order of 0.1 pF, and therefore the value of g_m/C_G is expected to be of the order of 500 MHz. The slope of the I_D-V_{DS} curve in the saturation region corresponds to an output resistence of 50 kOhm.
Note in Fig.3b that the square law approximation of the transcharacteristic fits very well the experimental points. This is due to the fact that this relationship is obtained for the simplified configuration of a delta charge distribution in the channel , which represents quite well the deep n+ implantation of the described JFET.
Note that the measurements shown in Fig.3a and 3b were performed with n+ contacts in proximity of the JFET connected to ground. These contacts collect most of the current from the depleted bulk and from the reverse biased p^+ junctions, which otherwise would be collected by the Drain, giving a non neglegible Id current in pinch-off conditions.
The Gate current is the algebraic sum of different components, as shown graphically in Fig.4a:

- I_1 - bulk current (electrons collected by the n^+ Gate pad),
- I_2 - reverse current of the junction formed by the p^+ Gate electrode and the n^+ channel (electrons emitted by the p^+ Gate electrode),
- I_3 - reverse current of the junction formed by the n^+ Gate pad and the surrounding p^+ guard (electrons collected by the n^+ Gate pad),
- I_4 - current exchanged between the n^+ Gate pad and the n^+ Drain and Source electrodes and pads (electrons emitted by the n^+ Gate pad).

Fig.4b shows the measurement of the Gate current as a function of the Gate to Source voltage of the JFET connected as a diode with the Drain short circuited to the Source. Curve A shows the case with external n^+ contacts connected to ground and p^+ guards reverse biased with respect to the Gate. Curve B shows the case with p^+ guards connected to ground. Curve A shows that in operating conditions for the JFET (V_{DG}<5V) the Gate current is less than one tenth of a nA. Curve B is presented in order to show,

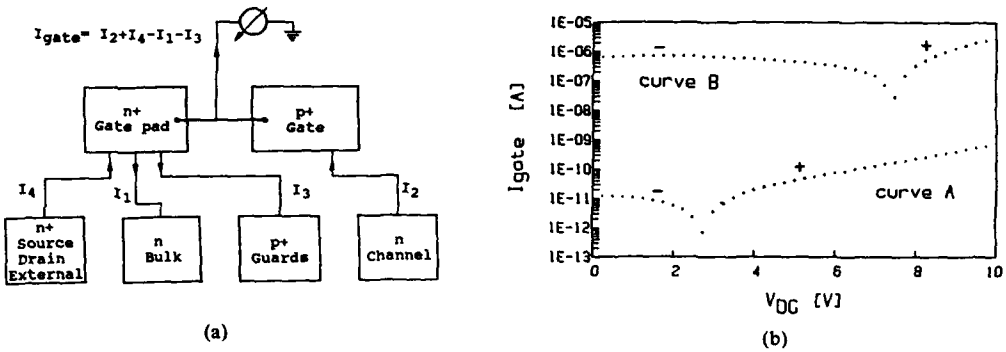


Fig.4 - a) Scheme of the different current contributions to I_{gate} to show their signs and origins. Arrows and signs are shown according the usual convention.
b) Gate currents of a diode-connected FET (D and S short-circuited) as a function of V_{DG} under different guard biases: curve A V_{guard}=-2V; curve B V_{guard}=0V. Guard potentials referred to the Gate.

in comparison with curve A, the effectiveness of a negative bias (with respect to the Gate) applied to the guard in screening the Gate n^+ pad from collecting the bulk current I_1 and in blocking the I_4 component. In the case of curve B, approximatively 1 uA of current is collected by the n^+ Gate pad, while in the case of curve A most of this current flows to the external n+ pads and only the I_2 and I_3 components (approximately one tenth of a nA) flow to the Gate pad.
In case of curve A, the Gate current is dominated by I_1 for V_{DG}<3V (V_{DG}<7.5V for curve B), while for higher V_{DG} it is dominated by I_4. These two components have opposite sign and the prevailing of one over the second determines the change in sign of the overall Gate current shown in Fig.4b. The components I_2 and I_3 are neglegible in the polarization conditions presented. The low value of the total gate current is surely not due to a compensation effect of much larger currents of opposite signs because of its relatively slow variation in a relatively large range of bias conditions. Such a low value of the gate current much less than the reverse current of the detector, will make its shot noise contribution negligible with respect to the one associated with the detector leakage current.

5 - CONCLUSIONS

In view of the integration of a low noise preamplifier on the high resistivity fully depleted silicon wafer of a radiation detector, the required JFETs have been produced and tested. Due to the operational and technological constraints posed by the detector, a new non conventional design of the transistors has been introduced. The measured performances of these JFETs meet the the design specification and show the possibility of an improved low noise processing of the detector signal.

Research supported by the U.S. Department of Energy (Contract No. DE-AC02-76CH00016) and by the Italian INFN and CNR, Progetto Finalizzato MADESS.

REFERENCES

/1/ E.Gatti, A.Longoni, M.Sampietro, 17 European Solid State Device Research Conference - ESSDERC 1987, Proceedings pg.9
/2/ J.Kemmer, Nucl.Instr. and Meth. A253 (1987) 365
/3/ V.Radeka, Nucl.Instr. and Meth. 226 (1984) 209
/4/ L.Struder, P.Holl, G.Lutz and J.Kemmer, Nucl.Instr. and Meth. A253 (1987) 386
/5/ E.Gatti,P.F.Manfredi, Riv. Nuovo Cimento 9, n.1 (1986)
/6/ V.Radeka, Nucl.Instr. and Meth. A253 (1987) 309
/7/ E.Gatti, P.Rehak, Nucl.Instr. and Meth. 225 (1984) 608
/8/ E.Gatti et al., Nucl.Instr. and Meth. 235 (1985) 224
/9/ P.Rehak et al., Nucl.Instr. and Meth. A248 (1986) 367
/10/ V.Radeka et al., IEEE Trans. Nucl. Sci. Vol.NS35-n.1-Feb.88

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

SELECTIVE EPITAXY BASE FOR BIPOLAR TRANSISTORS

J.N. BURGHARTZ, B.J. GINSBERG, S.R. MADER, T.-C. CHEN and D.L. HARAME

IBM Research Division, T.J. Watson Research Center, Yorktown Heights, NY 10598, U.S.A.

Un transistor dont la base est preparée par épitaxie sélective a été fabriqué avec une largeur de base de 1100Å et une résistance carrée de la base intrinsèque de 2800Ω/□. Nous démontrons qu'il est possible d'utiliser une largeur de base inférieure à 1000Å si les cycles thermiques après le dépôt de la base sont minimisés. L'utilisation de couches minces epitaxiées pour la base intrinsèque permet d'éviter les difficultés liées à la formation des espaceurs, en limitant la nucléation sur le silicium polycristallin de la base extrinsèque à la peripherie de l'emetteur. L'interface poly/epi est orientée suivant un plan (111). Ceci conduit à une degradation des characteristiques des dispositifs à cause de l'étalement lateral de la base extrinsèque en dessous des espaceurs.

A Selective Epitaxy Base Transistor (SEBT) is presented with a basewidth of 1100Å and an intrinsic base sheet resistance of 2800Ω/□. It is demonstrated that a sub-1000Å basewidth is possible if the temperature-time cycles after base deposition are minimal. Using thin epitaxial layers to form the intrinsic base avoids difficulties in sidewall formation caused by nucleations on the extrinsic base polysilicon at the emitter window perimeter. The poly/epi interface was found to be on a (111) plane. This leads to a degradation of the device characteristics due to extrinsic base encroachment underneath the sidewall.

INTRODUCTION

Advanced bipolar technologies with a narrow implanted intrinsic base have doping concentration profiles which are determined by channeling and defect enhanced diffusion during subsequent annealing steps /1/ and which therefore limit scalability. For scaled bipolar transistors of the future a rectangular profile is desired to achieve both a sufficient collector-emitter (*C-E*) punch-through voltage and a minimal base transit time. Unlike ion implantation silicon epitaxy does not suffer from the above restrictions and abrupt profiles, i.e. a steep out-diffusion tail, can be achieved. To integrate epitaxial base technology in advanced double-poly self-aligned structures /2/, the epitaxial deposition process has to be selective to substitute for ion-implantation without any additional process steps.

INTRINSIC BASE PERIMETER

The Selective Epitaxy Base Transistor (SEBT) was presented previously /3/. The results demonstrated the device quality of selective epitaxial films, grown in a conventional epitaxy reactor. In the SEBT structure the intrinsic base epitaxial layer is linked-up directly to the extrinsic base polysilicon which is different from technologies using implanted base layers (Fig.1). It is preferable to grow the epitaxial base bounded by polysilicon rather than oxide or nitride because facetting with high defect density occurs at the epi/dielectric boundary /4,5/. But nucleations have been observed on the polysilicon (Fig.2), which make emitter-sidewall formation difficult, i.e. sidewall etch-through is possible if nucleations are too large. It was demonstrated that for sufficiently thin epitaxial films, and consequently small nucleation sizes, sidewall formation becomes noncritical /3/. For the obtained basewidth of 1100Å the emitter sidewall covers completely the nucleations at the polysilicon. However it was observed not to be straight because of a slightly higher thickness of the epitaxial film close to the window boundary which is due to merging of the polysilicon and the epitaxial growth at the window perimeter (Fig.3). The poly/epi interface was observed to be on a (111)-plane. Since the boron diffusion coefficient is essentially higher in polysilicon compared to epitaxial silicon, lateral encroachment of the extrinsic base dopant to the emitter is possible underneath the sidewall. The consequence is a degradation of the current gain by a reduction of the effective emitter Gummel number at the emitter edge /6/.

INTRINSIC BASE VERTICAL PROFILE

As mentioned above, epitaxial base formation makes possible basewidths in the sub-1000Å range. To achieve an acceptable punch-through and Early voltage with such a narrow base the total base dopant, i.e. the boron concentration, has to be sufficiently high. A narrow more highly doped base is more sensitive to the epitaxy processing conditions and the subsequent termperature-time cycles which widen the basewidth. To quantify the impact of sidewall processing steps on profile broadening SIMS measurements were performed before and after emitter sidewall formation (Fig.4). The selective epitaxial film was deposited in an Applied Materials 7810 radiantly heated barrel type reactor running in reduced pressure mode using $SiCl_4$ at 900°C /7/. To protect the epitaxy silicon surface from dry-etching during sidewall formation, a 300Å buffer oxide was grown at 800°C after epitaxy. This oxidation is the dominant factor in broadening the profile because of oxidation enhanced boron diffusion /8/. The basewidth was 1100Å including sidewall processing and 950Å without emitter sidewall.

DEVICE CHARACTERISTICS AND DISCUSSION

The Gummel plot of the SEBT with 1100Å basewidth is shown in Fig.5. Compared with former results /3/ similar device performance was achieved also for small emitter sizes. This is due to the lowered *E-B* leakage current by a well performed emitter sidewall. Current gains up to 24 were measured. Such values are still somewhat low for the measured intrinsic base sheet resistance of 2800Ω/□. It is assumed that an encroachment of the highly doped extrinsic base, as described above, is the the reason for this reduction. At high currents the gain is even more degraded (Fig.6). In this range base current crowding takes place, increasing the current density near the emitter perimeter and decreasing the effective gain of the transistor.

Therefore lateral extrinsic base encroachment has to be lowered by either a wider sidewall or by modifications in processing. Since the sidewall width is determined by the height of the extrinsic base nitride/oxide/poly stack, encroachment has to be limited by a reduction of the final temperature-time cycles including epitaxy process itself.

CONCLUSION

A Selective Epitaxy Base Transistor (SEBT) was presented with a basewidth of 1100Å. Basewidths below 1000Å before emitter sidewall formation were measured. The incorporation of an epitaxial base and appropriate low temperature processing will maintain a narrow base profile and result in a sub-1000Å basewidth in double-poly structures. *E-B* leakage caused by emitter sidewall etch-through, as seen in former work, could be removed with deposition of thinner epitaxial films and consequently smaller nucleation sizes. Current gains up to 24 were found which are somewhat small values for the measured intrinsic base sheet resistance of 2800Ω/□. An encroachment of the extrinsic base dopant to the emitter was found to be the reason for this reduction. It is concluded that high device performance can be obtained in the SEBT device with a reduction of the temperature-time cycles and an optimized sidewall.

ACKNOWLEDGEMENT

The authors are thankful to G.P.Li for helpful discussions and to the members of the Yorktown Silicon Facility, especially to M. M. d'Agostino for processing selective epitaxy. Jerry Scilla is acknowledged for performing the SIMS measurements, Carol Stanis and Jeffrey Wetzel for TEM and Ted Ross for SEM.

REFERENCES

/1/ A.E.Michel,W.K.Chu,M.Neuman, Appl.Phys.Lett., to be published.

/2/ G.P.Li,T.H.Ning,C.T.Chuang,M.B.Ketchen,D.D.Tang,J.Mauer, IEEE Trans.El.Dev., Vol.ED-34,1987, pp.2246-2254.

/3/ J.N.Burghartz,B.J.Ginsberg,S.R.Mader,T.-C.Chen, D.L.Harame, IEEE Electron Device Letters, 9(5)(1988)259-261.

/4/ A.Ishitani,H.Kitajima,K.Tanno,H.Tsuya,N.Endo,N.Kasai, Y.Kurogi, Microelectronic Engineering, Vol.4, 1986, pp.3-33.

/5/ S.Nagao,K.Higashitani,Y.Akasaka,H.Nakata, IEEE Trans.El.Dev., Vol.ED-33,1986, pp.1738-1744.

/6/ G.P.Li,C.T.Chuang,T.C.Chen,T.H.Ning, IEDM Technical Digest, 1987, pp.174-177.

/7/ B.Ginsberg,G.Bronner,S.Mader, Extended abstr. 10th conf. chem. vapor dep., Electrochem. Soc., Honolulu, Hawaii, 1987, pp.991-992.

/8/ W.G.Allen,K.V.Anand, Solid State Electronics, 14(1971)397-406.

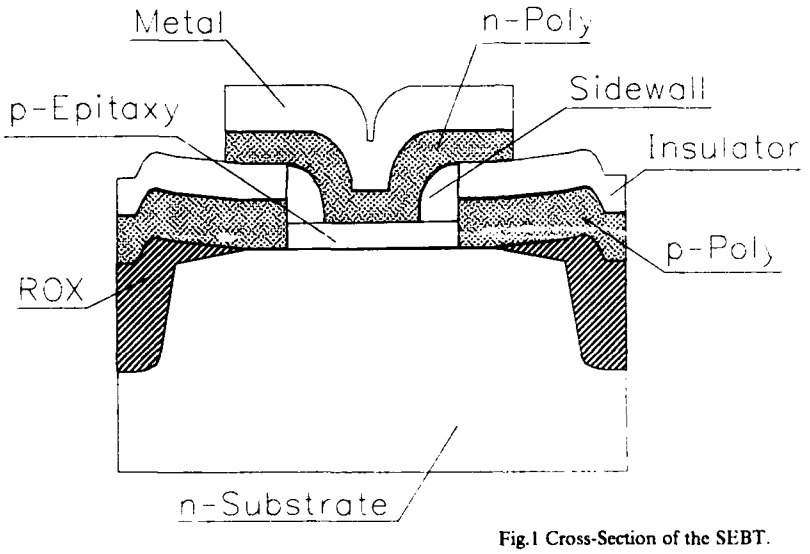


Fig.1 Cross-Section of the SEBT.

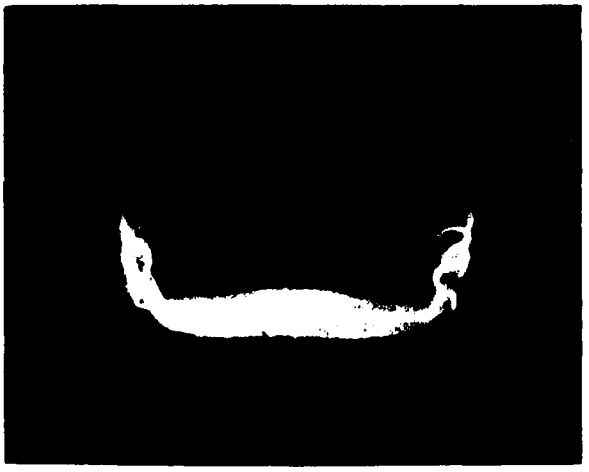


Fig.2 SEM of emitter opening after selective epitaxy.

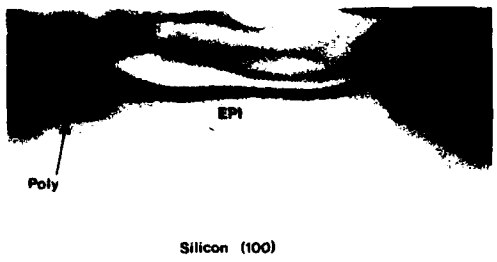


Fig.3 TEM cross-section of the emitter window after epitaxial base formation. Polysilicon nucleation occurs at the perimeter resulting in an (111)-interface with the single-crystalline layer. In this region the film thickness is somewhat higher. The epi thickness (1600Å) is essentially thicker than in device processing.

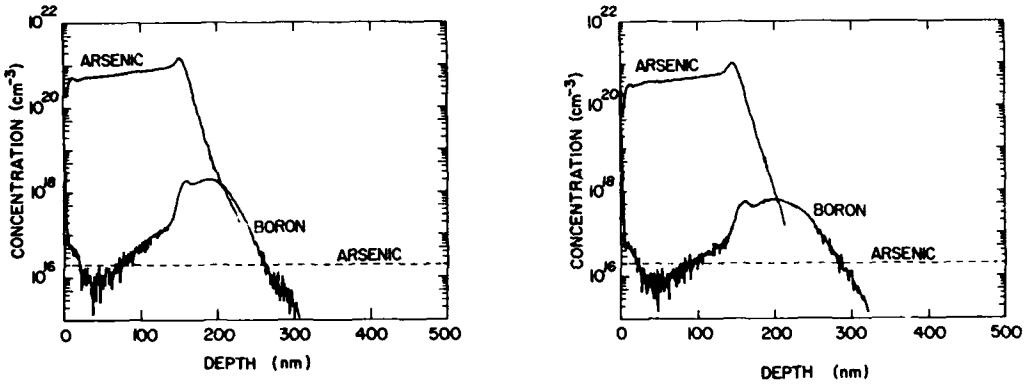


Fig.4 SIMS measurements of the intrinsic transistor profile without (left) and with (right) temperature-time cycles for emitter sidewall formation. The basewidths are 950Å and 1100Å, respectively.

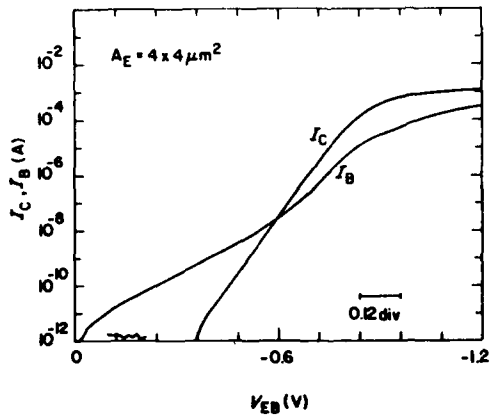


Fig.5 Gummel plot of a SEBT with 1100Å basewidth. Current gain reduction results from extrinsic base impurity encroachment underneath the emitter sidewall.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

UNIFIED MODEL FOR BIPOLAR TRANSISTORS INCLUDING THE VOLTAGE AND CURRENT DEPENDENCE OF THE BASE AND COLLECTOR RESISTANCES AS WELL AS THE BREAKDOWN LIMITS

F. HÉBERT* and D.J. ROULSTON

Electrical Engineering Department, University of Waterloo, Waterloo, Ont. N2L-3G1, Canada
**Avantek Inc., Advanced Bipolar Products, 39201 Cherry Street , Newark, CA 94560, U.S.A.*

Résumé - Un modèle pour transistor bipolaire, qui tient compte des variations des résistances de base et de collecteur avec la tension émetteur-base, la tension collecteur-base et le courant collecteur, ainsi que la dépendence en tension de la charge de base et le claquage par avalanche, est présenté. Un bon accord entre les simulations par ordinateur et expérimentations est obtenu.

Abstract - A unified bipolar transistor model, which takes into account the variation of the base and collector resistances with emitter-base voltage, collector-base voltage and collector current, as well as the voltage dependence of the base charge and the avalanche breakdown, is presented. The agreement between computer simulations and experiments is shown to be very good.

1 - INTRODUCTION

High performance bipolar transistors typically have narrow base widths and operate at high current densities. These two effects tend to reduce the useful operating region of the transistor since punch-through and avalanche effects reduce the maximum voltage, and saturation and ohmic quasi-saturation effects increase the minimum voltage limit for high current operation. Also, the base resistance varies with operating voltage and current which affects the device performance.

Accurate and physical transistor models are required in order to properly predict the device performance. A unified model which takes into account the base punch-through and avalanche breakdowns [1], as well as the base and collector conductivity modulation is presented.

2 - ASSUMPTIONS

The modeling of the voltage and current dependence of the base and collector resistances is based on the solution of the carrier distribution within the base and collector regions.

For an NPN transistor, with the doping profile shown in Fig. 1a, it is assumed that when in saturation, the electrons injected from the forward biased base-emitter junction will spread into the collector region (Fig. 1c). The base-collector space charge layer will then effectively disappear, and due to charge neutrality in the collector region, the hole concentration (p(x)) may be solved from the knowledge of the electron distribution (n(x)) and the epitaxial collector doping $n(x)=p(x)+N_{epi}$.

The base and collector conductivity modulation is then solved from n(x) within the transistor. Some workers have solved for the collector conductivity modulation through a solution of p(x) (holes) as a function of the base-collector voltage (V_{bc}) [2-5] but this neglects the effect of the base-emitter voltage (V_{be}) and the fact that the solution of p(x) is difficult [5] due to the uncertainty in defining the exact point of hole injection in the collector under high level injection (HLI). An other advantage in considering n(x) is that the base conductivity modulation is solved automatically (unified model).

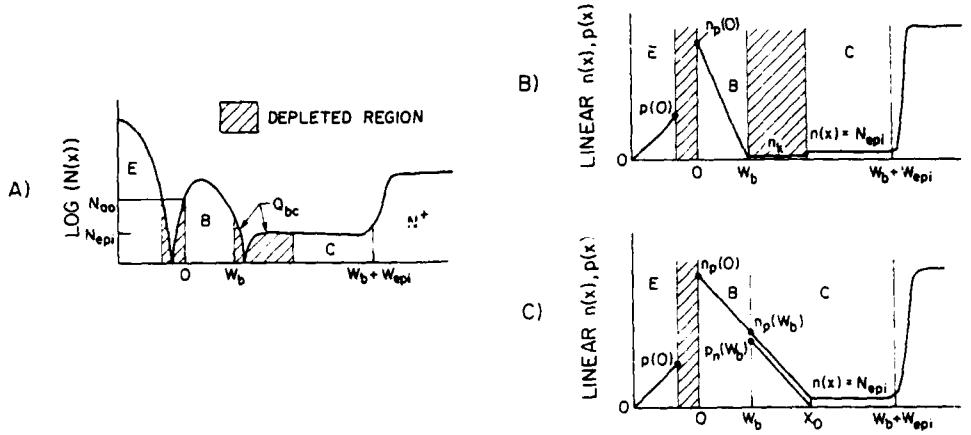


Figure 1 - Details of the emitter, base and epitaxial collector regions. a) typical doping profile. Carrier distribution in the normal mode (b) and in saturation (c).

3 - CARRIER DISTRIBUTION

For a reverse biased base-collector junction (Fig. 1b), the collector current may be computed by assuming negligible drift current:

$$Ic = K\ q\ A_e\ \mu_n\ V_t\ dn/dx = K\ q\ A_e\ \mu_n\ V_t\ n_p(0)/W_b \qquad (1)$$

where q is the electronic charge, Ae is the effective emitter area, μ_n is the average electron mobility, V_t is the thermal voltage (kT/q), W_b is the neutral base width, K is unity under low level injection (LLI) conditions and K is 2 in HLI (since Ic is made up of equal drift and diffusion components). $n_p(0)$ is the injected electron density in the base at the emitter-base space charge layer edge (Fig. 1). It is defined as, for LLI and HLI conditions:

$$n_p(0) = \{ [(n_i^2/N_{ao})\ \exp(V_{be}/V_t)]^{-1} + [n_i\ \exp\ (V_{be}/2V_t)]^{-1} \}^{-1} \qquad (2)$$

where n_i is the intrinsic carrier concentration, N_{ao} is the effective doping level at the emitter-base scl edge (Fig. 1a) and is obtained from CV measurements or from device simulations.

In the saturation mode of operation, the base-collector junction no longer acts like a sink for the electrons, and free carriers will extend from the base into the collection region. The electron distribution may be solved from I_c:

$$I_c = [1 + (n_p(W_b)/(n_p(W_b)+N_{hli}))]\ q\ A_e\ \mu_n\ V_t\ dn/dx \qquad (3)$$

The term in square brackets takes into account the gradual transition between LLI and HLI (K above). N_{hli} is the effective doping level at which HLI begins, $n_p(W_b)$ is the electron carrier density at the base-collector junction equal to $(p_n(W_b)+N_{epi})$.

The electron distribution is

$$n(x) = n_p(0) - (dn/dx)\ (x) \qquad (4)$$

where

$$dn/dx = I_c/\{[1 + (n_p(W_b)/(n_p(W_b)+N_{hli}))](q\ A_e\ \mu_n\ V_t)\} \qquad (5)$$

4 - COLLECTOR RESISTANCE

The width of the collector region over which the conductivity is modulated by free carriers is defined as x_o, as shown in Fig. 1c. It is defined as:

$$x_o = (n_p(0) - N_{epi})/(dn/dx) \qquad (6)$$

The conductivity modulated collector resistance (R_{cm}) is

$$R_{cm} = \{(x_o-W_b)/(q\ A_e'\ \mu_{nc}\ N_{cm})\} + \{(W_{epi}-(x_o-W_b))/(q\ A_e'\ \mu_{nc}\ N_{epi})\} \quad (7)$$

where A_e' is the emitter area modified to include lateral current spreading within the collector region [6] and μ_{nc} is the field dependent electron mobility [5]. N_{cm} is the average concentration over the modulated region of width x_o, computed using (4) above with $x=(W_b+(x_o/2))$.

The voltage dependence of R_c is evaluated by assuming a one-sided abrupt junction

$$R_c = R_{cm}\ ((W_{epi}-W_{scl})/W_{scl}) + R_{cext} \quad (8)$$

where W_{scl} is computed from integration of Poisson's equation and from the knowledge of N_{epi}, ($W_{scl}<W_{epi}$) and R_{cext} is the current and voltage independent resistance.

5 - BASE RESISTANCE

The base resistance is inversely proportional to the base conductivity and to the base width. HLI and base widening effects will therefore reduce the base resistance. For a given increase in base carrier concentration (proportional to an increase in base conductivity) defined as N_{bm}, and a given increase in base width, defined as W_k, the resulting reduction in base resistance is

$$R_{bbi} = R_{bbio}\ \{N_{avg}/(N_{avg}+N_{bm})\}\{W_b/(W_b+W_k)\} \quad (9)$$

where R_{bbio} is the low current intrinsic base resistance and N_{avg} is the average base doping. (9) may be expanded to include the base widening due to saturation effects and Kirk effect, and conductivity modulation due to HLI and Early effect (variation of Q_{bc} in Fig 1a). This is discussed in detail in [7].

6 - BREAKDOWN AND CROWDING EFFECTS

The avalanche and base punch-through breakdown effects are based on the work published in [1]. Crowding effects are included by using distributed equivalent circuits.

7 - MODEL PARAMETERS AND EXPERIMENTAL EVALUATION

Most of the model parameters may be obtained from measurements (sheet doping profile, sheet resistances, CV, etc.) or from device simulations. N_{hli} may be estimated as half the collector doping (N_{epi}) [2]. For devices with lightly doped bases, HLI will occur first in the base and the value of N_{hli} will tend to the average base doping.

Tne model has been implemented in the WATAND circuit simulator [8], and two different bipolar transistors displaying quasi-saturation effects have been considered. As shown in Figures 2 and 3, the agreement between simulations and experiment is very good. Figure 4 shows a comparison of the simulated and computed base resistance (using BIPOLE [9]). The reduction in intrinsic base resistance is properly simulated. The evaluation of the avalanche and base punch-through breakdown model has been carried out in [1].

8 - CONCLUSIONS

A physical model for the variation of the base and collector resistances of bipolar transistors with operating current, base-emitter voltage and base-collector voltage, compatible with CAD programs, is presented. The model is based on the solution of the electron distribution within the base and collector regions under LLI, HLI or saturation conditions. The model has been implemented in the WATAND circuit simulator and the agreement between simulations and experiment is very good.

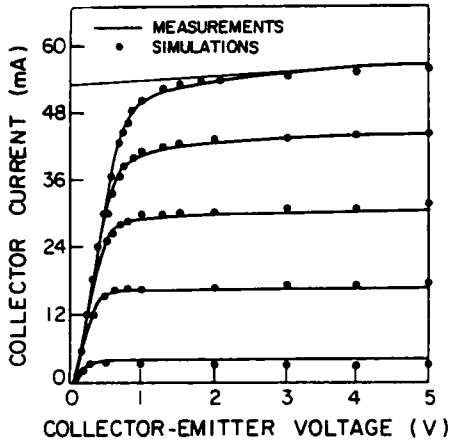


Figure 2 - I_c vs V_{ce} of a poly emitter transistor. The emitter area is 6x41 μm^2. Base current values of .1, .3, .5, .7 and .9 mA.

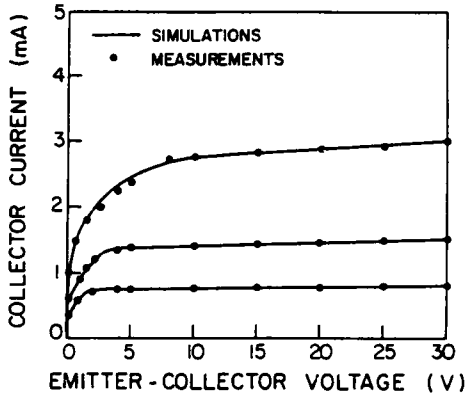


Figure 3 - I_c vs V_{ce} of a PNP device. Transistor and measurements are from [5]. Base current values of 10, 20, and 50 μA.

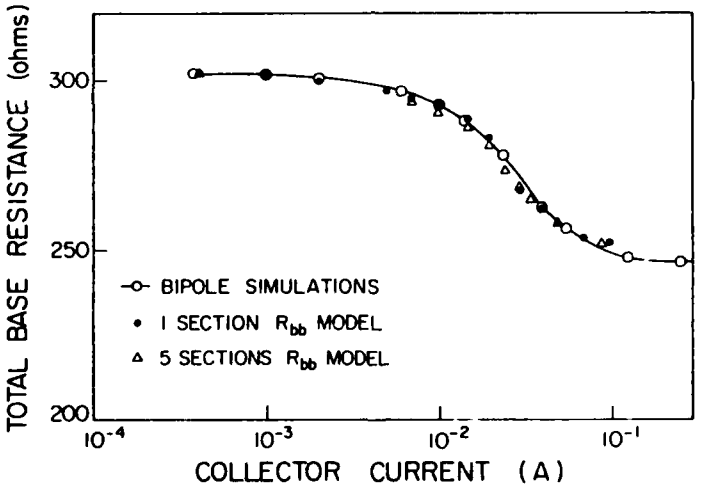


Figure 4 - Variation of the total base resistance with current for the poly emitter transistor. WATAND simulations compared to BIPOLE simulations. The distributed model is used to indicate that current crowding is not important.

REFERENCES

[1] F. Hébert and D.J. Roulston, IEEE T. Elec. Dev., ED-34, (1987), 2323.
[2] L.J. Turgeon et al. Proc. IEDM, (1980), 394.
[3] J.R.A. Beale et al. Solid-State Elect., 11, (1968), 241.
[4] A.W. Alden et al. Solid-State Elect., 25, (1982), 723.
[5] G.M. Kull et al. IEEE T. Elec. Dev., ED-32, (1985), 1103.
[6] P.R. Gray and R.G. Meyer, Analysis and Design of Analog Integrated Circuits, Wiley, (1977), 82.
[7] F. Hébert and D.J. Roulston, IEEE T. Elec. Dev. to be published.
[8] WATAND circuit Simulator, Elec. Eng. U. of Waterloo, Ont. Canada.
[9] D.J. Roulston, Proc. CICC, (1980), 2.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

PEDESTAL BIPOLAR TRANSISTOR WITH POLYSILICON ACTIVE BASE AND EMITTER WHICH ACHIEVES MINIMIZED CAPACITANCES

F. HÉBERT* and D.J. ROULSTON

Electrical Engineering Department, University of Waterloo, Waterloo, Ont. N2L-3G1, Canada.
**Avantek Inc., Advanced Bipolar Products, 39201 Cherry Street , Newark, CA 94560, U.S.A.*

Résumé - Des transistors bipolaires qui utilisent une couche de polysilicium doper de bore pour la base intrinsèque (sur silicium monocristallin) et extrinsèque (sur oxyde), et un émetteur polycristallin doper de phosphore, ont été realisés avec des gains en courant inverses jusqu'à 10. Le processus de fabrication est basé sur la déposition normale de couches polycristallines, suivie d'une recristallisation qui dépent du doping.

Abstract - Bipolar transistors using a boron doped polysilicon layer for the intrinsic (on monocrystalline silicon) and extrinsic base regions (on field oxide), and a phosphorous doped polysilicon emitter, have been realized with upward current gains up to 10. The process is based on standard polysilicon film deposition followed by doping dependent recrystallization.

1 - INTRODUCTION

Minimized parasitic junction capacitances of bipolar transistors may be achieved by self-alignment and by locating the extrinsic base region on top of a thick oxide layer [1-4]. The latter method, considered here, typically requires high temperature epitaxial film growth for the base formation [1,2,3] or photo-epitaxial equipment [4]. The results of such methods are similar: epitaxial silicon is formed over the monocrystalline Si (intrinsic region) and polysilicon forms where there is oxide (extrinsic region).

In this paper, a polysilicon rather than epitaxial base is used in order to simplify the processing requirements. Polysilicon base devices (non self-aligned PNP transistors) have been realized by workers at IBM [5] using an in-situ doped arsenic base, but low current gains and large capacitances have been obtained (rapid thermal anneal (RTA) was not utilized, and the extrinsic base region was not deposited on the field oxide). NPN transistors with minimized capacitances and recrystallized polysilicon base have been fabricated using RTA techniques. The structure, referred to as the PEDESTAL transistor from its appearance (see Figure 1), has been realized based on the doping dependence of the recrystallization of polysilicon films [6]. The recrystallization is performed prior to base doping in order to minimize the sensitivity of the base width to processing.

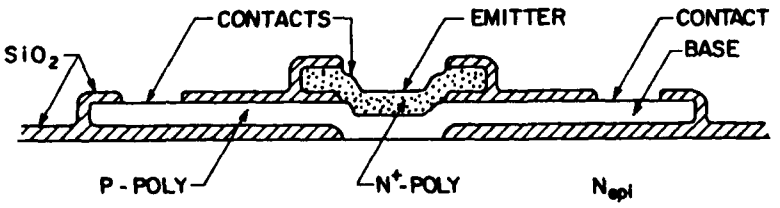


Figure 1 - Cross-section of an NPN PEDESTAL transistor. The base and emitter are formed of polysilicon. The downward mode emitter is labelled.

2 - PROCESS DESCRIPTION

The devices have been fabricated using a 3 μm thick, 0.5 ohm-cm N type epitaxial layer grown on an N^+ <111> substrate. After growth of a 0.5 μm field oxide, the intrinsic base area is patterned by etching of the oxide. The base, an undoped polysilicon layer of 0.5 to 0.6 μm is deposited at 590°C using silane in an LPCVD reactor. An HF interface treatment (to remove any native or chemical oxide grown during any RCA clean) is carried out prior to poly deposition.

Since undoped polysilicon does not easily realign epitaxially (even after a 120 seconds 1150°C heat treatment) [6], but phosphorus (Ph) doped poly films do (even after a 30 seconds 1100°C heat treatment), a low concentration of phosphorus is introduced in the undoped poly layer in order to enhance the recrystallization and epitaxial realignment with the substrate. The Ph furnace deposition conditions are 10 minutes at 650°C followed by an 80 minute drive-in at 900°C to redistribute the doping. It should be noted that an in-situ Ph doped layer (with concentration of the order of $1x10^{16}$ to $1x10^{17}$ cm^{-3}) could also have been used (this was not investigated since the gas flow meters used in our LPCVD in-situ doped reactor cannot yield such concentrations). Rapid thermal anneal (RTA) is used to recrystallize the poly film [6] and the conditions are 1150°C for 40 seconds in oxygen.

The base layer is then boron (B) doped (900°C process) yielding a base sheet resistance of approximately 150 ohms/square. An indication of the crystalline nature of the base layer is the time required for patterning similarly doped poly layers in KOH. Typical poly films etch in 4 minutes while the base layer of the PEDESTAL transistor required 42 minutes to clear, indicating that the layer morphology has significantly been altered. The combination of Ph doping and RTA is responsible for the changed layer morphology since an identical poly layer to the PEDESTAL base deposited on a patterned test wafer, undoped when annealed at 1150°C for 120 minutes in O_2, cleared in only 13 minutes using the same KOH solution.

The emitter window is defined through patterning of the base oxide followed by deposition of a 0.4 μm thick in-situ phosphorus doped N+ polysilicon film (deposited at 585°C, with 20-30 ohms/square). After RTA for activation of the emitter doping (40 seconds at 1050°C in nitrogen) the contact windows and metallization steps are performed.

3 - EXPERIMENTAL RESULTS

I_c-V_{ce} characteristics of a PEDESTAL transistor are shown in Figure 2 (upward mode). The device characteristics are adequate for I^2L operation: BVCEO of 2.5 volts, Early voltage of greater than 13 volts and current gain greater than 5. The measured base width of the transistor is approximately 0.14 μm, partly within the polysilicon and epitaxial regions, as obtained from groove and stain measurements. Current gains of 10 have been obtained with the trade-off of reduced Early voltage (4 volts). The Gummel plots of a 100x100 μm^2 upward transistor are shown in Figure 3. The collector current shows ideal behaviour and the leakage current is less than 5 $\mu A/cm^2$. The base current has an ideality factor of 1.62. In the downward mode of operation, Ic is ideal, with low leakage. The collector saturation current (I_s) of the transistor is near identical for either mode of operation, indicating a near symmetrical transistor (upward I_s of 0.99 fA and downward I_s of 1.22 fA) in cerms of effective emitter area. The base current in the downward mode is greater, with ideality factor of the order of 2 to 2.6 which is typical of polysilicon diodes [5,7,8]. The maximum downward current gain achieved is unity.

It has been observed through measurements of devices with equal emitter area but different perimeter, that the recombination is not only a function of the poly characteristics, but also of the oxide films covering the base layer (surface recombination). The base oxide has been grown using 850°C steam oxidation rather than a higher temperature dry oxygen process, and the poor quality oxide which results (cracking observed using a surface profiler) increases the surface recombination. This has been confirmed

through measurement of the base ideality factor of poly emitter transistors with diffused bases in mono silicon, fabricated using the same base oxide: ideality factor of 1.4 rather than 1.0 is obtained, indicating that the surface recombination is non-negligible (also, devices with large emitter periphery to area ratios have lower current gains).

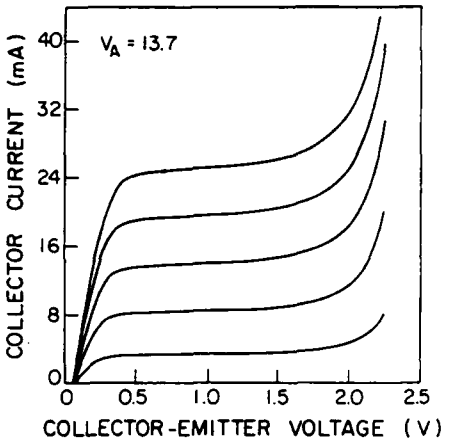


Figure 2 - I_c-V_{ce} characteristics in the upward mode of operation. The base current values are 1, 2, 3, 4 and 5 mA. 100x100 μm^2 emitter device.

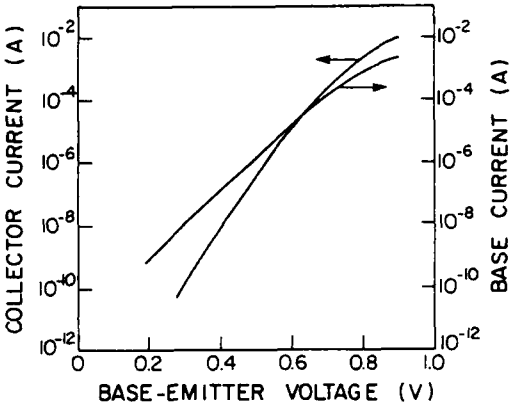


Figure 3 - Gummel plots in the upward mode of a 100x100 μm^2 device. The ideality factors for I_c and I_b are 1.0 and 1.62 respectively.

Test wafers have also been processed without RTA and Ph doping of the base polysilicon layer prior to boron doping. The resulting devices had current gains below unity and non-ideal Gummel plots (ideality factors for the base and collector current greater than unity). The poor performance is attributed to non-ideal characteristics of the interface between the base layer and the emitter and collector regions since boron doped poly does not easily recrystallize and epitaxially realign [6], even after the final RTA used to activate the emitter dopant. Similar results have been obtained for test devices which received an RTA of the boron doped poly prior to the emitter deposition.

4 - TYPICAL APPLICATION

A typical application considered for such a process is Integrated Injection Logic (I^2L). The lateral PNP device, required as injector, may be fabricated by diffusing the base dopant into the N type substrate rather than into the undoped polysilicon layer, as shown in Fig. 4. This would require patterning the base polysilicon prior to doping, such as to expose the collector and emitter regions of the PNP transistor (one extra non-critical mask, no extra diffusions). Improved speed is possible due to reduced dimensions of the active area and reduced injection in the extrinsic base when compared to diffused technologies [7]. Since the delay (due to the device) is proportional to the stored charge, the increase in switching speed is proportional to the reduction in active area of the transistor. For a typical 5 µm process, the reduction in active emitter area (upward mode) when using the PEDESTAL process is approximately 5 times.

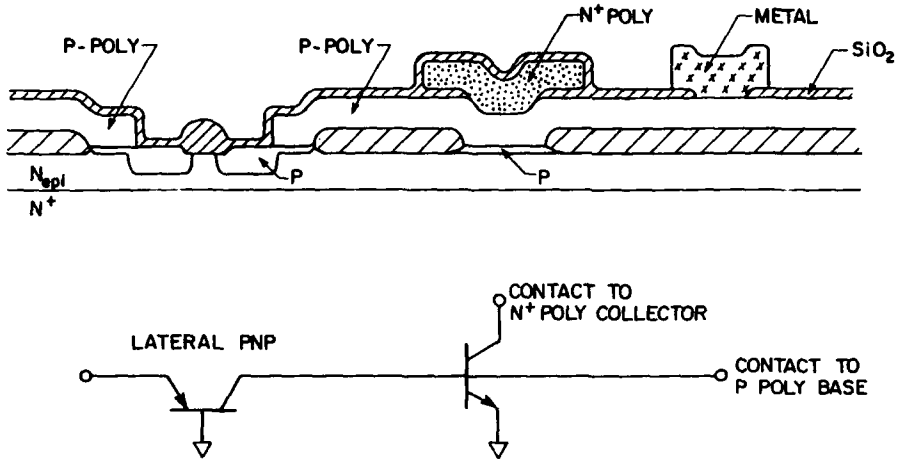


Figure 4 - Cross-section of a I^2L cell compatible with the PEDESTAL process.

5 - CONCLUSIONS

Bipolar transistors referred to as PEDESTAL transistors have been fabricated using a boron doped polysilicon layer for the intrinsic and extrinsic base regions. The extrinsic base is formed on top of a thick field oxide in order to minimize the parasitic capacitances and minimize the effective upward emitter area. The downward emitter is formed using an in-situ phosphorus doped polysilicon layer. Current gains up to 10 and ideal collector current behaviour have been realized through recrystallization of the polysilicon base layer. This has been achieved by depositing an undoped poly base layer and doping the layer very lightly with phosphorus prior to a rapid thermal anneal (1150°C for approximately 40 seconds) and prior to boron doping. The PEDESTAL device is an ideal candidate for high speed low complexity I^2L logic.

ACKNOWLEDGMENTS

The authors would like to thank Roger Grant of the SiDIC laboratory for the successful fabrication of the test wafers and for useful discussions regarding technological issues.

REFERENCES

[1] D.D. Tang et.al., IEEE J. Solid-State Cir., SC-17, (1982), 925.
[2] T. Nakamura et.al., ISSCC Dig. Tech. Papers, (1981), 214.
[3] N. Oh-uchi et.al., Proc. IEDM, (1983), 55.
[4] T. Sugii et.al. IEEE Elec. Dev. Let., EDL-8, (1987), 528.
[5] P.M. Solomon et.al., Proc. IEDM, (1979), 510.
[6] M. Tamura et.al., Jap. Jour. Applied Physics, 24, (1985), L151.
[7] H. Mikoshiba, IEEE J. Solid-State Cir., SC-13, (1978), 483.
[8] J. Manoliu and T.I. Kamins, Solid-State Elec., 15, (1972), 1103.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

HIGH FREQUENCY BASE RESISTANCE AND THE REPRESENTATION OF TWO AND THREE-DIMENSIONAL AC AND DC EMITTER AND BASE CURRENT FLOW OF BIPOLAR TRANSISTORS

F. HÉBERT* and D.J. ROULSTON

Electrical Engineering Department, University of Waterloo, Waterloo, Ont. N2L-3G1, Canada
**Avantek Inc., Advanced Bipolar Products, 39201 Cherry Street , Newark, CA 94560, U.S.A.*

Résumé - La résistance de base et l'impédance d'entrée à haute fréquence de transistors bipolaires sont simulées en utilisant une représentation à trois dimensions (3D). On observe que la réduction de la resistance de base ainsi que les courants 3D de transistors non-murés ne peuvent être correctement representées par des circuits équivalents standards. Une représentation parallèle du transistor est proposée pour mieux modeler les transistors non-murés à haute fréquence.

Abstract - The high frequency base resistance, input impedance and current flows of bipolar transistors are simulated using a quasi three-dimensional (3D) representation of the device. It is found that the reduction in base resistance as well as the 3D current flows of non-walled transistors cannot be properly simulated using standard distributed equivalent circuits. A parallel connected equivalent circuit is proposed for improved high frequency modeling of non-walled devices.

1 - INTRODUCTION

The base resistance (Rbb) and input impedance of bipolar transistors will effectively reduce at high frequencies due to ac current crowding within the intrinsic and extrinsic regions, since the capacitances and resistances are distributed. In this paper, a three-dimensional (3D) technique required for the simulation of the ac base resistance based on a 2D distributed network representation of the layout of transistors [1], with all the parasitic capacitances added, is described. The quasi 3D simulation of the input impedance of a poly emitter transistor is carried out to demonstrate the effect of ac current crowding. It is also found that the standard distributed equivalent circuit does not properly represent the two dimensional current flow of non-walled transistors; a parallel equivalent circuit is proposed as an alternative.

2 - QUASI THREE-DIMENSIONAL SIMULATIONS

In order to study the 3D ac behaviour of a transistor, the 2D distributed network representation of the transistor layout, as used in [1], is modified to include the vertical capacitive current: the peripheral (side) and plane capacitances of the base and emitter diffusions are distributed, and the current sources under the emitter, which represent the vertical base current flow [1], are replaced by a resistance (R_π), shunted by a capacitance (C_π) as per the hybrid-π equivalent circuit representation. This is shown in Fig. 1. C_π is the parallel combination of the diffusion (C_D) and the emitter-base capacitances (plane, C_{ep}, and sidewall, C_{es}). C_{cs} and C_{cp} are the sidewall and plane collector-base junction capacitances respectively.

This representation permits simulation of the input impedance and of the intrinsic and extrinsic base resistances (R_{bbi} and R_{bbx}, from power dissipation considerations, [1]) at various frequencies. The operating current is defined by the value of R_π and C_D used in the analysis.

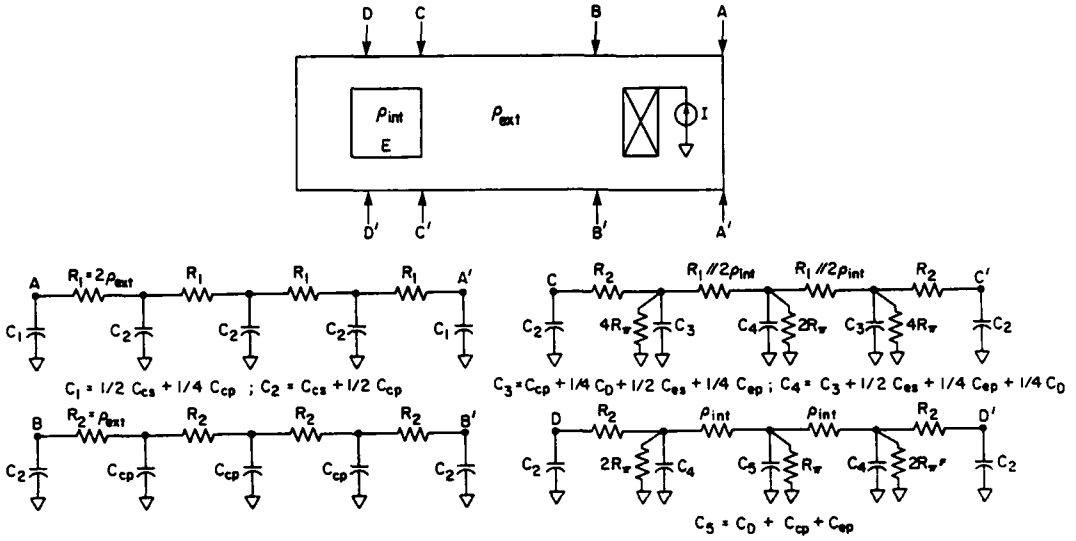


Figure 1 - 2D layout representation [1] modified to take into account the vertical ac current flow. The component values are indicated in the 4 cross-sections shown.

It is assumed that the collector is an ac ground and that dc current crowding is negligible. The sheet capacitances, specified in Farads per unit area, and the sidewall (edge) capacitances, specified in Farads per unit length, are represented by grounded capacitors at each node of the equivalent circuit, as shown in the various cross-sections of Figure 1. The capacitive values required for the analysis may be obtained from measurements or from device simulators, such as BIPOLE [2], for the bias voltage of interest.

3 - HIGH FREQUENCY SIMULATIONS

The ac input impedance has been simulated for a single base contact transistor with the characteristics listed in Table I. The transistor considered is an oxide isolated device with a non-walled poly emitter and an f_t of 3 GHz. The resulting "input impedance circle" [3] is shown in Figure 2. A semi-circle if formed by curve-fitting of the low frequency data points (as expected [3]) and the extrapolated high-frequency intercept (1400 ohms) yields exactly the low-frequency base resistance.

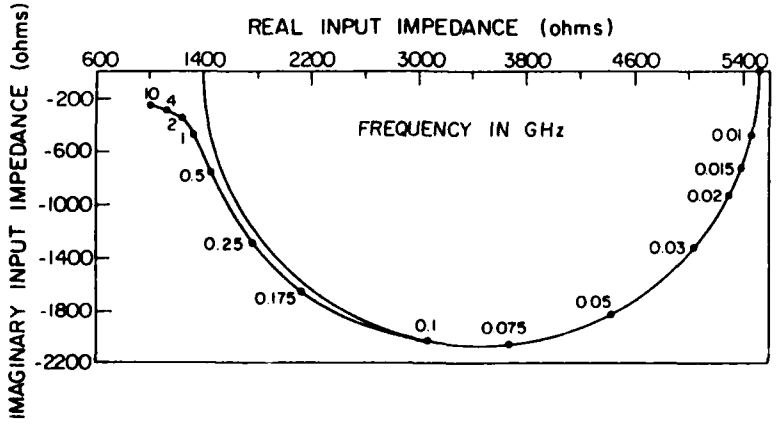


Figure 2 - Input impedance of a non-walled poly emitter bipolar transistor. Half of the layout view shown in Fig. 1 was simulated.

Table I - Poly emitter transistor characteristics

Parameter	Value	Units
Emitter area	2 x 2	μm^2
Base area	4 x 7	μm^2
Ext. sheet res.	548	ohms/square
Int. sheet res.	3020	ohms/square
Current density	151	amps/μm^2
R_{π}	1.28×10^{5}	ohms/μm^2
C_{π}	1.40×10^{-14}	F/μm^2
C_{es}	7.77×10^{-16}	F/μm
C_{cs}	2.75×10^{-17}	F/μm
C_{cp}	1.75×10^{-17}	F/μm^2

The divergence of the input impedance from the semi-circle at very high frequencies is a result of ac current crowding which effectively reduces the base resistance. Both Rbbi and Rbbx are observed to reduce, and the reduction in Rbbi is greater than that of Rbbx due to the larger capacitance in the intrinsic region.

4 - EQUIVALENT CIRCUITS

The variation of the input impedance of non-walled emitter transistors cannot be properly simulated using a standard distributed equivalent circuit (as shown in Figure 3) due to the multi-dimensional current flows. To illustrate this, the surface potential over the emitter of walled (Fig. 4) and non-walled (Fig.5) devices have been computed using the quasi 3D technique described above. The base current flow in the walled device is clearly one-dimensional (from a layout view) and the equivalent circuit of Fig. 3 may be used to accurately model the variation of the input impedance with bias or frequency. In the case of the non-walled structure, since the base current flows around the emitter, current crowding first occurs at the emitter edge which is farthest away from the base contact. Also, the emitter-base voltage at the emitter periphery will vary with current which results in 2D variations of the emitter-base capacitance This does not occur when the device is walled since the base current is one-dimensional in the layout plane.

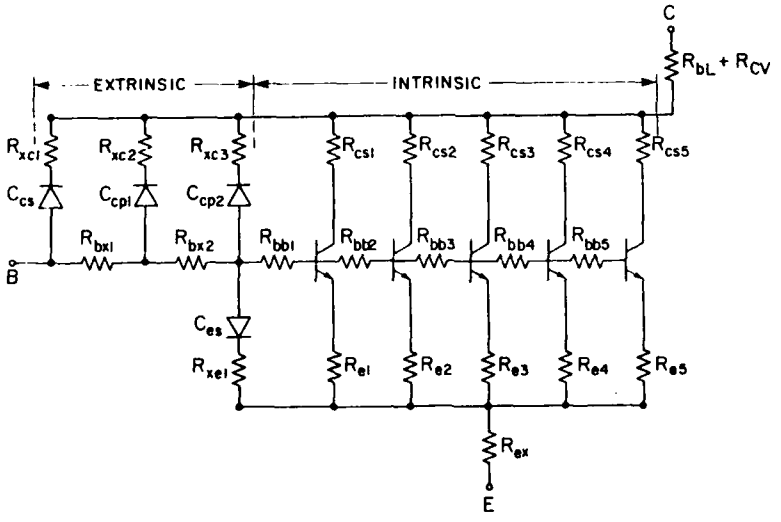


Figure 3 - Standard distributed equivalent circuit showing the plane and sidewall capacitances. 5 intrinsic and 2 extrinsic sections are shown.

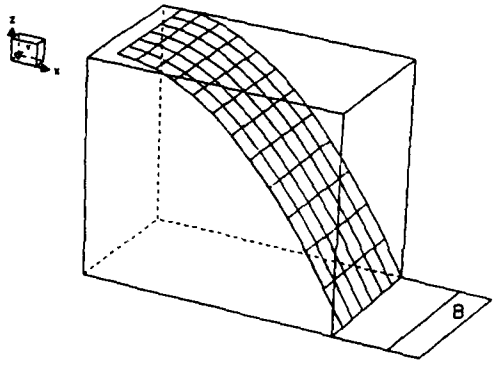


Figure 4 - Potential (z-axis) over the surface of a walled emitter one base contact bipolar transistor.

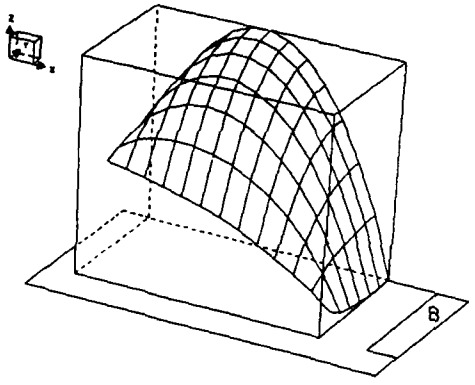


Figure 5 - Potential (z-axis) over half the surface of a non-walled emitter single base contact device.

In order to better represent the distribution of the emitter-base voltage over the emitter surface and the 3D currents, the parallel distributed circuit of Fig. 6 is used. Each of the three transistors shown (device "A" represents the quadrant near the base contact, device "C" represents the quadrant farthest away from the base contact and device "B" represents the remaining two quadrants) may be represented by the distributed equivalent circuit of Fig. 3. R_{bA}, R_{bB} and R_{bC} are the extrinsic resistances along the paths shown in Fig. 6a. These are computed using a 2D simulator [1] with the following requirements: R_{bA} is the minimum value of R_{bbx} (when most of the current flows in quadrant "A") and the sum of R_{bA}, R_{bB} and R_{bC} is equal to the total extrinsic base resistance computed at low currents and low frequencies [1].

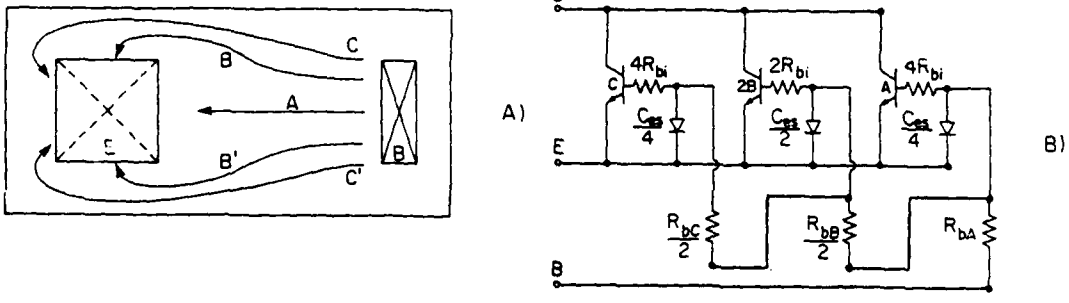


Figure 6 - Parallel equivalent circuit of non-walled emitter transistor. a) layout view showing the base current paths. b) equivalent circuit. The four emitter quadrants are assumed to carry equal currents. The area of the "B" transistor is twice that of the "A" and "C" devices.

REFERENCES

[1] F. Hébert and D.J. Roulston, Solid-State Elec., 31, (1988), 283.
[2] D.J. Roulston, Proc. CICC, (1980), 2.
[3] W.M.C. Sansen, R.G. Meyer, IEEE J. Solid-State Cir., SC-7, (1972), 492.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

ELECTRICAL BEHAVIOUR AND MODELLING OF A N-DOPED a-Si:H EMITTER BIPOLAR TRANSISTOR

O. BONNAUD, A. EL GHARIB and M. SAHNOUNE

Groupe de Microélectronique, Université Rennes I, Campus de Beaulieu, F-35042 Rennes Cedex, France

I - INTRODUCTION

In recent years, heterojunction bipolar transistors (HBT's) have attracted much attention because it is possible to get a high current gain with heavily base doping. Furthermore, heterojunction structures mainly employing the conventional silicon technology are particularly attractive, especially hydrogenated amorphous silicon/single crystalline silicon (a-Si:H/c-Si) heterojunctions [1 -4]. The use of n-doped a-Si:H/p-doped c-Si as emitter-base junction of a HBT allows to create an extra barrier to holes that minimizes the minority carrier injection in the emitter and thus allows to get a high current gain bipolar transistor in silicon technology. This paper is devoted to the study of the electrical behaviour of a n-doped a-Si:H emitter HBT, to the understanding and the modelling of the electrical mechanisms in order to improve the features of these devices.

II - FABRICATION OF THE DEVICES

The bipolar transistors are fabricated on an epitaxial substrate as follows. The n^+-doped substrate constitutes the collector layer. The base is fabricated from ion implantation of boron in the n^--epilayer. The emitter is made of a-Si:H deposited on the base layer from a silane decomposition at low temperature (250-275°C) with PH_3 as doping gas (glow discharge technique). The a-Si:H layer is very thin in order to minimize the series resistance of the emitter and is 50 nm thick. Following this deposition, an emitter contact is made of chromium that plays the role of a diffusion barrier to the aluminum atoms of the contact overlayer (fig.1). The choice of a chromium layer instead of titane layer is due to a better ohmic contact obtained with this element after checking the contact resistance on test structures (chromium or titane and a-Si:H layers deposited on glass substrates and n^+ degenerated silicon substrates). The finished devices are annealed in forming gas at low temperature temperature (250°C) ligthly lower than the a-Si:H deposition temperature. This annealing step improves the electrical characteristics of the transistor

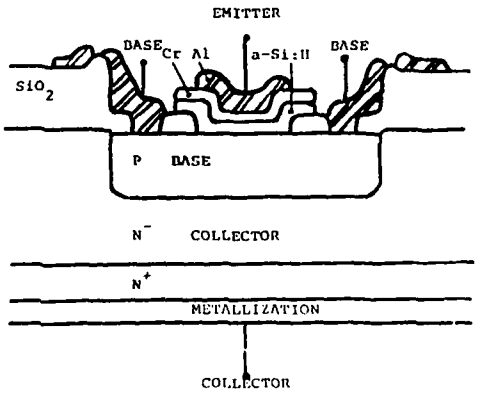


Fig. 1 : Crosssection of the structure

possibly because the aluminum contacts on silico and chromium are improved. The emitter area varie in the range 60-3500 um^2. Figure 2 shows a fina structure.

Fig. 2 : Photograph of the final transistor

III - DEVICE CHARACTERISTICS

The electrical behaviour of the devices is deduced from the electrical characteristics.

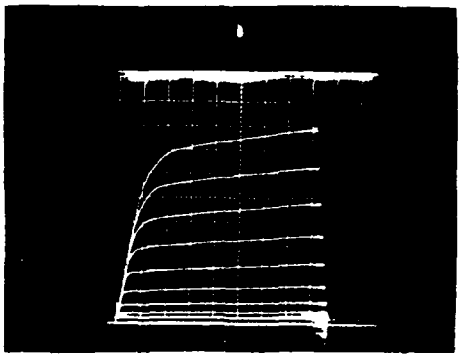


Fig. 3 : $I_C(V_{CE}, I_B)$ characteristics - I_C= 2 mA/div I_B= 5uA/div V_{CE}= 2 V/div 10 steps of I_B

Figure 3 shows the variation of the collector current versus collector emitter voltage with base current as parameter. In this case, the dynamic current gain equals to 600, but depending on the Gummel number of the base, G_B, and the quality of the amorphous layer, more especially the conductivity, this current gain varies in the range 80-1200.

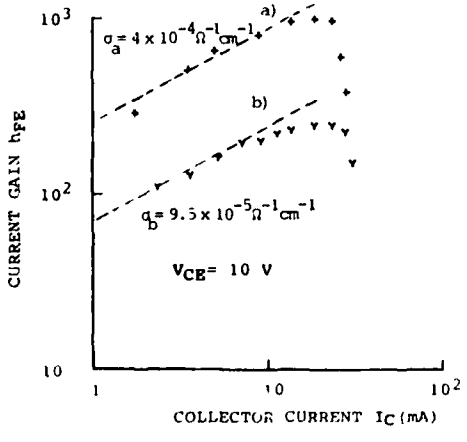


Fig. 4 : Dynamic current gain versus collector current.

Figure 4 shows the variation of the dynamic current gain versus collector current for two types of amorphous silicon layers. In the case of curve a), the saturation of the current gain is not observed and furthermore the law β proportionnal to $I_C^{1/2}$ is verified. Therefore, the current appears to be strongly infuenced by the recombination phenomenon in the space charge layer of the emitter-base junction. Curve b) corresponds to a poor conductivity (10^{-4} $Ohm^{-1}cm^{-1}$) of the a-Si:H layer ; the doping of the emitter is low enough to limit the maximum current gain.

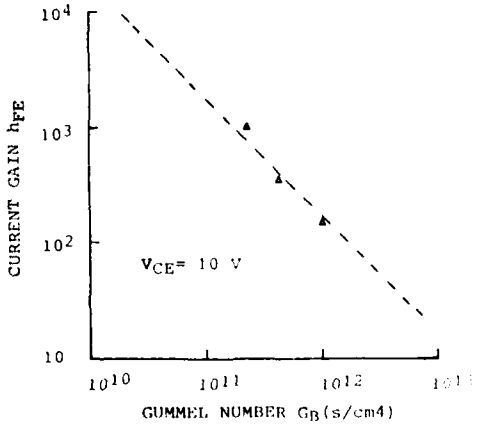


Fig. 5 : Dynamic current gain versus Gummel number.

Figure 5 provides the current gain results versus the Gummel number of the base deduced from the ion implantation dose. The law β proportionnal to $1/G_B$ is approximately verified.

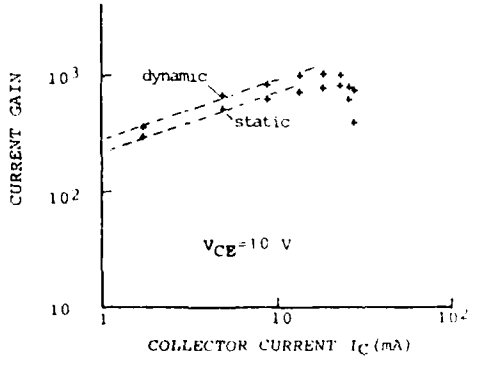
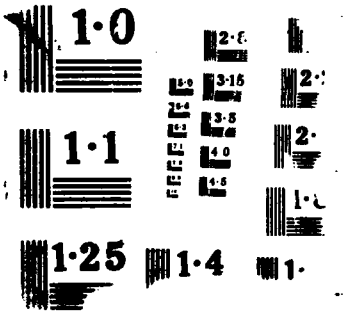


Fig. 6 : Comparison of the static and dynamic current gain versus collector current.


1·0
1·1
1·25
1·4
1·

A comparison between static and dynamic current gains versus collector current is given on figure 6. The difference can be easily explained by the variation of the total current gain versus the current density (Fig. 4). Figure 7 shows the total emitter-base junction current versus the base-emitter bias when collector-base junction is shortened (V_{CB}=0). Because the slope equals $1/V_T$, the collector current law is well described by the diffusion model in a three or four orders of

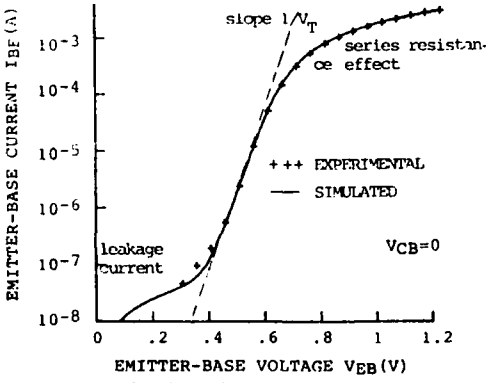


Fig. 7 : Total emitter-base junction current versus base-emitter voltage when collector-base junction is shortened.

magnitude. At high level, the conduction is governed by the series resistance of the emitter layer ; at low level, a leakage current occurs. Photograph on figure 8 shows the experimental base-emitter characteristics. The reverse breakdown voltage of the emitter-base junction is close to 7V, similar to its monocrystalline conterpart.This can be explained

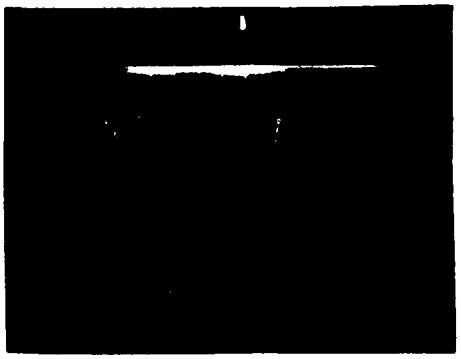


Fig. : 8 : Photograph of the base-emitter characteristics. I_B= 2mV/div V_{BE}= 2 V/div

from C(V) measurements. Because the doping profile of the base is directly deduced by tracing $1/C^2$ versus V, the space charge layer of the heterojunction mainly extends in the monocrystalline region. Therefore, the breakdown voltage depends mainly of the doping profile of the base as in heavily doped monocrystalline emitter bipolar tansistor. On the other hand, we may predict that the volumic recombination mainly occurs on the base side of the space charge layer of the junction.

IV - MODELLING

We propose a modelling of the collector current versus the base emitter voltage (fig. 8, simulated curve) and of the current gain versus the collector current taking into account the following parameters :

- surface recombination rate at the a-Si:H/c-Si interface (fig. 9)

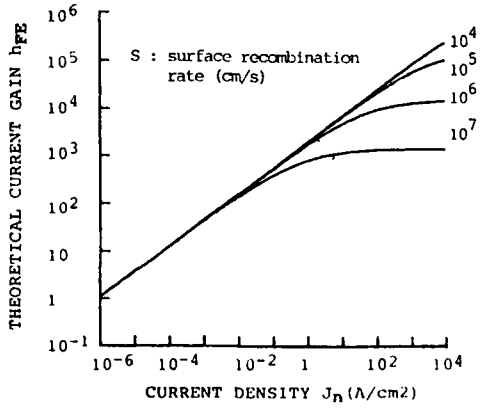


Fig. : 9 Simulated current gain versus collector current with surface recombination rate as parameter

- carrier lifetime in the space charge layer of the heterojunction (fig. 10)
- minority carrier density in the a-Si:H emitter (fig. 11)
- series resistance of the n-doped a-Si:H layer
- access resistance of the extrinsic base
- crowding effect in the case of large emitter areas.

V - DISCUSSION

The comparison between the modelled curves and the experimental results allows the following interpretations :

- for values as high as 1000, the current gain is not limited by the interface recombination rate,

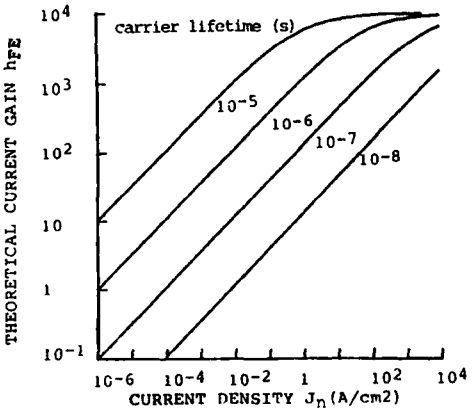


Fig. 10 : Simulated current gain versus collector current with carrier lifetime in the space charge layer as parameter

that means that the a-Si:H acts as a passivant layer for the c-Si surface. In a contrary case, the current gain would saturate at low level

- the minority carrier density in the a-Si:H is low enough to permit a high current gain even when the n-doped a-Si:H conductivity is as low as $10^{-4}Ohm^{-1}cm^{-1}$; thus, the extra barrier for the hole deduced from the energy band diagram is effective and close to 0.4 eV
- the current gain essentially depends on the recombination in the space charge layer of the emitter-base heterojunction because the $I_C^{1/2}$ variation law of β
- the emitter electron current density injected in the base is governed by the diffusion process (the ideality factor equals 1)
- at high level of the current density, we have to take into account of the emitter series resistance which is strongly function of the a-Si:H deposition conditions.

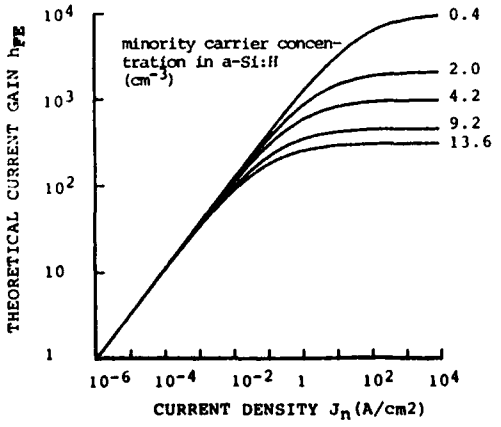


Fig. 11 : Simulated current gain versus collector current with minority carrier density in a-Si:H as parameter.

References :

|1| O. BONNAUD,P. VIKTOROVITCH, IEE Proceedings, vol.132, Part I, n°1, Feb. 1985
|2| R. MERTENS, J. NIJS, J. SYMONS, K. BAERT, ESSDERC 87, BOLOGNE (It.), Sept. 1987
|3| H. FUJIOKA, S. RI,K. TAKASAKI, K. FUJINO, Y. BAN, IEDM 1987, pp 190-93
|4| J. SYMONS, M. GHANNAM, A. NEUGROSHEL, J. NIJS, R. MERTENS, Solid State Elec., vol.30, n°11, Dec. 1987, pp 1143-45

This work is partially supported by GCIS CNRS France. The authors thank M. MORIN, J.L. FAVENNEC and Y. CHOUAN with CNET LANNION for a-Si:H deposition, and wish to thank all the team working in Technology Laboratory of C.C.M.O. (Centre Commun de Microélectronique de l'Ouest).

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, seotembre 1988

MAXIMUM OPERATING VOLTAGE LIMITATIONS DUE TO PARASITIC BIPOLAR ACTION IN VLSI CMOS

D. TAKACS and M. STEGER

Siemens AG, Corporate Research and Development, Microelectronics, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

Abstract-The influence of the PBTA (Parasitic Bipolar Transistor Action) induced by impact ionization on the maximum operating voltages as well as on the isolation performance in VLSI CMOS structures was investigated using the sensitivity of the nMOS- and bipolar-parameters to the operating voltages.Strong coupling between the MOS- and bipolar-currents exists due to the potential drop caused by the impact ionization currents of the nMOSFET.The base currents of the bipolar transistors reduce the measurable substrate current.The maximum operating voltage in CMOS is limited by the impact ionization induced PBTA.The isolation performance is reduced by the collector current of the vertical transistor.

1. INTRODUCTION

The driving force to reduce dimensions in VLSI is to increase the speed and the packing density of the circuits. The main difficulty in scaling down CMOS designs in the future will be the reduction of the effective channel length L_{eff} of the MOSFETs and the n^+/p^+ isolation spacing between them. By reducing L_{eff}, the breakdown properties of the MOS transistors are affected. Reducing the n^+/p^+ spacing will affect the electrical separation between the n- and p-MOSFETs.

When scaling down to submicron levels at constant operating voltages, the electric field in the channel and the related impact ionization rates increase rapidly. The potential drop ΔV_{BB} produced by the impact ionization currents in the electrical substrate of the nMOSFET gives rise to Parasitic Bipolar Transistor Action (PBTA). The resulting bipolar current is an injection current. It increases exponentially with ΔV_{BB} (≈1 decade/60 mV). The PBTA leads to interactions between the currents of the n- and p-channel areas that can be enhanced by the positive feedback of the npn and pnp bipolar transistors in CMOS.

Although potential drops ΔV_{BB} below 0.6V will only lead to relatively low bipolar currents, they can significantly influence the MOS currents and the isolation between the well and the substrate.The MOSFET parameters are affected by the impact ionization dependent change of the back gate bias V_{BB} and by the bipolar current I_B injected from the source of the nMOSFET.The isolation performance is then affected by the collector current of the npn transistors.The result is a limited operating voltage and a reduced isolation performance.

The purpose of this work was to investigate the influence of the impact ionization induced PBTA on the maximum operating voltage of the nMOSFETs and on the isolation performance in CMOS.Experiments were carried out on samples fabricated in p-well based twin-tub CMOS processes with 1.0µm to 1.5µm feature sizes and gate oxide thicknesses from 20nm and 25 nm,respectively.

2. MAXIMUM OPERATING VOLTAGE

Fig.1 shows the cross sectional view of a p-well CMOS structure with the n-channel MOSFET, the parasitic npn bipolar transistors and the current generator I_{ion} representing the impact ionization current. At sufficiently

large potential drop ΔV_{BB} caused by I_{ion}, electrons are injected from the n^+ source into the p-well leading to PBTA.Since the emitter efficiency of the n^+ area is very high,two effective npn parasitic bipolar transistors contribute to the injection current. The consequences of the PBTA in p-well CMOS can become increasingly important considering the voltage limitations and the hot carrier degradation of the nMOSFETs.

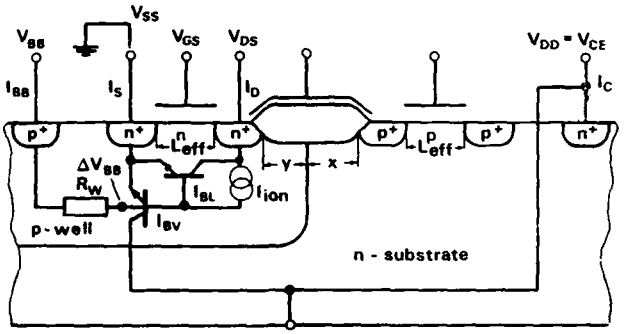


Fig.1:Cross sectional view of the test structure with MOSFETs,parasitic bipolar transistors and the current generator I_{ion}.

The lateral npn transistor can cause the well known drain induced barrier lowering (DIBL) for short channel lengths.It represents a drain to source voltage limitation for the nMOSFET.

The vertical npn transistor can cause:

* A reduction of the isolation between p-well and n-substrate due to its collector current.
* A reduction of the maximum operating voltage $V_{DD}=V_{CE}$ for a CMOS relevant functional unit (to retain the same current levels in the well and the substrate as without PBTA).
* A reduction of the maximum operating voltage V_{DS} of the nMOSFET because of an enhanced injection from the source. An increased injection of electrons from the source leads to increased drain current due to the attraction of the injected electrons by the gate field. An increased drain current leads to an increased impact ionization and thus to a positive feedback.

The two basic conditions for starting PBTA are: the impact ionization current I_{ion} and a certain collector to emitter voltage $V_{CE}=V_{DD}$. The current I_{ion} is determined by the operating point of the nMOSFET e.g. V_{GS}, V_{DS} and V_{BB}. V_{CE} is the CMOS relevant operating voltage V_{DD}. The degree of the V_{DS} and V_{DD} limitations by the PBTA will be demonstrated by the dependences of the more important MOSFET parameters on the operating voltages.

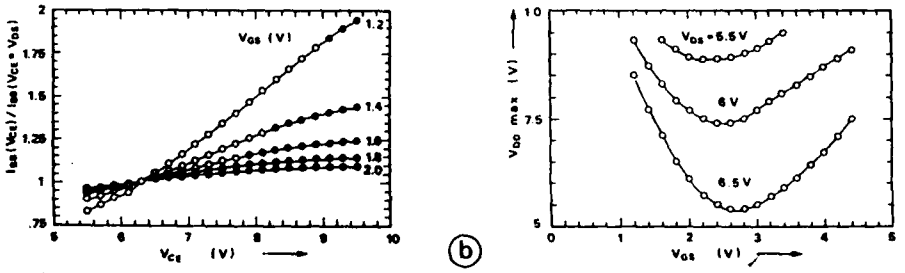


Fig.2:Influence of the operating voltages on nMOST-and CMOS-parameters.Fig.2a:substrate current I_{BB},normalized by I_{BB} at $V_{CE}=V_{DS}=6.4V$ vs. V_{CE}. Fig.2b:the maximum operating voltage V_{DD-max} vs. V_{GS} for different V_{DS}.

The normalized substrate current of the nMOSFET as a function of the operating voltage $V_{CE}=V_{DD}$ for different gate voltages V_{GS} at $V_{DS}=6.4V$ is shown in Fig.2a. The normalizing factor is the substrate current at $V_{CE}=V_{DS}$. The filled circles for each gate voltage correspond to the condition that injection from the source occurs. This condition is verified by measuring the

drain and source currents of the nMOSFET. The first filled circle going to higher V_{CE} voltages for each V_{GS} indicates the condition at which the value of the source current becomes just equal to the drain current. The last value of V_{CE} before injection represents the maximum operating voltage V_{DD-max} for the given CMOS structure. At low gate voltages the substrate current (=current in the p-well) of the nMOSFET is very low too so that it results in a high relative change. Increasing V_{GS}, injection occurs at lower $V_{CE}=V_{DD}$ voltages indicating a respective operating voltage limitation.

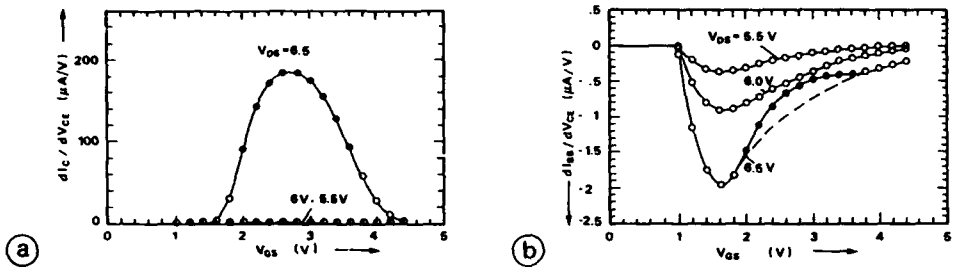


Fig.3:Operating voltage-sensitivities of the collector current I_C (Fig.3a) and of the n-channel substrate current I_{BB} (Fig.3b) as a function of V_{GS}.

Fig.2b shows the maximum operating voltage V_{DD-max} as a function of the gate voltage V_{GS} for different V_{DS} voltages of the nMOSFET. The shape of the V_{DD-max} curves reflects the well-known substrate current shape for all drain voltages V_{DS}. V_{DD-max} decreases with increasing V_{DS}.

The collector current of the vertical npn bipolar transistor I_C, the substrate and the drain currents I_{BB} and I_D of the nMOSFET characterized by their sensitivities to $V_{CE}=V_{DD}$ as a function of the V_{GS} for different V_{DS} voltages are shown in Fig.3a, 3b and Fig.4, respectively.

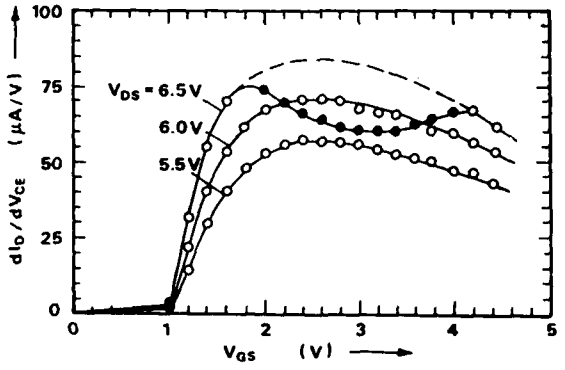


Fig.4:Sensitivity of the n-channel drain current vs.V_{GS} for different V_{DS}.

Fig.3a shows that injection occurs only for $V_{DS}\geq 6.5$ V. The curve follows the V_{GS}-dependent substrate current shape. Even at $V_{DS}=6.5$V the injection from the source is significant only in that gate voltage range where the substrate current is sufficiently high. Below $V_{DS}=6.5$V no injection into the substrate takes place. The nMOST parameters in this case are influenced by the body effect changing the threshold voltage V_T. Due to the modulation of the depletion width at the well/substrate junction, the well resistance R_W and thus the potential drop ΔV_{BB} are modulated too.

The influence of the PBTA on the substrate and drain currents is demonstrated in Fig.3b and Fig.4. Without injection (circles) both substrate and drain sensitivities are increased when V_{DS} is increased. With injection (filled circles) both substrate and drain sensitivities are reduced. The reduction can be estimated from the difference between the dashed line (i.e.the curve approximated from points without injection) and the measured curve in the V_{GS} range where injection occurs. The I_{BB} reduction shows clearly that a part of

the impact ionization current I_{ion} is consumed by the base current of the npn transistors. Therefore the impact ionization in MOSFETs will be underestimated using the measurable I_{sub} to monitor it.The reduction in I_D is a consequence of the decreased voltage drop ΔV_{sub} caused by the reduced I_{sub}.

3. ISOLATION PERFORMANCE

The influence of the PBTA on the isolation performance of a CMOS relevant functional unit e.g. inverter is demonstrated in Fig.5a. The substrate currents of the n- and p-channel MOSFETs in the inverter circuit , I_{subn} and I_{subp},respectively are shown vs. the input voltage V_{IN}. The regular substrate currents I_{subn} and I_{subp} generated by impact ionization in the respective MOSFETs are denoted by solid lines. However,an additional peak of the I_{subp} is observed (denoted by dashed lines).This is the collector current of the vertical npn transistor.Preventing PBTA by reverse biasing the p-well/source junction, the additional peak disappeares (dotted line).
The maximum values of the additional peaks for different doping levels in the p-well vs.V_{DD} are shown in Fig.5b. Reducing the PBTA by increasing the Gummel number,the collector currents I_{subp} are reduced effectively.

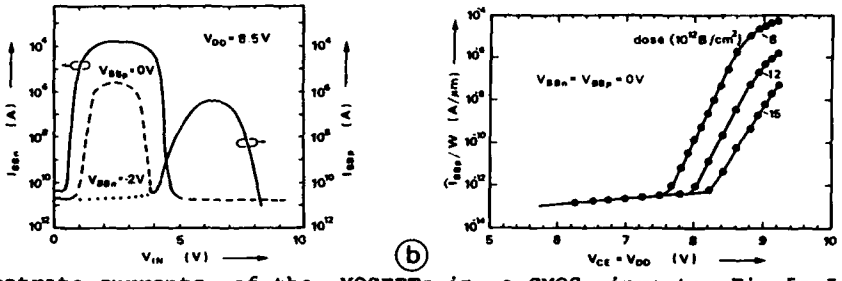


Fig.5:Substrate currents of the MOSFETs in a CMOS inverter.Fig.5a:I_{subn} and I_{subp} vs.V_{IN}.Fig.5b:Maximum values of the additional peaks I_{subp} normalized by the channel width W vs. V_{DD} for different doping levels in the p-well.

4. CONCLUSIONS

Voltage drops ΔV_{sub} below .6V caused by impact ionization currents in the electrical substrates of the respective MOSFETs lead to PBTA in VLSI CMOS.This action results in limitations of the maximum operating voltages of the discret MOSFETs and of the CMOS functional unit.These influences were investigated using operating voltage- sensitivities of the MOS- and bipolar-currents.

The obtained results demonstrate clearly that:
* The nMOSFET- and the bipolar-currents in CMOS are strongly coupled by the potential drop caused by the impact ionization currents.
* The maximum operating voltage in CMOS is limited by the PBTA induced by the operating point dependent impact ionisation in nMOST.
* Any increase of the well-doping effectively reduces this effect.
* The base currents of the parasitic bipolar transistors reduce the measurable substrate current of the corresponding MOST.
* The isolation performance of a functional unit in CMOS is reduced due to the collector current of the vertical bipolar transistor.

Session 4B Room B : Dielectrics

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

PLASMA ANODISATION OF SILICON FOR ADVANCED VLSI

S. TAYLOR, W. ECCLESTON, J. RINGNALDA*, D.M. MAHER*·(1), D.J. EAGLESHAM*, C.J. HUMPHREYS* and D.J. GODFREY**

Department of Electrical Engineering and Electronics, Liverpool University, PO Box 147, GB-Liverpool L69 3BX, Great-Britain
**Department of Material Science and Engineering, Liverpool University, PO Box 147, GB-Liverpool L69 3BX, Great-Britain*
***GEC Hirst, Wembley, GB-Middlesex HA9 7PP, Great-Britain*

Abstract - Inductively coupled, RF stimulated plasma anodisation of silicon is discussed in terms of both MOS electrical properties and the oxidation of Si_3N_4/SiO_2 /Si materials systems. The electrical properties of the plasma oxides grown at 400°C are comparable to those of thermal oxides grown at 1000°C. Preliminary results based on transmission electron microscopy observations prior to and after plasma anodisation indicate that Si_3N_4/SiO_2 strips on silicon exhibit interesting lateral oxidation behaviour and therefore Si_3N_4 may be a potential mask against plasma anodisation for advanced VLSI.

1 - INTRODUCTION

Plasma anodisation is an attractive technique for the oxidation of silicon due to the rapid oxidation rates observed at low temperatures /1/. Prolonged oxidation at high temperatures can generate defects in the silicon and these defects lead to degradation in both yield and performance of small geometry devices. An additional disadvantage of thermal oxidation is the so called "bird's beak" effect: lateral oxidation occurs underneath a masked area, thus limiting the minimum device separation which can be achieved. Although plasma anodisation has been widely investigated (see /1/ and references therein), previous studies have highlighted the severe difficulty of producing effective masks for this process /2/; most of the established masks against thermal oxidation appear to be consumed during plasma oxidation. Therefore an important issue with regard to plasma anodisation is to find materials systems in which the vertical oxidation rate of the mask is low compared to silicon and the lateral oxidation of both the mask and the silicon substrate under the mask are minimal. The present summary addresses this issue in the limited context of Si_3N_4/SiO_2 strips on Si and includes data which show that the electrical properties of plasma grown oxides at 400°C are acceptable for advanced VLSI applications.

2 - EXPERIMENTAL

The experimental apparatus used for the anodisation has been described elsewhere in some detail /3/. The anodisation chamber consists of a quartz tube containing oxygen at a pressure of approximately 150 mTorr. The plasma is initiated using an inductively coupled RF generator operating at 27 MHz and 1 kW. The sample to be oxidised is positioned on a heated metal pedestal 5-6 cm from the plasma discharge centre and forms the anode in an external d.c. circuit. Silicon wafers of up to 100mm in diameter can be uniformly oxidised.

(1) Present address: AT&T Bell Laboratories, Murray Hill, NJ 07974, USA

The response of several materials systems to plasma anodisation is being investigated by cross-sectional transmission electron microscopy (x-TEM). The materials systems include strip structures (~3μm wide) of silicon nitride/thermal oxide and silicon nitride/thermal oxide/polysilicon on (100) n-type silicon substrates. Both the silicon nitride and polysilicon structures were grown by CVD deposition. Specimens for TEM were prepared from the initial structures and from structures after oxidation by mechanical polishing to a thickness of ~30 μm, followed by ion-beam milling to perforation using 4kV Ar ions. The x-TEM observations were carried out at 200kV and all structures were imaged edge-on along the strip (i.e. [011]) under zone-axis, bright-field conditions.

3 - RESULTS

As shown in previous work /4/, the electrical properties of plasma oxides grown at 400°C are comparable to thermal oxides grown at 1000°C and the electrical breakdown strength of the plasma oxides exceeds that which is required for isolation of VLSI structures. A typical histogram of electrical breakdown data is presented in Fig. 1. The mean electrical breakdown strength is 9.5MV/cm over 90% of the probed area with a complete absence of low-field breakdown. In addition MOS capacitors fabricated using plasma oxides have electrical integrity similar to that attainable using thermal oxides. Figure 2 shows a comparison between a typical I-V curve for MOS capacitors with plasma- and thermally-oxidised dielectric layers. Conduction through the oxide in the high-field region prior to breakdown is thought to occur by Fowler-Nordheim tunneling. The differences in the curves are attributed to a higher level of positive charge either near to the interface or within the oxide. At the present time these positively charged traps can only be removed by higher temperature anneals. The slight inflection, or ledge, in the plasma grown case is probable associated with electron traps which are filled by the Fowler-Nordheim tunneling current.

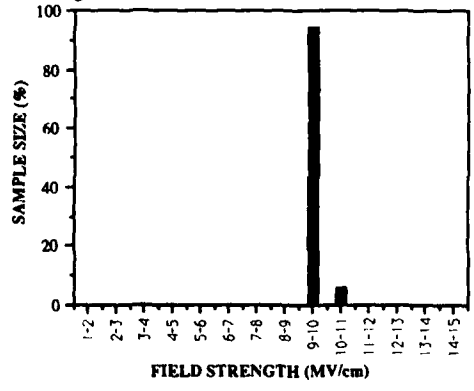


Fig. 1 - Histogram showing the field strength of 25 MOS capacitors of diameter 0.38 mm: the plasma oxide was grown at 400°C and the the oxide thickness was ~0.1 μm.

The structural features of Si_3N_4/SiO_2 strips on silicon before and after plasma anodisation at 400°C are illustrated by the x-TEM micrographs in Fig. 3. Figure 3a is before plasma anodisation and Figs. 3b and 3c are after plasma oxides of ~0.17μm and ~0.45μm thickness were grown, respectively. It can be seen from this sequence of micrographs that the silicon substrate was masked from plasma oxidation at the expense of the Si_3N_4 strip, the surface of which has been partially converted into an oxynitride layer. Other important features which are apparent from Fig. 3c include a marked reduction of the "bird's beak" effect usually produced during conventional LOCOS isolation, as well as the potential to achieve substantial improvements in the lateral encroarchment of a vertical edge, produced by thermal oxidation (e.g., as illustrated in Fig. 3d). This important phenomenon is presently under investigation. Lastly, an electrical field

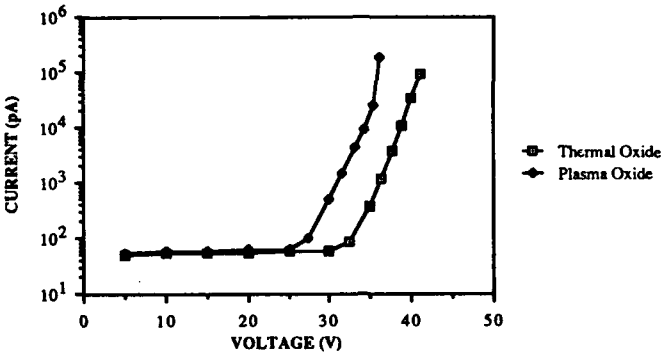


Fig. 2 - Current-voltage characteristics of plasma and thermal oxides showing comparable electrical integrity: plasma oxide was 420 nm thick and the thermal oxide was 460 nm thick.

induced variation in the oxide surface topography above the terminating edge of the silicon nitride strip is observed (Fig. 3c). This effect can be visualised quite dramatically from the low-

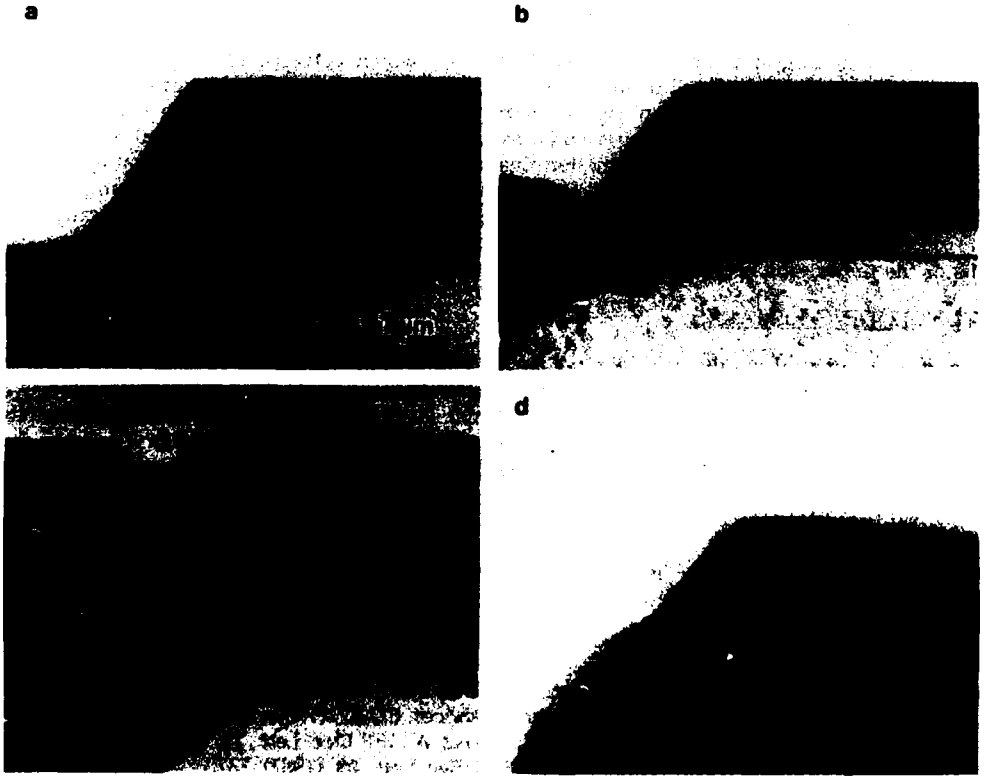


Fig. 3 - Cross-sectional electron micrographs showing similar portions of Si_3N_4/SiO_2 strips on crystalline silicon (a-c) and polysilicon (d): a) is before plasma anodisation; b) and c) are after plasma anodisation at 500°C; and d) is after thermal oxidation at 1000°C.

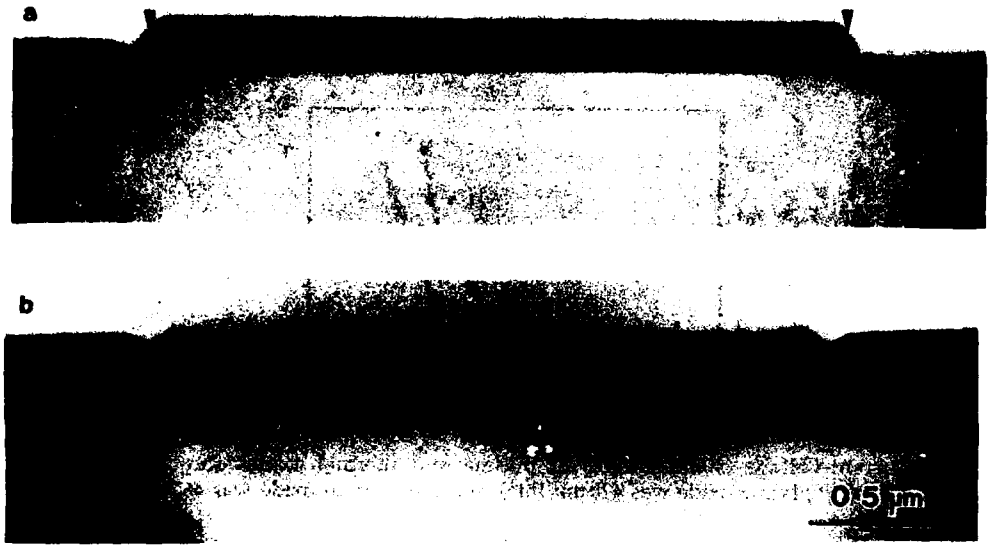


Fig. 4 - Cross-sectional electron micrographs showing the entire length of the Si_3N_4/SiO_2 strips on Si which appear in Figs. 3b and 3c. The terminations of the Si_3N_4 strip in 4a (downward arrowheads) map onto the inflections at the upper SiO_2 surface in 4b (upward arrowheads).

magnification micrographs shown in Figs. 4a and 4b which correspond, respectively, to the structures shown in Figs. 3a and 3b. It is quite clear from these observations that the end terminations of the Si_3N_4 strip in Fig. 4a map directly onto the inflections at surface of the plasma oxide in Fig. 4b. This result is taken as an indication that both the lateral and vertical oxide-field strengths are modified by the terminations of the Si_3N_4 strip. The fact that there is an anisotropy in the plasma-oxidation field strength is further supported by the observation that a SiO_2 thickness of ~0.45 μm can be achieved at the expense of ~0.06 μm of Si_3N_4 vertically and ~0.15 μm of Si_3N_4 laterally. The net result of the lateral Si_3N_4 movement is to define the vertical limits of the underlying SiO_2/silicon interface (i.e. minimal bird's beak effect).

4 - CONCLUDING REMARKS

The present structural results are taken as an indication that the lateral shift of a Si_3N_4 mask under a given plasma anodisation process will depend strongly on the geometry of the Si_3N_4 strip. Furthermore the realization of large anisotropic electric-field effects for different materials systems and varying geometries of a particular materials system may open the way for nanometre-scale structural engineering in the field of plasma anodisation.

ACKNOWLEDGEMENTS

DMM is grateful to the Royal Society for a Guest Research Fellowship. This work has been partly funded by the ALVEY Directorate.

REFERENCES

/1/ Kiermasz, A., Eccleston, W. and Moruzzi, J.L., Solid State Electron., **26**, (1983), 1167.
/2/ Ho, V.Q. and Sugano, T., IEEE Trans. Electron Devices, **ED-27**, (1980), 1436.
/3/ Taylor, S., Barlow, K.J., Eccleston, W. and Kiermasz, A., IEE Elec. Lett., **23**, (1987), 309.
/4/ Taylor, S., Eccleston, W. and Watkinson, P., IEE Elec. Lett., **23**, (1987), 732.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CHARACTERIZATION OF SiO_2 FILMS DEPOSITED BY PYROLYSIS OF TETRAETHYLORTHOSILICATE (TEOS)

W.S. WU, S. ROJAS, S. MANZINI, A. MODELLI and D. RE

SGS-Thomson Microelectronics, I-20041 Agrate Brianza, Milano, Italy

Résumé - La vitesse de déposition, l'uniformité d'épaisseur et la couverture de marche des couches de SiO_2 deposées sur plaquettes de Si par TEOS pyrolyse ont été étudiées comme fonction des paramètres de processus: température, flux de gas et pression. Des couches d'oxyde avec uniformité de ±1% sur plaquettes de 100 mm de diamètre et couverture de marche de 76% sur marches de 1μm × 1μm ont été obtenues. La caracterisation électrique des couches de SiO_2 de TEOS deposées sur simple crystal de Si ont montré une densité de charge d'environ 10^{11} cm^{-2}, une densité d'états superficiel a moitié du gap du silicium inférieure à 10^{10} eV^{-1} cm^{-2}, une barrière Si/SiO_2 de 2.6 eV et un champ de rupture de 7.5 à 9 MV/cm. Les couches de SiO_2 de TEOS deposées sur polysilicium dopé $POCl_3$ montrent proprieté d'isolation comparable aux couches d'oxyde thermique crû à la temperature de 1100°C.

Abstract - Deposition rate, thickness uniformity and step coverage of SiO_2 films deposited on Si substrates by pyrolysis of TEOS have been studied as a function of process parameters: temperature, pressure, and gas flow. Oxide films with uniformities of ±1% on 100-mm wafers and step coverage of 76% on 1 μm wide and deep trenches have been obtained. The electrical characterization of TEOS-SiO_2 films deposited on single crystal Si shows a charge density of about 10^{11} cm^{-2}, a surface state density at Si midgap lower than 10^{10} eV^{-1} cm^{-2}, a Si/SiO_2 barrier height of 2.6 eV and a breakdown field strength from 7.5 to 9 MV/cm. TEOS-SiO_2 films deposited on $POCl_3$ doped poly-Si show insulation properties comparable to that of SiO_2 films grown at 1100°C.

1. Introduction

Highly integrated semiconductor devices require insulating layers with low defectivity, high thickness uniformity and good electrical properties. Thermal oxide grown on single crystal Si has nearly ideal electrical properties, but when grown on polycristalline Si its qualities are degraded. Moreover, the high temperature required for Si oxidation give rise to severe constrains on process architecture. On the other hand, low temperature CVD oxides do not match the defectivity requirement and their electrical properties are very poor. Recently there is a renewed interest on TEOS based SiO_2 and several publications have pointed out that TEOS-SiO_2 films are promising for microelectronic applications /1,2/.

In this work we have investigated the TEOS-SiO_2 properties—deposition rate, thickness uniformity, etch rate, step coverage, stress— as a function of process conditions and we have identified two optimum processes for low and high deposition rate suitable for film thicknesses in the range from 300 to 5000 Å. The physical characteristics of these films will be briefly reported, while more space will be devoted to the electrical characterization of TEOS-SiO_2 layers on single crystal Si. Also, as one of the first conceivable application of this film is as interpoly insulating layer, we will provide an electrical characterization of the film deposited on $POCl_3$ doped polysilicon.

2. Process characteristics and film properties

The depositions were performed in a standard horisontal LPCVD system on 100-mm wafers. A schematic drawing of the system is depicted in Fig. 1. The gas flow rate from the TEOS bubbler, and thus the amount of TEOS vapor introduced into the reactor, is automatically controlled by a throttle valve and a pressure transducer while the bubbler temperature was kept constant at 50°C. The wafers were placed back to back with spacing of 9.5 mm in quarts boats with covers having no holes.

The films were characterised by measuring refractive index and thickness (Rudolph ellipsometer, Nanospec). etch rate in P-etch solution (30 cc HNO_3 70%, 51 cc HF 40% and 900 cc H_2O), film-induced stress (Canon LSF 500), and step coverage (SEM). Measurements were performed on films as deposited and after anneal (950°C, 15 min in N_2). Density and infrared absorption measurements were performed too.

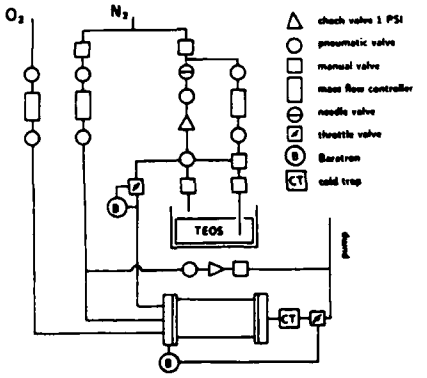


Fig. 1. Schematic drawing of the experimental setup used for the deposition of TEOS-SiO_2.

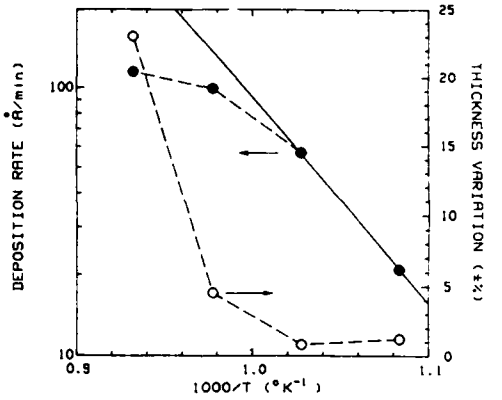


Fig. 2. Deposition rate (left) and thickness variation (right) as a function of deposition temperature.

Table I. TEOS processes (low - high deposition rate).

Process	Dept. rate Å/min	Temperature °C	Pressure mTorr	TEOS pressure Torr	Intra-wafer uniformity	Inter-wafer uniformity
A	60	700 ±1	700	3	1%	3%
B	115	720 ±1	700	4	1.5%	3.5%

Process optimisation has been accomplished by examining the reaction rate sensitivity against temperature, pressure, and TEOS flowrate. The initial choise for these parameters was: temperature = 700°C, pressure = 400 mTorr, TEOS line pressure = 3 Torr. Each parameter has been varied one at a time while keeping the other constant. Fig. 2 reports the Arrhenius plot of the deposition rate. From this figure it can be desumed an activation energy of 37 kcal/mol at low temperature while a saturating behavior is observed at high temperature, in agreement with previously reported data /1,3,4/. The thickness uniformity within wafer is also shown in Fig. 2. For temperatures lower than 745°C thickness variations are within ±5%.

The effect of pressure variations on deposition rate is small. The deposition rate increases linearly from 43 Å/min at 200 mTorr to 65 Å/min at 800 mTorr, while within-wafer uniformity is nearly constant and equal to about ±1%. At lower pressure the thickness variation within wafers of the same run is about ± 13%, while in the pressure range between 400 and 800 mTorr is about ± 6%.

The deposition rate increases sublinearly with the TEOS line pressure and is 21 Å/min at 1 Torr,57 Å/min at 3 Torr and 70 Å/min at 5 Torr. At 1 Torr the film thickness decreases along the wafer diameter from the nearest to the furthest position relative to the TEOS injectors in the reactor, indicating a depletion of reactants. At higher TEOS pressures the thickness uniformity is good (± 1%).

We have characterised the step coverage of TEOS-SiO_2 films by measuring the ratio of film thickness on the wall to the one on the top of 1-μm wide and deep trenches. Step coverage increases with increasing process pressure from 65% at 400 mTorr to 76% at 700 mTorr (deposition temperature = 700°C) and decreases with increasing temperature—at 745°C and a pressure of 700 mTorr the step coverage is 63%.

On the bases of the results of process characterization two optimum process conditions have been selected for low and high deposition rate (process A and B of Table I). Additional fine tuning of wafer boats position in the reactor and a sligthly tilted temperature profile (±1°C) has allowed to improve the thickness uniformity within the wafers of a run to about ±3%. Table II reports the physical properties of SiO_2 films obtained by the two processes compared with that of a SiO_2 film grown at 1100°C in dry O_2 ambient.

Table II. Physical film characteristics.

Process	Etch rate (P-etch) Å/min	Step coverage	Stress 10^9 dynes/cm^2	Refractive index	Density gr/cm^3	Si-O stretching band, cm^{-1}
A as dep.	15	76%	0.8 C	1.446	2.17	1063
A annealed	3	—	2.8 C	1.450	2.20	1074
B as dep.	10	74%	1.2 C	1.445	2.17	1063
B annealed	3	—	3.6 C	1.450	2.22	1074
Thermal	1.6	—	—	1.460	2.25	1080

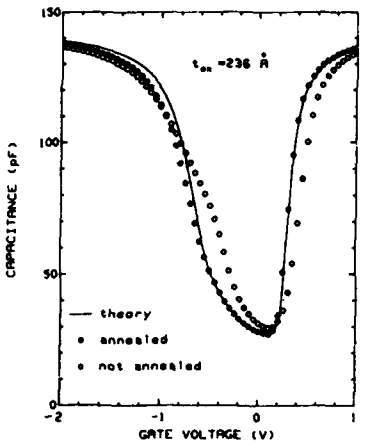


Fig. 3. Effect of annealing in inert ambient on the $C-V$ characteristic of TEOS-SiO_2 on single crystal Si. Continuous line: theoretical $C-V$ curve, shifted of 0.22 V along the V_G axis.

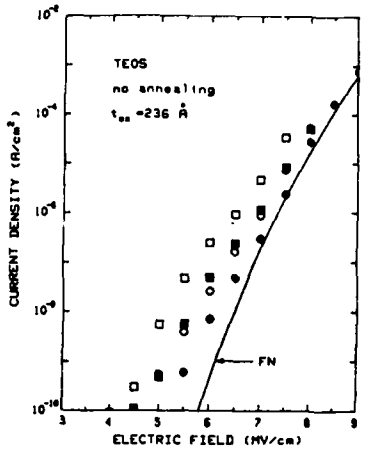


Fig. 4. $J-E_{ox}$ curves of TEOS-SiO_2 on single crystal Si. Closed symbols: T=23°C. Open symbols: T=200°C. Circles: $V_G > 0$. Squares: $V_G < 0$. Continuous line: $J-E_{ox}$ curve of thermal SiO_2.

3. Electrical characterization

The devices used for the electrical characterization were large area (10^{-3} cm²) MOS capacitors fabricated on (100)-oriented, p-type, 1.7 – 2.5 Ω cm single crystal Si, n-type (phosphorous implanted) single crystal Si, and gas phase ($POCl_3$) doped poly-Si. TEOS-SiO_2 layers 200- to 1000-Å thick were deposited by using both processes A and B. Some samples were prepared according to process A but at the higher temperature of 745°C. An annealing step (15 min at 950°C in N_2) was performed on some wafers immediately after the TEOS-SiO_2 deposition. A simple n^+ polysilicon-gate process was used in the device fabrication. For comparison, MOS capacitors with thermal SiO_2 as insulating layer were processed in parallel. The oxide thickness of TEOS-SiO_2 films was determined by $C-V$ measurements assuming $\kappa_{ox}=3.9$ for the dielectric constant. The agreement between electrical and optical measurements of the film thicknesses was within 5%.

In Fig. 3 are shown the low frequency $C-V$ curves for as deposited and annealed samples prepared according to process A. Also shown is the theoretical $C-V$ curve, shifted of +0.22 V along the gate voltage axis to match the annealed sample curve in the inversion region. The interface state density D_{it} was evaluated by both the quasi-static /5/ and the conductance /6/ method. The unannealed sample has a u-shaped distribution with $D_{it}=2\times10^{10}$ eV^{-1} cm^{-2} near midgap and $D_{it}=4\times10^{11}$ eV^{-1} cm^{-2} at 0.2 eV above the valence band edge. The capture cross section is of about 2×10^{-17} cm². Annealing in inert ambient reduces the interface state density near midgap of about an order of magnitude. The short time constant of the interface states near flatband and in weak accumulation limits the experimentally accessible region of the band gap, but the large deviation from the theory of the $C-V$ curve in weak accumulation indicates a high density of interface states just above the valence band edge, even for the annealed sample. The flatband voltage shift ΔV_{FB} obtained from 1-MHz $C-V$ characteristics corresponds to a negative effective charge of 3×10^{11} cm^{-2} for unannealed samples and of about 10^{11} cm^{-2} for annealed samples, independent of oxide thickness. This finding, together with the high interface state density near the valence band edge, suggests that the measured ΔV_{FB} is mainly due to the charge of these interface states—filled with electrons at $V_G=V_{FB}$—rather than to electrons trapped in the bulk of the oxide.

The current density – electric field $J-E_{ox}$ characteristics of the unannealed sample are reported in Fig. 4. Very similar characteristics were measured on annealed samples. The nearly temperature independent conduction suggests that charge transport in TEOS-SiO_2 is controlled by Fowler-Nordheim electron tunneling trough the Si/SiO_2 interface. The low-field deviation from the ideal Fowler-Nordheim conduction for positive gate voltage is likely due to a decrease of the barrier height from a value close to that of thermal SiO_2 near the Si/SiO_2 interface to a lower value in the bulk of the film. This hypothesis is confirmed by the higher conduction observed for negative gate voltage, when electrons are injected from the poly-Si/SiO_2 interface. The asymmetrical behavior with respect to the gate voltage polarity is enhanced for thicker oxides and depends on process conditions, as shown in Fig. 5 where the $J-E_{ox}$ characteristics of 900-Å thick samples are reported in a Fowler-Nordheim plot. For positive gate voltage only one curve is shown, the other two beeing nearly coincident. Measurments performed on thermal oxide of similar thickness are shown for comparison.

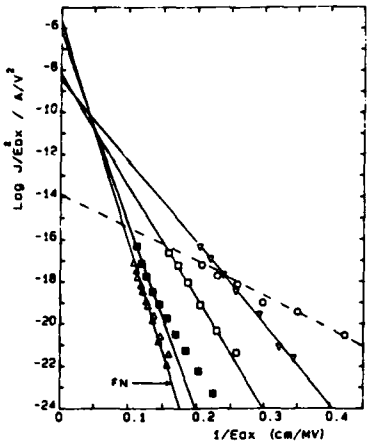


Fig. 5. Fowler-Nordheim plot of the $J - E_{ox}$ characteristics of TEOS-SiO_2 on single crystal Si. Closed symbols: $V_G > 0$. Open symbols: $V_G < 0$. Squares: $T_{dep} = 700°C$. Circles: $T_{dep} = 720°C$. Triangles down: $T_{dep} = 745°C$. Triangles up: thermal SiO_2.

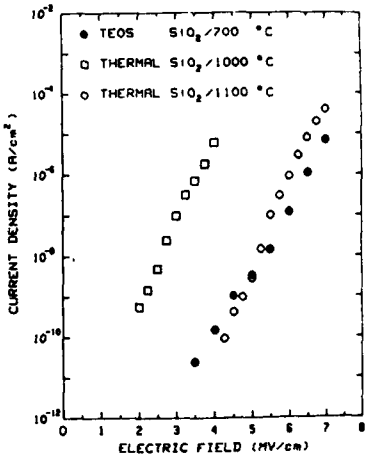


Fig. 6. $J - E_{ox}$ characteristics of SiO_2 films on $POCl_3$ doped polysilicon. Closed symbols: TEOS-SiO_2 deposited at 700°C. Open squares: thermal SiO_2 grown at 1000°C. Open circles: thermal SiO_2 grown at 1100°C.

The effective barrier height /7/ evaluated from the slope of the characteristics, assuming an electron effective mass equal to 0.5 the electron free mass, is of 2.9 eV for thermal oxide, 2.6 eV for TEOS-SiO_2 at high positive gate voltage, 1.9 eV for 700°C deposited TEOS-SiO_2 at negative gate voltage and 1.5 eV for 745°C deposited TEOS-SiO_2 at negative gate voltage. Samples deposited at 720°C (process B) show an anomalous behavior which is also apparent from the time evolution of the current at a fixed negative gate voltage: for these samples the current decreases with time even in the low field regime following a power law $J \propto t^{\alpha}$ with α between 1/2 and 1/3 depending on the electric field. For the other two samples the current was remarkably constant in time with the exception of the higher fields tested.

Breakdown measurements were performed using a staircase voltage ramp with positive gate polarity on TEOS-SiO_2 films deposited on n-type Si. The average breakdown field ranges from 9 MV/cm for thin (200 Å) films to 7.5 MV/cm for thicker (900 Å) films. No dependence on process conditions was observed.

In order to evaluate the performances of TEOS-SiO_2 films in practical applications we have performed defectivity, breakdown field and $I - V$ measurments on samples deposited according to process A on $POCl_3$ doped polysilicon. Current – field characteristics are compared in Fig. 6 with those of thermal oxide grown at 1000 and 1100°C. The oxide thickness was of about 500 Å for all samples. Film defectivity was tested on large area capacitors at an electric field of 3 MV/cm. To pass the test a leakage current density lower than 1 mA/cm² was required. A defect density of 4.5 defects/cm² for 500-Å thick samples and of 1.2 defects/cm² for 800-Å thick samples was found. Although thinner films have a higher defectivity, their dielectric breakdown strength is better than for thicker samples. The average breakdown fields are of 7 and 5 MV/cm for thin and thick films, respectively.

References

/1/ Becker,F.S., Pawlik,D., Ansinger,H. and Spitzer,A., J. Vac. Sci. Technol. B 5 (1987) 1555.
/2/ Levin,R.M. and Evans-Lutterodt,K., J. Vac. Sci. Technol. B 1 (1983) 54.
/3/ Adams,A.C. and Capio,C.D., J. Electrochem. Soc. 126 (1979) 1042.
/4/ Hupperts,H. and Engl,W.L., IEEE Trans. Electron Devices ED-26 (1979) 658.
/5/ Sah,C.T., Solid-State Electron. 19 (1976) 975.
/6/ Nicollian,E.H. and Goetzberger,A., Bell Syst. Tech. J. 45 (1967) 1055.
/7/ Weinberg,Z., J. Appl. Phys. 53 (1982) 5052.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

RAPID THERMAL PROCESSING OF ARSENIC-IMPLANTED POLYSILICON ON VERY THIN OXIDE

J.Y.-C. SUN, R. ANGELUCCI, C.Y WONG, G. SCILLA and E. LANDI*

IBM Research Division, Thomas J. Watson Research Center, PO Box 218, Yorktown Heights, NY 10598, U.S.A.
**CNR-LAMEL Institute, I-40126 Bologna, Italy*

Résumé - Dans le cadre d'un procédé CMOS à grilles en silicium polycristallin (n^+ et p^+), le recuit thermique rapide a été utilisé avec succès pour une grille en silicium polycristallin implanté avec de l'arsenic et deposé sur 7 nm de SiO_2. Le travail de sortie de la grille que nous avons obtenu correspond à la valeur théorique attendue. Les états d'interface et les charges fixes dus au recuit rapide peuvent être eliminés par un recuit à 500°C sous forming gas. Les mesures de tensions de claquage aussi bien instantanées qu'en fonction du temps montrent que l'integrité de l'oxide de grille (7 nm) est preservée après le recuit rapide. La diffusion de l'arsenic dans le silicium polycristallin lors du recuit rapide est en accord avec les résultats de la litterature pour des recuits conventionnels.

Abstract - We demonstrated the feasibility and advantages of using rapid thermal annealing (RTA) to achieve a proper work function for arsenic-implanted polysilicon gate on 7 nm SiO_2 in a dual work function (n^+ and p^+) poly-gate CMOS process. Interface states and fixed oxide charge due to RTA can be annealed out at 500°C in forming gas. Time-zero and time-dependent breakdown results show that the integrity of 7 nm gate oxide can be preserved after RTA. The diffusivity of arsenic in polysilicon under RTA is found to be consistent with literature data from conventional furnace anneals.

1 - INTRODUCTION

Doping of the polysilicon gate by using the source/drain ion implantation step for a dual work function (n^+ and p^+) poly-gate CMOS process is desirable and challenging under reduced implant dose and thermal budget for future sub-micrometer CMOS VLSI applications /1/. Recently rapid thermal processing (RTP) or annealing (RTA) has been studied as an alternative or supplement to the conventional furnace annealing for minimizing the heat treatment necessary to diffuse and activate dopants. It has also been explored for silicide, glass reflow, oxidation and nitridation processes /2/. In this paper we report the results of a comprehensive study on the rapid thermal annealing of arsenic-implanted polysilicon gate electrode on 7 nm gate oxide. It is demonstrated that RTA is required for the activation of arsenic to get a degenerately doped n^+ polysilicon, while maintaining shallow source and drain junctions in a dual work function poly gate CMOS process. SIMS profile (chemical concentration) should be supplemented with carrier concentration profile to investigate the dependence of gate work function on arsenic dose and annealing conditions. The effective diffusivity of arsenic in polysilicon under RTA conditions is extracted. The effects of RTA on the quality and integrity of the underlying 7 nm thin gate oxide have been identified and methods to anneal out RTA induced damage are reported.

2 - EXPERIMENTAL

Aluminum contacted polysilicon-gate capacitors were fabricated on 1 Ω-cm (100) p-type Si wafers. Immediately after 7 nm dry oxide was grown, 250 nm polysilicon was deposited at 610°C by low pressure chemical vapor deposition (LPCVD). The polysilicon gate was doped by arsenic ion implantation (2×10^{15} or $4\times10^{15}cm^{-2}$) through 10nm screen oxide at 30 KeV. Different combinations of furnace annealing and RTA (AG-2146) in nitrogen were performed to redistribute and activate arsenic. Splits in RTA gas atmosphere (N_2, Ar, and forming gas) and ramp rate conditions have been included. No post-aluminum annealing was performed. The flat band voltage (V_{FB}) and interface states density D_{it} were obtained by using high-frequency (HFCV) and quasi-static C-V (QSCV) measurements. Arsenic concentration was measured by SIMS. Carrier concentration and mobility were determined by Hall effect and resistivity measurements using a Van der Pauw geometry fabricated on polysilicon. Time zero oxide breakdown (TZBD) was measured by ramping the current and recording the snap-back voltage. Time dependent breakdown (TDBD) measurements were performed by constant current stress.

3 - ARSENIC DIFFUSION AND ACTIVATION

The difficulty of using conventional furnace annealing to achieve degenerately doped n-type polysilicon with reduced implant dose and annealing temperature is illustrated in Fig.1, case A, where the chemical (SIMS, solid line) and carrier concentration (dots) profiles of As implanted ($2\times10^{15}cm^{-2}$ dose) 250 nm polysilicon are shown after furnace annealing at 850°C for 30 min. High diffusivity of As along the grain boundaries /3/ at 850°C gives a high concentration of As (about $2\times10^{19}cm^{-3}$) at the gate oxide interface. However, the carrier concentration profile in Fig.1, case A, clearly shows low activation of As (4.2% of the implant dose) due to severe As segregation into the grain boundary at low temperature /4/. A high-resistivity polysilicon layer exists near the gate oxide interface due to limited dopant activation and low carrier mobility. The work function of this polysilicon deviates from that of a degenerately doped n^+ polysilicon as indicated by the measured value of V_{FB} (-0.6V) which is more positive than the theoretical value (-0.92V) for n^+ polysilicon on 7 nm oxide and $2\times10^{16}cm^{-3}$ boron-doped substrate. Moreover, a reduction of MOSFET transconductance can result from the formation of a depletion layer near the polysilicon/oxide interface under positive gate bias due to low active As concentration.

SIMS profile (solid line) of Fig.1, case B, shows that an additional RTA at 1100°C for 10 sec in N_2 redistributes As uniformly across the 250 nm polysilicon film. The nearly flat carrier concentration profile (dots) with an average value of $1.7\times10^{19}cm^{-3}$ clearly demonstrates the effectiveness of the additional RTA (high temperature, short time) in activating a large amount of As, up to 21% of the implant dose.

Fig.2 shows the As concentration (solid lines) and the carrier concentration (dashed lines) in polysilicon as a function of RTA temperature without any prior furnace anneals. RTA at 1100°C for 10 sec is necessary to distribute As uniformly throughout the 250 nm polysilicon film as indicated by curve C. The redistribution of As in 250 nm polysilicon is incomplete at lower RTA temperatures as evidenced by curve B (1050°C) and curve A (1000°C), where a considerable amount of As remains near the top surface of the polysilicon film. It is worth noting that the arsenic concentration increases towards the polysilicon-oxide interface after RTA at 1100°C as seen in the SIMS profiles of Fig.2, case C, and Fig.1, case B. This effect is due to the grain structure of the polysilicon. Cross sectional TEM reveals columnar grains with high density of smaller grains near the polysilicon/oxide interface, meaning increased grain boundary area toward that interface. This results in an increased amount of arsenic confined at or near the grain boundary and therefore an increase in total arsenic (SIMS) concentration toward the polysilicon/SiO_2 interface.

The average carrier concentrations are $0.9\times10^{19}cm^{-3}$, $1.1\times10^{19}cm^{-3}$, and $1.8\times10^{19}cm^{-3}$ after RTA at 1000°C, 1050°C and 1100°C, respectively. Although the average carrier concentrations are within a factor of two for the three RTA temperatures, the carrier concentration profile depends strongly on RTA temperature. As shown in Fig.2, RTA at 1100°C is required to activate a large amount of arsenic close to the oxide interface. The carrier concentration after RTA at 1000°C is high at the surface ($\simeq4\times10^{19}cm^{-3}$), but falls off sharply beyond a depth of 50 nm from the polysilicon surface, consistent with measured SIMS As concentration, grain size, and the segregation effect. It is therefore clear that neither SIMS profile nor average carrier concentration is sufficient, and they must be supplemented with carrier concentration profile measurements in order to investigate the dependence of polysilicon work function on implant dose and annealing conditions, even in the case of RTA.

The effective diffusivity of As along the polysilicon grain boundary was extracted from SIMS profiles using a large grain (200-300 nm) polysilicon film /5/ so that only a very small fraction of the implanted arsenic impurities near the surface, i.e. those within a diffusion length of the grain boundary can reach the grain boundary and diffuse rapidly through it. In such a case a simple one dimensional limited-source diffusion model applies, and the tail part of the SIMS profile fits well to a Gaussian curve. The RTA and furnace diffusivity data obtained in this way are plotted in Fig.3 together with some published data of arsenic in polysilicon /3,6/. The diffusivity data spreads at least over two order of magnitude due to the strong dependence of grain boundary diffusion on grain structure and impurity or dopant content. Our data from either furnace annealing or RTA is within the spread. The activation energy calculated from our limited data is 3.5 eV, consistent with reported values /3,6/, suggesting that the diffusion of arsenic under RTA and furnace annealing follow the same physical mechanism. The broadening of the arsenic peak in the SIMS profile at the surface of the same large grain (200-300 nm) polysilicon after RTA, mainly due to the movement of the dopant inside the grain, may be used to estimate the diffusion length of high concentration of arsenic in single crystalline silicon. A value of 15 nm for $\sqrt{Dt}$ under 1100°C 10 sec RTA was obtained, which agrees with extrinsic arsenic diffusivity data /7/. This is evidence that dopant redistribution in the substrate is very little during such polysilicon RTA processes.

4 - INTERFACE STATES AND OXIDE CHARGE

There is an increase in interface states density after RTA in N_2 as shown in Fig.4. Two peaks have been

observed in both poly and Al gate capacitors, one located at roughly 0.1eV above midgap energy with a density of 10^{13} $eV^{-1}cm^{-2}$ and the other at 0.35eV below midgap with a density of 2×10^{12} $eV^{-1}cm^{-2}$. This increase in D_{it} is consistent with published data on post oxidation furnace annealing in nitrogen /8/. It has little dependence on RTA temperature above 800°C, gate material, or RTA temperature ramp rates. There is no D_{it} increase if RTA is performed in forming gas. The RTA induced interface states may be explained in terms of the reduction of the concentration of water related groups and the increase of the density of Si dangling bonds at the SiO_2/Si interface during RTA in inert gas ambient. These interface states persisted after 400°C 30 min anneal in forming gas. They are completely annealed out after a 500°C pre-aluminum FG anneal for 30 min.

RTA also induced a positive V_{FB} shift that can be attributed to changes in the density of interface states and fixed oxide charge. A reduction in positive fixed oxide charge after high temperature anneal in an inert ambient was reported previously /9,10/. The change in fixed oxide charge is further confirmed by performing RTA in forming gas, where positive shift is observed without any increase in D_{it}. Another evidence of the change in fixed oxide charge is that D_{it} increase is independent of the RTA temperature ramp rate while V_{FB} shift is bigger for higher ramp rate. Our data shows that the positive V_{FB} shift can be completely recovered after a pre-metallization FG anneal at 500°C for As-doped n^+ polysilicon-gate capacitors.

Fig.5 shows V_{FB} and carrier concentration data as a function of thermal process. The first case (850°C 30 min) corresponds to case A of Fig.1 where the low (-0.6 V) V_{FB} value is due to low carrier concentration at the polysilicon-oxide interface. The second and third case correspond to RTA at 1100°C with or without a prior furnace anneal. In these cases the carrier concentration at the interface is high enough for degenerately doped polysilicon conditions but V_{FB} is still low, which is due to RTA induced changes in oxide charge and interface states. An ideal V_{FB} value (-0.92V) is achieved if an additional anneal is done after RTA, i.e., at 500°C for 30 min in forming gas (the one to anneal out interface states) or at 800°C for 15 or 60 min in N_2. In spite of the decrease of carrier concentration after 800°C furnace annealing, the V_{FB} is still the expected one, suggesting that the polysilicon is degenerately doped from V_{FB} point of view if carrier concentration at the oxide interface is at least $10^{19}cm^{-3}$.

5 - INTEGRITY OF 7 nm GATE OXIDE

Fig.6 shows that RTA at 1100°C for 10 sec in N_2 increases the average time-zero oxide breakdown field by roughly 0.4 MV/cm. Additional anneal at 800°C for 1 hr in N_2 has no extra effect on the breakdown strength. Fig.7 shows little change in the oxide breakdown charge density after RTA at 1100°C under constant current (Fowler-Nordheim tunneling) stress measurements. The improvement in breakdown strength is consistent with published data on post-oxidation annealing effects using conventional furnaces /8/. which can be attributed to the reduction of water-related groups and the modification of sub-oxide charge state concentration /8/.

6 - CONCLUSION

In conclusion, we demonstrated the need, feasibility, and advantages of using rapid thermal annealing to achieve a proper As-doped n^+ polysilicon work function on very thin gate oxide with little dopant redistribution in the substrate. The importance of carrier concentration data to supplement SIMS data has been demonstrated. Diffusivity data of arsenic in polysilicon for furnace annealing and RTA has been extracted. The effects of RTA on the interface states, fixed oxide charge, and breakdown strength have been identified. The integrity and quality of 7 nm gate oxide after RTA of the polysilicon gate can be preserved by means of a suitable low temperature furnace anneal.

REFERENCES

/1/ J.Y.-C. Sun, et al., IEEE Trans. Electron. Dev., ED-34, (1987) 19
/2/ See for example, Rapid Thermal Processing of Electronic Materials, ed. by S.R. Wilson, R. Powell and D.E. Davies, Mat. Res. Soc. Proc. 92, (1987)
/3/ B. Swaminathan, et al., Appl. Phys. Lett., 40, (1982) 795
/4/ A. Carabelas, et al., J. Phys., 43-C1, (1982) 187
/5/ R. Angelucci, et al., Mat. Res. Soc. Symp. Proc., 106, (1988)
/6/ M. Arienzo, et al., J. Appl. Phys., 55, (1984) 365
/7/ R. B. Fair, in "Silicon integrated circuits part B", D. Kahng, ed., (1981) 28
/8/ L. Dori, et al., Mat. Res. Soc. Symp. Proc., 76, (1987) 259
/9/ A. Kamgar, et al., J. Appl. Phys., 51, (1987) 1251
/10/ B.G. Deal, et al., J. Electrochem. Soc., 114, (1967) 266

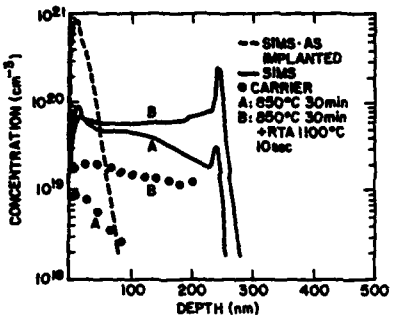


Fig.1: SIMS (solid lines) and carrier concentration (dots) profiles of As in 250 nm polysilicon *on top* of 7 nm gate oxide after 850°C 30 min furnace anneal (A) and additional RTA at 1100°C for 10 sec (B). Dashed curve represents the as-implanted As profile (2×10^{15}As/cm^2, 30 KeV)

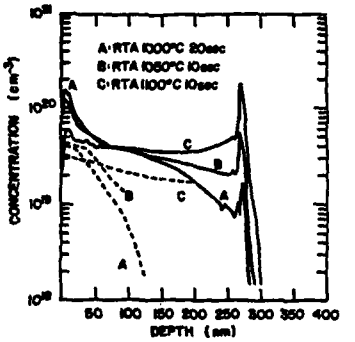


Fig.2: SIMS (solid line) and carrier (dashed line) concentration profiles of As after 1000°C (A), 1050°C (B), and 1100°C RTA (C) without any prior furnace anneal. 250 nm polysilicon; 2×10^{15}As/cm^2, 30 KeV.

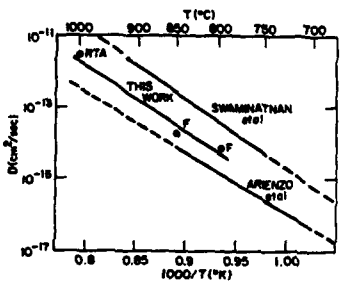


Fig.3: Effective diffusivity of As along the polysilicon grain boundary.

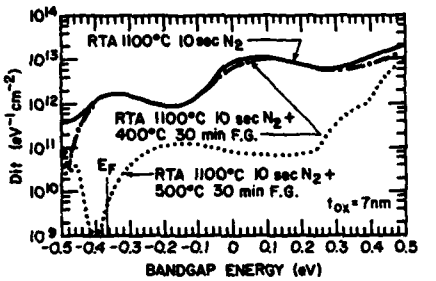


Fig.4: Interface state density after RTA and the effects of subsequent pre-aluminum anneals in forming gas.

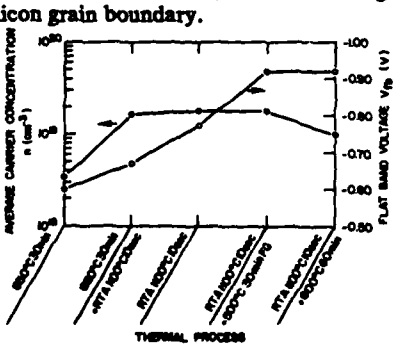


Fig.5: Flatband voltage and average carrier concentration vs. thermal process. 250 nm polysilicon; 2×10^{15}As/cm^2, 30 KeV.

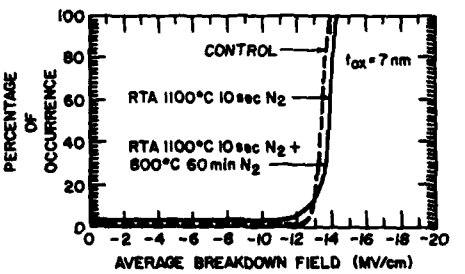


Fig.6: Time-zero oxide breakdown statistics as a function of thermal process history. The polysilicon of control devices was doped degenerately with 4×10^{15}As/cm^2 at 30 KeV and 950°C annealing for 30 min in N_2.

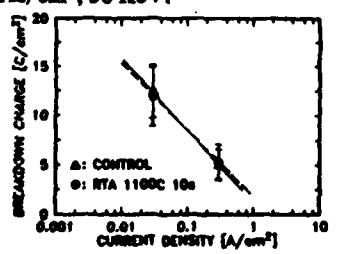


Fig.7: Breakdown charge density vs. stress current for the control and RTA samples of Fig. 6.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

INFLUENCE OF THE FABRICATION CONDITIONS ON THE p⁺-$TaSi_2$/POLY-Si GATE QUALITY

C. MAZURE, U. SCHWALKE, F. NEPPL, P. EICHINGER* and M. METZGER*

Siemens AG, Corporate Research and Development, Microelectronics, D-8000 München 83, F.R.G.
**GeMeTec, D-8000 München 60, F.R.G.*

Résumé - L'effet des différentes étapes de fabrication du poly-siliciure sur la qualité resultante du dopage p⁺ de la grille est examiné. En particulier on étudie la diffusion de bore en correlation avec la contamination en oxigène. Le rôle de l'interface $TaSi_2$/poly-Si et de l'atmosphere de recuit sur la redistribution du bore est montré.

Abstract - The effect of different polycide fabrication steps on the resulting p⁺ gate quality is investigated. In particular the boron diffusion is studied in correlation with the oxygen contamination. The role of the quality of the $TaSi_2$/poly-Si interface and of the annealing atmosphere on the boron redistribution are pointed out

1 - INTRODUCTION

The $TaSi_2$/poly-Si bilayers constitute a well suited system for the fabrication of p⁺-gates /1/. Nevertheless controversial results concerning the boron redistribution after implantation into the $TaSi_2$ layer have been reported /1,2/, questioning the p⁺ $TaSi_2$ polycide quality, i.e. the p⁺ work function ϕ_m reproducibility. Earlier experiments suggested a correlation of the boron profiles with the oxygen distribution in $TaSi_2$ /3/. Recently, evidence for TaB_2 formation after high dose implantation ($2x10^{16}cm^{-2}$) has been reported /4/. However, the effect of oxygen contamination in $TaSi_2$ or Ta-B compound formation on the B outdiffusion from $TaSi_2$ into the poly-Si is yet unclear.

In this paper we first discuss the effects of an interfacial oxide on the B transport across the $TaSi_2$/poly-Si interface and point out the importance of poly-Si surface cleaning prior to $TaSi_2$ deposition for good p⁺ polycide behavior. Next, the role of oxygen on the B diffusion within the silicide is discussed. Finally the effects of annealing ambient on the B redistribution in the p⁺ $TaSi_2$ polycide are investigated.

2 - EXPERIMENTAL

MOS-capacitors (0.01 to 8 mm^2) with $TaSi_2$/poly-Si gates were fabricated on n- or p-type (100) Si. The gate oxide thickness was 25 nm. The poly-Si cleaning prior to $TaSi_2$ sputtering consisted of a buffered HF dip and an in-situ sputter cleaning with Ar ions. The silicide films were prepared by Ta and Si cosputtering to a thickness of 200 nm. The Ta-Si sandwich thickness was ≈ 1.3nm. After $TaSi_2$ deposition on poly-Si (100 nm) the wafers were covered with a CVD-oxide (100 nm). B was implanted either into $TaSi_2$ or into poly-Si prior to Ta-Si sputtering. The B doses used were $5x10^{15}cm^{-2}$ and $2x10^{16}cm^{-2}$. Finally the samples were annealed at 900°C in a N_2 or Ar/H_2 atmosphere.

SIMS profiles of B and oxygen were recorded using oxygen (O_2^+) and cesium (Cs^+) primary ions, respectively. For Capacitance-Voltage C(V) measurements the oxide cap was removed. SIMS measurements were done before and after the oxide cap etching; the SIMS results for samples with and without oxide cap did not differ. B implantation standards for poly-Si and $TaSi_2$ were used to calibrate the B SIMS signal. Unimplanted polycide samples were used to determine the background SIMS signal. For B (O^+ primary ions) the background was $3-5x10^{17}cm^{-3}$ in $TaSi_2$ and $2-4x10^{16}cm^{-3}$ in poly-si. The oxygen profiles (Cs^+ primary ions) were limited by a background of about 10^3 counts.

3 - RESULT AND DISCUSSION

3.1 - TaSi2/POLY-Si INTERFACE

Earlier experiments with $TaSi_2$ on mono-Si have shown, that interfacial oxides impede B transport from $TaSi_2$ into mono-Si /5/ for T≤900°C. At the barrier brakes down /2,5/, but this temperature is not capatible with modern CMOS processes. *In order to elucidate the role of interfacial oxide on poly-Si on* the B transport across the $TaSi_2$/poly-Si interface at 900°C, samples with and without the poly-Si surface cleaning were investigated.

The B profiles for the samples with appropiate poly-Si surface cleaning and with interfacial oxide are shown in Fig. 1. The B profiles can be roughly divided into a region af low B diffusivity (the implanted region), and a tail region of fast diffusing B (grain boundary enhanced diffusion) /1,6/. There is no significant difference between the low diffusivity regions of both type of samples. In contrast the tail regions differ significantly (Fig. 1). For the clean interface sample the flat tail saturates at $1\text{-}2\times10^{19}cm^{-3}$ /1,6/. For the sample with interfacial oxide the B concentration in the tail region continuously decreases. Additionally a pronounced peak is observed at the $TaSi_2$/poly-Si interface. *In the poly-Si the B doping is very low.* At the poly-Si/gate-oxide interface the B concentration is as low as or lower than $1\times10^{17}cm^{-3}$.

A B pile-up is a characteristic of interfaces between thick oxides and silicide /7,8/. Fig. 2 shows an example of B accumulation at the oxide-cap/$TaSi_2$ interface. By increasing the implantation energy the B profile peak is moved deeper into $TaSi_2$, and the B accumulation near the oxide cap (100nm) is separated from the implanted region. In the case of interfacial oxide the resulting B profile exhibits a similar B pile-up at the $TaSi_2$/poly-Si interface (Fig. 1, curve b), although in this case the oxide thickness is only 3-5nm. Furthermore, the charactersitc B distribution in the tail region for the $TaSi_2$/poly-Si bilayers (Fig. 1, curve a) is altered. In the sample with interfacial oxide the B profile within the $TaSi_2$ layer strongly resembles those obtained for $TaSi_2$ between two thick oxide layers /7/. The fast diffusing B accumulates in $TaSi_2$ near and at the $TaSi_2$/poly-Si interface inducing a B concentration decrease in the tail region deeper in the $TaSi_2$ layer.

C(V) measurements confirmed the SIMS results. For the sample with interfacial

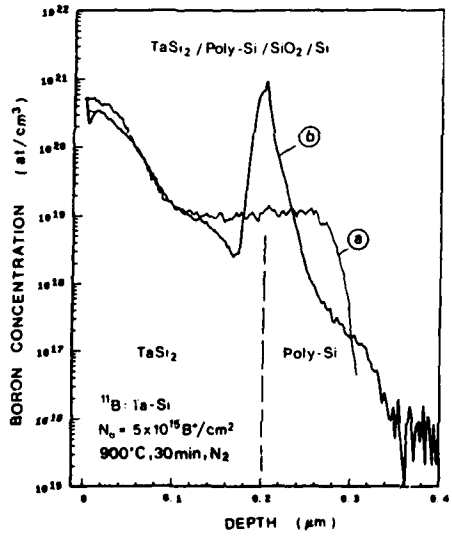


Fig. 1

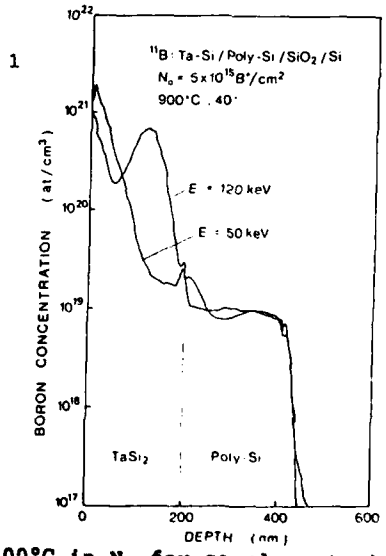


Fig. 2

Fig. 1 Boron profiles after annealing at 900°C in N_2 for samples a) with poly-Si surface cleaning and b) with interfacial oxide.
Fig. 2 Boron profiles after annealing in N_2 for two implantation energies. Boron was implanted into the as-deposited silicide.

oxide the C(V) characterstic are shifted to more negative gate voltage values indicating a lower p^+-ϕ_{MS}. This consistent with the insufficient doping of the underlying poly-Si from SIMS measurements.

3.2 - OXYGEN CONTAMINATION

Fig. 3 shows the oxygen profiles for the samples with poly-Si cleaning and with interfacial oxide. The corresponding B profiles are shown in Fig. 1. The oxygen concentration in the as-deposited $TaSi_2$ layer has a constant value throughout the layer and the poly-Si is largely oxygen free. For annealed samples with poly-Si surface cleaning a decrease of the oxygen concentration towards the oxide/$TaSi_2$ and $TaSi_2$/poly-Si interfaces is observed. Additionally, oxygen diffuses into the poly-Si increasing its oxygen content by several orders of magnitude. For the sample with interfacial oxide the oxygen concentration remains quite constant in $TaSi_2$ and increases near the oxide/$TaSi_2$ interface. The poly-Si remains practically oxygen free. The interfacial oxide appears as a small oxygen peak at the $TaSi_2$ interface. Apparently an interfacial oxide also hinders the oxygen diffusion across the interface.

The oxygen profiles for samples with poly-Si cleaning with and without B implantation were alike. Moreover for a B dose increase from $5x10^{15}cm^{-2}$ to $2x10^{16}cm^{-2}$ no dose effect related to the oxygen distribution is found. It follows that neither the presence of B nor the implantation affects the oxygen distribution in the polycide. A similar oxygen distribution is also expected in the case of B implantation into the poly-Si prior to $TaSi_2$ deposition. In this way implantation effects within $TaSi_2$ are avoided. As shown in Fig. 4, a widely uniform B concentration is obtained and again a B pile-up near the oxide/$TaSi_2$ interface is observed. The oxygen concentration near the oxide/$TaSi_2$ interfaces can decrease, increase or remain at a constant value

Fig. 3

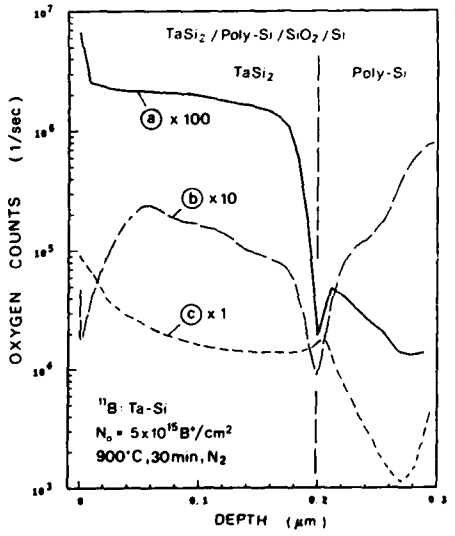


Fig. 4

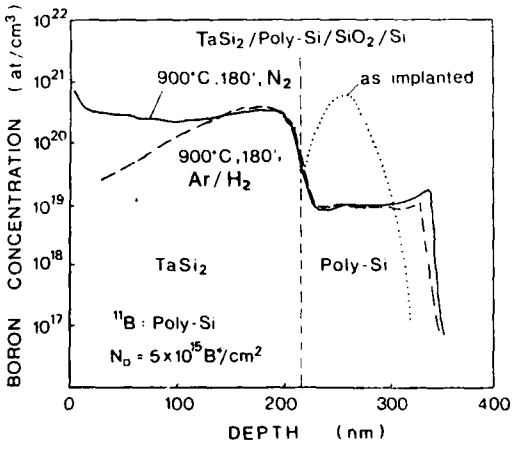


Fig. 3 Oxygen profiles for a) as-deposited $TaSi_2$, and after annealing at 900°C in N_2 for samples b) with poly-Si surface cleaning and c) with interfacial oxide.
Fig. 4 Boron profiles in case of boron implantation into poly-Si prior to $TaSi_2$ deposition: a) as-implanted and after annealing in b) N_2 and c) Ar/H_2.

though in either case B accumulates near the oxide/$TaSi_2$ interfaces. This suggests that oxygen does not significantly affect the B redistribution.

3.3 - ANNEALING AMBIENT

For samples annealed in Ar/H_2 dopant loss has been detected in spite of a 100nm oxide cap /3/. In order to investigate the dependence of the B redistribution on the annealing ambient (Ar/H_2 or N_2) independently of implantation effects in $TaSi_2$, B was implanted into undoped poly-Si prior to $TaSi_2$ deposition. The same poly-Si surface cleaning was used as described before.

Typically for CMOS processes, after the gate fabrication step the remaining 900°C processing time is 180-250 min. Fig. 4 shows the resulting B profiles after annealing in Ar/H_2 or in N_2 at 900°C for 180 min. After annealing most of the implanted B diffuses into the $TaSi_2$ layer. If annealed in N_2 a high and uniform B concentratin ($\approx 2x10^{20} cm^{-3}$) is achieved. For Ar/H_2 anneals a considerable dopant loss occurs in spite of the 100nm oxide cap. This is due to the increase of the B diffusivity in oxide by several orders of magnitude in case of H_2 containing annealing ambients /9/. Such a B loss to the oxide cap is only possible if a sufficient B supply is assured by the $TaSi_2$ layer. This suggests that the grain boundary diffusionis the dominant mechanism for the B redistribution even for long annealing times. Other mechanisms such as Ta-B compound formation do not seem to significantly affect the B redistribution. Nevertheless this does not rule out that B in $TaSi_2$ partly exists as TaB_2.

A rough estimate gives that approximately half of the implanted dose is lost if annealed in Ar/H_2. Although for B dose $\geq 5*10^{15} cm^{-2}$ the polycide p^+-ϕ_{ox} remains unchanged (B concentration in poly-Si $\approx 1x10^{19} cm^{-3}$). But, on the other hand, the B contamination of the furnace might be of concern. The poly-Si doping does not depend on the B dose for doses $\geq 5x10^{14} cm^{-2}$ and for low B doses the resulting B redistributin is quite independent of the implantation scheme /6/. If low doses are used, however, subsequent processing must avoid H_2 containing anneals in order to assure p^+-ϕ_{ox} reproducibility.

4 - CONCLUSION

In summary we haver shown that a native oxide constitutes an effective diffusion barrier for B. Therefore the cleanliness of the $TaSi_2$/poly-Si interface is a crucial factor for reproducible p^+-$TaSi_2$ polycide gates. The content and distribution of oxygen in the polycide does not affect the B diffusion. Annealing at 900°C in H_2 containing ambients leads to dopant loss in spite of a thick oxide cap. In case of long annealing times and/or of low B doses the dopant loss would reduce the p^+doping of the p^+ polycide, i.e. of the underlying poly-Si layer, deteriorating the MOS properties.

REFERENCES

/1/ U. Schwalke, C. Mazuré & F. Neppl, MRS Meeting on Poly-Si and its Interfaces; Dec. 1-4, 1987, Boston (USA).
/2/ N. Lifshitz, F.A. Baiocchi & D. Malm, Ext. Abst., Vol. 87-1, 243, ECS 1987.
/3/ H. Gierisch, F. Neppl, E. Frenzel, P. Eichinger and K. Hieber, J. Vac. Sci. Tech. B5, 508 (1987)
/4/ V. Probst, H. Kabza & H. Goebel, presented in this conference.
/5/ H. Gierisch, F. Neppl, unpublished.
/6/ U. Schwalke, C. Mazuré & F. Neppl, to be published in J. Vac. Sci. Tech. B.
/7/ P. Eichinger, E. Frenzel & F. Neppl, MRS Sympl. Proc., Vol. 25, 165 (1984).
/8/ F.C. Shone, K.C. Saraswat & J.D. Plummer, IEDM 1985, Vol. 15.3, p. 407.
/9/ C.Y. Wong & F.S. Lai, Appl. Phys. Lett. 48, 1658 (1986).

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

EFFECT OF DEPOSITION TEMPERATURE ON PLASMA GROWN ALUMINUM OXIDE FILMS

C. BOURREAU and Y. CATHERINE

Laboratoire des Plasmas et des Couches Minces, CNRS-UA 838, Institut de Physique et Chimie des Matériaux de Nantes, 2, rue de la Houssinière, F-44072 Nantes Cedex 03, France

Résumé.— Les propriétés physiques et électriques de couches minces d'oxyde d'aluminium déposées par plasma (13,56 MHz) à partir de triméthylaluminium sur du silicium sont fortement dépendantes de la température à laquelle est effectuée cette croissance. On montre que les couches se densifient lorsque la température de dépôt passe de l'ambiante à 300°C.

La caractérisation par la méthode C(V) des structures MOS fait apparaître une dispersion des résultats. Toutefois, un recuit dans Ar + O_2 à 350°C et pendant une heure apporte une amélioration sensible des caractéristiques de ces films (diminution de la tension de bande plate et de l'hystérésis).

Abstract.— The physical and electrical properties of aluminum oxide films deposited from aluminumtrimethyl under plasma conditions have been studied as a function of the silicon substrate temperature. It is shown that an increase of the temperature enhances the oxidation reaction and gives dense films.

The C(V) characterization of MOS structures shows a large scattering in the results. However higher temperature (up to 300°C) gives lower flat band voltage, lower hysteresis which indicates a lowering of the free charges in the oxide.

1 - INTRODUCTION

During the past years there has been a great interest in low temperature techniques for the deposition of insulator films for the purpose of insulation, charge storage, diffusion inhibition and processing of optoelectronic devices /1/. Silicon oxide is now frequently used in electronic devices but silicon nitride and aluminum oxide present several advantages compared with silicon oxide : a higher dielectric constant, a good radiation resistivity and are good barrier against the diffusion of impurities. These dielectric characteristics are necessary if one wants to obtain electronic components with stable and reproducible behaviour. The plasma method in other respects permits the deposition of dielectric films at low temperature (< 400°C) and enables the use of temperature sensitive substrate as III - V compounds.

The present work deals with the influence of the deposition temperature on plasma deposited aluminum oxide films using aluminumtrimethyl as aluminum carrier gas.

2 - EXPERIMENTAL METHODS

The figure 1 is a schematic diagram of the glow discharge apparatus used in this study.

Aluminum oxide films have been deposited by decomposition of aluminum trimethyl diluted in helium (~ 1% MA) in a radiofrequency discharge (13.56 MHz) with carbon dioxide (CO_2) or nitrous oxide (N_2O) as oxidizing agent. The influence of several parameters such as RF power, pressure, has already been studied /2/. Typical deposition pressures of 13.3 Pa (0.1 torr) are used and deposition powers are in the range 10 - 30 W. The deposition temperature can be choosen between room temperature and 450°C. More experimental details can be found in reference 2.

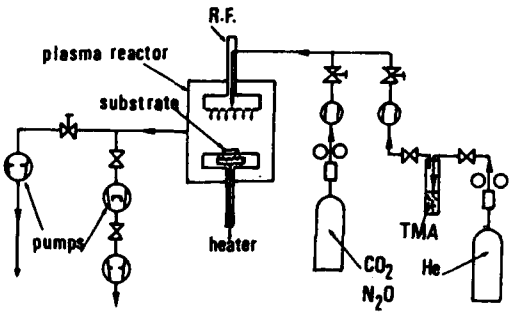


Fig. 1 - Schematic diagram of the glow discharge set up for AlO_x film deposition.

3 - INFLUENCE OF THE DEPOSITION TEMPERATURE UPON GROWTH KINETICS

As shown in figure 2 an Arrhenius plot of the deposition rate as a function of the temperature results in a decrease of the growth rate with the temperature rise. This plot also gives a low negative apparent activation energy of - 0.4 kcal/mole (- 0.017 eV). Such a low apparent activation energy is representative of a deposition rate controlled by radical flux /3/.

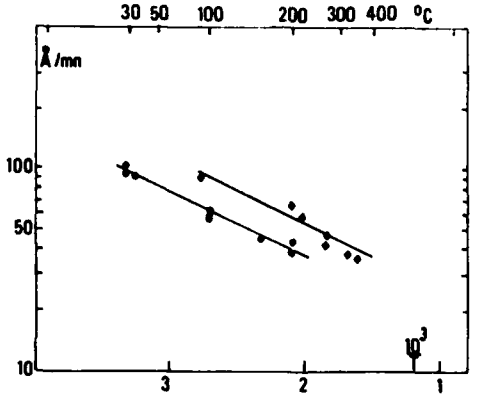


Fig. 2 - Arrhenius plot of the dependence of the growth rate on the substrate temperature : Pressure 13.3 Pa, flow rate : 25 sccm 60% CO_2 50 % CO_2.

The negative temperature coefficient can be interpreted by the competition of two processes : aluminum incorporation via adsorption of aluminum trimethyl radicals and the oxidation of aluminum which inhibits deposition sites. Thus the increase of the oxidation reaction rate with the temperature involves the decrease of the film growth rate by aluminum incorporation. Similar results have been observed with N_2O.

4 - EVOLUTION OF PHYSICAL PROPERTIES AS FONCTION OF THE DEPOSITION TEMPERATURE

The evolution of the refractive index and etch rate (room temperature, 1% HF) are reported in fig. 3.

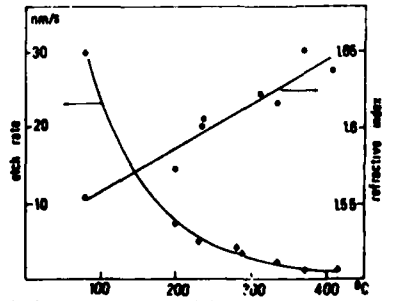


Fig. 3 - Etch rate and refractive index versus deposition temperature : Pressure 13.3 Pa, flow rate : 25 sccm, 60% CO_2.

It clearly shows the densification of the films with the higher deposition temperature. This densification is characterized by the increase of the refractive index n from 1.55 up to 1.65. This variation of n corresponds to an increase of the film density from 1.8 up to 3.2 g/cm³. These density values are always lower than sapphire density values : from 3.4 up to 3.7 g/cm³. ESCA analysis confirms this enhancement of the oxidation of Aluminum with the deposition temperature (Table 1).

Table 1 - Atom deposition of AlO_x layers measured by ESCA.

DEPOSITION TEMPERATURE		ATOM % WITHOUT CLEANING			ATOM % AFTER Ar ION CLEANING		
T°C		[O]	[Al]	[C]	[O]	[Al]	[C]
room	(50 % CO_2)	47,4	27,6	25	60,6	30	9,4
300°C	(50 % CO_2)	53,9	33	13,1	56,3	38,7	5
300°C	(60 % CO_2)	44,5	32,7	22,9	55,3	44,7	–

Most of the carbon of the high temperature films is due to pollution of the layers during storage while low temperature films exhibit Al – C bonds as was shown by IR spectroscopy /2/.

5 - ELECTRICAL PROPERTIES

5.1. As deposited films

C(V) curves of film deposited on n type silicon have been obtained by using a mercury probe in order to constitute a MOS structure. The results present a large dispersion and no significant evolution of the flat band voltage can be observed with the deposition temperature variation. A typical C(V) curve for as grown films is given in fig. 4 a and is dominated by a large hysteresis, that is the C(V) curve is dominated by large flat band voltage shifts as for sputerred Al_2O_3 films /4/ or CVD films from $Al(CH_3)_3$ /5/. Typical values for effective charge levels Q_T are given in Table 2 and compared with the results of other authors.

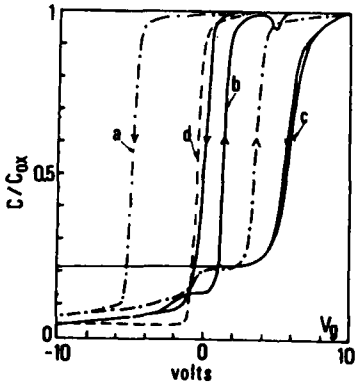


Fig. 4 - C(V) curves of Al_2O_3 MOS structure.
a : as deposited film, **b**: after one hour 350°C annealing in Ar/O_2 (1/1), **c**: after two hours 350°C annealing in Ar/O_2(1/1), **d**: theoretical curve.

Table 2 - Effective charge level for as grown Al_2O_3 films deposited by various methods.

Deposited Method	C V D $Al(CH_3)_3$	C V D $AlCl_3$	Plasma oxidation	This work
Charge density Q_T (cm^{-2})	$-1.5,-5\,10^{12}$	$-3\,10^{11}$	$8\,10^{11}$	(CO_2) $4\,10^{11}, 2\,10^{12}$ N_2O $-4\,10^{11}, 6\,10^{11}$
Reference	5	6	7	

5.2. Effects of annealing processes

We have studied the effect of a one hour annealing (Ar/O_2 (1/1) 350°C) on the electrical and physical characteristics of these layers. After one hour annealing a noticeable densification occurs : thickness varies from 2034 Å down to 1410 Å while refractive index grows from 1.55 up to 1.59.

The evolution of the C(V) curve shown in figure (**4 b**) gives a decrease of the hysteresis and of the flat band voltage which are characteristics of a lowering of the effective and free charge levels.

After a second annealing of one hour (Δt = 2 h) there is a slight variation of thickness 1410 Å down to 1380 Å but a deterioration of the film structure substantiated by the lowering of the refractive index from 1.59 down to 1.49 and of the oxide capacitance from 277 pF down to 162 pF. This effect was also observed in oxide films obtained by plasma oxidation /7/.

6 - CONCLUSION

Aluminum oxide layers obtained using plasma deposition from $Al(CH_3)_3$ at temperature of 300° and more present good physical and structural properties, however the electrical characterization shows the importance of the insulator, semiconductor interface and the need of a better control of trapped and free charges. This will be the goal of our future work.

REFERENCES

/1/ K.P. PANDE, V.K.R. NAIR and D. GUTIERREZ, *J. Appl. Phys.*, **54** *(1983), 5436.*

/2/ Y. CATHERINE and A. TALEBIAN, *J. of Electron. Mat.*, **17** *(1988), 127.*

/3/ D.J. ERLICH and R.M. OSGOOD, *Chem. Phys. Lett.* **79** *(1981), 381.*

/4/ R.S. NOWICKI, *J. Vac. Sci. Technol.*, **14** *(1977), 127.*

/5/ M.T. DUFFY, A.G. REVESZ, *J. Electrochem. Soc.* **117** *(1970), 372.*

/6/ K.M. SCHLESIER, J.M. SHAW and C.W. BENYON, *R.C.A. Rev.* **37** *(1976), 360.*

/7/ K.J. BURLOW, W. ECCLESTON and S. TAYLOR, *Appl. Surf. Sci.* **30** *(1987), 47.*

*

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

SUPERFICIAL-ENHANCED THERMAL NITRIDATION OF SiO_2 THIN FILMS

A. GLACHANT, B. BALLAND*, A. RONDA(1), J.C. BUREAU* and C. PLOSSU*

CRMC²-Département de Physique, Faculté des Sciences de Luminy, Case 901, F-13288 Marseille Cedex 9, France
**Laboratoire de Physique de la Matière (CNRS-UA 358), Bât. 502, INSA de Lyon, 20, Av. A. Einstein, F-69621 Villeurbanne Cedex, France*

Résumé - La nitruration superficielle de films minces (∿ 13 nm) de SiO_2 peut être stimulée en pratiquant des recuits répétitifs entre 900°C et 1200°C dans une faible pression ($\leq 10^{-1}$ mbar) d'ammoniac de pureté contrôlée. La zone superficielle nitrurée s'étend alors sur 3 à 4 nm avec une décroissance rapide de la concentration d'azote en fonction de la profondeur. Les résultats expérimentaux suggèrent que les espèces amines NH_x ($0 < x < 3$) sont essentielles dans le processus de nitruration de SiO_2 par NH_3. Des structures Al/SiO_2 nitrurée/Si(100)p ont été fabriquées in situ ou non. En optimisant les conditions de nitruration, les mesures électriques révèlent une faible densité d'états d'interface (10^{10} à 10^{11} eV^{-1} cm^{-2}), une bonne tenue à l'injection d'électrons caractérisée par un équilibre entre piégeage et dépiégeage dans la région nitrurée et un champ destructif parfois supérieur à 10 MV cm^{-1}.

Abstract - Superficial nitridation of thin (∿ 13 nm) SiO_2 films can be enhanced using annealing cycles of short duration at high temperatures (900°C-1200°C) in low pure ammonia pressures ($\leq 10^{-1}$ mbar). The nitrided surface region is 3 to 4 nm wide with a rapidly decreasing nitrogen concentration versus depth. The experimental data suggest that amine NH_x ($0 < x < 3$) species are essential in the nitridation process of SiO_2 using NH_3 gas. Al-gate/nitrided SiO_2/Si(100)p capacitors have been prepared *in situ or* not. Electrical measurements show that a low interface-state density (10^{10} to 10^{11} eV^{-1} cm^{-2}), a good stability during electron injection due to a balance between trapping and detrapping in the nitrided region of SiO_2 and a destructive breakdown field as high as 10 MV cm^{-1} can be achieved by optimizing nitridation conditions.

1 - INTRODUCTION

Superficial nitridation of thin SiO_2 films grown on Si would modify the insulator structure in such a way to make it impervious to contaminant diffusion and less susceptible to degradation during VLSI processes and high-field stresses, without damaging the electrical characteristics of the SiO_2/Si interface. The nitridation of SiO_2 thin films is generally accomplished in high ammonia pressures ($P \geq 1$ bar) and at high temperatures ($900°C < T < 1200°C$) /1/ using excessively long furnace times or plasma-enhanced processes. Rapid thermal nitridation at high temperatures /2,3/ has been also proposed to reduce nitrogen interfacial concentration, increase surface nitridation and overcome the difficulty of dopant redistribution. However, these various processes suffer from the inability to control the spatial nitrogen distribution in the oxide film and consequently the defect density in the oxide bulk and at the SiO_2/Si

(1) Present address : LISE, Facultés Universitaires N.D. de la Paix, 61 rue de Bruxelles, B-5000 Namur, Belgium.

interface. We show here that superficial nitridation of very thin SiO_2 films (thickness of $\simeq$13 nm) thermally grown on p-Si(100) can be achieved in low ammonia pressures ($P \leqslant 10^{-1}$mbar) by thermal activation (950°C $\leqslant T \leqslant$ 1150°C) using annealing cycles of short duration (sequential mode). The purpose of our study is to understand the changes in electrical properties correlated with the composition of the insulating films during the nitridation process. Moreover, this study should lead to a better understanding of the nitridation mechanisms of thin thermal SiO_2 films and then to a limitation of nitrogen incorporation in the surface region.

2 - EXPERIMENTAL METHODS

The nitrogen content of the thermal oxides as a function of nitridation parameters (T, P, annealing duration Δt, total nitridation time t) has been investigated using several analytical techniques. Nitrogen depth profiles in the nitrided films were measured by AES in conjunction with argon ion-sputtering and SIMS. The elemental composition and chemical bonding were studied by AES, XPS, Raman (R S) and infra-red (IRS) spectroscopies. Some electrical measurements by use of I(V), $C(V,\omega)$-$G(V,\omega)$ and DLTS techniques, time-dependent breakdown and trapping-detrapping measurements have been performed on Al gate MIS devices prepared in situ or not with thermal-nitrided oxide gate insulators to relate the electrical properties to the elemental composition.

3 - RESULTS AND DISCUSSION

Enhancement of surface reaction was observed for sufficiently large nitridation times ($4\ h \leqslant t \leqslant 10\ h$), $T \geqslant$ 950°C and $P \geqslant 10^{-3}$ mbar /4,6/ as shown in figure 1. Moreover, the reaction rate during the first nitridation stages increased for the shortest annealing periods ($\Delta t \leqslant$ 10 min). In these conditions, a temperature rise above $\sim$ 950°C did not affect markedly the reaction rate. The experimental data suggest that amine NH_x ($0 < x < 3$) species are essential in the nitridation process (Fig. 2) : these species have to diffuse through the surface layer to react with SiO_2. The thermal reaction proceeds via the replacement of oxygen atoms following probably successive reactions predicted in the literature : the final products expected are eventually nitride and/or higher order amine groups with the loss of oxygen in the form of water. The nitrogen pile-up at the SiO_2/Si interface was hardly detectable and the reaction was essentially localized in the surface region of SiO_2. In our case, surface reaction enhancement can be attributed to various complementary factors, including a dry ammonia ambient, a low concentration gradient of nitriding species and a rapid decrease of the diffusion coefficient of NH_x groups when increasing the nitrogen concentration.

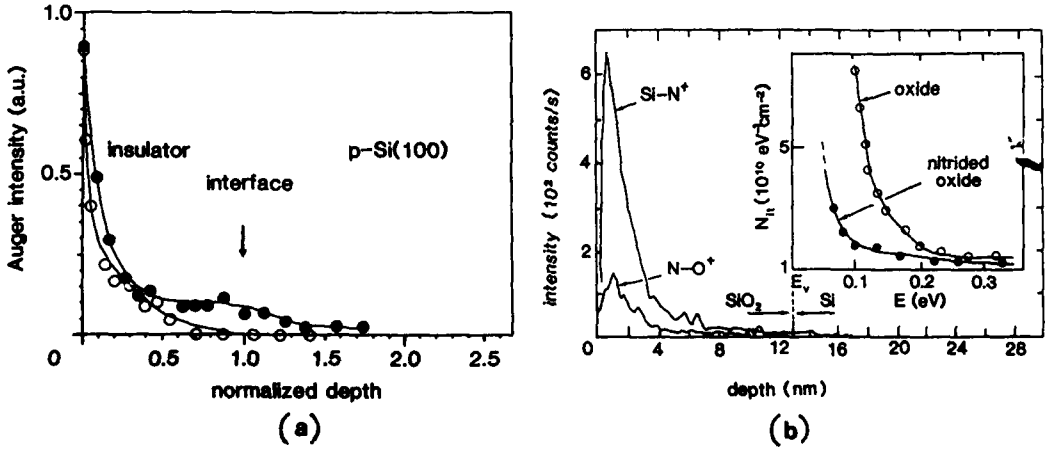


Fig.1 - (a) N Auger depth profiles from a nitrided SiO_2 film ($P = 10^{-1}$mbar - T = 1024°C - t = 10h) using different sputtering rates (1 nm.min^{-1} : ● and 0.2 nm.min^{-1} : o) ; (b) SIMS profiles of Si-N and N-O from a SiO_2 film reacted at T = 1045°C ($P = 10^{-3}$ mbar) during t = 4h and the corresponding interfacial-state density vs. energy. The nitrogen concentration is quite negligible in the interfacial region.

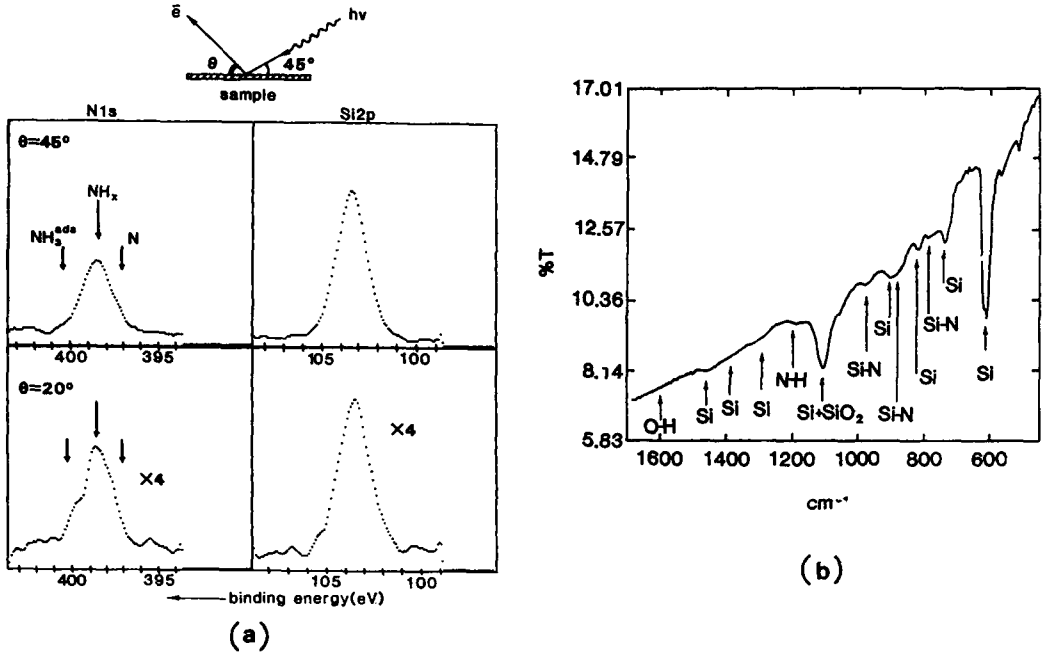


Fig.2 - (a) N1s and Si2p photoemission lines (T = 1024°C - P = 10^{-3} mbar - t = 5h) for two vertical sampled depths (about 2 nm (θ = 20°) and 4 nm (θ = 45°) with respect to the N1s line) ; (b) Infra-red spectrum of the same reacted sample. Nitride species are mostly localized near the insulator surface.

The effect of nitridation on the flat-band voltage shift ΔV_{FB} was analysed by comparison of simulated ideal and actual C-V curves. ΔV_{FB} was in the range -1 to 0.8 V depending on nitridation conditions and was related to the nitrogen distribution within the nitroxide. The small positive charge density created (negative ΔV_{FB}) was generated rapidly during the initial nitridation process and is correlated with trapping by Si-OH or Si-H groups resulting from Si-O bond breaking and/or Si-dangling bonds generated by in-diffusion of H species (as shown by SIMS, XPS and IRS analysis). After a long nitridation time (t > 90 min), the C-V characteristic was shifted in positive bias direction due to either a longer heat cycle or an increase of nitrogen concentration. The amphoteric and deep ≡Si centers produce a density of states peak near midgap (electro-optical measurements : UV photo-ionisation) ; they will thus show two levels : a lower +/0(≡Si°) trapped hole level and a higher 0/-(≡Si⁻) trapped electron level. The interface densities are evaluated from C (V,ω) and G(V,ω) curves or from DLTS. Before nitridation, the midgap surface-state density at the SiO_2/Si interface is 10^{10} $eV^{-1}.cm^{-2}$; a sufficiently low interface-state density (1 to 10.10^{10} $eV^{-1}.cm^{-2}$) can be achieved by optimizing nitridation conditions (Fig.1b). As observed from conduction and I-V characteristics, at high electric field, the Fowler-Nordheim (FN) tunneling remains the dominant mechanism in nitrided oxide, similar to the oxide (Fig.3). Electron injection from metal gate or from the substrate into insulating layers was accomplished by FN tunneling : the nitrided oxide having the highest nitridation rates did not show catastrophic dielectric breakdown (before reaching saturation) but rather exhibited a balance between injected electron trapping and tunnel detrapping (Fig.4). The final destructive breakdown field was measured by applying a negative bias voltage to the Al electroce at a definite sweep rate : it varied from 2 $MV.cm^{-1}$ to > 10 $MV.cm^{-1}$ depending on the nitridation parameters.

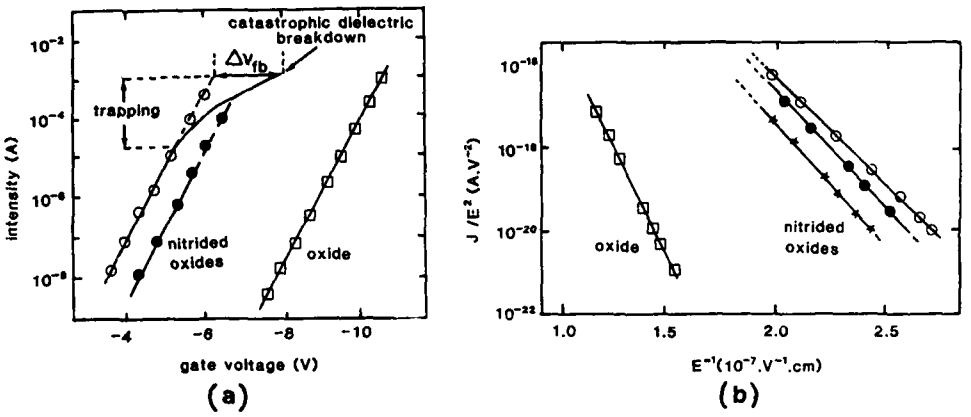


Fig.3 - (a) I(V) characteristics for different MIS structures (o,● : nitrided oxides, □:oxide). (b) J/E^2 vs. 1/E for different MIS structures (J = current density; E = electric field across the film) : the FN tunneling is the dominant conduction mechanism.

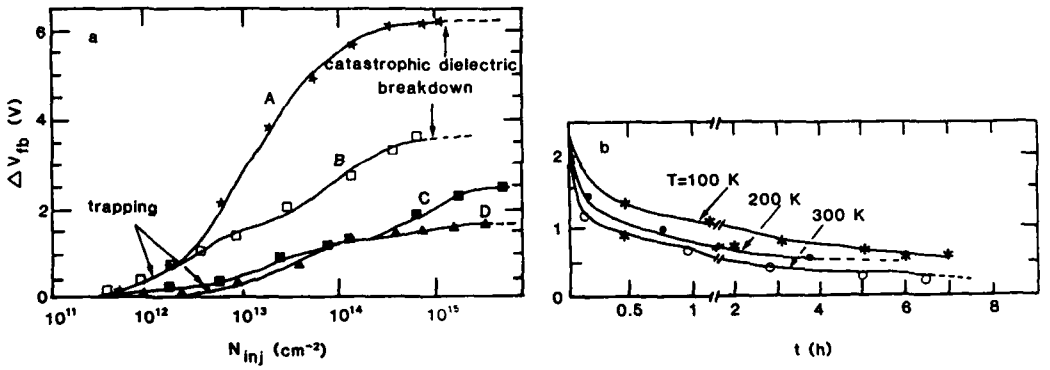


Fig.4 - (a) Flat-band voltage shift vs. injected electron number (curves A,B,C,D). (b) Flat band voltage relaxation (curves C,D). (A): nitrogen to oxygen ratio r = 0.5 (1045°C, 10^{-4}mbar, 7.5h), (B): r = 0.45 (1045°C, 5.10^{-4}mbar, 4h), (C) and (D): r = 2.3 (1045°C, 10^{-3}mbar, 4h).

4. CONCLUSION

Superficial-enhanced thermal nitridation of SiO_2 thin films has been carried out by thermal activation in low pure ammonia pressures. Amine NH_x species have to diffuse through the surface layer to react with SiO_2. The electrical properties measured with Al gate MIS structures show that nitrided oxide films will be good quality gate insulators for future scaled-down VLSI devices.

REFERENCES

/1/ Kato,I., Ito,T., Inoue,S.I., Nakamura,T. and Ishikawa,H., Jpn. J. Appl. Phys. 21 (1982) 153.
/2/ Moslehi,M.M. and Saraswat,K.C., IEEE Trans. Electron Devices ED2 (1985) 106.
/3/ Moslehi,M.M., Saraswat,C.K. and Shatas,S.C., Appl. Phys. Lett. 47 (1985) 1113.
/4/ Ronda,A., Glachant,A., Plossu,C. and Balland,B., Appl. Phys. Lett. 50 (1987) 171.
/5/ Balland,B., Dupuy,J.C., Glachant,A. and Ronda,A., Phys. Stat. Sol. A100 (1987) 187.
/6/ Benamar,A., Bureau,J.C., Balland,B., Ronda,A. and Glachant,A., Appl. Surf. Sci. 30 (1987) 216.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

PROCESS DEPENDENCE OF HOLE TRAPPING IN NITRIDED SiO_2 FILMS

M. SEVERI, M. IMPRONTA, L. DORI and S. GUERRI

CNR, Istituto LAMEL, Via de' Castagnoli n.1, I-40126 Bologna, Italy

Résumé - On a effectué une analyse sistématique du piégeage des trous dans des couches minces (20-30 nm) d'oxide nitruré avec la technique de l'injection à avalanche, en fonction des conditions de nitruration. La nitruration realisé à températures relativement basses (700-800°C) et pour des temps courts produit une augmentation des piéges des trous. La reduction du piégeage des trous peut être obtenue seulement avec des conditions de nitruration plus sévères. On a étudié aussi l'effet du récuit en atmosphère d'oxygène après nitruration.

Abstract - A systematic investigation of hole trapping in 20-30 nm nitrided oxides as a function of the nitridation conditions were performed using the avalanche injection technique. Nitridation carried out at relatively low temperatures (700-800°C) and for short times brings about an increase of hole traps. Hole trapping reduction can only be achieved for more severe nitridation conditions. The effect of postnitridation annealing in oxygen was also studied.

1 - INTRODUCTION

Thermally nitrided SiO_2 films are being extensively investigated in order to improve several properties of the insulator in metal-oxide-semiconductor structures for VLSI. In particular, it has been inferred from high field stress that hole traps in nitrided oxides are reduced to very low levels /1/. Very recently /2/, this has been confirmed by direct measurements using hole avalanche injection. By using the same technique, in this work we have performed a systematic investigation of hole trapping in 20-30 nm nitrided oxides as a function of the nitridation conditions and postnitridation annealing treatments.

2 - EXPERIMENTAL

n-type (100)-oriented silicon substrates with a resistivity of 0.1 ohm.cm were used. 20-30 nm gate oxide films were thermally grown in dry oxygen at 1000°C in quartz-tube furnace and in-situ annealed at the same temperature in N_2 for 10 min. Thermal nitridation was carried out in ultrapure ammonia gas at temperatures between 700 and 1100°C and for times ranging from 5 to 60 min in a cold-wall rf-heated reactor at atmospheric pressure. Some of the nitrided oxides were subsequently annealed in oxygen at 900-1000°C. Aluminum and gold were deposited by evaporation. Electrodes with areas of 2.46×10^{-3} cm^2 were formed by photolithography or using a metal mask. In some Al-gate samples, a post metallization anneal (PMA) in nitrogen at 450°C for 20 min was performed. Avalanche injection was achieved with a 150-kHz sawtooth signal. The average injected current density was kept constant at a value of 1.2×10^{-7} A/cm^2. All the measurements were done at room temperature. The oxide charge buildup was monitored by periodically interrupting the injection and performing C-V measurements.

3 - RESULTS AND DISCUSSION

Let us first discuss some experimental results obtained on standard SiO_2 films. Fig.1 shows the negative flatband voltage shift upon hole injection for Al and Au gate capacitors with 30 nm SiO_2. Two features are worth noticing: a) Au and Al gate capacitors before PMA show very similar hole trapping behavior. This result rules out that electron injection from the gate gives a significant contribution to the voltage shift. b) PMA treatment reduces hole trapping in Al gate samples. This result is in agreement with some indications already reported in the literature without discussion /2-4/. On the other hand, hole trapping in Au gate samples does not practically change by performing an anneal in forming gas before metallization. It is our opinion that the effect of PMA is related to the modification of the stress at the SiO_2/Si interface brought about by the PMA itself. Upon cooling to room temperature, thermal mismatch leads to a large tensile stress in the Al film /5/ which counteract the compressive stress in the SiO_2. In this way, the density of the strained Si-O-Si bonds near the Si/SiO_2 interface

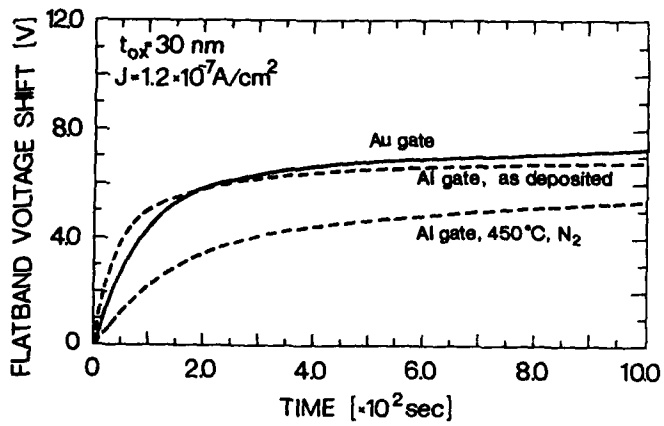


Fig.1 - Negative flatband voltage shift by avalanche injection of holes in 30 nm SiO_2 with as deposited Al (dashed curve) and Au (solid curve) gate. For Al, data are also shown after PMA.

might be reduced. Strained bonds near the interface have been suggested as one of the defects which can give rise to hole traps /3/. The possibility that mechanical stress in Al gate might reduce the strained bonds was first proposed by Chin and Ma /6/ in discussing the generation of interface states by ionizing radiations.
Now we turn to discuss the trapping properties of the nitrided oxides with Au gate electrodes. Similar results were obtained in Al gate samples. Fig.2 shows the effect of the nitridation temperature from 700°C to 1100°C on the hole trapping in 30 nm SiO_2 films.

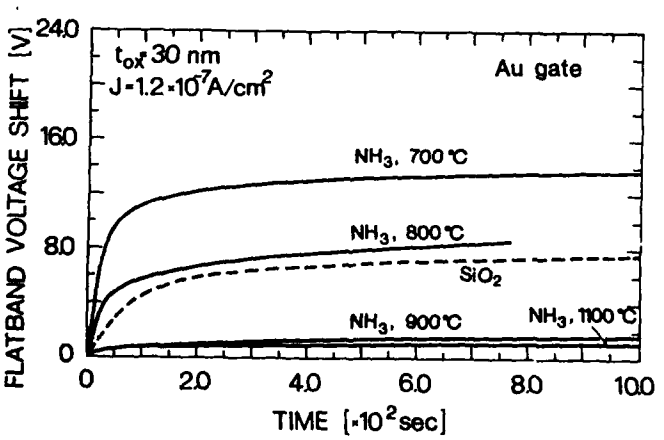


Fig.2 - Effect of nitridation temperature on the hole trapping in 30 nm SiO_2 with Au gate. Nitridation time was 30 min. Dot area = 2.46×10^{-3} cm^2.

As can be seen, nitridation at 700 and 800°C increases hole trapping with respect to the standard oxide, while at higher temperatures a remarkable decrease is obtained. We believe that the effect on hole trapping is strictly related to the nitridation kinetics /7,8/. For low nitridation temperatures (700-800°C) and/or for short times, there is a pile up of nitrogen at the Si/SiO_2 interface. In these conditions, hole traps increase. However, when the nitridation process is carried out at higher temperatures and/or for longer times, the nitrogen peak moves away from the interface and an oxygen-rich layer grows. The formation of this layer has been attributed to the oxygen released by the exchange reaction in the bulk of the oxide /8/. This oxygen may help in reducing the oxygen deficiency near the Si/insulator interface which has been suggested to be one of the cause for hole traps /9/.
To test this hypothesis we have performed some experiments by varying the nitridation time at

a fixed temperature (900°C). The results are reported in Fig.3. For a short nitridation time, hole trapping steeply increases, in correlation with the rapid increase of the nitrogen concentration at the Si/dielectric interface. With increasing nitridation time, oxygen diffuses to the interface and the hole traps decrease. It is interesting to note that saturation has been observed for a 60 min nitridation.

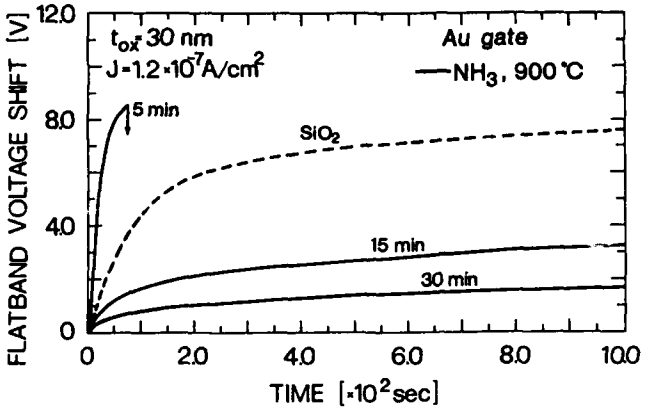


Fig.3 - Effect of nitridation time on the hole trapping in 30 nm SiO_2 with Au gate. Nitridation temperature was 900°C. The arrow indicates breakdown.

A decrease in the strain in the interfacial region with the increase of nitrogen, as already suggested /2/, may give a further contribution to the reduction of the hole trap density.
The reduction of the hole traps brings about a reduction of the interface states generated upon avalanche injection of holes, as can be seen in Fig.4, which shows quasi-static C-V curves after 100 sec of injection for pure oxide and for oxides nitrided at 900°C for 30 min. These results confirm those reported in /2/.

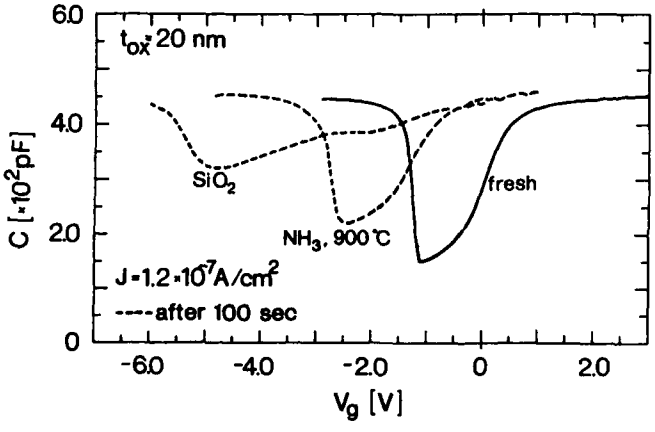


Fig.4 - Quasi-static C-V curves for 20 nm SiO_2 and for nitrided oxide after 100 sec of hole injection with a current density of 1.2×10^{-7} A/cm^2.

It is interesting to observe that similarities and differences exist between hole and electron trapping in nitrided oxides. Contrary to hole traps, electron traps increase with the nitridation temperature /10/; however, an interfacial phenomenon, like the "anomalous positive charge" (APC) generation during electron avalanche injection, shows the same nitridation temperature and time dependence as hole trapping, suggesting a possible correlation between trapped holes and APC.
Finally, we would like to discuss some preliminary results on the effect of postnitridation heat treatments. The dependence of hole trapping on high temperature annealings in oxygen is

illustrated in Fig.5 for oxides nitrided at 900°C. It can be seen that hole trapping increases dramatically for the annealing at 950°C for 15 min, while it gradually decreases by performing more severe oxidizing treatments. Eventually, for a 60 min annealing at 1000°C a hole trapping lower than that in as-nitrided oxides is obtained. The increase in the hole trapping is probably due to the extension of the strained-bond region towards the oxide bulk through a viscous flow process during the anneal at high temperature ($\geq$ 950°C) /3,4/. Only when oxygen succeeds in penetrating the SiON barrier and reaches the interface, the hole traps can be reduced.

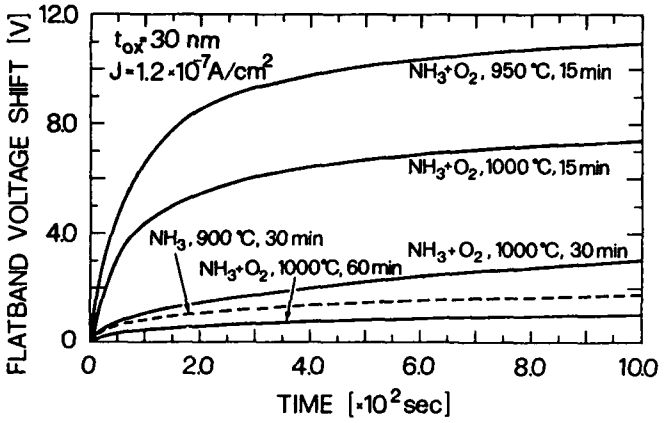


Fig.5 - Effect of postnitridation heat treatments in oxygen on the hole trapping in 30 nm SiO_2 nitrided at 900°C for 30 min.

4 - CONCLUSIONS

We have found that hole trapping in nitrided oxides depends on the nitridation conditions and on the postnitridation heat treatments. A reduction of hole trapping can only be achieved for relatively severe nitridation conditions (temperature $\geq$ 900°C and time $\geq$ 15 min). This reduction has been correlated with the formation of an oxygen-rich layer near the Si/insulator interface. A postnitridation O_2 anneal can increase or decrease the hole traps: its effect depends critically on the tempefature and time of the process.

ACKNOWLEDGEMENTS

The authors would like to thank P.Negrini, P.Castelli and G.Pizzochero for their assistance in sample preparation and M.Rancoita for avalanche measurements.
This work has been supported by the Progetto Finalizzato "Materiali e Dipositivi per l'Elettronica allo Stato Solido" of the Consiglio Nazionale delle Ricerche, Roma.

REFERENCES

/1/ Lai, S.K., Lee, J. and Dham, V.K., IEDM 83, p.190.
/2/ Yankowa, A., Do Thanh, L. and Balk, P., Solid State Electron. 30 (1987) 939.
/3/ Lai, S.K., J.Appl.Phys. 54, (1983) 2540.
/4/ Aitken, J.M. and Young, D.R., IEEE Trans. Nucl. Sci. 24 (1977) 2128.
/5/ Sinha, A.K. and Sheng, T.T., Thin Solid Films 48 (1978) 117.
/6/ Chin, M.R. and Ma, T.P., Appl.Phys.Lett. 42 (1983) 883.
/7/ Vasquez, R.P., Madhukar, A., Grunthaner, F.J. and Naiman, M.L., Appl.Phys.Lett. 46 (1985) 361 and J.Appl.Phys. 59 (1986) 972.
/8/ Han, C.J., Moslehi, M.M., Helms, C.R. and Saraswat, K.C., Appl.Phys.Lett. 46 (1985) 641.
/9/ Aslam, M. and Balk, P., in Proc. INFOS'83, p.103.
/10/ Severi, M. and Impronta, M., Appl.Phys.Lett. 51 (1987) 1702.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CMOS TECHNOLOGY USING PLASMA NITRIDED OXIDE AS A GATE DIELECTRIC

A. STRABONI, M. BERENGUER, B. VUILLERMOZ, P. DEBENEST, A. VERNA and P. DARS

Centre National d'Etudes des Télécommunications, CNET Grenoble, Chemin du Vieux Chêne, BP 98, F-38243 Meylan Cedex, France

Résumé - Une nouvelle technique de nitruration plasma d'oxyde a été développée pour les isolants de grille très minces. Les propriétés d'interface sont conservées après nitruration en plasma d'ammoniac à 950°C. Un excellent comportement des transistors à grille nitrurée lors d'expériences de vieillissement a été relevé ; l'absence de défauts dus au plasma illustrée par les rendements de mémoires SRAM indique la compatibilité du procédé avec la production.

Abstract - A new technique of plasma nitriding oxide has been developed for very thin gate insulators. The interfacial properties are preserved after plasma nitridation in ammonia at 950°C. Excellent behaviour is observed during aging experiments on transistors using nitrided oxide as the dielectric gate ; the absence of defects due to the plasma as illustrated by the yield achieved for SRAM memories suggests the process compatibility with production.

1. INTRODUCTION

In MOS-IC technology, interest devoted to very thin films of nitrided silicon dioxide is mainly due to their barrier properties (1). Furthermore, it has been reported that nitrided dielectrics show less wear-out and radiation sensitivity than silicon dioxide (2). Whatever the nitridation process, thermal or plasma, the key-point is to obtain surface nitridation without degrading the initial SiO_2 / Si electrical performance. In this paper, we report electrical characterization of MIS structures and MOS test transistors of SRAM devices using dielectrics prepared by direct plasma nitridation of dry thermal SiO_2.

2. EXPERIMENTAL PROCEDURES

The ammonia plasma treatments were performed in the Unit of Reactive Anodization for Nitridation and Oxidation of Semiconductors (U.R.A.N.O.S.) which is a multiwafer set-up compatible with IC production (3). The 13 MHz R.F. plasma is created through external electrodes located away from the reaction region so that no energetic particles could attain the wafers. Thermal oxides (from 10 to 25 nm) prepared at 950°C in dry O_2 were nitrided from one to six hours at 950°C using NH_3 at 3. 10^{-2} mbar (hPa) with an incident power of 600 W. Some of the nitrided samples were annealed in low pressure dry O_2 at the nitridation temperature. Capacitors electrodes were defined using photolithography. Device testing was performed after a postmetallization anneal in N_2/H_2 at 450°C for 30 min. High frequency and quasi-static C(V) measurements on structures prepared with aluminum and polysilicon as the gate material were used to determine the fixed charge Nf and interface trap densities Dit. The breakdown voltage of nitrided SiO_2 films were investigated by applying a fast ramp voltage (25 V/s). The breakdown voltage (Vbd) was defined as the gate voltage at which a current density of 1 x 10^{-4} A cm^{-2} flows through the oxide. Each histogram is the result of 85 measurements on the same wafer. Aging experiments based on transconductance degradation measurements were performed on NMOS transistors using 25 nm as grown and nitrided oxides as the gate insulator. Finally, SRAM devices including the plasma nitridation gate process were tested and their yields were compared to control wafers using the same 2 µ design rules CMOS technology.

3. RESULTS

Typical high frequency and quasistatic C(V) curves of a capacitor with 15 nm thick oxide after 3 hours nitridation on P-type Silicon using N^+ Polysilicon electrode is shown in Fig. 1. From the flat band shift as compared to the control oxide, an increase in Nf (dNf) up to 2. 10^{11} cm^{-2} is observed while the midgap Dit remains unaffected (2. 10^{10} $cm^{-2}.eV^{-1}$). The

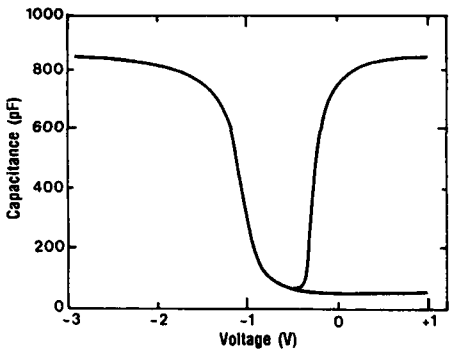


Figure I. high frequency and quasistatic C(v) curves for a 3 hours nitrided 15 nm dielectric using a N^+ Polysilicon electrode on P - Si with S = 0.38 mm2.

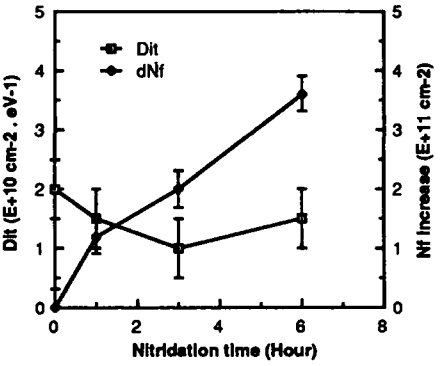


Figure II. Fixed charge and interface trap creation during the nitridation of 15 nm oxides.

kinetics of interface trap and fixed charge build-up are plotted in Fig. 2 as a fonction of the nitridation time. It indicates the continuous incorporation of positive charges with the time, an effect which has also been reported for pure thermal nitridation. But in contrast with thermal processes, interface trap levels are scaled down after plasma nitridation. These results were confirmed by precisely measuring the Nit distribution using DLTS for 10 nm nitrided oxide on N and P substrates. The data on P type silicon summarized in table I show that the average midgap Dit is reduced after 3 and 6 hours below the initial oxide

TABLE I. Dit ($eV^{-1}.cm^{-2}$)determined by D.L.T.S. for 10 nm oxides on P type Silicon.

Oxide as grown	1 Hour	3 Hours	Nitrided 6 Hours
$1\text{-}2\ 10^{10}$	$2\ 10^{10}$	$5\ 10^{9}$	$2\text{-}3\ 10^{9}$

TABLE II. Effect of O_2 post - nitridation anneal on the VFB distribution (264 capacitors/wafer) ‹ VFB › = mean value, σ(VFB)=std dev.

15 nm	Oxide as grown	2 Hours Nitrided	2H Nitrided + 0_2.anneal
‹VFB›	-0.66	-0.83	-0.73
σ(VFB)	.02	.10	.06
dNfx10^{11}		2.4	1.0

level. This low defect density can be explained by the low nitrogen incorporation in the bulk and at the interface already reported for plasma nitrided oxides (1). A low pressure oxygen anneal performed just after a 2 hours nitridation under the same pressure and temperature conditions was tested for a 15 nm oxide. As shown in table II, the fixed charges were mostly eliminated. Furthermore, the Vfb distribution over the wafer was also improved after this O2 anneal. Statistical measurements of the breakdown voltage Vbd in Fig. III reveal that nitridation does not induce defect in the dielectrics and that the breakdown

characteristics of the initial oxide (avg. Vbd = 15.5 V, σ = 0.3 V) are maintained or ligthly improved after nitridation (avg. Vbd = 15.7 V, σ = 0.2 V).

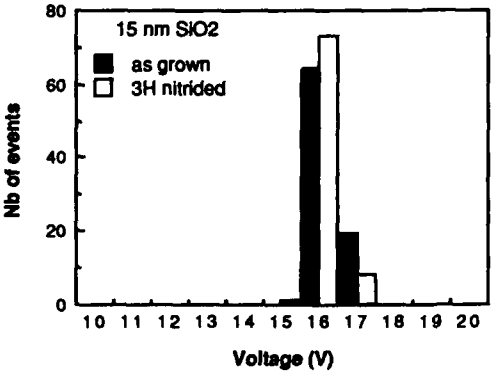


Figure III. Histogams of Vbd for a 3H nitrided as compared to the control oxide. bias (-)on the metal.

Figure IV shows the evolution of the flat band voltage Vfb for the original oxide as compared to the nitrided oxide structures during a Fowler-Nordheim injection from the gate into the dielectrics. The curves show that the nitrided oxide presents less modification during the experiment than

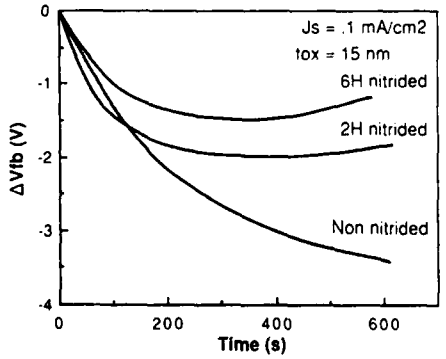


Figure IV. Change in flat band voltage during Fowler - Nordheim injection.

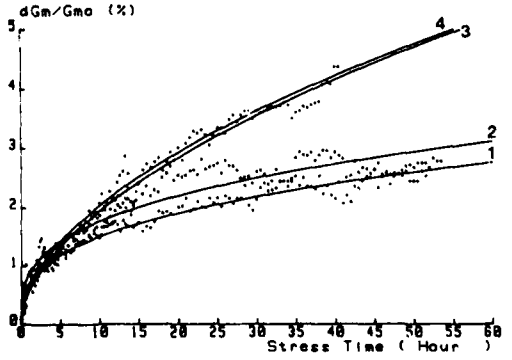


Figure V. Comparison of transconductance degradation (dGm/dGm_0 %) between nitrided (1 and 2) and non nitrided (3 and 4) transistors.

the thermal oxide. Similar results were obtained for the Dit measurements. Stress measurements were performed on NMOS transistors using 25 nm nitrided and standard oxides. Transconductance tests are reported on Fig. V for devices with channel lengths between 1 and 1.3 micron with the conditions indicated in table III. In spite of the channel length sligthly shorter for the nitrided dielectrics, the transconductance degradation is reduced with respect to the non nitrided oxide.

TABLE III. Voltage aging conditions, effective channel length and transconductance degradation (dGm %) for nitrided (1, 2) and non nitrided transistors (3, 4).

Transistors	1	2	3	4
Leff	1.14	1.12	1.31	1.31
VG	3.5	3.0	3.5	3.5
VD	7.0	7.0	7.0	7.0
dGm % (for 100H)	3.7	3.3	6.8	6.9

Based on the time dependent law $A.t^n$ (3), an extrapolation of the transconductance degradation (dGm %) for a 100 hours stress gives the results indicated in table III. This seems to indicate that aging reduction after plasma nitridation is related to the low interface trap density.
Plasma nitridation of the gate insulator was applied on 2 microns design rule CMOS technology including 60 SRAM (6 Transistors/point) memories and test transistors without any attempt to adjust the VT shifts due to the fixed charges and without post nitridation anneal. A strong correlation has been noticed between higher fixed charge and lower circuit yields regions over the wafer. The yield reduction (average 15% with 31% for the best nitrided wafer versus 60% avg. for the non nitrided set of wafers) is only attributed to the VT shift as no other defect was detected in the other test devices .

4. CONCLUSION

Our study shows that the plasma nitridation of SiO2 films used as dielectric gate insulators does not affect the electrical characteristics of the MIS structures. The reduction in interface trap density seems to be a characteristic of plasma nitridation which is also associated with lower trapping and better aging behaviour for N transistors. The threshold voltage of the transistors is shifted due to the presence of fixed charges which can be substantially reduced through a subsequent O2 anneal. Absence of other defects demonstrates the possible utilization of plasma nitridation processing in assembly lines.

ACKNOWLEDGMENTS

The authors wish to acknowledge A. Chantre for the DLTS experiments and G. Gimine and H. Mingam for useful discussions.

REFERENCES

(1) P. Debenest, K. Barla, A. Straboni and B. Vuillermoz. European M.R.S. Meeting , Strasbourg (May 31 - June 2) 1988.
(2) F. Terry and al., IEEE Elecron Dev. Lett. Vol EDL4, p 191, 1983.
(3) B. Vuillermoz, A. Straboni and R. Pantel, Electrochem. Soc. Meet.,Toronto, May 1985, Abstr. 248.
(4) E. Tadeka, N. Suzuki, IEEE Electron Dev. Letters, Vol EDL5,p50, 1984

Session 4C Room C : Process technology of compound s.c.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

DIFFUSION BARRIER LAYERS FOR OHMIC CONTACTS TO GaAs

D.A. ALLAN, J. HERNIMAN, M.J. GILBERT, P.J. O'SULLIVAN, M.P. GRIMSHAW* and A.E. STATON-BEVAN*

British Telecom Research Laboratories, Martlesham Heath, Ipswich, Great-Britain
**Dept. of Materials, Imperial College, GB-London SW7 2AZ, Great-Britain*

Résumé
Dans les circuits intégrés à transistors métal-semiconducteurs de AsGa, on s'est servi de barrières de diffusion par-dessus les contacts ohmiques alliés Ni-Au-Ge et "non alliés" Pd-Ge. Ceci empêche la diffusion des couches d'interconnexion suivantes et assure la stabilité et la basse resistance (0.05 ohm.mm) des contacts. L'évaluation de ces structures de contact du point de vue électrique ainsi que par AES et par TEM est présentée ici. Les résultats montrent que l'effet de barrière produit des contacts uniformes et faibles et un bon rendement de circuits numériques rapides à intégration à échelle moyenne (MSI).

Abstract
Barrier layers have been used on top of Ni-Au-Ge alloyed and Pd-Ge "non-alloyed" ohmic contacts for GaAs MESFET integrated circuits. They prevent diffusion of subsequent interconnection layers and so maintain stable low resistance (0.05 ohm.mm) contacts. The assessment of these contact structures, both electrically and using AES and TEM, is presented. The results show that the barrier action is effective in producing uniform and reliable contacts and good yields of MSI high speed digital circuits.

1 - INTRODUCTION

Ohmic contacts to GaAs using the Ni-Au-Ge alloyed system have been popular for many years and have been successfully used in MESFET integrated circuits for low resistance source and drain contacts. The contact morphology and reproducibility of these contacts can be rather variable and to improve these aspects we previously used an evaporated silicon dioxide dielectric encapsulating layer during the alloying process /1/. This prevented surface tension effects degrading the contact edge definition and also encouraged more even alloying of the contact. This approach had the disadvantage that the SiO_2 had to be removed after alloying, leaving the contact open to subsequent degradation during processing.
A more attractive option is to use a conductive barrier layer which can be left in place and, on top of which, an interconnect layer can be deposited in the same deposition cycle as the ohmic contact layers. The barrier has to be capable of remaining intact during ohmic contact formation, not allowing interdiffusion of the top interconnect metal. This maintains the stoichiometry of the underlying layers to preserve the integrity of the ohmic contact. The barrier layer also has to have good adhesion properties and must be stable during subsequent device processing. A further important requirement of the contact is good surface and edge morphology for subsequent photolithographic alignment of sub-micron features. We have investigated a large number of possible barrier materials but only three, zirconium diboride (ZrB_2), tungsten silicide (W_5Si_3) and tungsten nitride (WN) fulfilled these conditions.
The fabrication of Ni-Au-Ge alloyed and also Pd-Ge non-alloyed ohmic contacts is described in the following section. Pd-Ge contacts have the advantage that they are formed by solid phase epitaxy to give smoother interfaces and do not contain Au, a potential reliability hazard. The electrical characteristics of the contacts, their uniformity and stability and the resulting MESFET device results are summarised. The barrier integrity, on both large (100μm) and small (0.1μm) scales, was investigated by Auger Electron Spectroscopy (AES) depth profiling and Transmission Electron Microscopy (TEM) cross-sectional analyses respectively.

2 - EXPERIMENTAL

Semi-insulating GaAs substrates, ion-implanted with Si to give a peak doping level of $1*10^{18} cm^{-3}$, were used to give comparable conditions to the FET source and drain regions. The alloyed ohmic contact layers of Ni (5nm), Au (45nm), Ge (20nm) or the non-alloyed Pd (50nm) and Ge (100nm) layers were deposited by electron beam evaporation at a pressure < $1*10^{-6}$mbar. ZrB_2 was also evaporated, rather than sputtered as in initial reports for silicon technology /2/, so that it could be incorporated in the same deposition sequence. A thin Au layer was also deposited on top of the barrier layer to give a low resistivity surface for ease of measurements and subsequent layer contacting. Thus the total ohmic contact structure can be deposited in a single evaporation sequence and defined by conventional photolithographic lift-off techniques. The W_5Si_3 and WN were sputtered in order to retain a constant stoichiometry. Lift-off of these materials was still possible using a specially developed two level resist technology.
Heat treatment of the ohmic contacts was carried out on a carbon strip heater in flowing forming gas or in an optical annealing system in flowing argon /3/. The contact resistance was measured using the transmission line model (TLM) /4/ by extrapolation of resistance for zero contact pad separation and normalised to a 1mm wide contact and thus expressed in ohm.mm. Low contact resistivities (ohm.cm^2) are difficult to calculate accurately since the models assume specific current paths and doping values. The contact resistance is also more useful and easily applied to the parastic resistance of a MESFET.

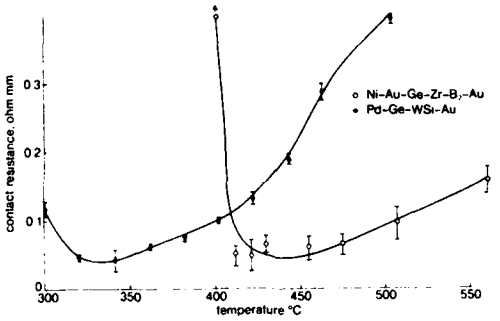


Fig 1 Contact resistance vs temperature of heat treatment for ohmic contact systems.

Figure 1 shows the contact resistance obtained as a function of heat treatment for both Ni-Au-Ge and Pd-Ge contacts with different diffusion barrier layers. The minima in contact resistance are in the same temperature range as Au-Ge-Ni /1/ and Pd-Ge /5,6/ without barrier layers, indicating that the barrier layers are effective and play only a secondary role in the contact formation. The minimum values of 0.06 ohm.mm at 440°C for Ni-Au-Ge and 0.04 ohm.mm at 330°C for Pd-Ge are amongst the lowest values ever reported. Particular care was taken to ensure that lithographic dimensional errors were accounted for since a difference of 0.4 µm in pad separation is equivalent to 0.03 ohm.mm which is significant. Each point represents the mean of 16 devices in the case of Ni-Au-Ge and 6 in the case of Pd-Ge and the error bars indicate the total spread in measurements. A similar uniformity is exhibited over a whole 50mm diameter wafer with the contact structure Ni-Au-Ge-ZrB_2-Au which has a mean of 0.07 ohm.mm and a standard deviation of 0.02 ohm.mm. A histogram showing similar results with a WN barrier layer is illustrated in figure 2.

AES sputter etch profiling was used to check the integrity of the barrier and in all cases abrupt interfaces between the layers were observed after ohmic contact formation. A typical depth profile for a Pd-Ge-WSi-Au stucture is shown in figure 3. The AES technique uses an analysis area up to 100µm in diameter, much greater than the average device area, and so only the average structure is revealed. In order to ensure that the barrier was effective and did not show small areas of breakdown which would be significant on a device (0.1 µm) scale, cross-sectional TEM (XTEM) samples were prepared using Ar^+ ion beam thinning and examined in a Jeol JEM 2000FX microscope with EDX analysis.

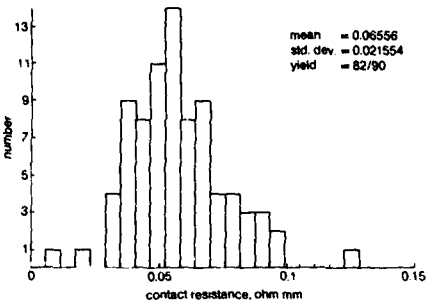


Fig 2 Histogram of contact resistance measurements across a wafer for the Ni-Au-Ge-WN-Au system.

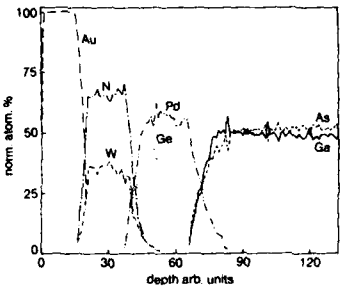


Fig 3 AES depth profile for the Pd-Ge-WSi-Au showing the separation of the PdGe from the Au by the barrier layer.

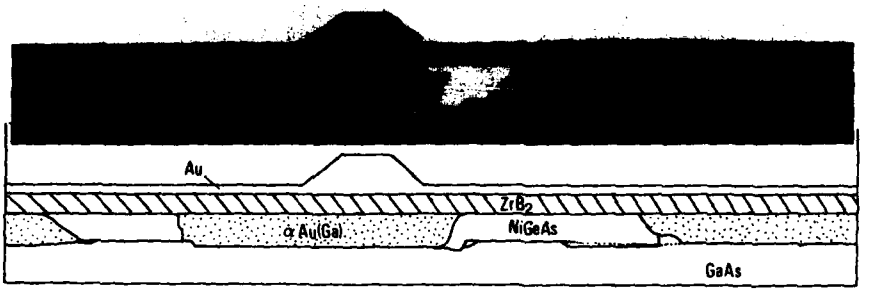


Fig 4 TEM cross-section of an optically annealed Ni-Au-Ge-ZrB_2-Au contact. The schematic diagram shows the phases present.

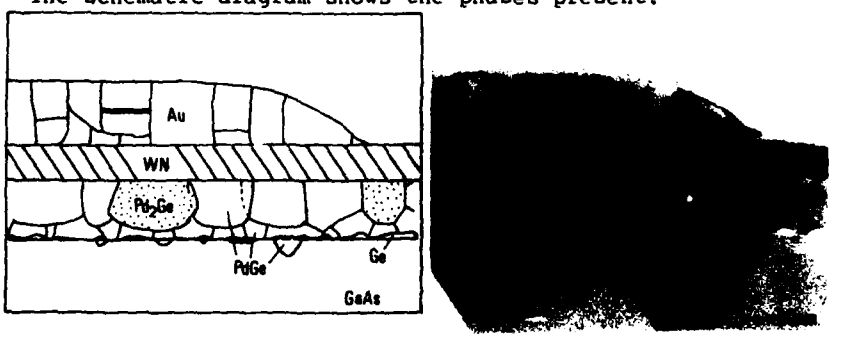


Fig 5 - TEM cross-section of an optically annealed Ge-Pd-WN-Au contact. The schematic diagram shows the phases present.

The XTEM study of the Ni-Au-Ge-ZrB_2-Au system, fig 4, showed the amorphous ZrB_2 layer to be an effective barrier. Occasionally, however, Zr from the barrier was observed to have undergone a limited planar reaction with adjacent Au rich phases forming Au_4Zr. The phases which were present in both the optically annealed and carbon strip annealed contacts included; orthorhombic NiGe, hexagonal NiGeAs, cubic epitaxial Ge, cubic α Au-Ga and hexagonal α' Au-Ga. The carbon strip heated samples contained additional, Ga-rich Au-Ga phases, hexagonal β + β' Au-Ga and orthorhombic Au_2Ga. These phases probably resulted from the longer anneal time of 60s (*cf.* 5s for optical annealing). The diffusion depth of the metallisation into the GaAs was 0-15 nm and 45 nm for the optically and carbon strip annealed samples respectively. These values compare favourably with the 70-110 nm values quoted for Ni-Au-Ge contacts without diffusion barriers /8/.
Fig 5 shows a typical XTEM of the Ge-Pd-WN-Au system. The fine polycrystalline WN barrier layer was found to be intact and unreacted. The top Au layer (f.c.c.) contained horizontal twin boundaries. The lower contact structure consisted of an upper layer of large PdGe and Pd_2Ge grains and a lower layer of smaller PdGe grains. A thin (<10nm) layer of epitaxial Ge exists at the GaAs interface through which some of the PdGe grains protruded up to 30 nm into the GaAs. It is thought that the rough two-layer morphology is a product of the short annealing time and not an effect due to the barrier.
The barriers have also proved to be effective during subsequent device fabrication which involves further high temperature processes and also during deliberate temperature stressing at 300°C for periods up to 100 h. No changes in contact resistance were observed within the measurement accuracy for the Ni-Au-Ge system but the Pd-Ge resistance increased to 0.23 ohm.mm. This is not surprising considering the sintering temperature was only 330°C. The incorporation of ZrB_2 barrier layer ohmic contacts in MSI scale circuits has proved satisfactory on over 100 wafers. These had additional Ti-Pt-Au layers on top of the ohmic contact for interconnection between devices. There is no evidence of interdiffusion of these different layers in contrast to the report of interdiffusion when no barrier layer was included /7/.
MESFETs with the following parameters were regularly achieved;-

Small signal transconductance	(g_m)≈151 mS/mm
Large signal transconductance	(G_m)≈104 mS/mm
Saturation threshold voltage	(V_T)≈ -2.05 V
Source-drain saturation current	(I_{dss})≈213 mA/mm

The drift in the above parameters was within +/- 2% after temperature stressing at 300°C for periods up to 100 h /3/. These MESFETs were used in 8:1 multiplexer and demultiplexer circuits operating at speeds of 1.8 Gbit/s or greater and which had fully functional yields of 32% - the failures being due mainly to lithographic defects.

3 - ACKNOWLEDGEMENTS

The authors wish to thank Dr M Hill of BTRL for the AES analysis and the Director of Research at BTRL for permission to publish this work.

REFERENCES

/1/ Allan, D.A. and Thorp, S.C., Physica 129B (1985) 445.
/2/ Shappiro, J., Sol. State Technology October (1985) 161.
/3/ Herniman, J., Allan, D.A. and O'Sullivan, P.J., IEE Proc Part I Sol. State and Electron Devices 135 (1988) 67.
/4/ Harrison, H. and Reeves, G., IEEE Elec. Dev. Letts. EDL-3 (1982) 111.
/5/ Marshall, E.D, Chen,W.X., Wu, C.S. and Lau, S.S., Appl. Phys. Lett. 47 (1985) 298.
/6/ Kaminska, E., Piotrowska, A., Barcz, A., Turos, A. and Pietrak, R., Crys. Prop. and Prep. 12 (1987) 277.
/7/ Lee, C.P., Welch, B.M. and Fleming, W.P., Electron. Lett. 17 (1981) 407.
/8/ Shih,Y.C, Murakami,M., Wilkie, E.L. and Callegari,A.C., J. Appl. Phys. 62 (1987) 582.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CONTROL OF THE FABRICATION STEPS OF InP MIS TRANSISTORS BY MEANS OF SCANNING PHOTOLUMINESCENCE MEASUREMENTS

B. COMMERE, M. GARRIGUES, S.K. KRAWCZYK, C. LALLEMAND, K. SCHOHE and B. CANUT*

Laboratoire d'Electronique, CNRS-UA 848, Ecole Centrale de Lyon, 36, Av. Guy de Collongue, F-69131 Ecully Cedex, France
**Dept. de Physique des Matériaux, Université Claude Bernard Lyon I, 43, Bd du 11 Novembre, F-69622 Villeurbanne Cedex, France*

Résumé - Nous montrons l'efficacité de l'imagerie de photoluminescence à suivre la qualité, l'homogénéité et la reproductibilité d'un échantillon dans la réalisation des transistors MIS sur InP, du substrat de départ et après chaque étape technologique. Grâce à l'utilisation de cette technique nondestructive, rapide, et n'exigeant pas de contacts avec la surface de l'échantillon, nous avons obtenu une amélioration considérable des dispositifs élaborés.

Abstract - We show the efficiency of scanning photoluminescence measurements for monitoring the quality, homogeneity and reproducibility of the starting wafers and processed substrates after each technological step during the realization of InP MIS transistors. Owing to this new technique, which is noninvasive, contactless and fast, we have obtained a considerable improvement of the fabricated devices.

1 - INTRODUCTION

The caracteristics of III-V semiconductor devices and the fabrication yield are critically influenced by nonuniform and nonreproducible properties of starting and processed substrates [e.g.1] . The control of the entire technological process is usually achieved a posteriori by electrical characterization of the completed devices. However, such approach is not sufficient, since it allows only to detect "that something again didn't work in the process", and that most of the devices on the wafer do not work properly. Thus, a tight control of the quality and the uniformity of the wafers at each process step is necessary in order to improve the device performance, to increase the scale of integration and to lower the costs by increasing the fabrication yield. Scanning photoluminescence (PL) measurements emerge today [e.g. 2, 3] as a promissing technique to assess the quality, the homogeneity and the reproducibility of the processed substrates in a contactless, noninvasive and fast way, conform to industrial requirements for routine quality check.

Here we report on the realization of InP MIS transistors improved and controled by room temperature scanning PL measurements, which were performed at each individual step during the process.

2 - EXPERIMENTAL

The InP MISFETs processed in this work had a classical structure of n-channel devices [e.g. 4] and were fabricated on Fe doped <100> oriented semi-insulating InP wafers. The nominal channel widths and lengths were 100µm and 20µm, respectively. The source and drain regions were defined by photolithography and selectively implanted with ^{28}Si ions (150keV energy, $2 \cdot 10^{14}$ cm^{-2} dose, room temperature). Post implant annealings were typically performed at 750°C for 10 min. using the "wafer-to-wafer" capping technique in a phosphorus rich atmosphere, provided by heated InP powder. After the annealing, the channel areas were slightly etched in an HIO_3 solution followed by various kinds of surface treatments such as HF etching, anodic oxidation [5] and others. After the deposition of the gate insulator (Al_2O_3, electron beam evaporation) the samples were annealed for 30 min. at 300°C in O_2 and then for 120 min. at 350°C in H_2/N_2.

Scanning PL measurements were performed at each fabrication step using the SCAT imaging system, which was developed in our laboratory (Fig.1). This system operates with an He-Ne laser as excitation source. The laser beam is focussed on the sample surface by a miscroscope objective. The PL intensity is measured with a Si photodiode. The sample is moved on an X-Y stepping motor stage, which is controlled by a computer, allowing to choose the scanning area (up to 2inch wafers) and the resolution (between 1μm and several tens of μm).

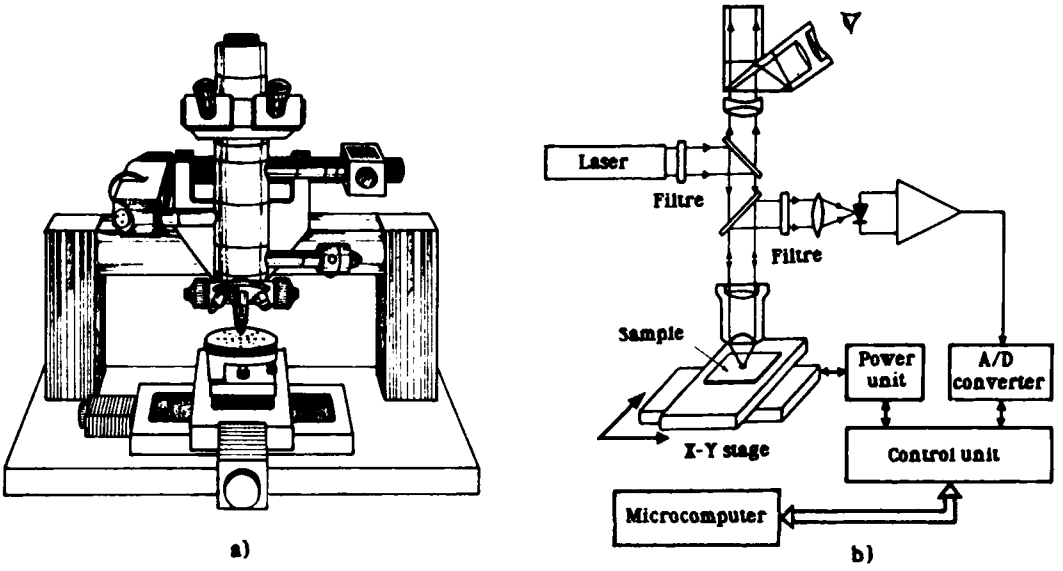


Fig. 1 The photoluminescence SCAT imager; a) general view, b) schematic of the system

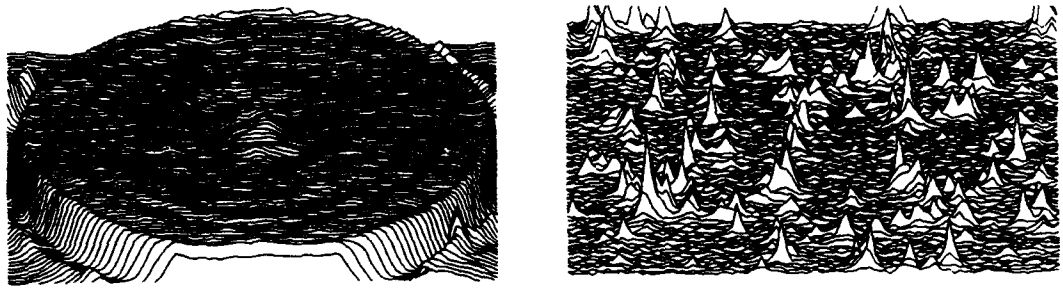
3.1 - STARTING SUBSTRATE QUALITY

We used scanning PL measurements to detect the presence of doping striations and dislocations in the bulk material and other extended defects in the InP surface region, which are probably left by the cutting and polishing of the substrate (Fig. 2). Destructive etching of the surface is normally necessary to reveal the defects and dislocations. In the case of semi-insulating InP substrates, dislocations locally result in a strong increase of the PL intensity , and thus, they can easily be detected and counted by means of scanning PL measurements. Although the impact of the striations, dislocations and other exteded defects on the performance of InP MISFETs is not yet definitively evaluated, we expect, that as in the case of the GaAs device technology, the choise of homogenous substrates will improve the homogeneity of the device characteristics on the wafer.

3.2 - IMPLANTATION AND POSTIMPLANT ANNEALING

The quality of implanted and annealed wafers depends on the degree of dopant activation, material recrystalization and on the thermally induced degradation of the semiconductor surface region. This degradation takes place on both implanted and not implanted areas.

We have found a monotonous correlation between the average PL intensity after implantation and subsequent annealing at various temperatures and the value of X_{min} deduced from aligned RBS spectra. These results indicate, that scanning PL measurements can be used to control the reproducibility of the annealing conditions and the uniformity of implanted and annealed wafers. A possible cause of long range nonuniformities is a nonhomogenous distribution of the temperature across the wafer during the post implant annealing.



a) b)

Fig. 2 Typical PI images obtained on a semi-insulating (Fe-doped) InP wafer; a) image of the entire 2inch wafer showing doping striation, b) image of a small area (2mm x 2mm) showing the presence of dislocations

We have identified two different types of thermally induced degradation. The first one, clearly observed in nonimplanted areas, results in a homogenously reduced average PL intensity after the annealing (Fig. 3). Scanning PL measurements, performed after successive etchings of the semiconductor (PL depth profiling) showed that this damage takes place within several hundred of Angstroms in the semiconductor bulk, and that its extension depends on the annealing conditions. This type of degradation was observed in all the samples. The second type of degradation results in a local increase of the PL intensity (apparition of "photoluminescence peaks" on the surface). This type of degradation was observed only on a few samples and could have been correlated with unefficient capping of the substrate during the annealing. Since this type of damage was easy to detect by scanning PL measurements, it was possible to elliminate bad wafers already at the very beginning of the process and to optimize the annealing conditions.

In some samples we have observed an anomalous and nonuniform increase of the PL intensity at the boundaries of the implanted areas. This effect was particularly pronounced in the channel region (Fig. 4) and was identified to be due to nonperpendicular walls of the implantation barrier. Optimization of the lift-off procedure was sufficient to improve drastically the definition of the implanted regions.

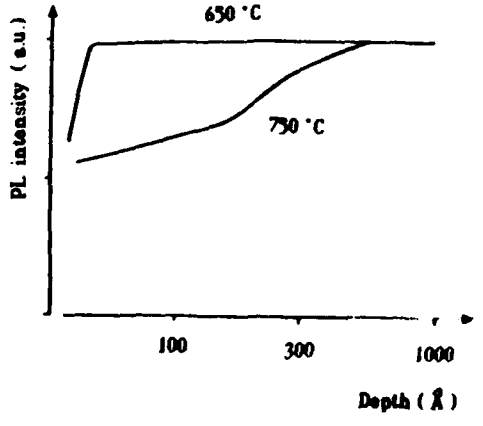


Fig. 3 Photoluminescence depth profiling obtained on not implanted area after post-implant annealing at 650° C and 750° C

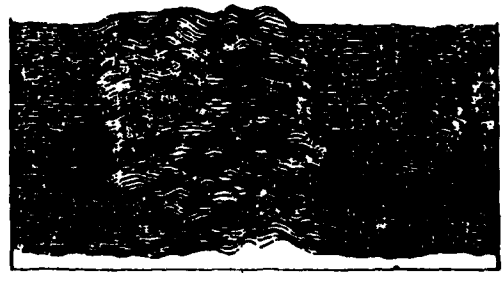


Fig. 4 Photoluminescence image of non-uniformities in the channel area of InP MIST observed after postimplant annealing. The scanned area is 50μm x 50μm

3.3 - ETCHING OF THE GATE AREA

Scanning PL depth profiling measurements allowed us to determine precisely the distance from the surface to which the substrate was degraded during the post-implant annealing. Removing of this damaged surface region prior to the deposition of the gate insulator dramatically improved the electrical properties of the transistors. In particular, we obtained an increase of the effective mobility of electrons in the channel from some 100 cm^2/Vs to more than 1000 cm^2/Vs. An individual transistor, characterized by PL imaging after implantation, annealing and local etching, is shown in Fig. 5. The higher PL intensity inside the etched areas of both implanted and nonimplanted regions proves, that the etching of the sample indeed improves the electronic properties of the semiconductor surface.

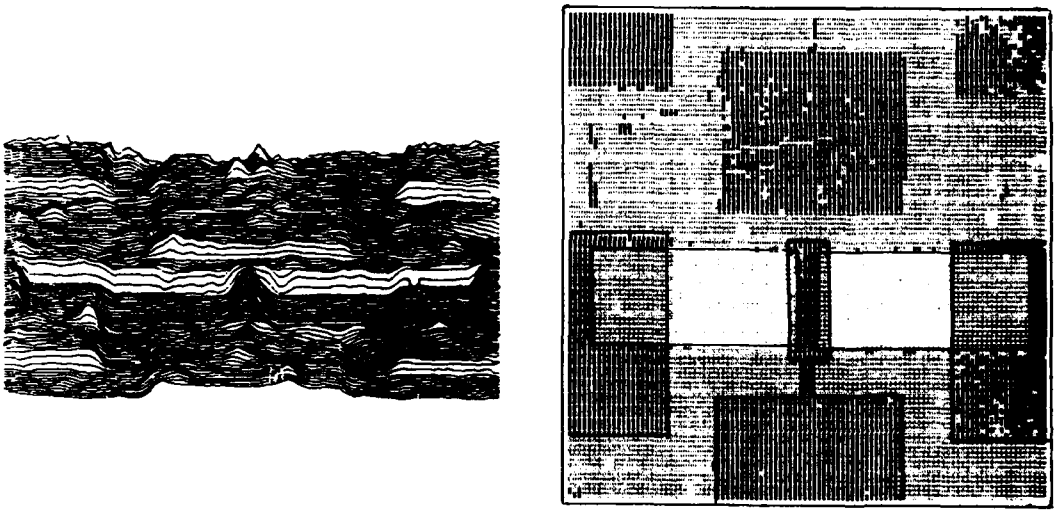


Fig. 5 PL image of an individual transistor in 3D and grey-level scale. This image was obtained after implantation, post-implant annealing and local etching. Dark pixels represent high PL intensity, bright pixels represent low PL intensity. The scanned area is 0.5mm x 0.5mm.

3.4 - CHEMICAL SURFACE TREATMENTS, GATE INSULATOR DEPOSITION AND POST-DEPOSITION ANNEALING

As we have shown in a previous work [6], the integrated PL intensity varies up to three orders of magnitude with chemical treatments of the InP crystal, which affect the composition of the native oxide on the surface. Also, we demonstrated in [7], that the PL intensity provides a reliable estimate of the surface quality in terms of the surface state density in the upper part of the gap. Thus, scanning PL measurements are useful to choose those treatments, which ensure low surface state densities. However, appropriate treatments should not only give good initial surface properties, but also should protect the crystal against the degradation induced during the subsequent technological steps, as in particular the insulator deposition. PL measurements, performed successively prior and after chemical surface treatments, after the insulator deposition and after the post-deposition annealing, allowed us to identify, to what extent each of the above steps contributes to the final quality of the interface and which interdependencies exist between these steps. Thus, scanning PL measurements offer the possibility to track the evolution of the surface/interface properties, providing crucial information for the process optimization.

Furthermore, we have found by means of scanning PL depth profiling measurements, that the insulator deposition not only results in the creation of interface states, but also may leed to the creation of defects and nonuniformities inside the semiconductor surface region. This effect depends on the deposition conditions, which can be optimized using scanning PL measurements as a guide. The power of scanning PL measurements for the control and the optimization of the InP/insulator interface properties will be discussed in detail elsewhere.

3.5 - CONTROL OF THE ENTIRE PROCESS

Fig. 6 shows the histogram of the PL intensity obtained on a small area of the wafer (covering a few individual transistors) after implantation, annealing and etching. The caracteristic distribution of the PL intensity is due to the presence of different zones in the scanned area (as indicated in the figure) and reflects the electronic quality of each zone. The comparison of such histograms obtained on different areas on the wafer provides information on the uniformity of the processed substrate at this stage of the process. The PL histograms obtained after each technological step can be considered as the "identity card" of the process, allowing the comparison of independent runs and the evaluation of the process reproducibility.
As a final example, Fig. 7 shows the evolution of the average PL intensity in the channel region for two different processes. In the first one the surface was simply treated with HF solution, in the second one an anodic oxide was formed on the surface prior to the gate insulator deposition. We have found, that if two independant runs of the same process type result in the same evolution of the PL intensity in the channel, the electrical characteristics of the devices are roughly the same. Furthermore, higher PL intensities in the channel are correlated with larger transconductancies of the transistors.

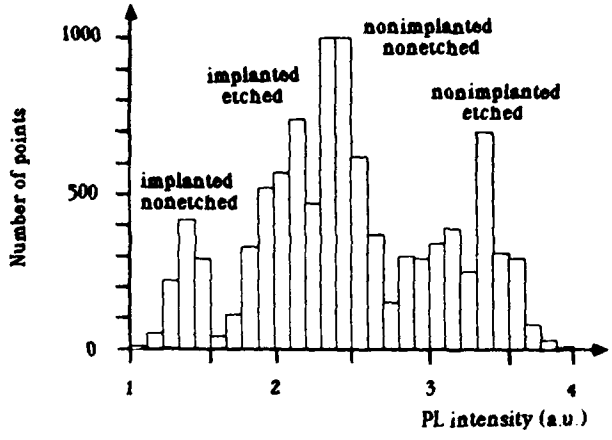


Fig. 6 Histogram of the PL intensities obtained on a small area covering a few individual transistors after implantation, annealing and etching. Photoluminescence from different zones are indicated on the figure.

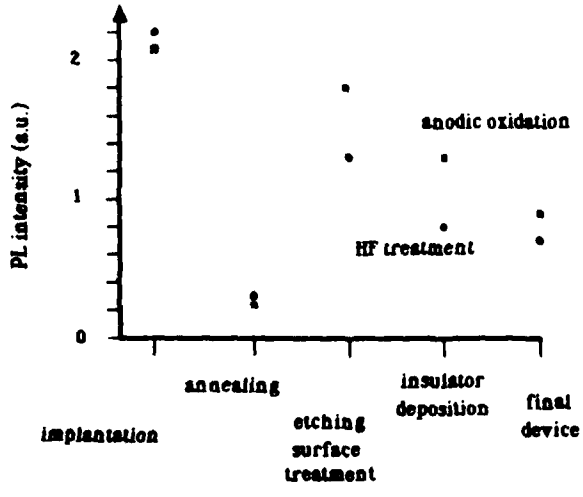


Fig. 7 Evolution of the average PL intensity in the channel area during two different processes: with HF etching and with anodic oxidation prior to the gate insulator deposition.

4 - CONCLUSION

Processing of InP MIS transistors was monitored and improved owing to scanning PL measurements performed after each technological step. Our results indicate, that the scanning PL measurements are well adapted, i) to extract pieces of information necessary for the optimization of individual steps in the process, and ii) to control the uniformity of starting and processed wafers as well as the quality and the reproducibility of each fabrication step. This new assessment technique is perfectly well compatible with the III-V technology.

ACKNOWLEDGEMENTS

The financial support to this work of the CNET (French Telecommunications, contract #86.1B.156.007.90.92.45.PAB) and of the Commission of the European Communities in the frame of STIMULATION (ST2*0265) is gratefully acknowledged.

REFERENCES

[1] MIYAZAWA, S., ISHII, Y., ISHIDA, S. and NANISHI, Y., Appl. Phys. Lett. 43 (1983) 853
[2] HOVEL, H. J. and GUIDOTTI, D., IEEE Trans. Electron. Dev. ED-32 (1985) 2331
[3] KRAWCZYK, S., GARRIGUES, M. and BOUREDOUCEN, H., J. Appl. Phys. 60 (1986) 392
[4] PANDE, K. P. and GUTIERREZ, D., Solid State Electronics 28 (1985) 1045
[5] HOLLINGER, G., BERGIGNAT, E., JOSEPH, J., and ROBACH, Y., J. Vac. Tech. A3 (1985) 2082
[6] KRAWCZYK, S. and HOLLINGER, G., Appl. Phys. Lett. 45 (1984) 870
[7] KRAWCZYK, S., BAILLY, B., SAUTREUIL, B., BLANCHET, R. and VIKTOROVITCH, P., Electronic. Lett. 20 (1984) 235

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

Mg^+/O^+ ION IMPLANTATION IN GaAs/GaAlAs HETEROSTRUCTURES

B. DESCOUTS, N. DUHAMEL and Y. GAO

Centre National d'Etudes des Télécommunications, Laboratoire de Bagneux, 196, Av. Henri Ravera, F-92220 Bagneux, France

Résumé - Des implantations de Mg^+ et O^+ ont été réalisées dans des hétérostructures GaAs/GaAlAs pour des applications de transistors bipolaires, le but de ces implantations étant de rendre la couche de collecteur fortement résistive (avec l'O) et de contacter la couche de base (avec le Mg). L'activation du Mg a été réalisée par des recuits rapides jusqu'à 900°C. Nous montrons que les conditions permettant d'obtenir une compensation par l'oxygène sont très dépendantes du matériau. De plus, pour des faibles doses d'oxygène nous avons noté une évolution dans le temps de la compensation. Enfin, nous détaillons l'intéraction entre Mg et Be (dopant de la base).

Abstract - Mg^+ and O^+ ion implantations have been performed in GaAs/GaAlAs heterostructures for bipolar transistor applications in order to form an isolation layer in the collector region (with O) and to contact the base layer (with Mg). Rapid thermal annealing with peak temperature up to 900°C has been employed to activate Mg. We show that the conditions to obtain compensation by oxygen are strongly dependant on the starting material. Moreover for low oxygen doses we have noted an evolution of the compensation with time. Finally we discuss the interaction between Mg and Be (dopant of the base).

I - INTRODUCTION

GaAs/GaAlAs heterojunction bipolar transistors (HBT's) are of great interest for application to high speed integrated circuits, but the performances of these devices are limited by extrinsic parasitic elements such as the external base capacitance. Several solutions to this problem have been reported in the litterature (1-4). Particularly, P.M. Asbeck et al. (2) have shown that oxygen ion implantation in GaAs/GaAlAs HBT's can lead to the formation of an isolation layer and to a reduction in the capacitance of the base collector region. Jointly to the oxygen implantation they performed Be implantation to make contact to the base region. However, they have characterized the effect of the oxygen implants on the device and the C-V and I-V measurements performed give information on the sample as a whole and not on each separate layer. They did not study the thermal stability of the compensation in such heterostructures.

We have already shown (5) that Mg is a suitable dopant to contact the base layer with a high hole concentration at the surface ($\sim 10^{19}$ cm^{-3}), avoiding the use of Be which is extremely toxic. In this work, we have performed O^+ and Mg^+ ion implantations in GaAs/GaAlAs heterostructures for bipolar transistor applications in order to form an isolation layer (with O) and to contact the base layer (with Mg). In addition to capacitance measurements, we have carried out electro-chemical profiling thus obtaining the carrier profiles in each layer of the heterostructure. We have studied the influence of the oxygen dose and the influence of the anneal temperature. During this work we have noted an anomalous diffusion of Mg and Be, so we have performed SIMS measurements to explain the behaviour of these impurities.

II - EXPERIMENTAL PROCEDURES

Mg^{24} and O^{16} or O^{18} were implanted at room temperature using a High Voltage Engineering 400 kV accelerator. The samples were 7° off the incidence direction to minimize channeling effects. In consideration to the thickness of the different layers of the heterostructures (table I) we have chosen the following conditions for Mg implantations : 30, 60, 200 keV at doses of 8×10^{13}, 2×10^{14}, 8×10^{14} Mg^+/cm^2 respectively. To reach the collector layer it has been necessary to perform the oxygen implantation with an energy of 550 keV ; we have used doubly charged oxygen ions. We varied the oxygen dose from 2.5×10^{12} to 1.5×10^{14} $O^+.cm^{-2}$. Following implantation, the samples were annealed in a commercial

halogen lamp furnace (6). The method to measure the temperature and the configuration used during R.T.A. have been described elsewhere (7). We varied the annealing temperature from 700 to 900°C.

TABLE I : EPITAXIAL LAYER STRUCTURES USED IN THIS WORK

LAYER	Al FRACTION	TYPE	DOPANT	CONCENTRATION	THICKNESS
CONTACT LAYER	0	n^+	Si	$1-3 \times 10^{18}$ cm^{-3}	0.50–0.60 µm
COLLECTOR	0	n^-	Si	$1-3 \times 10^{16}$ cm^{-3}	0.30–0.52 µm
BASE	0	p^+	Be	$4-5 \times 10^{18}$ cm^{-3}	0.10–0.23 µm
EMITTER	0.3	n	Si	$3-6 \times 10^{17}$ cm^{-3}	0.20–0.33 µm
CONTACT LAYER	0	n^+	Si	$3-5 \times 10^{18}$ cm^{-3}	0.10–0.15 µm

SIMS analysis were performed before and after anneal with a Cameca IMS 3F ion microanalyser equipped with an oxygen source or a cesium source, depending on the element to be analyzed. The compensation of the collector layer has been verified by two methods : we have measured the carrier concentration profiles in the heterostructures with a Polaron semiconductor profile plotter, and we have determined the capacitance of the extrinsic base-collector region. Mesa etching has been used to prepare individual diodes of area $(120\times120)\mu m^2$.

III - RESULTS AND DISCUSSION

a - Compensation by oxygen

Figure 1 shows the carrier concentration profile for a sample implanted with Mg only (curve a) and for a sample implanted with Mg and O (curve b). The two samples have been annealed at 900°C (peak temperature). On curve a, in the p-type region we see the hole profile corresponding to Mg (up to about 0.4µm) and the hole profile corresponding to Be (base dopant). We can already point out the "stairs-like" shape of the Mg^+ profile and the broad Be^+ profile. We will discuss these two points later. In the n-type region we see the doping level in the collector ($n^- \sim 2\times10^{16}$ cm^{-3}) and the electron profile of the collector contact layer ($n^+ \sim 4\times10^{18}$ cm^{-3}). After the oxygen implantation (curve b) the hole profile is identical but we do not have a p-n junction. We measure a very low hole concentration with a corresponding large depletion region. We can measure the carrier profile in the collector contact layer, but it is neccessary to etch all the collector layer which is highly compensated.

Figure 1 clearly shows that it is possible to form an isolated layer by oxygen implantation even after anneal at 900°C. However two facts must be pointed out : 1) the conditions (minimum oxygen dose, anneal temperature) are not reproducible from an heterostructure to another ; 2) for some heterostructures and for low oxygen doses ($<10^{-14}$ O^+/cm^2) the compensation is not stable with time and the highly resistive layer recovers its initial electrical activity within a few days after the anneal process. From these two observations we can conclude that compensation by oxygen is strongly dependant on the oxygen dose and on the material quality.

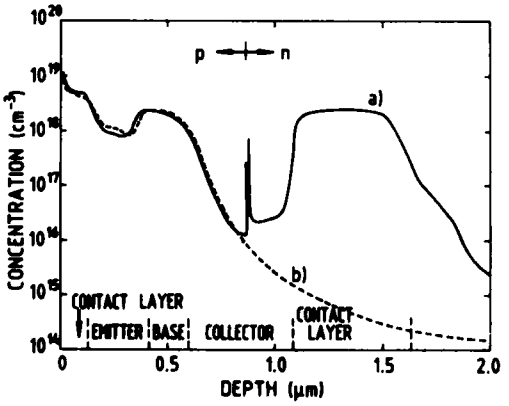


Fig. 1 : Carrier concentrations as a function of depth after anneal at 900°C, (a) for a sample implanted with Mg, (b) for a sample implanted with Mg and O (550 keV, 5×10^{13} O^+/cm^2).

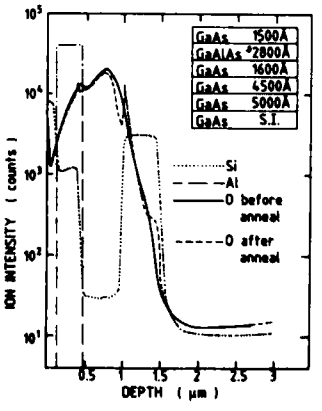


Fig. 2 : SIMS atomic concentrations as a function of depth before and after anneal at 750°C for a sample implanted with Mg^{24} and O^{18} (550 keV, 5×10^{13} O^+/cm^2).

In order to lower the detection limit for oxygen by SIMS we have implanted the isotope 18 of oxygen in some samples. The atomic profiles of O^{18} and Si^{28} are shown in figure 2. The important points to be underlined are the following : first, on the oxygen profile after anneal (750°C) one peak appears at about 1µm ; this peak can be attributed to an oxygen accumulation at the interface collector-collector contact layer during the anneal. The oxygen trapped at this interface will not be efficient to compensate the doping level of the collector, thus the quality of the interfaces will be of major importance. Secondly, in the collector contact layer oxygen diffuses toward the interface collector contact layer - S.I. substrate. These two points allow us to suppose that there is an interaction between Si and O as it has been reported in the litterature (8). However there is no evidence that the compensation is due to the formation of such complexes (9).

To obtain reproducible results it is necessary to implant high oxygen doses ($>1x10^{14} O^+/cm^2$). So we have carried out capacitance measurements for high oxygen doses. Figure 3 shows the behaviour of capacitance versus voltage for a sample implanted with Mg only (curve a) and for a sample implanted with Mg and O with a dose of $1.5x10^{14}$ O^+/cm^2 (curve b). The two samples have been annealed at 800°C. The capacitance of the sample without oxygen decreases with bias voltage. In the contrary, the capacitance of the sample implanted with oxygen is almost independant of bias voltage, indicating that we have formed an insulating layer. Moreover, at zero bias the capacitance with oxygen is equal to 3.5 pF, i.e. 1.5 times lower than the value obtained without oxygen (5.4 pF). This is in accordance with the results presented by Asbeck et al. (2).

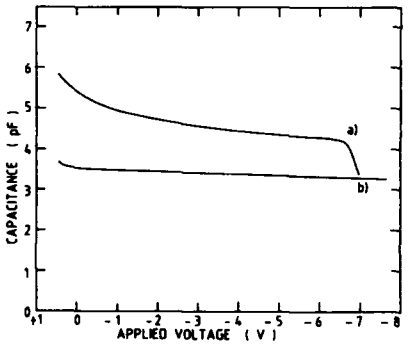


Fig. 3 : Capacitance versus voltage for diodes structures fabricated with a sample without oxygen (a) and a sample implanted with oxygen (b) (550 keV, $1.5x10^{14} O^+/cm^2$).

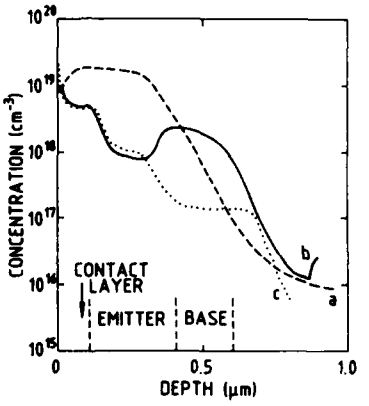


Fig. 4 : Carrier and atomic concentration as a function of depth for a sample implanted with Mg and O($1x10^{14} O^+/cm^2$) (a) Mg atomic profile before anneal, (b) hole profile after anneal at 850°C, (c) Mg atomic profile after anneal at 850°C.

b - Be - Mg interaction

As mentionned above, we have noted in the carrier profiles an anomalous behaviour of Mg and Be. To clarify this problem we have performed SIMS measurements before and after anneal for different cases.

Figure 4 concerns a sample implanted with Mg and O ($1x10^{14} O^+/cm^2$) and annealed at 850°C. It shows a strong outdiffusion of Mg leading to an important decrease of the Mg concentration, as compared to the atomic profile before anneal (although the Mg concentration at the surface is still high 10^{19} cm^{-3} ; and Mg reaches the base with a concentration of about $10^{18} cm^{-3}$). Moreover, the Mg atomic profile presents three steps and goes up to the tail of the Be profile. It is however interesting to note the good activation of Mg since the carrier profile is identical to the Mg atomic profile up to the base.

In figure 5 we have drawn the Be atomic profile before and after anneal for a sample which has also been implanted with Mg and with an oxygen dose of $1x10^{14} O^+/cm^2$ and annealed at 850°C. We clearly see the important diffusion of Be both towards the emitter layer and the collector layer. This diffusion is asymmetric, it is more important towards the emitter layer, i.e. towards the Mg implants. This effect is neither due to the high temperature anneal itself, nor to the oxygen implant. As a matter of fact, in the case of an unimplanted sample and a sample implanted with oxygen only, we do not see such an asymmetric diffusion. We can therefore suspect an interaction between the two p-type dopants Mg and Be. In order to verify this assumption we have performed Mg implantations with the two lower energies (30 and 60 keV, with doses of $8x10^{13}$ and $2x10^{14}$ Mg^+/cm^2 respectively) so that Mg does not reach the Be doped layer.

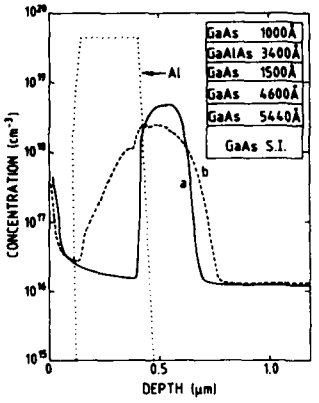


Fig. 5 : SIMS Be atomic concentrations as a function of depth for a sample implanted with Mg and O(550 keV, $1x10^{14}$ O/cm²) (a) before anneal, (b) after anneal at 850°C.

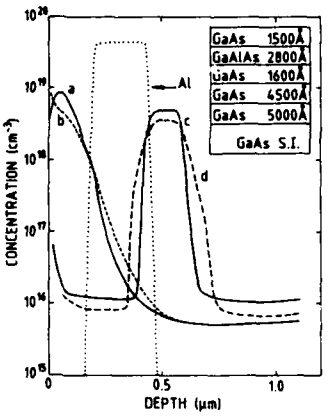


Fig. 6 : SIMS atomic concentrations as a function of depth for a sample implanted with Mg (30, 60 keV and $8x10^{13}$, $2x10^{14}$ Mg^+/cm^2 respectively). (a) Mg atomic profile before anneal, (b) Mg atomic profile after anneal at 850°C, (c) Be atomic profile before anneal, (d) Be atomic profile after anneal at 850°C.

Figure 6 shows that after anneal (850°C) there is still an outdiffusion of Mg and a slight indiffusion but we do not see any step in the atomic profile. On the other hand, there is also a diffusion of Be but it is much less important than in the previous case (with three Mg implantations), and it is symmetrical. This effect may result from competition of the two species for the available gallium sites as it has been reported by Houston et al. in the case of Zn and Be (10).

CONCLUSION

We have shown in this work that oxygen ion implantation allows to obtain high resistivity layers stable at high temperature (up to 900°C) in GaAs/GaAlAs heterostructures. But the compensation is very dependant on the quality of the starting material and to obtain reproducible results it is necessary to implant high oxygen doses (> 10^{14} O^+ cm^{-2}).

Another point which has been underlined is the strong interaction between Mg and Be resulting from the competition of these two p-type dopants for the Ga sites ; and leading to an anomalous diffusion of Mg and Be during the high temperature anneal.

The authors acknowledge Dr J. TASSELLI for her help in the technological steps, A. SIBILLE for C-V measurements, M. GAUNEAU for some SIMS measurements. They would like to thank Dr E.V.K. RAO for helpfull discussions and Dr F. ALEXANDRE and J. RIOU who furnished the heterostructures.

REFERENCES

/1/ Adachi S., Ishibashi T., IEEE Electron Device Lett. EDL-7, 32 (1986).
/2/ Asbeck P.M., Miller D.L., Anderson R.J., Eisen F.H., IEEE Electron Device Lett. EDL-5, 310 (1984).
/3/ Fischer R., Morkoç H., IEEE Electron Device Lett. EDL-7, 359 (1986).
/4/ Chang M.F., Asbeck P.M., Wang K.C., Sullivan G.J., Miller D.L., Electron. Lett. 22, 1173 (1986).
/5/ Duhamel N., Daoud-Ketata K., Descouts B., Krauz P., Gauneau M., Godefroy S., J. Electron. Mater., 15, 377 (1986).
/6/ Halogen lamps furnace from Atelier d'Electro Thermie : AET Grenoble (France).
/7/ Descouts B., Duhamel N., Daoud-Ketata K., Krauz P., Godefroy S., J. Appl. Phys. 60, 450 (1986).
/8/ Berth M., Venger C., Martin G.M., Electron. Lett. 17, 873 (1981).
/9/ Favennec P.N., J. Appl. Phys. 47, 2532 (1976).
/10/ Houston P.A., Shepard F.R., Spring Thorpe A.J., Mandeville P., Margittai A., Appl. Phys. Lett. 52, 1219 (1988).

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

NON-ALLOYED Ge/Pd OHMIC CONTACT FOR GaAs MESFET'S

A. PACCAGNELLA, C. CANALI*, G. DONZELLI**, E. ZANONI***, R. ZANETTI* and S.S. LAU*

Dipartimento di Ingegneria, Università di Trento, I-38050 Mesiano di Povo (TN), Italy
**Dipartimento di Elettronica e Informatica dell'Università, I-35131 Padova, Italy*
***Telettra SpA, Via Trento 30, I-20059 Vimercate (MI), Italy*
****Departimento di Elettronica ed Elettrotecnica dell'Università, I-70125 Bari, Italy*
**Dept. EECS, University of California at San Diego, La Jolla, CA 92093, U.S.A.*

Abstract - GaAs MESFET's with non-alloyed ohmic contacts have been achieved through a solid phase reaction of the Ge/Pd/GaAs(xtl) structure upon annealing at 325°C for 30 min. Different Au-based overlayers over Ge/Pd have been tested for device applications and compared with a conventional AuGeNi contact. The thermal stability of the contact resistivity has been evaluated through long-term storages at 300°C.

1 - INTRODUCTION

New ohmic contacts to n-GaAs have been developed with the aim of improving the contact properties as well as the technological features in comparison with the usual AuGeNi metallization. The fabrication process of AuGeNi, based on the formation of a liquid phase during a short time annealing, can still give repeatibility problems among successive runs and poor uniformity through the single wafer, in absence of an accurate GaAs surface preparation /1/. The search for a more controllable realization process pushed toward investigation of ohmic contacts achieved through solid state interactions, such as Ge epitaxy on GaAs /2/ or solid phase reactions between GaAs and metal overlayers /3/, which showed both excellent electrical and morphological properties.
Application of ohmic contacts based on the solid state reaction of the Ge/Pd/GaAs structure to GaAs MESFET's has been investigated in this work, with a particular attention to the technological realization process. The Ge/Pd ohmic contact resistivity and thermal stability, and the dc and rf MESFET performances have been compared with the results obtained in the case of conventional AuGeNi contacts, for 0.25 W devices.

2 - PdGe CONTACT REALIZATION

Ge(113 nm)/Pd(44 nm) bilayers have been e-beam deposited on conventionally chemically cleaned <100>n-GaAs substrates 1.5×10^{17} cm^{-3} S doped, patterned for measurements conform to the Transmission Line Method (TLM) /4/. After the contact definition by lift-off, ohmic contacts have been achieved through annealing at 325°C for 30 min in a forming gas flux. Such a treatment gives the lowest contact resistivity as shown in /5/.
The reaction kinetics has been monitored at different times during the annealing by AES, and results are schematically summarized in Fig. 1. After the shortest annealing period (5 min) the whole Pd is involved in the polycrystalline PdGe phase, through a solid state reaction with Ge. Pd_2Ge, which appears before PdGe in the "ideal" phase formation sequence, is already no more detectable. An inhomogeneous unreacted Ge layer is left at the surface, and a small Ge accumulation is detected at the PdGe/GaAs interface, attributed to the epitaxial growth of Ge over GaAs as shown in /6/. The thickness of the Ge epilayer increases with the annealing time, until all the excess surface Ge has been consumed, i.e. after 30 min. A non-flat PdGe/Ge(xtl) interface still remains at the end of the Ge epigrowth, due to the inhomogeneous columnar

growth of the PdGe grains /7/.

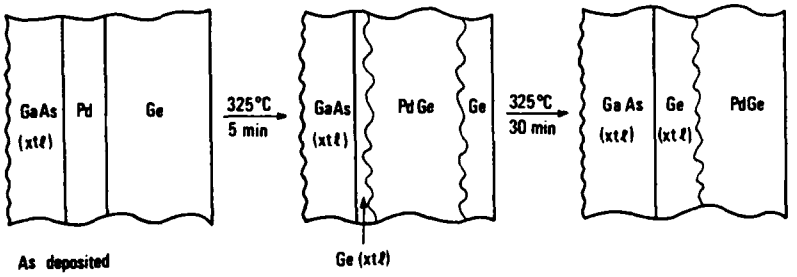


Fig. 1 - Schematic representation of the solid phase epitaxy of Ge over GaAs during the 325°C annealing of the Ge/Pd/GaAs system.

The formation of a crystalline Ge film is favoured due to the lower free energy than amorphous Ge. The Ga and As amount in the PdGe layer is below the AES detection limits, indicating that only a minor substrate dissolution might occur, and consequently no contact sinking into GaAs. Ohmicity is assessed through the Ge/GaAs interface by an electron tunneling conduction mechanism, due to the expected formation of a n^+-GaAs layer Ge doped during the annealing period /3/.
The contact resistivity of the Ge/Pd metallization has been evaluated in a test pattern conform to TLM, the spacing among the pads varying from 5 to 25 μm. From the specific contact resistivity ϱ_c derived through TLM, the linear contact resistivity ϱ_t can be deduced, more suitable for technological applications, as $\varrho_t = \varrho_c / L_t$. The transfer length L_t defines the width of the GaAs/metal interfacial stripe through which most of the current is injected from/in GaAs. L_t is typically 1 μm long for Ge/Pd contacts. In Tab. 1 the average value ϱ_t and the standard deviation $\Delta\varrho_t$ are reported for the Ge/Pd contact, and compared with the values obtained for a conventional Au(150 nm)/Ni(30 nm)/Au(60 nm)/Ge(30 nm) contact annealed at 450°C for 30 s. ϱ_t is much larger for AuGeNi than for Ge/Pd, while $\Delta\varrho_t$ is almost the same for both metallization, just after the ohmic contact formation.

3 - OVERLAYER METALLIZATIONS

In a double recessed channel structure, widely used for power MESFET's, source and drain ohmic contacts act as a mask for the channel first etching. Ge/Pd is not suitable for such application, due to its etchability by the GaAs chemical attacks. A Au-based overlayer represents a proper solution, useful moreover for bonding purposes, thermal dissipation improvement and decreasing of the contact sheet resistance. The overlayer is deposited in the same evaporation run of Ge/Pd and passes the same fabrication steps.

	Ge/Pd	AuGeNi	Au/Ge/Pd	Au/Pd/Ti/Ge/Pd	Au/Pt/Ti/Ge/Pd
ϱ_t (Ω mm)	0.16	0.24	1.19	0.14	0.12
$\Delta\varrho_t$ (Ω mm)	0.04	0.05	0.21	0.02	0.02

Tab. 1 - Average value ϱ_t and standard deviation $\Delta\varrho_t$ of the linear ohmic contact resistivity for different metallizations on n-GaAs.

The use of a bare Au layer is detrimental to the contact properties, as shown in Tab. 1, due to the metallurgical interactions with the underlying Ge/Pd layers during the 325°C annealing. The insertion of a proper diffusion barrier between Au and Ge, such as Pt/Ti or Pd/Ti, hampers the Au indiffusion, leading to excellent ϱ_t and $\Delta\varrho_t$ values (see Tab. 1). The ϱ_t values

slightly lower than for bare Ge/Pd, can be justified by observing that the microprobes used for TLM measurements are far from the contact edge much more than a transfer length. Hence the parassitic sheet resistance contributes to the total measured resistance and rises up the ϱ_t values. Such an experimental artifact is minimized by decreasing the contact sheet resistance, as shown in Tab. 1 for samples with the metal overlayer.

4 - THERMAL STABILITY

The thermal stability of the Ge/Pd contact is a fundamental requirement for applications in power devices where high temperatures, well over 100°C,are steadily reached during the operating life in the channel region. Accelerated aging tests at 300°C have been performed and results concerning the ϱ_t variations are shown in Fig. 2 for Ge/Pd with/without overlayers and for AuGeNi contacts. The Ge/Pd bare metallization shows a relatively fast increase of ϱ_t during the first tens of hours (Fig. 2a) followed by an almost linear slow growth, without any further increase of data spreading. The contact resitivity is almost twice after 200 hours, but is still acceptable for device applications. The physical mechanism determining the contact degradation has been tentatively attributed to the microstructural modifications affecting the PdGe/Ge(xtl) interface /7/. However, even impurity indiffusion from the annealing atmosphere could contribute to the observed contact resistivity degradation. On the contrary, the average ϱ_t remains roughly constant in the case of AuGeNi (Fig. 2a), but with a substantial increase of the data dispersion, up to $\Delta\varrho_t$=0.06 Ω mm.

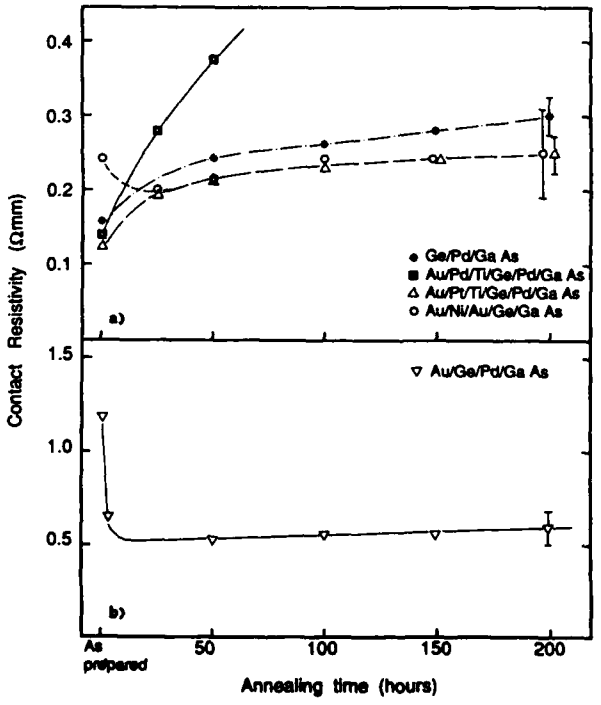


Fig. 2 - Variation of the linear contact resistivity ϱ_t versus annealing time at 300°C for different ohmic contacts. Lines are drawn for the eye guide.

The active role of Au in determining the contact resistivity is responsible for the peculiar behaviour of the Au/Ge/Pd metallization upon storages at 300°C, much different from Ge/Pd. The contact resistivity noticeably decreases, as well as the data spread, within the first hours at 300°C (Fig. 2b), indicating that even lower ϱ_t values could be achieved by optimizing the fabrication process. For longer annealing times,an average value larger than 0.5 Ω mm is main-

tained, without any noticeable further degradation,but not suitable for MESFET ohmic contacts. In the case of the Au/Pd/Ti overlayer (Fig. 2a), a dramatic and steep increase of ϱ_t is observed, suggesting the onset of detrimental metallurgical reactions between Ge/Pd and the overlayer. In this case the ϱ_t values quickly become not suitable for MESFET applications. The contact resistivity values for the Au/Pt/Ti overlayer (Fig. 2a) are always smaller than that of bare Ge/Pd, and very close to that of AuGeNi, but with a much narrower data spread. Thus, Au/Pt/Ti appears as a suitable overlayer for device applications.

5 - MESFET REALIZATION

Ge/Pd metallization with both Au/Pd/Ti or Au/Pt/Ti overlayers has been used for source and drain pads in 0.25 W, 4 double recessed gate fingers MESFET's fabricated by a lift-off photolithographic method. The substrate was a Cr-doped S.I. <100> GaAs wafer, over which a buffer undoped layer 2 μm thick and an active layer 0.4 μm thick and S doped at 1.5×10^{17} cm^{-3} were grown by VPE. 1μm long Au/Pd/Ti gate contact was deposited after the ohmic contact realization. A 200 nm SiN passivating layer was PECVD deposited prior chip scribing and mounting. For comparison, MESFET's have been realized with a similar procedure by using the AuGeNi ohmic contact previously described. Typical values of Ge/Pd MESFET parameters are Idss=150 mA, gm=60 mS, Gain=8 dB and Poutput=23 dBm at 8 Ghz at 1 dB compression, and are similar also for the AuGeNi devices. An evaluation of the parasitic source and drain resistances has been accomplished by using the Fukui method /8/ coupled with the measurements of the forward I-V diode characteristics at the gate-source and gate-drain terminal pads, on the AuGeNi and Au/Pd/Ti/Ge/Pd devices. As expected, the average source and drain resistances for Ge/Pd contacts, 1.59 and 1.95 Ω respectively, are lower than for the AuGeNi devices, 1.73 and 2.61 Ω respectively, owing to the lower ϱ_t value.

6 - CONCLUSIONS

The Ge/Pd ohmic contact offers a suitable option, alternative to the usual AuGeNi, for MESFET applications due to the excellent morphological and electrical characteristics. The solid phase reaction ruling the ohmic contact formation allows for a better process control resulting in good repeatibility with a small contact resistivity value dispersion. The use of a Au/Pt/Ti overlayer allows for a straightforward application of such ohmic contact in a conventional MESFET fabrication process, and moreover improves the contact thermal stability.

Work partially supported by CNR-P.F. Materiali e Dispositivi per l'Elettronica a Stato Solido.

REFERENCES

/1/ Shih, Y.-C., Murakami, M., Wilkie, E.L. and Callegari A., J. Appl. Phys. 62 (1987) 582.
/2/ Stall, R.A., Wood, C.E.C., Board, K., Daudekar, N., Eastman, L.F. and Devlin, J., J. Appl. Phys. 52 (1981) 4062.
/3/ Marshall, E.D., Zhang, B., Wang, L.C., Jiao, P.F., Chen, W.X., Sawada, T., Lau, S.S., Kavanagh, K.L. and Kuech, T.F., J. Appl. Phys. 62 (1987) 942.
/4/ Berger, H.H., Solid-St. Electron. 15 (1972) 145.
/5/ Sawada, T., Chen, W.X., Marshall, E.D., Kavanagh, K.L., Kuech, T.F., Pain, C.S. and Lau, S.S., Mat. Res. Soc. Symp. Proc. 54 (1986) 409.
/6/ Marshall, E.D., Chen, W.X., Wu, C.S., Lau S.S. and Kuech, T.F., Appl. Phys. Lett. 47(1985) 298.
/7/ Paccagnella, A., Migliori, A., Vanzi, M., Zhang, B. and Lau, S.S., Proc. 17th ESSDERC (1987) 605.
/8/ Fukui, H., Bell Syst. Techn. J. (1979) 711.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

RAPID THERMAL ANNEALING OF TiW SCHOTTKY CONTACTS ON GaAs

M. VAN HOVE, M. de POTTER, W. DE RAEDT, G. ZOU and M. VAN ROSSUM

IMEC vzw, Kapeldreef 75, B-3030 Leuven, Belgium

Résumé - La stabilité de l'interface TiW/GaAs durant le recuit rapide à des températures allant de 700°C à 1050°C a été étudiée avec les techniques SIMS, Auger, RBS, XRD, TEM à haute résolution, EDS, et des mesures IV et CV. Les mesures de contacts Schottky ont démontré une forte croissance de hauteur de barrière, allant de 0.70eV après déposition à 0.95eV après un recuit à 950°C. Les résultats indiquent une augmentation artificielle de la hauteur de barrière due au dopage de type p du substrat par une diffusion de Ti.

Abstract - The stability of the TiW/GaAs interface during rapid thermal annealing at temperatures between 700°C and 1050°C has been investigated using SIMS, Auger, RBS, XRD, cross-sectional TEM, EDS, IV and CV analysis. Schottky contact measurements showed a steady increase of the barrier height with annealing temperature from the as-deposited value of 0.70eV to 0.95eV after annealing at 950°C. The results are consistent with an artificial barrier height enhancement due to p-doping of the substrate by indiffusing Ti.

1 INTRODUCTION

Recently, there has been an increasing interest in the application of refractory gate metals in a self-aligned GaAs metal-semiconductor field- effect transistor (MESFET) technology. The processing requires the gate material to maintain a good rectifying contact with low leakage current and high breakdown voltage when subjected to high temperature annealing (~900°C) necessary to activate the n^+ implant. The refractory metal/GaAs interface has to be mechanically, chemically and electrically stable. Most problems are related with peeling off, chemical reaction and interdiffusion. Although TiW and TiW-based alloys have been frequently considered as appropriate candidates /1,2,3/, the stability of TiW during rapid thermal annealing (RTA) has not yet been ascertained. In this contribution we report an extensive study of the TiW/GaAs interface after RTA.

2 EXPERIMENTS

After degreasing and in situ Ar sputtercleaning, thin films (50nm-200nm) of TiW were deposited on Si-doped GaAs wafers (n ~ E17/cm^3) by dc magnetron sputtering from a 30%Ti/70%W compound target. The Ar pressure was optimized to obtain minimum film stress conditions. The composition of the deposited TiW films was determined by RBS as 25%Ti/75%W. The samples were subjected to RTA in forming gas ambient at 700°C-1050°C, 10s. Electrical measurements were performed on 90μmx90μm Schottky diodes patterned by standard lithography. Alloyed AuGe/Ni was used as backside contact. In order to avoid diode leakage at the edges about 200nm of GaAs was etched away in H2SO4:H2O2:H2O before measurement.

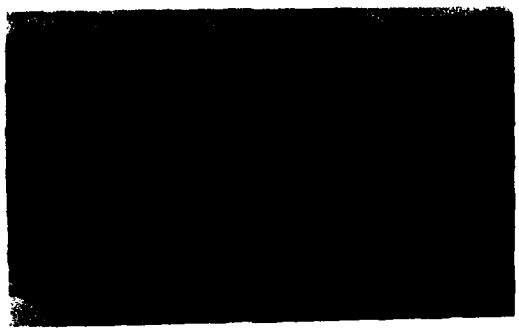


Fig.1 - High resolution cross-sectional TEM image of the TiW/GaAs interface after annealing at 900°C, 10s.

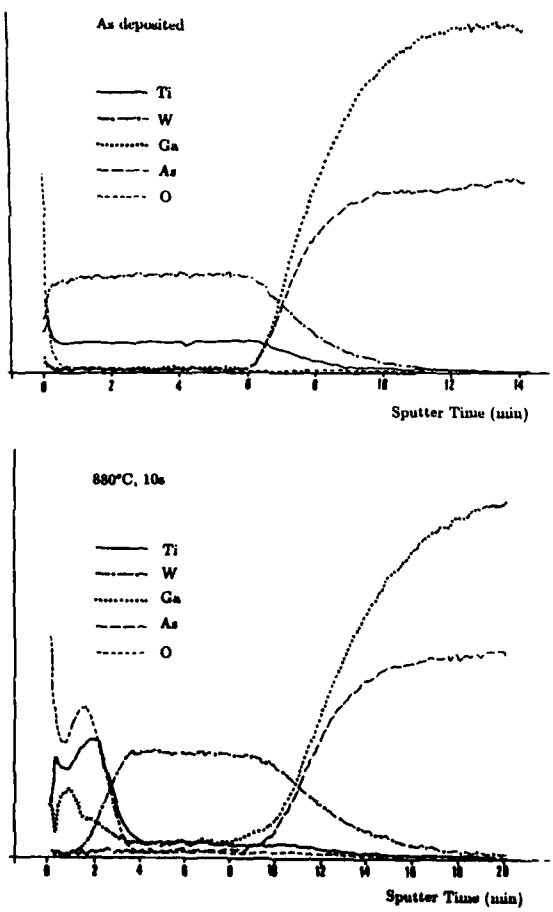


Fig.2 Auger depth profiles of Ti, W, Ga, As and O of as-deposited and annealed (880°C, 10s) TiW films on GaAs.

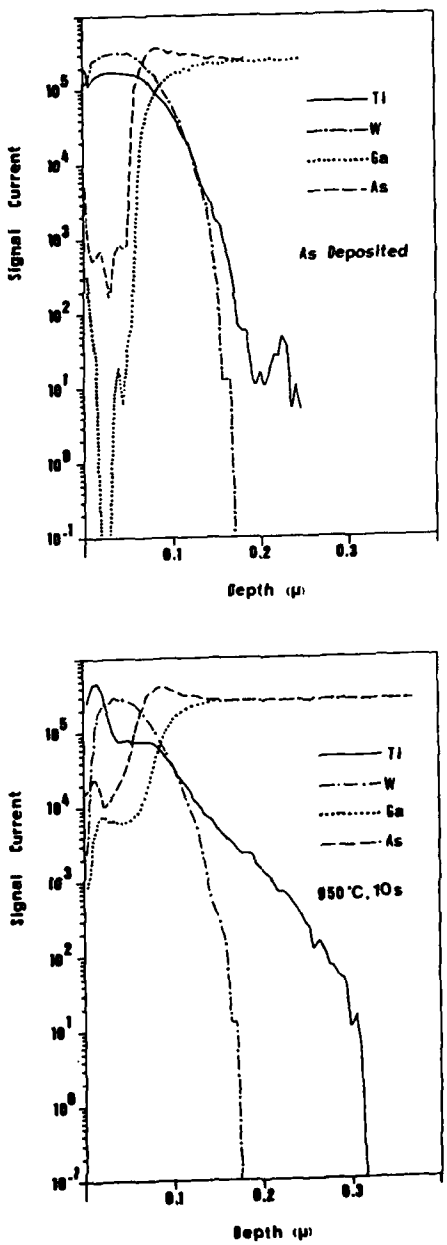


Fig.3 - SIMS depth profiles of Ti, W, Ga and As of as-deposited and annealed (950°C, 10s) TiW films on GaAs.

3 RESULTS AND DISCUSSION

All the films examined showed good adherence even after annealing at the highest temperatures. XRD analysis showed a clear W α-phase spectrum. After annealing, the peaks showed a reduction of the halfwidth, which can be attributed to grain size growth. High resolution cross-sectional TEM examination (fig.1) showed good TiW/GaAs interface stability. However, local island growth of epitaxial phases induced by RTA were found. No clear morphologic structure in the TiW film could be detected by TEM, due to the very strong absorption of the electron beam by the metal. Auger (fig.2), SIMS (fig.3), RBS (fig.4) and EDS show significant motion of the Ti resulting in surface accumulation as well as Ti diffusion into the GaAs substrate. The W/GaAs interface however remains stable for temperatures up to ~1000°C.

Forward and reverse current-voltage characteristics of as-deposited and annealed TiW Schottky contacts are shown in fig.5. They show a continuous increase in both the ideality factor and the Schottky barrier heights extracted from IV- and CV-data (fig.6). Diodes annealed at 950°C have a barrier height as high as 0.95eV (IV-value), which is considerably higher than the value of 0.70eV obtained for nonannealed diodes. Significant degradation of the diodes occurs after annealing at temperatures higher than 1000°C, coinciding with the onset of W/GaAs interdiffusion.

Breakdown characteristics of as-deposited and annealed TiW/GaAs contacts are depicted in fig.7. Unannealed diodes show a soft turn-on of the breakdown, which is characteristic for Schottky diodes. After annealing the breakdown voltage increases and the characteristics become avalanche-like. Measurements at higher temperatures show an increase in breakdown voltage, which confirms avalanche being the breakdown mechanism.

The RTA-induced properties of the contacts are consistent with the Shannon contact structure /4/ (metal/p^+-GaAs/n-GaAs). The p^+-formation is attributed to the indiffusion and activation of Ti which is known to have acceptor levels in GaAs /5,6/. Due to the p^+-layer formation a barrier height enhancement occurs. With increasing annealing temperature the p^+- layer becomes thicker and/or more highly doped. This explains the steady increase of the barrier height with annealing temperature. It is estimated that a 1.E18/cm^3 p^+-layer of 200$\AA$ thickness can account for the 250meV difference in the barrier height between the as-deposited and 950°C annealed diodes. The avalanche breakdown behaviour of the annealed diodes is consistent with the formation of a p^+-layer between the metal and the n-GaAs substrate. Breakdown will be initiated by avalanche multiplication at the p^+/n junction.

Our observations show that rapid thermal annealed TiW/GaAs contacts, and Shannon contacts in general, can have advantageous properties for application in a self-aligned MESFET technology.

ACKNOWLEDGMENTS

The authors are pleased to thank M. Meuris for SIMS, H. Bender for Auger, A. Demesmaecker for EDS and XRD and J. Vanhellemont for TEM analysis.

REFERENCES

/1/ Kohn, E., IEDM Techn. Dig. (1979) 775.
/2/ Yokoyama, N., Mimura, T., Fukata, M. and Ishikawa, M., International Solid-State Circ. Conf. Dig. of Techn. Papers (1981) 218.
/3/ Geissberger A.E., Sadler R.A., Balzan M.L. and Crites J.W., J. Vac. Sci. Technol. B5(6) (1987) 1701.
/4/ Schwartz, G.P.and Gaulieri, J. Electrochem. Soc. 133 (1986) 1266.
/5/ Kornilov, B.V., Marchukov,L.V. and Ergakov V.K., Sov. Phys. Semicond. 8 (1974) 141.
/6/ Ushakov, V.V. and Gippius A.A., Sov. Phys. Semicond. 16 (1982) 1042.

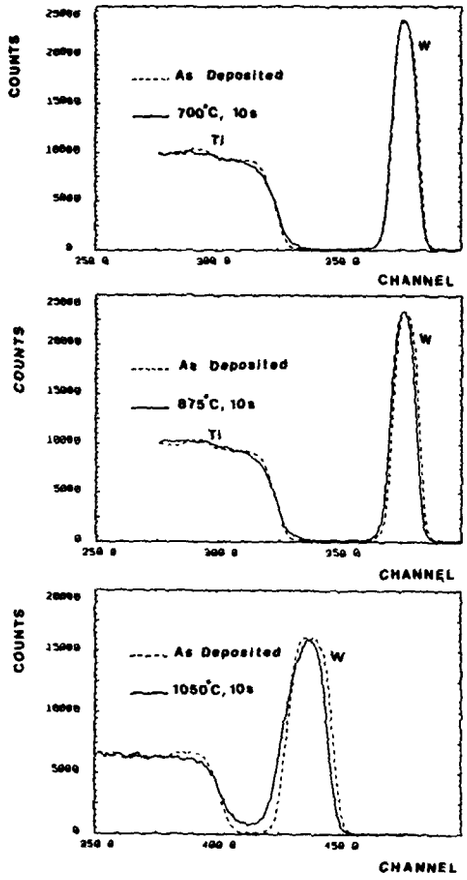


Fig.4 - RBS spectra of TiW/GaAs samples annealed at different temperatures.

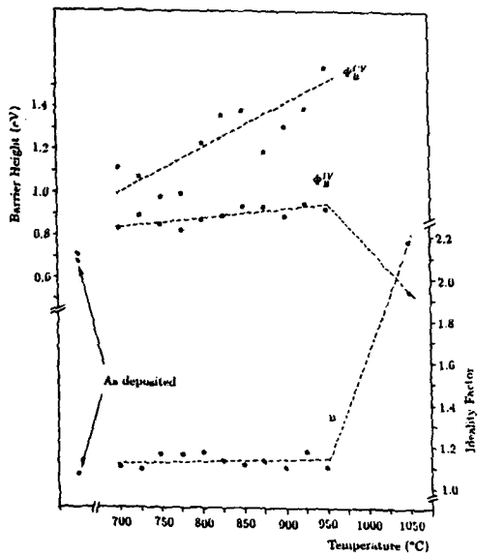


Fig.6 - Ideality factors and barrier heights for TiW/GaAs contacts annealed at different temperatures.

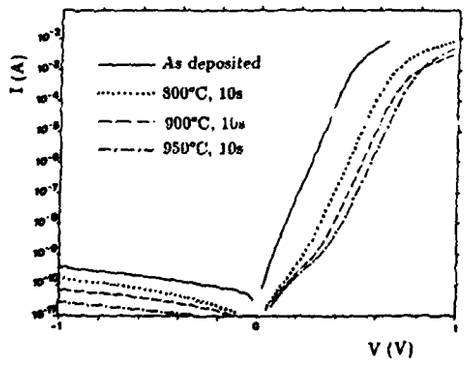


Fig.5 - Forward and reverse current- voltage characteristics for TiW/GaAs contacts annealed at different temperatures.

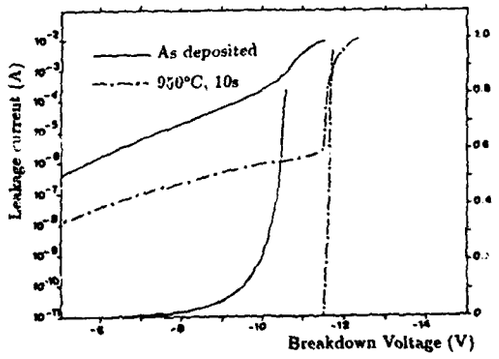


Fig.7 - Reverse leakage current versus reverse applied voltage of as-deposited and annealed (950°C, 10s) TiW/GaAs contacts.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

0.9 eV POTENTIAL BARRIER SCHOTTKY DIODE ON 0.75-0.5 eV GAP $Ga_xIn_{1-x}AS\backslash a\text{-}Si{:}H\backslash Pt$

A. DENEUVILLE, F. VALENTIN and S. BELKOUCH

Laboratoire d'Etudes des Propriétés Electroniques des Solides, CNRS, BP 166, F-38042 Grenoble Cedex, France, Associated with University Joseph Fourier, Grenoble, France

Résumé - Les courbes I(V,T) de ces structures pour 1000 Å de a-Si:H et x = 0.2 et 0.47 sont similaires à celles des diodes Pt/a-Si:H et en conséquence controlées par l'interface Pt/a-Si:H. Ceci ouvre de nouvelles perspectives pour les MESFET sur $Ga_xIn_{1-x}As$.

Abstract - The I(V,T) curves of such structures for 1000 Å of a-Si:H and x = 0.2 and 0.47 are both similar to those of Pt/a-Si:H diodes, so controlled by the Pt/a-Si:H interface. This opens new perspectives for MESFET on $Ga_xIn_{1-x}As$.

I - INTRODUCTION

The transmission wavelengths of optical fibers fit the gap of $Ga_{0.47}In_{0.53}As$ (0.75 eV) or $Ga_{0.2}In_{0.8}As$ (0.5 eV). Integrated circuits with detection and treatment of the photoinduced current are wished. At this moment neither MISFET nor MESFET structures are satisfactory respectively from problems to prepare good insulators on III-V semiconductors and from the pinning of the Fermi level at the metal III-V interface very close to the conduction band edge. Loualiche et al /1/ have shown that a 0.78 eV potential barrier can be obtained with the structure $Ga_{0.47}In_{0.53}As\backslash a\text{-}Si{:}H(1000\ Å)\backslash Pt$. The aim of this work is to reproduce this result, to improve the performance of the structure and to check if it works also on the 0.5 eV gap semiconductor $Ga_{0.2}In_{0.8}As$ of the same family.

2 - PREPARATION AND APPARATUS

The $Ga_xIn_{1-x}As$ films grown by MBE on semi-insulating InP are supplied by Loualiche of CNET Lannion.

The a-Si:H films (1000 Å) are deposited by glow discharge decomposition of SiH_4 in SiH_4 10 %/H_2 90 % mixture at a total pressure of 1.9 Torr and 230° C on the $Ga_xIn_{1-x}As$ and also on high resistivity Si monocristal for physicochemical characterisation. This is mainly done from the analysis of its infrared absorption by a Perkin Elmer 683 Infrared Spectrophotometer.

The height of the potential barrier is derived from the analysis of the I(V,T) characteristic which is recorded from 200 to 400 K by a Keithley 617 electrometer controlled by a microcomputer Apple II.

3 - EXPERIMENTAL RESULTS AND DISCUSSION

The a-Si:H is an amorphous wide band gap (~ 1.8 eV) semiconductor whose properties depends mainly on the hydrogen content. At low thicknesses, oxygen can also be included in the film which damages its properties. Hydrogen content as well as oxygen contamination of a-Si:H film can be studied through the specific absorption bands of Si-H and Si-O bonds. There are several absorption bands according to the mode of vibration of the bond. We check here the wagging mode of the Si:H bond around 630 cm^{-1} and the stretching mode of the Si-O bond around ~ 1050 cm^{-1}. Their oscillator strength is constant /2/, /3/, which allows a derivation of the H and O concentration from the area of the absorption bands after calibration. For instance, the I-R absorption band for the 1000 Å a-Si:H film is given on Fig. 1. From the aera of the 630 and 1050 cm^{-1} bands we deduce an hydrogen content of ~ 28 % and an oxygen contamination of ~ 0.2 %. The hydrogen content is as usually higher than in thick films, the oxygen contamination is at a quite satisfactory low level.

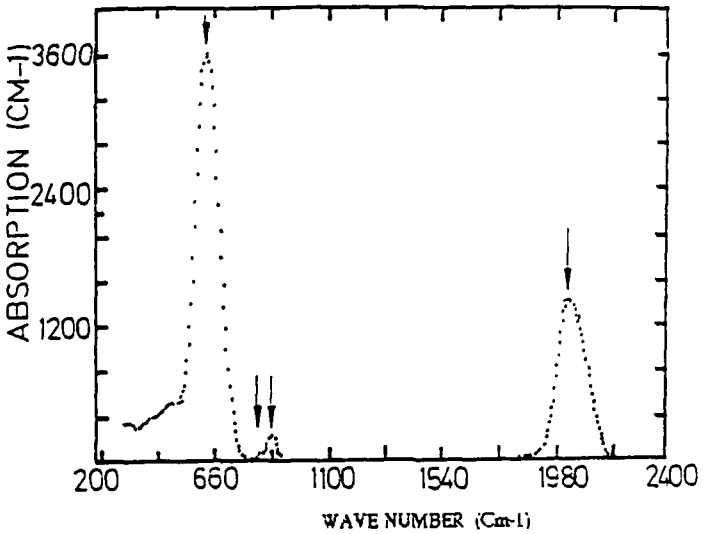


Fig. 1 - A typical infrared absorption curve for 1000 Å of a-Si:H deposited at 230° C by a 50 kHz glow discharge.

The I(V,T) curves of both the $Ga_{0.47}In_{0.53}As$ and $Ga_{0.2}In_{0.8}As$ structrues exhibit a Schottky behaviour as shown respectively on Fig. 2 and 3. In both cases, the current density increases from $\sim 10^{-11}A/cm^2$ to $\sim 10^{-4}A/cm^2$, and the ideality factor n decreases from 1.6 down to 1.1 as the temperature increases. There is also for both structures an inactivated current component appearing below ~260 K and whose relative intensity increases as the temperature decreases. This may be tentatively attributed to a tunnel component through the localized levels of a-Si:H

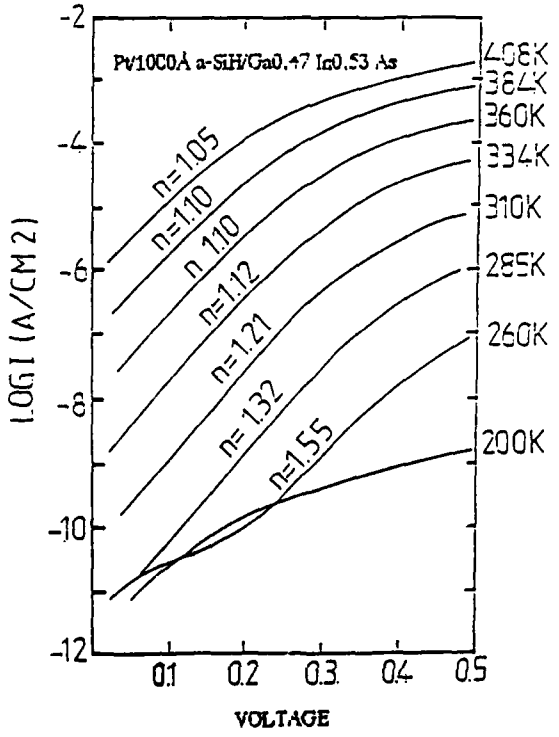


Fig. 2 - The I(V,T) curves of a Pt\a-Si:H\$Ga_{0.47}In_{0.53}As$ structure.

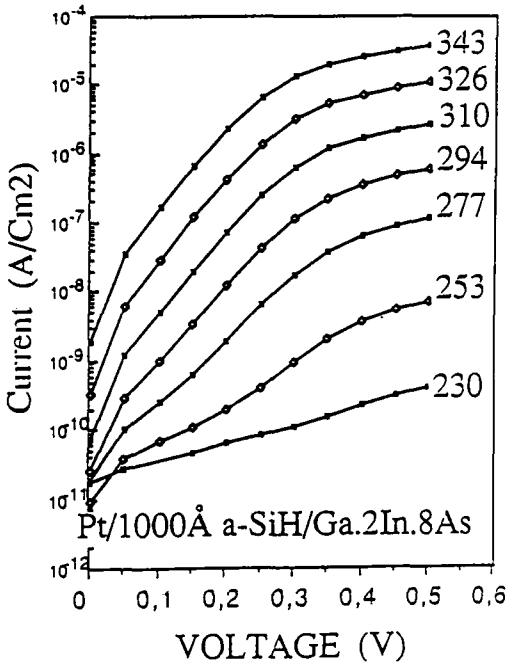


Fig. 3 - The I (V,T) curves of a Pt/a-Si:H/$Ga_{0.2}In_{0.8}As$ structure.

From the extrapolation of the linear part of the forward characteristic to zero voltage, we deduce the saturation density of current at each temperature. It is thermally activated, with respectively 0.94 eV (Fig. 4) and 0.84 eV for the $Ga_{0.47}In_{0.53}As$ and $Ga_{0.2}In_{0.8}As$ structures.

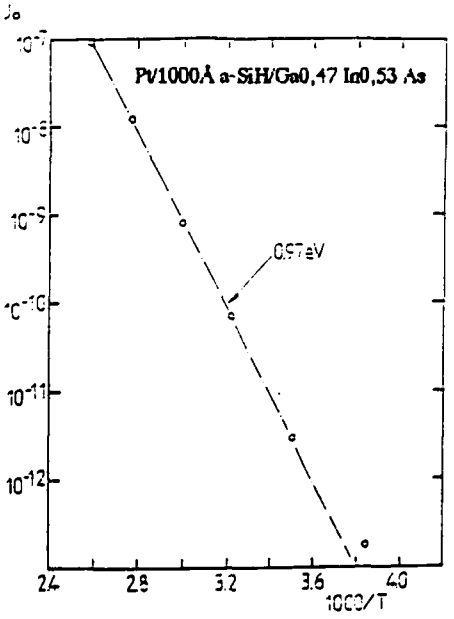


Fig. 4 - The saturation current Js versus the reciprocal temperature for the Pt/a-Si:H/$Ga_{0.47}In_{0.53}As$ structure.

So the I(V) curves of these structures can be represented by $J = A \exp - \frac{q\phi_B}{kT} \left[\exp \frac{qv}{nkT} - 1 \right]$ as for the usual Schottky diode, except that as for Metal/a-Si:H /4/ diode we cannot choice between a diffusion or thermoionic behaviour for the saturation current. The ideality factor and the saturation current are higher and the potential barrier slightly lower for the x = 0.2 than for the x = 0.47 structure.

If we compare to the results of Loualiche et al /1/ for x = 0.47 we get here at room temperature n = 1.2 instead of n = 2 Js $\sim$ 10^{-10} A/cm^2 instead of 10^{-7} A/cm2, $q\phi_B \sim$ 0.9 eV instead of 0.78 eV for the same a-Si:H thickness of 1000 Å.
for x = 0.2, n = 1.35 Js $\sim$ 10^{-10} A/cm^2 and $q\phi_B \sim$ 0.84 eV.

So, the structures have been improved in regard to the previous ones (lower n and Js, higher $q\phi_B$)and the results are similar for x = 0.2 and 0.47. The ideality factor, current density and potential barrier obtained here are also similar to those of Metal/a-Si:H Schottky diodes. So, this suggests, that the I(V,T) curves of this kind of structure are mainly controlled by the Metal/a-Si:H interface, which is confirmed by measurement on the same type of structure with various cristalline semiconductors /5/.

CONCLUSION

We have shown that Schottky type structures Pt/a-Si:H/$Ga_xIn_{1-x}As$ (x = 0.2 and 0.47) can be obtained for 1000 Å of a-Si:H with good ideality factors, low saturation currents high potential barriers, significantly improved in regard to the previous results on x = 0.47, similar to those of Pt/a-Si:H diodes. These caracteristics are controlled by the Pt/a-Si:H interface, and so, the potential barrier 0.9 eV is much higher than that of Pt/$Ga_xIn_{1-x}As$ ($\leq$ 0.3 eV) which will considerably help for MESFET structures.

REFERENCES

/1/ S. Loualiche, C. Vaudry, L. Henry, A. Le Corre, Electronics Letters 22, (1986), 897.
/2/ H. Shanks, C.J. Fang, L. Ley, M. Cardona, F.J. Demond, S. Kalbitzer, Phys. Stat. Sol.(b) 100, (1980), 43.
/3/ E. C. Freeman, W. Paul, Phys. Rev. B 20 (1979) 716.
/4/ A. Deneuville, M.H. Brodsky, J. Appl. Phys. 50 (1979), 1414.
/5/ A. Deneuville, E. Gheeraert, S. Belkouch, unpublished.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

VERY LOW RESISTIVITY AuMn GATE OHMIC CONTACTS FOR GaInAs DIFFUSED JFETs

P.E. HALLALI, P. BLANCONNIER, L. BRICARD and J.C. RENAUD

Centre National d'Etudes des Télécommunications, Laboratoire de Bagneux, 196, Av. Henri Ravera, F-92220 Bagneux, France

RESUME - Tant pour les transistors à effet de champ à jonction que pour les transistors bipolaires à hétérojonction de la famille GaInAs/InP, l'obtention de contacts ohmiques de type P à faible résistivité est une étape particulièrement critique en raison de la forte hauteur de barrière Schottky sur ces matériaux. La réalisation d'un surdopage p^+ par diffusion de Zn en boîte semi-fermée ainsi que l'utilisation de l'alliage MnAu ont permis de résoudre ces problèmes : une résistivité de contact aussi faible que 10^{-7} Ω cm^2 a en effet pu être obtenue.

ABSTRACT - For GaInAs/InP junction field effect transistors as well as heterojunction bipolar transistors, the achievement of very low resistivity P type ohmic contact is a very critical step because the Schottky barrier height on these materials is quite high. The realization of a highly doped P^+ layer by Zn diffusion in a semi-closed box and the use of MnAu alloy contact have allowed to solve these difficulties : in fact, a contact resistivity as low as 10^{-7} Ω cm^2 has been obtained.

1 - INTRODUCTION

GaInAs and AlGaInAs materials, lattice matched to InP, are currently of first importance for optoelectronic devices in the 1.3 - 1,55 µm wavelength range /1/. Furthermore, the combination of a high electron mobility and good quality heterojunctions predestinate them for high frequency systems /2/. To take advantage of these potentialities, it is essential to minimize all the parasitic RC components and so to reduce the specific contact resistances. Thus, for instance, microwave power and noise behaviour of field effect transistors (FETs) are strongly dependent on the gate resistance ; for junction FETs, this is particularly critical : besides the influence of the gate metal resistance (as for MESFETs), the contact resistance and the spreading resistance of the gate layer are very important. The situation is equivalent for heterojunction bipolar transistors (HBTs) for which the base resistance is a key point.

For $(GaInAs)_{1-z}(AlInAs)_z$, a Fermi level pinning has been found at about (0.5+0.3z) eV above the valence band, due to surface states of the metal semiconductor interface. This implies barrier heights on GaInAs close to 0.2 eV for Φ_{Bn} and 0.55 eV for Φ_{Bp}. Consequently, while extremely low specific contact resistances are achievable on n (Al)GaInAs, good ohmic contact on p (Al)GaInAs are very difficult to obtain.

The AuZn alloy is currently used to contact p-type Ga(Al)InAs /3/. However, Zn outdiffusion during the alloying step generates an increase of the metal resistance. Besides, for JFETs as well as HBTs, the contact must be shallow (< 1000 A) in order to maintain the quality of the underlying pn junction. The non alloyed TiAu metal is also very often used /4/ but the best resistivity obtained with this metal (≈2.10^{-6} Ω cm^2) is not a quite satisfactory value for high performances microwave transistors.

This paper presents the technique we have developed to obtain reproducible low resistivity p type ohmic contacts : it is based on the use of MnAu alloyed on a Zn diffused (Al)GaInAs layer. The performance of the contact will be described in terms of the d.c. contact resistance and two devices (JFETs and HBTs) using this process will be characterized.

2 - EXPERIMENTAL PROCEDURE

The $(GaInAs)_{1-z}(AlInAs)_z$ layers used in this work have been grown by Molecular Beam Epitaxy on semi-insulating InP /5/. The samples provided for evaluating the contact

resistivity by the transmission line model (TLM) /6/, are non intentionally doped. The JFET and HBT structures depicted in figure 1 are detailed elsewhere /7,8/. In all cases, the process can be divided in three main parts :

a) Zn diffusion : it occurs in a semi-closed quartz box /9,10/, with $ZnAs_2$ as diffusion source under a flow of purified H_2. The diffusion temperature is 500°C for GaInAs and 480°C for AlGaInAs.

b) 4 % Mn 96 % Au deposition : the metals are deposited using electron beam evaporation with a background pressure of 2.10^{-7} Torr. Deposition rates of 5 Å/s for Mn and 10 Å/s for Au are regulated by a controller. The Mn is deposited first. The proportion of Mn optimized for GaAs /11/ has been kept for (Al)GaInAs.

c) Annealing : the samples are annealed in a conventional furnace under a flow of $Ar-H_2$ gas. They are heated till they reach the chosen temperature and then immediately transfered in the cold part of the furnace. With such an annealing technique, a temperature of 300°C is reached with a thermal time constant of 25 s.

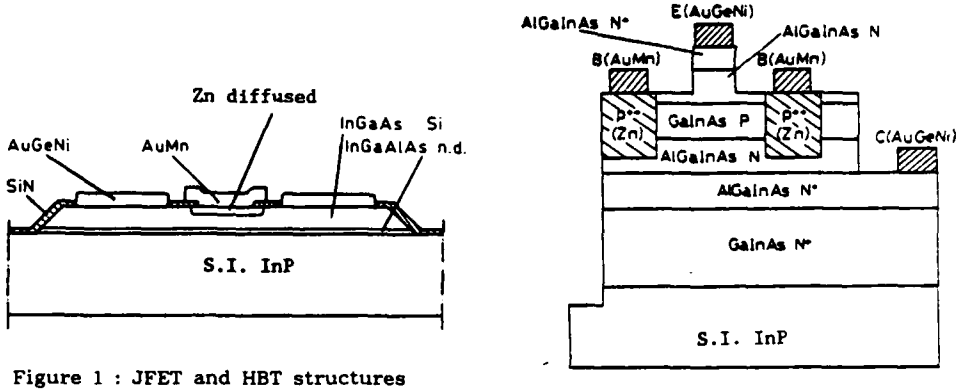


Figure 1 : JFET and HBT structures

3 - RESULTS AND DISCUSSION

The Zn diffusion has been investigated for different compositions of $(GaInAs)_{1-z}(AlInAs)_z$. Figure 2a illustrates the Zn diffusion speed for z = 0 and z = 0.4 ; a high surface concentration is obtained : 4.10^{19} cm^{-3} for GaInAs (figure 2b) which is essential for reducing the p contact resistivity. Besides, the reproducibility of the technique and the abrupt junction profile give a very good control of diffused depths. Finally, the high quality of these p-n junctions has allowed to fabricate GaInAs JFETs with low gate leakage currents (I_g < 25 nA at V_g = - 5 V), showing thus the under-layer is not damaged during the diffusion process.

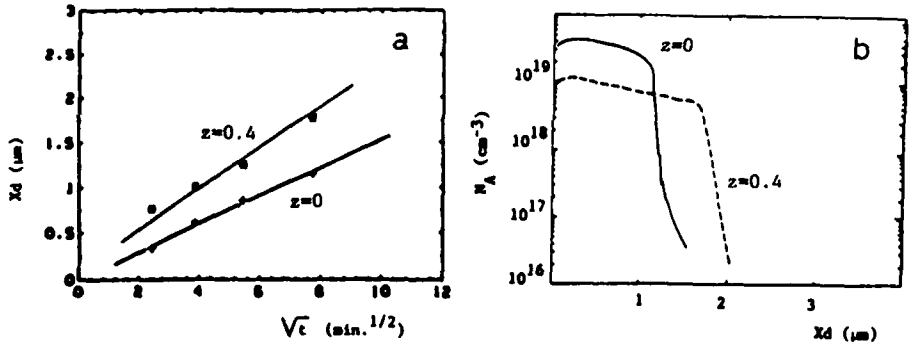


Figure 2 : a) Zn diffusion speed for z=0 and z=0.4
b) Zn Electrical profile (POLARON)

Several experiments with TLM stripes have been performed for determining the minimum of contact resistivity as a function of the annealing temperature (Fig. 3). For as deposited contacts, TLM evaluation gives specific resistances close to $10^{-5}\ \Omega\,cm^2$ for GaInAs and $10^{-4}\ \Omega\,cm^2$ for $(GaInAs)_{0.6}(AlInAs)_{0.4}$. The best contact resistivities R_C are obtained after an annealing at 290°C ; R_C is lower than $10^{-7}\ \Omega\,cm^2$ for GaInAs and $R_C = 10^{-6}\ \Omega\,cm^2$ for z = 0.4. These values, which are particularly reproducible, are the lowest ever reported on these types of materials.

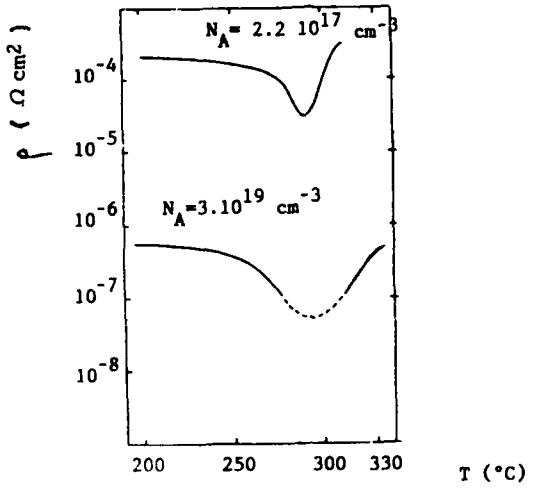


Figure 3 : Variation of the specific contact resistivity with the Zn doping level

Figure 3 illustrates also the interest to "over-dope" p type layers by a Zn diffusion : with a p doping of $2.10^{17}\ cm^{-3}$, the contact resistivity is not better than $3.10^{-5}\ \Omega\,cm^2$. The variation of R_C with the doping level is shown in Figure 4a. Such a behaviour can be expected referring to two main effects (figure 4b):

- the high doping level due to the diffusion process reduces the barrier width and makes easier the conduction by tunneling between the metal and the semiconductor.

- the shallow diffusion of Mn (≤ 1000 Å) reduces again the barrier width.

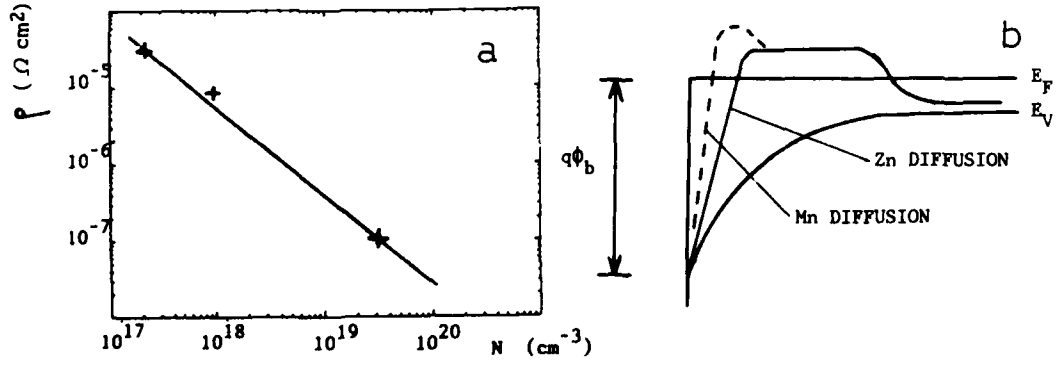


Figure 4 : a) variation of the contact resistivity with the Zn concentration
b) Schematic valence band diagram at the equilibrium, after Zn diffusion and after MnAu alloying

These two characteristics have been exploited for the processing of JFETs and HBTs. Even if the geometries of the devices are quite large : gate length of 2.5 µm for the JFET and emitter width of 10 µm for the HBT, their frequency behaviour is excellent : a maximum oscillation frequency of 14 GHz is reached for the JFET (which is the best value reported for diffused JFETs) and a transition frequency of 1.7 GHz for the HBT /12/.

4 - CONCLUSION

MnAu alloyed contacts with excellent ohmic characteristics have been fabricated on p^+ diffused (Al)GaInAs ; the technique presented here is an efficient method for obtaining very low contact resistivities required for high performances JFETs and HBTs as well as lasers and photodiodes. The lowest contact resistivity, $\leq 10^{-7}\ \Omega\ cm^2$, is found to occur with an annealing temperature near 290°C. This resistance value is the lowest ever reported for contact to p-type GaInAs.

REFERENCES:

/1/ J.C. Renaud, F. Lugiez, S. Vuye, M. Allovon, L. Nguyen, A.Scavennec, B. Bourdon
ECOC 1988 - Brighton

/2/ A. Scavennec
ESSDERC 1988 - Montpellier - Paper 2 IP1 - This issue

/3/ H. Kräulte, E. Woelk, J. Selders, H. Beneking
IEEE Transactions on Electron Devices, Vol ED-32 N°6 June 1985

/4/ R. Kaumanns, N. Grote, H.G. Bach, F. Fidorra
GaAs and Related Compounds 1987 - Heraklion

/5/ J.P. Praseuth, M. Quillec, M. Allovon, M.C. Joncour, J.M. Gérard,P. Hénoc
GaAs and Related Compounds 1987 - Heraklion

/6/ H.H. Berger
Solid-State Electronics 15, 145, 1972

/7/ L. Nguyen, M. Allovon, P. Blanconnier, E. Caquot, A. Scavennec,B. Bourdon
ESSDERC 1987 - Bologna pp 943

/8/ J.L. Pelouard, P. Hesto, J.P. Praseuth, L. Goldstein
IEEE Electron Device Letters, V. EDL 7, N° 7, Sept. 1986 pp 516

/9/ K. Kazmiersky, A.M. Huber, G. Morillot, B. De Crémoux
Jap. J. Appl. Phys 23 (1984) p 628

/10/ P.E. Hallali, P. Blanconnier, J.P. Praseuth
To be published in Journal of the Electrochemical Society

/11/ C. Dubon-Chevallier, M. Gauneau, J.F. Bresse, A. Izrael, D. Ankri
Journal of Applied Physics - V59 - 1 June 1986 - pp 3783

/12/ J.L. Pelouard, J.P. Praseuth, P.E. Hallali, G. Tremblay
Unpublished

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

DEFECT CHARACTERIZATION OF Si^+-IMPLANTED GaAs BY MONOENERGETIC POSITRON BEAM TECHNIQUE

J.-L. LEE, K.-H. SHIM, S. TANIGAWA*, A. UEDONO*, J.S. KIM, H.M. PARK and D.S. MA

Compound Semiconductor Department, Electronics and Telecommunications Research Institute, PO Box 8, Daedok Science Town, Choongnam, Korea
**Institute of Materials Science, University of Tsukuba, Sakura-mura, Ibaraki 305, Japan*

Abstract - Monoenergetic positrons with variable energies were used to study the depth distribution of implantation-induced vacancy-type defects in undoped GaAs and p-type Si. In B^+- and As^+- implanted Si substrates, parabolic-type distributions of vacancy-type defects were observed. In Si^+-implanted GaAs, the concentration of vacancy-type defect decreased continuously with increasing depth below the surface. The distribution of defects changed into parabolic-type in annealing the Si^+-implanted GaAs above the temperature of 900 °C.

1-INTRODUCTION

Processing of semiconductors often involves the use of ion implantation as a means of introducing dopant impurities. After implantation the semiconductor lattice will not only be damaged, but in addition, will be left a nonuniform lattice displacement /1/. Although a lot of researches have been theoretically proceeded to understand the mechanism of defect production during the ion-implantation into semiconductor/1,2/, their validities have not yet been confirmed experimentally because of the lack of suitable techniques. Especially little is known on the behavior of electron traps in the implanted region in activating the ion-implanted GaAs.

In the present work, we report the application of the monoenergetic positron beam technique for the study of vacancy-type defects in both B^+-and As^+-implanted Si and Si^+-implanted GaAs. The depth distributions of vacancy-defects below the surface were investigated in activating the Si^+-implanted GaAs by a two-step annealing technique in which a high-temperature main anneal step is followed by a second anneal at a low temperature /3/. The results were discussed in conjunction with those obtained by Hall measurements and deep level transient spectroscopy (DLTS).

2-EXPERIMENTAL METHODS

Undoped semi-insulating liquid encapsulated Czochralski (100) GaAs samples were implanted with the doses of 4.6×10^{13}, 3×10^{13}, and 3×10^{13} Si^+ cm^{-2} for the energies of 20, 30 and 120 keV, respectively. To compare the defect generation behavior in elemental semiconductors with that in compound ones, implantations with 80 keV B^+ and 150 keV As^+ were done on p-type (100) Si wafers, respectively. In order to examine the carrier generation, GaAs samples implanted with the dose of 3×10^{12} Si^+ . cm^{-2} at an energy of 120 keV were used. The implanted specimens were protected by a 2000 Å silox encapsulating layer to minimize As loss from the surface and then annealed by a two-step rapid thermal annealing (RTA). First stage annealing was done at 900 and 920 °C for a period of 5 s under flowing ultra-pure nitrogen gas. RTA temperatures for second stage annealing were 800 and 850 °C for periods of 10-30 s to examine the improvement of electrical properties by two-step annealing. After chemically removing the silox encapsulating layer, the wafers were characterized by mono-energetic positron beam measurements, Hall measurements and DLTS.

Monoenergetic positron beam technique enables to study an atomic scale disorder at the surface and in the sub-surface region/4-6/. The implanted positron eventually annihilates with an electron, producing two annihilation photons. A Doppler broadened profile of these annihilation photons can provide the momentum distribution of the annihilating electrons. A positron is repelled from positively charged ion cores by a coulomb interaction. In a perfect metal the

Doppler broadened energy spectrum consists of both a narrow parabolic part corresponding to the annihilation with conduction electrons and a broad one with core electrons. In a metal containing a high concentration of vacancy-defects positron annihilates selectively at vacancies of open-volume defects. As a result, the contribution of core electrons in the Doppler broadened energy spectrum is reduced and, in turn, that of conduction electrons is increased. The lineshape parameter S which is defined as the ratio of the central region to the total counts of the Doppler broadened energy spectrum shows the annihilation characteristics. From the change of the S parameter, one can derive the information about the fraction of positrons which annihilate in such defects and the concentration of defects through the positron trapping model /6/. Once we are able to get slow positrons, the acceleration of them with a desired energy makes it possible to adjust the implantation profile of positrons to restricted regions of interest in the specimen under study.

The present experiments were performed with a variable-energy positron beam line in ultra-high vacuum /5/. Positrons emitted from a ^{22}Na source are moderated by a venetian-type vane which is made of well annealed tungsten foils with a thickness of 25.4 μm. Implanted positrons tend to diffuse back to the surface, because of the negative work function of a positron for metals and/or the presence of a deep potential well at the surface image potential /7/. Positrons thermalizing within a diffusion length of the moderator surface can reach the surface and then they are emitted into vacuum. Slow positrons emitted from the moderator are guided by nine separated magnetic coils of 100 G with a diameter of 700 mm. Acceleration of slow positrons is performed by applying a high desired voltage between 0.1 ~ 50 kV to the source chamber. The source chamber is floated by a ceramic break.

The Doppler broadening profiles of annihilation radiations were measured by a high purity Ge detector with an energy resolution of 1.1 keV in FWHM for the 512 keV γ rays. Annihilation spectra with total counts of 5×10^6 were taken at various incident positron energies. The central region of the γ spectrum was defined from 510.5 to 511.5 keV. The incident energy of the positron beam was adjusted from 100 eV to 25 keV. All measurements were performed at room temperature and the vacuum in the present experiments was 8×10^{-8} Torr.

3-RESULTS AND DISCUSSIONS

Figure 1 shows the variation of the S parameter as a function of incident positron energy E for the three Si^+-implanted GaAs specimens. The relation between the positron incident energy E and the mean positron implantation depth α is given by/4/

$$\alpha\ (\text{Å}) = \frac{347}{\rho}\ E^{1.6}\quad (\text{keV}) \qquad (1)$$

where ρ is the density of the specimen. The values of α are shown in a horizontal axis together with the positron indicent energy in Fig.1. The value of S decreases continuously with increasing depth below the surface. Tanigawa et al/8/ reported that any detectable doping effect could not be found in Si by the angular correlation measurements. They considered that no trapping effect is due to the weak trapping power of impurities, that is, an interstitial-type defect does not have such a trapping potential for a positron. Thus the difference of the S parameters between the implanted and the unimplanted GaAs specimens is attributed to the introduction of vacancy-type defects by the ion implantation. The damaged region shifts toward inside of the specimen with applying the high energy of Si^+ to GaAs.

Gibbons et al/1/ theoretically suggested the generation of high concentration of vacancies by the recoils of Ga and As from near-surface region. They observed the stacking faults and the microtwins surrounded by dislocation networks extending the surface in the cross-section of Si^+-implanted GaAs /9/. It is generally accepted that V_{Ga} is of acceptor type and can exist in negatively charged or neutral states. On the contrary V_{As} is of donor type and its charge state can be negatively, neutral or positively charged depending on the position of the Fermi level /10/. It has been, however, reported that positron cannot be trapped at V_{As}. Therefore it is apparent that the S parameter change as a function of depth is related to the distribution of Ga vacancy and vacancy agglomerates.

The S parameters as a function of incident positron energies for B^+- and As^+-implanted Si specimens, respectively, are shown in Fig.2. The defects for both types of specimens are found to be distributed in a parabolic form, which are clearly compared with the defect profiles in GaAs, given in Fig.1. This implies that the vacancy generation mechanism for ion implants into GaAs is different from that into Si. According to the calculation given by Gibbons et al /1/, vacancies are produced from the origin points of high energy recoils. Consequently the concentration of implantation-induced vacancy decreases logarithmically from the surface to the inside of bulk which is clearly consistent with the present experimental results in Si^+-implanted GaAs. They suggested the distribution of atomic displacement with a parabolic shape which is originated from the final stopping points of the recoils /1/, namely low energy recoils. The distribution shape, resulting from the low energy recoils, agrees with that of the defect profiles in B^+- and As^+-implanted Si, which suggests the low energy recoils might dominate the vacancy generation in ion

implants into Si. However the reason why different mechanisms for the production of vacancy-like defects during ion-implantation are working in Si and GaAs remains open.

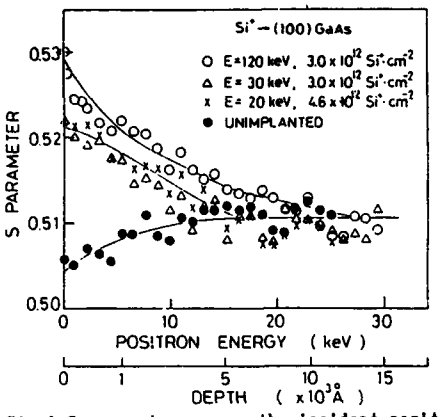


Fig.1 S parameter versus the incident positron energy and/or the depth for Si^+-implanted GaAs samples with different implantation energies and dose amounts.

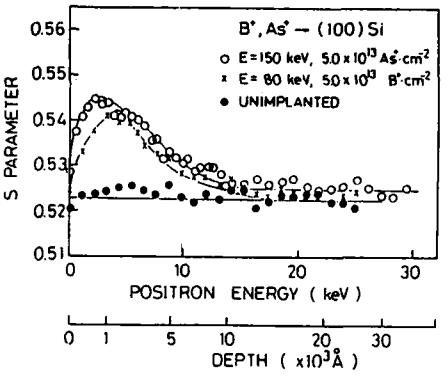


Fig.2 S parameter as a function of the incident positron energy and/or the depth for B^+- and As^+-implanted Si samples.

Variation of S parameters versus incident positron energy in samples activated by the two-step annealing process is illustrated in Fig.3. The damage profiles after RTA are parabolic-type distributions, in which defect concentration is not dependent upon the duration time of the second-stage anneal. Ralston et al/11/ suggested that the distribution of dislocation loops as a function of depth from the surface of ion-implanted GaAs was a parabolic-type, which is consistent with that indicated by slow positron measurement (see Figure 3). Potentially great interest in Fig.3 is the lack of recovery of Ga-related damage induced by ion implantation during RTA.

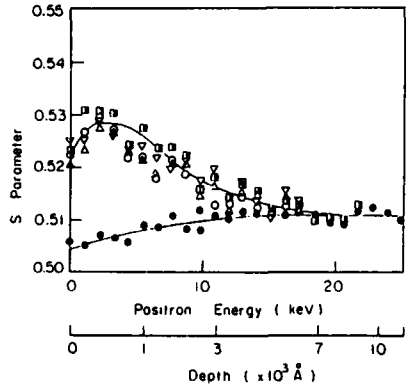


Fig.3 S parameter versus the incident positron energy and/or depth in Si^+-implanted undoped LEC GaAs activated by a two-step rapid thermal annealing. The conditi ns of the two-step RTA are 900 °C/5 s + 850 °C/0 s (□); 900 °C/5 s + 850 °C/10 s (▲); 900 °C/5 s + 850 °C/20 s (▼); 900 °C/5 s + 850 °C/30 s (○). Experimental results for unimplanted GaAs (●) is also included.

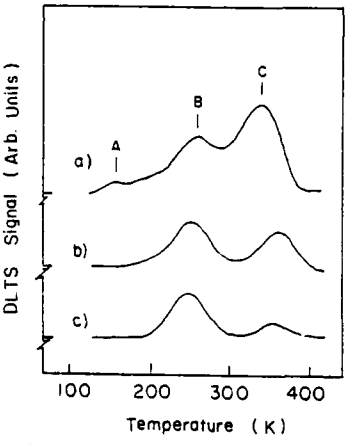


Fig.4 DLTS spectra of electron traps in Si^+-implanted undoped LEC GaAs activated by a two-step prapid thermal annealing at (a) 920 °C /5 s; (b) 920 °C/5 s + 800 °C/10 s; (c) 920 °C /5 s + 800 °C/20 s.

DLTS spectra of Si^+-implanted GaAs activated by a second stage RTA for different times are shown in Fig.4. In the first stage RTA, three distinct DLTS peaks are observed, which are denoted as A, B and C, respectively. The trap levels of the three peaks are 0.26, 0.48 and 0.64 eV, respectively. After the second-stage RTA, peak A diminishes. The signal C decreases with increasing the second-stage anneal time whereas the height of signal B is unchanged. The signal C with an activation energy of 0.64 eV might be related to EL3 (0.64 eV) or EN2 (0.63 eV) which were observed in SiO_2 capped GaAs by Kuzuhara et al /12/. Comparing DLTS results with depth profiles of vacancy-type defects given in Fig.3, it is believed that signal B with an activation energy of 0.48 eV is related to the Ga-related vacancy-type defects. A lot of DLTS works/13/ have

suggested that the trap level at 0.53 - 0.57 eV corresponds to implantation induced damage, which is consistent with the present results.

The results in Fig.5, obtained from Hall measurements, indicate that two-step RTA gives better activation efficiency than one-step RTA, which imply the enhancement of carrier generation with increasing the second-stage anneal time, although the vacancy-type defects produced by ion-implantation are not recovered regardless of the duration of second-stage anneal time, given in Fig.3. This implies the activation of Si^+-implanted GaAs proceeds from the exchange of interstitial Si (Si_I) with substitutional Ga (Ga_{Ga}) given in eq.(2), rather than the trapping of Si_I into Ga vacancy.

$$Ga_{Ga} + Si_I \rightarrow Si^+_{Ga} + Ga_I + e^- \qquad (2)$$

However the appropriateness of this mechanism in interpreting the activation phenomena of Si^+-implanted GaAs remains open.

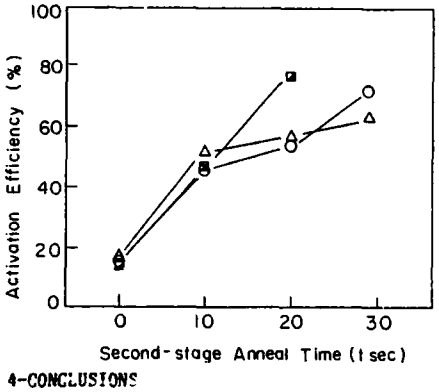


Fig.5 Dependence of activation efficiencies in Si^+-implanted undoped LEC GaAs on the two-step rapid thermal annealing at 900 °C/5 s + 850 °C /t s (△); 900 °C/3 s + 850 °C/t s (○); 920 °C/5 s + 800 °C/t s (□).

4-CONCLUSIONS

Slow positron beam technique was used to determine the distribution of vacancy-like defects in Si^+-implanted GaAs and B^+-and As^+-implanted Si. In B^+- and As^+-implanted Si substrates, the parabolic-type distributions for vacancy-type defects were observed. This implies the production of vacancy-like defects is dominated by low energy recoiling process in ion implants into elemental semiconductor. In compound semiconductor such as GaAs, the concentration of vacancy-type defects decreased continuously with increasing depth below the surface, which was due to the vacancy generation by high energy recoiling process. The continuous decrease of implantation-induced defects changed into the parabolic-type distribution in the depth below the surface by annealing the Si^+-implanted GaAs. The concentration of defects is not changed by increasing the second-stage anneal time whereas the activation properties detected by Hall measurements have been improved.

ACKNOWLEDGEMENT

This work was financially supported by Korea Telecommunications Authority. This work was also supported in part by a Grant-in Aid for Scientific Research of the Japanese Ministry of Education, Science and Culture, and Shimazu Science Foundation.

REFERENCES

/1/ Christel L.A. and Gibbons J.F., J. Appl. Phys. 52 (1981) 5050.
/2/ Brice D.K., J. Appl. Phys. 46 (1975) 3385.
/3/ Seo K.S., Dhar S. and Bhattacharya P.K., Appl. Phys. Lett., 47 (1985) 50.
/4/ Mills Jr. A.P. and Wilson R.J., Phys. Rev. A26 (1982) 490.
/5/ Tanigawa S., Iwase Y., Uedono A. and Sakairi H., J. Nucl. Mater. 133 & 134 (1985) 463.
/6/ Lynn K.G., Chen D.M., Nielsen B., Pareja R. and Myers S., Phys. Rev. B34 (1986) 144.
/7/ Mills Jr. A.P. in : Positron Annihilation, eds., Coleman P.G. (North-Holland, Amsterdam, (1982) P.142.
/8/ Tanigawa S., Hinode K., Nagai R. and Doyama M., Appl. Phys. 18 (1979) 81.
/9/ Opyd W.G., Gibbons J.F., Bravman J.C. and Parker M.K, Appl. Phys. Lett. 49 (1986) 974.
/10/ Hautojarvi P., Moser P., Stucky M., Corbel C. and Plazaola F., Appl. Phys. Lett. 48 (1986) 609.
/11/ Ralston J., Wicks G.W., Eastman L.F., Decooman B.C. and Carter C.B., J. Appl. Phys. 9 (1986) 120.
/12/ Kuzuhara M. and Nozaki T., J. Appl. Phys. 55 (1986) 3131.
/13/ Sriran S. and Das M.B, IEEE Elec. Dev. ED-30 (1983) 586.

SESSION 5

Session 5IP Room A : Invited papers

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

LASER SCANNING TOMOGRAPHY : A NON DESTRUCTIVE QUALIFICATION TEST FOR SEMICONDUCTORS

J.P. FILLARD

Laboratoire LINCS, Centre d'Electronique de Montpellier (CNRS-UA 391), USTL, F-34060 Montpellier, France

RESUME - Parmi les différentes méthodes de caractérisation des défauts dans les wafers de semiconducteur, les techniques par infra rouge doivent être remarquées pour leur caractère parfaitement non destructif. Le développement récent de la tomographie laser à balayage permet d'obtenir des images de haute résolution des dislocations décorées et des microprécipités. L'exploration du matériau dans les trois dimensions est réalisée avec une grande précision jusque dans le voisinage immédiat des zones actives des composants. On étudie l'observation des micro défauts et particulièrement les substrats GaAs de différentes natures. Des images à l'échelle microscopique révèlent des précipités de taille bien inférieure à la limite de diffraction optique. D'autres résultats ont aussi été obtenus sur InP, CdTe ou Si ; la plus part des semiconducteurs classiques peuvent être analysés par tomographie laser à balayage ce qui fait de cette méthode un outil précieux d'investigation.

ABSTRACT - Among the various physical approaches of the semiconductor wafer defect characterization a special attention is to be payed to the infra red imaging techniques which are prefectly non destructive . Recent developement of the Laser Scanning Tomography will be emphasized because it gives highly resolved images of decorated dislocations and microprecipitates . Three dimensional exploration of the material is allowed with a large accuracy and possibly close to the active regions of the circuits. Observation of micro defects are reviewed , especially concerning GaAs materials of various origins and constitutions. Microscale images enable us to reveal very small scatterers even if their size is much smaller than the optical diffraction limit. Results obtained on other semi conductors such as InP, CdTe or Si will also be reviewed ; most of the conventional semiconductors are relevant to this powerful technique which is a unique means of investigation.

1 - INTRODUCTION

It is known that compound semiconductors (such as GaAs for instance) suffer a large diversity of crystallographic defects induced by the delicate interface equilibrium during the growth process and also by the post growth cooling transient. These defects diversely affect the behavior of the elaborated circuits and leads to a scattering in the specifications /1,2,3/ .

In order to control such flaws and also to undertake a feedback on the growth technology , several inspection techniques were suggested which are more or less destructive of the surface of the wafers; also ingot or wafer annealing is often intentionally used to improve the uniformity of the technological process /4/.

Twelve years ago it was suggested by Tajima /5/ that light scattering ultra microscopy (Tyndall effect) could be a powerful tool for investigating defects and especially metallurgical faults in GaP crystals . This technique was afterwards largely improved by Ogawa and coworkers /6,7,8,9/ who invented a revamped experimental set-up incorporating new devices like the laser and the computer . They used it to explore quartz materials as well as GaAs , InP , CdTe etc...

This technique was called Laser Scanning Tomography (LST) because a focused laser beam is introduced in the sample and linearly moved in order to generate a virtual plane of illumination . The corresponding light beam image is collected by a camera having a line of sight at 90° to the beam direction, the successive images being assembled by a dedicated computer . This method is attractive because it gives three dimensional capability and is non destructive. The performances in contrast

sensitivity and spatial resolution are especially high. To our best knowledge we are presently the only laboratory outside Japan that is equipped and experienced for such analysis.

From 1985 up until now , several studies were reported /10,11,12,13/ from several laboratories , they all refer to large scale (even wafer scale) images on GaAs materials; they provide us with detailed and complex distributions of bright defects the origin of which is questionable even if a satisfying general correspondence is found with other classical investigations such as surface chemical etching SCE /14/ , X ray topography XRT /12/ , infra red transmission IRT /15/ or cathodo/photo luminescence C/PL /16/ .

We have used the LST method (see Ogawa /17/) with the essential improvement of the optical configuration in order to observe sub-micron scale particles: higher optical magnification (typically X 40 or larger) of the camera leads to a field of observation as small as 220x220 mm (frame memory 512x512 pixels) at a maximum distance of 5 mm from the laser beam input face. The tomographical "plane" thickness was also improved by a sharper focusing of the laser beam ; the minimum thickness achieved is 10 µm and this plane is adjusted with the focus plane of the camera objective (the thickness of which is also 10 µm). The wafer surface can be approached to within 10 µm without introducing major perturbations in the image. In that configuration of reduced field, the scanning of the laser beam can be introduced simply using a vibrating mirror thus providing a "live" image on the TV monitor. A more reduced scale of analysis (down to the diffraction optical limits) allows us to discover the ultimate objects giving rise to the scattering in GaAs bulk material : they are all reducible to small "droplets" of condensed foreign atoms in the matrix .

These aggregates are found either along the dislocation path as decoration precipitates (DP) or in the dislocation free zones as clouds of randomly distributed microprecipitates (MP) : these two kinds of defects are isotropic Rayleigh type /18,19/ scatterers ; they are the only objects revealed by LST , no evidence is found for a contribution of the EL2 centers /20,21/ in the scattering mechanism /22,23,24/ .

The presence of microprecipitates in LEC grown GaAs has been also revealed by several methods (see Stirland /25/). For precipitates much smaller than 1 µm laser scanning tomography (LST) is of special interest owing to its ability to detect individual scatterers which are smaller than the diffraction limit; it is also well suited for a 3 dimensional visualisation of these defects. We studied by LST technique different types of SI LEC GaAs materials, namely undoped, Cr or In doped, unannealed or ingot annealed.

In this review communication we intend to comment the present state of the art in this domain of defect imaging especially dedicated to semiconductor wafer inspection. Of course GaAs is not the only material amenable to such investigation methods : InP , GaP , CdTe or even Si itself give rise to unexpected features of hidden flaws .

2 - EXPERIMENTAL

The Laser Scanning Tomography (LST) experiment /17,26/ is sketched in Fig. 1- : a focussed infra red laser beam is introduced in the wafer through a flat section and parallel to the faces ; the light scattered by the internal flaws along the beam path is collected by a TV camera and the image of this "optical window" is sent to a computer which selects the central TV lines ("TV window") ; then the sample is moved by stepper motors ("step window") and the total image of the virtual scanning plane is reconstructed and displayed as sketched by Ogawa /17/ in Fig. 2- . Three dimensional information can be achieved by simply changing the position of the tomographical plane .

This method allows a large amplification of the contrast of the image , it is perfectly non destructive and does not require any vacuum or low temperature facility . A macroscopic scale image takes a matter of minutes whereas microscopic images 2x2 mm are live TV images . It still remains that the effective thickness of the observation plane is determined by the diametre of the laser spot on the sample (typically some 50 µm) .

It was shown /10/ that the scattered intensity obeys a $1/\lambda^4$ dependence which is typical of a Rayleigh isotropic scattering . The very low "absorption" coefficient a of semi insulating GaAs ($2 < a < 0.5$ cm-1) in the $1 > \lambda$ µm range enables penetration in the wafer to large distances (some cm) which makes it possible to inspect 2" wafer, using a YAG laser (1.06 µm) . Reducing the laser beam diameter down to 10^{1} µm allows generation of very thin tomographical planes. Lateral resolution for detecting the presence of a scatterer is very high provided that the scattering yield is effective: particles as small as 20 Å (macromolecules) were actually revealed by ultra microscopy in colloidal solutions /27/ .

This perfectly non destructive and accurate method was essentially used at a macroscopic scale ; it revealed very detailed cell structures in undoped GaAs and also dislocation clusters in GaAs-In. the comparison of these LST images with other images is quite satisfactory. keeping in mind that the referred volume is of limited thickness and adjustable in the bulk whereas other methods refer to the region close to the surface (EBM , C/PL , SCE and reflexion XRT) or to the whole thickness (transmission

XRT , TEM or IRT). The origin of the scattering was initially attributed to refractive index variations /28/ or to EL2 concentrations /29/. The present work definitively confirms with high resolution images that scatterers are microprecipitates condensed on dislocations or distributed in the volume as clouds of particles . As stated by Skolnick the deep center EL2 does not contribute in the scattering even if a correlation is found with a through a power 4 dependence /10/ .

If the beam path is kept very long in order to get cm scale images then absorption and light losses usually introduce a progressive reduction of the laser beam intensity along the distance and the LST image becomes unbalanced : this disadjustment can be compensated by a computer correction on the final display .

If a microscopical investigation is considered, the image field will be smaller : 250 µm would correspond to a lateral resolution 0.5 µm / pixel which could be considered as a satisfying outer limit . This is achieved with a microscope objective X40 on the camera. The immediate consequence is a camera focus depth some 10 µm wide. Of course the beam diameter has to match this dimension in order to remove "ghosts" of remote scatterers from the image. Such a reduction involves a more convergent laser beam and strongly limits the distance of maximum exploration see Fig.3- . Varying the beam focus and the camera objectives thus allows us to adjust the observation requirements between limits which become tighter as the magnification is higher. These optical constraints are in some way compensated by important advantages specific of these microscopical investigations :

i) the beam energy concentration is enhanced thus allowing us to reveal smaller scatterers
ii) the limited range of lateral exploration makes it possible to use vibrating beam or flat beam /8/ techniques instead of the stepper motors and a live image is obtained with remanent photocathod camera (N214) .
iii) the limited range of exploration corresponds to weak attenuation and no computer corrections are required .
iv) the reduced depth of focus allows us to locate the scatterers in depth with increased precision , in the micron range .

These new specifications of the experimental set up lead us to expect a more precise LST probing of interfaces and micro structures of epilayers .

3 - RESULTS

The origin of the scattering giving rise to the LST image in GaAs was related to condensed atoms on dislocations . It is likely that excess As atoms are concerned /30/ , especially in materials grown under As rich melt conditions ; these atoms were shown to precipitate and form small droplets some micron wide distributed along the path of the dislocations .

In Fig.4- we report investigations on a conventional LEC undoped material; the wafer comes from the central part of the ingot and the classical cell structure is observed at a macroscale (Fig.4a-) using a large laser beam (150 µm). Fig.4b-c- represent successive enlargements of windows indicated by white boxes; each view corresponds to an adapted thinner LST plane corresponding to a sharper focusing of the laser beam. Details of the walls become more precise to the point that small individual precipitates are resolved; they are crowded in the wall zone and arranged along curved lines entangled in the 3D space . Varying the vertical position of tomography allows verification that they effectively are point objects and not vertical string or rods .

The situation could be similar (Fig.5-) in InP crystals /31/,32/ but with a wider separation between the droplets: this should induce the illusion of non related defects. The density of defects strongly varies from sample to sample and in some cases inside the wafer itself. It is likely that the condensation mechanism of precipitates is very similar in InP and GaAs .

It is known that In doped GaAs materials show a reduced EPD density ; the corresponding LST images effectively reveal /22/ isolated ‹110› dislocations as shown in Fig.6- ; their correspondance with etch pits and also with X ray topography images can be satisfactorily verified /33/. A uniform cloud of microprecipitates is distributed in the bulk whereas a denuded zone (DZ) is surrounding the dislocation to a distance of some 50-200 µm apart . A graphical illustration of this situation is given in Fig.6b- These MP are presumably particles smaller than DP because the individual brightness is much lower and the density higher.

The only previous observation which can be related to these MP was performed by photoetching experiments /34/ /35/ and corresponds to a "micro roughness" of the etched surface ; we did not find evidence for these particles in the literature devoted to TEM analysis.
The collective effect of MP in the LST scattered intensity was discovered by Osaka /35/ and Moryia /36/; the high magnification of the microtomography allows us to resolve individual MP thus demonstrating

Fig.6c- that "speckle" artefacts as suggested by Osaka /35/ are not concerned in such images. Clouds of MP were observed in every GaAs-In samples we analysed and their density in the free space was estimated /32,35/ in the range 2.5 10^9 cm-3 which is a quite large value; then it is worth supposing that the size of these particles is rather small (› 1000 Å).

In ingot annealed materials a large cell structure is developed but the inner space of the cells is filled with MP as shown in Fig.7- ; three dimensional evaluation of the clouds was easily achieved revealing a sphere like arrangements selectively distributed; the local density of particles is in the range 3.5 10^9cm-3 which means that they are very small. Also a careful examination of the walls reveals a cloud of another kind of MPs which recovers and embodies the walls: these MPs are not easily observed because they superpose over the bright scatterers and ghosts of the walls which saturate the camera . It is to be supposed that the various cell MP (CMP) and the wall MP (WMP) are presumably of a different nature because they are differently distributed with respect to the cell structure; they seem to be specific of the manufacturer's fabrication process and likely of the high temperature annealing .

From all of these observations it turns out that LST is the more adapted technique for the study of microprecipitates and it is likely that their electronic influence on the GaAs IC specification could have been underestimated. The situation holds the same for Si where SiO_2 precipitates are clearly observed and numbered even at very low density.

We recently reached the experimental evidence that high magnification allows to obtain a more localised information consistent with the analysis of epilayers or micro circuits; an example is given in fig.8- showing the LST image of defects located in a 10 µm range beneath a test circuit on GaAs material; a wealth of work is now required to correlate the electrical specifications of the circuits with the presence of neighbouring defects.

Most of the investigations performed in the field of LST are devoted to GaAs bulk materials because there is a key problem of improving crystal growth technology; nevertheless this LST technique is as well available for InP as stated above but also for other materials such as CdTe or II-VI compounds such as ZnS , Zn Se /17/.

Surprisingly very few investigations (Ogawa /17/) were performed on Silicon may be owing to the established opinion that it is a well achieved and controled material. Si IC technology actually requires a precise control of the Oxygen concentration and of the gettering of metal impurity atoms by the Silicon dioxide precipitates . Our recent work showed that these particles can be easily observed by LST as observed in Fig.9-. This obviously means that Silicon will be a very promising field of research with LST inspection in the imediate future.

DISCUSSION AND CONCLUSION

Infra red imaging techniques of defects in III-V compounds was proposed some years ago by Ogawa /17/ and it was established that they can be compared to the traditionnal inspection techniques /33/ such as EPD or XRT . An important advantage is that they are perfectly non destructive , easy and fast to implement ; such inspection means must contributes in the data bases of wafer parametres already installed in some industries involved in the III-V ICs.

The LST method used for GaAs materials provide us with very good images of defect distribution . These images were found /22/ to agree fairly well with the dislocation pattern and assumed EL2 distribution. They were initially related to EL2 through different possible mechanisms of scattering; it now seems more realistic to directly interpret the LST image as a Rayleigh type scattering on microprecipitates of different metallurgical origins diversely arranged in the bulk depending on the local thermal situations .To our present knowledge there is not any evidence for LST observation of pure dislocations but only indirect evidence of decorated dislocations due to precipitates .

LST tomographes are currently used in Japan as systems dedicated to the GaAs wafer inspection at the laboratory level but also as "on line" non destructive test as well as surface etching or XRT controls . Presently two industrial instruments are developped in Japan and a second generation system is under development in France.

LST inspection of wafers leads to detailled and highly contrasted images of decorated dislocations and microprecipitates which can be possibly extended to a 3D information. It allows a very accurate evaluation of defects at a macroscopic as well as microcopic scale; extension of the method is expected to investigate interfaces or epilayer structures . LST also allows to obtain 3 dimensional information of the defect organization by accumulation of parallel tomographic planes ; some years ago japanese workers made some 3D mocke-up /38/ which clearly reveal the "sponge" texture of the intricated cells of dislocations.

It is likely that this imaging technique will be eventually used on various kind of semiconductors as a non destructive means of investigating microprecipitates.

ACKNOWLEDGEMENTS
The author likes to acknowledge all financial contributors to the implementation of the LST facilities in Montpellier: MRES, DRET, DAII, EN, Thomson, PREMAT.

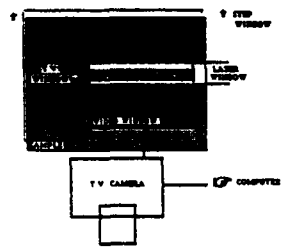
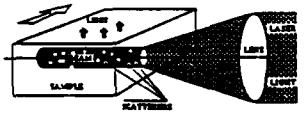
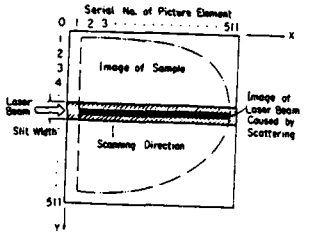


Figure 1- Organization of the LST experiment .





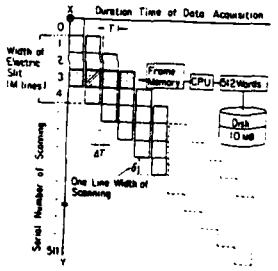


Image processing system of infrared light scattering tomography.
M: slit width of electric streak camera is chosen as 10 lines here, while $M=4$ in this figure for simplicity.
T: integration time at a given position of the laser beam, which is chosen as duration time of two TV picture frames and $N=2$ in this case.
dT: time of driving the laser beam by one scanning space and of storing the data integrated by the computer for 512 picture elements on one scanning line into the disk as the 512 words.

Figure 2- Image processing system of LST after Moriya /9/ .

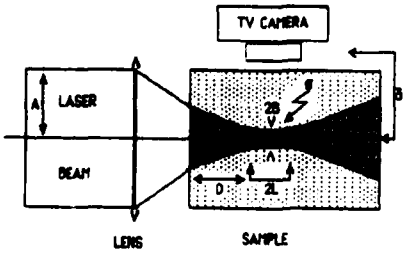


Figure 3- Beam focalisation in the sample .

-a- -b- -c-

Figure 4- LST images of defects in a GaAs conventional sample at different magnification (a - b - c).

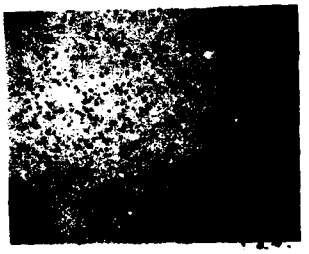


Figure 5- LST Tomography of InP material showing the distribution of microprecipitates

-a- -b- -c-

Figure 6- Indium doped GaAs : tomography of individual <001> dislocations(a) ; graphical illustration ; enlarged tomography of microprecipitates in the matrix (c) .

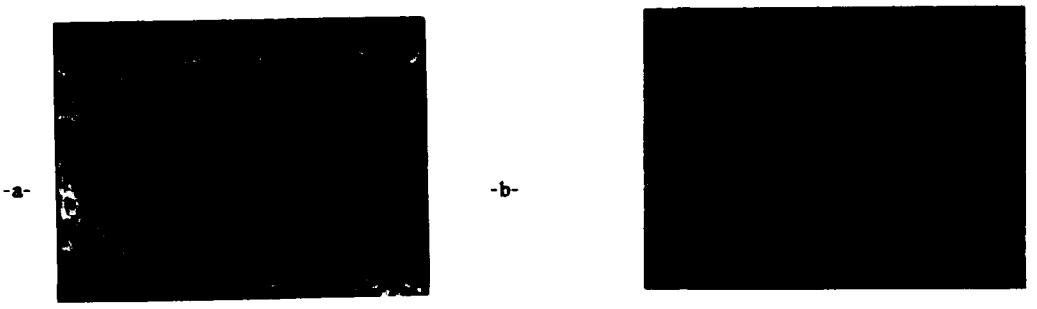


Figure 7- Ingot annealed GaAs (a) and enlarged tomography of microprecipitates located in the cells (b)

Figure 8- Defects in the vicinity of a GaAs circuit .

Figure 9- Silicon dioxide precipitates in Si materials

REFERENCES
/1/ Nanishi ,Y., Ishida, S., Honda, T., Ymazaki , H., Miyazawa , S. , J.Appl.Phys. 28 (1982) L335
/2/ Nanishi ,Y , Miyazawa ,S , Matsuoka , Y , DRIP I symposium - Montpellier - MRS monographs 31 JP Fillard Ed. Elsevier (1985) 225
/3/ Packeiser, G , NATO workshop - Oxford - D1 (1987)
/4/ Lohnert K, Wettling W, Koschek G , Sem. Ins. III-V Conf. Hakone , Ohmsha Ed. (1986) 267
/5/ Tajima M, Iizuka T, GaAs Related Compounds - C Hilsum Ed. - Inst.Phys.Conf. ser. 33a (1976) 123
/6/ Moriya K, Ogawa T , Phil.Mag. A 44 (1981) 1085
/7/ Moriya K, Ogawa T , Jap Jour.Appl.Phys. 22 (1983) L207
/8/ Moriya K- Jap Jour.Appl.Phys. (en japonais) 55 6 (1986)
/9/ Moriya K- DRIP I Symposium , JP Fillard Ed. (Elsevier) MSM 31 (1985) 27
/10/ Katsumata T, Okada H, Kikuta T, Fukuda T, Ogawa T, Semi Ins. III-V Conf. (Hakone) Ohmsha Ed. (1986) 149

/11/ Otoki Y, Nakamori M, Nakazono R, Kuma S , Semi Ins. III-V Conf. (Hakone) Ohmsha Ed. (1986) 285
/12/ Gall P, Fillard J P, Bonnafé J, Castagné M, Weyher J L, George A- DRIP II Symposium E.Weber Ed. (Elsevier) MSM 44 (1987) 215
/13/ Fillard JP, Castagné M, MRS Symp.Proc. 69 (1986)
/14/ Kidd P, Booker G R, Stirland D J, 5th Oxford Conf Micros. Sem. Con. Mat. - J5 (1987)
/15/ Proceedings Symposia DRIP I and II - Materials Science Monographs N°31 - N°44 Elsevier (1985) - (1987) .
/16/ Fillard J P, Rev.Phys.Appl. 42 149 (1987)
/17/ Ogawa T, DRIP I symposium - Fillard J P Ed. (Elsevier) MSM 31 (1985) 1
/18/ Lord Rayleigh - Phil. Mag. XLI 107 (1871) 274
/19/ Mie G, Annalen der Physik 25 (1906) 377
/20/ Skolnick M S, Brozel M R, Appl.Phys.Lett.48 (1986) 341
/21/ Skolnick M S , Brozel M R, Pitt A D, Maguire J, Semi Ins. III-V Conf (Hakone) Ohmsha Ed. (1986) 139.
/22/ Ogawa T, Jap.Jour.Appl.Phys.Lett. 25 L 316 (1986)
/23/ Ogawa T, Jap.Jour.Appl.Phys.Lett. 25 L 916 (1986)
/24/ Fillard J P, Montgomery P, Gall P, Asgarinia M, Bonnafé J, Jap. Jour. Appl. Phys.Lett.27 (1988) 384
/25/ Stirland D J , Sem. Ins III-V Conf. Hakone . Ohmsha Ed. 81 (1986) 133
/26/ Fillard , JP - NATO workshop - Oxford - April 1987
/27/ Policard A, Bessis M, Locquin M, Traité de microscopie , instruments et tech- niques - Paris (1957)
/28/ Castagné M, Fillard J P, Bonnafé J- Sol. State Com. 54 (1985) 7 653
/29/ Ogawa T, Nango N, Conf Soc.Phys. Jap. Tokyo novembre (1985)
/30/ Cullis AG , Augustus PD , Stirland DJ , Jour. Appl.Phys. Vol.51 (1980) 2556
/31/ Fillard JP , Baroudi A ,Gall P, Favennec PN, Bonnafé J , Coquillé R , Jap. Jour. Appl. Phys. 64 (1987) 1243
/32/ Fillard JP , Gall P, Baroudi A , George A , Bonnafé J , Jap. Jour. Appl. Phys. 26 (1987) L1255
/33/ Gall P, Fillard JP, Bonnafé J, Castagné M, Weyher JL, George A, DRIP II symposium - Weber E Ed. (Elsevier) MSM 44 (1985) 215
/34/ Stirland D J , Inst.Phys. Conf.Ser. 33 (1977) 150
/35/ Fillard JP, Gall P, Weyher JL, Asgarinia M, Montgomery P, Semi Ins. III-V Conf. (Malmö) to be edited (1988)
/36/ Osaka J,Okamoto H, Hyuga F, Watanabe K, Yamada K, Semi Ins. III-V Conf. (Hakone) Ohmsha Ed (1986) 279 .
/37/ Moriya K, Semi Ins. III_V Conf. (Hakone) Ohmsha Ed. (1986) 151
/38/ Kuma S, Otoki Y, Kurata K, DRIP I symposium - Fillard J P Ed. (Elsevier) MSM 31 (1985) 19

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

FUTURE BIPOLAR DEVICE STRUCTURES

H. GOTO

Bipolar Process Engineering Department, Fujitsu Limited, 1015, Kamikodanaka, Nakahara, Kawasaki 211, Japan

Abstract - Recently developed bipolar device structures including their problems and future trends are reviewed. Polysilicon emitter-base self-aligned structures and trench isolation techniques are becoming key elements for high performance bipolar ECL device structures, by which parasitic capacitances and resistances have been reduced drastically. In order to get further improved performance, smaller parasitic capacitances associated with the pull-up resistor as well as high cutoff frequency are required. Wafer-direct-bonded SOI structures are the promising candidate, while the base resistance and the cutoff frequency should be optimized moderately. The most serious problem is the power dissipation of ECL-type circuits. Smaller logic swings and low temperature operation should be also considered.

1 - INTRODUCTION

As the integrated circiut industry has come to the mature stage, the role of each device is getting more and more distinct. CMOS is the major VLSI and/or ULSI device because of the low power dissipation, while bipolar is in the high speed ranges with less integration level, and GaAs is in the small scale high frequency IC domains. Recently, between CMOS and bipolar ECL, BiCMOS has emerged with higher speed than CMOS and less power dissipation than ECL. With the advent of BiCMOS, conventional S-TTL or IIL digital devices seem to be swept out in the field of bipolar VLSI's such as gate arrays although standard LS- or ALS-TTL families are still in volume production. At present, the major role of bipolar devices has focused on very high speed logic devices with larger integration level than GaAs, which are achieved by ECL-type circuits. ECL has just become a major driving force for the progress of high performance bipolar device technologies. In this paper, the present status of the bipolar technologies is reviewed and by examining the problems of bipolar devices especially, ECL-type circuits, future bipolar device structures and suitable features are discussed.

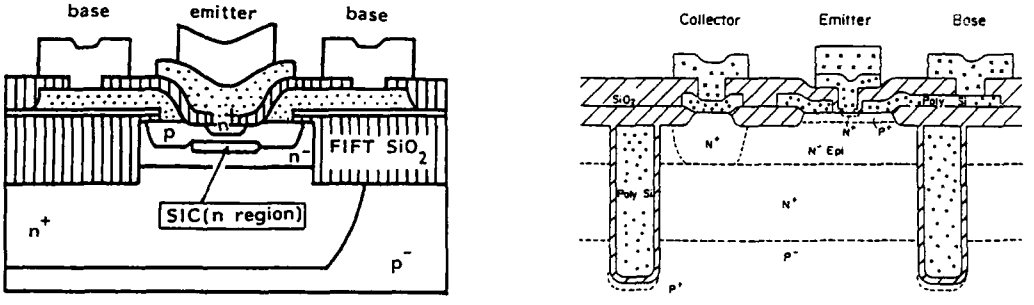
2 - PRESENT BIPOLAR DEVICE TECHNOLOGIES

There have been several good reviews about bipolar device and circuit technologies. Especially, in the article /1/ , almost all of bipolar device structures and circuit examples before 1980 were reviewed. In the articles /2,3/ ,advanced bipolar device structures developed recently were compared and discussed, while the similar topic was also discussed in the article /4/. In this section, recently developed bipolar device structures are rearranged and their essential properties are pointed out.

2 - 1 Device Structures

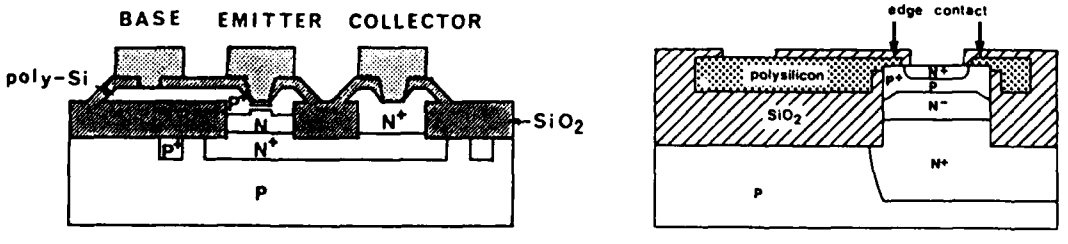
Since the development of DOPOS technique /5/ , polysilicon has been used as the emitter electrode as well as the emitter diffusion source. In the end of the seventies, a

polysilicon self-aligned (PSA) technique /6/ also utilized polysilicon as a base electrode. Based on these techniques, self-aligned techniques using the polysilicon as both the emitter and the base electrodes accompanied with anisotropic dry etching techniques made it possible to realize sophisticated self-aligned bipolar device structures /7,8/ . These techniques were mainly intended to achieve
(1) reduction of parasitic base-collector capacitances
(2) reduction of base resistances
without using tight lithography rules. The key structure is the stacking polysilicon electrodes separated by a sidewall insulating spacer. Another approach was also employed to achieve those two requirements by trying to make an ideal one-dimensional structure, or to make a symmetrical emitter and collector strucrure. This structure was formed by mesa-etching combined with a planarizing polysilicon technique /9/ or selective-epi-silicon growth in grooves /10/ . There is one more important structure implemented in bipolar devices, that is trench isolation /11,12,13/ . By introducing trench isolation, the collector-substrate capacitance has been reduced much smaller than that of oxide isolation. Concerning trench isolation there is a review in the article /14/. By improving these techniques and combining them, a lot of advanced bipolar device structures have been developed in the eighties. The stacking double polysilicon self-aligned structures are named such as SST /7/ , Self-aligned transistor /8/ , OXIS-III /15/ , SAPT /16/ , SDD /17/ , ExCL /18/ , MOSAIC-III /19/ and ESPER /20/ . Similar structures were also reported in the articles /21,22,23,24,25,26/ . These structures can be easily combined with trench isolation techniques. The most advanced examples are shown in Fig.1. The mesa-etched or selective-epitaxially grown sidewall-base-contacted structures are named SICOS /9/ or Symmetrical transistor /10/ . A sophisticated self-aligned transistor using selective growth of poly- and single- crystalline silicon was demonstrated /27/ . There is also an advanced version of SICOS as shown in Fig.2.



(a) (b)

Fig. 1 - Advanced double polysilicon self-aligned structures.
(a) SST-1B. A selectively implanted deep collector (SIC) region is formed beneath the intrinsic base /28/.
(b) ESPER combined with U-FOX. Typical combination with trench isolation /20/.



(a) (b)

Fig. 2 - Sidewall base contact structures.
(a) Self-aligned transistor using selective growth of poly- and single- crystalline silicon /27/ .
(b) SELECT/SICOS. Edge contact structure is applied to conventional SICOS /29/.

2-2 Device Characteristics

Ring oscillator results demonstrated by the previously mentioned device structures are listed in Table 1. The basic gate delay time of 20.5 ps was achieved by an NTL circuit, and sub-40 ps delay times were demonstrated by ECL circuits. Practical device performances are also listed in Table 2. Subnanosecond static RAM's, sub-50 ps gate arrays, 10 GHz frequency deviders have been demonstrated. Using advanced bipolar device technologies, not only high frequency small scale IC's but also very high speed LSI's or VLSI's have been realized.

Table 1 - Ring oscillator results by advanced bipolar technolgies.

Company	Circuit	tpd	Power or Ics	Emitter Size	Technology	Reference
NTT	NTL	20.5 ps	2.32 mW/G	0.35x13 um^2	SST-1B	/28/
NTT	ECL	34.1 ps	7.54 mW/G	0.35x 7 um^2	SST-1B	/28/
Fujitsu	ECL	38.8 ps	1.28 mA/G	0.35x10 um^2	ESPER **	/20/
Hitachi	ECL	46.7 ps	0.58 mA/G	2x 3 um^2 *	SELECT/SICOS	/29/
Plessey	CML	50 ps	---------	----------	HE1 **	/26/
NEC	ECL	52 ps	2 mA/G	1x 8 um^2 *	***	/30/
CETRI	CML	52 ps	0.16 mW/G	1.5x 3 um^2	***	/23/
Matsushita	CML	53 ps	3.2 mW/G	0.5x12 um^2	SDD **	/31/
Tektronix	ECL	54.6 ps	1.8 mA/G	0.5x10 um^2	***	/24/
SONY	---	69 ps	0.43 mA/G	0.6x 1 um^2	***	/25/
IBM	ECL	73 ps	12 mW/G	0.8x 2 um^2	Self-Aligned Transistor **	/32/
Motorola	ECL	75 ps	0.8 mA/G	1.5x 4 um^2 *	MOSAIC-III	/19/
Fairchild	ECL	80 ps	0.9 mA/G	1.5x2.5um^2	SAPT **	/16/

*; Mask size. **; Combined with trench isolation.
***; Double polysilicon self-aligned structures.

Table 2 - Recently developed high performance bipolar devices.

Device	Performance	Technology	Company	Reference
5K ECL RAM	t_{AA}= 0.85 ns	Self-Aligned Tr. and Trench Iso.	IBM	/33/
16K ECL RAM (& 1.2K G.A.)	t_{AA}= 2.8 ns	IOP-II	Fujitsu	/34/
32K RAM	t_{AA}= 3 ns	Trench Iso.	IBM	/35/
64K ECL RAM	t_{AA}= 5 ns	IOP-II	Fujitsu	/36/
2.1K CML G.A.	tpd= 43 ps	SST-1B	NTT	/37/
13K ECL G.A.	tpd= 100 ps	HE1	AMC	/38/
10K ECL/15K CML G.A.	tpd= 120 ps	OXIS-IIIH	Siemens	/39/
18K ECL G.A.	tpd= 150 ps	SCOT	Mitsubishi	/40/
16b Multiplier	tm = 4.3 ns	SST-1B	NTT	/37/
1/8 Divider	ft =10.38 GHz	SST-1A	NTT	/41/
8b A-D Convertor	fs = 400 MHz	SST-1A	NTT	/42/

t_{AA}; Address access time. tpd; Basic gate delay time.
tm ; Multiplication time. ft ; Toggle frequency.
fs ; Sampling frequency.

2-3 Properties of Present Bipolar Device Structures

The technique utilizing polysilicon is one of the key properties for the present advanced bipolar device technologies. The polysilicon is used as a contact material for both the emitter and the base as well as the diffusion sources. The polysilicon electrodes are easily separated by the oxide of polysilicon, realizing the minimum distance between the base and the emitter. This causes the small base-collector capacitance and the small base resistance. At the same time, sidewall spacer techniques also provide easily lower submicron emitter widths with only upper submicron lithography rules. In addition, the polysilicon emitter contact enhances the current gain, although the mechanism is not yet clearly understood. By optimizing the process condition, the current gain can be increased

without suffering from high emitter series resistances /43/ . Another major property is trench isolation. It not only enhances the packing density but reduces drastically the collector-substrate parasitic capacitance compared with conventional oxide isolation. By using these techniques, the switching speed as well as the power dissipation of ECL-type circuits has been improved. VLSI chips of ECL circuits are now coming in the market /44/.

3 - PROBLEMS OF PRESENT BIPOLAR DEVICE STRUCTURES

In this section, the problems concerning the advanced bipolar device structures of typical two cases are discussed.

3-1 Double Polysilicon Self-aligned Structures

The most serious problem of this structure is the sidewall spacer thickness control. There are two sequential ways to form the implanted intrinsic base region and the sidewall spacer; sidewall etching followed by base implantation and vice versa. In the case of the former, although there is little damage for emitter-window opening, the cross-link region between the intrinsic base and the extrinsic (graft) base should be formed by lateral diffusion from the polysilicon base electrode. Too much lateral diffusion enhances peripheral tunneling leakage, while less encroachment cause punchthrough /45/ . In the case of the latter, there is little cross-link problem because the intrinsic base region exists just beneath the sidewall spacer. However, careful emitter opening using anisotropic dry etching should be required. Even in this case the emitter junction depth, the sidewall spacer thickness and the graft base depth should be optimized to avoid too much lateral encroachment.

Another problem is a tight alignment control between the graft base mask and the intrinsic base opening mask when the base-collector capacitance is tried to be reduced. The double polysilicon self-aligned structures require one mask alignment step if the walled base structure is employed, even if lower submicron emitter widths and emitter-base separation can be achieved without using tight minimum lithography rules. In the case of /7,23,28/ , the graft base region is formed from the same mask as that of the intrinsic base formation without any alignment step. However, there is a disadvantage of the non-walled base structure that increases the outer peripheral parasitic base-collector capacitance. If a walled-base-like structure is introduced, a tightly aligned lithographic tool is needed.

3-2 Sidewall Base Contact Structures

The philosophy of sidewall base contact structures is mainly to realize an ideal symmetrical one-dimensional structure just as written in textbooks. It is the ultimate structure, but actual processes and device operations never result in one-dimensional behavior. The graft base diffusion control is the most serious problem in this structure. Sufficient graft base diffusion is necessary to achieve lower base resistance because only the sidewall regions are contacted by the doped polysilicon. The graft base diffusion occurs not only laterally towards the emitter but vertically along the oxide wall to the collector. The lateral diffusion enhances the emitter-base leakage current and the base-emitter capacitances just as the same situation as the double polysilicon self-aligned structures. Moreover, the vertical diffusion increases the base-collector capacitance. In order to avoid these problems some process modifications have been tried sacrificing process simplicity /28,46/ . Nevertheless, there are still contradictions between the base resistance and the base-collector capacitance.

4 - LIMITATIONS CONCERNING ECL CIRCUIT

The advantages of ECL circuits are not only their high speed operation but their excellent drivability of loaded capacitances. These are originated in the non-saturating operation of the current switch and the high transconductance of the emitter follower. However, their large power dissipation is the serious problem, that limits the integration level of ECL chips. A lot of circuit techniques have been tried to reduce the power dissipation of ECL-type circuits by using e.g. series gating, cascaded circuits of differential inputs and so on, but modified circuits usually have less flexibility for various applications of system designs. So, essentially the power dissipation of the current switch emitter follower should be decreased for VLSI or ULSI applications, whereas they never reduced to be nearly zero such as CMOS that dissipates the power only when switching. This means there is always a large gap between the largest integration level of CMOS chips and that of ECL chips. Nevertheless, the strategy of the optimization for ECL-type circuits is; the smaller the current, the lower the power dissipation.

How can the currents of ECL circuits be reduced without any degradation of the switching speed? The switching delay time (tpd) of ECL circuits is roughly expressed as

$$tpd = rb \cdot Cb + R_L \cdot Cc + k_1 \cdot t_F \qquad \text{--------(1)}$$

$$Cb = k_2 \cdot Cd + k_3 \cdot Cbe + k_4 \cdot Ccb \qquad \text{--------(2)}$$

$$Cc = k_5 \cdot Ccb + k_6 \cdot Ccs + k_7 \cdot C_L + k_8 \cdot Cef \qquad \text{--------(3)}$$

where rb is the base resistance, R_L is the pull-up resistor, t_F is the transit time, Cd is the diffusion capacitance, Cbe is the base-emitter capacitance, Ccb is the base-collector capacitance, Ccs is the collector-substrate capacitance, C_L is the wiring capacitance associated with R_L, Cef is the loaded capacitance associated with the emitter follower transistor, and k_1 to k_8 are the constants. If the logic swing is kept constant, the smaller switching current is achieved by larger R_L. This is the way to reduce the power dissipation especially for the larger integration chips. Lower power ECL circuits are more influenced by the second term of Eqn.(1) and Cc must be reduced much more than high power circuits to improve the switching time. These can be realized only by the improvement of the device structure and the process technology. There is also the other way to obtain smaller switching current by reducing the logic swing. R_L can be kept constant when the switching current is reduced in proportion to the logic swing while the circuit noise margins are decreased. However, supposing that low temperature operation such as 77k is introduced, the reduction of the logic swing from the value of room temperature to a few tenths is not a ridiculous approach. Additionally, there is a possibility to enhance the efficiency of the cooling system.

The first term and the last term of Eqn.(1) are strongly correlated. The first term can be reduced by smaller base resistance (rb). This means a highly doped intrinsic base region or a wide base width as well as smaller extrinsic base resistances are required. The last term can be reduced by smaller transit time, or higher cutoff frequency (f_T). The otimization of the effects caused by rb and f_T is very important because smaller rb often corresponds to lower f_T directly, although Cbe and Ccb can be reduced by shrinking the active area. The highest f_T reported to date corresponds to the t_F of nearly 6.2 ps (f_T=25.7 GHz), that is only 18% of the total switching time (tpd=34.1 ps) /28/ . If only the maximum f_T ($f_T$max) is raised to the twice value (>50 GHz), the reduction of the transit time will contribute at most 3.1-ps decrease of the switching time. Only too high $f_T$max has very little influence on the delay time. Accordingly, moderate values of f_T and rb should be chosen avoiding punchthrough and forward tunneling. Rather little contribution of $f_T$max upon the total switching time is the most different point from GaAs HBT. In the case of GaAs HBT, the transit time is 42% of the switching time (t_F=2.3 ps, tpd=5.5 ps) /47/ . The contribution of RC time constants of GaAs HBT is much smaller than that of Si BJT. This is mainly because parasitic capacitances corresponding to the substrate are nearly zero owing to the semi-insulating GaAs characteristics. This fact suggests that the parasitic capacitances associated with the substrate of Si BJT should be reduced drastically to the nearly zero value.

In addition to the properties of ECL circuits using Si BJT mentioned above, there are still a few problems. When the switching current is reduced, the emitter area should be decreased to keep the cutoff frequency high enough for the optimized circuit operation. This is a suitable situation for the reduction of junction capacitances, but contact resistances as well as base resistances will increase. In order to avoid this problem, the process and device structural optimizations are also required.

5 - FUTURE BIPOLAR DEVICE STRUCTURES

Double polysilicon self-aligned structures and sidewall base contact structures will go on making a progress in the near future. However, sooner or later, the most suitable structure will be selected from not only the performance point but the production cost point of view. The double polysilicon self-aligned processes have been already used in volume production. The most sophisticated device has been realized by the double polysilicon self-aligned and trench-isolated structure using submicron lithography rules. It seems that double polysilicon self-aligned structures will be the major selection in the future because of their valanced cost and performance properties. Of cource, present device structures are not optimized enough to achieve the ultimate performance of Si BJT. In this section, the future Si BJT structure to achieve the ultimate performance of ECL-type circuits in ULSI ranges are discussed.

As discussed in Section 4, the most crucial factor affecting the switching time and the power dissipation is $R_L \cdot Cc$. Cc consists of Ccb,Ccs,C_L and Cef. Among these capacitances Ccb can be reduced by the decrease of the base area using tightly aligned lithography tools, but never reaches zero. On the contrary, Ccs and C_L can be reduced if the parasitic capasitances associated with the silicon substrate are decreased. Ccs can be reduced to the nearly zero value while C_L may be less than a half if an insulating substrate is used. In other words, SOI is a very good choice to decrease both Ccs and C_L. A lot of SOI

techniques such as beam recrystallization, epitaxial lateral overgrowth and so on have been investigated for bipolar device applications, but there has always been a serious problem; the crystalline defects. However, recently wafer direct bonding techniques to obtain a thin SOI structure are coming up as a new face. According to these techniques, the crystalline quality of the SOI is essentially the same as that of single crystalline silicon wafers. This fact was proved by experimental 64K DRAM's /48/ . Now, the SOI structure of bipolar devices are not a dream but a realistic target.

Next point is the reduction of each resistance such as the base resistance (rb) and the emitter series resistance. In order to reduce series resistances, polysilicon electrodes should be covered by silicides or refractory metals /49/ . Concerning the base resistance, the intrinsic base resistance is the most dominant in the case of the self-aligned structures. The intrinsic base resistance is dependent on the base doping and the base width which also affect the cutoff frequency. So, the base profile must be optimized for the moderate intrinsic base resistance and high enough cutoff frequency without suffering from forward biased tunneling and punchthrough. The emitter series resistance, which is dominantly a contact resistance associated with the polysilicon emitter contact, should be kept small when the emitter area is decreased to sustain the sufficient cutoff frequency. For this purpose nearly recrystallized extended polysilicon emitter with a silicide contact is quite suitable.

Fig.3 shows the schematic cross-section of a future bipolar device structure. The BJT itself is fully isolated by the trench-isolated SOI structure. Both the polysilicon electrodes for the base and the emitter are covered by silicides or refractory metals, and the polysilicon emitter is a nearly single-crystallized extended emitter. The lateral device dimensions are optimized to achieve the best valance between the base resistance and the base-collector capacitance avoiding the extrinsic base encroachment trouble. The vertical profiles are also optimized for the emitter junction depth to avoid the narrow emitter effect and for the base junction depth to optimize the intrinsic base resistance and the cutoff frequency. In order to reduce the Kirk effect the collector doping profile should be tailored. Fig.4 is an example of such doping profiles of both the intrinsic and extrinsic base regions.

According to the two-dimensional physical DC and AC device simulation /50/ based on the device structure of Fig.3 and the impurity profiles of Fig.4, the cutoff frequency, the base-collector capacitance, and the AC base resistance were estimated. Fig.5 shows the cutoff frequency (f_T) versus the collector current (Ic) characteristics as well as the current gain dependence upon Ic. Maximum f_T of 56 GHz at the collector current of 0.3 mA/um is obtained and h_{FE} of more than 50 is achieved at the same time as shown in Fig.5. Fig.6 shows the base-collector capacitance (Ccb) and the AC base resistance (rbb') versus Ic, where Ccb of 1.1 fF/um and rbb' of 550 ohm·um are obtained at the small collector current. From these results the basic gate delay time of a typical ECL circuit (F.I.=F.O.=1) is estimated at least below 20 ps at a switching current of 0.3 mA/um in a condition of room temperature. If much lower temperature operation such as 77K is considered, the logic swing can be reduced without suffering from the decrease of circuit noise margins. In this case not only circuit designing but process conditions e.g. the doping profiles should be optimized /51,52/ . Low temperature operation can achieve the lower power dissipation, the enhancement of cooling efficiency, lower wiring resitances and longer MTF of wiring metals. All of these properties are quite suitable for ULSI chips that consist of ECL-type circuits.

When the basic gate delay time reaches around 20 or 30 ps, the most serious problem is the loaded wiring delay time owing to the wiring capacitances which are physically limited. ECL-type circuits are the best solution for driving large loaded capacitances whereas they are always accompanied by the large power dissipation. Low temperature operation is advantageous to relax this problem, while it is strongly dependent on the system and circuit designing strategy. Accordingly, the optimization of cost and performance valances in the total system designs including the software, the hardware and the individual chips are the key features for the system development in the future.

6 - SUMMARY

Today's advanced bipolar device structures, their problems and future trends have been reviewed. The major role of bipolar devices has focused on ECL-type circuits. In order to reduce the switching time and the power dissipation, the double polysilicon self-aligned structures combined with trench isolation have become the major structure. For further improvement RC time constants associated with the silicon substrate should be reduced, which can be achieved by wafer direct bonding SOI techniques. Low temperature operation should be also considered although it is strongly dependent on the system designing strategy. Basic gate delay times now have little influence because loaded wiring delay times are by far larger. The total cost and performance valances of the system designs are the key features for the future system development.

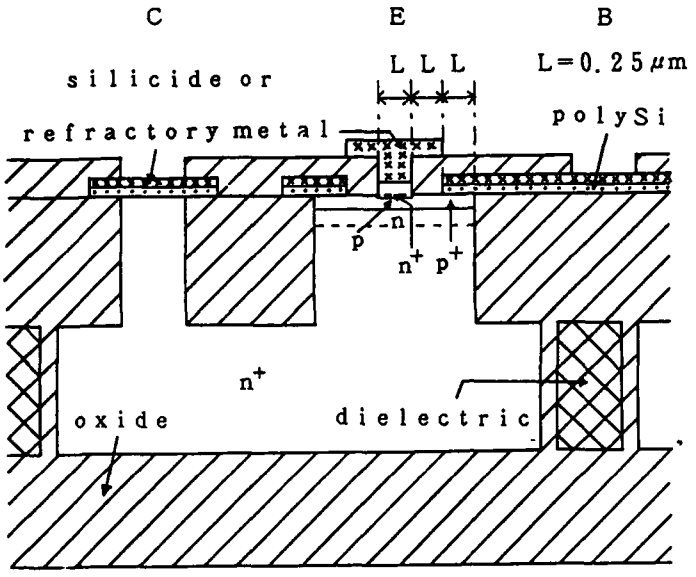


Fig. 3 - Ultimate future Si BJT structure. Effective quarter micron figures of the lateral dimension are considered. Nearly single crystallized 0.1-um thick extended polysilicon emitter is also assumed. (The vertical dimension is not in proportion to the lateral one.)

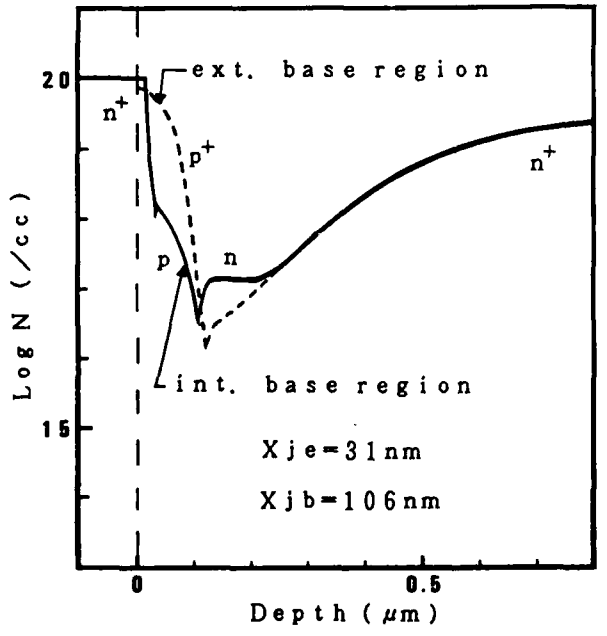


Fig. 4 - Example of doping profiles corresponding to the device structure in Fig.3. N is the impurity concentration. The maximum ionized impurity concentration is limited by 1.0E20/cc.

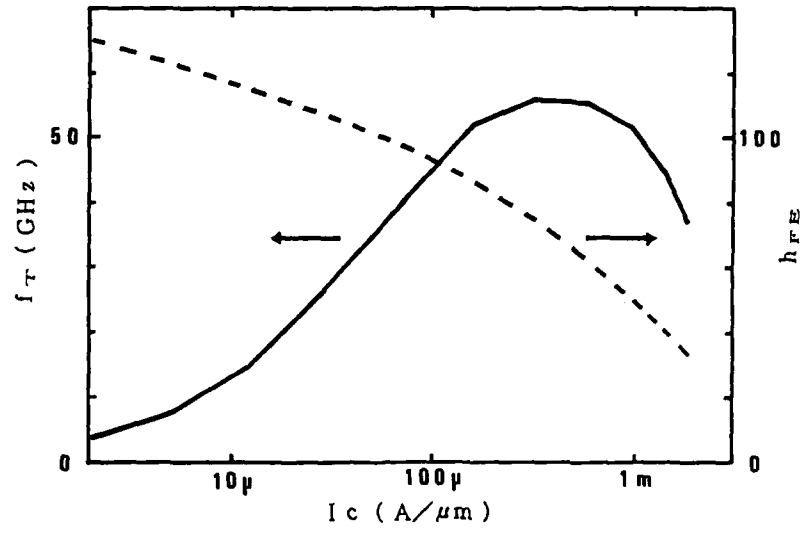


Fig. 5 - Cutoff frequency (f_T) and currnt gain (h_{FE}) versus collector current (Ic) at Vce=3v. The emitter width is standardized by 1.0 um.

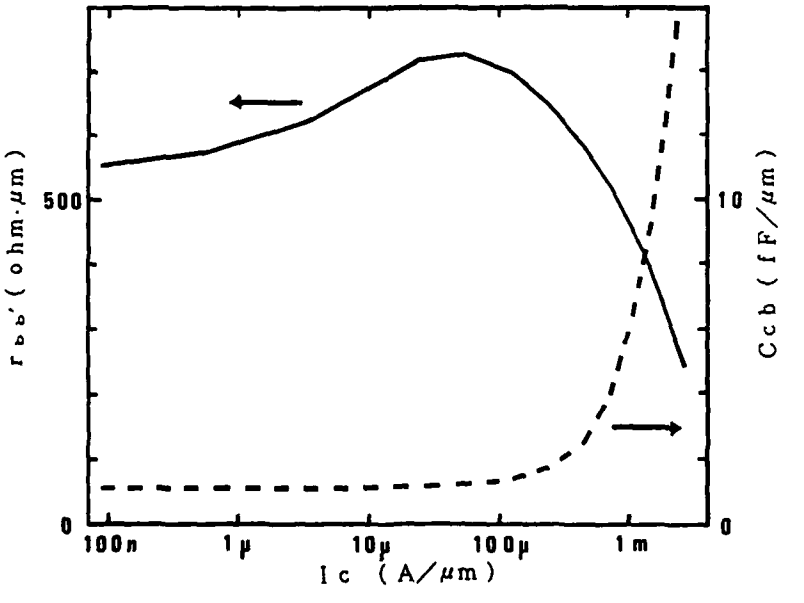


Fig. 6 - Base-collector capacitance (Ccb) and AC base resistance (rbb') versus collector current (Ic) at Vbc=0V. The emitter length is also standardized by 1.0 um.

REFERENCES

/ 1/ Lohstroh, J., Proc. IEEE 69 (1981) 812.
/ 2/ Ning, T.H., ESSDERC Tech. Dig. (1987) 277.
/ 3/ Sakai, T., Ext. Abst. 17th CSSDM, Tokyo (1985) 373.
/ 4/ Wiedmann, S.K., IEEE J. Solid-State Circuits SC-19 (1984) 282.
/ 5/ Takagi, M., Nakayama,K., Terada, C. and Kamioka, H., Proc. of 4th CSSD, Tokyo (1972) 101.
/ 6/ Okada, K.,Aomura, K.,Nakamura, T., and Shiba, H., IEEE Trans. Electron Devices ED-26 (1979) 385.
/ 7/ Sakai, T., Kobayashi, Y., Yamauchi, H., Sato, M. and Makino, T., Proc. of 12th CSSD, Tokyo (1980) 155.
/ 8/ Ning, T.H., Isaac, R.D., Solomon, P.M., Tang, .D.D. and Yu, H.N., IEDM Tech. Dig. (1980) 823.
/ 9/ Nakamura, T., Miyazaki, T.,Takahashi, S., Kure, T., Okabe, T. and Nagata, M. ISSCC Tech. Dig. (1981) 214.
/10/ Tang, D.D., Silvestri, V.J., Yu, H.N. and Reisman, A., IEDM Tech. Dig. (1980) 58.
/11/ Pogge, H.B., RNP ECS Meeting, May (1979).
/12/ Goto, H., Takada, T., Abe, R., Kawabe, Y., Oami, K. and Tanaka, M., IEDM Tech. Dig. (1982) 58.
/13/ Hayasaka, A., Tamaki, Y., Kawamura, M., Ogiue, K. and Ohwaki, S., IEDM Tech. Dig. (1982) 62.
/14/ Goto, H., and Inayoshi, K., ESSDERC Tech. Dig. (1987) 369.
/15/ Ullrich, H., Brackelmann, W., Fritzsche, H. and Wieder, A., ISSCC Tech. Dig. (1985) 200.
/16/ Vora, M., Ho, Y.L., Bhamre, S., Chien, F., Bakker, G., Hingarh, H. and Schmitz, C., IEDM Tech. Dig. (1985) 34.
/17/ Kikuchi ,K., Kameyama, S., Kajiyama, M., Nishio, M. and Komeda, T., IEDM Tech. Dig. (1986) 420.
/18/ Electronics Mar. 19 (1987) 73.
/19/ Zdebel, P.J., Balda, R.J., Hwang, B-Y., d.l.Torre, V. and Wagner, A., IEEE BCTM Tech. Dig. (1987) 172.
/20/ Ueno, K., Goto, H., Sugiyama, E. and Tsunoi, H., IEDM Tech. Dig. (1987) 371.
/21/ Tashiro, T., Takemura, H., Kamiya, T., Tokuyoshi, F., Ohi, S., Shiraki, H., Nakamae, M. and Nakamura, T., IEDM Tech. Dig. (1984) 686.
/22/ Sasaki, N., Anzai, A. and Uehara, K., IEDM Tech. Dig. (1983) 546.
/23/ Chai, S.H. and Lee, J.H., IEDM Tech. Dig. (1985) 26.
/24/ Park, H.K., Boyer, K., Tang, A., Clawson, C., Yu, S., Yamaguchi, T. and Sachitano, J., IEEE BCTM Tech. Dig. (1986) 39.
/25/ Yoshii, Y., Nakamura, M., Hirasawa, K., Kayanuma, A. and Asano, K., ISSCC Tech. Dig. (1987) 96.
/26/ Hunt, P.C. and Cooke, M.P., IEEE CICC Tech. Dig. (1988) 22.2.1.
/27/ Oh-uchi, N., Kayanuma, A., Asano, K., Hayashi, H. and Noda, M., IEDM Tech. Dig. (1983) 55.
/28/ Konaka, S., Amemiya, Y., Sakuma, K., and Sakai, T., Ext. Abst. 19th CSSDM, Tokyo (1987) 331.
/29/ Washio, K., Nakamura, T., Nakazato, K. and Hayashida, T., ISSCC Tech. Dig. (1987) 58.
/30/ Takemura, H., Kamiya, T., Ohi, S., Sugiyama, M., Tashiro, T. and Nakamae, M., IEDM Tech. Dig. (1986) 424.
/31/ Sakai, H., Kikuchi, K., Kameyama, S., Kajiyama, M. and Komeda, T., Symp. on VLSI Tech. Tech. Dig. (1987) 17.
/32/ Tang, D.D., Li, G.P., Chuang, C.T., Danner, D.A., Ketchen, M.B., Mauer,J.L., Smyth, M.J., Manny, M.P., Cressler, J.D., Ginsberg, B.J., Petrillo, E.J., Ning, T.H., Hu, C.C. and Park, H.S., ISSCC Tech. Dig. (1986) 104.
/33/ Chuang, C.T., Tang, D.D., Li, G.P., Franch, R.L., Ketchen, M.B., Ning, T.H., Brown, K.H. and Hu, C.C., Symp. on VLSI Tech. Tech. Dig. (1988) 91.
/34/ Sugo, Y., Tanaka, M., Mafune, Y., Takeshima, T., Aihara, S., and Tanaka, K., ISSCC Tech. Dig. (1986) 256.
/35/ Chan, Y.H. Brown, J.L., Nijhuis, R.H., Rivadeneira, C.R. and Struk, J.R., ISSCC Tech. Dig. (1986) 210.
/36/ Awaya, T., Toyoda, K., Nomura, O., Nakaya, Y., Tanaka, K. and Sugawara, H., ISSCC Tech. Dig. (1987) 130.
/37/ Suzuki, M., Hirata, M. and Konaka, S., ISSCC Tech. Dig. (1988) 70.
/38/ Coy, B., Mai, ., and Yuen, R., IEEE CICC Tech. Dig. (1988) 20.1.1.
/39/ Franz, M., Delker, K., Whang, T.C. and Wilhelm, W., IEEE CICC Tech. Dig. (1988) 20.2.1.
/40/ Nishimura, T., Kato, S., Tatsuki, M., Sato, H., Kohara, M., Sakaue, K., Hirao, T., and Kuramitsu, Y., ISSCC Tech. Dig. (1986) 76.

/41/ Sakai, T., Konaka, S., Yamamoto, Y. and Suzuki, M., IEDM Tech. Dig. (1985) 18.
/42/ Akazawa, Y., Iwata, A., Wakimoto, T., Kamato, T. Nakamura, H. and Ikawa, H., ISSCC Tech. Dig. (1987) 98.
/43/ Stork, J.M.C., Ganin, E., Cressler, J.D., Patton, G.L. and Sai-Halasz, G.A., IBM J. RES. DEVELOP. 31 (1987) 617.
/44/ Electronics Mar. 31 (1988) 99.
/45/ Li, G.P., Hackbarth, E. and Chen, T.C., IEEE Trans. Electron Devices ED-35 (1988) 89.
/46/ Tamaki, Y., Shiba, T., Ikeda, K., Nakamura, T., Natsuaki, N., Ohyu, S., and Hayashida, T., IEEE BCTM Tech. Dig. (1987) 22.
/47/ Nagata, K., Nakajima, O., Yamauchi, Y., Nittono, T., Ito, H. and Ishibashi, T., IEEE Trans. Electron Devices ED-35 (1988) 2.
/48/ Gotou, H., Arimoto, Y., Ozeki, M., and Imaoka, K., IEDM Tech. Dig. (1987) 870.
/49/ Kabza, H., Reisch., M., Probst, V., Böhm, W., Fertsch, J., Shcaber, H. and Eggers, H., ESSDERC Tech. Dig. (1987) 1047.
/50/ Nagase, Y., Hashimoto, K., Goto, H., Nakadai, T., Nawata, T. and Takada, T., NUPAD-II Tech. Dig. (1988) Nupad156.
/51/ Woo, J.C.S. and Plummer, J.D. IEDM Tech. Dig. (1987) 401.
/52/ Stork, J.M.C., Harame, D.L., Meyerson, B.S. and Nguyen, T.N., IEDM Tech. Dig. (1987) 405.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988 C4-481

III-V COMPOUND DEVICES ON SILICON : SITUATION AND PERSPECTIVES

J.P. HIRTZ and M.N. CHARASSE

Laboratoire Central de Recherche, Thomson-CSF, Domaine de Corbeville, F-91401 Orsay Cedex, France

Introduction - For several years, there has been a great deal of activity in the hetero epitaxy of GaAs and others III-V materials on silicon substrates. For the first time, we can now seriously consider hybridization of Si and GaAs technologies on the same wafer. Applications for GaAs on Si range from substitution of a GaAs substrate by a large area passive Si substrate with superior strength and thermal conductivity to the multi-function ICs were Si and GaAs are monolithically integrated on the same chip.

The achievement of device quality GaAs on Si and others mismatched systems is the consequence of several factors. Both Molecular Beam Epitaxy (MBE) and Metalorganic Chemical Vapor Deposition (MOCVD) permit preparation of oxide force silicon surfaces. The lattice mismatched and associated misfit dislocations are among the difficulties to be solved. The appropriate choice of substrate crystallographic orientation and a nucleation layer deposition regime helped to confine most of the misfits dislocations to a thin region close to the interface.

As a result the thicker portions of the epitaxial structure could have a crystal perfection that, though inferior to homoepitaxial GaAs, is suitable for majority carriers devices.

In the following, we will describe growth of device quality GaAs on Si and discuss material optimization through thermal annealing. We will then discuss the performance of GaAs devices which have been already grown on Si.

1 - MBE GROWTH AND MATERIAL OPTIMIZATION THROUGH THERMAL ANNEALINGS

Some publications have shown that rapid or long thermal annealings /1,2/ lead to a reduction of the dislocation density in the GaAs epilayer. We have studied the evolution of the defects in relation with these two thermal processes on a 400 nm GaAs film grown on Si /3/.
The samples are grown in a Varian Gen II MBE system. The (001) Si substrates are misoriented by 4° towards [110] or [100]. The substrate preparation consists in a degreasing followed by 10 min in ($HF:H_2O$) (1:10) and then 10 min in hot ($HNO_3:H_2O$) (1:1). Under UHV the sample is heated for half an hour at 800°C and 1 min at 1000°C. Then the temperature is lowered to 300-400°C and an As prelayer is deposited before the beginning of the GaAs growth. The samples consist in a 400 nm GaAs layer grown at 0.4 µm/h with a temperature ramp from 350°C to 450°C. The epilayer is encapsulated in a 150 nm AlAs layer, deposited at 630°C, to prevent GaAs from evaporating during the annealing process.

Long thermal annealing (LTA) has been made in situ at 720°C under As_4 overpressure during one hour. Rapid thermal annealing (RTA) has been performed ex situ on half of another specimen at 870°C for a few-seconds. Thus we have obtained three samples to compare: one as grown, one with ex situ RTA and one with in situ LTA just after growth.

The crystalline quality of the GaAs epilayer is checked by Rutherford Backscattering Spectrometry (RBS) (fig. 1) and Transmission Electron Microscopy (TEM) (fig. 2). Fig. 1a shows that the crystallographic quality of the as grown specimen is much inferior to the bulk GaAs. Fig. 1b indicates an improvement of quality for the LTA sample and Fig. 1c shows the closest signal to the GaAs bulk for the RTA specimen.
The count ratio between random and channelling spectra measured just after surface peaks, χ_{min} confirms these results. Its is equal to 35% for the as grown sample, 8% for the LTA-one and 3.3% for the RTA-one. In this last case, it shows that the surface quality of the sample is comparable to RBS observations of bulk GaAs.

In accordance with RBS observations, it is evident from TEM pictures on figures 2a.b that the dislocations density is largely reduced by RTA. The sample with LTA has a defects density equiva-

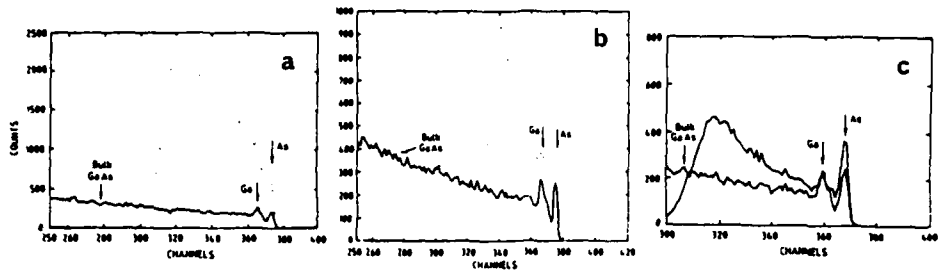


Fig. 1 - 2.3 MeV ion channeling of as grown sample (a), LTA-sample (b) and RTA-sample.

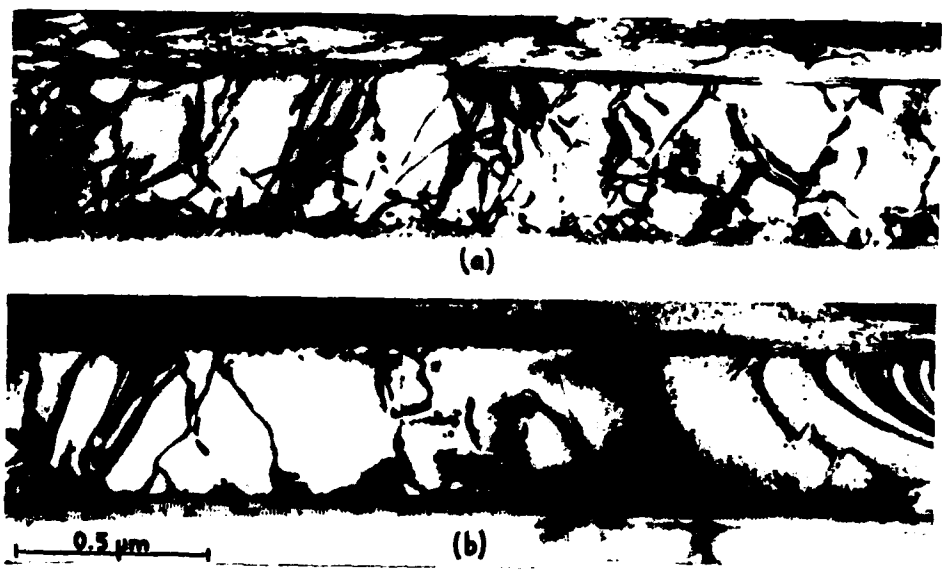


Fig. 2 - Brigh field image of a cross section of as grown sample (a) and RTA-sample (b).

lent to the RTA sample. Such results are in good agreement with those given by CHAND et al. /1/; they prove that the thermal annealing processes considerably improve the crystalline quality.

2 - DISCRETE GaAs DEVICES ON Si SUBSTRATES

2.1 - Solar Cells

Solar cells were among the first devices fabricated in GaAs on Si. The lower density of Si makes it an attractive substrate for GaAs space communications solar cells. As a minority carrier device, solar cells performances are highly dependent on crystalline perfection. Furthermore tensile stresses, related to a large difference in expansion coefficients of GaAs and Si, favors cracking for thick (> 4 µm) layers. The best GaAs on Si cells /4/ yielded 18% efficiency. The dislocation density was 2×10^6 cm^{-3}.

If the dislocation density and cracking could be controlled, cascade cells fabricated with a GaAs cell as top and a Si bottom cell could offer an optimized 38% theoretical efficiency by utilizing at its maximum the solar spectrum.

2.2 - LEDs

Large areas of LEDs have been suggested for printing applications because of the availability of large area, inexpensive Si substrates.

A GaInAsP light emitting diodes grown by LP-MOCVD on Si substrate has been recently reported /5/. The device has been operated CW at 200 mA for several hours without any significative degradation. Clearly, most GaAs on Si LED research is aimed at combined function circuitry such as optical interconnects and will be discussed later.

2.3 - Lasers

Injection layers emitting at room temperature, operating CW were reported for both GaAlAs/GaAs /6/ and GaInAsP/InP structures /7/. In spite of these very impressive successes, the demonstrate performance and reliability tests are not adequate for real applications. The greater silicon's thermal conductivity, may permit upright mounting, and consequently simplifies the technology associated with flip-chip mounting.

Lasers are subject to degradation as soon as one dislocation penetrates the active region. This place the upper limit on the dislocation density at about 10^3 cm^{-2}, 1000 times lower than the best values achieved with GaAs on Si. The residual stress also contributes to degradation. For reliable laser performance, the stress should not exceed 1×10^8 dyne cm^{-2} /8/, which is one order of magnitude lower than the values currently achieved. Both stresses and dislocation density are reduced with patterned substrates. When the substrate dimensions are close to the epilayer thickness, both the stress and dislocation density are reduced /9/. For a stripe layer the stress becomes essentially uniaxial and remains at a high level. If the laser dimensions can be reduced in both directions, then the expected reduction in stress should have a significative impact on reliability and performances.

2.4 - GaAs on Si Optical Waveguides

A lot of work is currently being done in the field of optical intrachip interconnections and therefore on integrated optical waveguides. III-V structures on Si are of wide interest due to the speed and electrooptic properties of multiquantum well (MQW) GaAs/AlGaAs waveguides. A structure consisting of three 100 periods GaAs/AlAs MQW layers has been grown on Si. MQW_1 (GaAs/AlAs respective thicknesses: 90/200 Å), MQW_2 (50/200 Å), MQW_3 (100/100 Å) are respectively the guide, cladding and buffer layers.

The sample cross-section has been observed by Scanning Transmission Electron Microscopy (STEM). Figure 3 shows a picture of MQW_2, which proves the very high quality of the multilayer.

The waveguide has been tested by monitoring the transmission of near infrared intensity filtered by a monochromator. Preparation of the end faces of the waveguide has been obtained by a standard cleaving procedure. The optical waveguide is monomode from 0.86 to 1.2 µm and further. Figure 4 shows the transmission spectra of the sample for two cavity lengths with an excitation with TE mode (similar results have been obtained with TM mode). The oscillations of the curves are caused by the coupling between modes due to birefringence. The propagation loss is about 14 dB/cm at 1 µm in TM mode, which is still higher than for GaAs on GaAs optical waveguide.

Fig. 3 - STEM dark field image of MQW3 cross-section. The mean distance between dislocations is approximately 1 µm.

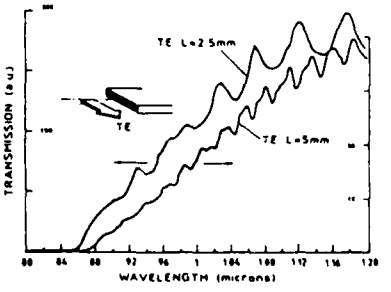
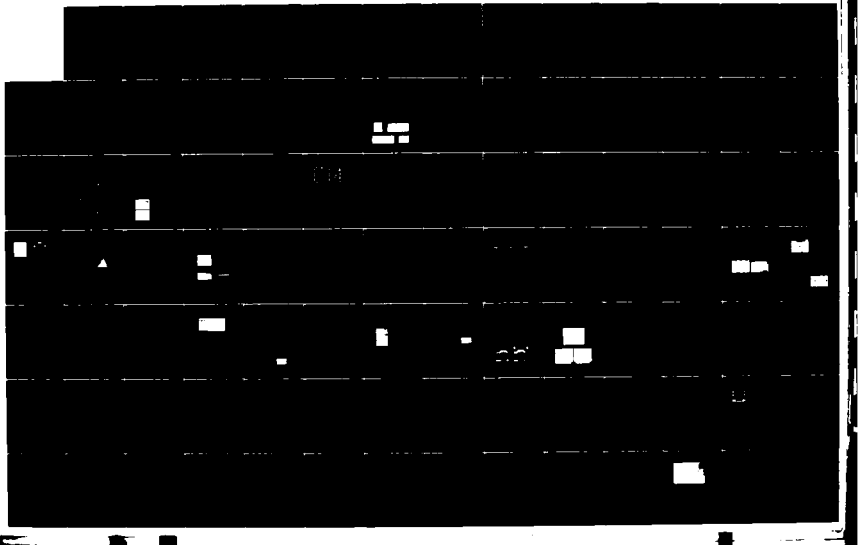
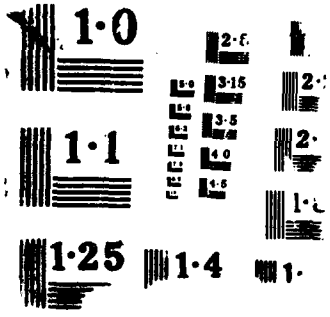


Fig. 4 - Transmission spectra of GaAs/Si waveguide for cavities of 2.5 and 5 mm.

AD-A210 618 EUROPEAN SOLID STATE DEVICE RESEARCH CONFERENCE (18TH) 6/9
ESSDERC 88 HELD IN.. (U) MONTPELLIER-2 UNIV (FRANCE)
J P NOUGIER ET AL. 16 SEP 88 R/D-5983-EE-02
UNCLASSIFIED DAJA45-88-M-0151 F/G 20/12 NL




1·0
1·1
1·25
1·4

2.5 - Field Effects Transistors

The majority carriers devices are relatively unsensitive to dislocations. The best GaAs on Si FETs present little degradation in performance from that expected from devices fabricated on GaAs substrates. Generally, FETs in GaAs-on-Si exhibit transconductances 80% of the values from all GaAs transistors. However, GaAs FET operation depends on high resistivity underlying substrate and buffer layer. In comparison with semi-insulating GaAs Si substrates, up to 4 inches, have a resistivity at least three order of magnitude below. The status is even worth over 4 inches. For GaAs on Si FETs applications, there is a requirement of a thick epitaxial high resistivity buffer layer. For buffer layers thinner than 3 μm, degradation of leakage currents and backgating effects are significative. Cut off frequencies up to 55 GHz have been measured with GaAs on Si MESFETs /10/. GaAs-on-Si has been proposed as a promising material for power FETs, to take advantage of the superior thermal conductivity of Si. Measurement of the thermal assistance under the gate attested of an improvement by a factor ~ 2 /11/. Simplified technology could be achieved for GaAs-on-Si power FETs on highly conductive Si substrates /12/.

Also it has proposed that GaAs-on-Si might be an attractive material by radiation resistance because carriers generated within the substrate by an ionizing particle may be trapped by dislocations close to the hetero interface.

3 - GaAs INTEGRATED CIRCUITS ON Si SUBSTRATES

3.1 - Digital Integrated Circuits - MESFET

T.I has demonstrated /13/ a functional 1K static RAM containing 7500 transistors. Threshold voltages both for enhancement and depletion mode MESFETs presents dispersion close to the devices fabricated entirely by an implantation on a GaAs substrate.

Wafer having both before and after implantation annealing were not sufficient to impede fabrication of such a dense circuit. If the access time are only slightly slower than those on circuits fabricated simultaneously on control GaAs wafers, the circuit yield (~ 10%) was much lower.

3.2 - Digital Integrated Circuits - Bipolar

Fifteen stage ring oscillators and ROMs have been fabricated in GaAlAs/GaAs on Si using the difficult emitter-down (HI^2L) configuration that allows high-density integration /14/. All layers were grown by MBE, the p^+ base regions were defined by ion implementation. In this HI^2L configuration, the emitter-base interface is the Si-GaAs hetero-interface and the current path is through the imperfect, Si/GaAs heterojunction. The observation of identical emitter-base characteristics for GaAs-on-Si and the all GaAs control devices indicates that the dislocations threading from the Si substrate do not appreciably degrade the performance of these heterojunction bipolar transistor (HBTs).

A novel concept of GaAs-Si heterojunction bipolar transistor that combines a highly developed Si bipolar technology with a GaAlAs emitter /15/. The device would require for motion of a critical GaAs/Si emitter base interface. Recent results at THOMSON /16/ concerning localized growth of GaAs on submicron SiO_2 windows provide encouragement for further investigation.

Calculations indicate an f_{max} of 108 GHz the AlGaAs-on-Si HBT as compared with 76 GHz for the conventional AlGaAs/GaAs HBT.

3.3 - Monolithic Microwave Integrated Circuits

Monolithic microwave integrated circuits (MMICs), consisting of GaAs components connected with metal microstrip transmission lines on a semi-insulating substrate, promise to be an important, large scale application for GaAs. Although these circuits are not densely integrated, they require large chips because of on-chip matching components such as inductors. The promise of using large diameter GaAs-on-Si wafers, which are resistant to breakage and have superior heat dissipation characteristics, is very attractive; but can the silicon substrate's relatively high electrical conductivity might be a problem. Indeed, the availability of GaAs in a high resistivity (10^8 ohm cm), semi-insulating form was a major factor in developing MMIC technology. Early work /17/ on the use of Si as a transmission medium for hybrid circuits showed that of the two transmission line loss mechanisms - substrate loss and conductor loss - the conductor losses dominate at higher frequencies (> 30 GHz) as long as the substrate resistivity exceeds 2000 ohm cm. At mm-wave frequencies (90 GHz) the attenuation of microstrip lines on Si with a resistivity of 10,000 ohm cm can be quite low (0.6 dB/cm) /18/.

The active elements in MMICs are usually majority-carrier MESFETs which are easily fabricated with acceptable performance in GaAs on Si. Application of GaAs on Si to MMICs then depends on the availability of high resistivity silicon. Such material is grown by the float zone process and traditionally has been available only in sizes less than 125 mm. Although float zone growth of 150 mm Si crystals is believed to be feasible, the cost may be excessive in comparison with the expected advantages of the larger wafers. Wafers of intermediate size (100 mm) would still be attractive for potential MMIC applications, but there remains an additionnal potential problem. High resistivity silicon is subject to resistivity reductions during high temperature processing. Fortunately, recent results /18/ indicate that Si with a resistivity as high as 10,000 ohm cm can be stable at diffusion and thermal oxidation temperatures.

A single-stage MMIC feedback amplifier in GaAs on Si has been tested /19/. Operating between 8.5 and 11.5 GHz, the GaAs-on-Si MMIC showed a lower gain and smaller bandwidth than the GaAs-substrate control. However, computer simulations showed that the lower performance was due mostly to a leaky Schottky gate on the GaAs-on-Si test circuit and that the lower resistivity of the Si substrate (7000 ohm cm) had only a negligible effect on the microwave performance. No resistivity degradation occurred during growth of the GaAs layer or during subsequent processing. The superior thermal conductivity of the Si substrate was evident in thermal resistance measurements which indicated that the junction temperature of the GaAs-on-Si MMIC was about 15°C lower than the all-GaAs control.

4 - MONOLITHIC INTEGRATION OF GaAs and Si

The most exciting potential of GaAs-on-Si occurs in applications which combine functional components for both GaAs and Si in a single chip. This can be an emerging technology that couples the high components density and process sophistication of Si with the high speed, high frequency and optoelectronics capability of GaAs. Significative technology evaluators examples of such integration include intra- and interchip optical interconnects, GaAs MMICs with Si signal processors and Si high density CMOS coupled with high speed GaAs processors.

Success depends on the development of special processing regimes, *e.g.*, growth of the GaAs on selected regions of a preprocessed silicon wafer containing fabricated devices or circuits. Because silicon processing generally requires higher temperatures than GaAs, the silicon circuits must be formed (at least up to metallization) before GaAs epitaxial deposition, during which the preprocessed Si circuitry must be protected by a inert mask. Ultimately, interconnects must be established between the Si and GaAs components. This requires that the upper surface of the GaAs layer be nearly coplanar with the Si surface (or the dielectric layer covering it). The approach being investigated to achieve such coplanarity employs deposition of the GaAs selectively into depressions or "trenches" etched to the optimum depth within the Si substrate /14/. When the GaAs is deposited by MBE, a polycrystalline film is simultaneously formed over the mask that protects the Si and defines the desired single-crystal GaAs area. The polycrystalline material must be removed during subsequent processing.

Deposition of GaAs into trenches etched into Si is fraught with potential problems. As mentioned previously, tilted substrate orientations are desired to minimize threading dislocations and eliminate antiphase domains; however, multiple orientations, exposed within the trench, are subject to simultaneous deposition. Depending on the trench etching process, the original tilted surface orientation may not be preserved at the bottom of the trench. The severity of these potential difficulties is being experimentally assessed.

4.1 - High Speed GaAs with High Density Si

Successful operation of a preliminary integrated circuit has been demonstrated with a composite ring oscillator containing 35 Si CMOS inverter stages and 12 GaAs MESFET buffered logic inverter stages all operating in concert /20/.

4.2 - Optoelectronic GaAs with Si

A device-level demonstration vehicle integrating a AlGaAs/GaAs light emitting diode with a Si MOSFET fabricated in the underlying substrate has already been tested /21/. The LED was modulated at rates up to 27 Mbit/s (limited by the large area of the particular Si test device used). But serious applications will require integration of lasers to obtain the requisite output powers and bandwith, and the barriers to laser performance and reliability imposed by the high dislocation density and stress of GaAs on Si must be circumvented. Presently the stresses measured in GaAs on Si are around 10^9 dyne cm^{-2}, about an order of magnitude greater than considered acceptable for lasers.

5 - CONCLUSION

We have reviewed the state-of-the-art of the MBE and MOCVD growth of GaAs directly on Si substrates and of GaAs devices that have so far been fabricated. We have shown the tremendous progress made in solving the major problems of the growth of GaAs on Si. Striking improvements have been made with regard to misfit dislocation control, and the achievement of reproducible GaAs surfaces with very low dislocation densities (i.e., 10^3 cm^{-3} range) is now in progress.

We have also shown that many of the important GaAs devices have already been grown on Si, with performance identical to their counterparts on GaAs for the majority carrier devices and with good performance for the minority carrier devices. In light of the rapidly increasing number of industrial groups that are now working on the fabrication of GaAs devices on Si substrates, one can consider that a first phase of the GaAs on Si research is completed and that the feasibility of this technique is now established. Further efforts should be directed toward two areas: i) the conception of new devices actually using the hybridization of GaAs and Si technology, ii) the extension of this new heteroepitaxy to other III-V compounds such as GaSb or InP and to II-VI compounds /22/.

REFERENCES

/1/ Chand, N., People, R., Bariocchi, F.A., Wecht, K.W. and Cho, A.Y., Appl. Phys. Lett. 49, 815 (1986).

/2/ Hull, R., Rosner, S.J., Koch, S.M. and Harris Jr., J.S. Appl. Phys. Lett. 51, 36 (1987).

/3/ Charasse, M.N., Georgakilas, A., Chazelas, J., Hirtz, J.P., Blanck, H., Heral, H., Rocher A., Siejka, J. and Ortega, C. Proceedings of E-MRS Meeting, June 1987, Vol. XVI, p. 343 (les Editions de Physique, Paris 1987).

/4/ Yamamoto, A. and Yamaguchi, M., to be published, Mat. Res. Soc. Symp. Proc. (1988).

/5/ Razeghi, M., Blondeau, R., Defour, M., Omnes, F., Maurel, P., to be published, APL (1988).

/6/ Turner, G.W., Choi, H.K., Mattia, J.P., Chen, C.L., Eglash, S.J. and Tsaur, B.-Y., to be published, Mat. Res. Soc. Symp. Proc. (1988).

/7/ Razeghi, M., Dufour, M., Omnes, F., Maurel, Ph., Chazelas, J., Brillouet, F., Applied Phys. Lett. Aug. 1988.

/8/ Goodwin, A.R., Kirkby, P.A., Davies, I.G.A., and Baulcomb, R.S., Appl. Phys. Lett. 34, 647 (1979).

/9/ Lee, H.P., Wang, S., Huang, Y-H., and Yu, P., Appl. Phys. Lett. 52, 215 (1988).

/10/ Aksun, M.I., Morkoc, H., Lester, L.F., Duh, K.H.G., Smith, P.M., Chao, P.C., Longerbone, M. and Erickson, L.P., Appl. Phys. Lett. 49, 1654 (1986).

/11/ Akiyama, M., Kawarada, Y., Weda, T., Nighi, S., Kaminishi, K., J. Cryst. Growth 77, 490, (1986).

/12/ Hirtz, J.P., to be published.

/13/ Shichijo, H., Lee, J.W., McLevige, W.V. and Taddiken, A.H., Proceed. 1986 Int. Symp. GaAs and Related Compounds, Inst. Phys. Conf. Ser. 83, 489 (1987).

/14/ Shichijo, H., Tran, L.T., Matyi, R.J. and Lee, J.W., Mat. Res. Soc. Symp. Proc. 91, 201 (1987).

/15/ Chen, J., Unlu, M.S., Morkoc, H. and Verret, D., Appl. Phys. Lett., 52, 822 (1988).

/16/ Charasse, M.N., Bartenlian, B., Hirtz, J.P., Chazelas, J., NATO Int. Conf. on heteroepitaxy on Si, Cargese, 15-21 May 1988.

/17/ Rosen, A., Caulton, M., Stabile, P., Gombar, A.M., Janton, W.M., Wu, C.P., Corboy, J.F. and Magee, C.W., RCA Review 42, 633 (1981).

/18/ Buechler, J., Kasper, E., Russer, P. and Strohm, K.M., IEEE Trans. Elec. Dev. ED-33, 2047, (1986).

/19/ Eron, M., Taylor, G., Menna, R., Narayan, S.Y. and Klatskin, J., IEEE Electron Device Lett. EDL-8, 350 (1987).

/20/ Shichijo, H., Matyi, R.J. and Taddiken, A.J., to be published, IEEE Electron Device Lett., (1988).

/21/ Choi, H.K., Turner, G.W., Windhorn, T.H. and Tsaur, B.-Y., IFEE Electron Device Lett. EDL-7, 500 (1986).

/22/ Bean, R.C., Zanio, K.R., Hay, K.A., Wright, J.M., Seller, E.J., Fischer, R., Morkoc, H., J. Vac. Sci. Technology, A4(4), 2153 (1986).

Session 5A Room A : Metallisation and shallow junctions

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

TiW AND Al CONTACTS TO SHALLOW p⁺ JUNCTIONS - A COMPARISON BETWEEN FURNACE AND RAPID THERMAL ANNEALING (RTA)

P. WIKLUND, U. WENNSTRÖM, H. NORSTRÖM, I. SUNI* and A. LINDBERG**

Ericsson Components, S-163 81 Stockholm, Sweden
**Techn. Res. Centre of Finland, SF-02150 Espoo, Finland*
***Inst. of Microelectronics, S-164 21 Stockholm, Sweden*

RESUME

L'utilisation de TiW (15/85 w.%) en contact direct avec du silicium p⁺ a été évaluée. La resistance des contacts a été étudiée en function de traitements thermiques différents, recuits rapides (RTP),ou recuits conventionels au four en atmosphère sèche ou humide. Des mesures parallèles ont été faites simultanément sur des échantillons metallisés avec de l'aluminium.
Les profiles de diffusion des éléments dopants ont été déterminés par mesure électrique (SRP). La hauteur des barrières de Schottky a été mesurée sur de silicium de type n. Les valeurs déterminées sont 0.55 eV pour TiW et 0.8 eV pour l'aluminium. Les résultats montent qu'il est nécessaire d'avoir des concentrations superficielles très élevées a fin d'obtenir des resistances de contact faibles.
Ces effets sont plus accentués pour l'electrode characterisée par une plus haute barrière de Schottky avec le silicium p⁺. Ainsi les contacts TiW/Si sont très sencibles à des petites variations de concentration de Bore. En dépit de ces observations, pour une concentration superficielle donnée, apres recuit rapide (RTP) les preformances des contacts TiW/Si p⁺ sont meilleures que celles des contacts Al/Si p⁺.
Dans le cas de ces contacts des précipitations de silicium réduisent l'aire effective des contacts. Il en résulte une augmentation des resistance de contact qui est d'autant plus marque que les dimensions sont réduites. Avec le TiW ces précipitations sont éliminées.

ABSTRACT

The use of sputter deposited TiW (15/85 w.t.%) as a direct contact material to p⁺ silicon is evaluated. The effect on the contact resistance of different post-implant activation conditions, e.g. rapid thermal processing and conventional furnace annealing in dry or wet ambients, has been studied. For comparison parallell measurements were done on identically processed wafers metallized with aluminium.
The resulting dopant profiles were determined by spreading resistance profiling (SRP). Schottky barrier height measurements were performed on n-type wafers metallized with TiW or Al. The respective values were Φ_{Bn}= 0.55 eV and Φ_{Bp} = 0.8 eV. The results from contact resistance and SRP measurements show that a high surface carrier concentration is necessary in order to achieve a low contact resistance.
This effect was more pronounced for the material with the higher barrier height to p⁺ silicon i.e. TiW was found to be more sensitive to a small variation in the boron surface carrier concentration. Despite of this it was observed that given a high surface carrier concentration, after RTP , the performance of TiW contacts to p⁺ silicon surpassed that of Al. In case of a direct Al-Si contact, silicon precipitation occurs in the contact hole thereby reducing the effective contact area. This will cause an increase in the resistance of contacts which becomes more pronounced for small dimensions. With TiW ,silicon precipitation is avoided.

INTRODUCTION

In modern IC-technology dimensions are approaching 1 micrometer. The reduced size of contact holes may result in unacceptably high contact resistances. A method often suggested to circumvent this problem is to form a silicide on the source and drain areas of the transistor. Alternatively, the specific contact resistivity can be lowered by increasing the dopant activation, since the resistance of a metal to semiconductor contact is a function of the active dopant concentration and the Schottky barrier height.
To maintain a high degree of dopant activation in a shallow junction, it is

necessary to use a relatively high temperature for reduced annealing times These demands are difficult to meet with conventional furnace annealing. However they can be met with the RTA technique (1). In boron implanted silicon the transient enhanced diffusion of dopants is severe below 900C (2). At low temperatures the dissolution of small defect clusters proceeds slowly thereby increasing the transient period and the resulting junction depth. The clusters, thought to act as primary source of interstitials, remain stable at 600C. Therefore the low temperature annealing step does not affect the diffusion tail of the final impurity profile. For complete defect removal, temperatures in the order of 1100C are needed (3). At this temperature a few seconds are sufficient to fully activate the implanted dopant atoms. For such annealing conditions the transient broadening of the impurity profile should be limited to tens of nanometers only.
We have compared conventional furnace annealing with RTA by measuring doping profiles, dopant activity and the resulting contact resistance propoerties of p+ contacts to silicon. In atechnology with shallow source/drain junctions, it can be necessary to improve the reliability of the Al/Si system by introducing a barrier metal, e.g. TiW, that prevents excessive reaction between Al and Si (4). Therefore we have carried out our tests on both Al/p+-Si and Al/TiW/p+-Si contacts.

EXPERIMENTAL

N-type wafers of 15-25 Ωcm resistivity (100) oriented were used as starting material. Active device areas were patterned by means of conventional LOCOS technique. After field oxidation a 200 Å thick screen oxide was grown. The samples were then implanted with BF_2 with a dose of 5 10^{15} cm^{-2} at 50 keV. A low temperature anneal at 600 °C for 25 minutes was used to recrystallize the damage caused by implantation before 4000 Å of undoped TEOS-oxide was deposited on the wafers.
The implanted species were then activated by Rapid Thermal Annealing (RTA) in an AG 210T, Heat-pulse equipment. The effect of different annealing time and temperature was studied. For sake of comparison a couple of as-implanted wafers were furnace annealed at 925 °C in either a dry or a wet ambient. Contact windows were defined by standard photolithography and opened by RIE in a mixture of CHF_3 and O_2.
Prior to metallization the wafers were subjected to a dip in buffered HF to remove any native oxide in the contact windows. Metallization was performed in an Electrotech MS 6200 D.C.magnetron sputter.
A trilayer structure consisting of 1500Å TiW/3400Å Al(Si,Cu)/1100Å $MoSi_x$ was sputter deposited on the wafers (4). The TiW-film, serving as a direct contact to silicon, was deposited from a target containing 15 % titanium by weight. The as-deposited TiW-film was found to contain 24 atomic percent of titanium as determined by Rutherford Backscattering Analysis (RBS). For reason of comparison a bilayer of Al(Si,Cu)/$MoSi_x$ was deposited on a few identically processed wafers to serve as a set of references.

The structures used for evaluation of the contact properties were four terminal Kelvin resistors with three different contact sizes viz 5x5, 3x3 and 2x2 µm. The actual doping profiles were determined by Spreading Resistance Profiling (SPR) technique.
The LOCOS technique was also used to define circular patterns for measurements of the metal-semiconductor barrier height. In this case N-type silicon wafers with a 10 µm thick epi layer (15-25 Ωcm) on top of a bas layer of 0.001 Ωcm was used in order to reduce the influence of the series resistance. These wafers were also aluminium metallized on the backside. The barrier height for both Al(Si,Cu) and TiW to silicon was determined with the C-V technique and the Norde plot (5).
Prior to electrical evaluation all samples recieved an hydrogen anneal at 400 °C for 70 minutes.

RESULTS

Figure 1 shows a comparison between boron profiles obtained by RTA and conventional furnace anneal. The RTA was performed at 1125C during 5s in inert ambient and the furnace anneal was done at 925C for 30 min in N_2, wich for both samples resulted in similar junction depths. The overall as well as the maximum active dopant concentration was a factor of 3 higher for the sample

where RTA was used. The corresponding sheet resistances measured with four-point probe was 42 Ω/square for the RTA sample and 126Ω/square for the furnace annealed sample. The dopant distibution is also markedly different, with the RTA profile showing considerably less dopant depletion towards the surface.

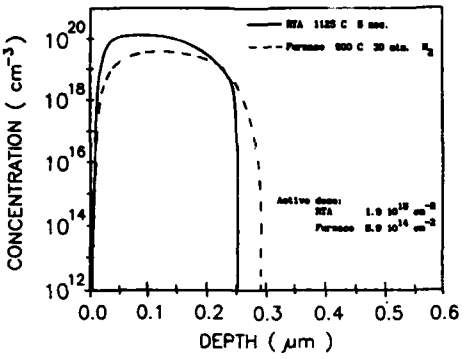


Fig.1: Boron profile for conventional anneal and RTA

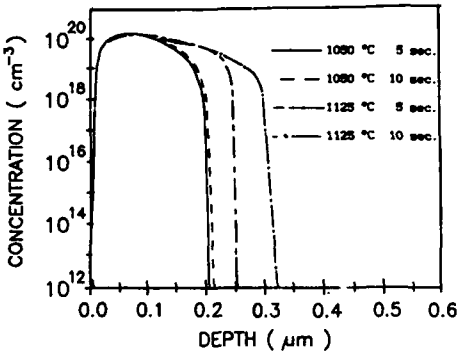


Fig.2: Boron profile obtanied by SRP for RTA samples anneald at different temperatures and times.

The recorded doping profiles after thermal treatment at two different temperatures and process times are shown in figure 2. From this figure it is evident that the maximum boron surface concentration is almost unaffected by a temperature increase from 1050C up to 1125C. This was further verified by contact resistance measurements made on Kelvin structures. The recorded contact resistivity values exhibited no dependence on the RTA temperature. This indicates that the surface carrier concentration remains almost constant throughout the studied temperature interval. The junction depth is however pushed deeper into the silicon bulk as the tmperature and time is increased whereas the electrical activity seems to be restricted by the maximum carrier concentration in silicon (6).
The contact resistance is an exponential function of the barrier height divided by the square root of the surface carrier concentration (7). It is therefore of prime interest to measure the barrier height for the respective contact system considered before interpreting the recorded resistivity data. By RBS, the as deposited TiW film was found to contain approximately 24 at.% Ti. The Schottky barrier height of this alloy at room temperature on n-type silicon, was found to be Φ_{Bn} ~ 0.55 - 0.57 eV as determined by capacitance measurements or the Norde plot (5).
Similar measurements on aluminium Schottky diodes gave a barrier height value of approximately Φ_{Bn} ~ 0.8 eV. The results are listed in table 1 below.

	Norde plot		C-V
	Barrier height (eV)	Ideal. factor	
Al/MoSi$_x$	0.838	1.13	0.88
TiW/Al/MoSi$_x$	0.870	1.12	0.85

Table 1: Schottky barrier height data of TiW and Al on n-type silicon.

	FURNACE		RTA
	925°C wet O_2	925°C N_2	1125 °C
	15 min.	15 min.	5 sec.
Al/p$^+$-Si	$2.5\cdot10^{-6}$	$2.2\cdot10^{-6}$	$7.0\cdot10^{-7}$
Al/TiW/p$^+$-Si	$5.5\cdot10^{-6}$	$2.4\cdot10^{-6}$	$8.0\cdot10^{-7}$

Table 2: Contact resistivity for different annealing conditions. All values are in Ωcm^2.

The recordede barrier height of the TiW alloy is in close agreement with the reported data by Aboelfotoh (8), using the forward I-V technique.
Table 2 gives a comparison between the recorded contact resistivities for MoSi$_x$/Al/p+ Si and MoSi$_x$/Al/TiW/p+ Si at different annealing conditions.

Furnace anneal in wet O_2, which is often used in standard processing schemes to enable flow of phosphorous glass, gives a higher contact resistivity than anneal in N_2 ambient. It is clearly seen that RTA results in lower contact resistivities for Al contacts than furnace annealing does. Due to higher dopant activation at the surface, values below 10^{-8} Ωcm^2 are achieved. The measurements on Al/TiW/p+-Si contacts yielded similar results in terms of improvement of contact resistance when RTA was applied. As the Schottky barrier for TiW is higher than that for Al on p+-Si, the corresponding contact resistivities were also higher. However, when RTA was applied for the TiW samples, contact resistivities below 10^{-8} Ωcm could be reached. The contact resistance as a function of the contact size, for differently processed wafers, is depicted in fig. 3. The highest resistance values are observed for TiW contacts to p+ areas which have been activated in a wet ambient at 925C. If however the gas mixture is altered so that annealing occurs at 925C in a dry ambient there is less influence of the contact material on the contact resistance as the surface doping level now has increased. It is worthwhile to notice that within this group 2x2μm contacts metallized with Al(CuSi)/$MoSi_x$ exhibits a higher resistance value than similar contacts metallized with TiW. This is likely to be a result of an enlarged silicon precipitation in the former case. Furthermore it is interesting to note that pure Ti-contacts to p+ areas, annealed in a dry atmosphere, results in a lower resistance than for corresponding TiW contacts, although the barrier heights of TiW correspond to that of pure Ti (8). A likely explanation for this is that Ti in the TiW alloy is less reactive than pure Ti and therefore less efficiently reduces the native oxide in the contacts.
Rapid thermal processing further lowers the contact resistance for both Al and TiW to p+ silicon as a result of an enlarged surface carrier concentration. As Al(Cu,Si) contacts to silicon are prone to cause silicon precipitation, which reduces the active contact area, it is not surprising that, in this latter case, TiW contacts result in a lower contact resistance value than Al.

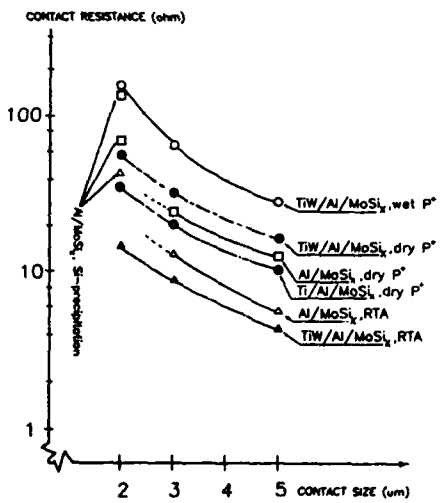


Fig.3: Contact resistance as a function of different contact sizes obtained under various process conditions.

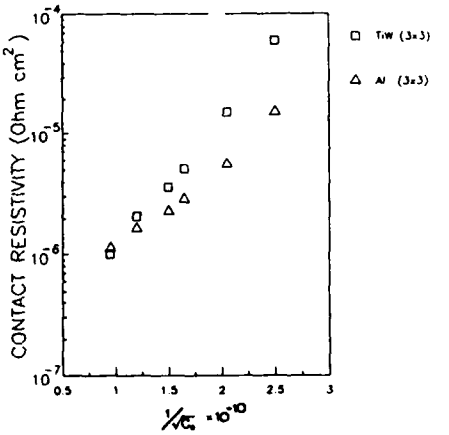


Fig.4: The contact resistivity as a function of the inverse of the square root of the surface carrier concentration.

In fig.4 contact resistivity is plotted versus the inverse of the square root of doping. It is observed that for the lower values of surface carrier concentration, the higher Schottky barrier of TiW, results in a contact resistivity three times the value achieved for Al. However, as the surface electrical activity of the boron increases and exceeds 1 $10^{20}cm^{-3}$, the resistivities for TiW and Al converges and reaches values below 10^{-8} Ωcm^2., Due to the high effective dopant concentration, the band bending is strong and the influence of the barrier height on the contact resistivity is reduced.

CONCLUSIONS

The contact properties of TiW and Al to shallow junctions in silicon have been studied by contact resistance measurements and SRP. The Schottky barrier height of TiW and Al to n-type silicon was also measured and found to be 0.55 eV and 0.8 eV respectively. Due to the higher barrier of TiW to p+ silicon compared with Al, the contact resistance of TiW to p+ is more sensitive to variations in surface dopant concentration. This is more pronounced when furnace annealing is used because the dopant activation is comparably low. In this case the absolute contact resistances are also higher for TiW to p+ than for Al to p+.
However, when RTA is used, a higher temperature can be tolerated, which yields a higher degree of dopant activation. This also causes the contact resistance of TiW to decrease substantially and reach the values achieved with Al.
In addition, the introduction of a TiW barrier layer prevents silicon precipitation that reduces the active contact area, giving unacceptably high contact resistances.

REFERENCES

(1) S.S.Cohen et al.; J.Appl.Phys, 57(1985), pp.1200

(2) T.O.Sedgwick, A.E.Michel, V.R.Deline, S.A.Cohen; J.Appl.Phys. 63, 1452(1988)

(3) A.Usami, M.Katayama, Y.Tokuda, T.Wada; in *Materials Issues in Silicon Integrated Circuit Processing*, MRS Symposia Proceedings Vol.71 (Materials Research Society, Pittsburgh 1986)p.211

(4) A.Lindberg, M.Östling, H.Norström, U.Wennström; Proc.17th ESSDERC -87 pp.191

(5) H.Norde; J.Appl.Phys. 50(1979)

(6) H.Ryssel et al.; J.Appl.Phys, 22(1980) pp.35

(7) S.M.Sze; Physis of Semiconductor Devices 2nd edition

(8) M.O.Aboelfoto; J.Appl Phys, 61(1987)pp.2558

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

RELIABILITY ASPECTS OF VLSI METALLISATION WITH DIFFUSION BARRIERS

G. RÖSKA and F. NEPPL

Siemens AG, Corporate Research and Development, Microelectronics, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

Abstract -

In this work we show the necessity of diffusion barriers for scaled VLSI and investigate the resulting changes of reliability risks. Accelerating stress tests at relevant test structures are presented and discussed in detail.

1 - INTRODUCTION

Diffusion barriers between aluminum based metal interconnects and the highly doped mono silicon source/drain regions (S/D) of MOS transistors are necessary for reliable contacts of 1µm² and below. The barrier decouples aluminum and Si substrate and thus reduces the contact resistivity and prevents aluminum spiking even at the high current densities of scaled VLSI technologies. In this paper we examine contacts with and without diffusion barriers by reliability experiments and compare the various reliability risks of typical MOS transistor contacting schemes in detail.

2 - EXPERIMENTAL

For wafer preparation phosphorus was implanted into p-Si and thermally activated to form the highly doped n^+ regions. The junction depth was about 0.5µm. The contact holes were etched by RIE. For interconnect metal Al-(1wt%)Si-(0.2wt%)Ti /1/ was D.C. magnetron sputtered. The thickness was 0.8µm. As diffusion barriers co-sputtered $TaSi_2$ /2/ or magnetron sputtered Ti/TiN /3/ was used. The barrier was patterned together with the interconnect metal. The wafers were finally annealed at 450°C for 30 minutes in H_2. Afterwards the samples were Al wedge bonded and encapsulated in ceramic packages.

The test structures were different contact stripes containing quadratic and long contacts. The structures are especially designed for reliability experiments.

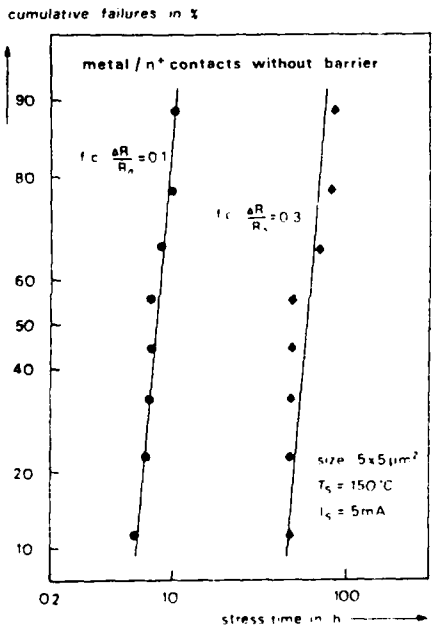
The accelerating stress tests were done at an ambient temperature of 150°C and at various current densities. As failure criterion resistance increase, metal open or a leakage current of > 10^{-10}A through the n^+/p diodes were considered.

To have a reference in terms of metal line reliability, stress tests were carried out with 500µm long metal interconnects using $2*10^6$A/cm² and 150°C. Here the failure criterion was a 50% increase of the line resistance.

3 - COMPARISON: WITH/WITHOUT BARRIER

To compare the reliability of contacts with and without a diffusion barrier contact chains of some tens of contact couples were used. The contact sizes were 5x5µm² and 0.9x0.9µm².

Contacts between metal and substrate can fail by aluminum spiking, resistance increase, open at the contact edges or interruptions of the interconnects near the contacts. Contacts without a diffusion barrier degrade especially by mechanisms connected to the aluminum/mono silicon interface. For metal/n^+ contact chains even moderate stress increases the resistance drastically /4/. In figure 1 the resistance increase $\Delta R/R_o$ is used as failure criterion. Since experiments with metal/p^+ contacts don't show this increase it is attributed to epitaxial regrowth of Al (p)-doped Si out of the aluminum/silicon alloy /5/. After about 2000h stress time the examined metal/n^+ contacts were so strongly degraded by aluminum spiking that they finally failed by leakage current even though junction depth was relatively high (0.5µm).



	without barrier	with diffusion barrier	
contact size in µm²	5 x 5	5 x 5	0.9 x 0.9
stress cond.			
I/contact in mA	5	12.5	5
T in °C	150	150	150
MTTF in h $\left(\frac{\Delta R}{R_0} = 0.3\right)$	60	>6000	> 6000

Fig. 1 - Degradation of metal/n^+ contacts without diffusion barrier by resistance instabilities at stress conditions of 5mA/(5*5µm² contact) and 150°C. A comparison with contacts containing diffusion barriers is added by the table right. (f.c.: failure criterion)

For comparison contact chains with diffusion barrier were examined (table in Fig. 1). For the same contact size of 5*5µm² the mean time to failure (MTTF) is more than 2 orders of magnitude higher although the stress current is increased by a factor of 2.5. Using the same stress current of 5mA also 0.9*0.9µm² contacts with barrier were tested. The MTTF was also at least 2 orders of magnitude higher and no failure by leakage current was detected.

As expected, introduction of a diffusion barrier eliminated the Al/Si interface as major reliability problem. Since $TaSi_2$ or Ti/TiN barrier has an excellent step coverage, also excessive metal thinning at contact edges due to poor step coverage is widely avoided. Therefore the narrow interconnect line might become the weakest link in the interconnect/contact system.

4 - LAYOUT RELATED EXPERIMENTS

To investigate this possibility and to take into account various types of layouts, the test structures of Fig. 2 were stressed. The peculiarity is a fixed contact area which can be used for only one single long contact (type A) or many quadratic minimum size contacts at minimum distance (type B). The latter approach is often taken to eliminate photolithography problems. The length of the contact areas are always 12µm. The width is varied by 3µm and 1.3µm in spite of a different size and number of the quadratic contacts. Metal/contact overlap and diffusion region/contact overlap were 0.8µm for all test structures. The test structures allow to seperate failures caused by poor step coverage and narrow line width, because the left parts should only show step coverage problems but the right parts both of them.

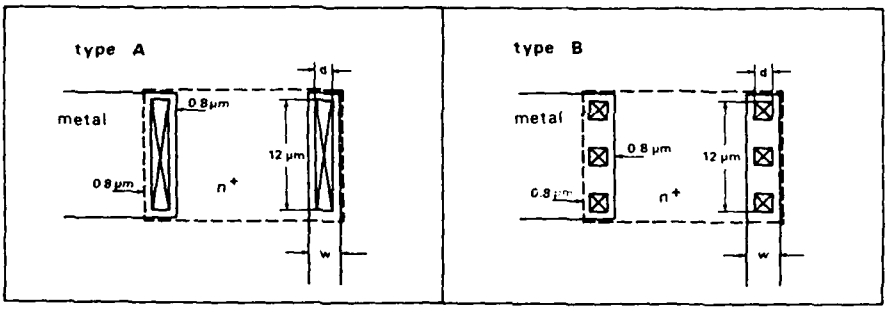


Fig. 2 - Test patterns to emulate 4 typical possibilities of source/drain contacting with long contacts (type A) and many minimum size contacts at minimum distance (type B).

For the accelerating stress tests at a stress current of always 60mA the failure criterion was metal open. As diffusion barrier Ti/TiN was used. Failures were only observed at the right part of the test structure (Fig. 2). and were located at the narrow metal interconnects near the contact regions. This proves that no step coverage but interconnect problems are present.

Therefore interconnects without contacts but $TaSi_2$ and Ti/TiN as underlayers were examined by an accelerating stress test (Fig. 3b). We found a strong line width dependence of the MTTF similar to results without an underlayer /6/. In comparison to interconnects of only AlSiTi the total open of the lines happens later but the resistance increases drastically. Therefore a circuit related upper limit must be defined as failure criterion - in our case it was 50%. As shown in figure 3b an underlaying material can influence the MTTF of the interconnects significantly.

As expected from figure 3b the MTTF of the interconnects leading to the contacts confine qualitatively the MTTF of the whole contact test structures (Fig. 3a). Although the current density for the 1.3µm wide contacts is substantially higher, for the interconnects it is about a factor of 2, the test structures with the wider contacts and with the larger interconnect widt w (Fig. 2) fail first.

Not well understood ist the MTTF decrease for the minimally designed quadratic contacts compared with the single long contacts. This is also true for the contacting areas with 6 contacts of $1.3*1.3\mu m^2$. The MTTF is nearly the same as of the $3*12\mu m^2$ contacts. On the first view the MTTF of both geometries should not differ, because the interconnect width w (Fig. 2) is the same. But the significant difference in lifetime show that the influence of current distribution and temperatur gradient must be taken into account. Therefore also the results of the figures 3a and b agree only qualitatively.

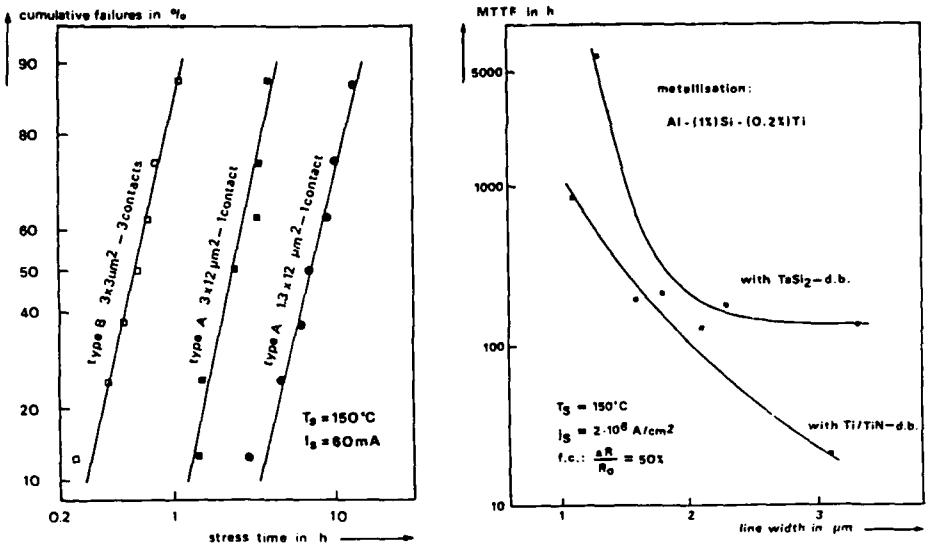


Fig. 3(a) - Stress time dependence of the cumulative failure of test structures described in fig. 2. The stress conditions were 60mA and 150°C in each case and the failure criterion is metal open.

Fig. 3(b) - Interconnect line width dependence of the MTTF of 0.8µm thick AlSiTi alloy layered on $TaSi_2$ and Ti/TiN.

5 - CONCLUSION

Effective diffusion barriers eliminate the aluminum/mono silicon interface as reliability risk widely. If the barrier additionally ignores step coverage problems, which can be achieved by an improvement of the aspect ratio by selective tungsten for example /7/ or a good step coverage of diffusion barriers like $TaSi_2$ or Ti/TiN, the interconnects close to the contacts become critical. Therefore the effects of the diffusion barriers on the reliability of the interconnects are also a problem of the contact system. If the interconnects get narrower and thus much more reliable the interface finally might become the weak spot again.

It was shown that minimum sized contacts favored by the layout have a lower lifetime than long contacts. On the other hand long contacts might require 0.2µm larger overlaps because of contact widening. For an optimal contact geometry both aspects must be considered.

6 - ACKNOWLEDGEMENT

We want to thank our colleagues for preparation of the wafers. This work was supported by the Federal Department of Research and Technology of the Federal Republic of Germany. The authors alone are responsible for the contents.

7 - REFERENCES

/1/ F. Fischer, F. Neppl, Proc. 22nd IRPS, p. 190 (1984)
/2/ F. Neppl, F. Fischer, U. Schwabe, Proc. 22nd IRPS, p. 185 (1984)
/3/ C.Y. Ting, M. Wittmer, Thin Solid Films, 96, 327 (1982)
/4/ G. Röska, H. Haudek, F. Neppl, Proc. 15th ESSDERC, Vol. 9H, p. 324 (1985)
/5/ G. Röska, v. Marinelli, F. Neppl, to be published (1988)
/6/ V. Teal, S. Vaidya, D.B. Fraser, Thin Solid Films, 136, 21 (1986)
/7/ P.A. Gargini, C. Tseng, M. H. Woods, Proc. 20th IRPS, p. 66 (1982)

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CONDUCTING DIFFUSION BARRIER : FORMATION AND CHARACTERISATION OF WN OBTAINED BY THERMAL ANNEALING UNDER NH_3 OF W FILMS DEPOSITED ON Si

A. DENEUVILLE, M. BENYAHYA, M. BRUNEL* and B. CANUT**

Laboratoire d'Etudes des Propriétés Electroniques des Solides, CNRS, BP 166, F-38042 Grenoble Cedex, France
**Laboratoire de Cristallographie, CNRS, BP 166, F-38042 Grenoble Cedex, France*
***Département Physique des Matériaux, Université Claude Bernard - Lyon I, Boulevard du 11 Novembre 1918, F-69621 Villeurbanne Cedex, France*

Résumé - A partir de la RBS et de la diffraction X, nous montrons qu'au dessus d'une épaisseur critique de W (inférieure à 2000 Å), WSi_2 couvert par γ WN peut être formé par recuit sous NH_3 au dessus de 800° C de films de W déposés sur C-Si. La résistivité du γ WN est ~ 100μΩ-cm.

Abstract - From Rutherford Backscattering and X-ray diffraction we show that above a critical W thickness (< 2000 Å), WSi_2 covered by γ WN can be formed by annealing under NH_3 of W films deposited on C-Si. Its resistivity is ≤ ~ 100μΩ-cm.

I - INTRODUCTION

As the size of the elementary devices decreases in Si integrated circuits the depth of the jonction decrases. An intermediate layer of silicide (usually of refractory metal) must be inserted between Al and Si to avoid the formation of Al spikes which can destroy the source and drain jonctions /1/.

However Al also react with refractory metal silicide at ~ 500° C and another intermediate layer acting as Al and Si diffusion barrier between Al and the refractory metal silicide have also to be added /2/. Amorphous Ti-W mixtures or WN or TiN can be used for this purpose /2/. However, their structure and their properties change by annealing when $T_A > 500°$ C.

Here we look for the direct formation of crystalline tungsten nitride on tungsten silicide on Si by direct thermal annealing under NH_3 of crystalline W films deposited on Si. Some results exists at this moment for annealing under NH_3 of Ti deposited on Si /3/, but nothing in the case of W.

If the W silicide have been studied, the results are scarce and controversial on tungsten nitrides /4/. Neitherthe N content corresponding to the various crystallographic phases β"W_2N" and γ"W_2N, nor the presence of γ"W_2N" only, which has the peaks of β"W_2N" plus additional ones, are clear.

Here, we look at what will be formed by annealing under NH_3 of crystalline (α phase) W deposited on Si. We use X-ray diffraction at glancing incidence, Rutherford Back Scattering and four-point probe measurements.

2 - PREPARATION AND APPARATUS

After the usual cleaning procedure, high resistivity Si samples are mounted in a cathodic sputtering group. After a base pressure of ~ 410^{-7} Torr, a 80 % A_r + 20 % H_2 mixture is admitted in the group with a total pressure of 410^{-3} Torr. After predepositions on a shutter, then on a sacrified sample, the W is deposited on the Si substrate at 200° C. W thicknesses of 200 and 2000 Å are measured by DEKTAK on a step after localized etching of W.
The annealing is performed under a classical joule oven between 600 and 1100° C during 2 hours. The X-ray diffraction is done at glancing incidence, typically O.5°, to increase the sensitivity and excited with the CuKα radiation.
The Rutherford Back Scattering use He ions of 2 MeV and normal (2000 Å of W) or 60° incidence (200 Å).
The square resistance $R_\square$ is measured by the usual four-point probe.

3 - EXPERIMENTAL RESULTS AND DISCUSSION

1) - Rutherford Back Scattering

a) 200 Å thick W films : Fig. 1.

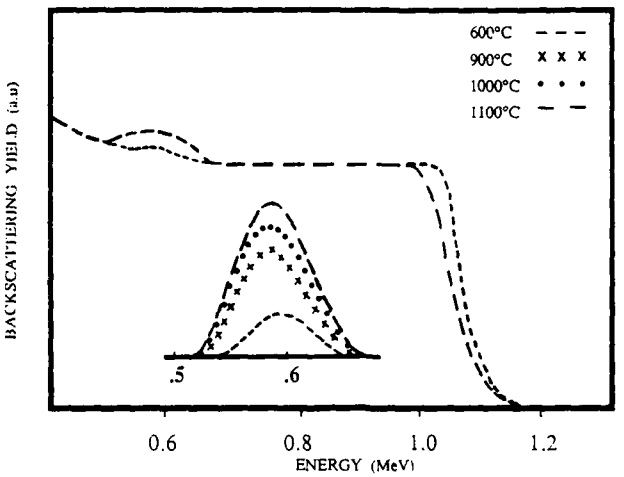


Fig. 1 - RBS signal for the 200 Å thick W on Si after 2 h annealing under NH_3 at 600° C -----, xxxxx 900° C, ... 1000° C and 1100° C ———

Although N is not favourable here as it appears on the Si background, it is clearly seen by RBS as soon as 600° C. The N content increases up to ~ 900° C where there is saturation . Formation of tungsten silicide (Plateau or shoulder) on the Si leading edge and W falling edge is not seen even with an incidence of 60°. From the height of the N and W signals approximate values of the N/W can be deduced. It increases from ~ 0.3 at T_A = 600° C up to a saturation value of ~ 1.16 as soon as 900° C

b) 2000 Å thick W films : Fig. 2.

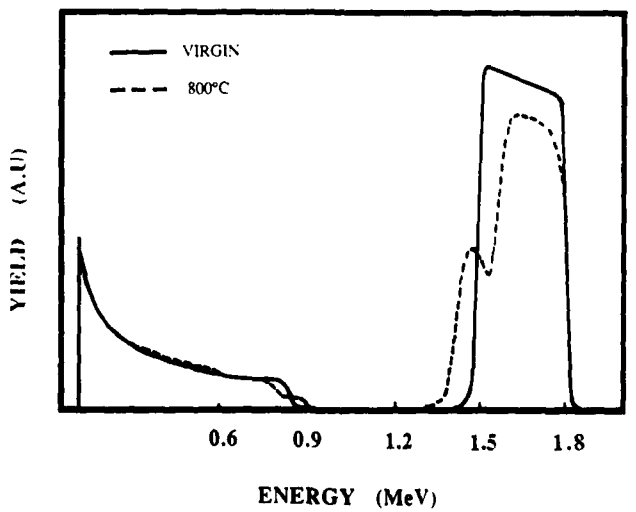


Fig. 2 - RBS signal for the 2000 Å thick W on Si virgin (——) and after 2 h annealing under NH_3 at 800° C.

It is here very difficult to determine the N signal height. Above $T_A \geq 800°$ C, there is a plateau on the Si signal and a hump in front of and smaller than the main W signal. This main signal is lower than that of pure W (——). There is formation of tungsten silicide (Si/W ~2) covered by a tungsten nitride of unknown composition.

2) - X-ray diffraction.

a) 200 Å thick W film.

The virgin films have only the peak of crystalline "pure" W (α-W) at 20.12 and 29.12°. At T_A = 600 and 700° C, there are additional peaks from "β" W_2N ((18.88, 21.94, 31.89°). At 800° C there is only "β"W_2N. At $T_A \geq 1000°$ C additional peaks at 10.78, 15.41, 24.73, 27.25° occur. All the peaks can be attributed to "γ"W_2N. There are also traces of a-W at ~ 20.12°.

b) 2000 Å thick W : Fig. 3.

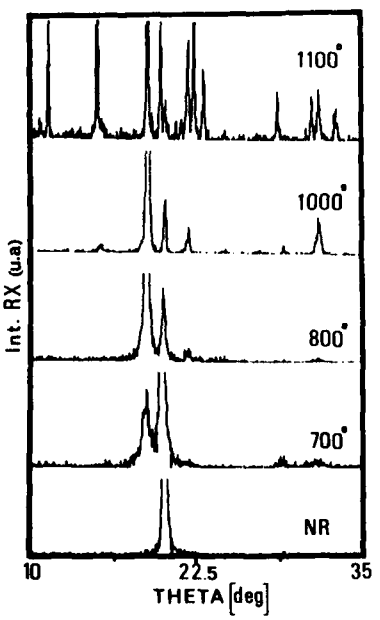


Fig. 3 - X-ray diffraction at glancing incidence (0.3°) for the 2000 Å thcik W on Si virgin (NR) and after 2 h annealing at 700, 800, 1000, 1100° C.

In the virgin sample (NR) the peaks are those of α-W. At 700° and 800° C there are additional peaks from "β" W_2N and the relative intensity of the W signal decreases. At 1000° C then 1100° C additional lines appears at 15.41 and 24.73 (γW_2N) and with a relative intensity increasing with temperature at 11.35, 15.02, 19.82, 22.40, 23.10, 28.70, 31.23° (tetragonal WSi_2). Actually with an higher incidence angle, the specific peaks of WSi_2 appears as soon as 800° C.

3) - Square resistance at room temperature : Fig. 4.

For the 200 Å virgin W film it starts from 30 Ω with a maximum of 115Ω at 900° C, then decreases to 50 Ω at 1100° C. As will be discussed later, only tungsten nitrides are formed. From their thickness we deduced resistivities of ~ 200 μΩ-cm β"W_2N and ~ 100 μΩ-cm γ"WN".

For the 2000 Å virgin W film it starts from 1.7 Ω has a maximum of 8 Ω at 900° then decreases around 2 Ω.

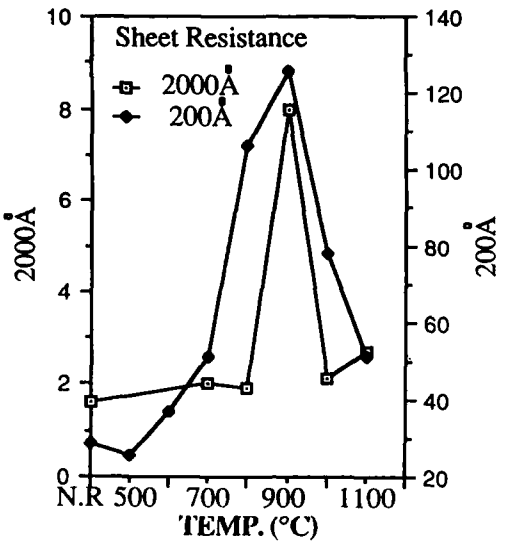


Fig. 4 - Sheet resistance of the 200 Å and 2000 Å thick W films 2 h annealed under NH_3

4 - DISCUSSIONAND CONCLUSION

- From RBS: There is a progressive increase of the N content from $WN_{0.3}$ (600° C) up to a saturation ~ WN at $T_A \geq$ 900° C without formation of tungsten silicide for the 200 Å virgin W film. For the 2000 Å W film, WSi_2 is formed for $T_A \geq$ 800° C with an increasing thickness as T_A increases. It is covered by a tungsten nitride.
- From X-ray diffraction increasing amount of β"W_2N" and γ'"W_2N" are formed respectively between 600 and 800° C and 1000 (perhaps 900° C) and 1100° C in both cases, and so their occurence depends only on T_A..

For the 200 Å W film, the whole film is filled by N prevent the formation of WSi_2 at $T_A \geq$ 650° C /1/. There is only β"W_2N" at T_A = 800° C which has to be with 0.5 < N/W < 1 which agrees with RBS, and would suggests W/N ~ 1, γ'"WN" rather than γ'"W_2N" for T_A = 1000 and 1100° C. From their $R_\square$, the resistivity of β"W_2N" and γ'"WN" would be respectively around 200 and 100 μΩ-cm.
For the 2000 Å W thick film, N does not reach the Si interface at 800° C. This allows the formation of WSi_2. But there is always α-W in the top part of the film which is never saturated by N for 2 h annealing between 600 and 1100° C.

REFERENCES

/1/ A. Hiraki in "Layered structures and interfaces kinetics. Their technology and application", Edited by S. Furukawa KTK Scientific publishers, Tokyo (1985).

/2/ M. Wittmer, Appl. Phys. Lett. 52 (1988) 1573, F.C.T. So, E. Kolawa, Xin-An Zhao, M.A. Nicolet, Thin Solid Films 153 (1987) 507.

/3/ A.E.T. Kuiper, J.C. Barbour, M.C. Willemsen, Proc. European Workshop on Refractory Metals and their silicides. Le Vide, Les Couches Minces 24, (1987) 29.

/4/ R. Kiessling, Y.H. Liu, Transactions AIME, J. of Metals (1951) 639.
V.I. Khithova, Z.G. Pinsker, Soviet Physics Cristallography 6, (1962) 712.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A SELF ALIGNED CONTACT PROCESS WITH IMPROVED SURFACE PLANARIZATION

K.H. KÜSTERS, W. SESSELMANN, H. MELZNER and B. FRIESEL

Siemens AG, Microelectronic Technology Center, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

A new self aligned contact technology has been introduced into a 4Mbit DRAM process. The contact hole is overlapping gate and field oxide. A thin nitride/thin poly-Si/oxide multilayer allows a contact hole etch, which does not significantly affect the oxide isolation of the gate and the field oxide. After acting as etch stop, the poly-Si is changed into oxide by selective oxidation. The new process offers an improved reflow of isolation oxide and contact hole rounding.

1. INTRODUCTION

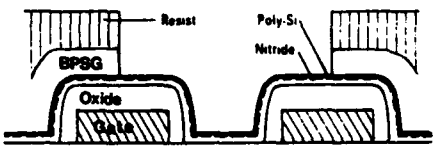
By using self aligned contacts, the packing density of integrated circuits can be increased /1/, lithography requirements are less stringent. A 25 % shrink of 4Mbit DRAM cell area was achieved by a fully overlapping bitline contact (FOBIC) /2/. The cell design allows the contact hole to overlap gate and field oxide. For etching the dielectric underneath the bitline a nitride etch stop layer has been used.

The self aligned contact technology can also be based on a poly-Si etch stop layer /3, 4/. In this paper we report on the integration of this technique into a 4Mbit DRAM process. The new process allows an improved surface narization. Improved planarization techniques are required, because surface topology is critical for the step coverage and for the patterning of submicron bitlines. The realization of submicron contacts by the proposed process is combined with improved techniques for the reduction of contact resistance, which is critical for device performance.

2. SELF ALIGNED CONTACT PROCESS

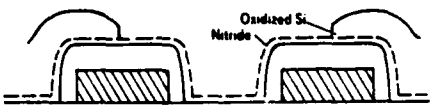
The self aligned contact process is employed for the fabrication of a 4Mbit DRAM trench capacitor cell. The 4Mbit DRAM process proceeds by the formation of trench capacitors, transistors, 1st level polycide and 2nd level metal interconnects /2/. The contact between polycide bitline and n^+ diffusion (source/drain of transfer gates) is self aligned; the process sequence for contact formation is outlined in Fig.1.

a) The gate is encapsulated by oxide ($\geq$ 150 nm) using a spacer technique: the oxide spacer covers the sidewalls of a double layer poly-Si/oxide. After the deposition of a thin nitride/thin poly-Si/oxide (BPSG) multilayer oversize contact holes are patterned. The top oxide (BPSG 350 nm) is etched anisotropically in a CHF_3/O_2 plasma. The thin poly-Si acts as a very efficient etch stop due to the high selectivity (20:1) of the dry etch process. A 30 nm poly-Si layer allows a long overetch time and efficient removal of BPSG sidewall spacers (Fig. 2). The gate isolation is not affected by the top oxide etch.

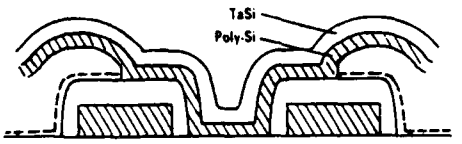


1) direct contact mask, oxide etch stopped by poly-Si layer

3) thin nitride/thin oxide etch



2) poly-Si etch; resist removal
poly-Si changed into oxide by wet oxidation



4) poly-Si deposition, poly-Si doping (As implantation), TaSi sputtering

Fig 1: Process flow for self aligned bitline contact

b) After etching poly-Si (selectively to nitride) and stripping resist poly-Si is wet oxidized. A simultaneous BPSG reflow occurs. The underlying thin nitride prevents the oxidation of active areas.

c) Only a thin dielectric is left for the final etch step by which the oxide isolation of the gate and the field oxide can be thinned. The thin nitride/thin oxide on contact areas is etched anisotropically in a CHF_3/O_2 plasma (Fig.3).

d) By LPCVD poly-Si deposition, As implantation doping and polycide (TaSi) deposition the bitline contact is formed (Fig.4). The contact area is defined by gate and field oxide edges. A distance contact to gate of only 0.25 µm can be realized, the oxide isolation between gate and bitline is at least 120 nm.

The new process with BPSG top oxide/wet reflow is advantageous for surface planarization and contact hole tapering. This is demonstrated by the comparison with the corresponding process using an undoped Teos top oxide (Fig.5). The reflow of BPSG top oxide removes the sharp contact hole edge

Fig 2: After dry etch of BPSG top oxide (← : poly-Si etch stop)

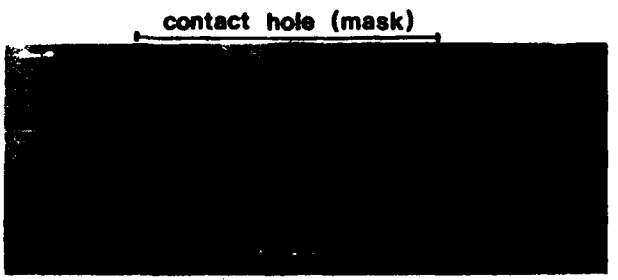


Fig 3: After contact hole etch (BPSG top oxide) (arrow: poly-Si etch stop layer, poly-Si is changed into oxide)

Fig 4: bitline contact after complete process (BPSG top oxide)

Fig 5: process with undoped Teos top oxide after contact hole etch

(see Fig.3). The wet reflow allows an effective reflow although BPSG doping level is kept low to avoid any influence of further temperature steps (2nd BPSG reflow) on the BPSG underneath the bitline.

The proposed process requires only a thin nitride layer (10-20 nm) as oxidation barrier. The thickness of the nitride layer, which remains on the wafer, has to be lower than 50 nm to avoid any effect on device performance e.g. enhanced leakage current due to defect generation /3/.

Different from a self aligned contact process /2,3/ based on nitride as an etch stop for top oxide etch the new process allows a pure dry etch of top oxide. In the case of a nitride etch stop a nitride thickness of 70 nm is required because of lower selectivity of the CHF_3/O_2 etch process; the nitride thickness can only be reduced by using wet etching techniques for the top oxide.

After serving as etch stop the poly-Si layer has to be oxidized to avoid problems with further processing (e.g. 2nd contact hole etch) or shorts. Fig.6 shows the oxide thickness grown on a poly Si buried under BPSG (500 nm)vs. oxidation time. The oxidation of the poly-Si layer can be performed without significant increase of the temperature budget of the process because the BPSG top oxide is less dense than thermal oxide.

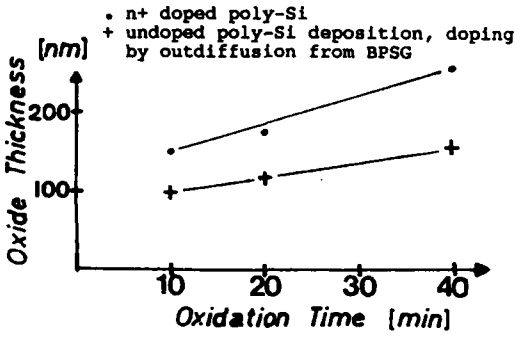


Fig 6: Oxide thickness on poly-Si under 500 nm BPSG

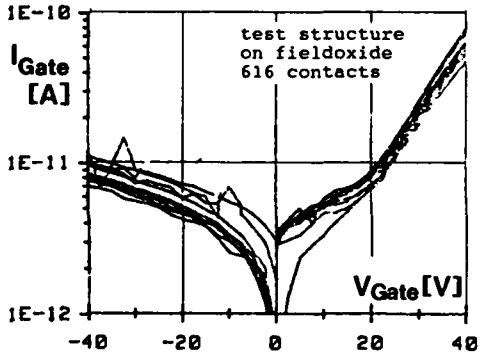


Fig 7: Leakage current across gate to bitline isolation

The new process has been introduced into the 4Mbit DRAM process without degradation of device yield compared to a conventional contact technique. The leakage current through the gate to bitline isolation was evaluated using a special test structure placed on field oxide to suppress the leakage current path through the gate oxide (Fig.7). The isolation of the gate encapsulation is superior to the isolation of thermal gate oxide. Investigation of LDD-Transistor properties exhibits no effect of adjacent self aligned contacts.

3. SELF ALIGNED CONTACT FOR STACKED CAPACITOR DRAM CELL

The proposed process can also be used for another type of self aligned contact: the contact is not only self aligned to gate and field oxide but also to an electrically active n^+ poly-Si layer like the upper cell plate of a Stacked Capacitor DRAM cell. The Stacked Capacitor /5/ consists of two poly Si plates, formed after the transistor. During the etching of the bitline contact to source/drain of the transfer gate, the upper cell plate can be used as poly-Si etch stop layer. After etching the contact through the n^+ poly-Si layer the contact isolation to the n^+ poly-Si is achieved by selective oxidation of the poly-Si inside the contact hole (Fig.8). Investigation of the oxide isolation between n^+ poly-Si and bitline contact shows a marked increase of breakdown voltage above 20 V for d_{ox} >160nm. Above this critical oxide thickness a reliable isolation is achieved.

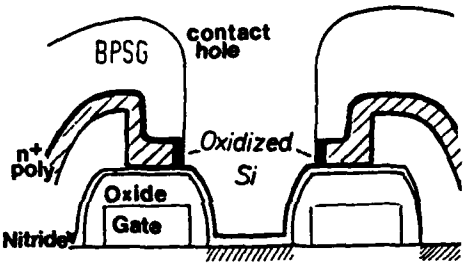


Fig 8: Self aligned contact through a n^+poly-Si layer (Stacked Capacitor DRAM) after poly-Si etch and poly-Si oxidation

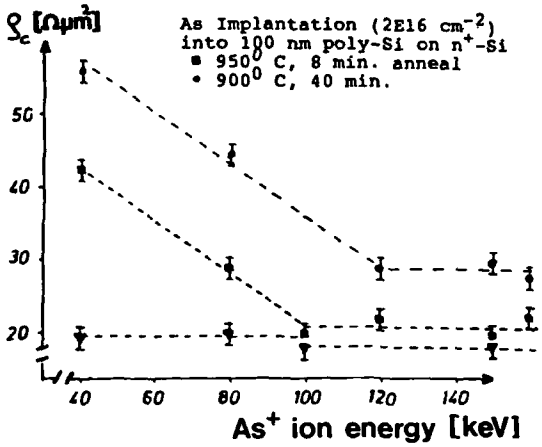


Fig 9: Specific contact resistivity as a function of As^+ ion energy (▼ additional Si^+ implantation)

4. CONTACT RESISTANCE POLY-Si/Si

For poly-Si/Si contacts it is well known /6/ that a thin native oxide layer layer of about 15 A between n^+-Si contact area and deposited poly-Si results in high contact resistivity. By ion implantation into poly-Si low specific contact resistivities ρ_c can be realized. Fig.9 shows the dependence of ρ_c on As^+ and Si^+ implantation energy. Low ρ_c values in the 20 $\Omega/\mu m^2$ range are found if the energy is high enough that As^+-ions reach the contact interface. High resolution TEM shows that the native oxide layer is broken up and epitaxial growth of the poly-Si layer occurs after a moderate heat treatment (950 C, 8 min). If Si^+-ions are used to break the native oxide layer low ρ_c values are obtained even after a low energy As^+ ion implantation (Fig.9). This technique turns out to be very promising since almost no degradation of the contact doping profile occurs by the additional Si^+ implantation.

Acknowledgement
The authors are grateful to W. Müller and G. Enders for their continuous support and encouragement. This work was supported by the Federal Department of Research and Technology (sign. NT 2696)

References

/1/ J.S. de Zaldivar, NTG Fachberichte Bd 77 (1981), p.22
/2/ K.H. Küsters, G. Enders, W. Meyberg, H. Benzinger, B. Hasler, G. Higelin, S. Röhl, H.M. Mühlhoff, W. Müller, Symp. on VLSI Tech. Dig. 1987, p.93
/3/ K.H. Küsters, H.M. Mühlhoff, G. Enders, E.G. Mohr, W. Müller, Proc. 1th Int. Symp. ULSI, 1987, Electrochem. Soc. Proc. Vol. 87-11, Ed. S. Broydo, C. M. Osburn, p.640
/4/ M. Kubota, S. Ogawa, S.Okuda, O. Shippou, T. Fujii, M. Inoue, T. Komeda, to be published, VLSI Symp. 1988
/5/ L. Risch, W. Sesselmann, R. Tielert, Proc. ESSDERC, 1987, p.757
/6/ G.L. Patton, J.C. Bravman, J.D. Plummer, IEEE Trans. El.Dev.ED-33, 1754 (1986)

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

FORMATION OF SHALLOW P⁺ JUNCTIONS BY IMPLANTATION INTO SILICIDE

K.J. BARLOW

Plessey Research Caswell Ltd., Caswell, Towcester, Northants, Great-Britain

Abstract

PMOS transistors with channel lengths down to 0.35μm have been fabricated by implanting boron into $TiSi_2$ and using a low temperature anneal (800°C) to outdiffuse the dopant and form the P+ junction. This method allows the formation of shallow, low resistance junctions (<0.26μm, 5Ω/sq) and retains any ion damage within the silicide. Electrical measurements show reverse leakage currents of ~1nA/cm² and PMOS characteristics comparable to or better than conventionally formed PMOS transistors.

Introduction

The formation of shallow, low resistance junctions is an important requirement in MOS VLSI in order to minimise short channel effects in MOSFETS. However, the formation of shallow P+/n junctions is difficult to obtain since boron tends to channel deeply into silicon during implantation and diffuses rapidly during further high temperature processing. Shallow P+ junctions can be formed by using RTA, preamorphisation and boron fluoride implants, whilst silicidation is used to reduce the sheet resistance of the source, drain and gate regions. A problem with using preamorphisation implants is the creation of defects in the silicon, whilst fluorine can increase leakage currents. Many of the problems associated with shallow junctions can be avoided by implanting the dopant into the silicide and then outdiffusing it using a fairly low temperature anneal to form the junction (1,2). Implantation into the silicide means that any ion damage is confined within the silicide - not in the silicon, uniform dopant profiles can be achieved since anisotropy and shadowing effects of the implantation are eliminated. Also since the implant process does not create damage in the silicon the temperature is not determined by that required to anneal out the ion damage, thus low temperature processing can be used.

Here the results of an investigation into the formation of PMOS transistors, with channel lengths down to 0.35μm, by implanting boron into $TiSi_2$ and using a low temperature anneal to outdiffuse the dopant are presented.

Experiment

An advanced CMOS process, employed in our laboratories was used to fabricate the devices. This employs trench isolation, oxide sidewall spacers, a 20nm gate oxide and silicided source, drain and gate regions. Fig. 1a,b shows a schematic of the two process schedules used to form the pmos devices. In the conventional process, fig. 1a after the cvd oxide sidewall spacer is formed a silicon implant is used to preamorphise the silicon to a depth of about 0.12μm, followed by a boron fluoride implant (1E15/cm², 40kV). The solid phase epitaxial regrowth anneal is used to regrow the single crystal silicon whilst minimising boron diffusion. RTA annealing is followed by deposition of 500Å of Ti and a two stage RTA process to form ~680Å of $TiSi_2$. Fig. 1b shows the reduced number of processing steps required when forming P+ junctions by implanting the dopant into the already formed silicide layer. Boron implantation into the silicide is followed by a fairly low temperature anneal (800°C) to form the P⁺ junction. Fig.2 shows the relationship between boron implant energy and junction depth for P+ junctions formed by implantation into $TiSi_2$. Shallower junctions can be formed by increasing the silicide thickness or reducing the implant dose and energy. Fig 3 shows a SIMS profile for a P+

junction formed by implanting boron at 3E15/cm^2, 10kV into 680Å of $TiSi_2$ and then outdiffusing it using an 800°C, 60 min, N_2 anneal. The junction depth is around 0.21μm, when taking the pmos n-well doping into account. After contact formation and metallisation the devices were sintered (425°C, 30min, H_2/N_2) and were then ready for testing. The sheet resistance was ~6Ω/sq for the P+ junctions formed by implantation into silicide, compared to ~5Ω/sq for the conventional silicided P+ junctions. P+/n diodes fabricated by this technique have shown low leakage currents of ~1nA/cm^2 at -5v and almost ideal forward current characteristics, for a range of boron doses and energies. Fig.4 shows a reverse I-V plot for a P+/n diode of area 47,000 μm^2.

PMOS transistors, with channel lengths down to 0.35μm have been fabricated by implanting boron into $TiSi_2$ and compared with conventional pmos devices. Measurement of the linear transconductance Gm, as a function of effective channel length and pmos transistor type is shown in fig 5, as can be seen the implant into silicide device has a slightly higher Gm, due to the reduced series resistance of the P+ junction. The low field gain of the implant into silicide device is ~ 23μA/V^2 compared to ~20.5μA/V^2 for the conventional device and the saturation current of the implant into silicide device is a few percent higher at the same channel length. Obviously by increasing the BF_2 dose the series resistance of the conventional device will be reduced, however the junction depth will also increase - by implanting into the silicide a shallow, high surface concentration junction can be formed. Fig.6 shows the output characteristics for a pmos transistor with an effective channel length of 0.35μm, formed by implanting boron into $TiSi_2$ and using an 800°C anneal to form the junction.

In the subthreshold region both the conventional and implant into silicide devices have the same subthreshold slope (91mV/dec), fig.7 shows a subthreshold plot for a pmos transistor with a channel length of 0.4μm formed by implanting boron into $TiSi_2$ to form the P+ regions. Offstate leakage (Vd=-5v, Vg=1v) was found to be less than 10pA for channel lengths down to 0.65μm, below this the leakage increases to .1μA for a channel length of 0.45μm as a punchthrough begins to control the maximum allowable drain voltage. At channel lengths above 0.65μm the maximum drain voltage is controlled by the breakdown voltage of the P+/n junction.

The reduction of the threshold voltage Vt with channel length was plotted for the conventional pmos and implant into silicide device, fig.8. As the plot shows the use of high dose boron (5E15/cm^2) implanted into the $TiSi_2$ layer does not enhance Vt reduction.

Conclusion

Implantation of boron into $TiSi_2$ and use of a fairly low temperature anneal (800°C) has enabled the fabrication of shallow P+/n junctions with low leakage currents. This method has been used to fabricate pmos transistors with channel lengths down to 0.35μm. The devices formed by this method have characteristics equal or superior to conventionally formed silicided pmos transistors.

The procedure is compatible with submicron CMOS processing.

Acknowledgements

The author would like to thank the Plessey Co. for permission to publish and the Alvey Directorate for partial funding.

References

1. F C Shone, K C Saraswat and J D Plummer. IEDM 1985, ext abs p.407.

2. R Lui, D S Williams and W T Lynch. IEDM 1986, ext abs p.58.

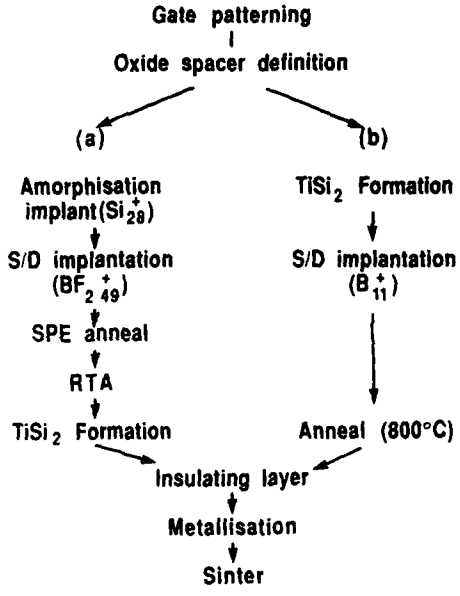


Figure 1

Schematic of the silicided source/drain schedule.

(a) Conventional first form junction then silicidation.

(b) Silicidation then implant into silicide and activation at 800C.

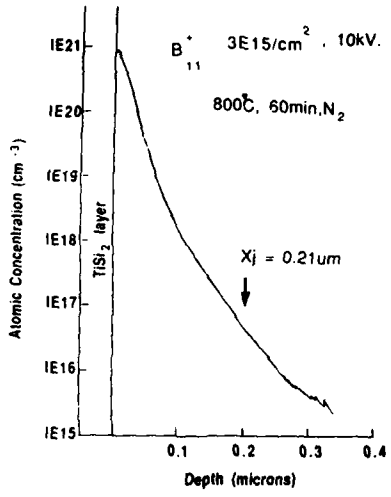


Figure 3

SIMS profile for a P^+ junction formed by boron implantation into $TiSi_2$.

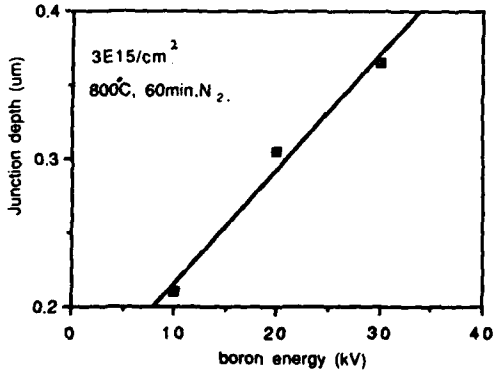


Figure 2

Junction depth as a function of boron energy.

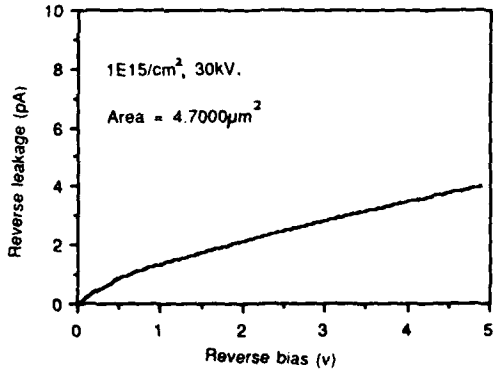


Figure 4

Reverse leakage plot for a P^+/n diode formed by boron implantation into $TiSi_2$

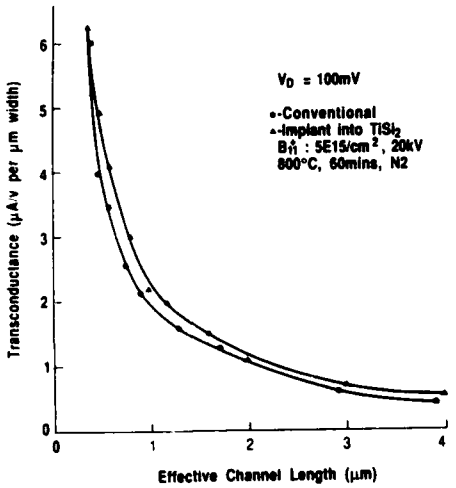


Figure 5

Linear transconductance as a function of effective channel length.

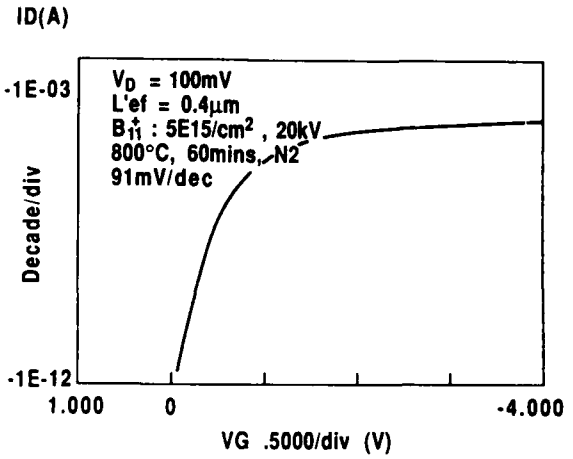


Figure 7

Subthreshold characteristics for a PMOS transistor formed by implantation into $TiSi_2$

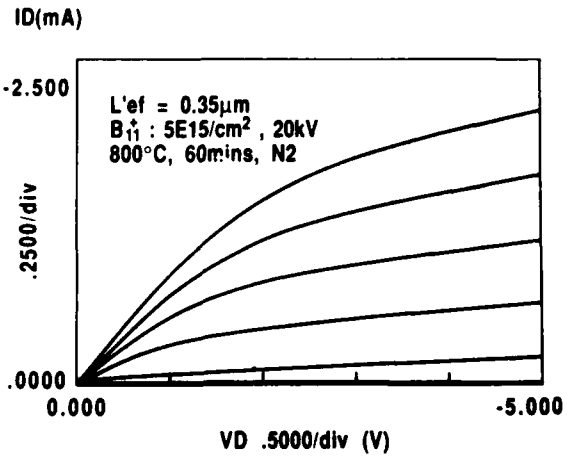


Figure 6

Output characteristics for a PMOS transistor formed by implantation into $TiSi_2$

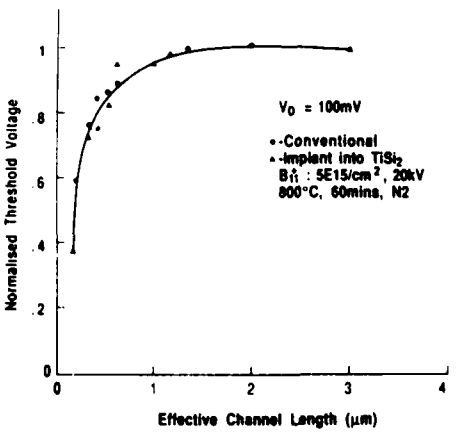


Figure 8

Channel length dependence of threshold voltage reduction.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

SHADOWING EFFECT IN SELF-ALIGNED CONTACTS

E. CAMERLENGHI, E. SERVIDA and M. TOSI

SGS-Thomson Microelectronics, Via C. Olivetti 2, Agrate Brianza (MI), Italy

Résumé. *Une comparaison entre le dopage avec arsenic et phosphore des contacts auto-alignés par implantation ionique est presentée. En utilisant la simulation bidimensional pour le profil de dopage il à été possible expliquer la haute défectuosité mesurée des contacts dopés arsenic. Son origine depend d'une insuffisant superposition entre la diffusion du contact et l'oxyde de champ, dû à un effet d'ombre dans une region critique du contact.*

Abstract. *A comparison between arsenic and phosphorus implanted self-aligned contacts is presented; by using 2D process simulations the high leakage defectivity measured on the As contacts has been explained. Its origin arises from an insufficient overlap between the contact diffusion and the field oxide, due to a shadowing effect in a critical zone of the contact.*

1 Introduction

To enhance compactness of a CMOS process architecture, self-aligned contacts are often used; a self-aligned contact (SAC) is defined as a contact larger than the active area (fig 1a), therefore in a SAC two zones are present in which the bird's beak of the field isolation may be removed during the contact opening, leading to a short between the metal and the underlying p-isolation. To avoid this problem, that can appear also in conventional contacts when misalignment between active areas and contacts is present, a n^+ implant is usually performed after the contact opening operation, to self-align the contact itself.
In the particular case of an EPROM process, to minimize lateral diffusion and the contact-gate distance, an arsenic implant is generally considered to be the better choice [1]. However, experimental results show that similar contacts have to be considered critical due to their high leakage defectivity.

2 Experimental

The structure used to verify the leakage of SACs consist of an array of about 60000 $1.4 \times 1.6\mu m^2$ contacts, that are self-aligned to an active area whose width is $1\mu m$. The leakage of the structure is collected by a metal cover that contacts all the SACs simultaneously. The junction depth of the As active area diffusion is about 2500Å. All the contacts have the same orientation on wafer.
To make comparisons with the behavior of standard contacts, similar structures are available in which an active area head is added to the contact.
Different contact implants have been tested, and all have been performed using a tilt angle of 7°, to minimise channeling problems; leakage currents have been measured when a reverse bias of 10 V is applied to the structure.
To avoid shadowing effects in the region below the lateral side walls of the SAC, the ion beam projection has been set parallel to the active area.
In figure 2 the leakage distributions of both self-aligned and standard contacts with an arsenic contact implant are reported; the implant dose is $5 \times 10^{15} cm^{-2}$ at 80 KeV, and the metalisation is a simple AlSi. As it can be noticed, a high defectivity is exibited by the SACs, and it disappears if a head is added to the contacts.
No significant differences are introduced when the As contacts are metalized with barrier metal; as it can be shown in figure 3 the use of a double layer of Ti and TiN does not remove the high leakage tail in the SACs.
Better results are obtained by implanting the contacts with phosphorus; in figure 4 the leakage distribution is reported for a $5 \times 10^{15} cm^{-2}$ 80KeV P implant and standard metalization. A strong reduction of the SAC defectivity is evident, together with a sharper distribution of the leakage current.
In figure 5 the behavior of As contacts with barrier metal is shown , in which the As dose has been implanted half at 7° and half at $-7°$, to avoid residual shadowing effects. A heavy improvement in comparison with simple 7° implant is manifest.

3 Simulation and discussion

As it appears from figure 2 and 3, in which no leakage is present in contacts with head, the problem of the leakage in As implanted SACs is strictly related to the presence of the self-aligned part of the contact. Besides, the particular kind of metalisation utilised, does not play any role in determining the leakage mechanism.
The experimental data could be explained if we suppose some residual shadowing effect in spite of the appropriate wafer orientation used during the contact implantation.
To verify this assumption, 2D simulations of the self-aligned contact have been performed with TITAN-4 2D Process Simulator [2]; cross sections of arsenic and phosphorus doped self-aligned contacts have been obtained in the worst case of 0.4μm misalignment between active area and contacts (mask to mask to mask misalignment). In particular, doping profiles simulations have been performed along cross sections normal and parallel to the active area. In the last case the doping profile has been obtained along a line 1500Å away from the side wall of the misaligned contact (see line A in fig. 1b).
Data are reported in figure 6 for both arsenic and phosphorus. For the arsenic case, considering the normal cross section, a small and critical overlap between the diffusion and the field oxide may be noted. This critical overlap indeed becomes a heavy miscoupling if the parallel cross section is analyzed.
Therefore the 2D simulation confirms that in a SAC exists a critical point in which the p-isolation may be contacted, leading to a leakage of the contact itself. This zone is indicated with C in figure 1b.
This problem is completely removed by the use of a phosphorus implant; as it can be noted, the deeper junction of phosphorus gives place to a larger plug below the contact, that guarantees a good overlap also in presence of misalignment and shadowing effect.
The simulated trends are widely confirmed by the SEM cross sections reported in figure 7. Also it is important to note that the As implant at 7° and −7° is sufficient to eliminate the leakage current, since it eliminates the miscoupled zone . Unfortunately a similar solution is not proposable because of the related low throughput and high handling.
A 7° phosphorus implant seems to be the best choice to reduce the leakage defectivity; in fact experimental data from wafers with different P contact implants do not point out any critical dependence of the leakage current on the particular dose and energy in a wide range of variation. On the other side, the necessity of a P contact implant imposes to keep into account the spurious effects on the effective length of the adjacent transistor and on the EPROM cell writing efficiency; in fact the presence of phosphorus near the drain region of the cell (that becomes particularly important when contact to gate misalignment is present) could reflect either on a higher drain electric field either on a lower one, depending on the particular dose utilized for the implant; an example is reported in figure 8. Therefore the right dose and energy has to be carefully tuned in order to maintain a good writing performance.
Similar observations should be valid also for 0° arsenic implants; this situation, that has not been evaluated in this work, should lead to results similar to those of a ±7° As implant. Anyway the presence of vertical and lateral channeling could affect the performance of the nearby transistors too.

4 Conclusion

The high leakage defectivity of the arsenic implanted SACs has been interpreted as due to a residual shadowing effect in a critical corner of the contact; the 2D TITAN-4 Process Simulator has turned out very useful to visualize geometries and dopings in different zones of the self-aligned contact.
It is important to underline that the entity of the shadowing non- overlap effect is strictly dependent on the particular process architecture and contact technology that is utilized. The extent of the contact over-etch, the contact reflow temperature and time and the lithographic misalignment are the most critical parameters that can affect the leakage current. The effects of critical overlap of the arsenic contact diffusion will become heavier and heavier for future process architectures in which higher compactness and new technologies with reduced diffusions will be required.
The high lateral diffusion that is obtained with P implants allows to overcome the shadowing problems; on the other side, spurious effects could be introduced. A fine tuning of the phosphorus dose and energy seems to be the safer solution. A 0° arsenic implant could also be a suitable one but a careful preliminary evaluation of the lateral channeling mechanisms is necessary.

Acknowledgments

We should like to thank N.Circelli and L.Polignano for their assistance to this work and CNET, Grenoble, France, for TITAN-4.

[1] S.Mori et al. *IEDM Technical Digest*, 835, 1986

[2] A.Gerodolle et al. *Proc. NASECODE IV conf.*, 477 - 482, Dublin, Jun. 1985

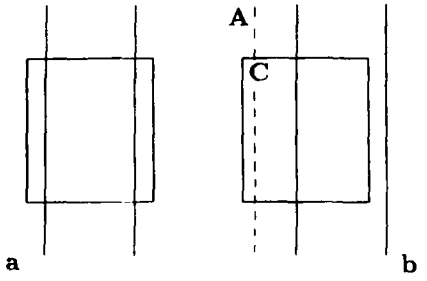


figure 1 1a) Self- aligned contact schema.
1b) The self-aligned contact when 0.4 μm misalignment respect to the active area is present. Along the line A, 2D doping profile simulations have been performed; with C the critical corner is indicated in which the shadowing effect takes place.

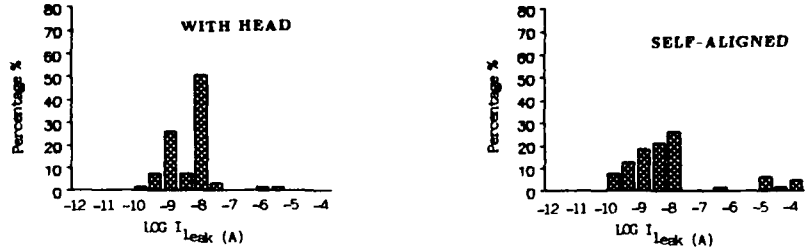


figure 2 Leakage distribution at T=300 K of 10^4 arsenic implanted contacts metalized with AlSi; the implant dose is 5×10^{15} cm^{-2} at 80 KeV and has been implanted with a 7° tilt angle.

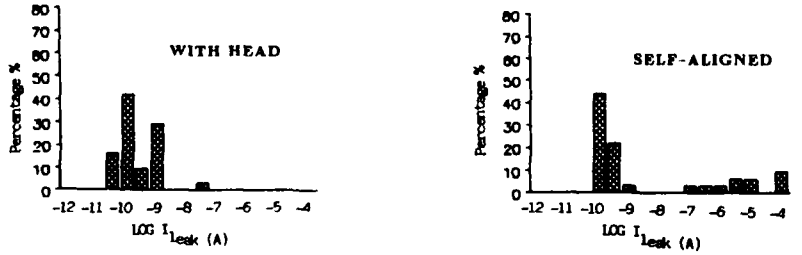


figure 3 Leakage distribution at T=300 K of 10^4 arsenic implanted contacts metalised with barrier metal; the implant dose is 5×10^{15} cm^{-2} at 80 KeV and has been implanted with a 7° tilt angle.

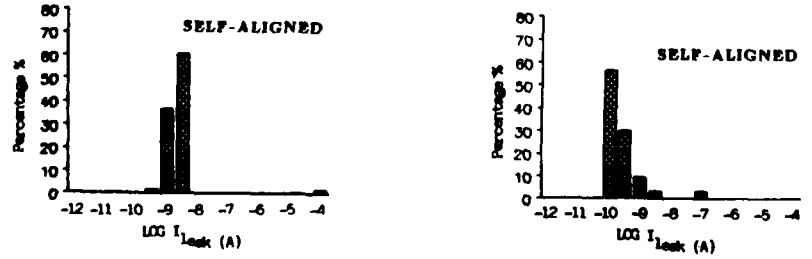


figure 4 Leakage distribution at T=300 K of 10^4 phosphorus implanted contacts metalised with AlSi; the implant dose is 5×10^{15} cm^{-2} at 80 KeV and has been implanted with a 7° tilt angle.

figure 5 Leakage distribution at T=300 K of 10^4 arsenic implanted contacts metalised with barrier metal (Ti TiN); the implant dose is 5×10^{15} cm^{-2} at 80 KeV and has been implanted with a $\pm$7° tilt angle.

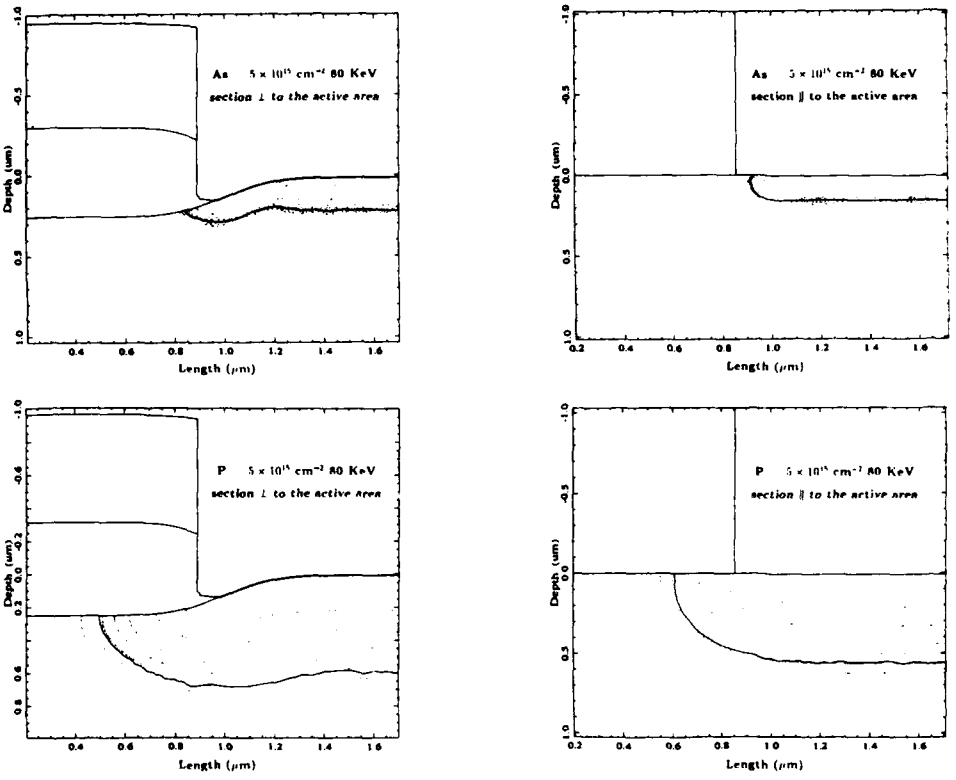


figure 6 In the upper and lower part of the figure, the simulated doping profiles are reported for As and P doped SACs respectively. On the left, the cross sections normal to the active area are shown, the parallel ones on the right.

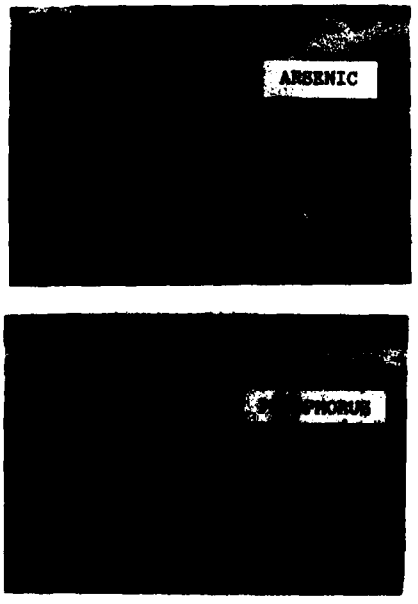


figure 7 Cross section SEM photos for arsenic and phosphorus implanted self-aligned contacts.

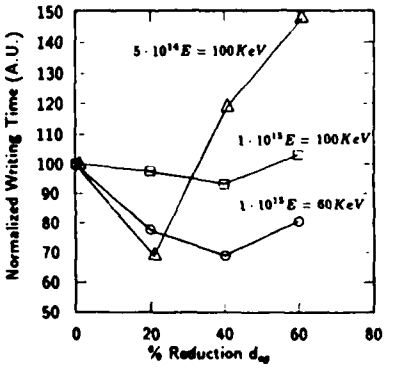


figure 8 Normalised writing time of an EPROM cell vs percent reduction of the contact-gate distance, for different doses and energies of the P contact implant.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

2D BORON DISTRIBUTIONS AFTER ION IMPLANT AND TRANSIENT ANNEAL

P.J. PEARSON and C. HILL

Plessey Research Caswell Ltd, Towcester, Northants, GB-NN12 8EQ, Great-Britain

Abstract - A novel 2D profiling technique of high spatial resolution (20 x 20 nm) is used to measure the asymmetry in implanted and transiently annealed boron implants caused by 7° tilt of VLSI wafers during implantation.

1 - INTRODUCTION

A new experimental technique, described elsewhere /1/, has enabled high resolution measurements of two-dimensional dopant distributions in small geometry structures. The principle of the technique consists of fabricating identical planar resistor structures, electrically isolated from the substrate and from each other, and then monitoring the conductivity changes as thin layers of different geometries are removed from each structure. Mathematical analysis shows that with a minimum of three differently sectioned structures, sufficient data is obtained to reconstruct the original conductivity distribution, and hence by use of published mobility data, the ionised dopant distribution. In order to realise this technique experimentally, four essential components have been developed:

(1) A test chip used to define 12 special 4-terminal resistor structures. The resistors are shown in Fig.1; the set (C, G, K) produce the minimum data set required, set (D, F, L, J) enable the effects of implant asymmetry on resistor boundaries to be measured, resistor edge (B, H) and end (A, I) effects can be assessed and the progress of the anodisation can be monitored (E).

(2) A fabrication sequence produces the resistor connections and windows. The desired implant and anneal is incorporated in these windows and the starting geometry for sectioning is defined. Schematic sections through 7 of the windows used to obtain results reported later in the paper are shown in Fig.2. One resistor (K) is retained complete to the anodic sectioning stage, and yields the familiar one-dimensional dopant profile representative of the planar portion of the resistor. All other resistors have the planar portion of the implant removed by plasma etching. Half of the structures then have the dielectric mask removed (G, L, J): in (C, D, F) it is retained. Since the implantation is not normal, but at 83° to the silicon surface, along the line A-A' in Fig.1, resistors close to this alignment (C, G) will have symmetrically implanted edges, whereas those nearly at right angles (D, F, L, J) will be asymmetric. The processing removes one resistor edge from two structures (D, L), and the other edge from another two structures (F, J), as shown in Fig.2., so that this asymmetry can be measured.

(3) A computer-controlled anodisation cell and probe card allow a cyclic anodisation and etch process up to 20 nm thick sequential layers from the resistor structures. Conductance measurements are made on each structure automatically. The initial and final topographies are determined by Dektak and SEM measurements.

(4) The 2D conductivity distribution is determined by modelling a section through each resistor as a finite matrix of rectangular elements. The topography of the sequential sections are mapped on to this matrix, and different conductances assigned to each matrix element until a set consistent with all three conductance profiles is found. A computer minimisation procedure is used to fit the conductance data to a 2D exponential function with up to 12 fitting parameters.

2 - EXPERIMENTAL

The dielectric layer in which the resistor windows were cut consisted of 200nm. of thermally grown SiO_2 overlaid first by 100nm. of Si_3N_4 and then 200 nm. of deposited oxide. Plasma-etching to halfway through the bottom oxide layer, followed by a wet-etch to the clear windows resulted in a mask edge shape as shown in Fig.5. Half the samples were then implanted with ^{28}Si ($1x10^{15}$ ions/sq.cm. at 70keV + $2x10^{15}$ ions/sq.cm. at 180 keV) with wafer cooling, to amorphise the silicon under the window to a depth of 200 nm. All samples were then implanted with boron ($5x10^{14}$ ions/sq.cm. at 40keV), using 22° rotation and 7° tilt to minimise ion channeling as illustrated in Fig.1. A matrix of transient anneals in 10% O_2 at 950°C (2 to 300

seconds) were carried out, one matrix condition per slice. After defining the metallisation and carrying out the silicon and dielectric layer removal stages described earlier, the starting values of conductance of the resistors over a large number of chips were measured. These values and the statistical variation about them are summarised in Table 1. Several slices were chosen for detailed 2D analysis; the conductance of all 12 resistors were continuously monitored as 10 or 20nm. layers were anodically and sequentially stripped from them.

3 - RESULTS AND DISCUSSION

The initial conductance values are in themselves informative. (Table 1). The 1D resistor (K) values show that while full anneal is achieved after 2 seconds in preamorphised (PA) samples, the same conductance is only reached in crystalline matrix (CM) samples after 300 seconds anneal. A similar comparison of the time - dependence of conductance of resistors C, G, D+F, L+J, shows suprisingly that the lateral dopant in all samples is fully activated after only 10 seconds anneal. The sum of conductances of the left and right hand edges of the asymmetric resistors (D + F) is close to that of the two symmetric edges measured in C, but the ratio D/F shows that there is strong asymmetry where the implant is at right angles to the mask edge. The excellent control of the overall processing is shown by the narrow spread of data about the mean values (about 2% for structure K, about 8% for the plasma etched structures) corresponding to a maximum variation in the etched topography of about 200 nm.

An example of the differential conductances obtained during the anodic sectioning sequence for resistors C, G and K is shown in Fig.3. The evolution of topography in structures C and G is shown in Fig.4. Data sets such as these were input to the computer analysis programme to produced the best fit boron distributions. Full conductance data sets, and subsequent 2D boron distributions, have been obtained for slices corresponding to rows 4 to 7 of Table 1. The left hand portion of Fig.5 shows the best fit boron distributions for the two sides of the asymmetric resistors; the accompanying table summarises the lateral spread values for symmetric and asymmetric plus the vertical penetration. The mask topography at the implant stage is also shown, to the same scale, and it can be seen that the angle of the ion beam would be expected to produce an asymmetry in the same direction as is actually measured. This assymetry in every case corresponds to 0.10±0.01 microns difference in lateral spread between the left and right hand sides of the resistor windows. The PM samples show no significant change in lateral redistribution between 2 and 30 seconds; both the lateral and vertical spreads of the 300 second CM sample are considerably greater.

These results establish (i) the reliability and high resolution (20nm.) of this novel 2D measurement technique, (ii) the value of preamorphisation for low temperature activation of boron implants, (iii) the strong asymmetry that can arise from angled implants defined by thick masking layers. The latter result has strong implications for submicron CMOS technology, since just these conditions are likely to be obtained in source/drain fabrication, where these structures are particularly sensitive to lateral doping asymmetry.

ACKNOWLEDGEMENTS

This work was supported by the Plessey Co. plc, The MOCVD and the Alvey Directorate, and thanks are due to them for funding and for permission to publish.

REFERENCES

/1/ C. Hill, P.J. Pearson, B. Lewis, A.J. Holden, R.W. Allen, Proceedings ESSDERC 87 (Bologna, Italy) C4.1.3 932-926 (1987).

TABLE 1 - Mean conductance values of special resistor structures, at the end of processing, from measurements over 140 chips. All resistors are 16 microns long; K is 8 microns wide. Conductance values and standard deviations are in microsiemens. First set of data refers to implants into crystalline silicon, second set to implants into preamorphised silicon.

Anneal time	Unetched	Planar Region of Resistor Removed			Planar region and Mask Removed		
	K	C	D+F	D/F	G	L+J	L/J
2	1308 + 25	31.2 + 2.4	31.0 + 0.8	2.51 + 0.29	20.4 + 1.9	21.8 + 1.4	3.46 + 0.73
10	1677 + 24	40.0 + 1.7	43.7 + 3.1	2.40 + 0.35	25.9 + 1.4	30.0 + 2.5	3.91 + 0.47
30	1889 + 22	41.3 + 3.8	44.2 + 3.0	2.55 + 0.27	27.9 + 2.6	30.9 + 2.5	3.72 + 0.46
300	2142 + 49	41.6 + 3.4	45.1 + 3.8	2.80 + 0.34	25.4 + 2.0	28.3 + 2.1	3.95 + 0.74
2	2029 + 53	27.0 + 2.2	31.1 + 1.9	3.96 + 0.60	16.9 + 1.7	20.7 + 1.7	6.60 + 1.25
10	2138 + 39	28.2 + 3.6	30.8 + 2.7	3.63 + 0.66	16.8 + 1.5	19.4 + 2.1	7.44 + 1.37
30	2063 + 10	31.3 + 1.7	33.4	2.44	21.6 + 1.8	20.6	4.11

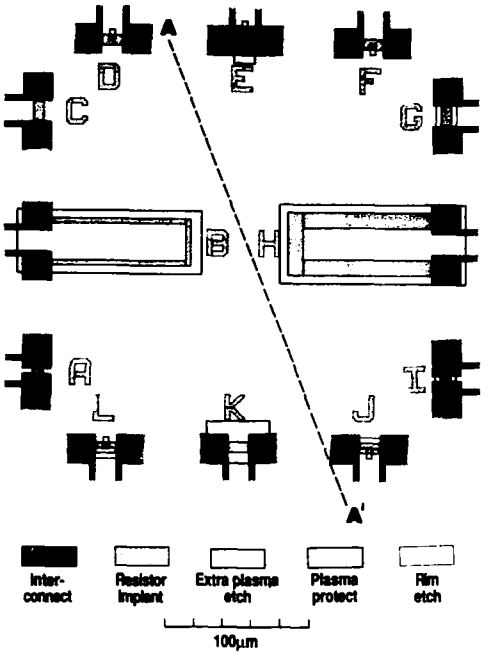


Figure 1 Detail of the 2D test mask central area

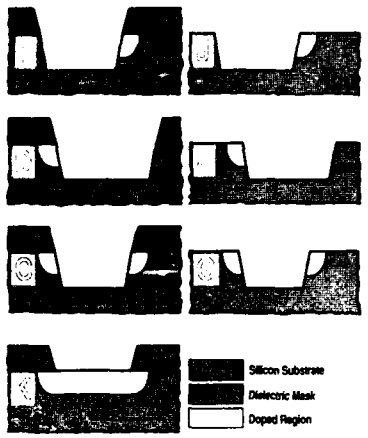


Figure 2 Schematic sections through selected resistor structures

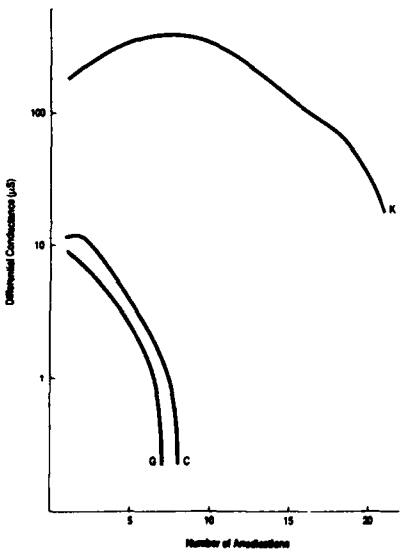


Figure 3 Experimental conductance-depth profiles measured during sequential anodic removals of boron implanted and annealed silicon layers from resistor structures C, G and K

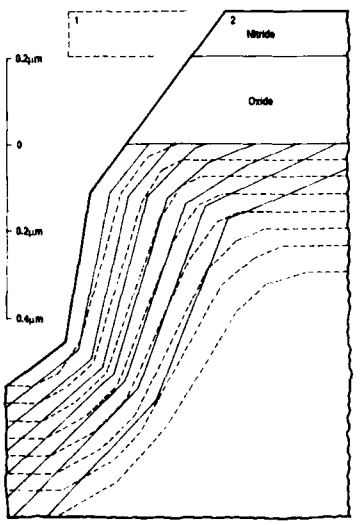
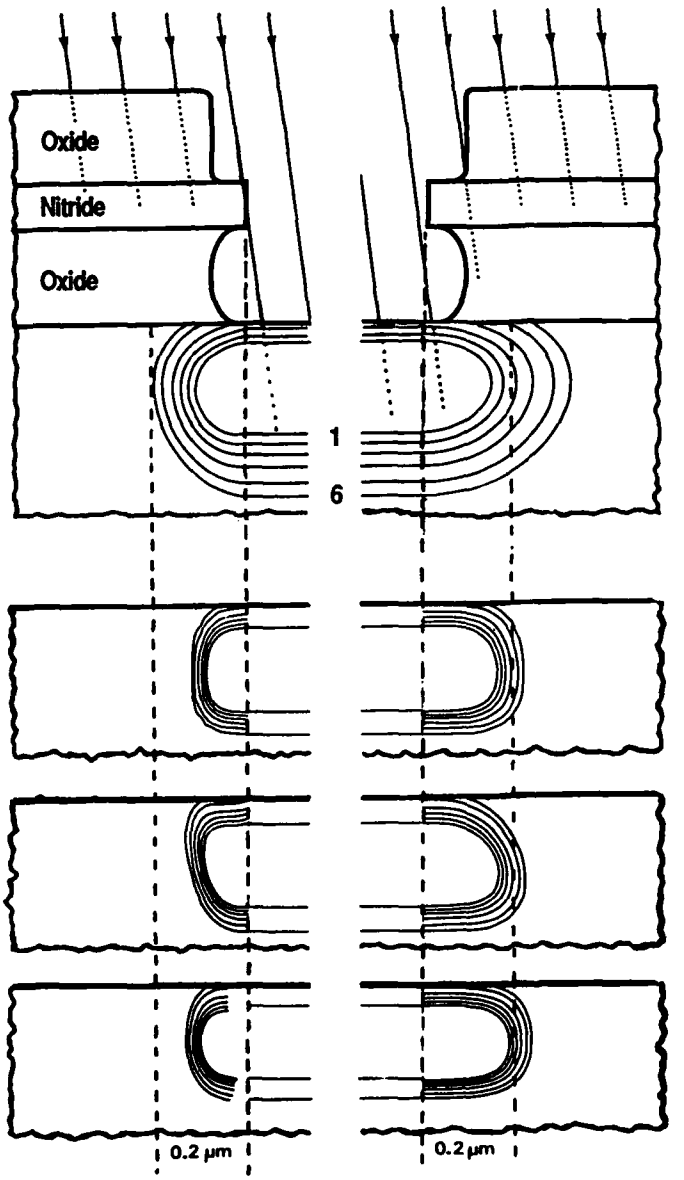


Figure 4 Experimental evolution of surface topography after anodic sectioning of structure C (solid lines) and structure G (broken lines). The positions of the masking dielectric edge after implant and anneal (1) and after substrate plasma etch (2) are shown



Penetration distance of the 10^{18} Boron contour (microns)			
Lateral spread Left asymmetric	Lateral spread Right asymmetric	Lateral spread symmetric	Vertical Penetration
0.20	0.31	0.23	0.37
0.12	0.22	0.17	0.27
0.13	0.22	0.16	0.28
0.13	0.23	0.19	0.24

Figure 5 Experimentally determined 2D boron concentration contours in sections through the edge of asymmetrically implanted and annealed resistor structures. Concentration represented by contours in ions/cc are:- 1, 1E19; 2, 8E18; 3, 6E81; 4, 4E18; 5, 2E18; 6, 1E18.
a) crystalline matrix, 300 sec. anneal
b) preamorphised matrix, 30 sec. anneal
c) preamorphised matrix, 10 sec. anneal
d) preamorphised matrix, 2 sec. anneal

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

MIGRATION OF FLUORINE ATOMS AND INFLUENCE ON SHALLOW P⁺N JUNCTION IN BF_2^+ IMPLANTED SILICON UNDER RTA

T.-Q. ZHANG, J.-L. LIU and X.-Y. YANG

Microelectronics Research Institute, Xidian University, Xi'an, China

Abstract-The migration of fluorine atoms in BF_2^+ implanted silicon under RTA has been analysed using SIMS, the microstructural defect of BF_2^+ implanted silicon before and after RTA has been observed using TEM, and the reverse leakage current of BF_2^+ implanted diodes after RTA has been measured. The results show that the amorphous layer and the damaged crystalline region strongly influence fluorine redistribution during RTA, but the anomalous migration of fluorine atoms has no measurable influence on the electrical properties of shallow P^+N junction.

1- INTRODUCTION

Scaled-Down VLSI CMOS circuits require shallow source-drain juctions with low sheet resistance. This gives rise to difficulty especially for P^+N junctions formed by B^+ implantation, because of ion channeling effect and fast boron diffusion during anneal, which necessaitates high temperature to obtain sufficiently high activation. To solve the problems, BF_2^+ implantation combined with rapid thermal annealing (RTA) technology is a better choice. Several works have been performed on this subject /1-7/, but the migration of fluorine atoms during RTA and it's influence on the electrical properties of BF_2^+ implanted silicon shallow P^+N junction have not been previously studied in detail. In this paper we report our experimental results on this project.

2- EXPERIMENT

Single-crystal, 4-6Ω.cm, N-type, (100) oriented silicon wafers are used in present study.Bare silicons used for SIMS and TEM analysis are implanted using BF_2^+ with energy of 45kev and doses of 1×10^{14}, 2×10^{15} and $5\times10^{15}cm^{-2}$. P^+N diodes with area of $10^4\mu m^2$ are made using BF_2^+ implantation with the same energy and doses. All the wafers are oriented 7° off the incident beam direction to minimize the channeling effect, and annealed in the Heat-pulse KST-1 type tungsten halogen lamp heated system. The annealing temperature of RTA ranges from 900°C to 1200°C

and the annealing time is 10 sec. The RTA is carried out in a dry N_2 atmosphere.
The migration of fluorine atoms in BF_2^+ implanted silicon after RTA is analysed using SIMS (CAMECA IMS 3F), the microstructural defects of BF_2^+ implanted silicon before and after RTA are observed using TEM (PHILIPS EM-430), and the reverse leakage current of BF_2^+ implanted diodes after RTA is measured using a FJ-356 electrometer.

3- RESULTS AND DISCUSSIONS

Fig.1 shows fluorine distribution profiles obtained from SIMS measurements for BF_2^+ implanted samples with 45kev and $2\times10^{15}cm^{-2}$ before and after RTA. The vertical dashed line in Fig.1 respresents the interface between amorphous and crystalline (a/c) silicon after implantation. The thickness of the continuous amorphous layer obtained from TEM measurements is about 600 Å. After RTA at 900°C for 10 sec, the peak of the fluorine distribution profile approachs to but is slightly less than the implantation peak.With the annealing temperature increasing, the amount of fluorine atoms diffused out increases, the broadness of fluorine distribution shrinks, the peak concentration of fluorine atoms drops, and the fluorine concentration peak moves towards a/c interface or below the interface.
Fig.2 shows fluorine distribution profiles of BF_2^+ implanted sample with 45kev and 5×10^{15} cm^{-2} before and after RTA. Since a dose of $5\times10^{15}cm^{-2}$ is higher, the damage of Si crystalline is seriouser. After RTA at 1100°C, most of the fluorine atoms still remain in silicon. The shallower peak corresponds to the implantation peak. The deeper peak occurs in the region below I which is heavily damaged but not amorphized by the implantation. It can also be seen from Fig.2 that boron atoms at the second peak suggest an enhanced diffusion effect.
Fig.3 shows fluorine migration of BF_2^+ implanted sample with 45Kev and $1\times10^{14}cm^{-2}$. Since a dose of $1\times10^{14}cm^{-2}$ is well below the critical dose needed to form an amorphous layer in silicon, the damage of Si crystalline is not serious. As a result, migration of fluorine atoms is apparently due to a simple outdiffusion process. After 1100°C annealing, fluorine atoms are almost diffused out, only few remain below the surface.
Fig.4 shows the reverse leakage current of BF_2^+ implanted diodes with 45kev and $2\times10^{15}cm^{-2}$ after RTA. It can be seen that the reverse leakage current annealed at 1200°C is the biggest, and the reverse leakage current annealed at 1050°C is the least. Comparing with Fig.1, it can be considered that the reverse leakage current is not directly related to fluorine atoms remaining in silicon.

4- CONCLUSION

The amorphous layer and damaged crystalline region strongly influence fluorine redistribution during RTA. Under lower dose implantation, the damage of Si crystalline is not serious, therefore fluorine atoms after RTA are almost diffused out. Under higher dose implantation, the damage of Si crystalline is serious, therefore most of the fluorine atoms after RTA are remained in silicon. But they have no measureble influence on the electrical properties of shallow P^+N junction.

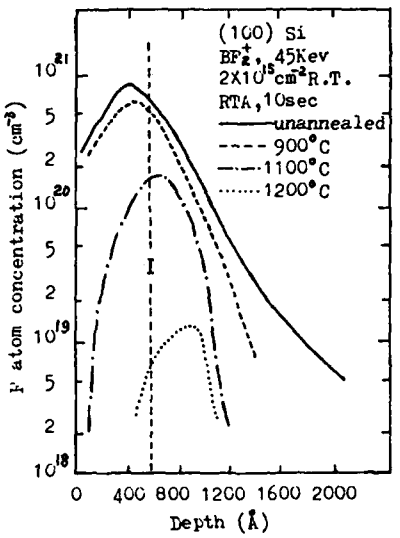


Fig.1-Profiles of fluorine atoms measured by SIMS for 45kev and $2\times10^{15}cm^{-2}$ BF_2^+ implanted sample before and after RTA. A dashed line (I) represents a/c interface after implantation.

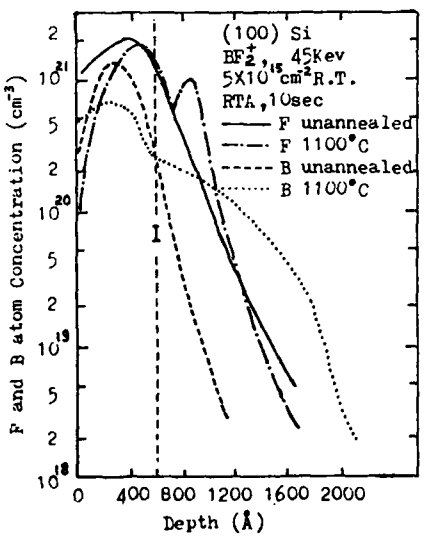


Fig.2 —Profiles of fluorine and boron atoms measured by SIMS for 45kev and $5\times10^{15}cm^{-2}$ BF_2^+ implanted sample before and after RTA. A dashed line (I) represents a/c interface after implantation.

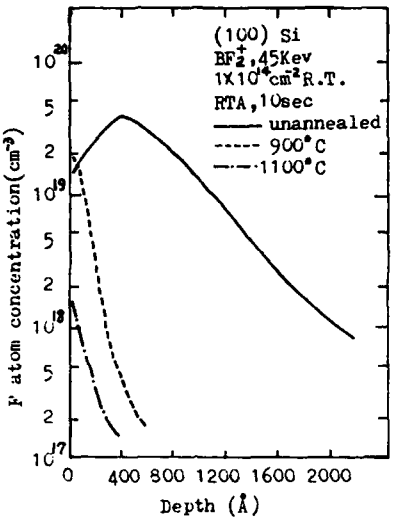


Fig.3-Profiles of fluorine atoms measured by SIMS for 45Kev and $1\times10^{14}cm^{-2}$ BF_2^+ implanted sample before and after RTA.

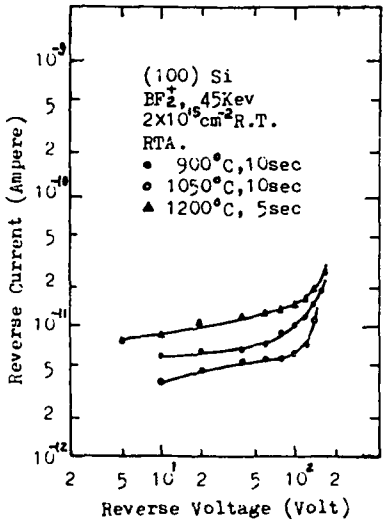


Fig.4- Reverse I-V characteristics for BF_2^+ implanted diodes with 45kev and 2×10^{15} cm^{-2} after RTA.

REFERENCES

/1/ Kwor,R., Kwong, D.L. and Yeo, Y.K., Appl.Phys. Lett. 45 (1984)77.

/2/ Lunnon, M.E., Chen, J.T. and Baker, J.E., Appl. Phys. Lett.45(1984) 1056.

/3/ Ozguz, V.H., Wortman, J.J., Hauser, J.R., Simpson, L., Littlejohn, M.A., Chu, W.K. and Rozgonyi, G.A. Appl. Phys. Lett. 45(1984) 1225.

/4/ Tung, N.C., J.Electrochem. Soc. 132(1985) 914.

/5/ Cohen, S.S., Norton, J.F., Koch, E.F. and Weisel, G. J., J.Appl.Phys. 57(1985) 1200.

/6/ Narayan, J., Halland, O.W., Christie, W.H. and Wortman, J.J., J.Appl.Phys. 57(1985)2709.

/7/ Mikoshiba, H. and Abiko, H., IEEE. Electron. Device. Lett. EDL-7(1986) 190.

Session 5B Room B : CMOS isolation

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

SCALLING OF TRENCH CAPACITOR CELL FOR NEXT GENERATION DRAMs

H.-M. MÜHLHOFF, C.M. ROGERS, P. MURKIN, M. ELAHY and S. RÖHL

Siemens AG, Corporate Research and Technology, Otto-Hahn-Ring 6, D-8000 München, F.R.G.

Abstract: When scaling a trench capacitor cell developed for the 4Mbit DRAM further down, both process and device limits are encountered. Device related topics are the subject of this paper. Issues to be discussed are: (1) narrow width effects of pass transistors, (2) short channel effects, (3) effect of storage region on pass transistor, (4) isolation between neighbouring cells.

Introduction: A 4Mbit DRAM using a Trench Cell with overlapping Bit-line contact has been developed and is now being sampled to customers[1]. Fig. 1 shows a layout of the memory cell and an SEM of its cross-section. Using this memory cell, a chip size of 92mm^2 has been achieved. In order to increase speed, reduce chips size or to possibly use this concept for the next generation DRAM, further scaling is required. Reducing the width of the active area increases the doping level in the transistor due to outdiffusion from the field region. This raises the threshold voltage and enhances degradation after hot electron stress due to higher field peaks at the channel borders. Shrinking transistor gate length degrades short channel behaviour and also aggravates the hot electron problem. Moving the cell plate too close to the pass transistor can either cause the highly doped varactor region to diffuse into the channel region decreasing the effective channel length and thus lowering V_T, or in order to avoid this problem the storage cell may no longer be properly connected to the pass transistor. The isolation between trenches can be maintained with shrinking trench-trench distance by increasing the substrate doping [2]. However, moving the trenches closer together also decreases the distance between trench and field isolation which can be attacked during the trench etch process.

The following sections are devoted to these problems and the minimum design rules are discussed. The actual design rules are obtained using these parameters and also taking the registration tolerance of the stepper and other process tolerances into account.

Narrow Width Effects: The increase of threshold voltage with decreasing transistor width is shown in Fig.2. Due to the already very high p-well doping ($4x10^{16}cm^{-3}$) in comparison to the channel stop implant ($10^{17}cm^{-3}$) the narrow channel effect is not very pronounced at widths down to 0.7um. Because of the higher doping at the channel borders, the pn-junction of the LDD region is more abrupt giving rise to higher field peaks. Therefore the narrow transistor degrades faster after hot electron stress as is evident from Fig.3. This problem is likely to get worse for even shorter and narrower transistors.

Short Channel Effects: In order to increase speed and save space, thinner gate oxide and shorter gate length will be employed. Fig.4 demonstrates the improved short channel behaviour when 18nm thick gate oxide is used instead of 20nm. A 0.9um transistor with 18nm thick gate oxide has identical V_T and dV_T/dL as a 1.0um transistor with 20nm thick gate oxide. A thinner gate oxide, however, causes a higher field peak

and enhances the generation of hot electrons under stress conditions. Degradation of transistors with 18nm gate oxide is worse than of transistors with 20nm thick gate oxide and identical drain structures. Lifetimes for 10% degradation of current in the linear operating regime are plotted in Fig.5 as function of gate lengths. Stress conditions were Vds=8V, Vgs=3V and Vds=-2.5V. Lifetimes for transistors with t_{ox}=18nm are a factor of 5 lower than for transistors with t_{ox}=20nm. Optimizing the drain structure, however, improves degradation and the disadvantage of thinner gate oxide can be eliminated by using a higher LDD dose ($4x10^{13}cm^2$ instead of $2x10^{13}cm^{-2}$) at a lower implantation energy (40keV instead of 60keV).

Effect of Storage Region on Pass Transistor: When decreasing the distance between trench capacitor and pass transistor, the outdiffusion of the highly doped varactor into the channel region of the transistor has to be avoided. At the present generation of 4Mbit DRAMs, a separate As-varactor implant is used to dope the planar part of the storage capacitor and to connect the transistor to the storage region. Fig.6 shows a cross-section where a is the overlap of the varactor implant and the cell plate and b the distance to the transfer gate. The dependence of V_T on b is plotted in Fig.7. It shows that no overlap between transfer gate and varactor can be tolerated. Otherwise the effective channel length is reduced and the pass transistor cannot be turned off properly. Data obtained to analyze the problem of sufficient overlap between cell plate and varactor is shown in Fig.8. Due to lateral diffusion of the varactor as well as lateral diffusion of the LDD region a negative overlap of 150nm can be allowed. The numbers obtained for distance between varactor and transfer gate and overlap over cell plate are worst case and cannot be exceeded under the maximum misalignment condition. Fig.9 shows the misalignment extracted from electrical measurements on symmetrical test structures. Further scaling can be achieved by omitting the varactor implant and using the lateral diffusion of the LDD implant as contact.

Isolation: In the trench capacitor cell, leakage due to punchthrough between adjacent trenches is a serious problem. However, previous results [2] show that at present p-well doping levels 1.5um separation can be realized. To improve the isolation further, a higher p-well concentration is used. The effect on the transistor V_T is compensated by thinner gate oxide thickness. The reduced trench-trench spacing is accomplished by decreasing the distance between trench edge and LOCOS. This causes some attack on the field oxide during the etching of the trench. Fig.10 shows a sketch of the worst case situation. For 0.9um LOCOS isolation problems arise, however, it is not a problem at 1.1um (Fig.11). Further scaling of cell size requires either a reduction of field isolation or a trench overlapping LOCOS.

Conclusion: Several issues affecting scaling of the trench capacitor cell have been illustrated. At this stage, it appears possible to develop a 4Mbit DRAM with an area of less than $70mm^2$ using optical lithography. Because of the many layers using minimum linewidth, achieving minimum registration tolerances is as important as achieving minimum linewidth. Using this concept for 16Mbit DRAM requires a process in which field isolation overlaps the trench or in which the trench is self aligned to the isolation. Also transistors with gate length and width of 0.6um will be required.

1 K.H. Küsters et al,
VLSI Techn. Symp., p.93 (1987)

2 P.A. Murkin et al,
Proc. ESSDERC, p.761 (1987)

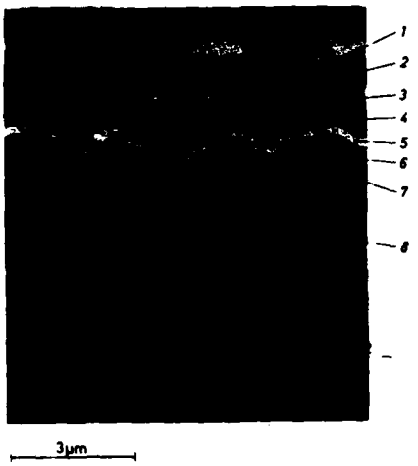
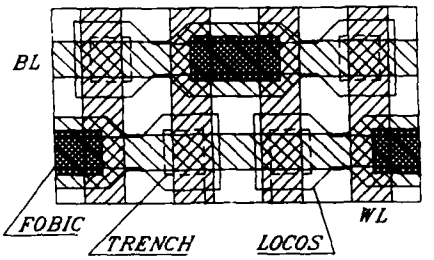


Fig.1 Cell layout and SEM of cross-section of 4Mbit DRAM cell.



1: Passivation Layer
2: Al/Si(1%)
3: Ti/TiN-diffusion barrier
4: BPSG
5: Silicided bitline
6: Poly-Si wordline
7: Poly-Si cell plate
8: Trench capacitor
9: Bitline contact
10: Pass transistor
11: Memory cell
12: Isolation between neighbouring cells

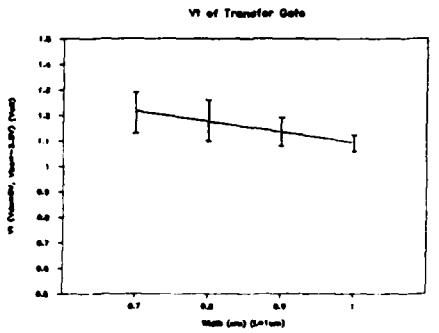


Fig.2 Dependence of Vt of transfer gate on width.

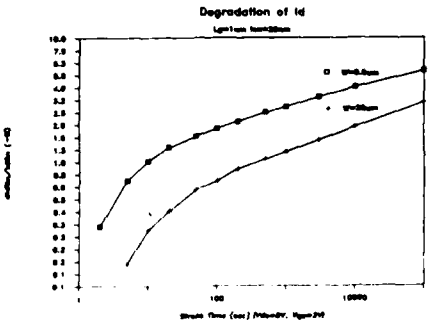


Fig.3 Enhanced Degradation of narrow width transistor.

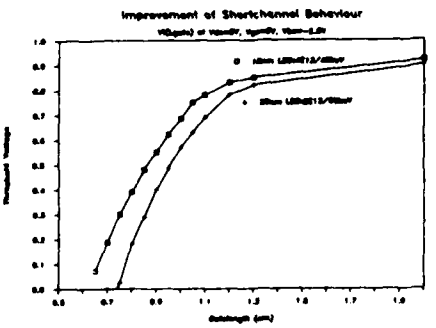


Fig.4 Short channel behaviour of transistors with t_{ox}=20nm and t_{ox}=18nm.

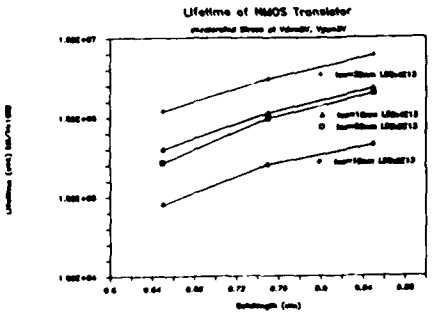


Fig.5 Lifetime after hot electron stress of transistor with t_{ox}=20nm and t_{ox}=18nm.

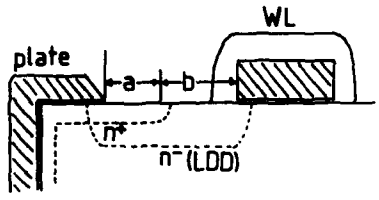


Fig.6 Cross-section of region between cell plate and transfer gate. n^+ is varactor.

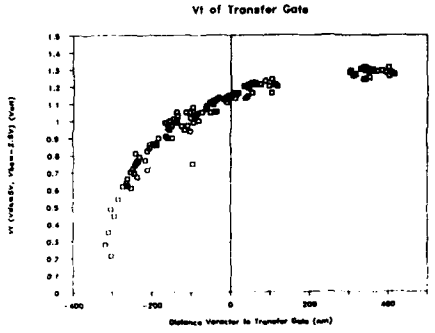


Fig.7 Deterioration of transfer transistor with overlapping varactor implant.

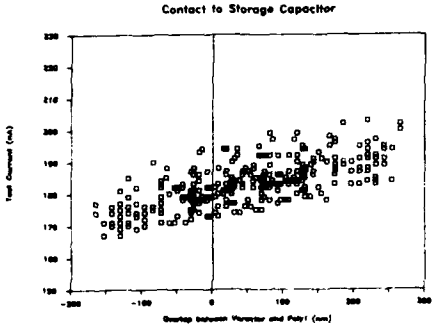


Fig.8 Contact to storage capacitor versus overlap of varactor.

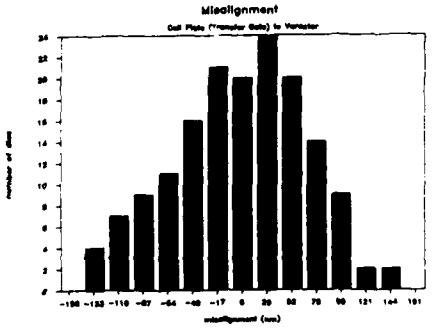


Fig.9 Misalignment of varactor to cell plate and transfer gate.

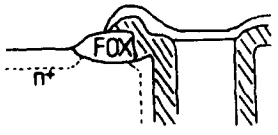


Fig.10 Worst case etch of FOX during trench etch process.

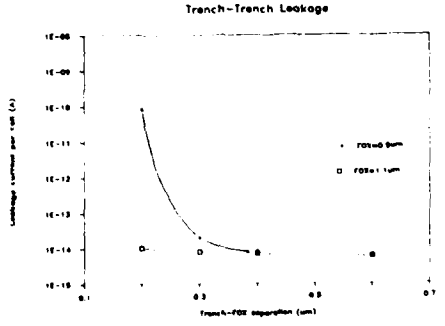


Fig.11 Dependence of trench leakage on length of LOCOS isolation. Vplate=2.5V, Vbs=-2.5V, Vtr-tr=5V.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CMOS 1 MICRON ISOLATION TECHNOLOGY USING INTERFACE SEALING BY PLASMA NITRIDATION : PLASMA SILO

P. DELPECH, B. VUILLERMOZ, M. BERENGUER, A.STRABONI and T. TERNISIEN

CNET Centre National d'Etudes des Télécommunications, Chemin du Vieux Chêne, BP 98, F-38243 Meylan Cedex, France

RESUME - Nous avons évalué l'apport de la technique d'isolation par SILO PLASMA dans une filière CMOS 1 µm par comparaison à un isolement par LOCOS classique. Le SILO PLASMA permet de réduire de 0.4 µm les pertes liées au LOCOS, ainsi que l'effet de canal étroit, tout en conservant les principales caractéristiques électriques de cette technique (courant sous le seuil, intégrité de l'oxyde grille, etc,)

ABSTRACT - The improvement of a 1 µm CMOS process using PLASMA SILO as an isolation technique has been evaluated by comparison with a classical LOCOS. The PLASMA SILO provides a reduction of 0.4 µm in the channel width loss, and a gain on the narrow channel effect. The other electrical characteristics are maintained (subthreshold characteristics, gate oxide integrity, etc,)

INTRODUCTION

For submicron design rules, new approaches have to be proposed for better control of lateral isolation. In order to avoid LOCOS limitations (i.e. Bird's Beak and lateral diffusion of channel stop implant) SILO has already been proposed [1]. In this paper we report an evaluation of the PLASMA SILO technique in a 1 micron CMOS process. This isolation technique uses a plasma sealed interface with few modifications in the standard LOCOS procedure; this constitutes a significant advantage for a rapid introduction into an industrial technology.

Application of plasma nitridation to the SILO [2] presents the main advantage to realise a very effective and reproducible barrier against lateral oxidation, by avoiding native oxide present between Silicon and deposited Si_3N_4. In addition, it is possible with this technique to grow the field oxide at a lower temperature, which reduces the dopant diffusion while keeping a short Bird's Beak value.

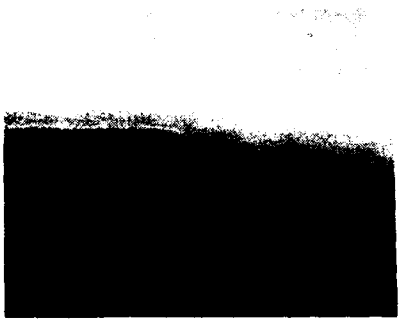
EXPERIMENTAL PROCEDURE

The sealing nitridation is achieved in a 13.56 MHz plasma with external capacitive coupling. The set up allows batch treatment and is now available as a production machine [3]. 4 to 5 nm thick Silicon Nitride is grown at 950°C in an ammonia flow at a pressure of 3.E-2 mBar(hPa) during 6 hours with a power of 800 W. These very thin layers were able to resist the severe oxidizing conditions (steam at 980°C) used to grow a 400 nm oxide on bare (100) Silicon. In the SILO process, the standard pad oxide (25 nm) / Si_3N_4 (90 nm) LOCOS mask is used over the sealing film, except that the thermal pad oxide is replaced by an LPCVD oxide.

RESULTS AND DISCUSSION

The Figure 1 shows the morphology of the resulting field oxides (700 nm) after steam oxidation at 980°C for the PLASMA SILO (a) and at 1050°C (reduced Bird's Beak) for the conventionnel LOCOS (b). The bird's beak length to oxide thickness ratio (L_{BB} / Tox) is 0.45 for the former as compared to 0.70 for the latter, as measured before mask removing.

a) PLASMA SILO BEFORE MASK REMOVING



b) STANDARD LOCOS AFTER MASK REMOVING

FIG. 1. SEM MICROGRAPHS OF BIRD'S BEAK EXTENSION

We have compared electrical characteristics of devices fabricated with both PLASMA SILO and optimised LOCOS (Field oxide LOCOS: 700 nm).
The main result is the reduction in the loss of channel width DELTAW = W_{mask} - W_{eff} by 0.4 µm, from 1.1 for optimised LOCOS to 0.7 µm for PLASMA SILO, with slight variation between N and P type (channel stop implanted NMOS side only).
The narrow channel effect is substantially reduced as seen in Fig 2, mainly due to the decrease in DELTAW. Thus, a transistor with a 0.5 µm effective width (1.2 µm designed width) is well controlled.

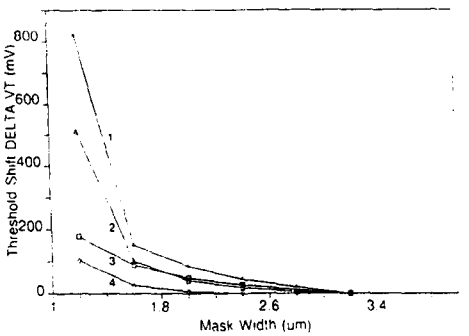


FIG. 2. NARROW CHANNEL EFFECT
COMPARISON SILO/LOCOS
SILO PMOS (1) SILO NMOS (2)
LOCOS PMOS (3) LOCOS NMOS (4)

A statistical analysis performed over 400 devices, shows (Figure 3) that the narrow channel effect is very weak, even for an effective width as low as 0.2 µm.

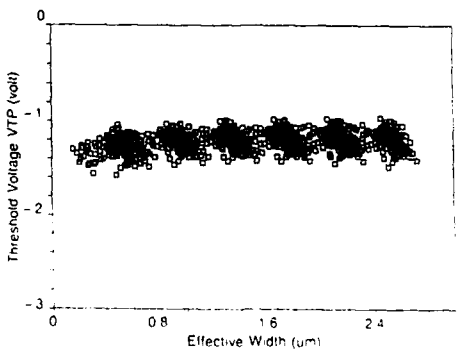


FIG. 3. NARROW CHANNEL EFFECT SILO PMOS

The subthreshold characteristics as measured on P-channel transistors , with L=1.6 µm and W between 3.2 and 1.2 µm are shown in Figure 4 for SILO and LOCOS respectively at Vd=-7 Volt. These curves show similar characteristics and in particular, no additional leakage current for the SILO process. The same behaviour is found for N-channel transistors.

Moreover, no kink appears showing the absence of parasitic sidewall transistors.

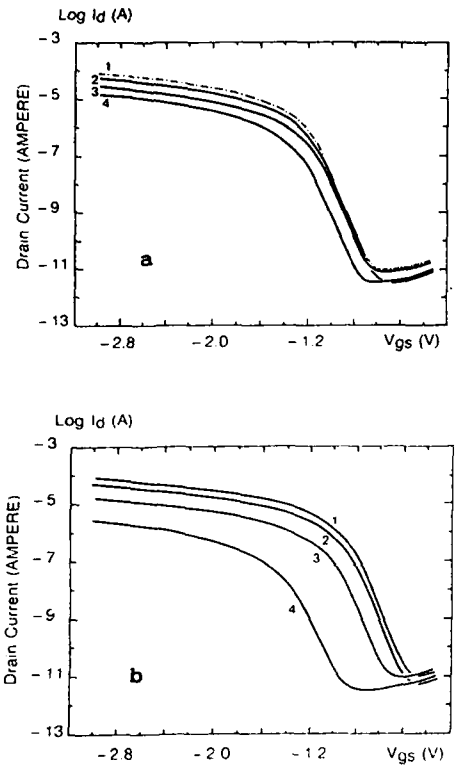


FIG. 4. SUBTHRESHOLD CHARACTERISTICS OF PMOS TRANSISTORS a) SILO b) LOCOS, V_D=-7 VOLT L=1.6 µM, W=3.2 µM (CURVE 1), W=2.4 µM (CURVE 2), W=1.6 µM (CURVE 3), W=1.2 µM (CURVE 4)

Junction leakage measurements were performed on diodes which exhibit different area(A)/perimeter(P) ratios : finger-type diodes with A=2.6E-4 cm^2 and P=1.7 cm, and rectangular-type diodes with A=6.8E-4 cm^2 and P=0.1 cm. The values of the perimeter leakage current at ± 5 V are reported in Table I. For the SILO process, an insignificant increase in the leakage value can be noticed.

TABLE I. PERIMETER LEAKAGE CURRENT I_R AT ±5 V FOR SILO AND LOCOS PROCESSES AND THE TWO TYPES OF DIODES N+/P AND P+/N

ISOLATION PROCESS	I_R (pA/cm) N+/P	P+/N
SILO	21.4	18.8
LOCOS	16.2	16.3

SEM observations on SIRTL etched samples oxidized at 980°C confirm the absence of crystalline defects at the boarder of the devices. This was, however, not the case on samples oxidized at 920°C, i.e.below the viscous flow transition temperature.
The SILO technique allows the same field threshold voltage $|V_{TF}|$ as the LOCOS, for N and P type, from 17 V for a 2.4 µm isolation spacing to 13 V for 1.2 µm spacing. The Id(Vg) characteristics of a field transistor are shown in Fig.5 for Vd=1 et 5 V. The leakage current is less than 1 pA/µm for both types of field transistors.

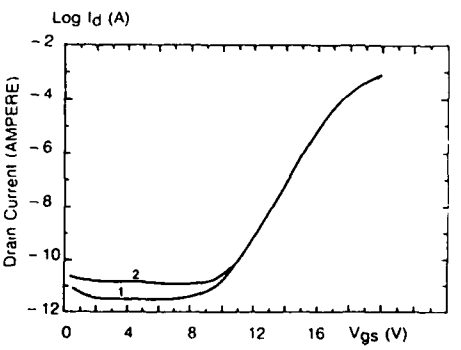


FIG. 5. SUBTHRESHOLD CHARACTERISTICS OF A NMOS FIELD TRANSISTOR V_D=1 V (1) AND 5 V (2)(W=1000, L=2 µM)

The effect of PLASMA SILO on the thin oxide gate integrity (25 nm) has been evaluated on capacitors with polycide gate. The defect density is not affected by such a technology, indicating that the mask has been efficiently removed.

CONCLUSION

We have presented an evaluation of the PLASMA SILO isolation which provides a significant gain in both the width loss and the narrow channel effect by comparison with the classical LOCOS, while the main electrical characteristics are preserved. Due to its simplicity (no additional masking level) this technology seems promising for submicrometer processes .

ACKNOWLEDGEMENTS

We would like to thank the technical staff of the pilot line for the processing of the wafers, and D.T.Amm and M.Haond for their critical review of the manuscript.

REFERENCES

[1] J.Hui, T.Y.Chiu, S.Wong, W.G.Oldham , IEDM p 220, 1982
[2] B.Vuillermoz, A.Straboni, R.Pantel, E.C.S. Meet., Toronto, Abstract 248, 1985
[3] French patent N° 83 18299

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A SUB-MICRON CMOS PROCESS EMPLOYING TRENCH ISOLATION

M.C. ROBERTS, P.H. BOLBOT and D.J. FOSTER

Plessey Research Caswell Ltd., Caswell, Towcester, GB-Northants NN12 8EQ, Great-Britain

Abstract A high density trench isolated CMOS process has been developed. Circuit designs have been initially fabricated at 1μm dimensions. Excellent device performance is demonstrated at channel lengths of 0.7μm and channel widths of 0.2μm showing the potential for fabricating circuits with 0.7μm and 0.5μm design rules.

INTRODUCTION

In order to achieve a high density CMOS process a trench isolation technique with both n+ and p+ diffusions abutted directly to the trench wall has been developed. The process has been initially developed with 1μm dimensions. The trench width, gate length, contact size, via size and metal spacing all being 1μm. Trench technology however allows the design rules to be easily scaled down to 0.5μm dimensions. The excellent performance of 0.7μm devices are demonstrated in this paper. By adopting trench isolation a packing density improvement greater than two was achieved over a 1μm LOCOS isolated process giving a packing density of approximately 4000 gates/mm^2. At 0.7μm geometries the packing density increases to 8000 gates/mm2. An improvement of a factor of two can also be achieved in performance since the improved packing density results in a large increase in speed due to reduced track capacitance.

Process Schedule

Fig.1 shows a schematic cross-section of the process after one layer of metal. The process is fabricated on p/p+ epitaxial material in order to eliminate latch-up. A 1μm wide and 2μm deep trench is used to isolate the devices. The trench is etched using an oxide/nitride/oxide mask. This ensures that the nitride over the active area is self aligned to the edge of the trench. The trench walls are passivated with layers of oxide and nitride and the trench filled with poly-Si by depositing a thick conformal layer followed by an etch-back. An active area mask is used to remove the nitride from the field regions prior to growing the field oxide and the trench capping oxide. The trench sidewall nitride prevents the encroachment of the capping oxide into the active area. This maintains a low value for delta W and also ensures that the source/drains abut directly to the trench sidewall.

The twin wells are formed by high energy P+ and B+ implants followed by a high temperature drive-in. The well profiles suppress the trench sidewall parasitics and also produce high punchthrough voltages for both nmos and pmos devices.

A thin gate oxide (200A) is used to improve the drive current capability. The poly-Si gate is etched undoped to allow the formation of both n+ and p+ poly-Si. This allows the threshold voltage of both nmos and pmos devices to be optimized without the need for threshold adjust implants. The p+ poly-Si also ensures that the pmos device has a surface channel and therefore is controlled more effectively by the gate.

A deposited oxide sidewall spacer is formed on the side of the poly-Si gate. This allows the electrical channel length to be maintained close to the gate length. It also allows a self aligned titanium silicide process to be incorporated. A doubly diffused n+ source/drain using phosphorus only was used in order to minimize hot electron effects.

A combination of low temperature oxide and spin-on-glass is used to planarise the wafers prior to metallisation. Two layers of metal are used at three micron pitch with polyimide used to planarise and isolate the metal tracks.

Device Results

The latch-up performance of the process was measured for a range of epitaxial thicknesses and resistivities (fig.2). The holding voltages were found to be considerably greater than the supply voltage.

One of the main concerns of a trench process is the existence of MOS parasitics along the trench wall. The poly-Si within the trench acting as the parasitic gate. Fig.3 illustrates the current paths of the lateral and vertical trench parasitics.

The behaviour of the trench parasitics is assesed by making contact to the poly-Si within the trench. This allows the subthreshold characteristics of the four parasitics to be measured. The results are shown in fig.4. Due to the well architecture used the threshold voltage of both nmos and pmos parasitics are maintained above 10 volt. The behaviour of the trench parasitics has been assesed at elevated temperatures. Fig.5 shows the change in the trench potential (measured at 1nA/μm and 10nA/edge) as a function of temperature. Even at 150°C the trench potentials are considerably higher than the supply voltage.

The performance of the nmos and pmos devices are summarised in table 1. The threshold voltages are optimised by making use of both n- and p-type poly-Si. This means that a high n-well and p-well dopant concentration profile ($5E16cm^{-3}$) can be used which maintains a high punchthrough voltage for the nmos and pmos device down to an Leff of 0.5μm. Fig.6 shows how the threshold voltage is controlled down to an Leff of 0.5μm. The figure also shows the change in threshold voltage with electrical width. Good control is maintained down to Weff of 0.2μm. The threshold voltage increases rapidly below 0.2μm due to the small birds beak that exists between the trench sidewall oxide and the active area. This increase in threshold voltage is accelerated for the pmos device due to the accumulation of negative charge at the sidewall oxide/Si interface.

The drain breakdown values of the nmos and pmos device are shown in table 1. The nmos drain breakdown is due to parasitic bipolar behaviour while the pmos drain breakdown is a result of junction breakdown. The IV characteristics of nmos and pmos devices are shown in fig.7. The bipolar parasitic is evident in the saturation curves of the 10/0.7 nmos device at voltages above 7 volts. The very narrow width devices (Weff=0.3μm) shows the ease with which the device width can be scaled using trench technology.

An important consideration in the performance of a sub-micron CMOS process is the degredation of the nmos device due to hot electron injection into the gate oxide. This can be reduced by minimising the electric field between the drain and well. The effectiveness of the DDD (doubly diffused drain) structure in reducing hot electron injection can be monitored by measuring the substrate current. Fig.8 shows how the substrate current varies with Leff. At an Leff of 0.7μm the substrate current is less than 1uA/μm width which correlates to a projected lifetime of greater than 10 years (delta Vt = 10mV).

Finally the performance of the process was assesed by fabricating 49 stage ring oscillators. The delay time was measured as a function of supply voltage for an unloaded inverter chain and for a loaded NAND chain (loaded with three gates). The results are shown in fig.9. A best stage delay time of 120 psec was measured for the unloaded inverter chain (Leff=0.9μm) .

To conclude, a trench isolated process offers an excellent improvement in packing density as well as providing good circuit performance. The process also allows circuit designs to be easily scaled from the present designs at 1μm down to 0.5μm dimensions.

Acknowledgements

The authors would like to thank the Plessey Company for permission to publish and the Alvey Directorate for their financial support.

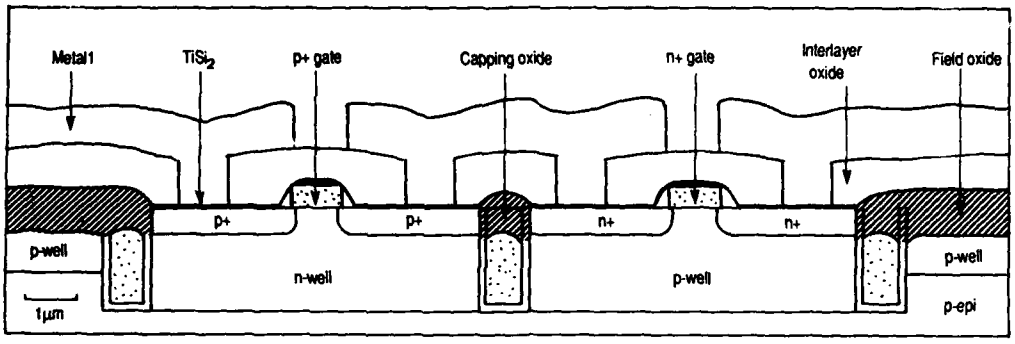


Fig. 1 CROSS-SECTION OF THE PROCESS AFTER 1 LAYER OF METAL

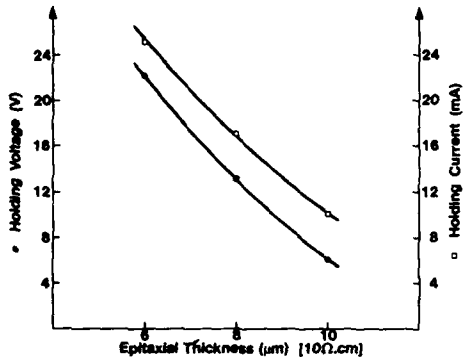


Fig. 2 LATCH-UP PARAMETERS AS A FUNCTION OF EPITAXIAL THICKNESS

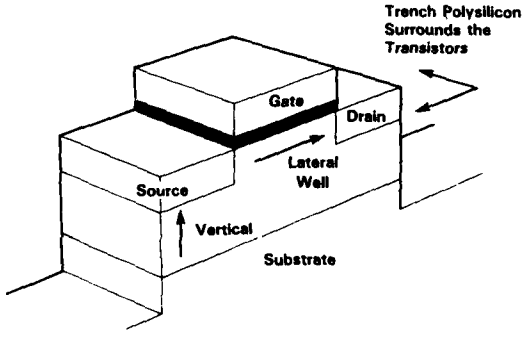


Fig. 3 SCHEMATIC SHOWING THE CURRENT PATHS OF THE TRENCH PARASITICS

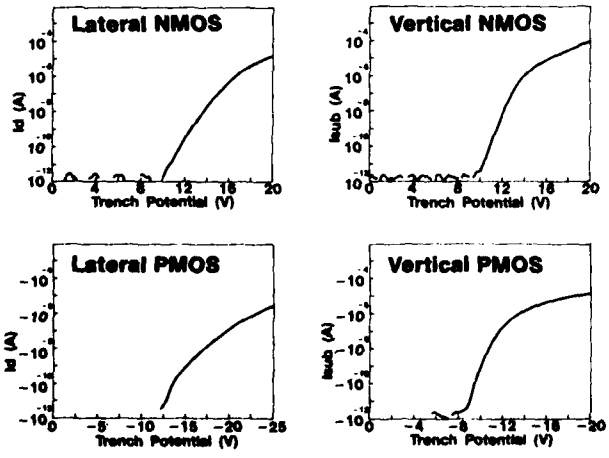


Fig. 4 SUBTHRESHOLD CHARACTERISTICS OF THE FOUR TRENCH PARASITICS

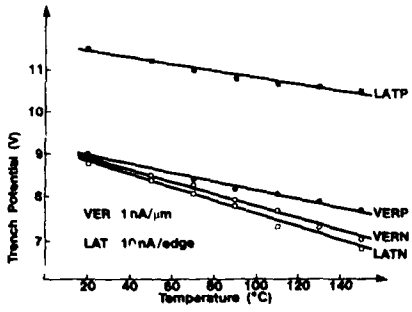


Fig. 5 DEPENDENCE OF THE TRENCH POTENTIAL ON TEMPERATURE

Table 1

	NMOS	PMOS
Vt(V)	0.7	−0.65
BETA00	85μA/V2	−23μA/V2
Drive Current (W = 10 μm)	3.0 mA	−1.8 mA
SUBVTS (mv/dec)	95	87
Off − state Leakage	<1E − 13A/μm	<1E − 13A/μm
Drain Breakdown (Leff = 0.7)	7V	−8V
Delta L (μm)	0.1	0.1
Source/Drain Sheet Res.	4Ω/sq.	4Ω/sq.

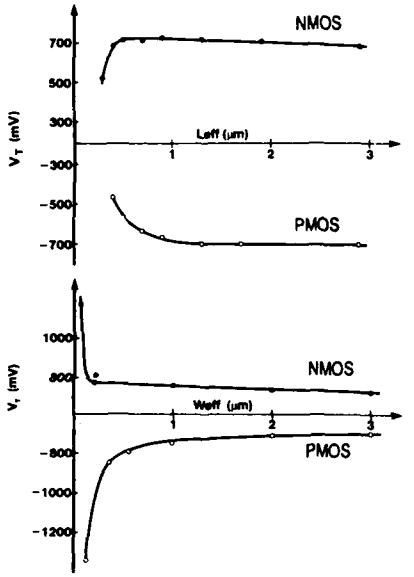
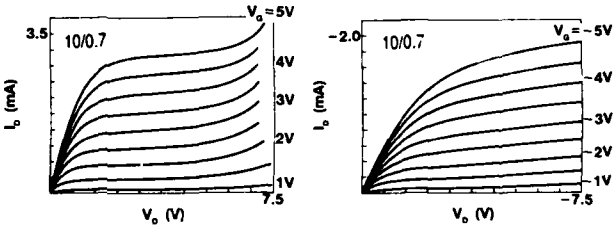


Fig. 6 DEPENDENCE OF THE THRESHOLD VOLTAGE ON Leff AND Weff



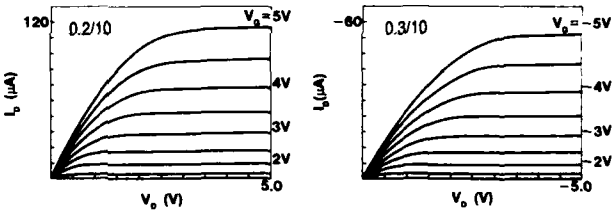


Fig. 7 IV CHARACTERISTICS OF NMOS AND PMOS DEVICES

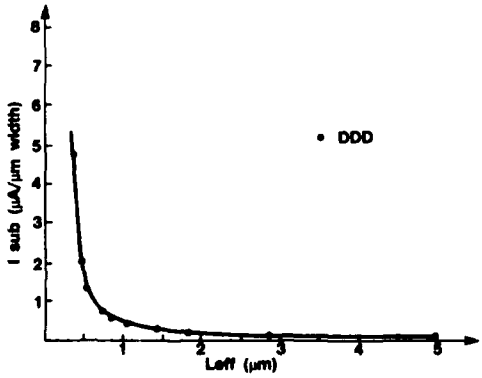


Fig. 8 DEPENDENCE OF SUBSTRATE CURRENT ON Leff

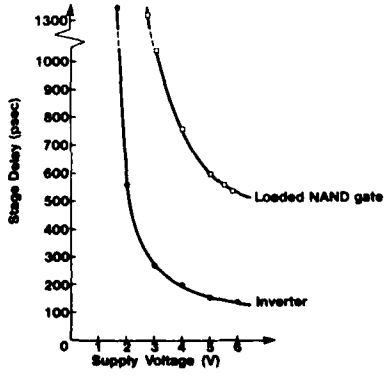


Fig. 9 RING OSCILLATOR STAGE DELAY AS A FUNCTION OF SUPPLY VOLTAGE (Leff = 0.9μm)

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A COMPARISON OF TRENCH FILLING MATERIALS FOR SUB-MICRON CMOS

P.H. BOLBOT, M.C. ROBERTS and P.L. MEDHURST

Plessey Research Caswell Ltd., Caswell, Towcester, Northants., GB-NN12 8EQ, Great-Britain

Abstract - This paper discusses the fabrication of trench isolated CMOS using different filling materials. Preliminary studies of dielectric films have been compared with polysilicon for trench-filling ease; resistance of trench material to subsequent erosion, thermal stability, stress generation and parasitic transistor supression. Polysilicon filling is less sensitive to trench shape for ease of processing well-filled trenches. Corner effects seen with oxide-filled trenches (similar to those seen on SOI) can be eliminated using the polysilicon filling technique and processing modifications can be made to totally supress parasitic sidewall devices. The additional use of polysilicon as resistors or capacitors makes it the most favoured technique.

1 - INTRODUCTION

In order to increase the packing density of sub-micron CMOS integrated circuits an isolation technique superior to the conventional LOCOS approach is required. A trench isolated CMOS process offers the advantages of considerable reduction in minimum n+ to p+ spacing together with improved electrical performance. This paper discusses the fabrication of trench isolated CMOS using different filling materials. Polysilicon, which is currently the most frequently used trench filling, has been extensively characterised. Oxide, with its lower parasitic capacitance, may have advantages over polysilicon for sub-micron CMOS and preliminary studies of several dielectric trench filling techniques are investigated.

Properties of a successful trench isolation technique are considered to be: ease of filling trenches; resistance of trench material to erosion during subsequent wet chemical etches; chemical stability of filling during additional high temperature treatments; low stress in the surrounding silicon substrate (doped oxides which flow more easily induce less stress than undoped oxides); vertical and lateral parasitic transistor suppression.

2 - TRENCH FILLING MATERIALS

2.1 Dielectric

Thermal Oxide. To prevent dimensional increase, a sacrificial polysilicon layer was deposited in trenches following the growth of a sidewall oxide (1500A). After polysilicon oxidation,the effects of severe stress were seen in silicon regions near trenches with defect densities as high as $10^6/cm^2$.

Spin-On Glass. Up to three layers of glass were spun on to wafers with etched trenches. Densification in N_2 took place after deposition of each layer. The technique showed excessive cracking of the oxide between layers and at oxide-silicon interfaces.

Low Pressure CVD Oxide. Films showed trench coverage to be no better than 20% of top thickness after high temperature reflow (1000°C) in steam. Voids were present which were opened by subsequent etches.

Plasma CVD Oxides. Poor step coverage was also seen with plasma deposited and low pressure plasma enhanced CVD oxide films. Post-deposition reflow anneal (800 - 950°C) in steam improved trench filling but films became unstable during a subsequent high temperature N_2-anneal with large, irregular voids appearing in the trench oxide.

Atmospheric Pressure CVD Oxide. Better filling was seen with atmospheric pressure CVD oxide: sidewall to top coverage being around 50% (Fig.1). Films were of sufficient thickness to close off trenches up to 1.5μm wide but elongated voids remained in trenches as a result of oxide cusping over top corners of trenches. BSG (3.5 or 4.5 wt%) and BPSG (3.5/3.5, 4.0/1.5, 4.5/1.5, 2.5/1.5 wt%) films were evaluated using post-deposition re-flow conditions of 1000 - 1100°C in N_2. All films flowed sufficiently between these temperatures. After reflow, voids either disappeared or became buried deeply enough to avoid being opened up on etch-back. Film surfaces became smooth after reflow but variations in oxide thickness between field regions and areas with dense trench patterning made the etch-back stage more complex (Fig.2). Where processing could not compensate for regional thickness variations, poor filling of trenches resulted; this showed up electrically as parasitic corner transistor action (similar to that seen on SOI). Both void formation at deposition and regional film thickness variations after reflow were seen to be a function of trench shape; thus, sidewall profile optimisation is necessary for processing simplicity.

Film resistance to buffered hydrofluoric acid etches during later stages of processing increases with boron and decreases with phosphorous content. Of the films investigated, only 3.5/3.5 wt% BPSG showed a slightly enhanced etch rate over thermal oxide; all other films etched at a slower rate. Stress generated defects were not seen on fully processed samples with trench spacings down to 0.3μm.

Devices have been fabricated with well-filled atmospheric pressure CVD oxide trenches (Fig.3). Good electrical characteristics have resulted with gains of $84\mu A/V^2$(NMOS) and $25\mu A/V^2$(PMOS) for transistors with unit aspect ratios; threshold voltages were 0.75 and -1.0V respectively. Fig.4 shows devices with good subthreshold slope characteristics. These results, together with junction breakdown voltages, were comparable with those seen on polysilicon filled trenches.

TEOS. BPSG deposition by the pyrolytic decomposition of tetraethylorthosilicate (TEOS) showed better coverage of trench sidewalls than all other dielectrics, being about 2:3 (trench sidewall:top coverage). As a result, flowed film thickness variations between field and active/trench areas were reduced. Trenches separated by 0.3μm showed no stress induced defects after the reflow stage. This technique therefore shows potential although full device processing has not yet been completed.

2.2 Polysilicon

LPCVD polysilicon was deposited on 200nm of trench sidewall oxide and a thin film of nitride used to maintain sub-micron dimensions during isolation heat treatments. This technique takes advantage of the conformal coating properties of polysilicon, its ease of deposition and good planarisation control during etch-back (Fig.5). A capping oxide is grown on the polysilicon after etch-back (Fig.6); the film being suffiently thick to allow for some erosion during subsequent etches without causing detrimental parasitic transistor action. Trench widths and spacings can be as low as 0.3μm using this technique; stress generated defects occur only below this value and are found to correlate with junction leakage currents.

The parasitic action of the trench polysilicon acting as an unwanted gate electrode has been fully evaluated. The sidewall dielectrics together with optimised well doping have ensured that effects are minimal. Trench polysilicon voltages for a vertical trench leakage current of 1pA/μm trench length are typically 8.0V and -9.5V for NMOS and PMOS transistors respectively.

3 - DISCUSSION & CONCLUSION

BSG and BPSG oxide films can be flowed easily to fill trenches at temperatures compatible with device fabrication. The processing ease of the etch-back stage is largely dependent on the shape of the trench; thus, under optimum conditions with a slight positive bevel on the sidewall, good control over the oxide level in the trench should be possible. Initial filling of well-shaped trenches with oxide is therefore not considered to be a problem and erosion by subsequent wet chemical etches can be minimised by dopant-tailoring. However, some reduction of oxide level in the trench is seen to be inevitable during processing and has the effect of enhancing corner parasitic action. In addition, residual polysilicon

fillets may be left at trench edges after gate etching due to thickening of the gate over trench edge steps; short circuiting between devices could result.

Good filling of trenches using polysilicon films is less sensitive to trench shape due to its conformal coating properties. Corner parasitic action caused by poorly filled trenches at end-of-process is not a problem as the capping oxide can be tailored to allow for a known amount of in-process erosion. By filling trenches with polysilicon, parasitic sidewall devices can be totally suppressed through the use of optimised well doping and sidewall dielectrics. Stress is minimised and transistors with active device widths as low as 0.3μm have been fabricated. Polysilicon also provides increased flexibility over oxide filling by having additional use as resistors or capacitors. Thus polysilicon trench filling remains the most favoured trench isolation technique.

Acknowledgements

The authors would like to thank the Plessey Company for permission to publish and the Alvey Directorate for their financial support.

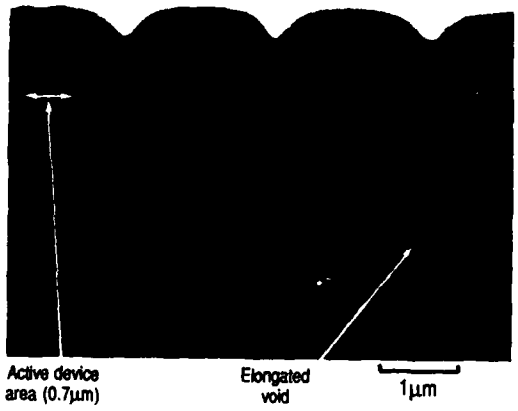


Fig.1 - As deposited atmospheric pressure CVD oxide filled trenches (1μm film)

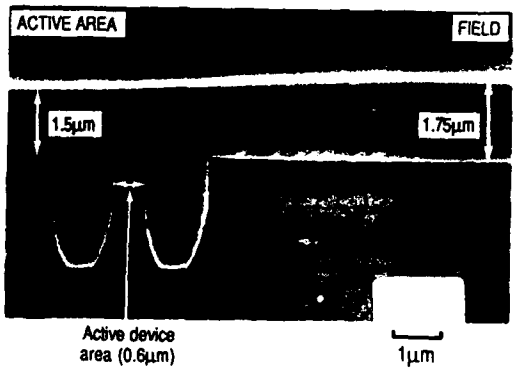


Fig.2 - BPSG after high temperature reflow showing variations in oxide thicknesses

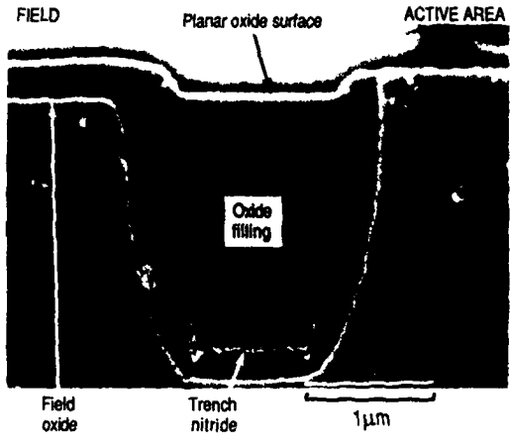


Fig.3 - BPSG filled trench on fully processed wafer

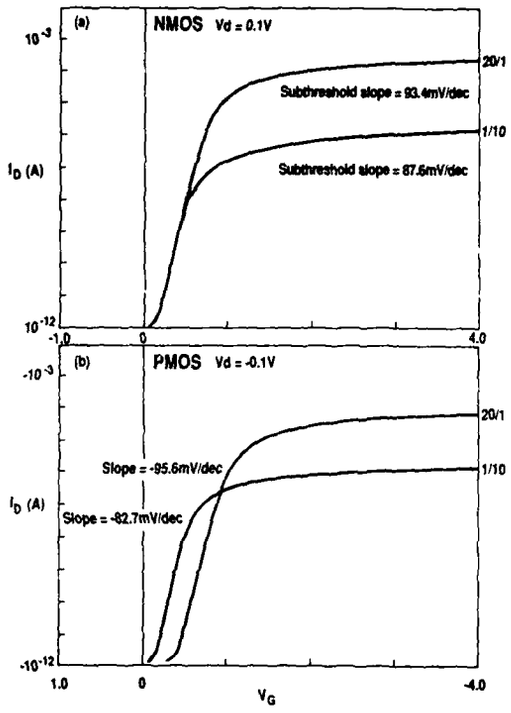


Fig.4 - Subthreshold characteristics for transistors with oxide filled trenches

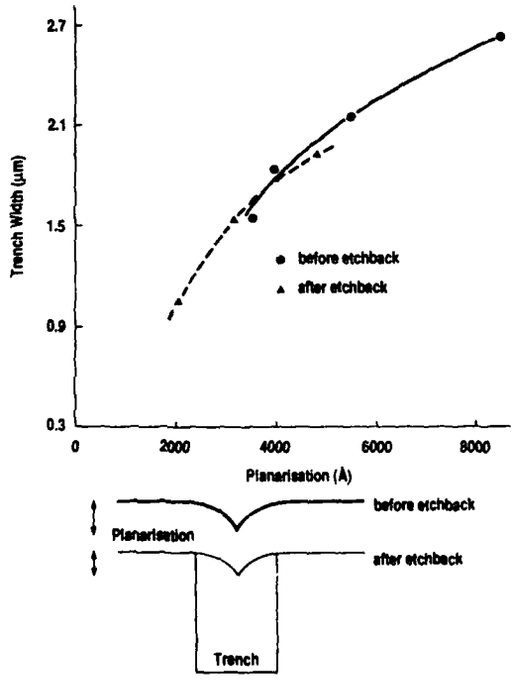


Fig.5 - Planarisation of 1.5µm polysilicon film before and after etch-back

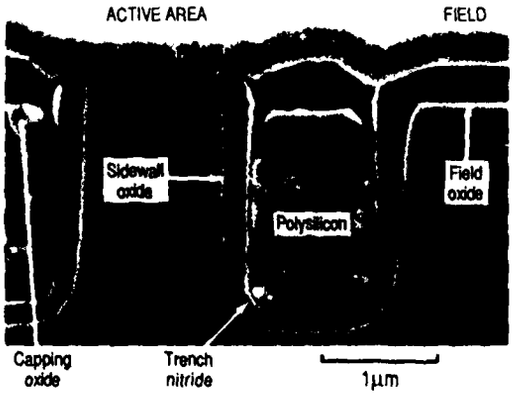


Fig.6 - Polysilicon filled trench structure

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A NOVEL BOROSILICATE GLASS (SiOB-BSG) BY LOW PRESSURE DECOMPOSITION OF A MONOMOLECULAR LIQUID PRECURSOR

H. TREICHEL, F.S. BECKER, D. FUCHS and T. KRUCK*

Siemens AG, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.
**Inst. f. Inorganic Chem., University of Cologne, Greinstr. 6, D-5000 Köln, F.R.G.*

Résumé

Couches minces de BSG avec du composition constante (ca. 4.7 wt% B) sont préparés en LPCVD, la source était Tris(trimethylsiloxy)boron $B[OSi(CH_3)_3]_3$ une molécule se compose de bore et de silicium. Les conditions optimales pour le dépôt sont 800 °C, 105 Pa (800 mTorr) et l'addition de 280 sccm O_2. Les couches sont stable exposées à l'air atmospheric, la couverture d'arêtes aiguës est excellente et la vitesse d'une corrosion humide est trés lente. Il est prouvé qu'ils sont des sources dotant effectifs dans une zone de 800 °C à 1000 °C.

Abstract

BSG with constant composition (ca. 4.7 wt% B) was deposited using Tris(trimethylsiloxy)boron $B[OSi(CH_3)_3]_3$ in a standard LPCVD system. The optimum deposition conditions are 800 °C, 105 Pa (800 mTorr) and 280 sccm O_2 purge. The films are chemically stable in atmosphere and exhibit good step coverage and low wet-etch rates. They are effective dopant sources in the 800-1000 °C range.

Introduction

The progress in silicon technology has revived the interest in doped oxides as diffusion sources. This is because ion implantation can damage the silicon lattice or suffers from geometrical restriction when the doping of trench walls is attempted / 1 /. This problem has been solved e.g. by using arsenosilicateglass (AsSG) / 2 / deposited by the use of two liquid precursors, a simple and straightforward LPCVD process / 3 /.
Another application proposed recently / 4 / is the realization of shallow junctions in bipolar technology. A borosilicateglass (BSG) is deposited and acts as a diffusion source during a subsequent heat-treatment. The investigations have demonstrated that the electrical results obtained depend critically on the boron concentration in the deposited layer.
Organic molecules have shown to be superior to silane / 5-9 /. Thus the pyrolysis of tris(trimethylsiloxy)boron $B[OSi(CH_3)_3]_3$ was used to produce high quality BSG-films. The necessity to control two liquids with different vapor pressures (e.g. TEOS and TMB) is in this case circumvented because the molecule contains already silicon, oxygen and the dopant boron. This results in films of inherently constant composition. The feasability of using a monomolecular precursor for the depostion of a binary glass had been demonstrated for the case of alumosilicate glass before / 10 /. To our knowledge, however, no attempts have been made to extend this principle to glasses like e.g. BSG or AsSG used in highly integrated circuits.
The process characteristics and film properties of the BSG obtained by the novel process will be described in the following section.

Experimental

The experiments were carried out in a conventional low pressure chemical vapor deposition (LPCVD) setup described before / 7 /. The BSG layers were deposited in a 235 mm diameter quartz tube on 100 mm, <100>, 5 Ωcm, p-type wafers. Open boats with a wafer spacing of 4.7 mm were used. The highly purified tris(trimethylsiloxy)boron was supplied by the University of Cologne (Inst. f. Inorganic Chemistry). The metal bubbler containing the liquid source was kept at temperatures between 25 °C and 35 °C. The main process parameters were tem-

perature (500-900 °C), pressure (80-400 Pa) and the amount of oxygen added (40-280 sccm).
The films were characterized regarding deposition rate, film thickness variation, etch rate, step coverage, particles, boron content in the layer and the dopant effectiveness as described before / 6-7 /.

Process characteristics

The dependencies of the deposition rate and the film thickness variation on the main process parameters are illustrated in the first three figures. Figure 1 shows the influence of a variation of the deposition temperature for a relatively high pressure and O_2-flow. The optimum working point is 800 °C, about 80 °C higher than in the case of TEOS / 6-8 /. Higher temperatures result in severe non-uniformities, probably due to depletion effects of the precursor. The temperature dependence of the deposition rate does not follow an Arrhenius type behaviour which hints at a complexe dissociation mechanism. The boron content of the films remains nearly unchanged, with a slight tendency towards higher boron concentration for higher temperatures. This is in contrast to results obtained with alumosilicateglasses, where a strong dependence of the film composition was observed / 10 /.

Figure 2 illustrates the influence of pressure. The process is stable up to 160 Pa with particle counts typically below 10 ($\leq 0.5\ \mu m^2$). Note the distinct increase in thickness non-uniformity above 105 Pa. At 400 Pa a very high number of particles indicate gas phase reaction.

The role of the O_2 purge is demonstrated in figure 3. Additional oxygen clear ly enhances the film growth. At higher O_2 flow tis effect levels off, probably due to dilution of the SiOB-compound. The intermediate maximum of the thickness variations is in contrast to results observed for TEOS/TMB-BSG / 5 / and defies a simple explanation. The boron content of the samples remains at a quite constant level at 4.7 wt% boron, independant of the amount of oxygen added.

For any doped glass chemical stability in atmosphere is a matter of constant concern / 9 /. In this respect, SiOB-BSG films compare favorably with TEOS-BSG which can develop particles when stored in humid atmosphere if the boron weight fraction is ≥ 5 wt%. SiOB-BSG is perfectly stable when deposited at 800 °C. Cristallization of boric acid has been observed on films produced at 700 °C or by a mixed SiOB/TEOS deposition. Therefore it can be assumed that the glass formation process is rather completed in the SiOB-films deposited at 800 °C. This assumption is corroborated by studies of the wet-etch rates in different media / 7 /. It is evident that the rates of densified TEOS-BSG correspond to the values of undensified SiOB-BSG films.

Infrared spectroscopy (IR) is a fast and non-destructive tool to determine the dopant level in glasses / 9 /. The IR spectrum of SiOB-BSG does not exhibit any specific feature when compared to TEOS-BSG.

In modern integrated circuits with high aspect ratios, the step coverage of deposited films is a very important property. The superior results achieveable with organic compounds have been demonstrated for doped and undoped TEOS-oxides by several authors / 2, 3, 5-9 /. SiOB-BSG fits well into this pattern with its step coverage of about 70 % as shown in figure 4.

To test the effectiveness of the BSG as a dopant source, films were deposited on wafers which had been given a HF-dip before being loaded into the furnace. After the BSG deposition an in-situ drive-in was performed by ramping the temperature up to 900 °C for 30 minutes. Figure 5 demonstrates that this thermal budget is sufficient to obtain a boron concentration of about $2*10^{20} cm^{-3}$. By variying time and temperature of the drive-in this could easily be raised.

Conclusion

The novel SiOB-BSG process described produces films of superior quality while eliminating problems stemming from the regulation of different vapors. The films are effective dopant sources with inherently constant composition. Monomolecular liquid precursors should be of interest for CVD of doped insulating films (e.g. intermetalinsulators like BPSG) as well / 9 /.

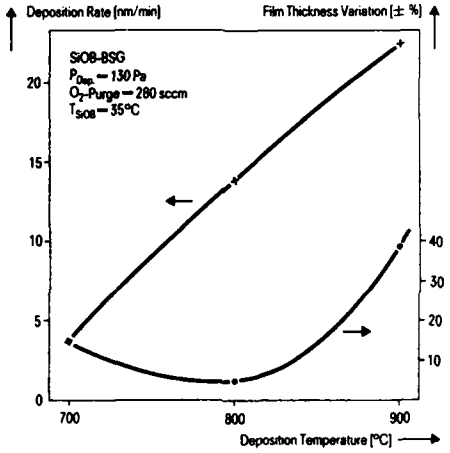


Fig. 1 **Deposition Rate and film thickness variation (SIOB-BSG) vs deposition temperature**

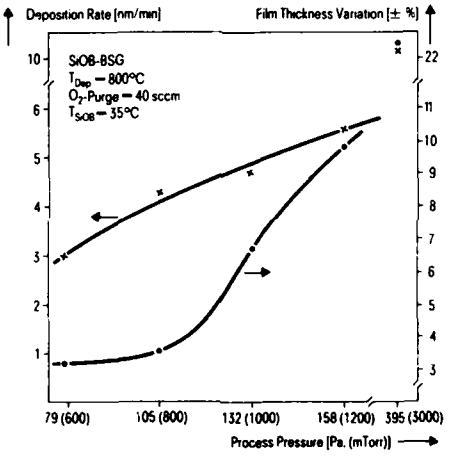


Fig. 2 **Deposition Rate and film thickness variation (SIOB-BSG) vs process pressure**

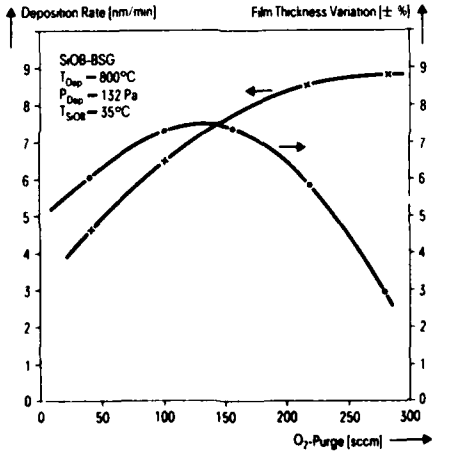


Fig. 3 **Deposition Rate and film thickness variation (SIOB-BSG) vs additional O_2 purge**

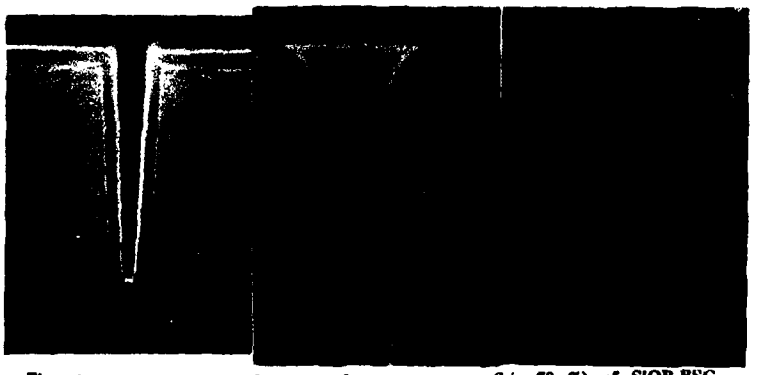


Fig. 4 **Example of the good step coverage (b/a 70 %) of SIOB-BSG in a deep trench**

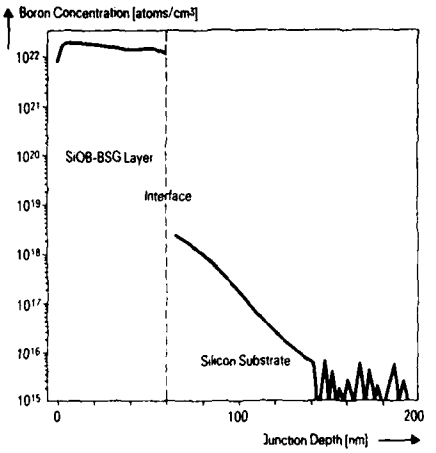


Fig. 5 ***SIMS-concentration profile in SIOB-BSG and silicon substrate (after in-situ anneal of 30 min/900 °C)***

References

/ 1 / H. Binder, H. Geiger, R. Kakoschke, H.M. Mühlhoff, S. Röhl
18th Conf. Sol. State Dev. and Mat., Tokyo
Proc. Vol., p. 299 (1986)

/ 2 / K. Yamada, K. Yamabe, Y. Tsunashima, K. Imai,
T. Kashio, H. Tango
IEDM Technol. Dig. 1985, p. 702 (1985)

/ 3 / F.S. Becker, H. Treichel
Proc 6th European Conf. on CVD, p. 207 (1987)

/ 4 / M. Nakamae
Proc. of ESSDERC, Bologna (9/1987), p. 361-363

/ 5 / R.M. Levin, K. Evans-Lutterodt
J. Vac. Sci. Technol., B1, 54 (1983)

/ 6 / F.S. Becker, D. Pawlik, H. Schäfer, G. Staudigl
J. Vac. Sci. Technol., B4 (3), 732 (1986)

/ 7 / F.S. Becker, D. Pawlik, H. Anzinger, A. Spitzer
J. Vac. Sci. Technol., B5 (6), 1555 (1987)

/ 8 / F.S. Becker, S. Röhl
J. Electrochem. Soc. 134, 2923 (1987)

/ 9 / F.S. Becker
Plenum Publ. Comp., New York (Nato ASI K-87 8684)
to be published

/ 10 / Y. Koga, M. Matsushita, K. Kobayashi, Y. Nakaido,
S. Toyashima
Thin Film Dielectrics
Electrochem. Soc., New York, p. 355 (1969)

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

OPTIMISATION OF SELECTIVE POLYSILICON OXIDATION FOR 0.8µm-TECHNOLOGY

R. BURMESTER, M. KERBER and C. ZELLER

Siemens AG, Corporate Research and Development, Microelectronics, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

Abstract-The dependence of bird's beak length on process parameters has been studied for selective polysilicon oxidation. The gate oxide thinning at the field oxide edge is correlated with the voltage drop across the gate oxide for constant current stress. We show, that with an optimised set of process parameters the bird's beak length can be reduced to 0.15µm without deteriorating device reliability.

1. INTRODUCTION

In VLSI technology isolation regions with approximately 1µm feature size have to be realised in order to achieve high packing density. On the other side the field oxide thickness must not be reduced to avoid increase of leakage currents and parasitic capacitances. Rather sophisticated isolation processes like BOX or trench have been suggested, which have the disadvantage of a complicated and costly process flow and an increased risk of defect generation. In our work we studied, to which extent a technologically simple process like selective polysilicon oxidation /1/ can be applied for defect free isolation in VLSI.

2. PROCESS FLOW

Selective polysilicon oxidation is explained in fig.1: A pad oxide is grown thermally, a polysilicon layer and a CVD-nitride layer are deposited successively. The nitride layer is patterned by photolithography and dry etching (fig.1a). The field oxide is grown in a wet ambient. Nitride mask and remaining polysilicon are removed. Due to the interface between polysilicon and nitride a second beak is formed in addition to the normal bird's beak (fig.1b). In order to remove any nitride residues in the active areas a reoxidation of 100nm is performed (fig.1c). This oxide is stripped together with the remaining pad oxide in HF. During the same step the length of the normal bird's beak is reduced and the second beak is etched off. Eventually the gate oxide is grown (fig.1d). Fig.1e shows a SEM-picture of an isolation oxide, where a pad oxide thickness of 50nm and a polysilicon thickness of

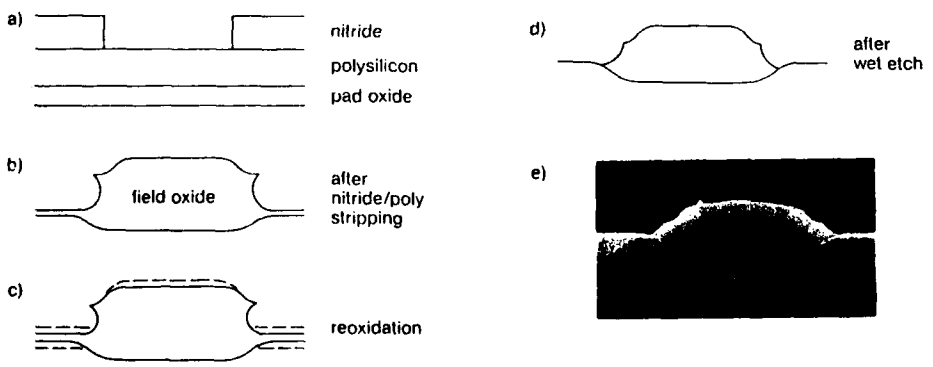


fig. 1.a-d Process flow of Selective Polysilicon Oxidation
fig. 1.e Field oxide (pad oxide 50nm, polysilicon 150nm, nitride 250nm)

150nm have been used. The as-grown field oxide thickness was 850nm, the wet etch time 100sec. The resulting bird's beak length is 0.3µm.

3. *EXPERIMENTAL RESULTS*

In order to optimise this technology for a 0.8µm-process, the bird's beak length has to be reduced to 0.1 - 0.15µm. In our experiments we varied the process parameters according to an orthogonal table, which provides reliable results with a minimised number of experiments /2/. The dependence of bird's beak length on the thickness of pad oxide and polysilicon layers is found by a statistical evaluation /3/ of the data (fig.2). The final field oxide thickness is 600nm for all sets of parameters. A reduction of bird's beak length is possible by using

a) thinner pad oxide, because oxygen diffusion in the pad oxide and lateral oxidation are suppressed. A minimum thickness of 30nm is necessary, to provide an etch stop and sufficient reduction of nitride stress.

b) thicker polysilicon, because a larger part of the field oxide is grown in the polysilicon, where lateral oxidation is small. In consequence step height and step angle of the field oxide are increased (fig.3a).

c) longer wet etch time. The bird's beak length can be reduced very effectively (fig.3b), but field oxide thickness is lost by the same amount.

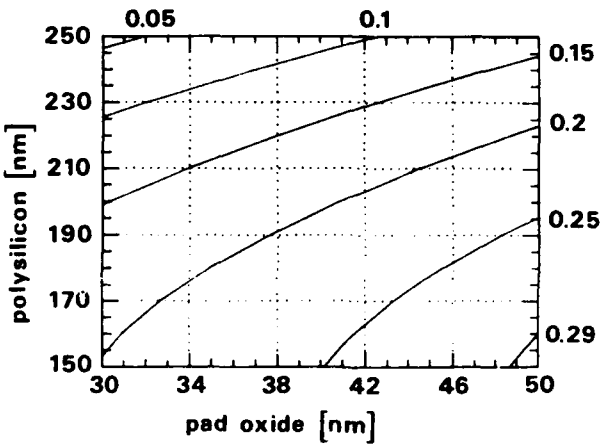


fig. 2 Bird's beak length as a function of pad oxide and polysilicon thickness (nitride 250nm)

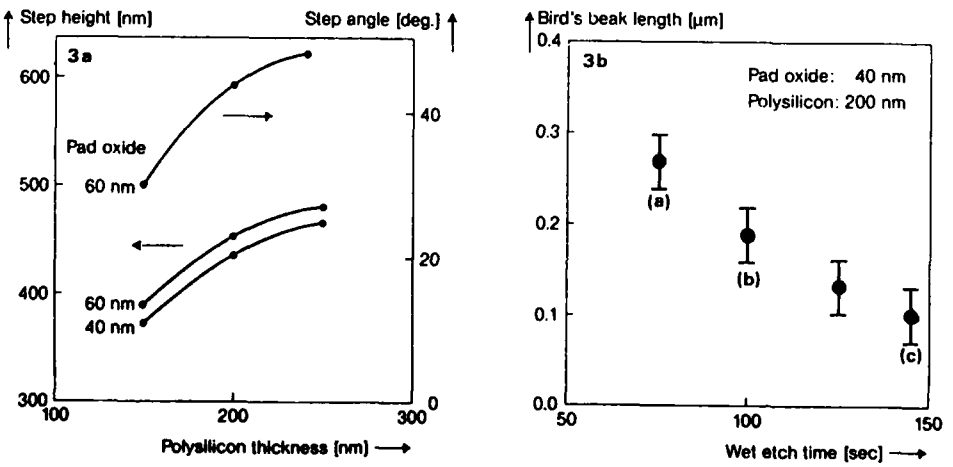


fig. 3a Step height and step angle of field oxide
fig. 3b Dependence of bird's beak length on wet etch time

At the field oxide edge a thinning of the gate oxide is observed, which is caused by stress induced reduction of oxygen diffusivity /4/ and by a reduced oxygen supply at the field oxide edge during gate oxidation /5/. If process parameters are varied in order to minimise bird's beak length, care has to be taken, that device reliability is not deteriorated by increased gate oxide thinning. The thinning becomes more pronounced, if the step angle of the field oxide or the wet etch time is increased. In fig.4a-c we show TEM-pictures for samples with different gate oxide thinning.

An increased thinning can be correlated with a higher Fowler-Nordheim tunneling current through the gate oxide for a given test geometry. Consequently the voltage drop U_{inj} across the overall gate oxide under constant current injection depends sensitively on the thinning (fig.4d). Thus it is possible to determine the gate oxid thickness at the field oxide edge by a simple electrical measurement of the voltage U_{inj}. In our experiments we used test structures with a gate area of 8mm² and a boundary length of 3000mm.

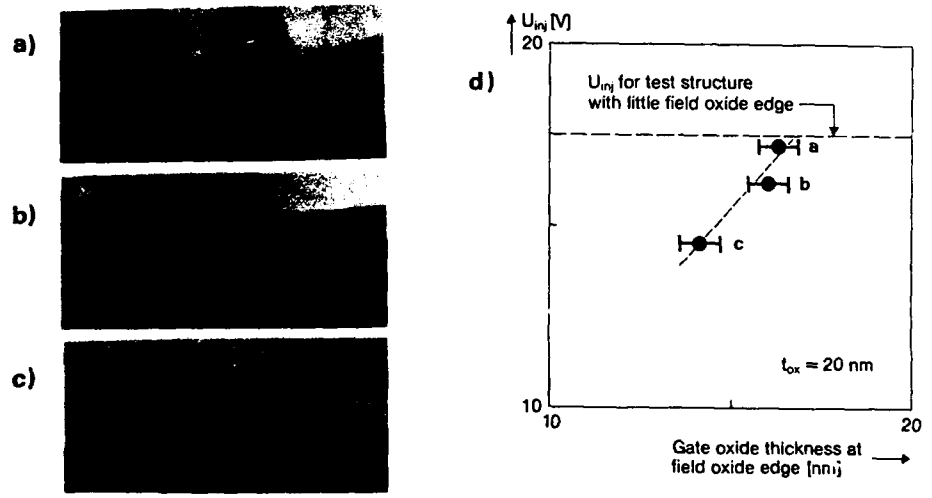


fig. 4a-c Gate oxide thinning for different wet etch times (see fig. 3b)
fig. 4d Voltage drop over gate oxide for constant current injection

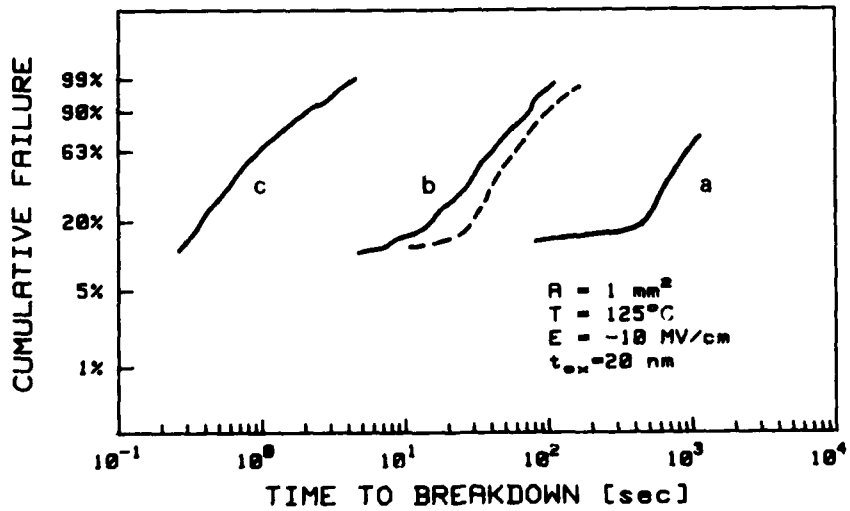
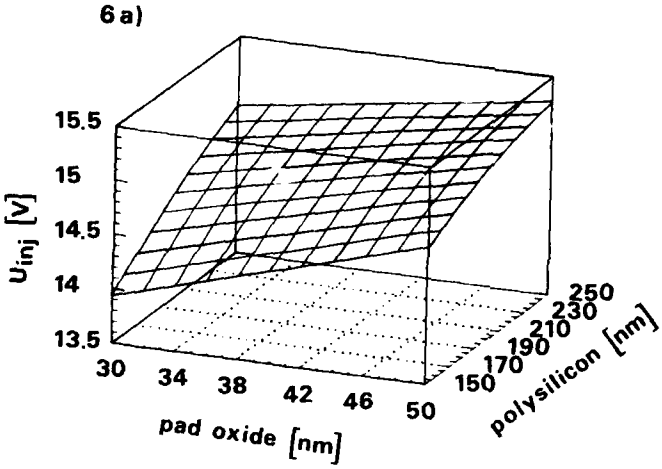


fig. 5 Time dependent dielectric breakdown due to constant bias stress.
Gate oxide thinning at field edge: a: 16.9% b: 21.0% c: 26.3%

Constant voltage stress leads to enhanced field strength in the thinned regions. Consequently the locally enhanced injection current reduces the intrinsic time to breakdown (t_{BD}) by orders of magnitude (fig.5). The dashed line in fig.5 represents the cumulative failure of sample c with 30% thinning but reduced field strength of -9 MV/cm. The data coincide with the cumulative failure of sample b with 21% thinning (field strength -10 MV/cm). This shows that the reduced t_{BD} is in fact due to the thinned regions. Therefore increased gate oxide thinning may cause reliability problems, if the intrinsic t_{BD} limits device lifetime.

U_{inj} is reduced by a thinner pad oxide, because the bird's beak angle is increased, and a thinner polysilicon, because the field oxide grows deeper into the substrate (fig.6a). Comparing figs.2 and 6 we find, that the bird's beak length can be reduced without increasing the gate oxide thinning by simultaneously varying pad oxide and polysilicon thickness. Choosing a pad oxide of 30nm and a polysilicon layer of 200nm the bird's beak is reduced to 0.15µm (fig.6b). The gate oxide thinning remains unchanged, however, compared with the isolation process shown in fig.1d. The step height increases by only 15%.



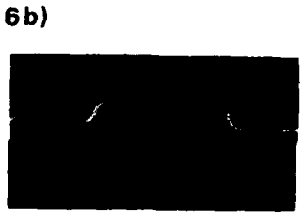


fig. 6a Voltage drop as a function of pad oxide and polysilicon thickness (nitride 250nm)
fig. 6b Field oxide (pad oxide 30nm, polysilicon 200nm, nitride 250nm)

4. CONCLUSION

In summary we have shown, that it is possible to find an optimised combination of pad oxide and polysilicon thickness, which adapts the bird's beak length to the requirements of 0.8µm-technology without increasing the gate oxide thinning and without essential aggravations of topography. A correlation between gate oxide thickness at the field oxide edge and voltage drop U_{inj} during constant current stress has been found, which allows to characterise an isolation process with regard to its gate oxide thinning. Time dependent dielectric breakdown and its impact on device reliability is correlated with the gate oxide thinning at the field oxide edge.

REFERENCES

/1/ J.Matsunaga et al., IEDM Techn. Dig. 565 (1980)
/2/ E.P. Box et al., Statistics for Experimenters, Wiley, New York (1978)
/3/ Statistical Graphics System STATGRAPHICS, STSC - PLUS*WARE
/4/ T.Kusaka et al., IEEE El. Dev. Lett. 8, 61 (1987)
/5/ L.O. Wilson, J. Electrochem. Soc. 129, 833 (1982)

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

AN ANALYSIS OF THE ADVANTAGES OF A BOX ISOLATION TECHNIQUE

E. FIGUERAS*, S. HAZEBROUCK and F. VAN DE WIELE

Université Catholique de Louvain, Laboratoire de Microélectronique, Place du Levant, 3, B-1348 Louvain-la-Neuve, Belgium
**Centro Nacional de Microelectrónica, CSIC-UAB, Universitat Autònoma de Barcelona, SP-08193 Bellaterra, Barcelona, Spain*

Résumé - Dans ce papier nous présentons un procédé BOX et les resultats électriques obtenus avec cette technique. Il utilise un contre-masque de polysilicium plutôt que d'utiliser un procédé avec une double couche de résine. Les caracteristiques électriques des transistors parasites, le courrant de fuite des diodes avec périmètres variables et la variation de la largeur de canal sont presentés.

Abstract - In this paper we present a BOX process and the electrical results obtained with this technique, in which a polysilicon counter-mask is used instead of a double resist layer to perform the field oxide. The electrical characteristics of active and parasitic transistors, the leakage current measured on diodes with various perimeters and the channel width measurements are presented.

1 - INTRODUCTION

In recent years, new isolations technologies have been developed to overcome the limitations presented by the conventional LOCOS in VLSI circuits. Most of them are based on the use of a thermal field oxide. However BOX technologies(1) use a deposited oxide and therefore allow to get rid of the problems related to bird's beak formation, channel-stop implantation encroachment and stress generation. In order to eliminate the oxide deposited on the active areas, an additional photolithographic step seems suitable /1,2/. In this work we continue to develop a polysilicon counter-mask process/2/. Even if the BOX process is more complicated than a LOCOS or SILO one, it has some advantages that are discussed in this paper.

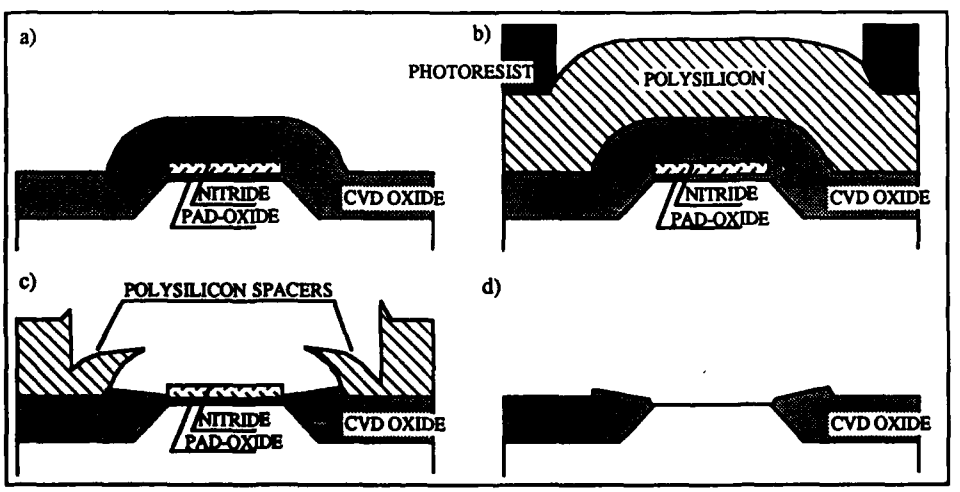


Fig. 1 - Fabrication steps of the BOX process using a polysilicon counter-mask .

(1) The autors use the therm BOX to nominate all processes which use a deposited oxide as field oxide.

2 - PROCESS DESCRIPTION

The starting material is <100> orientated silicon, 18Ω.cm p-wafers. First, a pad-oxide is grown and a nitride layer is deposited. After a photolithographic step to delineate the active regions, both layers are etched. Then grooves are etched in the silicon substrate with a KOH solution up to a depth of 0.5 µm, using the nitride film as a mask. After the channel-stop implantation, a SiO2 layer, with an equal thickness (or little more) to that of the Si groove depth, is deposited by CVD (Fig.1.a). A thick polysilicon layer is then deposited and whith an additional photolithographic step, a photoresist mask is left in the field areas (Fig.1.b). Then the exposed polysilicon is etched by RIE up to the oxide surface. The anisotropic etching allows the formation of spacers . Fig.1.c shows the structure after the wet etching of the oxide. During the last step, the nitride, the polysilicon and the pad-oxide layers are removed to obtain the desirable structure, as shown in Fig.1.d.
Fig.2 shows a SEM cross-section of an intermediate step of the fabrication process. In this picture we can see that even if a certain misalignement (less then 0.7 µm) of the resist mask exists, the polysilicon counter-mask is well formed. Fig 3 shows a SEM cross-section of a BOX structure with W= 1.3µm. No bird's beak is created and therefore no reduction of the dimensions are observed.

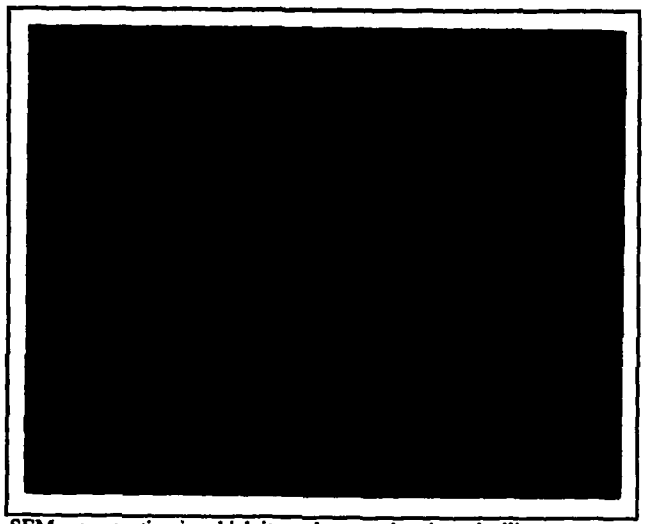


Fig. 2 - SEM cross-section in which it can be seen that the polysilicon counter-mask is auto-aligned in a range of 0.7µm

Fig. 3 - Cross-section of a 1.3µm device. Left : BOX, right : LOCOS

3 - RESULTS AND DISCUSSION

Due to the absence of thermal oxidation, the devices fabricated with a BOX process have noreduction of the active area dimensions. That is especially evident when we compare a BOX structure with a LOCOS one, as in Fig. 3. Both structures were made with the same mask. In this case the bird's beak formation in the LOCOS structure has rendered useless the active area. If we compare the BOX process with a SILO one /3/ the dimension loss is comparable, but, due to the high temperature step necessary to grow the SILO field oxide, the boron encroachment is more important with the SILO process, and so the electrical channel width is reduced. Experimentally, the effective channel width is determined by plotting the drain current of the MOS transistors (at Vd = 50mV) v.s. design mask width. The intercept with the abscissa yields a correction ΔW shown in Table I (The SILO values are obtained from ref.4)

	Φ at/cm2	Vt (V) field ox.	ΔW(1) (μm)	ΔW(4) (μm)	Δ(ΔW) (μm)
BOX	2E12	12.4	0.33	0.25	0.08
SILO	2E13	12.5	1.1	0.6	0.5

Table I. - ΔW(1) is the channel width variation for Vg-Vt = 1V
ΔW(4) is the channel width variation for Vg-Vt = 4V
Δ(ΔW) = ΔW(1) - ΔW(4)

These values demonstrate that the BOX process provides a smaller loss of channel width due to a minor boron encroachment.
The subthreshold characteristics shown in Fig.4 correspond to polysilicon gated parasitic field transistors with a width of 100μm and a length of 2 and 10μm . As it can be seen, both 2 and 10μm length field transistors have the same subtreshold characteristic. One explanation of this behaviour can be that, as demonstrated by S.H. Goodwin et al./5/ , the smaller is the radius of the transition region , the better is the subthreshold characteristiccontrol by this region. Therefore it is independent of the length of the field region.
The leakage currents have been mesured on N+P diodes with the same area (A = 1.6E-3cm2) and three different perimeters (P1 = 0.16cm, P2 = 0.96cm, P3 = 4.8cm). These currents are plotted in Fig.5. As expected, the leakage current is very low.

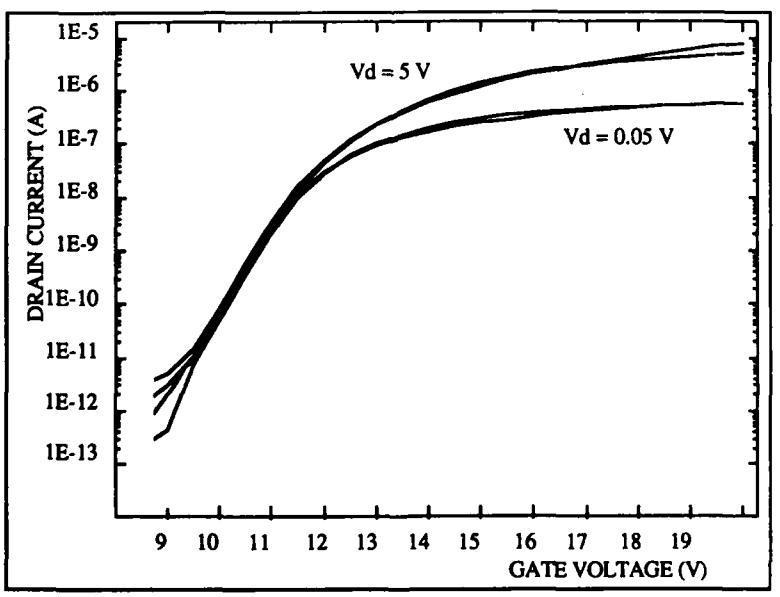


Fig.4 - Subthreshold characteristics of parasitic field transistors with L = 2μm and L= 10μm

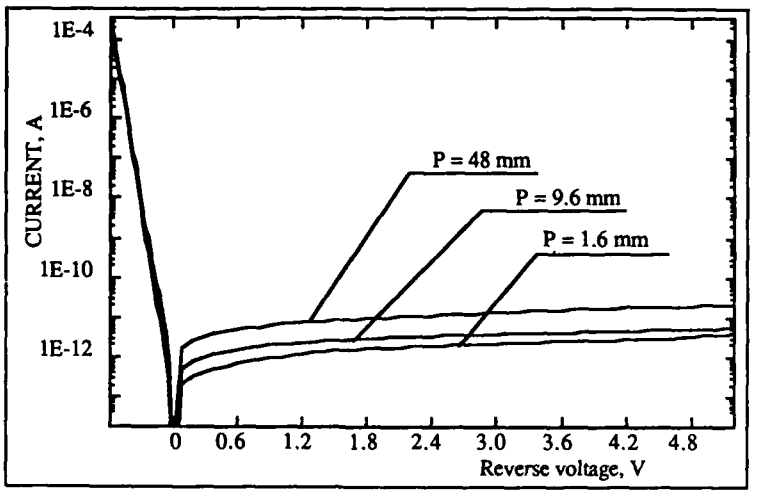


Fig. 5 - Leakage current of N^+P diodes with the same area and diferent perimeters.

CONCLUSIONS

In this work we have presented some electrical results obtained from devices fabricated with a BOX process based on the use of a polysilicon counter-mask. Even if the BOX process is more complicated than a LOCOS or SILO one, it presents some advantages : bird's beak free structures, boron encroachment reduction, very good isolations properties and a low leakage current; these adventages are related to the elimination of the long high temperature step necessary to grow a thermal field oxide.

REFERENCES

/1/ Shibata, T., Nakayama, R., Kurosawa, Onga, S., Konaka, M., Iizuka,H., IEEE IEDM Tech. Dig., paper 2.3, 1983.
/2/ Figueras, E., Coppée, J.L., Van de Wiele, F., Proceedings of the 17th ESSDERC, 1987, 473
/3/ Coppée, J.L., Figueras, E., Goffin, B., Gloesner, D., Van de Wiele, F., presented at the 18th ESSDERC,1988.
/4/ Coppée, J.L., Van de Wiele, F., S.S.E. Vol 31, N°5, p.887, 1988
/5/ Goodwin, S.H., Plummer, J.D., IEEE ED-31, N°7, p.861, 1984

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

LIMITATIONS ON n^+/p^+ SPACING DUE TO SHADOWING EFFECTS IN A 0.7μm RETROGRADE WELL CMOS PROCESS

M.G. PITT and P.A. van der PLAS

Philips Research Laboratories, PO Box 80000, NL-5600 JA Eindhoven, The Netherlands

Résumé

L'utilisation d'un implanteur d'ions à haute énergie afin de former des puits rétrogrades requiert une couche de résine photo-sensible très épaisse afin de prévenir la pénetration de l'implantation. Un ombrage de la région frontière du puit survient au cours de l'implantation, dû au déplacement de cette région en fonction de la position sur la gaufre et de l'angle d'implantation. Ce phénomène requiert donc une relaxation des règles de design. Cet effet fut étudié à l'aide de mesures électriques de transistors à effet de champs n^+/puit n^- en fonction de leur dimension et de leur orientation, tout en ayant un angle d'implantation nominal de 7°. Ces effets d'ombrage varient entre 0,15 et 0,25μm sur la surface d'une gaufre de 100mm de diamètre.

Abstract

Use of high energy ion implantation for retrograde wells requires thick resist layers to prevent implant penetration. Shadowing of the n-well implant results in a displacement in the position of the n-well edge, dependent on the position across the wafer and implant angle, thereby requiring a larger minimum design rule. Electrical measurement of n^+/n-well field transistors as a function of spacing and orientation has been used to investigate the amount of shadowing which occurs for nominal 7° implants. Shadowing effects were found to vary from 0.15 to 0.25μm across a typical 4 inch diameter wafer.

1 Introduction

In advanced CMOS technologies a small n^+ to p^+ spacing is required in order to achieve high packing densities. The use of suppressed LOCOS for oxide isolation (1,2) provides a technology for narrow field isolation regions whilst retrograde wells (3) are successful in providing high parasitic threshold and punchthrough voltages. An additional advantage of retrograde well technology is that the low thermal budget allows the use of thin epitaxial layers to suppress latch-up. The design rule minimum n^+ to p^+ spacing of 2.5μm is comprised of 1.4μm n^+ to n-well and 1.1μm p^+ to p-well spacings. The use of high energy implants to produce retrograded wells requires the use of a thick (here 2.6μm) resist layer. Shadowing by this thick resist layer results in a displacement in the position of the edge of the n-well, relative to the field isolation, as is shown schematically in figure 1. At the well concentrations used there is a considerable reduction in field transistor threshold voltage as the spacing is reduced. This provides a means by which the relative alignment of the edge of the n-well can be related to the threshold voltages observed and thus the effect of shadowing can be studied.

2 Process summary

The process described is based on that for the production of 1Mbit SRAMs (4). The relevant features of this process are:

1. Field oxide formation using a suppressed LOCOS process (1). The field oxide is planarised using an etch back procedure to a thickness of ~500nm.
2. Blanket p-well implantation consisting of 4×10^{12}B/cm^2 at an energy of 200keV.

3. Masked n-well implantation consisting of $9 \times 10^{12}P/cm^2$ at an energy of 500keV with a 2.6μm thick resist layer.

The remainder of the process consists of gate oxidation, polysilicon deposition and definition, source-drain implantations, silcidation and an advanced double level metallisation scheme (4).

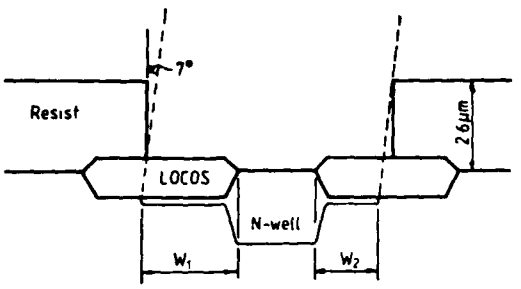


Figure 1(a): Schematic of n-well showing shadowing from high energy retrograde implant.
$W_1 \approx W + \Delta R_p + t_{LOCOS} \cdot \tan 7°$
$W_2 \approx W + \Delta R_p - (t_{resist} + t_{LOCOS}) \cdot \tan 7°$
where W is the nominal N+/N-well spacing.

Figure 1(b): Schematic showing cross-section after polysilicon definition and source-drain implantation.

3 Results

Figures 2 and 3 show respectively typical characteristics for n^+/n-well and p^+/p-well field transistors. Because of the relatively high threshold voltages as compared to gate oxide breakdown voltages ($\sim$ 20V) all measurements were restricted to the subthreshold regime. Threshold voltages were defined as the gate voltage required to give a drain current of 1nA at a drain voltage of $\pm$1V. The field transistors had a width of 100μm. Figure 4 shows the threshold voltage as a function of n^+/n-well or p^+/p-well spacing. Due to short channel effects the threshold voltage is reduced for smaller spacings especially for n^+/n-well devices.

The variation in threshold voltage with spacing provides a method by which shadowing effects across a wafer can be quantitatively evaluated. A test structure consisting of four n^+/n-well field devices (nominal spacing= 1.4μm) orientated at angles of 0, 90, 180 and 270° to the wafer flat was used to investigate the effect of orientation on threshold voltage. In order to remove the effect of implant dose and field oxide thickness variations the n^+/n-well field transistor threshold voltages were correlated with those of an n-well/n-well polysilicon gate field device as is shown in figure 5. This variation was then used to modify the threshold voltages of the n^+/n-well devices to that corresponding to the mean threshold voltage of the n-well/n-well devices. Calculation of the n^+ to n-well spacing was then made by fitting the modified threshold voltages for each orientation to the curve of V_t versus n^+ to n-well spacing at that wafer location. The results thus obtained are plotted in figure 6 in the form of vectors at each die position representing the displacement of the n-well from its nominal position relative to the field oxide edge. The degree of shadowing was estimated here as varying systematically from 0.15 to 0.25μm across the wafer As this uncertainty must be added separately to both n^+/n-well and p^+/p-well field transistor spacings the minimum n^+/p^+ spacing is increased by around 0.5μm. Increasing the well dose is effective in reducing the

minimum achievable n^+/n-well spacing (2) but at the cost of increased junction capacitances. The well doses may have to be as much as 50% higher in order to provide adequate threshold voltages under worst case shadowing conditions. Implanting the wafers with zero tilt would possibly be effective in considerably reducing the degree of shadowing which occurs. For a vertical resist profile shadowing of $\sim t_{resist} \cdot \tan 7° = 0.32\mu m$ is expected (i.e. assuming perfect masking by the resist). The actual resist edge profile was about 80° and this is probably why the observed shadowing has a lower mean value of 0.20μm. Angular variation of offset angle across the wafer of ±2.3° is partially responsible for the ±0.05μm variation in shadowing observed. Variations in layer-to-layer alignment are also visible where two rows of die positions are offset in the Y-direction in an opposite direction to the remaining die, possibly as a result of wafer warpage. Also visible is the amount of twist given to the wafers away from the flat at the implantation stage; in this case approximately 23°.

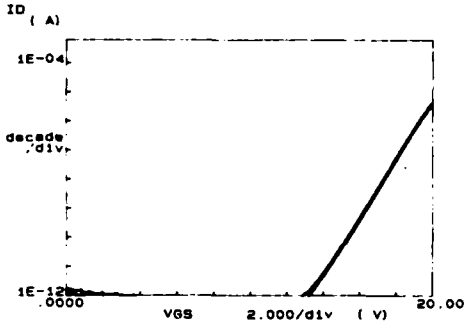


Figure 2: Characteristics of typical n^+/n-well polysilicon gate field transistor (V_{ds} =1,3,5V, nominal spacing=1.4μm).

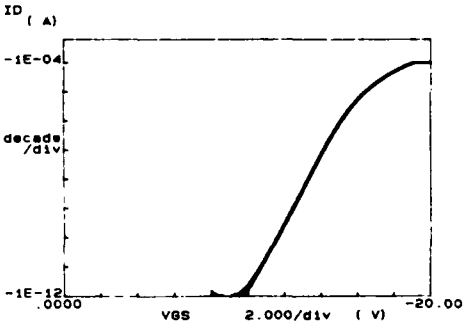


Figure 3: Characteristics of typical p^+/p-well polysilicon gate field transistor (V_{ds} =-1,-3,-5V, nominal spacing=1.1μm).

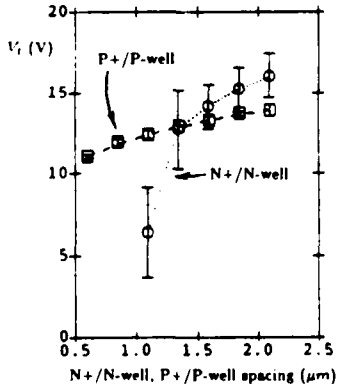


Figure 4: Threshold voltages (measured at V_{ds} = 1V, I_d = 10pA/1μm width) of n^+/n-well and p^+/p-well polysilicon gate field transistors versus spacing.

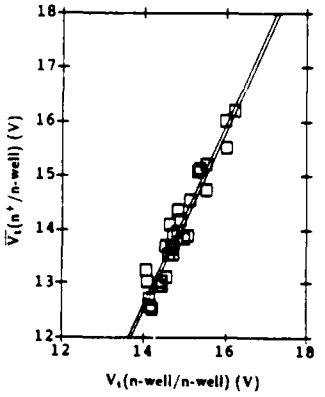


Figure 5: Mean threshold voltage of four n^+/n-well field transistors (nominal separation=1.4μm) orientated at 0, 90, 180 and 270° to the wafer flat versus threshold voltage of n-well/n-well polysilicon gate field transistors (separation=2.5μm).

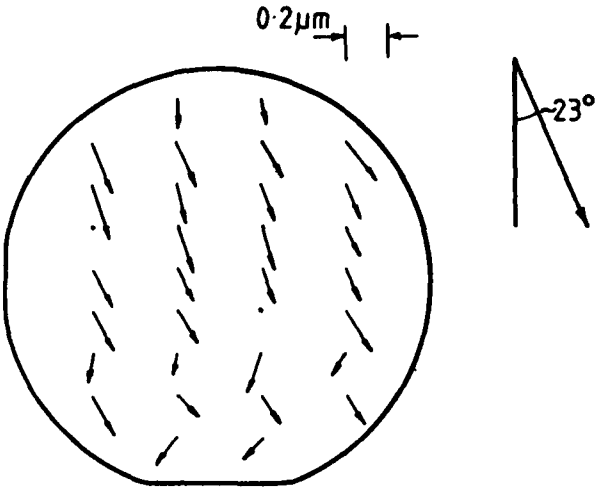


Figure 6: Wafer map showing calculated vector displacements of n-well edge. The twist of around 23° given to the wafer is indicated.

4 Summary

Electrical evaluation of field transistors has been used to estimate the amount of shadowing occuring during n-well implantation. Under certain conditions the n^+ to n-well field transistor threshold voltage is a strong function of n^+ to n-well spacing and by measuring n^+ to n-well polysilicon gate field transistors under different orientations the displacement of the n-well mask relative to the field oxide definition mask was estimated. Increasing the well dose is effective in reducing the minimum achievable n^+/n-well spacing (2) but at the cost of increased junction capacitance. In order to study shadowing effects, however, a somewhat low well dose is advantageous as it gives rise to larger variations in threshold voltage with spacing. The technique used provides a very sensitive method of establishing layer-to-layer misregistrations and could be extended to study lithographic misalignments by, for instance, construction of transistors with a separately implanted source drain regions with a thin gate oxide and a heavily overlapped gate defined after source-drain implantation.

References

1. P. A. van der Plas, W. C. E. Snels, A. Stolmeijer, H. J. den Blanken and R. de Werdt, 1987 VLSI Technology Symposium, p19

2. P. A. van der Plas, P. H. J. Spijkers and F. M. Klaassen, to be presented at ESSDERC 1988.

3. R. D. Rung, C. J. Dell'oca and L. G. Walker, IEEE Trans. Electron Devices, ED-28(1981)p1115

4. R. de Werdt et. al., 1987 IEDM Technical Digest, p532

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

SIMULATION OF SOURCE/DRAIN STRUCTURES FOR SUBMICRON MOSFETs WITH AND WITHOUT PREAMORPHIZATION

M. ORLOWSKI, C. MAZURÉ and L. MADER

Siemens AG, Corporate Research and Development, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

Résumé - Un modèle autoconsistant est proposé pour la diffusion des impuretés et des defauts ponctuels décrivant avec succès la diffusion a haute concentration de phosphore et bore y compris le cas de préamorphisation.
En particulier, les mécanismes générants des interstitiels et l'absorption d'interstitiels localisée dans la couche des boucles de dislocation - un résultat de la préamorphisation - sont pris en consideration d'une façon consistante.

Abstract - A selfconsistent model for the impurity and point defect diffusion is proposed and applied successfully to high concentration phosphorus and boron diffusion with and without preamorphized substrate. In particular the generation of the interstitials by high phosphorus and boron diffusion, the absorption of the interstitials at the damaged layer consisting of dislocation loops - a remnant of the preamorphization -, and the generation of the interstitials by the decay of the precipitated phase of the impurities above the solubility limit is taken into account in a consistent way. The present model is an important tool for advanced optimization for submicron MOSFETs dealing with involved interstitial dynamics as in the presence of preamorphization effects.

1 - INTRODUCTION

Recently substrate amorphization prior to source/drain implantation has been used to fabricate shallow junctions for submicron MOSFETs /1/. Already for a 900 ^{o}C 40 min inert anneal a reduction of source/drain junctions by 0.15 or 0.20 μm can be achieved for n^+ and p^+ regions, respectively. Due to the shallow junctions the minimum channel length with long channel behaviour is reduced by about 30 % for both channel types. Also the subthreshold swing can be improved significantly at small channel lengths when preamorphization has been used. For both types of transistor no deterioration of the junction quality has been observed.
The effects of the preamorphization are highlighting the importance of the point defect dynamics for the impurity diffusion in submicron devices. Therefore for successful simulation supporting MOSFET optimization it is of paramount importance to have a consistent model for impurity and point defect diffusion. The present model describes successfully the involved interaction between impurity and point defect dynamics.

2 - SIMULATION OF THE PREAMORPHIZATION

The amorphization of the top layer (thickness 0.2 μm) of the silicon crystal is performed by a blanket high dose Si implant prior to source/drain implantation. In Fig. 1 the damage profiles $C_{si}(x)$, the number of vacancies per unit volume in the silicon crystal, as a function of the depth, simulated with the Monte-Carlo simulator TRIM /2/ right after the Si-implantation is shown. By comparison with the experiment it is found that the critical damage density required to amorphize the Si-crystal is $C_{Si}^{crit} = 1.55 \cdot 10E^{22} cm^{-3}$. During the subsequent drive-in the amorphized layer recrystalizes very quickly, but there remains residual damage layer in a depth x_{si} corresponding to C_{Si}^{crit} such that $C_{si}(x_{si}) = C_{Si}^{crit}$. The gradient $\partial C_{si}(x)/\partial x|_{x=x_{si}}$ determines the width w_{si} of transiton layer from amorphous to cristalline phase which transforms during the recrystallization into the residual damage layer of corresponding width. It can be seen from Fig. 1 that for higher implantation energy the transition layer becomes broader in agreement with experimental date from TEM-pictures prior and after the recrystallization. The residual layer acts as an effective sink for the interstitials during the subsequent drive-in. The efficiency to absorb the interstitials has been determined from our diffusion simulations described below.

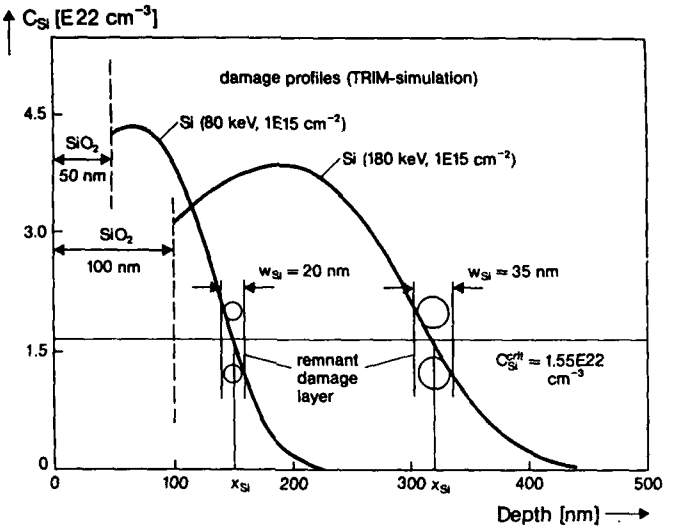


Fig. 1: Calculated damage profile for two Si-implantation energies

3 - DIFFUSION MODEL FOR PHOSPHORUS AND BORON WITH AND WITHOUT PREAMORPHIZATION

In order to simulate phosphorus and boron high concentration diffusion with and without preamorphization we have extended the interstitial-impurity pair diffusion model by Mulvaney and Richardson /3/ adding source and sink terms for interstitials and discriminating between active and clustered impurity concentration. The corresponding equations are given in the Table.
These equations are now part of a more complete theory for impurity diffusion in silicon /4/. Here j denotes phosphorus or boron, D_j is the corresponding intrinsic diffusivity of the impurity j, z_j is charge state, f_j the fractional interstitial diffusion component ($f_{phosphorus} = 0.9$, $f_{boron} = 0.75$), C_I denotes the interstitial concentration, C_I^{eq} the intrinsic interstitial concentration, D_I the diffusivity of the interstitials, ψ the electric field, and c_i the carrier concentration, being electron density for phosphorus and hole density for boron. The difference between chemical and electrically active impurity concentration is described by a simple cluster model with the clustering and declustering coefficients, k_{cl} and k_D, respectively. In this model it is assumed that both phosphorus and boron above the solubility limit generate interstitials with the efficiency β_j.

Table: Model Equations

$$\frac{\partial C_I}{\partial t} = \frac{\partial}{\partial x}\left[D_I f_I \frac{C_I}{C_I^{eq}} \frac{\partial C_I}{\partial x} + D_I f_I \frac{C_I}{C_I^{eq}} \frac{\partial C_I}{\partial x} + Z_I C_I D_I \frac{q}{kT} f_I \frac{C_I}{C_I^{eq}} \frac{\partial \psi}{\partial x}\right] - k_{cl}^{I} C_I^{m_I} c^k + k_D^{I} C_I^{clust} \quad (1)$$

$$\frac{\partial C_I^{clust}}{\partial t} = \frac{1}{m_I}\left[k_{cl}^{I} C_I^{m_I} c^k - k_D^{I} C_I^{clust}\right] \quad (2)$$

$$\frac{\partial C_I}{\partial t} = \frac{\partial}{\partial x}\left[\left(D_I + f_I D_I \frac{C_I}{C_I^{eq}}\right) \frac{\partial C_I}{\partial x} + f_I D_I \frac{C_I}{C_I^{eq}} \frac{\partial C_I}{\partial x} + Z_I f_I C_I D_I \frac{C_I}{C_I^{eq}} \frac{q}{kT} \frac{\partial \psi}{\partial x}\right] + \alpha D_I C_I g(x) - \beta_I \frac{\partial C_I^{clust}}{\partial t} \quad (3)$$

j ≙ phosphorus or boron I ≙ interstitials

This term is necessary in order to describe properly high concentration phosphorus and boron diffusion with the boundary condition for interstitials $C_I(x=0) = C_I^{eq}$, when the maximum of the impurity is not at the Silicon-surface.
The recombination of the interstitials at the damage layer is described by the last but one term in eq. (3) with $g(x) = \exp(-(x-x_{si})/w_{si})^2)$, where x_{si} and w_{si} have been taken directly from TEM-pictures and served for the calibration of the gradient $\partial C_{si}/\partial x|x = x_{si}$, see Fig. 1. The eqs. (1-3) have been evaluated using the 1-D process simulator ZOMBIE /5/, which proved to be an indispensable tool for this investigation.

4 - RESULTS

In Fig. 2 and Fig. 3 the profiles for the impurities, phosphorus and boron, and interstitials are shown with and without the preamorphization effect. As seen from Figs. 2 and 3 the agreement with the experiment is excellent. We consider first the impurity diffusion without the preamorphization. Fitting the experimental phosphorus and boron SIMS-profiles after a drive-in at 900 °C for 40 min the interstitial diffusivity, D_I, and the equilibrium interstitial concentration given by Tan and Gösele /6/ turned out to be the most reasonable. The impurity profiles are very sensitive to these values and to the boundary condition for the interstitials at the surface. The increase of D_I leads to the formation of the characteristic diffusion tail at higher impurity concentrations. The increase of C_I^{eq} entails a shorter and steeper tail. In case of phosphorus we had to enhance C_I^{eq} (Tan, Gösele) by 3.4, in case of boron by 1.2. D_I (Tan, Gösele) did not need any readjustment. However, the exact magnitude of these enhancement factors is of limited relevance. They indicate only the correct range of magnitude. More precise analysis requires more complete theoretical framework /4, 7/.
Finally the phosphorus and boron profiles with preamorphization have been obtained by simply activating the sink term $\propto D_I C_I g(x)$ using the same sink efficiency $\propto$, same parameters x_{si} and w_{si} without any further fitting. It could be surmised, that the sink efficiency $\propto$ can be traded against the width of the damaged layer w_{si}, since both parameters together determine the overall sink efficiency. However, it turned out that the shape of the impurity profile with preamorphization is more sensitive to $\propto$ and w_{si} than the interstital concentration in the bulk which depends on an overall sink efficiency only. The experimental value w_{si} with adjusted has provided the best fit to the experiment.
From Fig. 2 and 3 it can be seen that in case of phosphorus the residual damage layer reduces the interstitial supersaturation produced by high phosphorus concentration diffusion, while in case of boron the damaged layer supresses the weak interstitial supersaturation to undersaturation values.

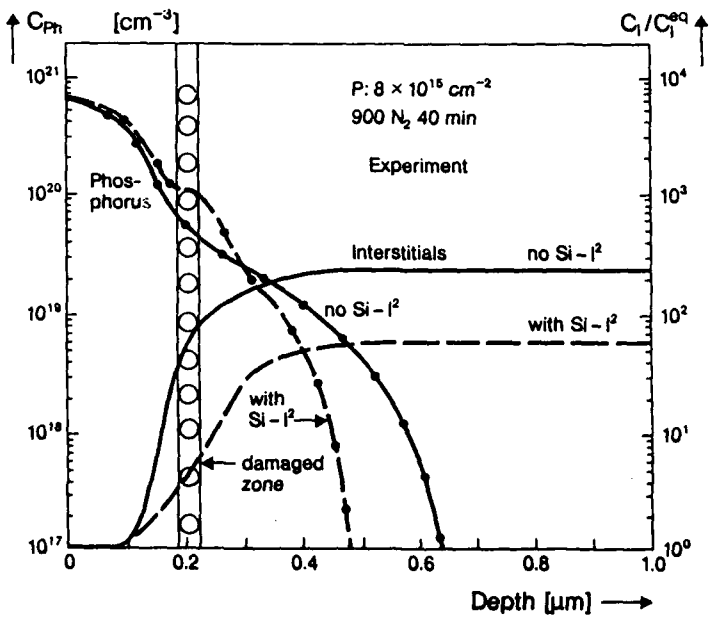


Fig. 2: Similuted phosphorus and interstitial concentration distribution with and without preamorphization

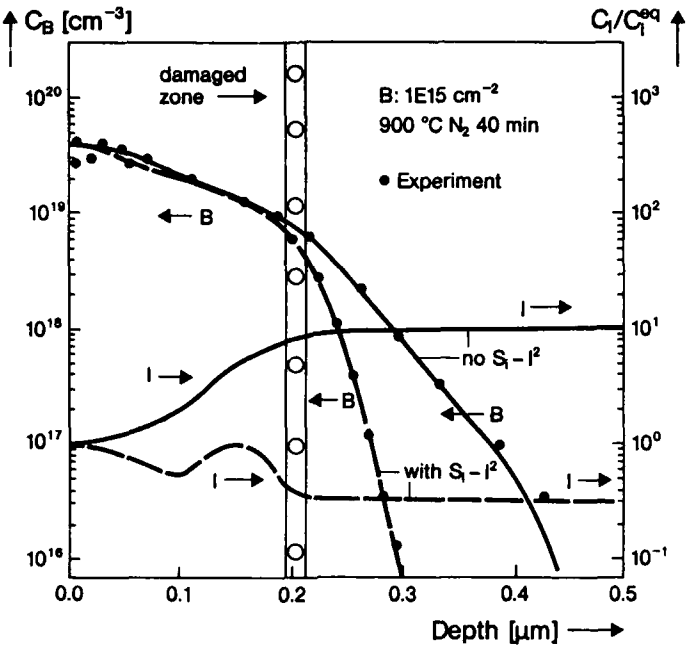


Fig. 3: Simulated boron and interstitial concentration distribution with and without preamorphization

5 - CONCLUSIONS

A model being part of a more general theory /4/ has been presented and compared successfully with rather involved experimental structure of high concentration phosphorus and boron diffusion in presence of preamorphization effects. The key feature of the presented equations is the consistent treatment of the point dynamics and its coupling to the impurity diffusion. Such models are badly needed for successful process simulation for submicron devices not only for source/drain regions but also for the simultaneous and consistent treatment of the channel region. In Ref. /8/ it has been shown how drastically interstitial generation during source/drain formation can influence the channel doping distribution and change electric properties of the device. The model given here and its generalization in Ref. /4/ can deal with such circumstances.

REFERENCES

/1/ C. Mazuré, J. Winnerl and F. Neppl, ESSDERC, Proceedings 1987, Bologna, p. 585
/2/ J. F. Ziegler, J. B. Biersack and U. Littmark, "The Stopping and Range of Ions in Solids", Pergamon Press 1985
/3/ B. J. Mulvaney and W. B. Richardson, Appl. Lett. 51 (1439) 1987
/4/ M. Orlowski, "Unified Model for Impurity Diffusion in Silicon", submitted for publication to Appl. Phys. Lett.
/5/ W. Jüngling et al., IEEE ED-32, (156), 1985
/6/ T. Y. Tan and U. Gösele, Appl. Phys. A 37 (1), 1985
/7/ M. Orlowski, in preparation
/8/ M. Orlowski, C. Mazuré and F. Lau, IEDM-Proceedings (632) 1987

Session 5C Room C : Device modeling

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

AVALANCHE AND TUNNELING BREAKDOWN MECHANISMS IN HEMT'S POWER STRUCTURES

Y. CROSNIER, F. TEMCAMANI, D. LIPPENS and G. SALMER

Centre Hyperfréquences et Semiconducteurs, CNRS-UA 287, Université des Sciences et Techniques de Lille Flandres Artois, F-59655 Villeneuve-d'Ascq Cedex, France

Résumé - Cet article expose une étude sur le phénomène de claquage dans les HEMTs de puissance AlGaAs/GaAs. Il fait usage, d'une part, de résultats expérimentaux obtenus sur différents composants test et, d'autre part, de modélisations prenant en compte soit l'ionization soit l'effet tunnel. Cette approche apporte une nouvelle compréhension sur les rôles respectifs de ces deux effets dans le déclenchement du claquage et permet de définir des règles de construction.

Abstract - The present paper reports a study on the breakdown phenomenon in AlGaAs/GaAs power HEMT'S. It uses, on one hand, experimental results carried out on various test devices and, on the other hand, modeling taking into account either ionization or tunneling effects. Such an approach gives a new understanding on the respective roles of these two effects in the breakdown occurrence and allows to define design rules.

INTRODUCTION

AlGaAs/GaAs High Electron Mobility Transistors (HEMT's) are now able to operate up to 70-80 GHz for small signal amplification. They have also considerable potential in the fields of power amplification provided a high drain current and a high breakdown voltage can be achieved. The first requirement can be satisfied by multiple channel structures [1, 2, 3, 4]. The second one presents much more difficulty. Indeed, high doping level in AlGaAs layers is necessary to get high electron sheet density in the quantum wells, but a great attention must be paid to the design of the structure in order to avoid excessive electric field values in the gate vicinity while keeping an efficient charge control and a good dynamic behaviour under E.H.F. operations.

I - TYPICAL EXPERIMENTAL HEMT BREAKDOWN VOLTAGES

Figure 1a shows typical breakdown voltages V_{BR} of conventionnal single heterojunction HEMT's. As it can be noted V_{BR} is a decreasing function of the AlGaAs doping level N_D and reaches very low values above 10^{18} cm^{-3}. Measurements carried out on multiquantum well HEMT's have revealed a similar dependance on the layer lying underneath the gate. The design of this layer constitutes a key problem for power HEMT's. Nevertheless, a great improvement is achieved using pulse-doped structures [5] where a thin undoped AlGaAs layer separates the gate from the normally doped AlGaAs layer. As shown in figure 1b, such an arrangement can give a breakdown voltage of about 15V with only a slight dependance on the undoped layer thickness.

II - MODELING OF BREAKDOWN BY IONIZATION

In answer to the question about what determines the above breakdown behaviour, the first idea coming in mind is that of an ionization mechanism. This assumption has been investigated with a numerical model that computes the potential shape and gives the electric field distribution in the transistor under pinchoff conditions. This model solves Poisson's equation and current equation together

$$\nabla^2 V(x,y) = -q/\varepsilon [N_D(x,y) - n(x,y)]$$

$$\vec{J}_n = q[-\mu_n . n(x,y)\, \vec{\text{grad}}\, V(x,y) + D_n \vec{\text{grad}}\, n(x,y)] = 0$$

and then computes on every electric field line the ionization integral $\int \alpha\, dl$, until this one is equal to unity. The ionization coefficient is expressed by $\alpha = A \exp -(B/E)^C$, E being the electric field amplitude

with $A = 1.8\ 10^7$ m^{-1}, $B = 5.75\ 10^7$ V/m, $C = 1.82$ for GaAs.

and $A = 4.1\ 10^7$ m^{-1}, $B = 10^8$ V/m, $C = 2$ for AlGaAs($x = 0.26$).

Neglecting, in a first approach, the gate fringing effect and, therefore, using a one dimensional description along the y axis (normal to the gate), one obtains the results shown in figure 2a, with breakdown voltages of about 35 volt, for a homogeneous AlGaAs layer, whatever the doping concentration and the thickness are. Similar results are given in the case of a pulse-doped AlGaAs

layer, as shown in table 2b. All these results are far from the experimental ones which indicates that this one dimensional approach does not agree with reality. In a second approach, we have then used a true twodimensional description. Figure 3 shows the results obtained for a homogenous AlGaAs layer. A satisfying agreement with experimental data appears for $N_D < 7\ 10^{17}$ cm^{-3} but, above this value, there is a discrepancy which is more and more important as N_D increases. Applying the twodimensional treatment to the case of pulse-doped AlGaAs layers, we have found results very similar to those of figure 3, which is somewhat surprising. A detailed analysis has revealed that, despite important variations in the electric field paths under the major part of the gate, the avalanche condition $\int \alpha dl = 1$ is always realized close to the gate edge where the electric field reaches very high values because of the geometrical singularity and exhibits evolutions that depend only slightly on the layer arrangement. This situation is illustrated with typical examples of field configurations in figure 4.

III - MODELING OF BREAKDOWN BY TUNNELING EFFECT

As seen above, the gate edge is the locus of very high electric field values, so the possibility of tunneling effect at the Schottky contact must be taken into account. Two cases have to be considered. The first one concerns homogeneous doped AlGaAs layers and the second pulse-doped ones.

For the former case the electric field can then be assumed to decrease steeply in the gate vicinity. Therefore the emitter current can be considered as a pure tunneling current (thermionic contribution is negligible) and expressed by the following relation derived from Padovani and Stratton works [6]

$$J_T = (A\,\Pi\,W_{oo}/kT)\,\frac{\exp\left|-\frac{2q\varnothing_B}{3W_{oo}}\left(\frac{\varepsilon E_M}{2N_D q\varnothing_B}\right)^{-\frac{1}{2}}\right|}{\left|\frac{\varepsilon E_M^2}{2N_D q\varnothing_B}\right|^{-\frac{1}{4}}\sin\left|\frac{\Pi kT}{W_{oo}}\left(\frac{\varepsilon E_M^2}{2N_D q\varnothing_B}\right)^{-\frac{1}{2}}\right|}$$

where A = Richardson constant in metal

$W_{oo} = q\hbar/2\,(N_D/\varepsilon m^*)^{1/2}$ $\qquad$ $\varnothing_B$ = Schottky barrier potential

E_M = electric field amplitude at the Schottky contact.

For the latter case the electric field can then be considered almost flat near by the gate. In this case, the tunneling effect results from both thermionic and field emissions and it can be expressed by the relation established by Chivé et al [7] for similar problems.

$$J_T = A\exp\left(-\frac{q\varnothing_B}{kT}\right)x\left|\frac{1}{kT}\int_0^{q\varnothing_B}\exp\left(\frac{u}{kT} - c_o u^{3/2}\right)du + 1\right|$$

with

$$c_o = \frac{4}{3\hbar q E_M}(2m^*)^{\frac{1}{2}}$$

Figures 5a and 5b show results of computations carried out with the above expressions assuming simply a one dimensional description. The comparison with experimental data gives a satisfying agreement and shows clearly that the tuneling effect is dominant for heavily doped layers underneath the gate. A more complete description would take into account gate fringing effect, space charge reaction and ionization alltogether. We are developping such a treatment. Preliminary results indicate that breakdown voltages are probably slightly lower than shown in Figures 5a and 5b.

IV - TEMPERATURE TEST

Heating is the commonly used test that permits to distinguish between breakdown by ionization and/or tunneling effect. Indeed, the variation of the breakdown voltage with temperature is positive for the first case and negative for the second one. We have experimented this method on a lot of HEMTs, all of them having the uppest AlGaAs layer homogeneous and highly doped. As expected, heating has proved that the involved breakdown mechanism is undoubtly the tunneling effect. This is shown in Figure 6 with, for comparison, the corresponding behaviour of a typical MESFET (NE 720) indicating that, in this device, ionization is dominant.

CONCLUSION

The above analysis shows clearly that tunneling effect instead of ionization is the dominant mechanism of the breakdown phenomenon in HEMT's when the uppest AlGaAs layer of these devices is heavily doped ($N_D > 7\ 10^{17}$ cm^{-3}). Therefore, the achievement of high breakdown voltages requires

that this layer presents a pulse-doped profile with a thin undoped (≃ 200 Å) zone underneath the gate. The designer must pay great attention to this requirement when optimizing power HEMT structures.

REFERENCES

1. N.H. Sheng et al, IEEE EDL, 307, June 1985.
2. K. Hikosaka et al, IEEE Trans ED, 583, May 1986.
3. P. Saunier, IEEE EDL, 503, Sept. 1986.
4. K. Hikosaka et al, IEEE EDL, 521, Nov. 1987.
5. Y. Crosnier et al, JNM III V, Aussois France, Jan. 1987.
6. F.A. Padovani et al, SSE, Vol. 9, 695, 1966.
7. M. Chivé, Thèse Doct. es Sc. Phys., Lille France, 1978.

This work was supported by DRET under Contract N° 85-176.

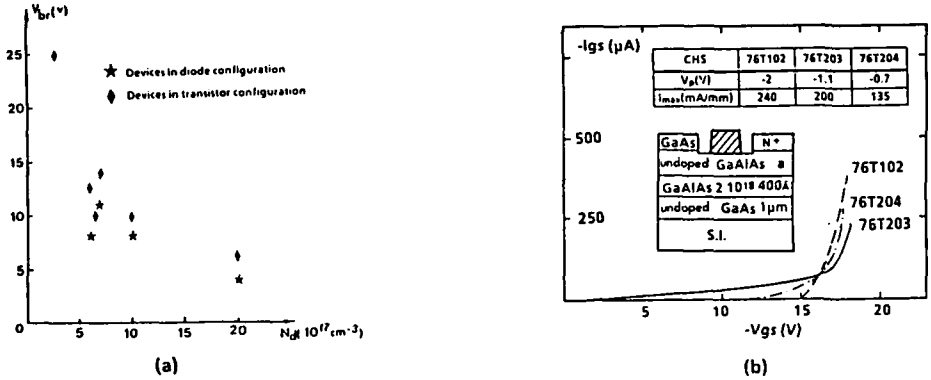


Fig. 1 : Experimental Breakdown voltage for conventional HEMTs with an homogeneous AlGaAs layer (a) and for pulse-doped HEMTs (b).

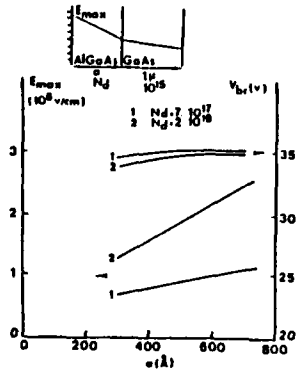
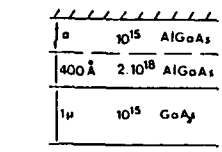


Fig. 2a : Calculated breakdown voltage and maximum electric field of a conventional HEMT as a function of the AlGaAs thickness for two. doping concentrations. One dimensional modeling that only considers ionization.



a (Å)	E_{max} (10^6 v/cm)	V_{br} (v)
100	1.57	33.6
200	1.55	34.3
300	1.54	34
400	1.51	32.4

Table. 2b : Calculated breakdown voltage and maximum electric field as a function of the thickness of the undoped AlGaAs layer for a MIS HEMT structure. One dimensional modeling that only considers ionization.

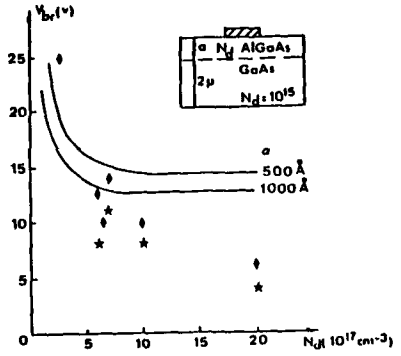


Fig. 3 : Calculated breakdown voltage of a conventional HEMT as a function of the AlGaAs layer doping concentration for two thickness. Two dimensional modeling that only considers ionization. Comparison with experimental measurements.

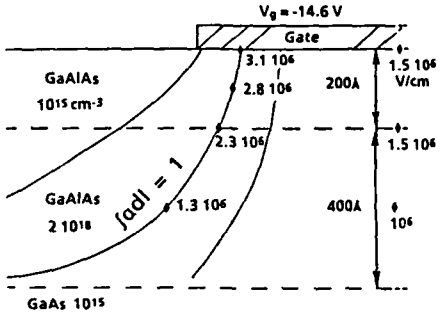


Fig. 4 : Typical example of electric field configuration, at breakdown, for a pulse-doped HEMT.

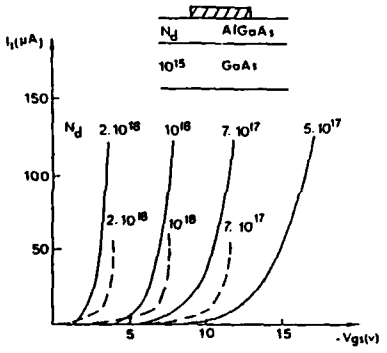


Fig. 5a : I(V) characteristic of the Schottky gate of a conventional HEMT, for various AlGaAs doping concentrations :

——— Calculated with a modeling that only considerers tunneling effect.

---------- Measured.

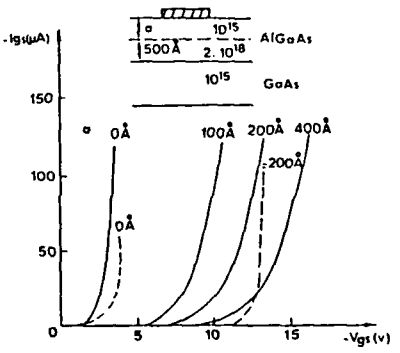


Fig. 5b : I(V) characteristic of the Schottky gate of a MIS HEMT structure, for various thickness of the undoped AlGaAs layer.

——— Calculated with a modeling that only considers tunneling effect.

---------- Measured.

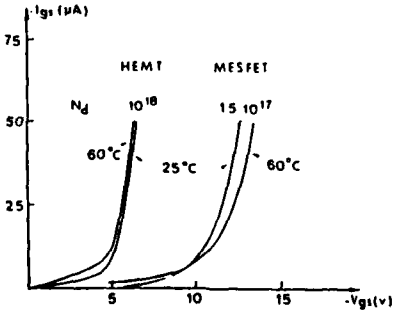


Fig. 6 : I(V) characteristic of the Schottky gate of a GaAs MESFET and a AlGaAs/GaAs HEMT for two temperatures.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988 C4-567

REAL-SPACE TRANSFER IN HETEROJUNCTION FET's : MONTE-CARLO SIMULATION AND ANALYTICAL MODEL

M. MOUIS, F. PAVIET-SALOMON, P. DOLLFUS and R. CASTAGNÉ

Institut d'Electronique Fondamentale, CNRS-UA 22, Université Paris-Sud, F-91405 Orsay Cedex, France

Résumé - Nous analysons les résultats d'une simulation Monte-Carlo du NERFET, dispositif à effet de champ où le transfert spatial des électrons du canal au dessus d'une hétérojonction permet d'obtenir un effet de résistance différentielle négative (RDN) sur le courant de drain. Sur la base de cette simulation, nous établissons une expression analytique du courant de transfert spatial. Les résultats obtenus sont en bon accord avec les résultats de la simulation Monte-Carlo sans qu'il soit nécessaire d'introduire de paramètre ajustable.

Abstract - We present the results of a Monte-Carlo simulation of the NERFET, a field effect transistor where the real-space transfer of channel electrons over a heterojunction barrier gives rise to a negative differential resistance (NDR) effect on drain current. An analytical expression of the real-space current is built upon our simulation results. By this way, we construct an analytical model of the NERFET which shows good agreement with Monte-Carlo results without the need of adjusting any parameter.

1-Introduction:

Field-effect devices using conduction along a heterojunction are of growing interest for high speed integrated circuits. They combine two advantages: the high electron density which can be obtained in an accumulation or inversion layer, and a high electron mobility which can be achieved both because of the high purity of the active layer and because of the high quality of the heterojunction interface. However, the confining potential barrier is lower than in classical MISFET's. in heterojunction FET's (HFET's), this barrier is equal to the conduction band discontinuity at the interface and is of the same order of magnitude as the kinetic energy of hot electrons in short devices. Thus, hot electrons can transfer from the channel into the barrier semiconductor. This real-space transfer (RST) can be used in order to obtain a negative differential resistance (NDR) effect on drain current as it is the case in the NERFET [1,2].

2-The model:

The active region of the NERFET is shown in figure 1-a, drawn on the basis of the data in reference 1. The assumed band discontinuity between GaAs and $Ga_xAl_{1-x}As$ (x=0.34) is 0.3eV. The mid-gap Fermi level pinning at the GaAs surface is taken into account using the convenient surface charge. The source to drain distance is L = 1µm.

As it has been pointed out by several authors [2,3,4], electrons in the channel of submicron devices are strongly out of equilibrium. Furthermore, real-space transfer is governed both by the energy and angular distributions of channel electrons and by the local conditions allowing transmission over the heterojunction conduction band discontinuity. A 2D Monte-Carlo simulation is then very well suited for studying real-space transfer, since it accurately describes non stationnary transport phenomena and two-dimensional effects. We used the many-particle Monte-Carlo model which was presented in a previous paper [5]. The motion of each electron is divided into a succession of random free-flights in the 2D self-consistent electric field and of interactions with lattice phonons, ionized impurities and alloy disorder. For each electron impinging on the interface, the transmission probability is calculated with a quantum mechanical model which ensures the continuity of the parallel component of the wave-vector, the continuity of the total (kinetic+potential) electron energy and the continuity of the probability current[6]. However, we do not account yet for a possible tunneling of electrons with an energy lower than the conduction band discontinuity.

In figure 1-b, we present the source, drain and substrate characteristics versus drain and gate voltages. As expected, a substrate current (due to RST) appears for a given V_{DS} value. For bias voltages where source current saturates, the increase in substrate current leads to a symmetrical decrease in drain current. This induces the NDR effect. The electron flux through the heterojunction is especially important at the drain end of the channel, where electrons are hot. However, the substrate current which arises from RST is not just related to heating phenomena: it is shown in figure 2 that the evolution of the electron energy along the channel for a given V_{DS} is not dependant on substrate voltage, whereas I_{sub} is (fig 1-b). Several mechanisms are involved and will now be detailed.

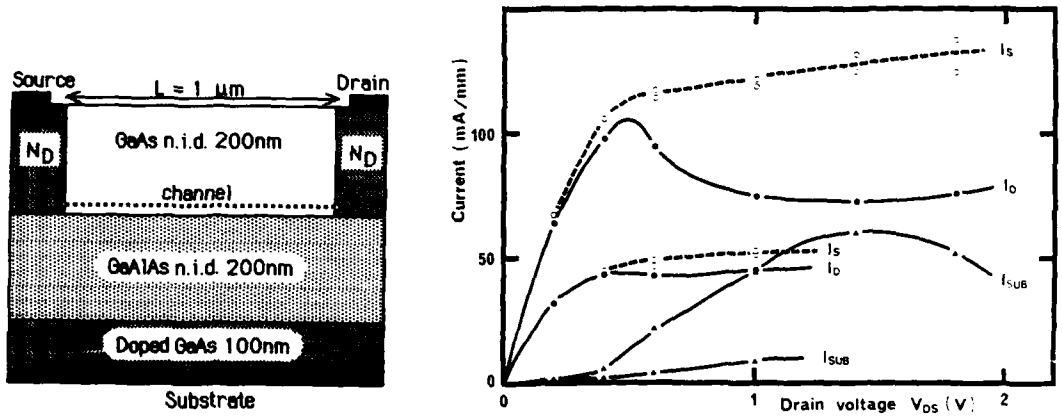


Fig.1- Device geometry and current-voltage characteristics deduced from Monte-Carlo simulation.

Let us note firstly that increasing V_{sub} results naturally in a higher confining field $E_\perp$ at the upper GaAs/GaAlAs interface. This implies modifications both in the current saturation mechanism and in the amount of real-space transfer. Figure 2 shows the evolution of the transverse electric field along the interface for two values of V_{sub} under bias conditions leading to source current saturation (V_{DS} = 1V). For the low substrate voltage (V_{sub} = 1V), $E_\perp$ becomes negative near the drain, causing channel pinch-off and moving hot electrons away from the interface. For a higher substrate voltage, (V_{sub} = 2V), $E_\perp$ has a confining influence all along the channel and current saturation is due to another mechanism. From the electron concentration, we find that a dipolar domain is trapped at the drain N^+ region boundary, clearly indicating a velocity saturation mechanism [7,8]. In the accumulated region of the domain, population inversion of the electrons from the Γ valley and the lateral valleys occurs. This allows the presence of hot electrons, pertaining to the L valleys, near the interface.

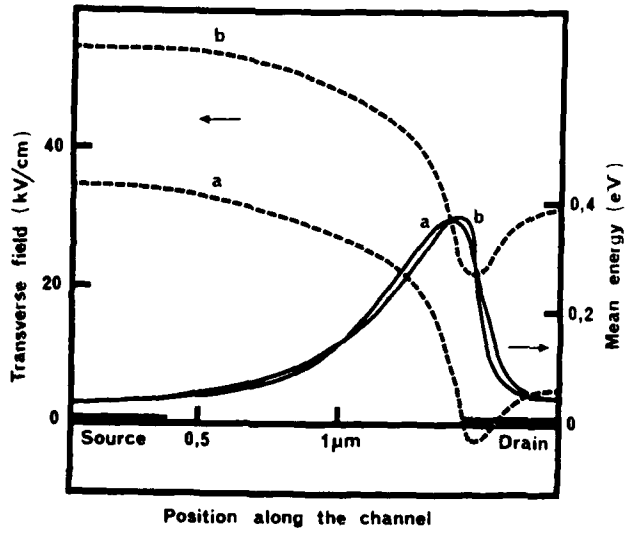


Fig.2-Evolution of the electron energy and of the transverse electric field along the channel; V_{sub}=a)1V, b)2V.

Furthermore, RST is easier for these L valleys electrons. This comes first from the local conditions which allow the transmission of an electron over the conduction band discontinuity. The transmission is impossible unless the kinetic energy associated with the transverse wave-vector component ($\hbar^2 k_\perp^2/2m^*$ neglecting non-parabolicity) exceeds a minimum value. In other words, an electron from GaAs impinging on the interface with a kinetic energy $\mathcal{E}$ will surmount the interfacial potential discontinuity only if the incidence angle is lower than a maximum angle $\theta_{lim}(\mathcal{E})$ defined by $\sin^2\theta_{lim}=(1-\Delta\mathcal{E}_c/\mathcal{E})\, m^*(GaAlAs)/m^*(GaAs)$. Due to the barrier heights differences, the maximum incidence in the L valleys is higher than in the Γ

valley. Moreover, in submicron devices, the velocity distribution of hot electrons in the Γ valley is known to be very anisotropic, with a main orientation parallel to the channel interface [4,9]. On the contrary, in the L valleys, the velocity distribution of electrons is randomized by the frequent interactions between equivallent satellite valleys. This again favours the real-space transfer of previously k-space transfered electrons.

To sum up, for the low or medium substrate voltages, the current saturation mechanism is involved in the amount of real-space transfer: we have shown that thermoionic injection is facilitated by electron velocity saturation and by k-space transfer, whereas channel pinch-off induces an antagonist transverse field.

3-Analytical model for the real-space current:

The basis of our analytical model are derived directly from the Monte-Carlo results. We saw that the amount of RST is strongly enhanced when k-space transfer occurs. Therefore, we will only investigate the case where the source current saturation is due to velocity saturation of electrons and is associated with the formation of a dipolar domain trapped near the drain. The substrate current is mostly carried by electrons from the accumulation region of the domain which obey the transmission conditions. Electrons in the domain all belong to the L valleys of the conduction band. Due to the frequent intervalley interactions, they are supposed to obey a Maxwellian distribution (with a temperature T_e).

The substrate current is calculated by adding the contributions of electrons, according to their energy and to the direction of their velocity. With Boltzmann statistics, the number of electrons in a range of $d\mathcal{E}$ in the L valleys is:

$$dn(\mathcal{E}) = \frac{n_0}{kT_e} \exp\left(-\frac{q\mathcal{E}}{kT_e}\right) d\mathcal{E} \quad \text{where } n_0 \text{ is the electron density in the accumulation domain.}$$

Within this population, the electrons which have their velocity directed within a differential solid angle around θ contribute to the substrate current in the ratio of $v_\perp = \sqrt{2q\mathcal{E}/m_L^*}\cos\theta$. The resulting differential current is:

$$di_{sub} = q Z L_d^e \left\{ \frac{n_0}{kT_e} \exp\left(-\frac{q\mathcal{E}}{kT_e}\right) \right\} \left\{ \sqrt{\frac{2q\mathcal{E}}{m_L^*}} \cos\theta \right\} \left\{ \frac{2\pi \sin\theta \, d\theta}{4\pi} \right\} d\mathcal{E}$$

Here, Z is the device length and L_d^e is the length of the accumulated region of the domain.

Continuity equations in the heterojunction plane (with equal effective masses for electrons in the L valleys of GaAs and GaAlAs) show that an electron can be injected in GaAlAs if its energy $\mathcal{E}$ is greater than the band discontinuity in the L valleys ($\Delta\mathcal{E}_{cL}$) and if the velocity is directed within a solid angle which is defined by $\sin^2\theta_{lim} = 1 - \Delta\mathcal{E}_{cL}/\mathcal{E}$. Integrating over θ and $\mathcal{E}$ leads to:

$$I_{sub} = \frac{qn_0 L_d^e Z}{4} \sqrt{\frac{2q\Delta\mathcal{E}_{cL}}{m_L^*}} \, f\left(\frac{1}{x}\right) \text{ with } \qquad x = \frac{kT_e}{q\Delta\mathcal{E}_{cL}}$$

$$f\left(\frac{1}{x}\right) = \sqrt{\pi}\sqrt{\frac{x}{2} - 1}\ \text{erfc}\left(\frac{1}{\sqrt{x}}\right) + \exp\left(-\frac{1}{x}\right)$$

Considering that any increase in V_{DS} above the saturation value V_{DSsat} has the effect of reinforcing the dipolar domain, we obtain the relative length of the domain (L_d^e/L) by solving Poisson equation. On the other hand, kT_e/q can be determined using the energy relaxation equation. The only unknown quantity is then $n_0 L_d^e$, which accounts for the domain geometry. This geometry can be specified for the NERFET: the domain is confined on an almost constant thickness between the heterojunction plane and the surface space-charge layer resulting from the Fermi-level pinning at the GaAs surface. The whole set of equations can then be expressed only in terms of x and of constant parameters:

$$Q_D = qn_0 L_d^e aZ = x \, \frac{3\varepsilon_0\varepsilon_r \, aZ}{v_{sat} \, \tau_\varepsilon} \Delta\mathcal{E}_{cL} \ (eV) \quad (v_{sat} \text{ is the saturation velocity and } \tau_\varepsilon \text{ the energy relaxation time})$$

$$\frac{L_d^e}{L} = \frac{2}{L} \frac{\varepsilon_0\varepsilon_r \, aZ}{Q_0} \frac{V_{DS} - V_{DSsat}}{Q_0} \frac{1}{x} \sim \frac{Q_0}{qN_D aZL} x \qquad \text{where } Q_0 = \frac{3\varepsilon_0\varepsilon_r \, aZ}{v_{sat} \, \tau_\varepsilon} \Delta\mathcal{E}_{cL}$$

$$\boxed{I_{sub} = \frac{Q_0}{4a} \sqrt{\frac{2q\Delta\mathcal{E}_{cL}}{m_L^*}} \, x \, f\left(\frac{1}{x}\right)}$$

On the other hand, the drain current is obtained as the ratio of the stored charge Q_D over the transit time in the accumulation domain (L_d^a/v_{sat}) whereas the source current is a saturated current controlled by V_{sub} with a transconductance g_m. Therefore:

$$I_D = \frac{Q_D\, v_{sat}}{L} \frac{x}{L_d^a/L} \quad \text{and} \quad i_S = g_m\,(V_{sub} - V_T)\,\frac{1}{1 - L_d^a/L}$$

Current conservation allows the determination of x. Figure 3 shows the characteristics obtained by using the electrical parameters deduced from the simulation ($Z = 100\mu m$, $g_m = 6.2mS$, $V_T = 0.28V$, $V_{DSsat} = 0.4V$). Remenbering that these characteristics are valid for V_{DS} lying in the range $[V_{DSsat}, V_{GS}-V_T]$, we show the Monte-Carlo characteristics for comparison. The maximum error in drain and source currents is less than 15%. This is quite satisfying since there is no adjustable parameter. Furthermore the intermediate results of the analytical model (T_e, L_d^a, Q_D...) are found also to correlate well with the Monte-Carlo results.

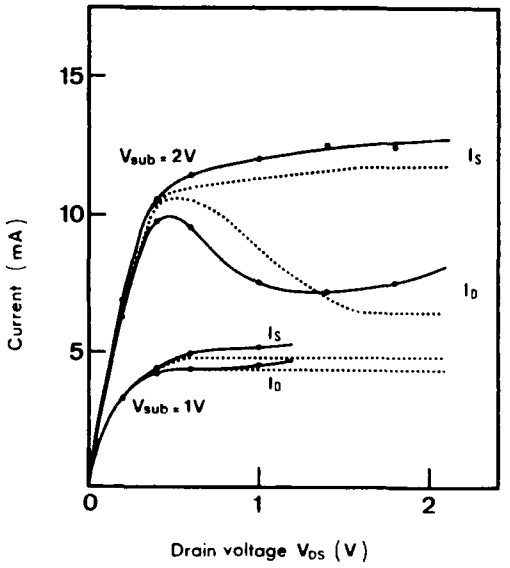


Fig.3- Comparison between the analytical (dashed lines) and the Monte-Carlo (solid lines) characteristics for two values of the substrate voltage (V_{sub} = 1V and 2V)

4-Conclusion

We have studied real-space transfer in the NERFET using a self-consistent Ensemble Monte-Carlo simulation which accurately describes the thermoionic injection range. Based on these results, we propose an analytical expression for the RST current which is well correlated with our Monte-Carlo results.

References:

[1] A. Kastalsky et al.,Solid-State Electron. 29 (1986) 1073-1077
[2] M. Mouis et al, Proceedings of the IEEE/Cornell Conference (1985) 144
[3] U. Ravaioli, D. K. Ferry, IEEE Transactions on Electron Devices, ED33 (1986) 677
[4] T. Wang, K. Hess, J. of Appl. Phys. 57 (1985) 5336
[5] P. Hesto et al., Proceedings of Nasecode IV, Boole Press (1985) 315
[6] see for instance L. Landau, E. Lifchitz, "Mécanique Quantique" (Editions Mir, Mosrou, 1966)
[7] see for instance P.J. Bulman, G.S. Hobson, B.C. Taylor, Transferred Electron Devices (Academic Press, 1972)
[8] P. Bonjour et al , IEEE Transactions on Electron Devices ED27 (1980) 1019

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

USE OF A GATE DELAY EXPRESSION TO COMPARE SELF-ALIGNED SILICON BIPOLAR AND AlGaAs/GaAs HETEROJUNCTION BIPOLAR TECHNOLOGIES

P. ASHBURN, A.A. REZAZADEH*, E.F. CHOR** and A. BRUNNSCHWEILER

Department of Electronics and Computer Science, University of Southampton, Southampton, Great-Britain
**GEC Hirst Research Centre, East Lane, Wembley, Middlesex, Great-Britain*
***Department of Electrical Engineering, National University of Singapore, Kent Ridge, Singapore*

Abstract - A comparison is made of the performance of silicon bipolar and AlGaAs/GaAs heterojunction bipolar technologies for high-speed ECL circuits. Gate delays are calculated for state of the art technologies using a quasi-analytical equation which expresses the gate delay in terms of all the time constants in the circuit. Transistor parameters are used as input to the gate delay expression and these are calculated using either device simulation programs or approximate analytical expressions. A one to one comparison is made possible by the use of an idealised but realistic, transistor layout compatible with both technologies. For an emitter width of 1μm, a collector current of $2 \times 10^4 A/cm^2$, and a unity fan-out, gate delays of 26.9 and 12.3ps are predicted for silicon and AlGaAs/GaAs technologies respectively. On scaling to 0.4μm geometries, these delays decrease to 17.2 and 11.4ps. The gate delay expression is used to identify the dominant time constants of the circuit, and hence the most promising options for process and circuit optimisation.

1) INTRODUCTION

The recent introduction of self-aligned fabrication techniques[1], [2] into silicon bipolar processes has led to a dramatic improvement in circuit performance. Gate delays of 25.8ps for NTL and 46.3ps for CML ring oscillators have been reported [3], with divider operation up to 10.4GHz [3], [4]. An analogous approach has been used in AlGaAs/GaAs heterojunction bipolar technology, with gate delays of 14.2ps being reported for CML ring oscillators [5], and divider operation up to 20.1GHz [6]. In view of these similar circuit results, it would clearly be of interest to quantify the difference in performance for circuits with identical geometries, fabricated under equivalent conditions.

In this paper, a quasi-analytical expression for the gate delay of an ECL gate [7] is used to compare the performance of state of the art silicon bipolar and AlGaAs/GaAs heterojunction bipolar technologies. Transistor parameters are used as input to the gate delay expression, and these are in turn related to the fabrication data through approximate analytical equations and/or device simulation programs. The advantage of this approach is that it provides physical insight into the dominant time constants of the circuit, and hence is an invaluable aid to process and circuit optimisation. An idealised, but realistic, self-aligned transistor layout is used for both technologies, thereby allowing a one to one comparison.

2) PROCESS AND DEVICE CHARACTERISTICS

Because silicon bipolar and AlGaAs/GaAs heterojunction bipolar transistors are used in the same logic families, there is inevitably a high degree of commonality in the processing conditions used to fabricate the two types of transistor. This can be seen in table I which summarises the fabrication data for state of the art transistors [3], [5]. Of particular interest are the identical values of basewidth and collector doping concentrations. The major difference between the two types of device is the very high base doping in the AlGaAs/GaAs transistor. This is made possible by the use of a heterojunction, where the enhanced gain can be traded for a high base doping. Equally important, a lower emitter doping can also be used (5×10^{17} cm^{-3}), thereby avoiding an unwanted increase in emitter/base capacitance.

The transistor geometries of silicon and AlGaAs/GaAs transistors are determined by the particular fabrication techniques used in their manufacture. Although there are major differences between the two technologies, analogous self-aligned fabrication techniques are used in both cases, as can be seen from the transistor cross-sections in figure 1. It is therefore possible to devise an idealised self-aligned transistor layout, which makes possible a direct comparison between the two technologies. The most critical dimensions are the oxide spacer width (or undercut of the emitter mesa), which determines the value of extrinsic base resistance, and the area of the extrinsic collector region, which determines the extrinsic collector/base capacitance. A spacer width of 0.2μm has been chosen, and the extrinsic collector area has been assumed to be equal to the intrinsic collector area.

Having decided upon a transistor layout, device simulations (BIPOLE in this case) can be used to provide the important transistor parameters, as summarised in the bottom half of table I. The predicted values of f_T of 54 and 20GHz are in good agreement with measured values of 55 and 17.1 GHz [5], [4], thereby confirming that the process parameters in table I are reasonable for state of the art transistors. It is also interesting to note that the peak f_T occurs at a higher collector current for the AlGaAs/GaAs transistor. This is probably caused by the suppression of high-level injection effects in the base of the heterojunction device, due to the higher base doping.

3) ECL GATE DELAY

From the electrical data in table I, the transistor parameters can be evaluated using standard textbook expressions, and hence the gate delay calculated from the weighted sum of all the time constants of the circuit [7]. The results are summarised in table II for an ECL circuit incorporating identical transistors with an emitter geometry of 1.0×4.0 μm^2 operating at a collector current density of 2×10^4 A/cm^2. The predicted gate delay of 12.3 ps for the AlGaAs/GaAs transistor is in good agreement with the measured value of 14.2 ps obtained by Chang et al [5] on CML ring oscillators fabricated under similar conditions to those in table I. For silicon technology, a gate delay of 26.9 ps is predicted, which is approximately 2.2 times slower than heterojunction technology.

From the results in table II, it is a simple matter to identify the limiting time constants, and hence gain a valuable insight into how the process might be improved. For the silicon transistor, there is no single dominant time constant, which implies that the transistor is well optimised, and that scaling to sub-micron geometries would be the most effective way to improve performance. For the AlGaAs/GaAs transistor, the largest time constants are the forward transit time and the load resistance terms. These latter terms could be decreased by operating the transistor at a higher collector current. This is possible for the heterojunction transistor, since the peak f_T is maintained up to a collector current density of 4×10^4 A/cm^2. The results of this modification are illustrated in figure 2, along with predictions of the gate delay as a function of fan-out. It can be seen that the increase in collector current has given rise to the expected improvement in gate delay, with the gate delay for unity fan out decreasing to 8.6 ps. This is a factor of 3.1 times faster than that obtained for silicon.

Scaling the emitter width to sub-micron geometries is the most obvious way of reducing the capacitances associated with bipolar transistors and hence improving the performance. The results of this procedure are summarised in figure 3 for an ECL circuit incorporating 0.4×4.0 μm^2 transistors. A comparison with the results in figure 2 indicates a marked improvement in the gate delay for silicon technology; 17.2 ps compared with 26.9 ps at a geometry of 1 μm. This has come about largely through a reduction in the intrinsic and extrinsic base resistance time constants. In contrast, scaling the emitter width of the AlGaAs/GaAs transistor has given only a marginal improvement in the gate delay; 11.4 ps

compared with 12.3 ps. This can be explained by noting that the reductions in capacitance obtained by scaling have been offset by increases in the series resistances. This has occurred because contact resistance is the dominant component of most of these resistances, and contact resistance increases as geometries are reduced. Effective scaling of AlGaAs/GaAs transistors therefore requires lower contact resistances than the value of $7 \times 10^{-7}\ \Omega cm^2$ used in this work.

ACKNOWLEDGEMENTS

This research was supported partly by ESPRIT project 971 "Technology for AlGaAs/GaAs heterojunction bipolar ICs".

REFERENCES

[1] T H Ning, R D Isaac, P M Solomon, D D Tang, H N Yu, G C Feth, S K Wiedmann, "Self-aligned bipolar transistors for high performance and low power delay VLSI", IEEE Trans. on Electron Devices, vol. ED28, pp.1010-1013)1981).

[2] A Cuthbertson, P Ashburn, "Self-aligned transistors with polysilicon emitters for bipolar VLSI", IEEE Jnl. of Solid State Ccts, vol. sc 20, pp.162-167 (1985).

[3] T Sakai, S Konaka, Y Yamamoto, M Suzuki, "Prospects of SST technology for high-speed LSI", IEDM Technical Digest, pp.18-21 (1985).

[4] M Suzuki, K Hagiomoto, H Ichino, S Konaka, "A 9 GHz frequency divider using Si bipolar super self-aligned process technology", IEEE Electron Device Lett, vol. EDL6, pp.181-183 (1985).

[5] M F Chang, P M Asbeck, K C Wang, G J Sullivan, N H Sheng, J A Higgins, D L Miller, "AlGaAs/GaAs heterojunction bipolar transistors fabricated using a self-aligned, dual-lift-off process", IEEE Electron Device Letters, vol EDL8, p303 (1987).

[6] K C Wng, P M Asbeck, M F Chang, G J Sullivan, D L Miller, "A 20GHz frequency divider implemented with heterojunction bipolar transistors", IEEE Electron Device Letters, vol EDL8, p.383 (1987)

[7] E F Chor, A Brunnschweiler, P Ashburn, "A propagation delay expression and its application to the optimisation of polysilicon emitter ECL processes", IEEE Jnl Solid State Ccts, vol SC 23, p.251 (1988).

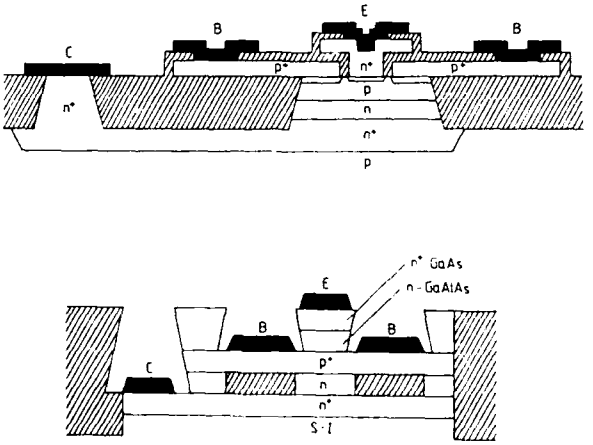


Fig. 1 Cross-sectional views of silicon bipolar and AlGaAs/GaAs heterojunction bipolar transistors

PROCESS/DEVICE PARAMETER	AlGaAs/GaAs TECHNOLOGY	SILICON TECHNOLOGY
Emitter doping cm^{-3}	5×10^{17}	3×10^{20} peak
Base doping cm^{-3}	1×10^{20}	5×10^{17} peak
Collector doping cm^{-3}	6×10^{16}	6×10^{16}
Basewidth µm	0.06	0.06
Intrinsic base sheet resistance kΩ/sq	0.17	10.1
Peak f_T (V_{CB}=1.5V) GHz	54	20
I_c for peak f_T A/cm^2	$2\text{-}4 \times 10^4$	$1\text{-}2 \times 10^4$
Forward transit time ps	2.6	5.0

Table 1

Process and device parameters for state of the art transistors

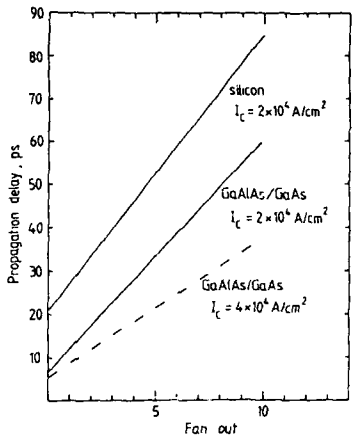


Fig. 2 Propagation delay as a function of fan-out for silicon bipolar and AlGaAs/GaAs heterojunction bipolar ECL circuits. Emitter geometry = 1.0 x 4.0µm^2, logic swing = 0.4V

TIME CONSTANT	PROPAGATION DELAY ps	
	AlGaAs/GaAs TECHNOLOGY	SILICON TECHNOLOGY
τ_F	3.12	6.00
$r_{bi}C_{jci}$	0.03	1.86
$r_{bi}C_{je}$	0.02	1.12
$r_{bi}C_d$	0.04	4.07
$r_{bx}C_{jci}$	0.11	1.07
$r_{bx}C_{jcx}$	0.10	1.01
$r_{bx}C_{je}$	0.06	0.76
$r_{bx}C_d$	0.13	2.50
R_LC_{jci}	1.40	1.40
R_LC_{jcx}	1.90	1.90
R_LC_{je}	0.75	0.90
R_LC_L	3.00	3.00
TOTAL	12.3	26.9

Table 2

Dominant components of the propagation delay for ECL circuits incorporating 1.0 x 4.0µm^2 transistors

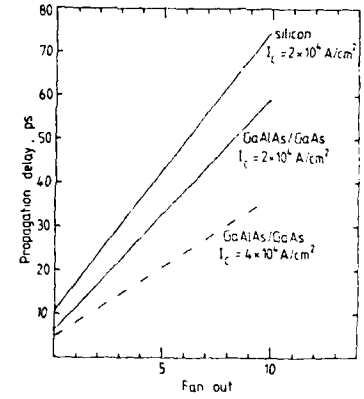


Fig. 3 Propagation delay as a function of fan-out for silicon bipolar and AlGaAs/GaAs heterojunction bipolar ECL circuits
Emitter geometry = 0.4 x 4.0 µm^2, logic swing = 0.4v.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

COMPARISON OF ECL GATE PERFORMANCES USING DIFFERENT HETEROJUNCTION BIPOLAR TRANSISTORS PROCESS

J. DANGLA and D. HAVOND

Centre National d'Etudes des Télécommunications, Laboratoire de Bagneux, 196, rue H. Ravera, F-92220 Bagneux, France

Résumé - Les performances en commutation d'un transistor bipolaire à hétérojonction GaAlAs/GaAs (TBH) sont examinées dans le cadre d'une application logique ECL à partir d'un modèle de type schéma électrique équivalent. Ce modèle a été validé à l'aide de résultats expérimentaux. L'influence sur le temps de propagation des paramètres de couche, des règles de dessin et des procédés technologiques est étudiée. Les simulations montrent l'existence d'un optimun à la fois sur l'épaisseur de base et sur la largeur du doigt d'émetteur. Finalement la technologie la mieux adaptée aux performances ultimes se révèle être la technologie mésa.

Abstract - The switching performance of GaAlAs/GaAs n - p - n Heterojunction Bipolar transistors (HBT) has been investigated for Emitter Coupled Logic (ECL) circuit operation using a CAD model witch has been validated by experimental results. Switching time is discussed in conjunction with layer parameters, design rules, and technological process. The simulation shows the existence of an optimun both for the base thickness and the emitter width. Finally Mesa process give lower switching time than the implanted one.

1 - INTRODUCTION

The HBT offers a unique combination of advantages. It drawns on existing bipolar logic integrated circuit design and performances, the speed of GaAs and the gain enhancement of a wide band gap emitter. The aim of this paper is to optimize the HBT structure /1,2/ and process parameters in order to find the minimun delay time taking into account the technological state of the art. In this study the propagation delay time is the only investigated parameter. A single - stage ECL gate was employed as a vehicule for the simulation work. The logic swing is 400 mV, and the current across the current source is 2 mA.

This work can be of a great interest in order to drive process investigations, find ultimate performances compatible with technolgical process, find the best advanced process.

The HBT has been chosen because is is a very fast device. Maximun available frequency up to 150 GHz /3/, divider by four toggling at 22 GHz /4/ and switching time below 5 ps with ECL logic /5/ have been reported, thus verifying the high potential of HBT. The HBT have also high driving capability at low current and good threshold reproducibility. This advantages come from the bipolar device, the high mobility of the gallium arsenide material and the possibilities given by emitter bas heterojunction and deep proton implantation. The epitaxial technique permits to have base grading, and modified collector structure which decreases the total transit time.

Emitter coupled Logic has been chosen for the developement of the digital circuit in our laboratory because it is the fastest, it have a good nose marging and it is widely used in silicon bipolar circuit design.

This work is partially supported by the European Economic Community (ESPRIT Project No 971)

The methodology used for the simulation is the following : the standard HBT fabricated in our laboratory is chosen has the starting point of the simulation work, the influence on propagation delay time of each parameter separately is studied. Finally a global optimisation of the ECL gate is performed. The parameter have been shared in three classes : layer parameters, design rules and technological process. The last parameter permits to choose the advanced process which will be used. Figure 1 presents the schematic drawing of the standard GaAlAs/GaAs HBT developed in our laboratory. This structure is now used in the fabrication of integrated circuit /6/. In order to have a good yield, the dimensions are quite still large in particular the emitter base distance is 4 um while the emitter finger metallization size is 3 X 8 um^2. On this device, the base is contacted from the surface through the emitter layer by use of a Mg implantation. The base doping level is only 5 10^{18} cm^{-3} in order to avoid Be diffusion during the implantation annealing.

For the logic gate simulation work, a large signal CAD model /7/ based on the well known Ebers Moll model has been developed. In order to reproduce the behaviour of the HBT, external diodes, recombination diodes and external resistance are added. The input parameters of this model are the layers parameters, the design rules and the technological process. This model has been widely validated by comparison with experimental results obtained either in our laboratory or already published.

2 - LAYER PARAMETERS OPTIMIZATION

2 - 1 Doping levels optimisation

Figure 2 shows the variation of the propagation delay time versus the different doping levels. On each curve the dot represents the reference data parameter which corresponds to a delay time of about 160 ps. These curves show that the delay time decreases when the emitter and collector doping levels raise and the base doping level falls. Such variation can be easily explained by decrease of the junction capacitances when the emitter and collector doping levels decrease; on the other hand, the base resistance decreases when the base doping level increases. Some interesting information can be deduced from these curves. Firstly the base doping level is the most interesting parameter on switching speed, secondly there is no significant change in switching speed for emitter and collector doping levels smaller than 2 10^{16} cm^{-3}.

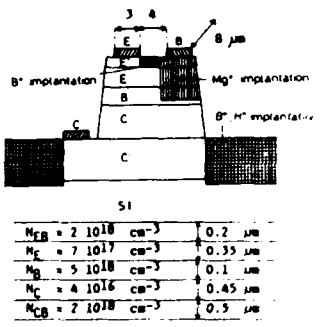


Figure 1 Schematic drawing of the standard GaAlAs/GaAs HBT developed in our laboratory.

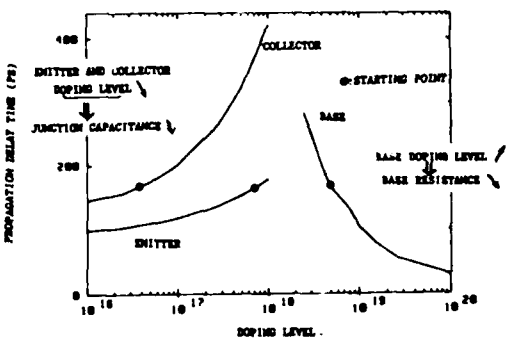


Figure 2 Variation of the propagation delay time versus the different doping levels

2 - 2 Layers thickness

The next step is devoted to the study of thickness influence. The important point is the existence of an optimun for the base layer layer thickness which is in this case 0.4 um. This results of the opposite effects of

this thickness on the base resistance and on the base transit time. At a thickness of 0.4 um the current gain is very low and the ouput signal may be fall down for a fan out different to zero. So similar computation is performed with a fan out at 1. The results are the same. The others thicknesses have no significant influence on the switching performances.

2 - 3 Global optimization

Combining the previous optimizations of doping levels and layer thickness, the result of a global optimization is shown on figure 3. By use of a base doping level of 5 10^{19} cm^{-3}, emitter and collector doping levels of 2 10^{16} cm^{-3}, the propagation delay time has been reduced from 160 ps to less than 20 ps. Furthermore, if the base thickness is optimized as shown previously, the delay time can be as low as 20 ps using the large design rules of the reference HBT.

3 - DESIGN RULES OPTIMIZATION

Figure 4 represents the influence of the design rule on delay time. The most important parameter is the distance between the emitter and the base implantation because it directly influences the external base resistance. The second important parameter is the emitter metalization width. Beyond two microns, its reduction doesn't reduce the transit time. This result is important, because high performance HBT can be processed with conventionnal lithography for ohmic contact realization. The two others distances, lateral diffusion of the implantation and the distance between emitter contact and mesa have a very low influence on the delay time. This is due to the weak influence of the emitter base and collector base capacitance on the delay time.

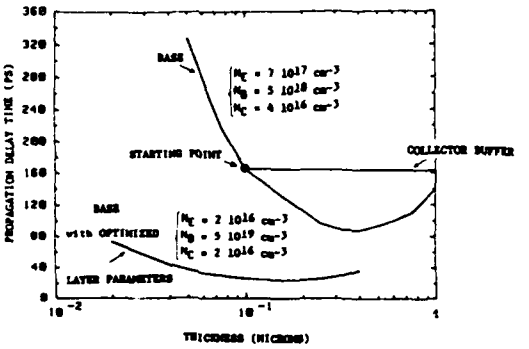


Figure 3 Global optimization of layers parameters.

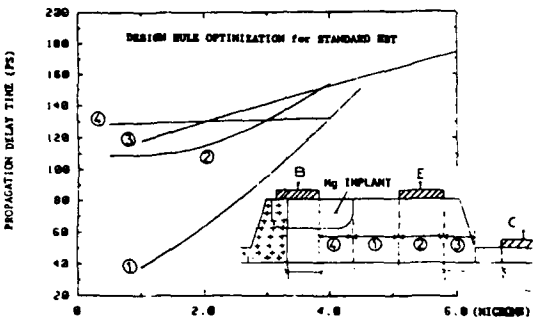


Figure 4 Influence of the design rule on delay time.

4 - TECHNOLOGICAL PROCESS OPTIMIZATION

Previous simulation have shown that the base resistance is one of the most important parameter for reducing the delay time. In order to reduce the base emitter distance, self-aligned processes can be used. The first one can be achieved by use of a refractory emitter metallization which is used as a mask for the base implant. The base acces resistance is reduced by this way. Unfortunaly it presents the common disadvantage for all implanted technology to limit the value of the base doping level. The second one is to use the emitter metallization to as a mask to etch the emitter layer and to obtain the self aligned double mesa technology. Its main difficult is the selective etching of the GaAlAs/GaAs. In order to increase the performances, a deep proton can be added into the external collector region which reduces the capacitance in this region.

On the table 1 are reported the optimized propagation delay times for various processes with and without deep proton implantation. All process present very close results. The delay time of a non self-aligned HBT

process with very high base doping level (2 $1O^{20}$ cm^{-3}) is lower than both the implanted self-aligned process and double mesa self-aligned with base doping level smaller than 5 10^{19} cm^{-3}. The better result (7.7 ps) is obtained with the double mesa self-aligned process with deep proton implantation. For all process the delay time pourcentage between the HBT process with and without deep proton does not exceed 18 %.

Technological process	Base doping (cm^{-3})	Emitter base distance (microns)	Delay time without deep proton implantation (ps)	Delay time with deep proton implantation (ps)
Implanted	10^{19}	1	16.7	13.9
Implanted self-aligned	10^{19}	0.2	13.5	11.5
Double mesa	5.10^{19}	1	13.7	11.8
	2.10^{20}	1	10.7	8.9
Double mesa self-aligned	5.10^{19}	0.2	11.6	9.6
	2.10^{20}	0.2	9.3	7.7

Table 1 Optimized propagation delay times for various processes with and without deep proton implantation.

5 - SUMMARY

Emitter coupled logic gates is simulated and the influence on the delay time of each parameter is given. The major parameters are the base doping level and the emitter base distance. There are two parameters which present an optimun value : the base thickness and the emitter width. The emitter and collector doping levels, the self aligned process, the deep proton implantation have a quite relative influence. Finally the minimun value of the transit time (7.7 ps) is obtained by double mesa technology with very high base doping level.

/1/ R. KATOH et Al.: "Numerical CML switching analyses for herojunction GaAs/(GaAl)As bipolar transistors" Solid State Elect., Vol. 29, No 2.pp 151 - 157,1986.
/2/ J. TASSELLI et Al.: "Etude théorique des performances de l'inverseur ECL à base de transistors bipolaires à hétérojonction GaAlAs/GaAs" Revue Phys. Appl. 23 (1988) 1117 - 1125
/3/ P. M. ASBECK et al. :" Heterojunction Bipolar Transistors for Microwave and Millimeter - wave Integrated Circuits" IEEE Transaction on Electron Device, Vol. ED - 34, No. 12, December 1987.
/4/ Y. YAMAUCHI et al. : "22 GHz 1/4 frequency divider using AlGaAs/GaAs HBTs" Electronics Letters, 13th August 1987, Vol. 23, No. 17.
/5/ Y. YAMAUCHI et al. : "Ultra - High - Speed Emitter Coupled Logic Circuits using AlGaAs/GaAs HBTs" to be published.
/6/ P. DESROUSSEAUX et al: "GaAlAs/GaAs heterojunction bipolar transistor models HECL circuits design", IEEE BTCM 87, pp 200 - 202, MINNESOTA.
/7/ J. DANGLA et al. :"A CAD Model For Heterojunction Bipolar Transistor", ESSDERC 1987, Bologne.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

MICROWAVE POWER GaAs/AlGaAs HETEROJUNCTION BIPOLAR TRANSISTOR MODELLING

J.G. METCALFE, R.C. HAYES, A.J. HOLDEN and A.P. LONG

Plessey Research Caswell Ltd., Caswell, Towcester, Northants., GB-NN12 8EQ, Great-Britain

Abstract - A high power heterojunction bipolar transistor has been designed and fabricated. A model of this transistor has been developed for the SPICE simulation package. A program has been written to use SPICE to run simulations on the model and calculate matching conditions for optimum output power.

The aim of this work was to model HJBT power transistors and demonstrate their advantage over Si bipolar and GaAs MESFETs. Microwave power transistors have traditionally been of two types: Si bipolar for higher power applications and GaAs FETs for higher frequency. In applications such as radar systems at X-band and above, both high power and high frequencies are required and the GaAs/AlGaAs bipolar heterojunction transistor offers a discrete device solution in these areas [1].

Microwave power heterojunction bipolar transistors have been fabricated from high uniformity MOCVD GaAs/AlGaAs layers grown on 2 inch n+ substrates. A multiple interdigitated emitter-base arrangement is used to achieve high current operation (Fig. 1). The collector contact is formed on the back of the thinned wafer. Base and isolation implants are used to maintain a near planar surface geometry for the transistor. A key element of the transistor design is the use of an extended metal contact scheme to eliminate debiasing of the long emitter stripes (4μ x 100μ).

We have developed a large signal model for the power GaAs HJBT to run in the SPICE circuit simulation program. This is based on Plessey's established SPICE bipolar model and has been fitted to measurements made from prototype devices manufactured at Caswell.

The SPICE model is shown in figure 2. The transistor is biassed at a desired collector current and collector-base voltage by means of the two current sources, the use of which allows convergence of the output power in a low number of periods. The matching is achieved by series resistor, inductor and capacitor elements on the input and output circuit (Fig.2a). The basic SPICE transistor model has been enhanced by including distributed and non-ideal effects appropriate for a large transistor structure (Fig.2b).

Provided that heat dissipation is not a problem (pulsed operation), the main limitations on the biassing of the transistor and hence the available output power are the collector-base breakdown voltage and the maximum collector current determined by the Kirk effect in the collector. The maximum collector current density is set at $5 \times 10^4 A/cm^2$ by biassing Ic to half that value (Fig.2a). This value agrees with theoretical considerations and device measurements. The

breakdown condition is modelled by a separate SPICE sub-circuit (Fig.2c). It is modelled as an avalanche effect; the relationships between the labelled values in the diagram are:

$$I_1 = \frac{V_{CB}^N \times I_S(\text{diode})}{V} \qquad I_{CB} = \frac{I_2 \times I_C}{I_S(\text{diode})}$$

$$I_{CB} + I_C = \frac{I_C}{1-(V_{CB}/V_{BB})^N}$$

The value of the breakdown voltage V was taken as 27V and the power, N, as 5 by fitting to measurements of open base and open emitter breakdown voltages.

The object of the modelling is to predict the values of input and output impedance which give optimum output power. The program starts with values of input and output impedance and calculates the output power at 1dB gain compression. The impedances are varied and the maximum power is found using a minimisation routine. The 1dB gain compression point is found each time by using a root finding routine and comparing the gain with the maximum value achieved. The output power is determined by using SPICE to simulate the circuit and determining the Fourier component of the output voltage waveform at the frequency of interest. The curve of gain against input power at the optimum impedance is shown in figure 3; the kink in the region 30dBm, 5.3dB is due to the influence of the breakdown effect.
The model currently predicts a power density of 4.3W/mm of emitter periphery for common-emitter operation at 10 GHz, which is consistant with [1]. This compares with up to 1W/mm available from conventional power MESFETs and a frequency limit of 3GHz for power Si bipolar. The real part of both the input and output impedances was calculated to be 3.5 and 5.4ohms respectively, for a 1.2 mm periphery. This simulation was performed with the transistor in the common-emitter configuration (Fig.2A) and operating as class A amplifier, though the model can be used with any common terminal and in any class.

This work demonstrates the advantage of the GaAs/AlGaAs HJRT and shows that it is possible to successfully model its operation and design appropriate matching networks.

ACKNOWLEDGEMENT
This work has been supported by the Procurement Executive, UK Ministry of Defence (DCVD), through RSRE Malvern

REFERENCE

[1] B.Bayraktaroglu, N.Camilleri and H.O.Tserng, IEEE Advanced Concepts in High Speed Semiconductor Devices and Circuits (1987), "AlGaAs/GaAs HETEROJUNCTION BIPOLAR TRANSISTORS FOR POWER APPLICATIONS"

Figure 1: Electron micrograph of Fabricated Power HJBT

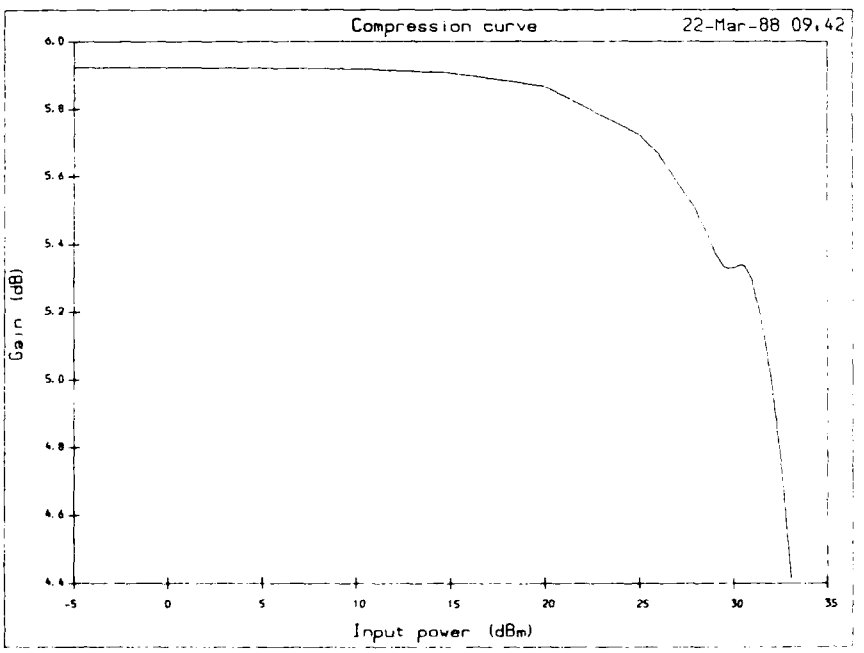


Figure 3: Compression Curve

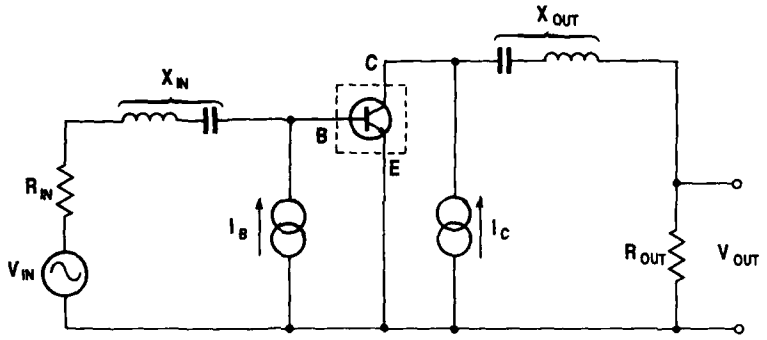


Figure 2a Biassed and Matched Transistor - Common Emitter

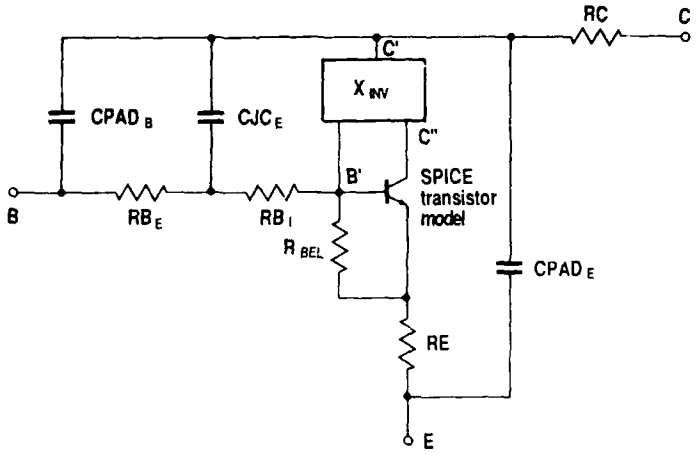


Figure 2b Enhanced SPICE Transistor Model including Breakdown

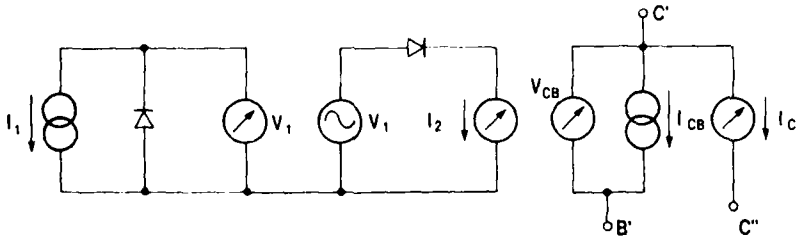


Figure 2c X_{INV} - SPICE Breakdown Circuit

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988 C4-583

NOISE OF GaAs DIODES

A. MOATADID, D. GASQUET, M. de MURCIA and J.P. NOUGIER

Centre d'Electronique de Montpellier (CNRS-UA 391), Université des Sciences et Techniques du Languedoc, F-34060 Montpellier Cedex, France

RESUME L'etude à faible polarisation d'une diode n⁺nn⁺ GaAs à profil de dopage réel montre, par comparaison entre le bruit expérimental et le bruit modélisé, que la méconnaissance actuelle de la variation precise du coefficient de diffusion D(E), en fonction du champ électrique E, peut conduire à une modélisation erronnée, en particulier des caractéristiques de bruit de composants GaAs Les profils de champ electrique et de densité de porteurs libres, sont étudiés également en mode d'oscillations Gunn

ABSTRACT By comparing experimental and modelled noise results of a n⁺nn⁺ GaAs diode, we show that the lack of precise knowledge on the variation of the diffusion coefficient D(E), versus the electric field E, may lead to erroneous predictions, in particular as concerning the noise behaviour of GaAs devices The electric field and free carrier density profiles are also studied in Gunn oscillation operating regime

1. INTRODUCTION:

In spite of numerous studies of hot carriers in GaAs the variation of the diffusion coefficient D(E), versus the electric field E is not well known ([1] to [4]) The aim of this paper is to demonstrate that this may lead to erroneous prediction of the behaviour of devices, particularly as concerning noise characteristics For this purpose, we shall model the simple but realistic case of n⁺nn⁺ diodes, with a doping profile $N_D(x)$

2. STEADY STATE CHARACTERISTICS:

The devices modelled are about 10 µm long. Then the classical transport equations may be used, we do not need using the dynamic equations for submicron devices, but one should of course take into account hot carrier effects) The total current I(t) is then the sum of drift, diffusion, and displacement currents With obvious standard notations, this writes, with $q = +1.6\times10^{-19}$ Cb, $V(x=0)=0$, $V(x=L)=V_L>0$, $E(x)<0$, $v[E(x)]>0$, $I<0$

$$I(t) = q\,n(x,t)\,v[E(x,t)]\,A - q\,A\,D[E(x,t)]\frac{\partial n(x,t)}{\partial x} - \varepsilon\,A\frac{\partial E(x,t)}{\partial t} \qquad (1)$$

$$\frac{\partial E(x,t)}{\partial x} = \frac{q}{\varepsilon}[n(x,t)-N_D(x)] \qquad (2)$$

Eliminating n(x,t) gives eq (3), where v and D stand for v[E(x,t)] and D[E(x,t)]

$$\frac{\partial E}{\partial t} = -D\frac{\partial^2 E}{\partial x^2} + v\frac{\partial E}{\partial x} + \frac{q}{\varepsilon}v\,N_D(x) - \frac{qD}{\varepsilon}\frac{dN_D(x)}{dx} - \frac{I(t)}{\varepsilon A} \qquad (3)$$

In the present section, we are interested in studying the steady state regime ($\partial E/\partial t=0$) We drop the diffusion current, we found negligible From eq (3), we get then the electric field profile $E_0(x)$ as a solution of

$$\varepsilon\,A\,v[E_0(x)]\frac{dE_0(x)}{dx} + q\,A\,N_D(x)\,v[E_0(x)] = I_0 \qquad (4)$$

The diode extends from x=0 to x=L These points should be chosen far enough inside the n⁺ electrodes, so that the conduction is ohmic at x=0 and at x=L One should then have, since $N_D(x=0) = N_D(x=L)$

$$E_0(x=0) = E_0(x=L) = I_0 / [q\,A\,N_D(x=0)\,\mu_0(0)] \qquad (5)$$

The first order differential equation (4) was solved, for a given value of the bias current I_0 using a predictor-corrector method, with the initial condition given by eq (5) The characteristic field E_c and the saturation drift velocity v_s of the v(E) law [6], were taken as $E_c = 4$ kV/cm and $v_s = 1.1\times10^7$ cm/s The discretization step Δx was not constant The commercially available diode presented here, labelled G2, had an active thickness of 10.6 µm The origin x=0 and the extremity x=L were taken 2 µm far from the n⁺n junction, hence L=14.6 µm The diameter was 125 µm, the ohmic resistance R_0=11 Ω, the ohmic mobility μ_0=4800 cm^2/Vs

Figure 1 shows the doping profile of the diode G2, so as the electric field profiles obtained at different d.c. current biases The electric field intensity always exhibits a spike, even at low bias

As shown figure 2, the agreement, between the experimental and the computed $I_0(V_0)$ characteristics, is excellent at every bias below the Gunn oscillation threshold.

3. GUNN EFFECT:

For bias current higher than 200 mA, oscillations appeared. It was then interesting to model them, since we had the tools, although a lot of work has already been done in that field (see for example [6]).

For modeling the Gunn effect, one should take into account the output circuit namely a voltage supply $e_g(t)$ *in series with a resistor R. The time dependant regime is governed by eqs.* (3), where $E = E(x,t)$, and eq. (6):

$$V(t) = -\int_0^L E(x,t)\,dx \qquad \text{and} \qquad V(t) = e_g(t) - RI(t) \tag{6}$$

$e_g(t)$ was assumed to be linear with t, from 0 at t=0 to E_g at time 100 ps, and then constant. If the electric field profile $E(x_i,t)$ is known, at each point x_i, at a given time t, a numerical integration gives V(t), then I(t) (eq (6)). A fourth order predictor-corrector method gives then $E(x_i,t+\Delta t)$, since the right hand side of eq. (3) is numerically known.

The electric field profile at t=0 is solution of eq. (3), where $I(t=0) = 0$ and $\partial E/\partial t = 0$, with the boundary conditions given by eqs. (5) and (6). This system can be easily solved using a double sweep iterative method.

As an example, results obtained in the Gunn oscillation regime are displayed figure 3 . The free carrier density profiles, drawn every 10 ps on fig. 3a , show the formation and the propagation of domains, and fig. 3b shows the time evolution of current through the diode.

As a comparison, we show figure 4 the same quantities for a diode with a uniform doping profile in the active n region: the domains form in that case in a much more regular way than in the real diode (compare fig. 4a and fig 3a). As a consequence, the current in the diode with a uniform doping (see fig. 4b) is much more sinusoidal, with a period better defined.

4. NOISE MODELING AT LOW BIAS:

We are interested in modeling the "low frequency" noise, i.e. at frequency lower than 100 GHz, corresponding to time constants much larger than the dielectric relaxation time, below the Gunn oscillation threshold. then we start again from eq. (3) and we apply the impedance field method [7]. First, eq (3) writes, when neglecting the diffusion and the displacement currents:

$$\varepsilon A\, v[E(x,t)] \frac{\partial E(x,t)}{\partial x} + q A N_D(x)\, v[E(x,t)] = I(t) \tag{7}$$

One then sets $I(t) = I_0 + \delta I \exp(i\omega t)$ and $E(x,t) = E_0(x) + \delta E(x)\exp(i\omega t)$. These expressions are carried into eq. (9). The zero order terms give back eq (4). The first order terms give

$$a(x)\frac{d\delta E(x)}{dx} + b(x)\,\delta E(x) = \delta I(x) \qquad \text{with} \qquad a(x) = \varepsilon A v_0 \qquad \text{and} \qquad b(x) = \frac{I_0}{v_0}\frac{dv_0}{dE_0} \tag{8}$$

The quantities a(x) and b(x) are known, numerically, from section 2 above. The Green function of eq (9) can then be found, leading to the impedance field $\nabla Z(x')$, given as

$$\nabla Z(x') = -\frac{1}{a(x')}\int_{x'}^{L} K(x,x')\,dx \qquad \text{with} \qquad K(x,x') = \exp -\int_{x'}^{x}\frac{b(u)}{a(u)}\,du \tag{9}$$

The differential impedance Z, the noise voltage S_V and the noise current S_I for diffusion noise, are then given by:

$$Z = \int_0^L \nabla Z(x')\,dx' \,, \qquad S_V = 4q^2 A \int_0^L |\nabla Z(x')|^2\, n_0(x')\, D_0(x')\,dx' \qquad \text{and} \qquad S_I = S_V/|Z|^2 \tag{10}$$

The variation of the differential impedance Z, versus the bias current I_0, computed using eq.(10), is shown fig 5, so as the differential impedance obtained by computing the derivative of the $I_0(V_0)$ characteristic. Both quantities are in excellent agreement with the experimental differential impedance.

As can be seen from eqs. (10), the determination of the diffusion noise implies the knowledge of the variation of the diffusion coefficient versus the electric field. Unfortunately, very few results have been published in the litterature till now, as concerning GaAs ([1], [8] to [10]), and the results available differ quite significantly. Theoretical models also exhibit quite different variations of D(E) versus E, according to the values chosen for the coupling constants ([4], [11] to [13]) the results obtained are quite similar at low field, but differ by a considerable amount at fields higher than 2 kV/cm.

Figure 6 shows the experimental noise of the diode G2, measured using a pulse technique in order to avoid thermal heating. We verified that the noise was white in the range 220 MHz - 10 GHz, so that we actually deal with diffusion noise. Also are shown fig. 6 the theoretical noise computed through eq. (10), using the variations [4][12]

of D(E) available in the litterature. As can clearly be seen on fig 6, the experimental and the computed results are in good agreement at low bias, but none of the two theoretical models is able to account for experimental results at higher bias. This figure clearly shows that the noise predictions strongly depend on the variation law of D(E), and can be quite erroneous according to the law choosen: this demonstrates the usefulness of a precise knowledge of D(E), and also exhibits a lack of available data in the litterature as concerning GaAs.

Obviously, not enough precise data D(E) are now available, this determination needs both theoretical and mostly experimental efforts. Of course, this effect, pointed out in the present paper in the case of diodes, also remains valid in the case of any GaAs device exhibiting diffusion noise.

REFERENCES:

[1] J.G. RUCH, G.S. KINO, Phys. Rev. 174, 921 (1968)
[2] N. BRASLAU, P.S. HAUGE, IEEE Trans. Electron Dev. ED17, 616 (1970)
[3] J.G. RUCH, W. FAWCETT, J. Appl. Phys. 41, 3843 (1970)
[4] A. KASZINSKI, Thèse de 3ème cycle, Lille (France), 1979, available on request.
[5] H.W. THIM, J. Appl. Phys. 39, 3897 (1968)
[6] G.S. HOBSON, "The Gunn effect", Monographs in Electrical and Electronic Engineering, ed. P. HAMMOND and D. WALSH, Clarendon Press Oxford, 1974
[7] W. SHOCKLEY, J.A. COPELAND, R.R. JAMES, in "Quantum theory of atoms, molecules and the solid state", P.O Lowdin ed., Acad. Press, New York, chap. 8 (1966)
[8] D. GASQUET, M. DE MURCIA, J.P. NOUGIER, C. GONTRAND, Physica 134B, 264 (1985)
[9] D. GASQUET, M. DE MURCIA, J.P. NOUGIER, M. FADEL, "Noise in Physical Systems and 1/f Noise 1985", ed. A D'AMICO and P. MAZETTI, Elsevier Science Publisher, Amsterdam, p. 227 (1986)
[10] V. BAREIKIS, V. VIKTORAVISHYUS, A. GAL'DIKAS, A. MILYUSHITE, Sov. Phys. Semicond 14, 847 (1980)
[11] A.J. POZELA, A. REKLAITIS, Solid State Commun. 27, 1073 (1978)
[12] A.J. POZELA, A. REKLAITIS, Solid State Electron. 23, 927 (1980)
[13] M.A. LITTLEJOHN, J.R. HAUSER, Th. GLISSON, J. Appl. Phys. 48, 4587 (1977)

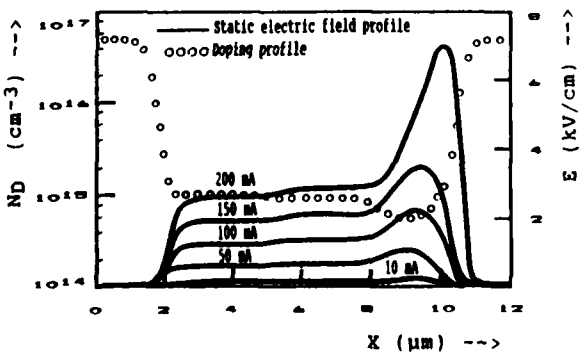


Fig. 1 n^+nn^+ GaAs diode G2 at 300 K, doping profile $N_D(x)$, and electric field profiles at different d.c current bias

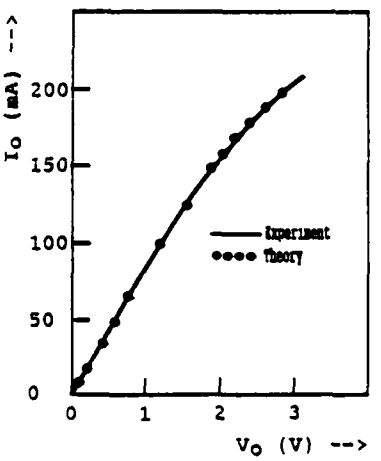


Fig. 2 Static characteristics $I_O(V_O)$ of diode G2

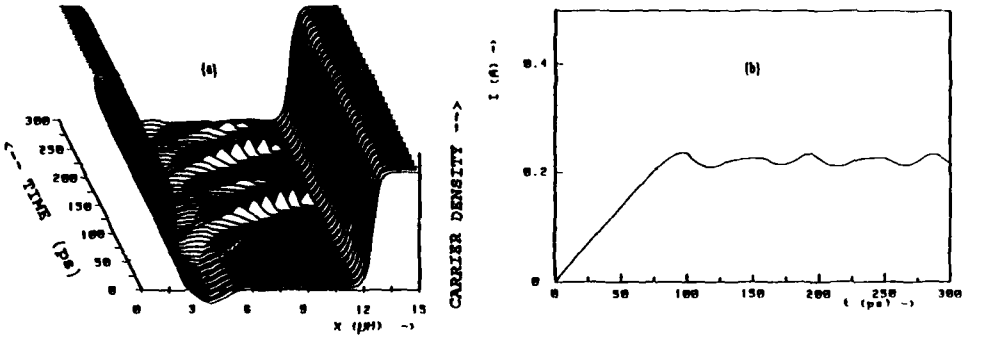


Fig. 3: Diode G2 at 300 K, in Gunn oscillation mode R=25 Ω, E_g=10 V. Time evolution of:
(a) The free carrier density profile (b) The current through the diode.

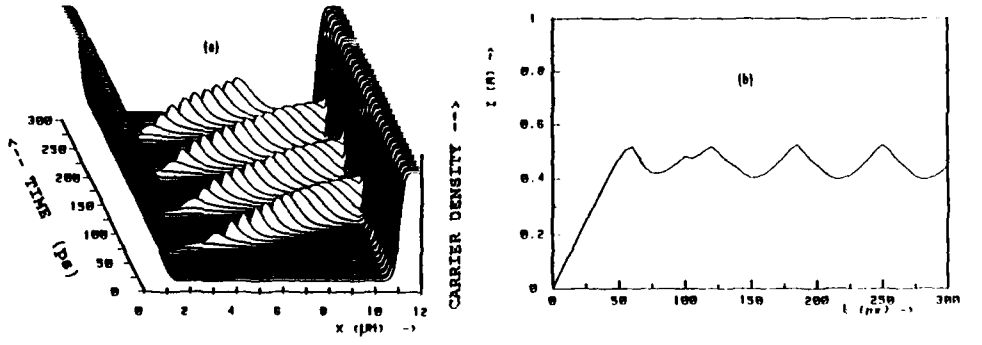


Fig. 4: n^+nn^+ GaAs diode at 300 K, uniformly doped in the active region R=20 Ω, E_g=25 V. Time evolution of
(a) The free carrier density profile (b) The bias current

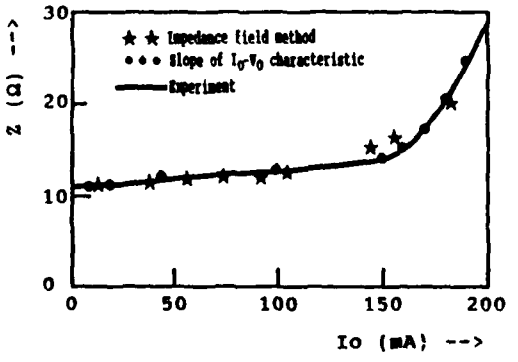


Fig. 5: Differential impedance Z of diode G2 at 300 K versus the d.c. current bias I_0

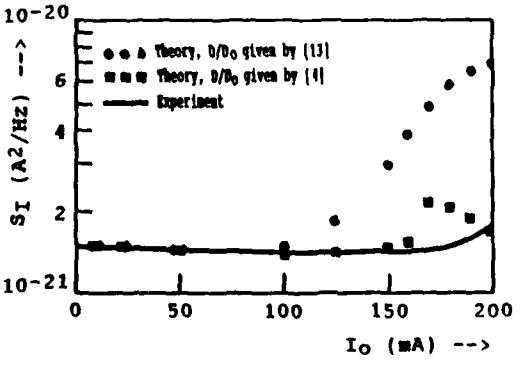


Fig. 6: Noise current spectral density S_I of diode G2 at 300 K, versus the d.c. current bias I_0

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988 C4-587

PHOTOVOLTAIC INFRARED SENSOR ARRAY IN HETEROEPITAXIAL NARROW GAP LEAD-CHALCOGENIDES ON SILICON

J. MASEK, C. MAISSEN, H. ZOGG, S. BLUNIER, H. WEIBEL, A. LAMBRECHT*, B. SPANGER*, H. BÖTTNER* and M. TACKE*

AFIF (Industry Research Unit) at Swiss Federal Institute of Technology, ETH-Hönggerberg, CH-8093 Zürich, Switzerland
**Fraunhofer-Institut für Physikalische Messtechnik, Heidenhofstr., 8, D-7800 Freiburg, F.R.G.*

Abstract - Linear arrays of photovoltaic IR-sensors for thermal imaging applications have been fabricated for the first time in narrow gap semiconductor layers grown heteroepitaxially on Si-substrates. Heteroepitaxy was achieved using intermediate stacked CaF_2-BaF_2 bilayers to overcome the large lattice- and thermal expansion mismatch between Si and lead-chalcogenides. Sensors fabricated in narrow gap PbTe have ≈5.7 µm cut-off wavelength at 90K and quantum efficiencies around 70%. Resistance-area products are up to 500 Ωcm^2 with mean value of ≈150 Ωcm^2 for 66 element linear arrays, well above the room temperature photon background noise limit. Sensor arrays with shorter cut-off wavelength were fabricated in the same manner in epitaxial $Pb_{1-x}Eu_xSe$ on fluoride covered Si-substrates.

INTRODUCTION

The ultimate goal of infrared focal plane array (IR-FPA) development for thermal imaging applications is a fully staring array with a flux integrating sensor for every image pixel. The array includes signal multiplexing circuitries for processing the large numbers of simultaneously generated electrical signals. The preferred spectral range for such arrays are the 3-5 µm or 8-12 µm atmospheric window. Several approaches are currently being developed or already applied towards this goal [1]:

- Extrinsic Si- and PtSi-sensor arrays with CCD read-out are fabricated in Si with standard Si-processing. However, they require very low operating temperature, their useful spectral range does not coincide with the atmospheric window and quantum efficiencies especially of silicide sensors are very small.
- Thermal IR-detector arrays may operate at room temperature, but suffer from low sensitivity and low speed. Hybrid combinations with Si-signal processing are needed.
- Intrinsic sensors fabricated in narrow gap semiconductors seem to be most suited for IR-detection. They have high quantum efficiencies, their cut-off wavelength can be tailored to the specific application by using a suitable composition of ternary alloys, and they have lowest noise allowing operation at higher temperatures compared to other types of sensors. However, fabrication of signal processing circuitries in these materials is a very difficult task or even impossible, especially for long wavelength materials. Therefore, most of the present focal plane arrays with intrinsic sensors are realized in a hybrid manner: A Si chip containing the multiplexing and read-out electronics is mated with a narrow gap semiconductor chip which contains the linear or square array of sensors. Since an electrical connection is needed for every pixel, such an approach becomes rather complicated and expensive for large 2-dimensional arrays and is limited in the pixel size.

To combine the advantages of narrow gap semiconductors for IR-detection with the signal processing capabilities of Si a fully monolithic approach would therefore be highly advanta-

geous. In this presentation, we report on our new results towards this goal. We have grown epitaxial narrow gap lead-chalcogenide layers (PbTe and $Pb_{1-x}Eu_xSe$) on Si-wafers by using intermediate fluoride buffer layers and have realized whole sensor arrays for the 3-5 µm range in these structures for the first time.

We already described single IR-sensors fabricated in PbTe as well as in $Pb_{1-x}Sn_xSe$ on fluoride covered Si-substrates which exhibited cut-off wavelengths up to ≈9.5 µm and which operated near or at the background noise limit (BNL) for room temperature radiation [2]. A stacked intermediate CaF_2-BaF_2 buffer layer was used to obtain epitaxy of the desired quality because such graded fluoride buffers allow to overcome the large lattice as well as thermal expansion mismatch [3] between the Si-substrate and top semiconductor layer which may be a IV-VI lead-chalcogenide [4], a II-VI-compound [5] or even a further material. The lattice mismatch between PbTe and Si is 19%, while the thermal expansion coefficients are different by a factor of 7. A further reason why we use fluoride intermediate buffer layers is that **bulk** BaF_2 single crystals have been successfully used by different groups as a substrate for epitaxial growth of lead-chalcogenides and fabrication of photovoltaic sensors in these layers [6,7].

EXPERIMENTAL

Si(111) wafers were used for the present work since fluoride growth is easiest with this orientation. However, we have recently grown lattice mismatched CaF_2-SrF_2-BaF_2 stacks on (100)Si-wafers [8] too, which is the preferred orientation for CMOS-technology. The substrates were cleaned using a modified Shiraki method which needs an in situ anneal up to ≈900°C for ≈5 min immediately prior to deposition (This step can be omitted by using a low temperature cleaning procedure [9]). A stacked CaF_2-BaF_2 buffer layer of ≈2000 Å thickness was grown first by molecular beam epitaxy (MBE) at substrate temperatures up to 700°C. Approx. 3 µm thick PbTe layers were then deposited at ≈400°C by a hot-wall-epitaxy (HWE) method in a separate system. Ternary $Pb_{1-x}Eu_xSe$ was grown by MBE on the fluoride covered substrates. Arrays of photovoltaic IR-sensors were fabricated using vacuum deposited blocking Pb-contacts on p-type layers with carrier concentrations in the low 10^{17} cm^{-3} range. Ohmic contacts were applied by Pt-sputtering. Fig. 1 shows the layout of the array. It consists of 66 elements arranged in a staggered way with active areas of 50 x 100 μm^2 of the individual sensors. Electrical isolation of the fan-out pattern to the Si-substrate is performed with evaporated polycrystalline BaF_2. Illumination is from the backside through the IR-transparent Si-substrate and CaF_2-BaF_2-layer. Due to the high permittivity of lead-chalcogenides, the width of the space charge region is only a few tenths of a µm. No surface passivation layer or antireflection coatings was applied for simplicity.

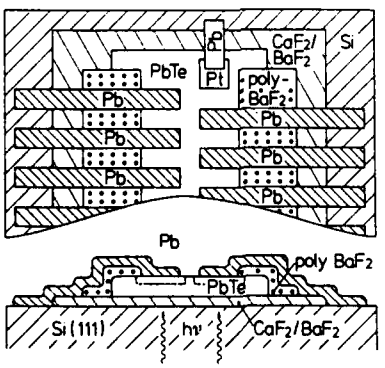


Fig. 1. Layout of linear photovoltaic IR-sensor arrays on Si-substrates with epitaxial lead-chalcogenide layers and step graded fluoride buffers. The active area below the Pb-blocking contact is indicated by a dashed line.

PHOTOVOLTAIC PbTe ON Si ARRAYS

Figs. 2 and 3 show the I-V-characteristics and spectral response of a typical PbTe sensor at 90K and 200K, respectively. At 90K, the differential resistance at zero bias is R_0 = 8 MΩ, which corresponds to a resistance-area product R_0A = 400 Ωcm^2. The spectral responses merely reflect a constant quantum efficiency up to the cut-off wavelengths but modulated due to interference effects. The cut-off wavelengths are about 5.7 µm and 4.6 µm at 90K and 200K.

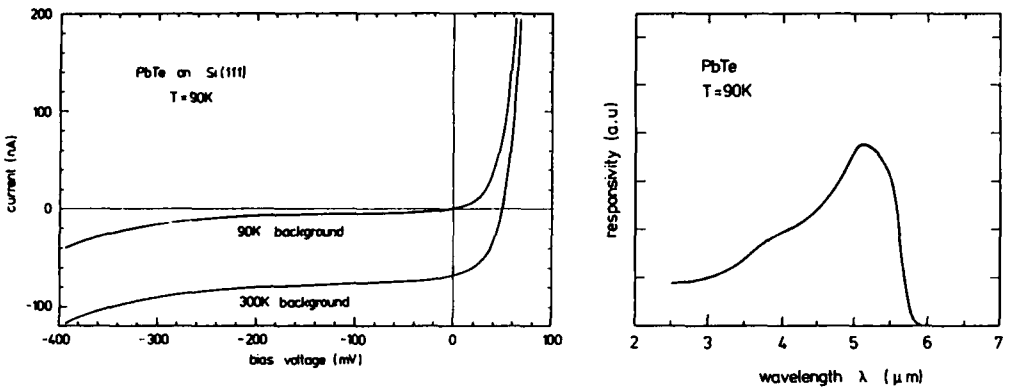


Fig. 2. I-V-characteristics and spectral response (in Amps/Watt scale) for a typical PbTe sensor on Si of the array shown in Fig. 1 at 90K.

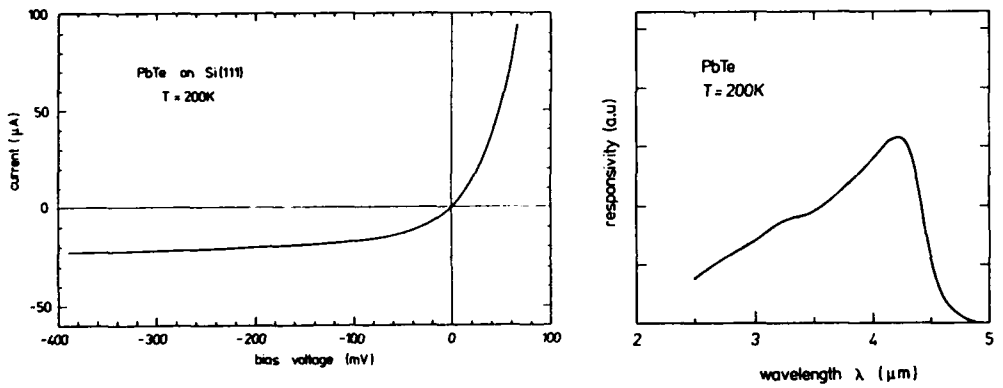


Fig. 3. I-V-characteristics and spectral response of the same sensor as in Fig. 2, but at 200K.

respectively, due to the positive temperature dependence of the bandwidth of lead-chalcogenides. Measured absolute quantum efficiencies are ≈70%, which means that nearly every photon which reaches the interior of the PbTe layer is registered.

In most arrays, all except 1 or 2 elements are properly working and have resistance-area products above 50 Ωcm^2 at 85K. Highest values are R_oA = 500 Ωcm^2, while the mean resistance area product of a whole array is typically 150 Ωcm^2 at 85K (the lowest temperature attainable in the cryostat used for the measurements). These values are up to 2 orders of magnitude above the BNL-limit for room temperature radiation, but still appreciably lower than the best values reported for photovoltaic PbTe sensors on bulk BaF_2-substrates [6] or state of the art IR-sensors with comparable cut-off wavelength and operating temperature [1], but which are of course not monolithically fabricated on Si-substrates. However, since most of the fabrication steps were performed with rather crude techniques, still much improvement seems possible.

The temperature dependence of the resistance-area product indicates a depletion limited noise current behaviour below 250K and a satisfactory agreement with the experimental data is obtained assuming a temperature independent Shockley-Read recombination lifetime of 2 ns. This lifetime is high enough to allow carrier diffusion over the few microns needed for collection of most of the photogenerated carriers. Above 250K, the R_oA products fall below 0.1 Ωcm^2 but come close to the theoretically calculated diffusion limit. Details on these measurement will be published elsewhere.

PHOTOVOLTAIC $Pb_{1-x}Eu_xSe$ ON Si ARRAYS

While $Pb_{1-x}Sn_xSe$ is suitable for sensors for wavelengths above 7 μm, shorter variable cut-off wavelengths are covered with $PbS_{1-x}Se_x$ and $Pb_{1-x}Eu_xSe$. A large number of investigations have been performed with $PbS_{1-x}Se_x$ Schottky barrier sensors on bulk BaF_2 substrates [6,7] but not with $Pb_{1-x}Eu_xSe$. We therefore investigated sensors of this material in some preliminary

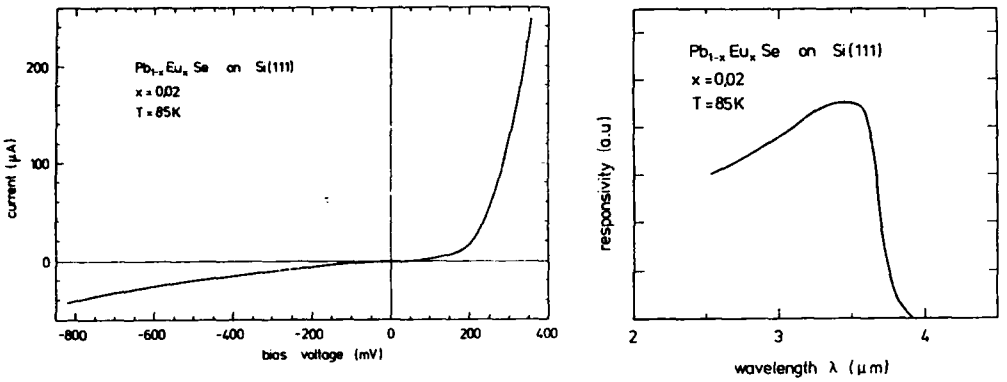


Fig. 4. I-V-characteristics and spectral response (in Amps/Watt scale) for a typical $Pb_{1-x}Eu_xSe$ photovoltaic sensor on Si at 85K.

experiments. We found that Pb indeed forms barriers on p-type layers, and that whole arrays on fluoride covered Si-substrates work consistently. Fig. 4 shows an I-V-characteristics of a typical $Pb_{0.98}Eu_{0.02}Se$ sensor element of an array with the same lay-out as shown in Fig. 1. True rectifying behaviour is observed. R_oA values are $\approx 2\,\Omega cm^2$, which is rather low for this temperature. However, this differential resistance is already attained at 200K and becomes independent of temperature below. The reasons for this saturation are not known at present. The cut-off wavelength is about 3.8 µm at 85K, in accordance with the expected value resulting from the chemical composition x=0.02 chosen. Shorter cut-off wavelength down to 2 µm are feasible by increasing the Eu-content [10].

CONCLUSIONS

We have demonstrated that photovoltaic lead-chalcogenide on Si IR-sensor arrays for the 3-5 µm range can be fabricated reproducibly and with sensitivities well above the background noise limit.

Growth temperatures are below ≈700°C for deposition of the buffer layer needed to maintain epitaxy and below ≈400°C for growth of the narrow gap semiconductor layer and the further fabrication steps. The processing appears compatible to active silicon substrates which contain signal processing electronics, thus opening the way to a heteroepitaxial, but fully monolithic approach of IR-focal plane arrays.

The 8-12 µm range can be covered with $Pb_{1-x}Sn_xSe$ and by using the same technique as we already demonstrated with single photovoltaic PbSnSe on Si sensors [2]. Shorter variable cut-off wavelengths between 2 µm and 7 µm are attained with $PbS_{1-x}Se_x$ or $Pb_{1-x}Eu_xSe$.

Acknowledgment:
The work is sponsored by the Swiss Defense Technology and Procurement Agency, the Swiss National Science Foundation, and the European Space Agency. The authors from F-IPM, Freiburg, also acknowledge the support from the German Federal Ministry for Research and Technology.

REFERENCES

[1] For reviews see: Proc. SPIE **443**, 1983; or Optical Engineering **26**, No. 3, March 1987.
[2] H. Zogg, P. Norton, Techn. Digest Int. Electron Devices Meeting IEDM, Washington D.C. Dec. 1985, p. 121.
[3] H. Zogg, Appl. Phys. Lett. **49**, 933, 1986.
[4] H. Zogg, W. Vogt, H. Melchior, Proc. SPIE **587**, 170, 1985.
[5] H. Zogg, S. Blunier, Appl. Phys. Lett. **49**, 1531, 1986.
[6] H. Holloway, Physics of Thin Films, Academic Press, G. Haas, M. H. Francombe ed., Vol. **11**, 1980, p. 105.
[7] R.B. Schoolar, J.D. Jensen, G.M. Black, Appl. Phys. Lett. **31**, 620, 1977.
[8] S. Blunier, H. Zogg, H. Weibel, Mat. Res. Soc. Symp. Proc. **116**, 1988, to be published.
[9] P.J. Grunthaner, F.J. Grunthaner, R.W. Fathauer, T.L. Lin, F.D. Schowengerdt, B. Pate, J.H. Mazur, Proc. 2nd Int. Symp. on Si-MBE, Honolulu, Oct. 1987, Electrochem. Soc. Proc. Vol. 88-8, p. 375.
[10] P. Norton, M. Tacke, Journal Cryst. Growth **81**, 405, 1987.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

SURFACE PLASMON POLARITON ENHANCED LIGHT EMISSION AND PHOTORESPONSE IN SCHOTTKY DIODES

A. KÖCK, W. BEINSTINGL and E. GORNIK

Institute for Experimental Physics, University of Innsbruck, A-6020 Innsbruck, Austria

Abstract – We have investigated the light emission from forward and reverse biased sinusoidally structured Ag/n–GaAs Schottky diodes. These Schottky junctions provide increased light emission due to the radiative decay of excited surface plasmon polaritons, resulting in drastically enhanced quantum efficiency. Conversely surface plasma modes excited at the surface of a semiconductor-metal-insulator structure provide selective coupling of light into the semiconductor with increased quantum efficiency. It is shown, that Schottky diodes can be used as photodetectors selective to polarization, frequency and angle of incident light.

1 – INTRODUCTION

For applications in numerous technical fields there is a great demand for selective emission and detection of visible light. Schottky barrier devices, such as Ag/n–GaAs junctions, can be used both as light emitters, when appropriately biased /1,2,3/, and as photodetectors, when exposed to light. If the surface is periodically structured, surface plasma modes (SPM) can be excited at the metal–air interface. The coupling condition is

$$k_L = \frac{\omega}{c} \sin \vartheta = k_{SPM} + nk_g$$

with k_L the wavevector of incident light, k_{SPM} the wavevector of SPM and $k_g = \frac{2\pi}{\Lambda}$ the grating vector; n is an integer. P–polarized light (magnetic field vector parallel to the grating) can excite only transverse magnetic modes (TMi, i = mode index), s–polarized light can excite only transverse electric modes (TEi). As the excitation of these modes depends strongly on wavelength, polarization and angle of incident light, the basic idea is to use these modes both for selective coupling of light into and out of the semiconductor. In emission experiments we confined to the TM_0 mode, in photoresponse to TM_0 and TE_0 modes.

2 – EXPERIMENTAL METHODS AND RESULTS

We present results from sinusoidally structured Ag/n–GaAs Schottky diodes providing both increased light emission and enhanced photoresponse due to the excitation of SPM. The sinusoidal gratings were fabricated holographically on photoresist and subsequently transferred to GaAs substrates by reactive ion etching. The Schottky junctions are formed by a Ag top layer (thickness 30 nm, 6 mm^2 area) evaporated on n–GaAs substrates. The geometry of the samples used is shown in the insert of Fig. 1a. The amplitude of the grating was H = 25 nm and the period Λ = 650 nm. The investigations of the emission spectra and the photoresponse measurements were performed at room temperature.

The emission intensity of a reverse biased Ag/n–GaAs Schottky diode (current 80 mA) as a function of wavelength is plotted in Fig. 1a. Curve (a) represents the emission spectrum of a flat junction at angle $\vartheta = 0^0$. The emission spectrum is broadband and completely unpolarized. Curve (b) shows the spectrum of s–polarized light emitted from a structured Schottky diode at angle $\vartheta = 0^0$. The s–polarized emission

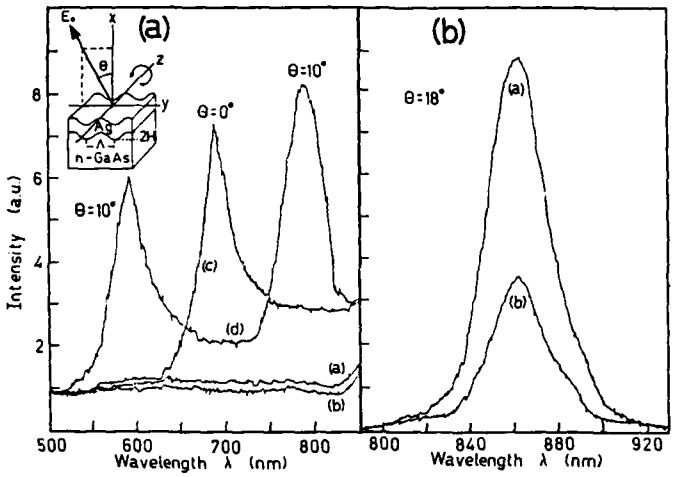


Fig. 1 a – Emission spectra of a.reverse biased Ag/n–GaAs junction. Curve (a) shows the spectrum of an unstructured junction. Curve(b) shows the s–polarized spectrum of a structured junction at $\vartheta = 0^0$. Curves (c) and (d) show p–polarized emission spectra at $\vartheta = 0^0$ and 10^0.

Fig. 1 b – Emission spectra of a forward biased structured junction at $\vartheta = 18^0$. Curve (a) shows the p–polarized, curve (b) the s–polarized emitted light.

spectrum of the structured sample (curve (b)) has the same characteristics as the emission spectrum of the unstructured sample (curve(a)). Curve (c) and (d) represent the p–polarized emission spectra at angles of $\vartheta = 0^0$ and $\vartheta = 10^0$, respectively.

The p–polarized emission maxima are due to the radiative decay of the TM_0 mode at the Ag–air interface. For an unambiguous interpretation we have calculated the dispersion relation of the TM_0 mode at a Ag–air interface and found good agreement with the experimental values.

Fig. 1b shows the electroluminescence spectra of the same Ag/n–GaAs Schottky diode when forward biased (current 100 mA). At an angle of 18^0 the wavelength of the TM_0 mode ($\lambda = 862$ nm) coincides with the energy of the bandgap of GaAs. Curve (a) represents the emission spectrum of p–polarized light, curve (b) of s–polarized light.

In addition to the known SPM excitation processes (hot electron mechanism /4/, conversion of slow–mode plasmons to fast–mode plasmons /5/), we found a third process responsible for the observed emission spectra. We interpret the light emission process for reverse and forward bias by the same physical mechanism: Photons generated in the GaAs substrate with emission angle larger than the critical angle of internal total reflection (16^0) can excite a TM_0 mode, which decays radiatively due to the presence of the grating. Therefore via excitation of the TM_0 mode more photons are coupled out of the semiconductor resulting in drastically enhanced quantum efficiency (QE). For a flat junction we obtained in reverse bias an external QE of $1 \cdot 10^{-7}$. For a corrugated junction the QE is 7 times larger. In forward bias excitation of the TM_0 mode resulted in enhancement of the QE by a factor 2.

In addition we have investigated the photoresponse of the same devices when either the TM_0 modes or the TE_0 modes are excited.The excitation of these modes can be detected by investigation of the reflectance and photoresponse spectra of the diodes. If a SPM is excited, a narrowband dip occurs in the reflectance spectrum, whereas the photoresponse shows a maximum /6,7/. This is due to the energy transfer of the incident light to SPM. The electromagnetic field of the SPM penetrates the metal film and generates electron–hole pairs in the semiconductor. These additional carriers enhance the photocurrent and therefore narrowband maxima are observed in the spectral response.

The TM_0 mode can be excited with p–polarized light at the metal–air interface of an uncoated Schottky diode. Fig. 2a shows the reflectivity (A) and the photoresponse (B) of a uncoated, structured Ag/n–GaAs junction as a function of wavelength at an angle of incidence $\vartheta = 10^0$ for p– and s–polarized light (solid and dashed line, respectively). The dips and maxima are due to the excitation of the TM_0 mode.

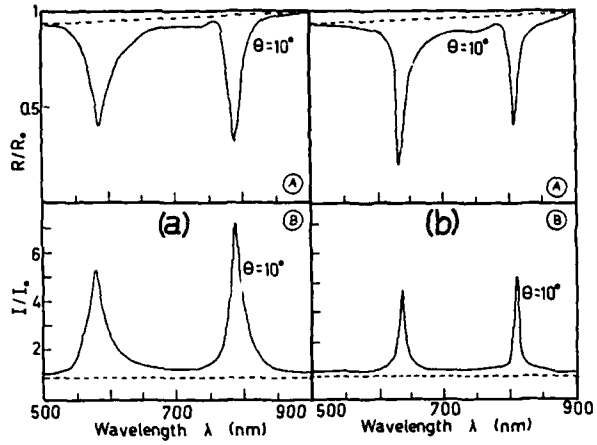


Fig. 2 a – Reflectivity and photoresponse spectra of an uncoated Ag/n–GaAs detector (TM_0 mode excited) at $\vartheta = 10^0$. P–polarized: solid line, s–polarized: dashed line.
Fig. 2 b – Reflectivity and photoresponse spectra of a coated (MgF_2, thickness 250 nm) Ag/n–GaAs detector (TE_0 mode excited) at $\vartheta = 10^0$. S–polarized: solid line, p–polarized: dashed line.

To excite the TE_0 mode with s–polarized light, the device has to be coated with a dielectric layer. Fig. 2b shows the reflectivity (A) and photoresponse (B) of a coated (i. e. MgF_2, thickness 250 nm), structured Ag/n–GaAs junction, when excited with s– and p–polarized light (solid and dashed line). The linewidth of the TE_0 mode has been found to be smaller than that of the TM_0 mode. The QE of the device, when the TM_0 modes are excited, is higher (14%) than in the case of TE_0 modes (11%). The best result found until now has been a QE of 31% at TM_0 resonance. The best signal to background ratio was 7:1 for TM_0 and 5:1 for TE_0 modes.

3 – APPLICATIONS

For technical applications the signal to background ratio of 7:1 is not sufficient for quantitative measurements. We have developed a method which improves the ratio to 50:1. The idea is to illuminate two detectors by the same light, but excite a SPM only on one of them. The two signals are subtracted from each other, and the remaining background is nearly zero. Two adjacent detectors were produced on the same substrate.

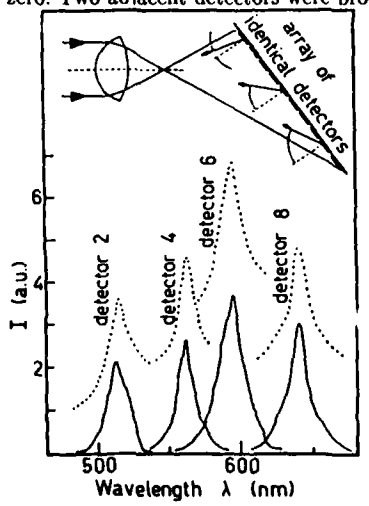


Fig. 3 – Spectrometer formed by an array of identical detector pairs. The dashed lines represent the photoresponse maxima due to TM_0 mode excitation of several detectors. The full lines show the photoresponse peaks after subtraction of the background signals. The improvement of the signal to background ratio is evident.

one of them was coated with MgF_2. The TM_0 resonance of the coated detector is shifted to longer wavelengths and thus the two detectors cannot produce a TM_0 signal at the same wavelength.

An application of these detectors is shown in Fig. 3. Ten identical pairs of detectors (one with grating and one without grating) were produced on one substrate by photolitographic techniques. Divergent light is incident at different angles on each pair. So on each detector pair the SPM resonance occurs at different wavelength and the array can serve as a simple, integrated spectrometer.

Small scale integration of the detectors can be realized by using standard technology. Various applications in several technical fields are possible, such as color, pattern and image recognition, optical data transfer, and spectroscopy of short light pulses.

Concerning selective light emission, further work on combining SPM excitation with light emitting devices is under progress. The QE of LED heterostructure devices might be enhanced by the use of an appropriate sinusoidal grating. Application to laser heterostructure devices seems to become posssible.

REFERENCES

/1/ B. Bayraktaroglu and H. Hartnagel, Electron. Lett. 14, 470 (1978)
/2/ A. A. Gutkin and V. E. Sedov. Sov. Phys. Semicond. 15, 348 (1981)
/3/ H. Adachi and H. Hartnagel, J. Vac. Sci. Technol. 19, 427 (1981)
/4/ J. R. Kirtley, T. N. Theis and J. C. Tsang, Appl. Phys. Lett. 37, 435 (1980)
/5/ S. Ushioda, J. E. Rutledge and R. M. Pierce, Phys. Rev. Lett. 54, 224 (1985)
/6/ K. Berthold, W. Beinstingl, R. Berger and E. Gornik, Appl. Phys. Lett. 48, 526 (1986)
/7/ K. Berthold, W. Beinstingl and E. Gornik, Opt. Lett. 12, 69 (1987)

SESSION 6

Session 6IP Room A : Invited papers

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

THE HOT-ELECTRON PROBLEM IN SUBMICRON MOSFET

W. HÄNSCH, M. ORLOWSKI and W. WEBER

Siemens AG, Corporate Research and Development, Microelectronics, Otto-Hahn-Ring 6, D-8000 München, F.R.G.

ABSTRACT - A review of the hot electron problem in MOSFET is given. This includes: Key experimental features, the problem of modelling hot carrier transport in Si and SiO_2 after injection into the oxide, and an evaluation of technological measures to obtain hot carrier resistant structures.

I) INTRODUCTION:

Long-term stability is a key to the successful design of a MOSFET in the submicron regime. An important factor in achieving this aim is a quantitative analysis of hot electron effects. Hot carrier induced damage was known long before the submicron region was entered. Here we deal with the always present high energetic electrons (holes) that can surmount the Si/SiO_2 barrier of 3.2 (4.8) eV and then get trapped in the oxide. These charges now interfere with the controlling gate and will produce threshold shifts, transconduction changes, saturation current modifications etc. It is clear that all this can only be tolerated to a certain degree if the device is to operate properly in its circuit environment.

Secondly, we have the influence of the elevated fields on the carrier transport and the carrier injection into the oxide. It is a well known fact that the conventional drift diffusion approximation is only valid for small current levels, e.g. close to the thermal equilibrium.

The link between the latter and the former is that Poisson's, continuity, and current equations provide carrier distributions and electrical fields which are used to calculate the carrier injection into the oxide.

To prevent the undesired effects of hot-electron damage, a careful analysis of the experimental material is indispensable in an early state of device development. Measurement and simulation, as an analytic tool, have to go hand in hand.

There is quite a variety of effects caused by hot-electron (carrier) injection. Trapped electrons and/or holes in the oxide (Q_{ox}), fast surface states at the Si/SiO_2 interface (N_{it}), series resistance effects in LDD, mobility degradation due to oxide charges, time evolution of damage, etc. To monitor all these phenomena there are only four terminal currents measurable in a MOSFET: Source, drain, bulk, and gate current. Because of the high degree of complexity it is absolutely essential to have sufficient experimental data.

In the next chapter this review we will address the key experimental features for a hot-carrier analysis and their evaluation. In Chapter III a detailed discussion will be presented on modelling hot-electron effects on the basis of the drift-diffusion approximation and its extensions. In Chapter IV we will discuss how simulation is used to prevent hot-electron damage in submicron MOSFET structures.

II) THE HOT-CARRIER EFFECT FROM THE EXPERIMENTAL STANDPOINT

In the real operation environment, devices can change their properties to such a high degree that the functioning of the whole circuit may be affected after a certain time of operation. An example of this is shown in Fig. 1. The current-voltage behaviour and as a consequence the current-time path are changed after a certain time of operation.

In order to compare different technologies with respect to their hot-carrier stability the question of the right parameter arises. One may propose the threshold voltage shift at $V_G = V_D = 5V$ to qualify a technology as was done some time ago, since this is a circuit-relevant parameter. However, technological developments require a fast evaluation of the stability so that sensitive parameters were sought at the expense of circuit relevance, since the aforementioned one is not sufficiently sensitive. Empirically it was found that the characterization of the damage at low drain voltage, in the linear region, can provide the desired fast evaluation of device stability (Later on it turned out that for the inverter operation these new linear region parameters are very relevant and in fact proportional to the frequency shift of an inverter chain /1/).

Now the question arises whether we can, having found sensitive characterization parameters, immediately proceed to heuristic optimizations of technology parameters? Is there a conclusive way to obtain optimal drain profiles and oxide growth conditions to design a device with a high hot-carrier stability without a detailed physical understanding?

The issue has been approached in most labs in a more or less empirical way. However, a physical insight into degradation phenomena can provide guidance for optimization, accelerating it and making it more effective. This approach will be addressed in Chapters III and IV. Furthermore, the danger is too great that without understanding the physical background, false conclusions about stability are drawn for the operating conditions which can never be used in stress test experiments because of the slowing down of the hot-carrier effect at low operating voltages . As a consequence we conclude that a detailed understanding of the hot-carrier effect is necessary.
A second question arises: how far are we in understanding hot-carrier effects? Have we understood the basic features of the hot-carrier effect or are we still right at the beginning? In the following we try to obtain an answer to this question on the basis of a few experimental examples. These examples led to understanding but new open questions arose.

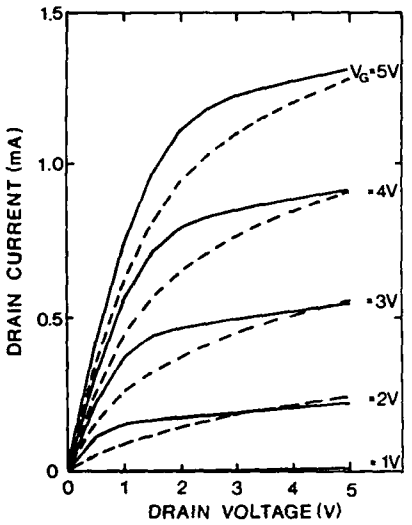


Fig. 1: $I_D(V_G,V_D)$ - characteristics before (full lines) and after (broken lines) a hot-carrier stress. A conventional nitride passivated 2.3um device with t_{ox} = 42nm is stressed with t_{stress} = 14h, V_D = 8V, and V_G = 3V.

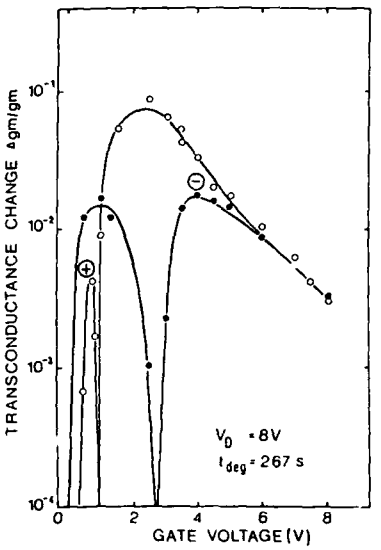


Fig. 2: Degradation vs. stress gate voltage is shown for devices as in Fig. 1 (open symbols) and devices from the same process but without a final nitride passivation (full symbols).

1) A few years ago a number of puzzling degradation results for different conventional n-MOSFETS led to a degradation model /2/ in which electrons and holes play an important part in explaining the resulting degradation. Holes are injected and trapped inside the oxide close to the interface and fixed positive charge is formed (visible by a positive $\Delta g_m/g_m$ in Fig. 2). Injected electrons are trapped with a large cross-section (due to Coulomb interaction) at the trapped holes and interface states are formed. These are evident through their negative charge in the negative $\Delta g_m/g_m$ branch in Fig. 2. A simple theory using rate equations was set up giving a reasonable fit to the experimental data with a proper choice of the two parameters of hole trap density and electron capture cross section. This model not only worked for different conventional MOSFETs (Fig. 2), but also for modern devices with reduced drain field such as LDD and DID.
Recently, however, R. Bellens et al. /3/ performed charge pumping measurements on devices under hole conditions (low V_G) and discovered the direct formation of interface states at the hole-injection condition. They claim that the subsequent electron injection merely causes a charge compensation. In order to decide finally on these findings, more data, especially space and energy resolved charge pumping measurements and 2D-simulations, are needed. Furthermore, essential features of the previous model, like the importance of the hole injections, are not questioned. - Nevertheless we can state that a precise degradation mechanism still remains unsettled.

2) A few years ago a relation ship between the lifetime and the substrate current of a device was found by Takeda and Suzuki /4/ and explained by C. Hu et al /5/. Although derived from an electron injection model, the result is also valid for a hole-injection model /6/. In Fig. 3 typical results are shown for different technologies, both LDD and conventional ones with a slope of 2.9 on the double-logarithmic scale. Agreement with the theory is good.

However, data have recently been measured at LDD devices which show a much steeper slope around 5 which cannot be explained by this theory (Fig. 3). Neither a technological reason is known for this difference nor a satisfactory physical explanation available for this phenomenon.

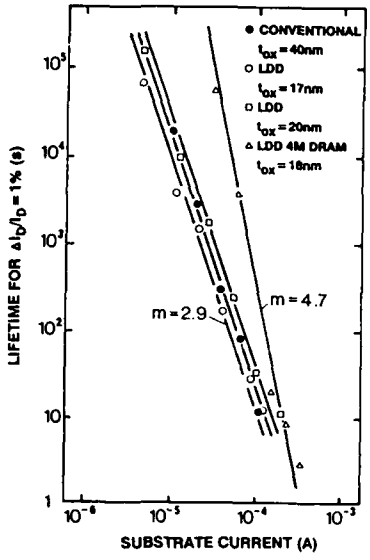


Fig. 3: Lifetime vs. substrate current for different technologies. One technology shows a steeper slope than the others.

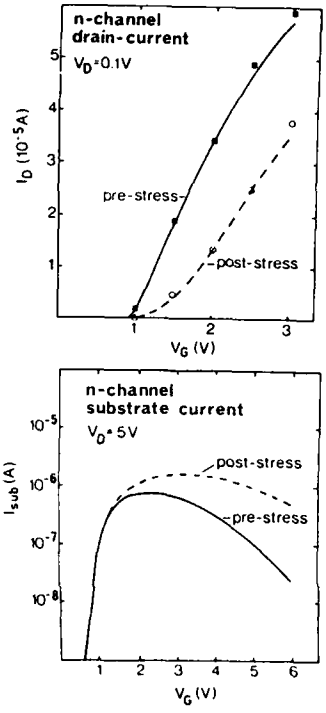


Fig. 4: Drain- and substrate currents of an n-MOSFET of the process in Fig. 1 before and after hot-carrier stress. The stress conditions were $V_G = 3V$, $V_D = 8V$ and $t_{stress} = 5x10^4$s.

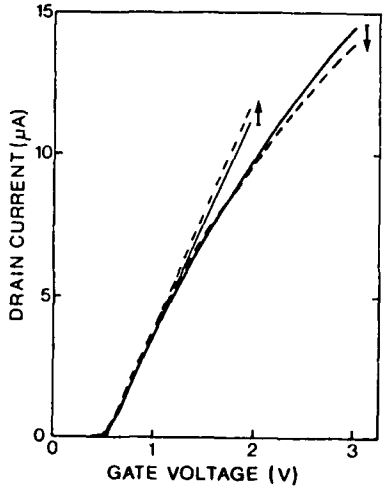


Fig. 5: $I_D(V_G, V_D = 0.1V)$ for a LDD-n-MOSFET with $t_{ox} = 25nm$, $l_{spacer} = 0.2um$ and $l_{gate} = 1.0um$ before (full lines) and after (broken lines) a stress with $V_G = 0.6V$, $V_D = 8V$ and a stress time of 12h. To clarify the result a straight line is drawn through the steepest parts of the $I_D(V_G)$-curve.

3) It can be seen in Fig. 1 how the current-voltage characteristics change after hot-carrier stress. A discussion has been going on for years on the question of the phenomena causing these changes. Are they due to fixed negative charges inside the oxide, to acceptor-type interface states acting through their potential, or to acceptor- or donor-type interface states acting through their mobility effect? It

is not easy to decide this question on the basis of electrical measurements alone, and surely the results will be different for different technologies. Recently Schwerin et al. /7/ have performed careful comparisons of 2D simulations and electrical data including the effects of stress-generated surface charges. This work revealed the complicated interplay of different types of oxide charges on drain and substrate currents. The essential result is this: without complicated simulations, statements on N_{it} and Q_{ox} can be made just by looking at I_D (V_G, V_D) and I_{sub} (V_G, V_D) characteristics together. An example of this is shown in Fig. 4. An n-channel MOSFET shows a decreasing drain current but an increasing substrate current. This cannot be explained by a mobility effect since this would not lead to an increased substrate current. Fixed negative charges would essentially lead to a parallel shift of the I_D (V_G) characteristics. Only acceptor-type interface states can explain changes of the curves in Fig. 4 satisfactorily.
Meanwhile, the experiment has brought new and puzzling results. Can we explain the behaviour shown in Fig. 5 with the above method? Here the I_D (V_G) curves before and after hot-hole stress intersect, possibly due to a combination of mobility and potential effects.

In the preceeding paragraph, several results and new questions have been discussed, showing that our understanding of the hot-carrier effect is more or less at an early state and fragmentary. These questions can become important for the way the device stability under operating conditions is extrapolated or calculated, especially for the coming generations of devices. There the safety margins will be further decreased as the technological possibilities are most likely exhausted.

III) HOT-ELECTRON MODELLING

The ultimate aim of hot-electron modelling is to reduce the hot carrier damage by choosing an optimal device structure. As a very first step to accomplish this task, models enabling the experimental data to be verified must be developed on the basis of the relevant physics. The key idea is to incorporate simple, but physically appropriate models into a realistic simulation program that allows to consider technologically relevant structures. This requires that source/drain and channel doping profiles to be accurately known. Process simulation, which itself is supported by experimental techniques to verify the doping profiles, can provide this data /8/. Another technological aspect concerns the quality of the oxide to be taken into account, in the gate and spacer regions. It is generally believed that near the gate edge the oxide quality decreases due to mechanical stress and/or implantation damage. The degradation properties certainly depend sensitively on the trap distribution in the oxide.
The problem of hot electron-modelling splits into two parts: (a) Calculation of the internal distributions in the device in the presence of very high fields; and (b) calculation of the parasitic effects such as oxide injection and charge transport in the oxide, that eventually lead to trapped charges which interfere with the controlling gate. The simulation can be used to analyze degraded samples to get information about the amount and location of the trapped charges. Or the degradation process itself, can be monitored.
We will initially discuss a modified description of carrier transport which is appropriate for submicron structures:

A) EXTENDED DRIFT-DIFFUSION APPROXIMATION

In an ideal device the field and carrier distributions are the solutions of the following semiconductor equations with appropriate boundary conditions.

$$\mathrm{div}(\epsilon \underline{E}) = -q(n-p+C) \qquad \text{(III.1)}$$

$$\mathrm{div}\underline{j}_n - \frac{q}{c}\frac{\partial n}{\partial t} = G-R \qquad \text{(III.2a)}$$

$$\mathrm{div}\underline{j}_p + \frac{q}{c}\frac{\partial p}{\partial t} = R-G \qquad \text{(III.2b)}$$

$$\underline{j}_n = q\mu_n n\underline{E} + qD_n \mathrm{grad}n \qquad \text{(III.3a)}$$

$$\underline{j}_p = q\mu_p p\underline{E} - qD_p \mathrm{grad}p \qquad \text{(III.3b)}$$

Poisson's equation Eq. (III.1) and the continuity equations Eq. (III.2a/b) are justified by classical electrodynamics and are valid as long as fluctuations in the system are negligible. The current equations Eq. (III3a/b) are based on the general assumption valid in macroscopic systems that the particle current has a field driven drift term and a concentration gradient driven diffusion term/8/. A justification of the current equations Eq. (III3a/b) can be given by the Boltzmann equation if only small deviations from thermal equilibrium are considered. In this case the momentum relaxation time τ_m is not field dependent and therefore the mobility $\mu_{n/p}$ is a constant and Einstein's relationship holds/9/.

$$qD_{n/p} = kT\mu_{n/p} \qquad \text{(III.4)}$$

The quasi-Fermi potential $\varphi_{n/p}$ is now a very useful concept for transforming the current equations.

$$j_n = -q\mu_n n \operatorname{grad}\varphi_n \qquad \text{(III.5a)}$$

$$j_p = q\mu_p p \operatorname{grad}\varphi_p \qquad \text{(III.5b)}$$

It was, however, realized long ago that the linear response regime is not sufficient to obtain satisfactory results for MOSFET. Velocity saturation had to be included to explain the experimental data. This was done by replacing the mobility in Eq. (III.5) by a field dependent function $\mu_{n/p}(F_{n/p})$. The choice of the saturation force $F_{n/p}$ is not unique and has been disputed in the literature/10/. Of several possibilities $F_{n/p} = |\operatorname{grad}\varphi_{n/p}|$ is the most often used. Replacing μ by a field-dependent mobility already takes a situation in which the electrons (holes) are no longer in equilibrium with the lattice partly into account . The physical reason for this is the onset of optical phonon scattering of carriers with elevated energies. Optical phonons carry away some of the energy that the carriers gain in the electrical field. For very high fields this scattering process is no longer sufficient to maintain thermal equilibrium. As a consequence the carriers will have an elevated average energy compared with that of the phonons and their average velocity will saturate because optical phonons randomize the motion of the carriers very effectively. Therefore if phonon scattering in high fields is properly taken into account an increased carrier temperature has to be included.
Assuming a drifted Maxwellian as an Ansatz for the microscopic distribution function f_n in $\underline{x}$- and $\underline{k}$-space,

$$f(\underline{x},\underline{k}) = \exp\left(-\frac{\varepsilon(\underline{k})+\underline{k}\cdot\underline{v}_d(\underline{x})}{kT(\underline{x})}\right) \qquad \text{(III.6)}$$

in the Boltzmann equation Blotekjar/11/ obtained an alternative set of transport equations by using the moment method. Blotekjars equations quoted for one carrier type, the stationary state, and an isotropic system are

$$\underline{j} = q\mu n\underline{E} + \frac{2}{3}\mu \operatorname{grad}\left(\epsilon - n\frac{1}{2}mv_d^2\right) \qquad \text{(III.7)}$$

$$q\underline{v}_\epsilon = -\frac{5}{3}\underline{j}\frac{\epsilon - nmv_d^2/5}{n} \qquad \text{(III.8)}$$

$$\frac{\epsilon-\epsilon_o}{\tau_\epsilon} = \underline{j}\cdot\underline{E} - \operatorname{div}\underline{v}_\epsilon \qquad \text{(III.9)}$$

Blotekjar does not consider the determination of the momentum (τ_m) and energy (τ_ϵ) relaxation times. These have to be suitably chosen and added to complete the formulation. This and the need to add the missing heat current in Eq. (III.8), are the major drawbacks of this formulation. By now there is clear evidence that a drifted Maxwellian is a poor approximation of the distribution function/12/. Before we discuss the physical contents of Eqs. (III.7 - III.9) we will present the results obtained by one of the authors. Because the drifted Maxwellian is a poor approximation, Hänsch and Miura-Mattausch/13/ proceeded differently. Utilizing the moment method for the Boltzmann equation as well their approach focuses on the self-consistency in the four lowest moments of the distribution functions, which are: n, $\underline{j}$, ϵ , $\underline{v}_\epsilon$. This is achieved by expressing the distribution function as:

$$f(\underline{x},\underline{k}) = \alpha(\underline{k})n(\underline{x}) + \underline{\beta}(\underline{k})\underline{j}(\underline{x}) + \gamma(\underline{k})\epsilon(\underline{x}) + \underline{\delta}(\underline{k})\underline{v}_\epsilon(\underline{x}) \qquad \text{(III.10)}$$

They derived an alternative set of equations to Eqs. (III.7 - III.9) that reads:

$$\underline{j} = q\mu n\underline{E} + \frac{2}{3}\mu \operatorname{grad}(\epsilon) \qquad \text{(III.11)}$$

$$q\underline{v}_\epsilon = -\frac{5}{3}\underline{j}\frac{\epsilon}{n} - \frac{10}{9}\mu\epsilon \operatorname{grad}\left(\frac{\epsilon}{n}\right) \qquad \text{(III.12)}$$

$$\frac{\epsilon-\epsilon_o}{\tau_\epsilon} = \underline{j}\cdot\underline{E} - \operatorname{div}\underline{v}_\epsilon \qquad \text{(III.13)}$$

In addition they can also provide the mobility μ and energy relaxation time τ_ϵ in a self-consistent way:

$$\frac{1}{\tau_m} = A + B\frac{\underline{j}\cdot\underline{v}_\epsilon}{j^2} \qquad \text{(III.14)}$$

$$\frac{1}{\tau_\epsilon} = \frac{1}{\tau_{\epsilon^0}} = \text{const.} \qquad \text{(III.15)}$$

Here A, B and $\tau_\epsilon{}^0$ no longer depend on the active variables n, $\underline{j}$, ϵ and $\underline{v}_\epsilon$. The heat current, the second term on the right side of Eq. (III.12), does not vanish. Although its actual form depends on the truncation scheme of the moment method, it is constructed such that it is exact for small disturbances from equilibrium.

Eqs. (III.7 - III.9) and Eqs. (III.11 - III.15) constitute possible extensions of the classical drift-diffusion Eqs. (III.3a/b). Both contain a finite energy relaxation time τ_ϵ, which is the characteristic time for the energy transfer from the carrier system to the lattice. In the classical drift- diffusion approximation the carriers transfer their excess energy instantaneously to the lattice, which means $\tau_\epsilon = 0$. Therefore the carrier system remains in thermal equilibrium with the lattice. However, in real systems momentum changes occur more frequently than energy losses, the system will heat up. Comparing the current Eq. (III.3) with Eq. (III.7) or Eq. (III.11) we notice that the diffusion current is now driven by the energy gradient. The current Eq. (III.7) agrees with Eq. (III.11) if the kinetic energy $n\frac{1}{2} m v_d^2$ is small compared to the total energy ϵ. This is the case if the variation of the electric field over a distance $v_d \tau_\epsilon < 100$ A is moderate, which is anyway a limit of the classical field concept. Adding in Eq. (III.8) a suitable heat current and neglecting $n\frac{1}{2} m v_d^2$ again Eq. (III.8) coincides with Eq. (III.12) for the energy current v_ϵ. More essential differences in their solution are attributed to the models utilized for τ_m or μ and τ_ϵ. Whereas in Eqs. (III.11 - III.15) they are included self-consistently, Eqs. (III.7 - III.9) demand an input from other sources.

Before we discuss a possible influence on the device performance we would like to point out a possible simplification of Eqs. (III.11 - III.15) which turns out to be very convenient for practical purposes. At first we generalize the concept of the quasi-Fermi-potential, which no longer constitutes the driving potential of the current. Instead we introduce a driving force $\underline{F}_n$ for the current $\underline{j}$:

$$\underline{F}_n = \underline{E} + \frac{1}{n}\text{grad}(u_{T,n}n) \qquad \text{(III.16)}$$

Here we use $\epsilon = q\frac{3}{2}u_{T,n}n$

$$\underline{j}_n = q\mu_n n\underline{F}_n \qquad \text{(III.17)}$$

Since the classical drift-diffusion approximation is justified as an expansion in φ_n we will now expand Eqs. (III.13) and (III.14) in $\underline{F}_n$.

$$u_{T,n} = u_o + \frac{2}{3}\tau_{\epsilon,n}\mu_n(F_n)F^2(1+\theta) \qquad \text{(III.18)}$$

$$\mu_n = \mu_{o,n}\{1-\eta[u_o - u_{T,n}(F_n)(1-\Omega)]\} \qquad \text{(III.19)}$$

As shown in /15/ θ and Ω are negligible small contributions almost everywhere in the limit F_n -----> oo. Therefore Eqs. (III.18) and (III.19) together with Eq. (III.16) constitute a singular perturbed system which is solved to a very good approximation by the reduced system. This is obtained by neglecting θ and Ω. Rewriting the reduced system we finally obtain:

$$\underline{j}_n = q\mu_n n\underline{F}_n \qquad \text{(III.20)}$$

$$u_{T,n} = u_o + \frac{2}{3}\tau_\epsilon v_{sat,n}^2\left(\frac{1}{\mu_n} - \frac{1}{\mu_{o,n}}\right) \qquad \text{(III.21)}$$

$$\mu_n = \frac{2\mu_{o,n}}{1+(1+(2\mu_{o,n}F_n/v_{sat,n})^2)^{1/2}} \qquad \text{(III.22)}$$

Eqs. (III.20 - III.22) represent a simplified version of Eqs. (III.11 - III.15).

All three alternatives for calculating the effect of an elevated carrier temperature on the carrier distribution and the electric field have been realized in a numerical code. We thus find Eqs. (III.7 - III.9) with variations realized in /16,17/, Eqs. (III.11 - III.13) together with a simplified mobility model (which corresponds to the approximation which leads to Eqs. (III.20 - III.22)) in /18/, and finally Eqs. (III.20 - III.22) are used by Hänsch and Selberherr/19/.

All these approaches have an enhanced diffusion constant (coefficient of grad n) in common. This leads to an increased diffusion current which causes the carriers to be pushed further away from the surface and therefore reduces the surface concentration: whenever $u_{T,n} << u_0$. On the other hand, the sharp drop of the carrier density in the pinch-off region is, including energy balance, less pronounced. However, for non-exotic MOSFET structures hot-electron effects will only little affect the peak electric field because it is located near pinch-off where the free carrier density is low compared to the doping. Therefore the redistribution of carriers does not affect the solution of Poisson's equation. This modified carrier distribution will, however, influence the generation rate due to impact ionisation and also the oxide injection of carriers. This is because both depend on the relative position of the maximum current with respect to the maximum electric field (see Chapter IV). The influence on the I-V characteristic is small for a channel length L > 0.5 μm, but depends on the mobility model, especially on the modelling of the gate field reduced mobility.

These remarks will close the overview on possible and feasible extensions of the classical drift-diffusion approach.

B) THE PARASITIC EFFECTS OF OXIDE INJECTION AND CHARGE TRAPPING:

We are now dealing with the carriers in the high energy tail and not with an average of the distribution function. If the energy range considered is far above the average energy it can be modelled independently of the lower moments of the distribution function discussed in the previous section. The problem is to estimate the number of carriers that can overcome the Si/SiO_2 barrier and to calculate the distribution of electrically active trapped charge. This charge can either be present as fixed negative or positive charge which is located away from the interface and only has a very weak coupling to the Si (slow states Q_{ox}) or is located in the vicinity of the interface and has a strong coupling to the Si (fast interface states N_{it}). The charge on the N_{it} is related to the position of the quasi-Fermi-level $\varphi_{n/p}$ at the interface. The origin of these fast interface states is still heavily disputed. There is some evidence that a hole trapping with a successive electron capture is a possible mechanism/2/. However, a model consistent with the experimental findings is required if the degradation process is to be investigated/3/.

The second major problem is the calculation of the number of high energetic carriers in Si. The energetic range we are interested in is 1.2 eV (~50 kT) and higher. The average energy of the carriers that we can calculate with one of the approaches outlined above is of the order of 10 kT, at room temperature. This permits an independent modelling of the high energy tail of the distribution function. These carriers are experimentally accessible only indirectly by substrate and gate current measurements. Both currents test a different energy range. The threshold energy for impact ionisation is between 1.2 eV...1.8 eV. For oxide injection an energy of 3.2 eV (4.8 eV) for electrons (holes) is required. In contrast to the previous problem an explicit form of the distribution function in terms of space and energy has to be known.

The lucky electron model is based on the Shockley model/20/ in which only those electrons are considered that follow the electric field lines until they reach the required energy without experiencing a collision. This ballistic approach implies that scattering is weak. This is not confirmed by Monte Carlo calculations/21/. In contrast, it is shown that once the particles exceed the energy of an optical phonon ω_{opt}~ 60 meV scattering is very effective. Particles with $\varepsilon >> \omega_{opt}$ are distributed very isotropically. The physical reason for this is that strong optical phonon scattering randomizes the motion of the carriers and does not allow a path to be followed along the field line. An analytical solution of the Boltzmann equation in the limit $\varepsilon >> \omega_{opt}$ in a constant electrical field is available/22/ and can serve as a foundation for further modelling efforts/23/. Ridley/24/ established the lucky-drift-mode model which serves as a bridge between the pure ballistic approach and the statistical approach by Keldysh. The principal ideas of how to model the high energetic carrier distribution are developed for constant electric fields. The MOSFET provides a very inhomogeneous field profile and therefore models have to be developed that take account of the fact that the field varies over the ionization length $d_{ion} = \varepsilon_i / q\,E$ /18,23,25/.

A suitable model for the high energetic carrier distribution should be able to explain substrate and gate current. The substrate current is easily calculated by introducing a generation rate due to impact ionisation on the right hand side of the continuity equation. The situation is not so simple for the gate current. Once the carriers are injected into the SiO_2 they will not necessarily end up at the gate electrode to be detected as gate current. If for instance $V_D >> V_G > V_{th}$ there is a strong electron injection after the pinch-off point in an n-channel device. However, the vast majority of these electrons will not reach the gate electrode because the electric field opposes their motion. Nevertheless they can still be very effective in filling possible traps in the oxide near the interface. A rigorous approach to calculating the gate current can only be to consider the transport of the injected carriers (electrons and/or holes) in the gate oxide. This can still be done by using a modified drift-diffusion approach because a stationary state of carriers and field exists for

the oxide thickness used in submicron MOSFET (t_{ox} > 8 nm). It has to be taken into account, however, that the injected carriers are far away from a stationary state with the gate field. They need a distance of 3 nm to relax to the oxide field/26/.
Once the transport problem in the oxide is solved the distribution of carriers in the oxide region is known. This is the precondition for studying the development damage caused by different mechanisms of trapping and detrapping of charge in the oxide. At this point, enter certain assumptions of the trap distribution in the oxide (oxide quality, material properties) and models how this traps are active. A very important feature here is the time development of a static stress experiment. Several attempts to study this can be found in the literature/2,27/. However, none of these takes into account the feedback of the trapped charges into the electric field in Si, which in turn will modify the number of injected carriers/26/. This is especially important if high-voltage, short-time stress is compared with low-voltage, long-time stress.
Another feature that can be studied if the correct distribution of carriers in the oxide is known, is the drift of carriers into regions of low injection but high trap density. An example is given by trapped charges above the n^--region in a LDD device. These are responsible for a series resistance degradation.

IV) OPTIMIZATION OF SOURCE/DRAIN STRUCTURES

After having considered the experimental degradation data and theoretical aspects of hot electron modelling we turn in this section to the question of how an efficient design of a refined source/drain structure can improve the device lifetime. Most concepts aimed at improving the device reliability have focused first on the reduction of the lateral electric field strength by making the p-channel to n^+-drain transition less abrupt. These concepts are represented by phosphorus, and arsenic-phosphorus source/drain junctions /28/, and most prominently and successfully by the lightly doped drain (LDD) structure /29/. The LDD region is introduced between the p-channel and n^+-drain to spread the high electric field at the drain-pinch-off region. This measure diminishes the impact ionization and consequently reduces both the substrate current and hot carrier injection into the gate and sidewall spacer oxide.
However, a number of puzzling effects have recently been identified in LDD structures and explained by Orlowski et al /30,31/ which render device optimization more difficult. The new effects can be attributed to additional, LDD-specific, peaks of the lateral field in addition to the well-known conventional peak at the lateral drain junction.

The LDD peaks are caused by a sudden variation in vertical band bending, at the silicon surface due to the rapid decrease of the gate field strength beyond the gate edges. On the drain side the situation is aggravated by the fact that for small gate-drain overlaps (depending on LDD concentration) the conventional and LDD field peak are no longer separated but overlap, and as result of this enhance one another considerably. This enhancement of both peaks is detrimental to the device reliability for two reasons: i) the enhanced lateral fields are located just at the surface where they heat up the carriers, thus facilitating injection into the oxide; ii) the enhanced field peak is located at least partly under the spacer oxide, which is probably of lower quality than the gate oxide and thus displays an increased cross section to capture and trap the injected carriers. The oxide damage is locally distributed above the n^--region. The relevant surface states and/or fixed charges deplete the surface of the n^--region and cause early degradation of the transconductance and the drain current. Note that, assuming the same oxide damage distribution the degradation mentioned above will be larger for a lower LDD dose than for a large one. The reason for this is that a n^--region with lower concentration can be more easily depleted than a high concentration region.

We like to point out that the aforementioned overlap and enhancement of the field peaks can already occur at such high LDD doses as 4E13 cm^{-2} and standard drive-ins at 900 °C for 40 - 150 minutes. This is due to high surface channel doping $N_{chan} \sim 10^{17} cm^{-3}$ for T_{ox} = 15 - 20 nm and V_{Th} = 0.7 - 0.9 V leading to strong compensation effects with the source/drain structures and partly due to the additional sidewall spacer for the LDD implantation, a necessary modification of the conventional LDD geometry (see discussion below).
For a conventional LDD structure the only way out of this predicament is to increase the LDD dose. This action enhances the overall field distribution, except at the gate edge where the field peaks are separated and mutual enhancement avoided. The simulation predicts higher substrate currents for the higher LDD dose, but smaller electric fields at the surface beneath the gate edge and therefore a longer device lifetime - in perfect agreement with the experiment.
With a modified LDD structure in which the gate electrode overlaps the entire LDD region (known as inside LDD realized in structures such as ITLDD/32/ and GOLD /33/) an entirely opposite procedure would be adopted: instead of increasing the LDD dose, the LDD dose would be lowered. The reason is simple: since the gate overlaps the entire n^--region, there is no gate edge effect and consequently no LDD field peaks and therefore no enhancement of the conventional peak. On the other hand the conventional field peak will decrease with lower LDD dose, leading to a smaller substrate current and smaller degradation at the same time. This has been also confirmed experimentally /32,33/.
Another important issue within the context of conventional LDD structures (but equally important for refined LDD structures) is the sidewall spacer for LDD implantation, or LDD spacer for short. In contrast to the conventional sidewall spacer for n^+-implantation,

which is a CVD oxide, the LDD offset spacer is due to the reoxidation of the gate. The purpose of the LDD-spacer is to remove the LDD implantation damage in the oxide from the gate edges. It seems that the implantation damage leads to a greater degradation. MOSFET transistors have been fabricated with and without the LDD spacer under otherwise identical conditions. The stress experiment shows that the variant without the LDD-spacer degrades more strongly than the variant with the LDD spacer, in conflict with the theoretical prediction /30/. On the other hand, all differences in device lifetime for different LDD structures among transistors with and without LDD- spacer were in perfect agreement with theoretical considerations. It was therefore concluded that the implantation generates additional traps in the oxide, thus enhancing the overall capture probability of the injected carriers. Thus, although in the case of an LDD structure without LDD spacer, fewer hot carriers are injected into the oxide, there is more oxide damage and consequently stronger degradation due to a significantly higher capture cross section in the damaged oxide than in the case of an LDD structure with LDD spacer.

Thus the use of an LDD-spacer is an essential precondition for a degradation resistant MOSFET. It should also be noted that an LDD-spacer alleviates the short channel effects too. At the present state of conventional LDD structures, an LDD-spacer (by reoxidation) of about 50 nm with a sufficient gate-drain overlap, 100 - 150 nm, offers optimum reliability.

To summarize this part devoted to conventional LDD structures including the ITLDD-/32/, GOLD-/33/, and PLDD /34/ versions, it can be said that the theory /30/ and simulation give useful and practical guidelines for prevention of hot electron damage in submicron MOSFETs. However, if the aim is to further improve the transistor reliability, or to maintain it with a reduced transistor geometry, conventional LDD structures must be abandoned in favor of refined LDD structures such as profiled LDDs /35/ or buried LDDs /36/ or graded-buried LDDs /36/ or consider a buried channel structure /37/.

Before embarking on a discussion of these modifications it should be pointed out, that although these structures clearly offer more parameters to construct more sophisticated LDD structures and possibly to improve the long-term stability of the device, they also possess many more pitfalls which render the devices worse than their conventional cousins. This plight is aggravated by the circumstance that in contrast to the case for conventional LDD devices, the theory and simulation up to now provide only vague guidance for purposefully optimizing the refined structures.

The profiled LDD structure (PLDD) /34/ results from the insight provided by the simulation that the LDD n^--region is divided into two parts. One is a graded profile region near the n^--channel junction, which must reduce the electric field sufficiently and suppress hot carrier generation, so that the lateral impurity profile of the junction is not changed. The other one is a flat n^--surface profile region. Most of the generated surface states are distributed above the latter region, and give rise to an increase of the parasitic resistance. Therefore, in the latter region, a sufficiently high concentration is required. This is achieved by an additional shallow arsenic implantation to form a double diffused highly doped drain. Although more reliable protection from hot-carrier injection has been reported than for conventional LDD-devices, it might be doubted whether this structure is really better than an optimized conventional LDD structure. This positive judgement on PLDDs is based on a comparison with conventional LDD transistors without the LDD spacer, which entails - as we know - considerably enhanced degradation. Furthermore, the additional As implantation tends to increase the electric fields at the surface dramatically unless the subdiffusion of the phosphorus part of the n^--region is large enough. For this reason this design should be discarded for practical purposes.

A more subtle and more promising approach is represented by the buried LDD (BLDD) structure /35/. In contrast to PLDD the BLDD employs a deep arsenic n^--implant instead of a shallow one. The arsenic dose is chosen such that a peak in the n^--impurity profiles occurs appreciably (100 nm) below the $Si\text{-}SiO_2$ interface forming a "buried" n^--layer whereas in the conventional LDD structures the n^--doping peaks at the surface. Since the maximum lateral field occurs where the n^--doping is heaviest, the maximum field is removed from the $Si\text{-}SiO_2$ interface as verified by the simulation. Thus the impact ionization occurs further away from the surface and the hot carriers near the surface experience a reduced electric field and are thus less likely than surface-heated electrons to overcome the oxide energy barrier and to damage the interface and the oxide. In addition one should attempt to separate the current path and the maximum electric field, such that the maximum current density detours the field peak resulting in a decrease of the bulk current and giving further relief to the oxide damage.

The buried channel structure /37/ exploits a similar idea of shifting the carriers away from the $Si\text{-}SiO_2$ interface. The experiments suggest that buried channel devices are more resistant to hot carrier effects than surface channel devices. This is attributed to the deeper and broader current path in the buried channel structure using p^+-polysilicon gate technology /38/. The probability that hot carriers with an energy greater than the barrier height can reach the $Si\text{-}SiO_2$ interface is supposed to be smaller in the buried channel than in a surface channel device. At the same time, buried channel devices have a higher effective mobility than surface channel devices.

In our opinion, a buried channel in conjunction with a buried LDD structure may be a promising candidate for submicron MOSFETS.

We now turn to the question of how to optimize the refined LDD structures. The experimental data and simulation show that when comparing refined LDD structures with optimized conventional LDD structures it is no longer sufficient to study the magnitude and lateral location of the relevant field peaks alone. The magnitude of the field peaks may be very similar but the degradation may nonetheless be significantly different. There are two main reasons for this circumstance: i) the degradation is extremely sensitive to

even small differences in the field peaks above $3 \cdot 10^5$V/cm, ii) even assuming the magnitude of the field peaks to be identical, the degradation will still depend considerably not only on the lateral but also on the transverse location of the field peaks, or more generally, it will depend on many subtle differences in the entire distribution of the lateral and transverse electric fields between the pinch-off point and the drain contact. It will, for instance, depend on the extent and location of the overlap between the current path and the high field distribution, see Chapter III. There are unfortunately, no clear criteria for weighting such differences with respect to degradation.

This situation calls for a refined model and theoretical insight to guide the optimum design for modified LDD structures. The model given above for conventional LDD structures /30/ cannot discriminate the aforementioned subtle differences in the field distribution, which are vital for the degradation properties of the device. A model is needed which provides clear criteria for process parameters to achieve an optimum design for hot carrier resistant source/drain structures.

Though we still lack such a model, it seems to be clear in which direction the research effort must proceed. Firstly the distribution of injected hot electrons and hot holes into the oxide must be consistently calculated. Secondly the capture, trapping probabilities (depending on interface, gate oxide and spacer oxide quality) and interaction between holes and electrons /2/ have to be reliably described. Thirdly the transient build-up of the oxide damage at different operating conditions has to be taken into account. Fourthly the influence of the oxide damage on the transistor performance must be known. The latter point was already included in the simulation /7,26/.

Finally, by reversing this chain of the analysis, it should be possible to determine the optimum process parameters and the most suitable geometry for the device. The theoretical work outlined in Chapter III shows that significant developments towards such a comprehensive model have already been implemented or are in progress.

REFERENCES

/1/ L. Risch and W. Weber, ESSDERC 1988, 6B3

/2/ K. R. Hofmann, C. Werner, W. Weber and G. Dorda, IEEE ED-32, 691, 1985

/3/ R. Bellens, P. Heremans, G. Groeseneken and H. E. Maes, IEEE EDL-9, 232, 1988

/4/ E. Takeda and N. Suzuki, IEEE, EDL-4, 111, 1983

/5/ C. Hu, S. C. Tam, F.-C. Hsu, P.-K. Ko, T.Y. Chan and K. W. Terril, IEEE, ED-32, 375, 1985

/6/ W. Weber, C. Werner and A. v. Schwerin, IEDM Digest 1986, pp.390

/7/ A. v. Schwerin, W. Hänsch and W. Weber, IEEE ED-34, 2493, 1987

/8/ W. Hänsch and F. Lau, Lecture Notes of an Inter. Short Course - NASECODEV, Dublin 1987,ed. J. J. H. Miller (BOOLE PRESS Ireland 1987) pp. 50

/9/ O. Madelung, Festkörpertheorie II, (Springer Berlin 1972), chp. VIII B

/10/ S. Selberherr, Analysis and Simulation of Semiconductor Devices, (Springer Wien 1984), chp. IV

/11/ K. Blotekjar, IEEE ED-17 38, 1970

/12/ C. Jaccoboni and L. Reggiani, Rev. Mod. Phys.55, 645 1983

/13/ W. Hänsch and M. Miura-Mattausch, J. Appl. Phys. 60 650, 1986

/14/ G. Baccarani and M. R. Wordemann, Solid State Electr.28, 407 1985

/15/ W. Hänsch and Ch. Schmeiser, Zeitschrift für angewandte Mathematik und Physik, submitted for publication

/16/ M. Fukuma and R. H. Uebing, IEDM Digest 1984, pp. 621

/17/ A. Forghieri, R. Guerrieri, P. Camolini, A. Gnudi, M. Rudan and G. Baccarani, NUPAD Symposium in Santa Clara 1986

/18/ B. Meinerzhagen and W. L. Engl, IEEE ED-35 689, 1988

/19/ W. Hänsch and S. Selberherr, IEEE ED-34 1074, 1987

/20/ W. Shockley, Solid State Electr. 2 1537, 1958

/21/ N. I. Tisnek and A. P. Dimitrev, Sov. Phys. Semicond. 31 93, 1987

/22/ L. V. Keldysh, Sov. Phys. JETP 21 1135, 1965

/23/ W. Hänsch, 1988 International Electronic Devices and Materials Symposium, Kaohsing Taiwan

/24/ B. Ridley, Journ. of Phys. C16 3373, 1983

/25/ R. Kuhnert, C. Werner and A. Schütz, IEEE ED-32 1057, 1985

/26/ A. v. Schwerin and W. Hänsch, ESSDERC 1988, Montpellier, 6B1

/27/ S. Baba, A. Kita and J. Veda, IEDM Digest 1986, pp. 734

/28/ H. Kume, E. Takeda, T. Toyabe and S. Asai, Japan I. Appl. Phys., vol. 21, suppl. 21-1, 67, 1982

/29/ P. J. Tsang, S. Ogura, W. W. Walker, J. F. Shepard and D. L. Critchlow, IEEE ED-27, 1540, 1980

/30/ M. Orlowski, Ch. Werner and J. Klink, "Model for the Electric fields in LDD-MOSFETs", Part I and II, submitted for publication in IEEE ED

/31/ M. Orlowski, Ch. Werner, W. Weber, H.-P. Mühlhoff and P. Murkin, VLSI-Symp. 1987, p. 51

/32/ T. Huang, W. W. Yao, R. A. Martin, A. G. Lewis, M. Koyanagi and J. Y. Chen, IEDM Digest 1986, pp. 742

/33/ R. Izawa, T. Kure, S. Iijima and F. Takeda, IEDM Digest 1987, pp. 38

/34/ Y. Toyoshima, N. Nihira and K. Kanzaki, ULSI-Symp. 1985, pp. 118

/35/ H. R. Grinolds, M. Kinugawa and M. Kakumu, IEDM Digest 1985, pp. 246

/36/ C.-Y. Wei, J. M. Pimbley and Y. Nissan-Cohen, IEEE ED-7, 380, 1986

/37/ M.Nakahara, Y.Hiruta, M.Yoshida, K.Maeguchi and K.Kanzaki, IEDM Digest 1985, pp.238

/38/ C. Mazure, U. Schwalke and F. Neppl, ESSDERC 1988, Montpellier 4B4

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

ELECTRON BEAM SOURCE MOLECULAR BEAM EPITAXY OF $Al_xGa_{1-x}As$ GRADED BAND GAP DEVICE STRUCTURES

R.J. MALIK, A.F.J. LEVI, B.F. LEVINE, R.C. MILLER, D.V. LANG, L.C. HOPKINS and R.W. RYAN

AT and T Bell Laboratories, Murray Hill, NJ 07974, U.S.A.

ABSTRACT

A new method has been developed for the growth of graded band-gap $Al_xGa_{1-x}As$ alloys by molecular beam epitaxy which is based upon electron beam evaporation of the Group III elements. The metal evaporation rates are measured real-time and feedback controlled using beam flux sensors. The system is computer controlled which allows precise programming of the Ga and Al evaporation rates. The large dynamic response of the metal sources enables for the first time the synthesis of variable band-gap $Al_xGa_{1-x}As$ with arbitrary composition profiles. This new technique has been demonstrated in the growth of unipolar hot electron transistors, graded base bipolar transistors, and M-shaped barrier superlattices.

The optical and electronic properties of $Al_xGa_{1-x}As/GaAs$ heterojunctions and quantum wells have been extensively studied in materials grown by molecular beam epitaxy (MBE). There are, however, a number of interesting structures which have been proposed which require precise alloy grading over atomic dimensions. The ability to tailor band structure to obtain novel electrical and optical properties is termed "band-gap engineering".[1] Previous attempts to obtain alloy grading by varying the temperature of MBE effusion cells have been only marginally successful and have serious limitations.[2-4] This is related to the large thermal inertia and restrictions in heating and cooling rates of the MBE effusion cells which leads to slow modulation of the beam flux along with a time lag in response. In addition, variations in the growth rates determined by the Group III molecular beam fluxes seriously complicates the problem of obtaining arbitrary alloy grading. Thus Group III sources with large dynamic response are required for band gap engineering of $Al_xGa_{1-x}As$. Several different approaches could be used which include electron beam heated sources which are typically used to evaporate Si and refractory metals,[5] gas sources,[6] and ion beam sources.[7]

We recently reported on the use of electron beam source MBE growth of III-V compounds.[8] In this paper, we discuss the use of this technique for precise analog grading of $Al_xGa_{1-x}As$ and its application in the growth of unipolar hot electron transistors, graded base bipolar transistors, and M-shaped barrier superlattices.

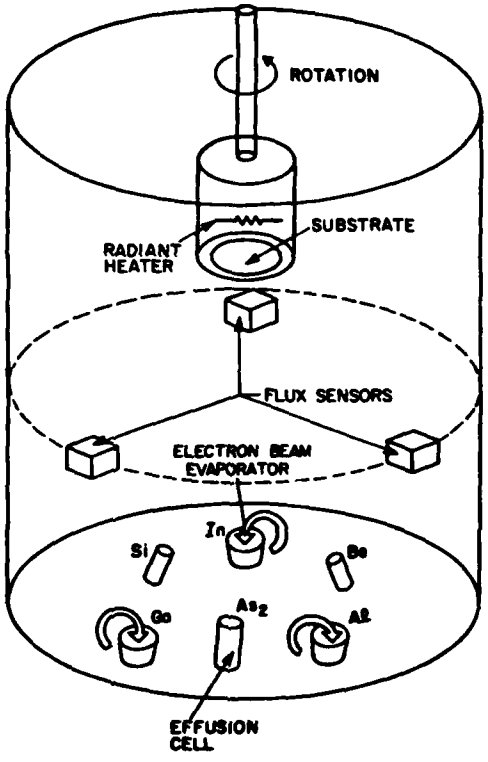


Fig. 1. Schematic diagram illustrating the geometry of the electron-beam source III-V molecular beam epitaxy system.

The geometry of the e-beam source MBE system is shown schematically in Figure 1. The MBE system is a vertical design which is necessitated by the use of the electron beam evaporation sources and is similar in design to systems used for silicon or metals epitaxy. Three electron beam sources are used to evaporate Ga, Al, and In since these elements have unity sticking coefficients and determine the alloy composition in III-V compound semiconductors grown by MBE. A large capacity, 300 cm^3 arsenic source with a high temperature cracker is used to generate an As_2 molecular beam. A Si strip heater and small 2 cm^3 Be effusion cell are used as n- and p-type doping sources, respectively. All the molecular beam sources have separate mechanical shutters which are computer controlled. A substrate manipulator which allows substrate heating and rotation is located 250 mm above the source flange. Indium-free wafer holders for 50 mm and 75 mm substrates can be fitted on the manipulator.

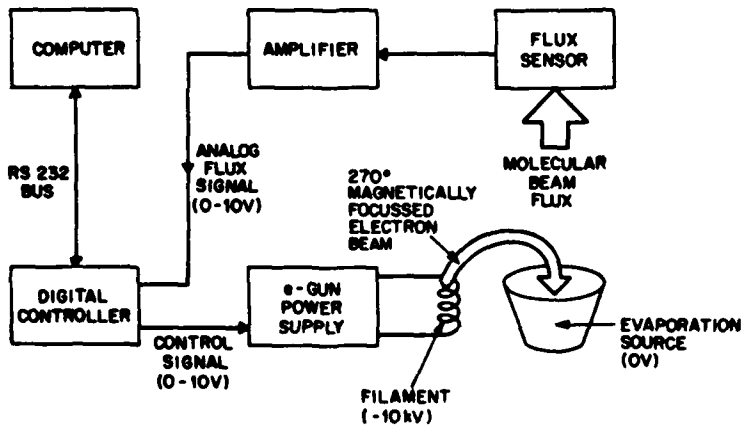


Fig. 2. Block diagram of the feedback control loop for the electron-beam evaporation sources.

A block diagram of the feedback control loop used to measure and program the evaporation rates for the Ga and Al e-beam sources is shown in Figure 2. Both a modulated ion gauge and a commercially available Inficon Sentinel III flux sensor[9] have been used successfully to measure the molecular beam fluxes. They both have adequate sensitivities to measure and control evaporation rates as low as 0.01 Å/sec. The Sentinel III flux sensor which uses a transducer based upon electron impact emission spectroscopy is superior from the standpoint of materials selectivity and its immunity from charged particles causing measurement error. An amplified analog signal which is directly proportional to the molecular beam fluxes or evaporation rates is input to a digital controller which interfaces with the e-gun power supply. The bandwidth for the feedback control loop of the Sentinel III flux sensor is 8 Hz. A computer communicates with the digital controller on a RS 232 bus and updates the evaporation rate setpoints every second. An absolute calibration of the molecular beam fluxes is made in situ by use of reflection electron diffraction oscillations.[10]

Precise alloy grading of $Al_xGa_{1-x}As$ is accomplished by the following method. The $Al_xGa_{1-x}As$ growth rate is kept constant during epitaxy and is just the sum of the individual growth rates for AlAs and GaAs. The normalized growth rate for $Al_xGa_{1-x}As$ can be expressed as

$$r_{AlGaAs} = r_{AlAs} + r_{GaAs} = 1$$

The Al mole fraction of the alloy is given by the ratio of the AlAs growth rate to the total growth rate

$$x = \frac{r_{AlAs}}{r_{AlAs} + r_{GaAs}}$$

Now if the individual growth rates are varied such that

$$r_{AlAs}(t) = f(t)$$

$$r_{GaAs}(t) = 1 - f(t)$$

then the time dependent alloy composition during growth is simply given by

$$x(t) = f(t).$$

Also since r_{AlGaAs} remains constant, then the position dependence of Al alloy composition can also be expressed as

$$x(d) = f(t)$$

where $d = r_{AlGaAs} \cdot t$ is the position within the epilayer. Thus it is evident that even very complicated mathematical expressions for alloy grading such as

parabolic, sinusoidal, or exponential can be readily obtained by computer programming.

An $Al_xGa_{1-x}As$ unipolar hot electron transistor was grown by e-beam source MBE using analog grading for the emitter and collector barriers.[11] A SIMS profile of the Al alloy composition in the electron barriers is shown in Figure 3. The structure was designed with a collector barrier peak of $x = 0.35$ and an emitter barrier peak of $x = 0.3$. The long and short arms of the graded barrier were 1000Å and 90Å, respectively with a barrier peak width of 50Å. The emitter, base, and collector were doped n-type to $1\times10^{18}cm^{-3}$ and the graded barriers to $1\times10^{16}cm^{-3}$. Mesa transistors were fabricated using anodic etching to the base and wet chemical etching to the collector followed by an alloyed Au/Ge/Ni metallization for contacts. Hot electron spectroscopy[12-13] was performed on a transistor with a base width of 400Å at 4.2° K. The derivative spectrum of collector current shows clear evidence of a "ballistic" electron peak. This is the first time that ballistic transport has been seen in analog graded $Al_xGa_{1-x}As$ structures. It should be noted that in a number of attempts to simulate alloy grading by using variable duty cycle chopped $Al_xGa_{1-x}As/GaAs$ superlattice,[14] no ballistic peak was observed. We infer from this that the multiple interfaces or trapped impurities at the interfaces in the chopped superlattices leads to enhanced electron scattering suppressing ballistic transport.

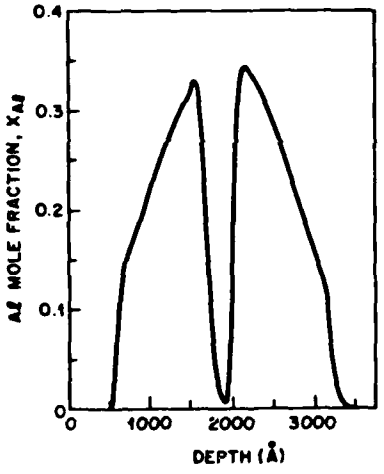


Fig. 3. Secondary-ion-mass-spectrometer measurement of the Al alloy composition in the analog-graded, unipolar, hot electron electron transistor.

The Al beam flux signal in the growth of a graded band-gap base $Al_xGa_{1-x}As/GaAs$ heterojunction bipolar transistor is shown in Figure 4. The Al composition is graded from x = 0.04 to x = 0.2 in the 1000 Å base region. The emitter junction is also graded from x = 0.2 to x = 0.33 over 300 Å and then graded back down to x = 0 over 500 Å for the emitter contact region. It has been previously demonstrated that alloy grading in the base results in a quasi-electric field in the base which reduces the base transit time.[15-16] The precise grading control is clearly illustrated here.

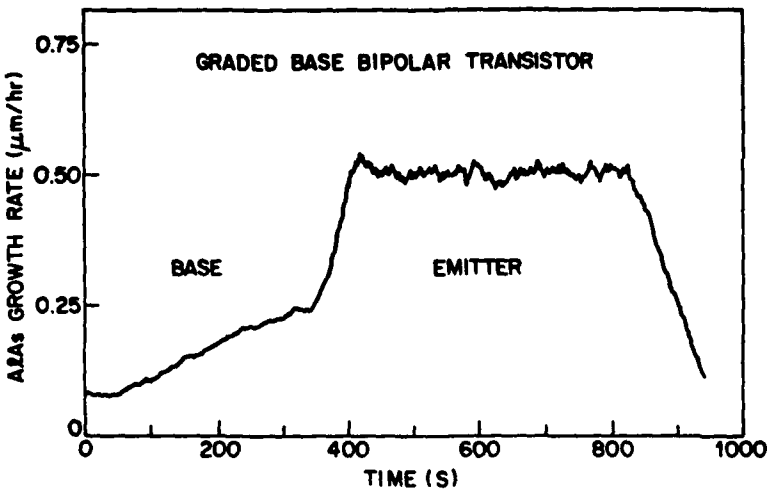


Fig. 4. Al beam flux signal in the growth of a graded base heterojunction bipolar transistor.

The Al beam flux signal in the growth of an M-shaped barrier superlattice is shown in Figure 5. This superlattice is used in an intersubband absorption quantum well photodetector.[17-18] The Al composition was linearly modulated from x = 0.3 to x = 0.15 and back to x = 0.3 over a distance of 150 Å. The sample contained a 50 period superlattice and was grown continuously without interruption. The excellent linear grading profile is apparent, and it is obvious that such a structure would be impossible to grow using conventional effusion cell sources.

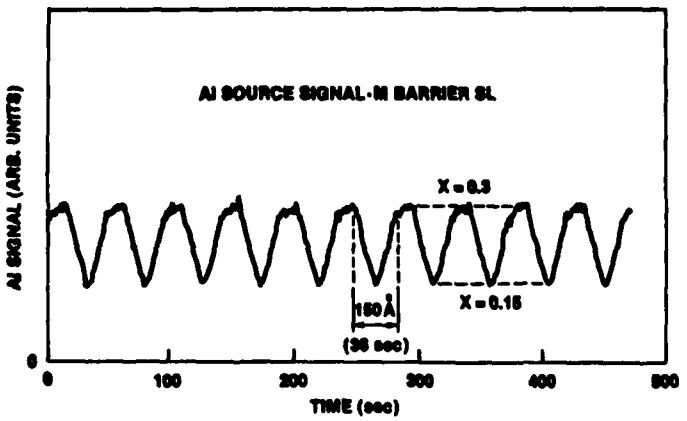


Fig. 5. Al beam flux signal in the growth of an M-shaped barrier superlattice.

In conclusion, a new technique has been developed for the growth of graded $Al_xGa_{1-x}As$ structures with arbitrary alloy profiles by molecular beam epitaxy. The system uses Ga and Al electron-beam heated evaporation sources which are measured real-time and feedback controlled using beam flux sensors. The MBE system has been used to grow novel $Al_xGa_{1-x}As$ graded alloy devices.

We gratefully acknowledge partial support of this work through NASA Goddard Space Flight Center contract NAS5-28602.

REFERENCES

[1] F. Capasso, Ann. Rev. Mater. Sci., *16*, 263-91 (1986).

[2] J. P. Harbison, L. D. Peterson, J. Leskoff, Proc. Fourth Int. Conf. MBE, U. of York, September 1986 (to be pub. J. Cryst. Growth).

[3] N. J. Sauer, T. Y. Chang, A. H. Dayem, E. H. Westerwick, J. Vac. Sci. Tech. B, *5*, 718 (1987).

[4] K. Alavi, A. Y. Cho, F. Capasso, and J. Allam, J. Vac. Sci. Tech. B, *5*, 802 (1987).

[5] J. C. Bean, Proc. Fourth Int. Conf. MBE, U. of York, September 1986 (to be pub. J. Cryst. Growth).

[6] M. B. Panish, H. Temkin, and S. Sumski, J. Vac. Sci. Tech. B, *3*, 657 (1985).

[7] S. Shimizu, O. Tsukakoshi, S. Komiya, and Y. Makita, Jap. J. Appl. Phys., *24*, 1130-40 (1986).

[8] R. J. Malik, J. Vac. Sci. Tech. B, *5*, 722 (1987).

[9] Inficon Leybold-Heraeus Co., Syracuse, NY 13057.

[10] J. H. Weave, B. A. Joyce, P. J. Dobson, and N. Norton, Appl. Phys. A, *31*, 1 (1983).

[11] R. J. Malik and A. F. J. Levi, Appl. Phys. Lett., *52* (1988).

[12] J. R. Hayes and A. F. J. Levi, IEEE J. Quantum Electron., *QE-22*, 1744 (1986).

[13] M. Heiblum, M. I. Nathan, D. C. Thomas, and C. M. Knoedler, Phys. Rev. Lett., *55*, 2200 (1985).

[14] M. Kawabe, M. Kondo, N. Matsuura, and K. Yamamoto, Jpn. J. Appl. Phys., *22*, L64 (1983).

[15] B. F. Levine, C. G. Bethea, W. T. Tsang, F. Capasso, K. K. Thornber, R. C. Fulton, and D. A. Kleinman, Appl. Phys. Lett., *42*, 769 (1983).

[16] R. J. Malik, F. Capasso, R. A. Stall, R. A. Kiehl, R. W. Ryan, R. Wunder, and C. G. Bethea, Appl. Phys. Lett., *46*, 800 (1985).

[17] B. F. Levine, K. K. Choi, C. G. Bethea, J. Walker, and R. J. Malik, Appl. Phys. Lett., *50*, 1092 (1987).

[18] K. K. Choi, B. F. Levine, C. G. Bethea, J. Walker, and R. J. Malik, Appl. Phys. Lett., *50*, 1814 (1987).

Session 6A Room A : Miscellaneous

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

POSITION RESOLVED CARRIER LIFETIME MEASUREMENTS IN SILICON POWER DEVICES BY TIME RESOLVED PHOTOLUMINESCENCE SPECTROSCOPY

G. BOHNERT, R. HÄCKER and A. HANGLEITER

4. Physikalisches Institut, Universität Stuttgart, Pfaffenwaldring 57, D-7000 Stuttgart 80, F.R.G.

Abstract

We present a new versatile method for measuring position resolved carrier lifetimes in semiconductor power devices, using time resolved photoluminescence spectroscopy. The experimental results obtained at ASCRs and silicon diodes are in good agreement to what is expected from the doping profile. In addition the effects of lifetime reducing processes, e.g. gold diffusion, e^-- or H^+- irradiation can be clearly demonstrated. The resolution in position is only limited by carrier diffusion.

Introduction

For a quantitative understanding of fast semiconductor power devices a detailed knowledge about carrier lifetime and its position dependence is essential. Especially the lifetime along the base region is of interest since it determines several important properties of the device, like the forward voltage, the reverse current or the reverse recovery time. With steadily increasing requirements on such devices new manufacturing techniques are needed in order to get an optimized doping profile and a well defined carrier lifetime profile across the device. Therefore a thorough understanding of the influence of the diverse processing parameters on carrier lifetime is needed.

While it is a rather simple task to measure bulk lifetimes in homogeneously doped samples, things get more complicated for devices. In the latter case one can only determine an effective lifetime since the decay of excess carriers is now determined by a superposition of recombination and diffusion processes. In addition the excact doping profiles as well as the distribution of recombination centers which may be present wantedly or unwantedly are not known beforehand.

In this paper we will present a new way of measuring the position dependency of carrier lifetimes in semiconductor power devices directly by using time resolved

photoluminescence. We will discuss results obtained from measurements on various prepared asymetric silicon controlled rectifiers (ASCRs) and silicon diodes.

Experiment

The basis of our method is the excitation of electron- hole pairs in the device by a pulsed laser beam (Ar- Ion- Laser, λ= 514 nm). These eh-pairs will return to the ground state via band-to-band Auger recombination or via recombination centers. The lifetime τ is obtained by observing the temporal development of the decreasing luminescence signal since signal intensity is a measure for the density of minority charge carriers.

The position resolution is achieved by using a translation stage (see fig. 2), which can be moved in steps as little as 1.25 μm. Thus a focussed laser beam can be scanned across the device. The scan will yield either lateral or cross-sectional (dependent on the preparation) position dependencies of the lifetime. Since the focus can be made very small (1 μm approx.) the only limitation in position resolution is the principal resolution limit given by the diffusion length.

A typical result of such an experiment is shown in fig. 1: immediately after the excitation a rapid non- exponential decrease of the luminescence signal occurs due to surface recombination and possibly due to band-to-band Auger recombination because of the high carrier density in that time intervall. After the transition to the low injection case an exponential decrease with time constant τ is observed.

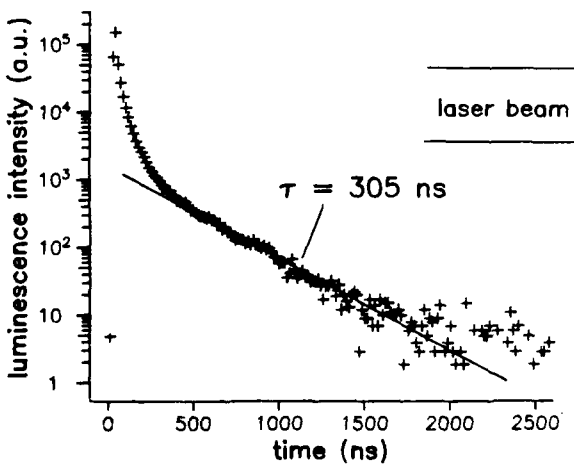


Fig. 1: Example for time development of photoluminescence.

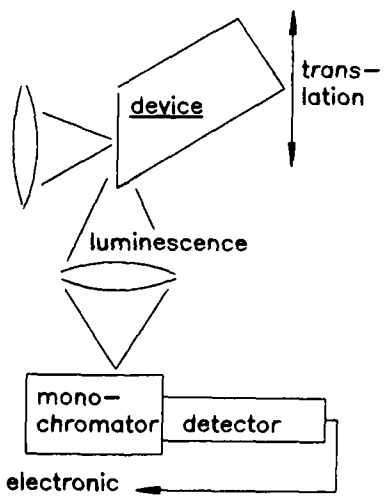


Fig. 2: Experimental set-up.

Results and Discussion

We have done our measurements on differently prepared ASCRs and silicon diodes. The schematic cross-sectional view of an ASCR is depicted in fig. 3a. The results are shown in fig. 3b-d. Fig. 3b shows the cross-sectional lifetime profile for the ASCR with no additional recombination centers. At a first glance a variation of carrier lifetime over three orders of magnitude is recognized. By comparison with the schematic diagram the lifetime profile can be understood in terms of the doping profile. The rapid decrease at the ends of the device section is due to the heavy doping of the p^+ and n^- emitters. Here the lifetime is limited by the band-to-band Auger recombination /1/. The lifetime gradient along the n-base is explained by the inverse gradient of the doping concentration (transition from n- to n^--doping). In this region the lifetime is not limited by the bulk properties but rather by diffusion from the bulk to its surfaces and interfaces. The pure bulk lifetime for a doping level of 10^{13} cm^{-3} and in absence of recombination centers would be of the order of some milliseconds /2/. One also recognizes a significant enhancement of carrier lifetime at the p^+-n and p-n^+ junctions. This can be explained by a carrier separation of the excited eh-pairs by the build-in electric field, which will reduce the recombination probability.

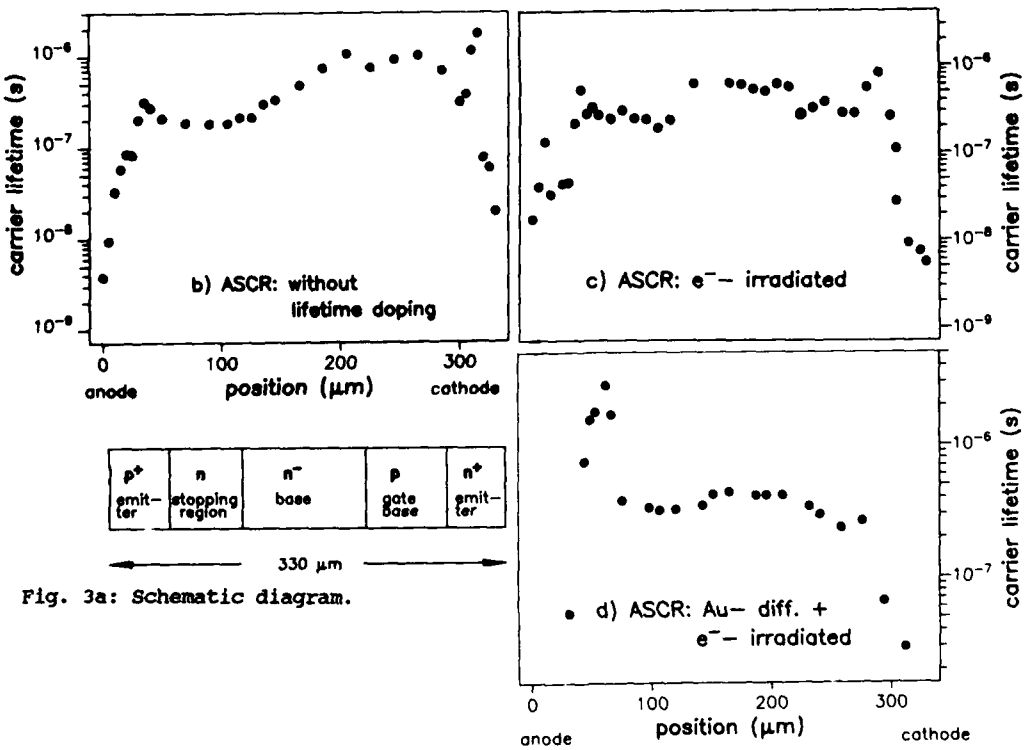


Fig. 3a: Schematic diagram.

Fig. 3a-d: schematic diagram and measured lifetime profile for three differently prepared ASCRs.

For comparison we have also measured cross-sectional lifetime profiles for two other ASCRs, which are differently prepared in order to reduce the lifetime in the base region (fig. 3c and d). One is e^--irradiated and the other e^--irradiated and gold diffused. It is clearly seen that after each device preparation step the effective lifetime in the base is reduced by roughly a factor of two. In the case of gold diffusion and e^--irradiation the lifetime gradient in the n-base disappears since now the recombination via deep level impurities dominates /3/.

In fig. 4 the measured lifetime profile and the schematic diagram for an e^-- and H^+- irradiated silicon diode is shown. From the lifetime profile it is clear that the H^+- irradiation occurs at the anode side and that the penetration depth for protons at this specific energy is about 100 μm. The lifetime gradient is again correlated with the inverse gradient of the defect concentration.

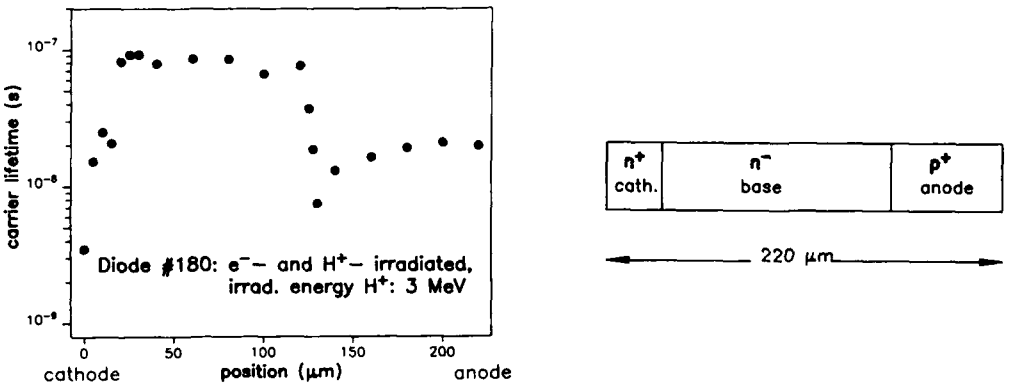


Fig. 4: Schematic diagram and measured lifetime profile for a H^+- irradiated silicon diode

In conclusion we have demonstrated the usefulness and versatility of time resolved photoluminescence measurements in order to obtain spatially resolved lifetime profiles of semiconductor devices. Results can be obtained directly without any complicated parameter fitting.

Acknowledgement

We are indebted to E.Borchert of AEG Warstein-Belecke, W.D.Nowak, and W.Wondrak of the AEG-Forschungsinstitut Frankfurt for providing the samples used in our work.The financial support by the Bundesministerium für Forschung und Technologie is gratefully acknowledged.

References

/1/ J.Dziewior, and W.Schmid, Appl.Phys.Lett. 31, 346, (1977)

/2/ A.Hangleiter, and R.Häcker, Proc. of the 18th Int. Conf. on the Phys. of Semiconductor Devices Vol.2, edited by O.Engström, World Scientific(1988)

/3/ A.Hangleiter, Phys. Rev. B 35, 9149 (1987)

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

POWER MOS FET MODELS FOR "SWITCHING" CIRCUITS

P. ROSSEL*, R. MAIMOUNI***, M. BELABADIA*, H. TRANDUC*,
C.E. CORDONNIER** and M. BAIRANZADE**

*LAAS du CNRS, 7, Av. du Colonel Roche, F-31077 Toulouse Cedex, France
**MOTOROLA Semiconducteurs, Le Mirail, BP 1029, F-31023 Toulouse Cedex, France
***Faculté de Sciences, Oujda, Maroc

RESUME: *Un modèle compact du transistor VDMOS de puissance, compatible avec le logiciel "SPICE2", est proposé En tant qu'applications, le transistor est "simulé" en régime de commutation sur des charges résistives et inductives et les résultats obtenus sont validés par des mesures expérimentales.*

ABSTRACT: *A compact model of the Power VDMOS Transistor compatible with the circuit simulator "SPICE2" is described in this article. This model is applied to the simulation of switching circuit with resistive and inductive loads; comparisons with experimental results are presented.*

I. INTRODUCTION:

In the Power Mosfet area, several suggestions on equivalent circuit have been recently published by International Rectifier [1], Motorola, RCA [2] Siemens [3] and the Washington University [4]. It consist in rather specific models obtained by identification of some electrical characteristics: for a given model, the application domain is not valid simultaneously for DC and dynamic operation, as well as for resistive and inductive loads. These models can use two active components from the SPICE library which are the MOS and the JFET; the use of the later for simulating the "quasi-saturation" phenomenon [5, 6] is no more useful, today,for the modeling as actual structures are optimized in which this phenomenon now appears only for drain current level over the nominal current rating.

In this paper, we describe a basic but accurate model for switching circuits with resistive and inductive loads. This model is compatible with the well known "SPICE" software. It works as well as with low voltage structures (one hundred of volts) of both standard and Logic Level (L^2FET [5] with a nominal gate source voltage of 5 volts) types, with medium voltage (100 - 500 volts) and high voltage devices (>500 volts). Its configuration is based on a more complete but also more difficult to handle model [8-10] whose parameters are all related to the physical properties and topology of the VDMOS transistor. Their values can be obtained from datasheets and thanks to some "classic" measurements. It must be pointed out that this model takes into account the high degree of non linearity of the gate-drain and drain-source capacitors and also some short channel effects (variable mobility, etc...) which mainly prevail in the low and medium voltage structures.

II - DESCRIPTION OF THE VDMOS TRANSISTOR MODEL:

II-1- VDMOST structure and model topology:

The model structure is directly related to the power VDMOS geometrical topology (see Fig.1). We can notice on the schematic: i) the intrinsic transistor (conduction channel) represented by the current generator Jd, ii) the n-type layer between the p wells with the access resistance Ra to the channel and the highly non linear gate-drain capacitor Cgd, i.e. the depletion capacitor Cd in series with the thin gate oxide capacitor Cgdmax, iii) the low doped n-type epitaxial bulk accounted for by the drift resistance Rd, the pn junction capacitor Cds with the p well and the internal body diode Dbody, iv) the gate metallization covering over the N^+ source diffusion which creates, through the thin and thick oxide, a constant parasitic capacitor Cgso. When the MOS transistor goes through a switching phase, all the three different modes: Blocked, Ohmic and Saturation, are involved. The model (see Fig. 2) takes into account these three functioning modes. In the figure 2, the conduction channel under the gate, i.e. the intrinsic transistor, is described by the current generator Jd and the non linear capacitors C1 and C2.

In the "Ohmic" mode (i.e. Vd < Vp), the current Jd as approximated from the general formulation , is given by:

$$(1)\quad Jd = \frac{Kp}{\left(1+(Vg-VT)/psi\right)\left(1+Vd/L.EO\right)}\left((Vg-VT)Vd - 0.5\,Vd^2\right)$$

with $Kp = \mu_0(Z/L).Cox$, where μ_0 is the carrier mobility at low electric field level, Z the channel perimeter, L the channel length, Cox the gate oxide capacitor per unit area, Vg and Vd the gate-source and drain-source potentials, VT the threshold voltage. "psi" and L.EO are the potentials related to the effect of the transverse and longitudinal electric fields on the carrier mobility in the channel (i.e. transverse roll-off limitation and longitudinal velocity saturation). The capacitors C1 and C2 can be considered as equal and are approximated to (Cox.Z.L)/2 [10].

In the "Saturation" mode (i.e. Vd ≥ Vp), the drain voltage Vp corresponding to the onset of the channel pinch-off, assuming a perfect current saturation i.e. there is no effect of channel length modulation [11], is given by:

(*) Ce travail est soutenu au L.A.A.S. du C.N.R.S. partiellement par le C.N.E.T. (contrat 87 8B 075 790 9245 LAB/ICM/TOH) et par le GRECO/CNRS Electrotechnique.

(2) $Vp = L.EO\left(\left(1+2\frac{Vg-VT}{L.EO}\right)^{1/2}-1\right)$

In that mode, the channel is pinched and "disappears" on the drain side, so does the capacitor C1, and C2 takes a value close to 2/3 Cox.Z.L [8,9].

In the "Blocked" mode (i.e. Vg < VT), the drain current will theoretically be null as the "subthreshold" current is neglected, and the intrinsic transistor is reduced to one element: the oxide capacitor Cgsb. Although this is a distributed capacitor over the whole P region, we suppose it, for the model, located between the gate and source electrodes (Cgsb=Cox.Z.L).

We complete the model with the source inductance Ls which can induce an important feed-back effect from the output to the input. The static parameters are Kp, VT, psi, L.EO, Ra and Rd; they can be obtained by a classical way with low and high level drain voltage measurements [12]. The dynamic parameters are Rg, C1, C2, Cgso, Cgsb, Cgd, Cds and Ls, they can be partially obtained with an original method developed previously [10], based on the constant current gate charge relation.

II-2- Implantation of the model for SPICE program:

This VDMOS transistor model cannot be described by a single component of the SPICE library so it is necessary to build a macro-component (sub-circuit) in which each element (non linear capacitor, current generator, etc.) is a part of this library. Then the current generator Id is represented by the MOS model level 1 or 3, and the non linear capacitors Cds and Cd thanks to the transition capacitances of the diode model. The elements Cgsb, Cgso, C2 are represented by an unique capacitor Cgs; this one is constant in a first order approximation as for Ra and Rd too. The capacitor C1 acts only in the "ohmic" mode so it can be neglected for the purpose of simulation. The remaining SPICE parameters are defined as: VMAX / MUO = EO and THETA=1/psi. The macro model obtained is shown in Fig. 3. We propose two solutions: i) the first one is what we call the "Switch Solution" because in order to simulate the non linearity of the Miller capacitance Cgd, we use alternately two components: a diode Cd' and a capacitor Cgdmax which are switched thanks to two transistors $M_{2,3}$ of the model level 1, ii) the second one requires less components but an initialization procedure to set up the static bias of the middle node of the two capacitors in series Cgdmax and Cd. This second model is much faster in CPU time than the first one as the transistor number is decreased by a three fold, but because of its initialization step, it cannot be set up as a "black-box" sub-circuit. The related listings can be supplied by the authors.

III. APPLICATION OF THE MODEL IN A SWITCHING CONFIGURATION:

III-1- Basic considerations on the device operation:

Beside the direct mode, the model takes account of the behaviour of the device within the third quadrant of the output electrical characteristics (Fig.4); this operating mode is useful for the modeling of applications such as synchronous rectifiers, power energy converters and controls, in particular, those without "free-wheeling" diodes. In this case, the current is shared by both the MOS conduction channel and the "body" diode. It is worthy to notice that, at "low level" current (i.e. compared to the device rated continuous current), the channel contribution decreases noticeably the conduction losses whereas, at higher level, the resistivity modulation of the drain epilayer by the diode holes injection allows an overall series resistance noticeably lower than the MOS on-state resistance RON from a 1:2 to 1:20 factor for 60V to 1000V device voltage ratings. Thus the "body" diode is differentiated from the element Cds in the model.

III-2- Switching operation :

We use here our model in a switching configuration (see Fig.5) and look at the gate-source VGS, drain-source VDS voltages and drain current ID. The electrical characteristics are simulated under SPICE "transient analysis" option, and compared to experimental temporal waveforms, as well as for low voltage products (MTP25N06L - 60 volts) and for high voltage ones (MTH5N100 -1000 volts).

III-2-1- Resistive switching (LD=0) :

As one can see on the figure 6, the simulation waveforms agree well with experimental results. If we look closer at the switching times ton and toff, we notice a difference less than 10% between the measured and simulated values for all the different devices studied.

III-2-2- Inductive switching (RD=0) :

This test is more realistic since it covers a lot of applications : motor drive, converters etc... The obtained waveforms are quite different from those studied previously on resistive loads and depend on how big the time constant LD/(RD + RON) is. Furthermore, as the inductive load LD on the transistor drain can induce very high voltage spikes at turn off, the transistor is usually clamped by several different ways : "free wheel" diode, SNUBBER, Zener clamp etc....

At first, we have verified that the model, without any protection on the drain, allows to forecast the voltage spike value on the drain. But usually the device is protected by a clamp between drain and source, so the overvoltage is limited to a value below the breakdown voltage of the MOS'T. The waveforms we got in this configuration are shown in Fig.7a. Fig.7b demonstrate the good precision of the simulation for both the voltage-current values and waveforms. The overvoltage is accounted for by the transient factor LD.dID/dt which can be very high.The drain voltage remains equal to Vclamp as long as there is energy to dissipate in the load LD. As soon as it reaches its clamping value, the drain current falls very sharply . The time needed for this current to fall to zero (as it is setting up in the clamp) depends on the device characteristics (mainly its source inductance Ls), on the initial peak value of the current and on the clamp used. Depending on how big the overvoltage is, we can notice some oscillations of the drain voltage whose period is related to the RLC series circuit elements: the load LD, the drain-source Cds and gate-drain Cgd capacitors, the damping resistances RD, RON and RG.

From a practical point of view, it is important to notice that a transient current flows into the gate resistance RG; the current value is proportional to dVDG/dt. Thus a transient gate voltage is set up by the ohmic drop through RG. Particularly, if the RG

resistance has a rather big value, the oscillations can bring a too high (dangerous) gate voltage well beyond the absolute ratings of the transistor.

IV. CONCLUSION:

We have shown here a model for VDMOS power transistors (working in switching mode) applicable to the analog circuit simulator U.C. Berkeley's SPICE2; it is obtained from a previous and more complete model [8-10].The SPICE macro-component requires five static "elements": M1(Jd), Dbody, Rd, Ra, Rs, and six dynamic ones: Cgdmax, Cd, Cgs, Cds, Rg, Ls. The way of determining them is explained in the several reports noted in the references.

The future work will consist in: i) extending these results to other Power Mosfets devices from the Motorola TMOS™ family (actually -june 1988- a preliminary library can be furnished on request), ii) taking into account the thermal effects (at least a global manner), iii) to transfer these results to end users to check if this model feeds their needs, iv) finally, to compare this model with those existing from other laboratories: especially for implantation method, parameters extraction, accuracy, calculation speed. About this last criteria, some comparisons, even if not systematic, have been performed and show that our model is better with a factor varying from 30 to 80% compared to other proposed configurations.

ACKNOWLEDGEMENTS:

The authors thank M. NAPIERALSKA, D. ALLAIN of LAAS (Toulouse, FRANCE) for all the electrical parameters determination and switching measurements.

REFERENCES:

[1] HEXFET Power MOSFET Designer's Manual, "SPICE Computer Models for HEXFET Power MOSFETs.", International Rectifier, Chapter 24, pp 1-191, sept.1987.

[2] H. R. RONAN, F.WHEATLMEY, GE/R.C.A. SSD Mountaintop, PA 18707, "Méthode de modélisation pour MOS de puissance", Revue Electronique de Puissance, N°23, pp.32-37, Oct.1987.

[3]J. M. HANCOCK, "Enhanced Techniques for SPICE Modeling of Power MOSFETs", PCI Proceedings, pp 268,284, 1988.

[4] H.P.YEE, P.O.LAURITZEN, "SPICE Models for Power MOSFETs: An Update", IEEE, CH2504-9/88/0000-0281, 1988.

[5]J.L. SANCHEZ, M. GHARBI, H. TRANDUC, P. ROSSEL, "Quasi-Saturation Effect in High-Voltage V.DMOS Transistors", IEE Proc., Vol. 132, Pt 1, N°1, pp. 42-46, Feb. 1985.

[6] M.N. DARWICH, "Study of Quasi-Saturation Effect in V.DMOS Transistors", IEEE Trans. on Elec. Dev., Vol. ED-33, N°11, pp. 1710-1716, Nov. 1986.

[7] C.E. CORDONNIER "Le MOS de Puissance à niveau logique", Electronique de Puissance, N°20, pp.45-51.

[8] S. LATRECHE, "Le Transistor M.O.S. à canal vertical en régime nonlinéaire d'amplification haute fréquence de puissance", Thèse de 3ème Cycle, Université Paul Sabatier, Toulouse, 1985.

[9] G. TARDIVO "Le Transistor D.MOS vertical en amplification haute fréquence de puissance", Thèse de 3ème Cycle, Université Paul Sabatier, Toulouse, 1987.

[10] M. BELABADIA, R. MAIMOUNI, P. ROSSEL, D. MOUSSAID, G. TARDIVO, "Modèles dynamiques du transistor V.DMOS de Puissance - Application aux régimes de commutation", Rapport LAAS N° 87149, Mai 1987.

[11] M. GAMBOA, P. ROSSEL, H. TRANDUC, T. PHAN PHAM, "Le Transistor : M.O.S. de Puissance en régime de saturation : : la résistance de saturation et les effets de faible multiplication", Rev. de Physique Appliquée, Vol. 17, N°1, pp. 65-74, Fév. 1982.

[12] P. ROSSEL, H. TRANDUC, J.L. SANCHEZ, A. BELLAOUAR, "Détermination expérimentale des paramètres des transistors MOS", Rev. de Physique Appliquée, pp. 487-493, Août 1983.

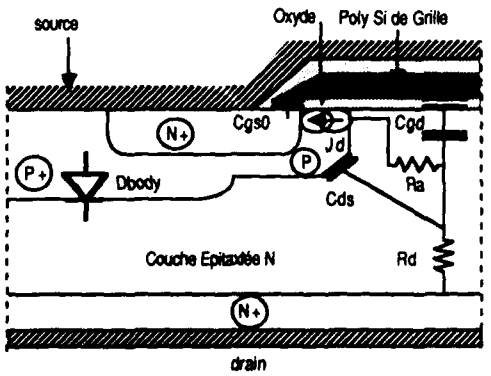


Figure 1: Cross section of the VDMOS power transistor. Localization of the simplified model elements .

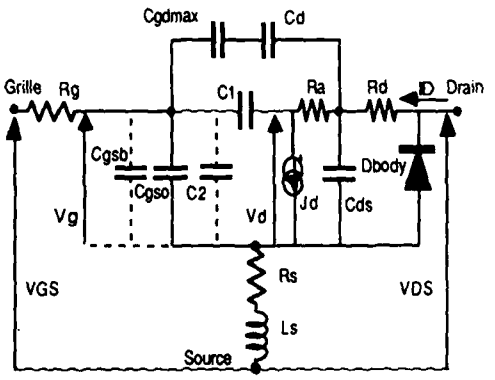


Figure 2: Simplified model of the VDMOS power transistor in conduction (in blocked mode, Cgsb is paralleled to Cgso).

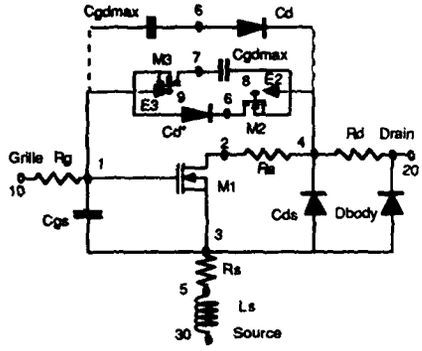


Figure 3: SPICE macro-model for VDMOS power transistors. Switch model and faster model requiring an initialization step.

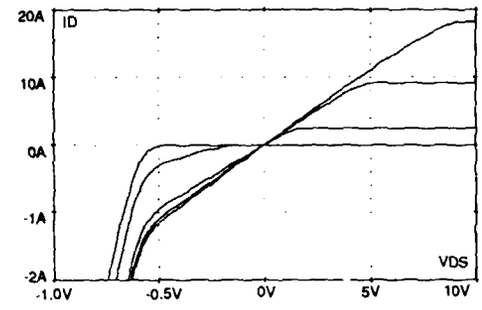


Figure 4: VDMOS'T output characteristics. 600 volts transistor MTP6N60 - direct and reverse operating mode-. (VGS step: 1V, max: 6V)

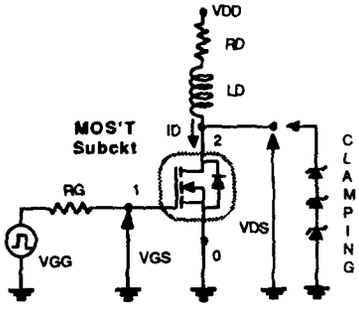


Figure 5: SPICE electrical circuit for switching modeling.

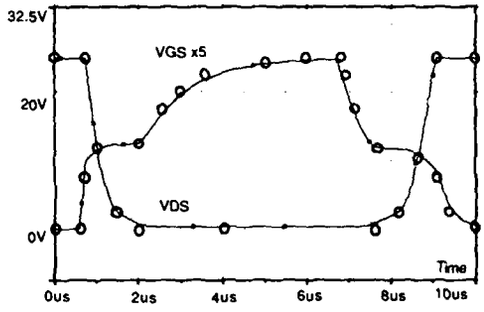
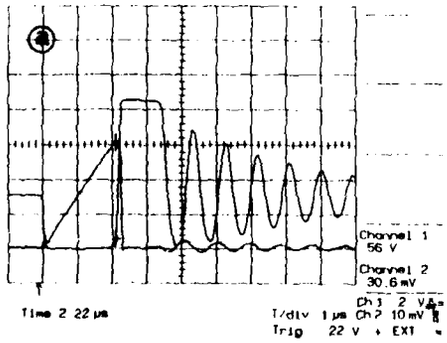


Figure 6: Drain source and gate source voltage waveforms for a resistive switching. 60 volts transistor MTP25N06L. VDD=25V, RD=9 Ohms, VGG=5V and RG=250 Ohms.
O Measurement ▫ ▪ Simulation



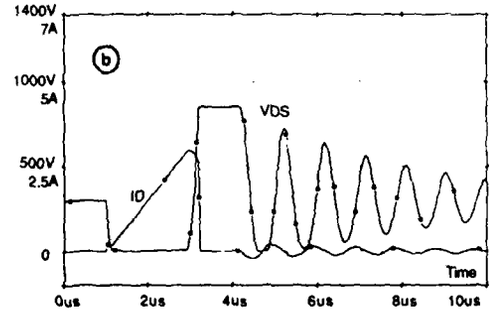


Figure 7: Drain source and gate source voltage waveforms for an inductive switching with a CLAMP protection on the drain. 1000 volts transistor MTH5N100 , VDD = 300V, LD = 188µH, VGG=10V, RG=50 Ohms, Vclamp = 850V, IDmax=3A.
(a) Measurement (b) Simulation

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

PHYSICAL BEHAVIOUR MODELLING OF VDMOS DEVICES

J. PAREDES, J. FERNANDEZ, F. BERTA, S. HIDALGO, J. REBOLLO and J. MILLAN

Centro Nacional de Microelectrónica, CSIC-UAB, SP-08193 Bellaterra, Barcelona, Spain

Résumé - Nous proposons un modèle analytique pour expliquer le comportement physique des transistors VDMOS quel que soit le niveau de courant. L'effect de quasi-saturation est pris en compte au travers de la saturation de la vitesse des porteurs à fort champ électrique. Des transistors VDMOS interdigités ont été réalisés et les simulations bidimensionelles effectueés en vue de vérifier le modèle.

Abstract - An analytical model is proposed in order to explain the physical behaviour of VDMOS devices at any DC current level. The quasi-saturation effect is included considering the carrier saturation velocity at high electric fields. Interdigitated VDMOS devices have been fabricated and 2D simulations have been carried out to check the model.

1 - INTRODUCTION

The vertical DMOS transistor is one of the most usual drivers in power MOS ICs /1/. The linear region of its I-V characteristic has been modelled considering a surface accumulated layer, the JFET resistance between cells and a bulk epilayer resistance together with the active channel /2/. Moreover, in previous works /3, 4/ the current pinching was responsible for the quasi-saturation (q-s) effect or limitation in the current handling capability typical at high current levels. In contrast, 2D simulations /5/ have not shown current pinching between cells but, contrarily, there is a majority carrier excess in the JFET region, suggesting that q-s effect is due to carrier velocity saturation. The carrier excess was explained by means of a dipole similar to the one formed in a JFET channel under carrier velocity saturation conditions. However, none of the previous models has quantitatively explained these phenomena.

This work is aimed to study the detailed physical behaviour of the VDMOS to avoid the tedious 2D CAD. Interdigitated VDMOS structures have been fabricated and 2D simulations /6/ have been carried out in order to check the proposed model.

2 - MODEL

Due to the axial symmetry of the VDMOS structure (see fig 1), current and electric field equations can be simplified to one-dimensional analysis:

$$\frac{I}{A(y)} = q\,\mu\left[n\,\mathcal{E} + \frac{kT}{q}\,\frac{dn}{dy}\right] \qquad (1)$$

$$\frac{d}{dy}\,[A(y)\,\mathcal{E}] = \frac{q}{\epsilon_s}\,A(y)\,(n_o - n) \qquad (2)$$

where n_o is the epilayer doping level and $A(y)$ is the cross-section area of the current path. Suitable approximations in the different regions of the epitaxy have been taken into account in order to obtain analytical solutions for the electric field distribution and for the majority carrier concentration; as shown in fig. 1, these regions are, namely: a) the surface accumulated layer induced by the gate effect of the MOS structure; b) the JFET region between cells and c) the drift region up to the drain contact.

a) SUPERFICIAL REGION. The limits of this region are the Si-SiO_2 interface and the null electric field point inside the epilayer (y_b), which is a consequence of the device bias; i.e., V_G and V_D are both positive while the source is grounded, so it must exist a point of

minimum potential between gate and drain. This region has been splitted into two zones since the electric field changes from very high values at the interface to zero at the other edge. In the first zone under the gate electrode, S_1, the electric field and the carrier concentration gradient are very sharp so that the drift (J_a) and the diffusion (J_d) currents are both much greater than the total current flowing through the VDMOS. Consequently, $n \gg n_o$ and $J_d = J_a$ are appropiate approaches, the solution being:

$$E = \left[\frac{1}{E_s} + \frac{q}{2kT} y \right]^{-1} \tag{3}$$

$$n = \frac{\epsilon_s}{2kT} E^2 \tag{4}$$

where E_s is the surface electric field.

The second zone, S_2, lies by the zero electric field point; therefore, J_a is small and a linear dependence on the electric field has been assumed, $J_a = q \mu n_b E$, n_b being the carrier concentration at the y_b point. Under these conditions the solution is:

$$E = A \exp(-ay) - \frac{I}{q \mu n_o \ell} \tag{5}$$

$$n = n_o + a \frac{\epsilon_s}{q} A \exp(-ay) \tag{6}$$

where A is an integration constant and $a = (n_b q^2/kT\epsilon_s)^{1/2}$. The boundary point of S_1 and S_2 zones, the constant A, and y_b are determined imposing the continuity condition of their solutions.

b) JFET REGION. This region spans from the edge of the former zone up to the point where current is no longer confined by the depletion layer of the body-epitaxy junction. This point is found assuming this junction to be plane in punchthrough mode at the onset of the junction curvature (see fig. 1). Obviuosly, the length of this region depends on the drain voltage but its width remains constant and equals the cell spacing (ℓ). In this sense, and according to this model, there is a modulation of the JFET lenght instead of a modulation of the JFET width as considered in previous works /3,4/.

In this region, the carrier concentration is close to n_o and consequently J_d is negligible. An accurate solution is:

$$n = n_o \frac{-1 + \left[1 + 4 \frac{n_b}{n_o} \left(\frac{n_b}{n_o} - 1 \right) \left[1 - \exp(-\Gamma (y - y_b)) \right] \right]^{1/2}}{2 \left(\frac{n_b}{n_o} - 1 \right) \left[1 - \exp(-\Gamma (y - y_b)) \right]} \tag{7}$$

$$E = \frac{I}{q \mu n \ell} \tag{8}$$

where $\Gamma = (q^2 \mu n_o \ell)/(\epsilon_s I)$. The carrier concentration (7) tends asymptotically to n_o at a rate which depends on Γ and, therefore, on the current level I. According to this analysis the tail of the accumulated layer can extend inside the epitaxy and it could even reach significant depths.

c) DRIFT REGION. Once the current flow leaves the JFET region, it spreads inside the epilayer with a 45º angle /2/ forming a trapezoidal zone followed by a rectangular zone if the epitaxy is thick enough. Solutions for this region are obtained similarly to the former case but considering a cross section A(y) linearly depending on the y coordinate in the trapezoidal

zone.

The above formulation is valid for electric fields below the saturation value, Esat. For higher fields we have supposed a carrier mobility $\mu(E) = v_{sat}/E$ according to /7/. The solutions of the general equations (1) and (2) under this condition in the different regions, with its respective approximations, are straightforward calculated.

Finally, the voltage drop across any region and its resistance are analytically obtained integrating the electric field distributions.

3 - RESULTS

Electric field distributions obtained from the proposed model and from 2D simulations are plotted on fig. 2, for V_D=5 and 20 V, the last value corresponding to the q-s mode. Note that the peak electric field moves inside the epitaxy as the drain voltage increases, showing the JFET lenght modulation effect predicted by the model. A good agreement between theory and 2D simulations is achieved not only inside the epitaxy but at the accumulated surface as shown in the insert.

Normalized carrier concentration is plotted on fig. 3. Let us remark that the tail of the accumulated layer extends up to the body depth even though the device is not in the q-s region. Similarly to the field distribution, the formulation fits the 2D simulations all through the epilayer.

The I-V characteristics of the fabricated devices are shown in fig. 4, where the gate voltage has been choosen high enough to reach the q-s region. Experimental as well as 2D simulations points are in good agreement with the model. On the same figure, the theoretical results without considering carrier velocity saturation are plotted. Note that this effect is just responsible for the current limitation when the transistor is under q-s conditions. The proposed model has been successfully checked as well for the VDMOS devices referred in /3,5/.

4 - CONCLUSIONS

An analytical model for VDMOS devices has been presented. It is based on the solution of the general semiconductor equations with suitable approximations for the different regions into which the epitaxial layer has been divided. This model accounts for the electric field and for the majority carrier concentration distributions which have been checked using BAMBI. This formulation also explains the quasi-saturation effect considering the carrier velocity saturation at high electric fields. Interdigitated VDMOS devices have been fabricated and their output characteristics are in good agreement with the proposed model as well as with the 2D simulations performed.

REFERENCES

/1/.- B.J. Baliga. IEEE Trans. Electron. Dev., ED-33, 1936 (1986).

/2/.- C. Hu, M. Chi and V.M. Patel. IEEE Trans. Electron. Dev., ED-31, 1693 (1984).

/3/.- J.L. Sanchez, M. Gharbi, H. Tranduc and P. Rossel. IEE Proc., 132, 42 (1985).

/4/.- J. Rebollo, J. Millán and F. Serra-Mestres. Proc. of the Symposium on High Voltage and Smart Power Devices. (The Electrochemical Society), 87-1, abs. 92, 130, Philadelphia (USA) (1987).

/5/.- M.N. Darwish. IEEE Trans. Electron. Dev., ED-33, 1710 (1986).

/6/.- A.F. Franz and G.A. Franz. IEEE Trans., CAD-4, 177 (1985).

/7/.- D.M. Caughey and R.E. Thomas. Proc. IEEE Let., 55, 2192 (1967).

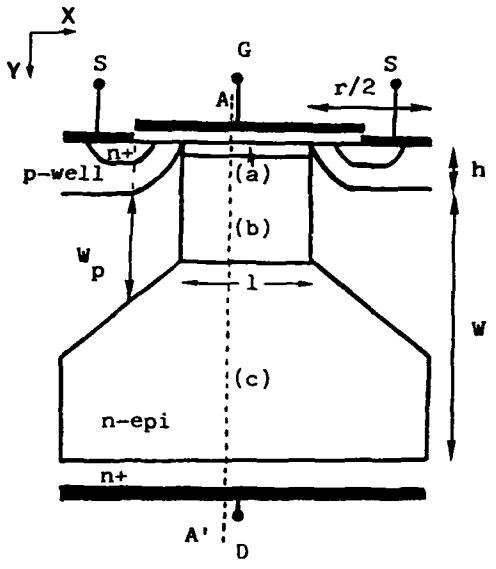


Fig. 1.- VDMOS cross-section. Fabricated structure parameters: W=40μm, h=5μm, r=60μm, l=10μm, n_o=2x10^{14} cm^{-3}, channel length = 2μm and channel width = 25 cm.

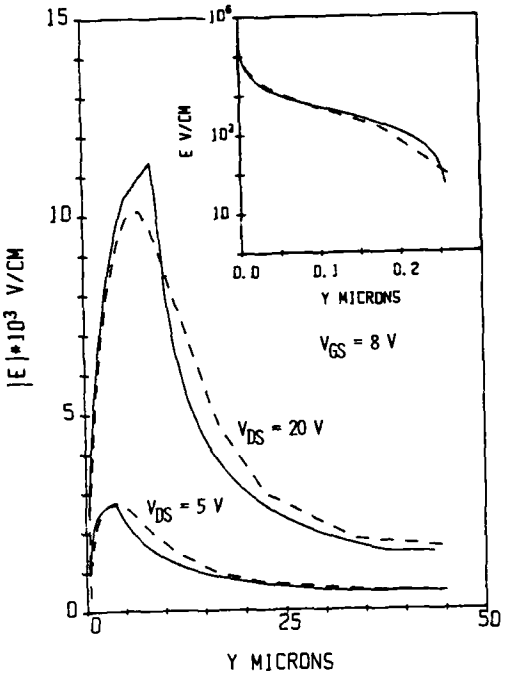


Fig 2.- Electric field along the AA' line
—— proposed model
---- 2D simulation
Insert: Electric field in the accumulated layer.

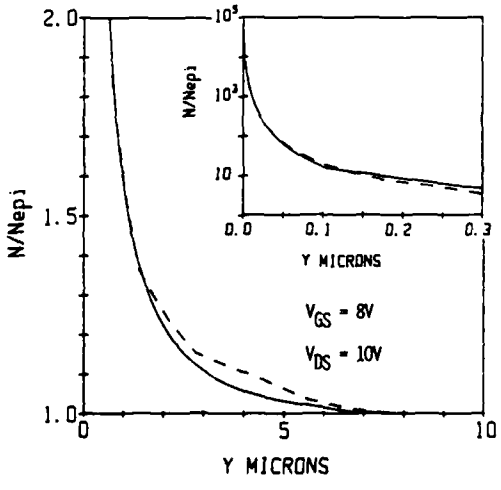


Fig 3.- Normalized carrier concentration along the AA' line before q-s.
—— proposed model
----- 2D simulation
Insert: Carrier concentration in the accumulated layer.

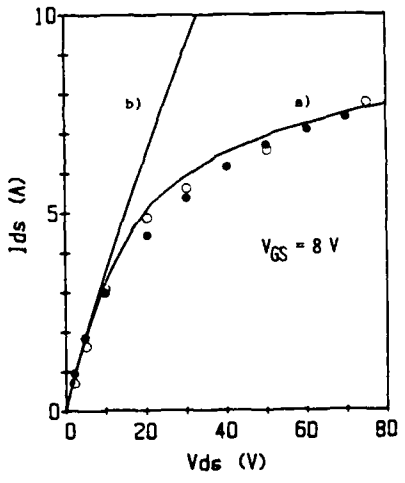


Fig 4.- I-V characteristics
—— model a) with / b) without carrier velocity saturation
○ 2D simulation
● experimental points

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

ON-RESISTANCE IN THE ALDMOST

G. NANZ[1], P. DICKINGER[2], W. KAUSEL[1] and S. SELBERHERR

Institut für Allgemeine Elektrotechnik und Elektronik, Technical University Vienna, Gusshausstrasse 27-29, A-1040 Vienna, Austria

Abstract — Recently a new lateral power MOSFET named *accumulation lateral DMOS* transistor (ALDMOST) has been proposed. We have investigated the dependence of the ON-resistance of this type of device on the oxide thickness and the additional semi-insulating layer along the surface of the gate oxide above the drift region. This layer has been introduced in order to lower the high ON-resistance which is in general a disadvantage of this type of MOS transistors.

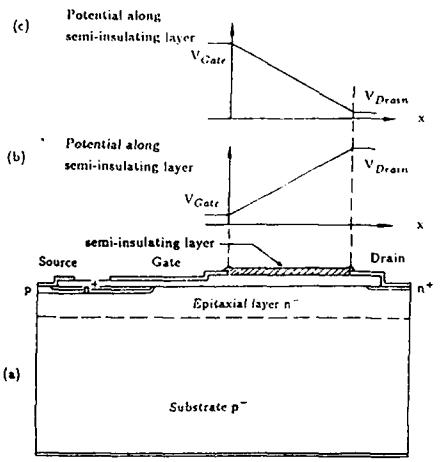


Fig. 1 Geometry and voltage distribution along semi-insulating layer
(a) Geometry
(b) Voltage distribution in OFF-condition
(c) Voltage distribution in ON-condition

(1) Supported by Siemens AG, Munich, West Germany
(2) Supported by Siemens AG, Villach, Austria

1 - INTRODUCTION

In June 1987 a new lateral power MOSFET structure has been presented /1/. It is a modified lateral DMOS device with a semi-insulating poly-Si (SIPOS) layer along the surface of the gate oxide above the drift region. This additional layer has been introduced to lower the high ON-resistance of this type of device which is in general a disadvantage of high voltage power MOS transistors.

We have performed a comparison between a conventional LDMOS structure which we have investigated previously /2/ and the ALDMOST /3/. The geometry of our device and the voltage distribution along the semi-insulating layer in ON- and OFF-condition can be seen in Fig. 1. It is an n-channel device with a semiconductor area of $82\mu m \cdot 49\mu m$. A significant decrease of the ON-resistance could be observed for the ALDMOST.

We have simulated the behaviour of three different devices: First we have taken the geometry from Fig. 1 with an variable oxide thickness of $0.4\mu m$ and $0.2\mu m$, respectively. Since in /1/ an oxide thickness of $0.1\mu m$ or less is requested for a significant gain in efficiency we have reduced the oxide thickness to $0.2\mu m$ and 0.1μ (constant, without step). The doping profile is approximated by Gaussian distribution functions (maximum values: n^+: $2.0 \cdot 10^{20}$ cm^{-3}, p: $5.0 \cdot 10^{16}$ cm^{-3}, n^-: $3.0 \cdot 10^{14}$ cm^{-3}, p^-: $1.2 \cdot 10^{14}$ cm^{-3}).

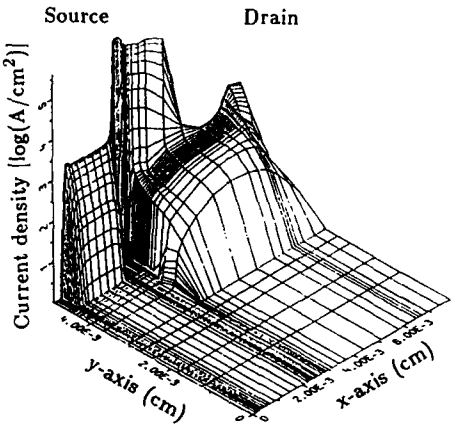


Fig. 2 Current density in ON-condition LDMOST
(V_{Drain} = 12.0V, V_{Gate} = 15.0V, oxide thickness $0.2\mu m$)

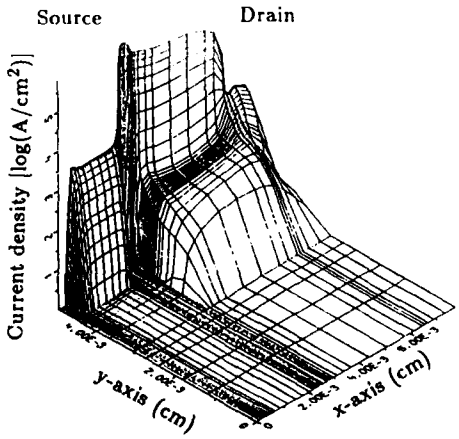


Fig. 3 Current density in ON-condition ALDMOST
(V_{Drain} = 12.0V, V_{Gate} = 15.0V, oxide thickness $0.2\mu m$)

2 - ON-RESISTANCE

The conventional LDMOS transistor has an oxide of variable thickness as it is shown in Fig. 1. A minimum oxide thickness is requested to guarantee the voltage stability by limiting the electrostatic field in the dielectric. This thickness also strongly influences the threshold voltage of the device.

For an oxide thickness of 0.4μm the threshold voltage will be approximately 6.0V, for 0.2μm the threshold voltage will only slightly change, for 0.1μm the threshold voltage will be about 2.2V.

The ON-resistance of the device mainly depends on the doping and the length of the drift region. High drift region doping lowers the ON-resistance but it results in punch-through /2/, a short drift region will lower the breakdown voltage /4/.

By the positive bias of the semi-insulating layer a strong electron accumulation under the drift region oxide will be enabled thus reducing the ON-resistance. Furthermore an additional path for the current flow close to the interface of the dielectric and the semiconductor region will be provided.

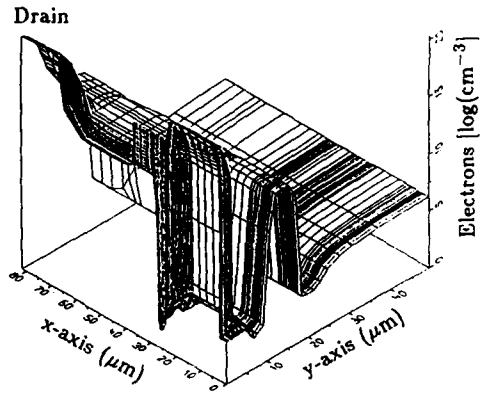


Fig. 4 Electron concentration LDMOST
(V_{Drain} = 5.0V, V_{Gate} = 6.0V, oxide thickness 0.1μm)

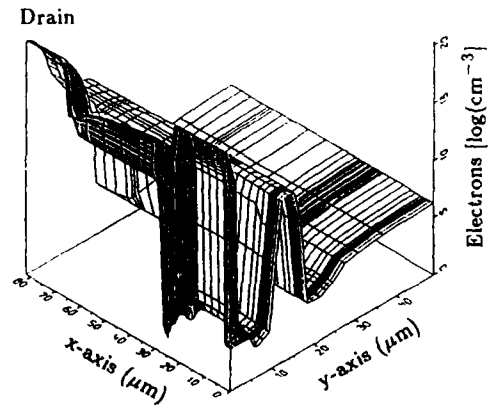


Fig. 5 Electron concentration ALDMOST
(V_{Drain} = 5.0V, V_{Gate} = 6.0V, oxide thickness 0.1μm)

3 - NUMERICAL RESULTS

The mechanism of the SIPOS layer has been discussed in /5/. We have handled the SIPOS layer as an additional Ohmic contact at the surface of the oxide between the gate and the drain contact with a linear variation of the applied potential with respect to space. It should be remarked that this smooth transition of the boundary conditions furthermore enhances the convergence speed of the Newton cycle.

During the simulation it has turned out that even for an oxide thickness of 0.4 μm above the drift region the ON-resistance is reduced by a factor of 1.5. This gain of efficiency will be increased for an oxide thickness of 0.2μm and 0.1μm. According to our results the ON-resistance will be lowered in these devices by a factor of 2 and more.

In Fig. 2 and Fig. 3 the total current densities of a conventional LDMOST and the ALDMOST, respectively, with an oxide thickness of 0.2μm are shown (V_{Gate} = 15V, V_{Drain} = 12V). It can clearly be seen how the current density leaks far into the drift region because of the additional bias provided by the semi-insulating layer above the depletion region.

In Fig. 4 and Fig. 5 the electron concentrations of the two devices are shown (oxide thickness: 0.1μm, V_{Gate} = 6V, V_{Drain} = 5V). The new path for the current flow because of the strong electron accumulation along the interface between the dielectric and the semiconductor can nicely be seen.

Our computations have been carried out with our two-dimensional device simulator BAMBI which solves the three basic semiconductor equations simultaneously in a totally selfconsistent way utilizing a 'Finite Boxes' grid /6/.

4 - CONCLUSIONS

The additional SIPOS layer provides a significant reduction of the ON-resistance in lateral DMOS transistors. Even for an oxide thickness greater than 0.1μm there is a gain in efficiency. Because the avalanche breakdown behaviour will almost not be changed by the SIPOS layer /3/ the lower ON-resistance of the ALDMOST is a real advantage compared to a conventional lateral DMOS transistor.

ACKNOWLEDGEMENTS
This work was supported by DIGITAL EQUIPMENT COPP. at Hudson, U.S.A., and by the 'Fonds zur Förderung der wissenschaftlichen Forschung', project S43/10. The authors are indebted to Prof. H. Pötzl for many helpful discussions.

REFERENCES

/1/ Habib, S.E.-D., *The ALDMOST: A new power MOS transistor*, IEEE Electron Device Letters EDL-8, No. 6, pp. 257-259 (1987)

/2/ Nanz, G., Dickinger, P., Kausel, W., Selberherr, S., *Punch-through in resurf devices*, Proc. of the AMSE-conference at Karlsruhe 1987, pp. 63-70

/3/ Nanz, G., Dickinger, P., Kausel, W., Selberherr, S., *Avalanche Breakdown in the ALDMOST*, to be presented at SISDEP 1988, September 26–29, 1988, Bologna, Italy

/4/ Colak, S., Singer, B., Stupp, E., *Lateral DMOS power transistor design*, IEEE Electron Device Letters EDL-1, No. 4, pp. 51-53 (1980)

/5/ Mukherjee, S., Chou, C.J., Shaw, K., McArthur, D., Rumennik, V., *The effect of SIPOS passivation on DC and switching performance of high voltage MOS transistors*, IEDM Tech. Dig., pp. 646-649 (1986)

/6/ Selberherr, S., *Analysis and simulation of semiconductor devices*, Wien, New York: Springer 1984

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

ULTRAFAST SINGLE PHOTON DETECTOR WITH DOUBLE EPITAXIAL STRUCTURE FOR MINIMUM CARRIER DIFFUSION EFFECT

A. LACAITA, M. GHIONI and S. COVA

Politecnico di Milano, Dipartimento di Elettronica and Centro di Elettronica Quantistica e Strumentazione Elettronica CNR, Piazza L. da Vinci 32, I-20133 Milano, Italy

Résumé. Photodiodes à avalanche specifiquement projetés pour travailler avec tension de polarisation superieure à la tension de breakdown sont employés pour détecter photons isolés et mesurer leur temps d'arrivée avec résolution à picosecondes. La fonction de résolution temporelle de ces dispositifs (dénommés Single Photon Avalanche Diodes SPADs) est composée d'un pic étroit et d'une queue indésirable, due à diffusion de porteurs de charge. Nous avions précédemment projeté et fabriqué un dispositif avec une structure épitaxiale, qui avait reduit l'effet de diffusion à une brève queue de forme exponentielle, avec constante de temps 1.7 ns. En vue d'obtenir un ulterieur progrès, nous avons ideé et fabriqué une structure à double épitaxie. Les mesures effectuées avec ces dispositifs ont démontré une drastique réduction de l'effet de diffusion. La fonction de résolution est characterisée par une largeur a mi-hauteur du pic de 45 ps et par une très brève queue exponentielle, avec constante de temps 270 ps.

Abstract. Avalanche photodiodes designed to work biased above the breakdown voltage are employed for timing single photons with picosecond resolution. The time resolution curve of these devices (called Single Photon Avalanche Diodes SPADs) shows a fast peak and an undesirable tail, due to carrier diffusion. We had, previously, designed an epitaxial SPAD structure that succeeded in reducing the effect to a short, exponential-like tail with 1.7 ns lifetime. In order to achieve a further improvement, we have devised and fabricated a double epitaxial structure. The measurements performed on these new devices show a drastic reduction of the diffusion effect. The time resolution curve of the new detector is characterized by a peak with better than 45 ps full-width at half-maximum (FWHM) and a short, exponential tail, with 270 ps lifetime.

I. INTRODUCTION

Single Photon Avalanche Diodes (SPADs) are used as ultrafast detectors in single photon counting techniques. They are p-n junctions with uniform breakdown voltage and low density of traps and generation centres in the depletion layer (1-8). The devices work biased above the breakdown voltage in the triggered avalanche mode. In this condition, the electric field at the active junction is high enough to sustain the flowing of an avalanche current. However, the device remains quiescent and the inverse current is pratically zero until a carrier succeedes in triggering the avalanche. The current then rises to a detectable value; if the carrier is photogenerated, the leading edge of the avalanche current is synchronous with the arrival time of the photon. The self-sustaining avalanche current continues to flow until an external circuit quenches the diode by lowering the bias voltage close to or below the breakdown voltage. After a finite dead time, the working voltage of the device is finally restored, in order to make possible the detection and timing of another photon. The avalanche can also be triggered by carriers thermally generated or emitted from deep levels in the depletion layer of the active junction. The corresponding dark count rate must be minimized for avoiding impairment of the device performance. Careful processes are employed in the fabrication of the devices for minimizing the concentration of generation centres and deep levels in the depletion layer. In order to completely avoid the degradation of the device resolution due to the dark count rate, Cova and Longoni devised the active quenching method (6-8). By using active quenching circuits, SPADs can work in accurately controlled bias conditions and the dead time following each pulse can be made very short. Remarkable performance is thus achieved in various applications, such as optical fiber characterization, laser ranging, fluorescent decay analysis, etc. (8). The SPADs are, therefore, a solid state alternative to photomultiplier tubes, providing wider spectral sensitivity range, higher resolution in the measurement of the photon arrival time, low detector noise.

The time resolution curve of a single photon detector is given by the statistical distribution of the delay from the true arrival time to the actual detection time of the photon (6-10). It is measured in a time-correlated single photon counting set-up (10) employing picosecond optical pulses.

We had previously implemented and tested prototype SPAD devices, having the geometry described by Haitz (2). The active junction was formed by a shallow (0.3 μm) n^+ layer in a p bulk substrate. A deep

diffused n^- guard ring surrounded it, avoiding edge breakdown. The resolution curve of the prototype SPAD devices is characterized by a peak and a slow tail. The fast peak is due to photons absorbed in the depletion layer and its width is mainly related to the statistical fluctuations in the avalanche build-up time. We observed and reported full-width at half-maximum (FWHM) values down to 60 ps (6-8). The experiments also showed that the time resolution improves by increasing the maximum electric field at the active junction (8). The tail is due to the minority carriers, photogenerated in the neutral region beneath the junction, that succeed in reaching the depletion layer by diffusion. Due to the wavelength dependence of the optical absorption coefficient, the intensity and the duration of the diffusion tail change with the wavelength of the detected photons. This effect represents a serious drawback to high resolution measurements of fast optical signals. In fact, the true signal waveform is obtained by a deconvolution analysis of the experimental data, which requires a very accurate knowledge of the shape of the resolution curve. In all cases where the optical signal is non-monochromatic, the actual resolution shape depends on the actual detected optical spectrum. Therefore, it may be impossible or at least unpractical to obtain it with sufficient accuracy.

II. THE PREVIOUS EPITAXIAL STRUCTURE

In order to improve the time response of the detector, the diffusion tail must be reduced. Our approach was to design epitaxial device structures in which the substrate acted as a sink for photogenerated carriers and the active junction was located in a suitable epistrate. A Monte Carlo program, previously developed in our laboratory (9), was used to compute the time-dependent diffusion effect in the devised structures and, consequently, to define the actual device design.

In Fig. 1, a schematic cross section of the previous single epitaxial structure is shown. The device was fabricated in a p epitaxial layer over a n substrate (11). In this geometry the electrons photogenerated in the substrate cannot reach the epilayer. Furthermore, the reverse biased substrate-epistrate n-p junction competes with the active junction in collecting the electrons generated in the neutral p epilayer. The device structure is quite different from the previous non-epitaxial design, since the avalanche current flows now laterally to the side ohmic contact. The series resistance of the device depends, therefore, on the thickness of the neutral epilayer beneath the guard ring. In order to avoid impairing the output signal, a few kOhm value should not be exceeded. This requirement sets a limit to the reduction of the epilayer thickness and, therefore, to the reduction of the diffusion effect. A shallow guard ring (2 μm) was, therefore, employed and a 12 μm thick epistrate was adopted. In order to achieve an improvement in the time resolution of the devices, the breakdown electric field was increased. A boron implantation increased the p doping in the active area, lowering the breakdown voltage down to 13 V. At the operative bias the maximum electric field attained values higher than 550 kVolt/cm.

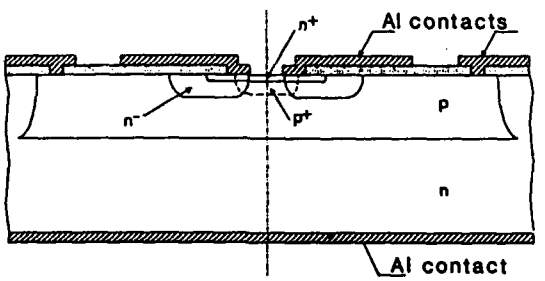


Fig.1 - Schematic cross section of the previous epitaxial SPAD device.

The experiments fully confirmed the reduction of the diffusion tail, in excellent agreement with the Monte Carlo simulation.The tail is exponential-like and its lifetime is 1.7 ns. A remarkable improvement was also found in the peak FWHM. At 6 V excess bias, the device resolution is better than 30 ps at room temperature and 20 ps at -70 C. To our knowledge, this is the highest time resolution so far reported in single-photon-timing measurements. It is worth noting, however, that the very high values of the maximum electric field employed entail a significant drawback. The dark count rate is strongly enhanced, due to the Frenkel-Poole effect and to the phonon-assisted tunneling effect, that enhance the carrier emission from generation centers and deep levels in the depletion layer (12). In fact, at room temperature the dark count rate of the devices was remarkably higher than that of the prototype SPADs, attaining several kpps already at 1 V excess bias and rapidly increasing with the bias voltage. Therefore, the high resolution of these SPADs can be fully exploited only by using active quenching arrangements.

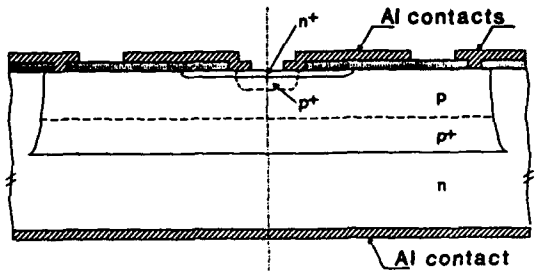


Fig.2 - Schematic cross section of the new, double epitaxial SPAD device.

III. THE DOUBLE EPITAXIAL STRUCTURE.

A further strong reduction of the diffusion tail requires to achieve a further reduction of the neutral epilayer thickness, but without increasing the series resistance beyond a few kOhms. The simple adoption of a thinner, higher-doped p epilayer would imply a marked reduction of the guard-ring breakdown voltage. The excess bias that could be applied to the active area would thus be reduced, and the output signal and the performance of the detector would correspondingly be impaired. We have, therefore, devised a new SPAD structure with two different p epilayers grown over the n substrate (Fig. 2). The buried epilayer, which is only 2 μm thick, provides a low resistivity path (0.3 Ohmcm) to the side contact. The active n^+p junction is built in the upper, low-doped p epistrate (10 Ohmcm), which is only 2.5 μm thick. In order to reduce as far as possible the epilayer thickness, a different type of guard ring is used to dilute the electric field at the device periphery. Instead of using a deep n diffusion, a "virtual" guard ring structure is implemented: a boron implantation in the central part of the active junction defines the sensitive area of the detector. The low p doping in the outer part of the n^+p junction and a field plate raise the edge breakdown voltage up to 50 V. Devices with this geometry have been fabricated with different diameters of the sensitive area, in the range from 8 to 20 μm. The boron implantation was deliberately made lighter than that used in the previous epitaxial devices, so that a slightly higher breakdown voltage was obtained, about 16 V. This limited the maximum electric field, thus avoiding the enhancement of the dark count rate at the cost of a slightly lower time resolution.

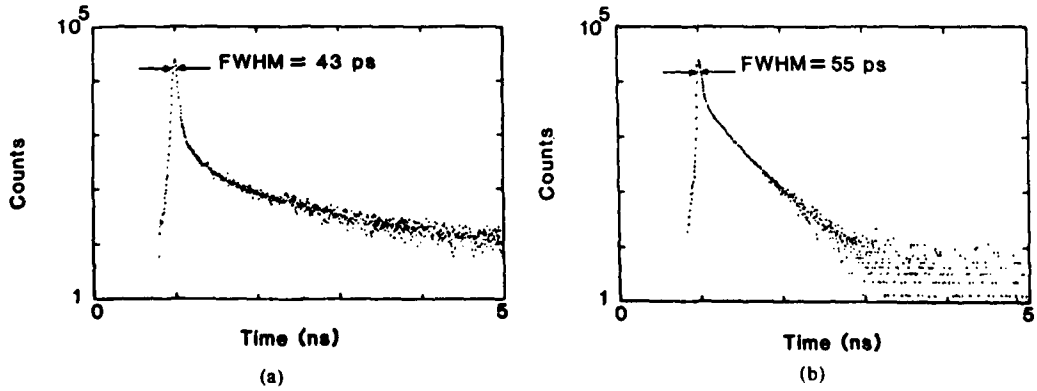


Fig.3 - Optical pulse of a gain-switched laser diode (13) as measured in a time-correlated single-photon counting set-up with: a) the previous epitaxial SPAD; b) the new double epitaxial SPAD.

The experimental tests confirm the expected behaviour of the double epitaxial devices. Fig. 3a shows, in logarithmic scale, the pulse of a gain-switched laser-diode (13) measured using the previous epitaxial SPAD. As above mentioned, the diffusion tail has 1.7 ns lifetime. The peak width of 43 ps FWHM is due to the laser peak width (28 ps FWHM), to the detector time resolution (28 ps FWHM) and to the electronic noise and jitter (15 ps FWHM). The same measure performed with the double epitaxial SPAD is reported in Fig. 3b. It is characterized by a slightly larger peak width (55 ps FWHM) and by a much shorter diffusion tail. The actual value of the detector time resolution is 45 ps FWHM. It is evaluated by quadratically subtracting the 28 ps laser FWHM and the 15 ps electronic time jitter FWHM from the 55 ps FWHM of the experimental curve. The diffusion tail is exponential, with only 270 ps lifetime. In comparison to the previous epitaxial device, however, the initial tail intensity is higher. This behaviour is in agreement with the results of the Monte Carlo simulation. It can be intuitively explained, by taking into account the different geometry of the two devices. In the epitaxial device, the neutral region adjacent to the active junction had a longer, narrow tubular shape, with side walls acting as sinks for the diffusing carriers. In the new, double epitaxial device, the corresponding neutral region is thinner, but has a more open, almost planar shape. The carriers generated in close proximity to the active junction have, therefore, a higher probability of reaching it. These carriers contribute to the first part of the tail.

IV. CONCLUSIONS.

It can be concluded that the double epitaxial SPAD structure succeedes in producing a marked shortening of the diffusion tail. The physics of the phenomenon is well understood. The question naturally arises whether the tail can be further reduced or, better, completely eliminated. Work aiming to this goal is in progress in our laboratory. Anyhow, the high time resolution and the well-behaved, short diffusion tail of the device structures so far fabricated and reported make the Single Photon Avalanche Diodes an attractive solid-state alternative to photomultiplier tubes in all applications of the single-photon counting techniques that do not require large sensitive area detectors.

ACKNOWLEDGMENTS

The authors wish to thank the research and development laboratory of SGS-Thomson Microelectronics, Castelletto-Milano, for the fabrication of the epitaxial devices.

This work was supported in part by Consiglio Nazionale delle Ricerche under programs "PF-MADESS Materiali e Dispositivi per Elettronica allo Stato Solido" and "PS Fotorivelazione" and by Ministero Pubblica Istruzione.

REFERENCES

1. A. Goetzberger, B. Mc Donald, R.H. Haitz and R.M. Scarlett, J.Appl.Phys., vol. 34, pp. 1591-1600, 1963.
2. R.H. Haitz, J.Appl.Phys., vol. 36, pp. 3123-3131, 1965.
3. H. Melchior, A. Goetzberger. E. Nicollian and W.T. Lynch, Solid State Electron., vol.12, pp. 449-461, 1969.
4. R.J. McIntyre, IEEE Trans. Electron Devices, vol. ED-19, pp. 703-713, 1972.
5. W.O. Oldham, R.R. Samuelson and P. Antognetti, IEEE Trans. Electron Devices, vol. ED-19, pp. 1056-1060, 1972.
6. S.Cova, A. Longoni and A. Andreoni, Rev.Sci.Instrum., vol. 52, pp. 408-412, 1981.
7. S.Cova, A. Longoni, A. Andreoni and R. Cubeddu, IEEE J.Quantum Electron., vol. QE-19, pp. 630-634, 1983.
8. S.Cova, G.Ripamonti and A.Lacaita, Nucl.Instr. and Meth., vol.A253, pp. 482-487, 1987.
9. G. Ripamonti and S. Cova, Solid State Electron., vol. 28, pp. 925-931, 1985.
10. D.V. O'Conner and D. Phillips, "Time-Correlated Single Photon Counting," Academic Press, London and New York, 1983.
11. M.Ghioni, S.Cova, A.Lacaita and G.Ripamonti, IEEE Int. Electron Device Meeting IEDM 87, Washington D.C., December 6-9, 1987, Technical Digest, pp. 452-455.
12. G.Vincent, A.Chantre and D.Bois, J.Appl.Phys., vol. 50, pp. 5484-5487, 1979.
13. Picosecond Light Pulser C1308, 833 nm wavelength, Hamamatsu Tv. Co. Ltd.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A MINIATURIZED APPROACH TO THE CRYOELECTRONIC MAGNETIC FIELD EFFECT TRANSISTOR

W. CLAUSS, J. PARISI, J. PEINKE, U. RAU and R.P. HUEBENER

Physikalisches Institut, Lehrstuhl Experimentalphysik II, Universität Tübingen, Morgenstelle 14, D-7400 Tübingen, F.R.G.

Résumé - Le contrôle magnêtique du courant filamentaire pendant le percement avalanchal dans semiconducteurs à basses temperatures représente un principe d'opération intéressant pour le dessin d'un transisteur basé sur l'effet d'un champ magnétique. Nous avons vérifié experimentalement des configurations réduites en utilisant des lines superconductives pour la production d'un champ magnétique et de tous les intermédiaires. L'évaluation des performances caractéristiques importantes de notre conception du transisteur paraît être très prometteuse.

Abstract - Magnetic control of the filamentary current flow during avalanche breakdown in bulk semiconductors at low temperatures represents an interesting operation principle for the design of a magnetic field effect transistor. We have experimentally verified miniaturized device configurations using superconducting lines for the generation of the magnetic control field and all interconnections. The evaluation of the important performance characteristics of our transistor concept looks highly promising.

1 - INTRODUCTION

Electrically driven into the regime of low-temperature avalanche breakdown, homogeneously doped semiconductors have been demonstrated to display spatially inhomogeneous current flow via high-conductivity filament channels generated by impurity impact ionization /1-3/. If an external magnetic field is applied perpendicular to the current direction, the current filaments are structurally changed due to the Lorentz force acting upon the moving charge carriers. The corresponding shift of the current-voltage characteristic (magnetoresistance) results from the "cooling effect" of the magnetic field on the hot charge carriers. At low temperatures where the carrier mobility and, hence, the Lorentz force is high, a relatively large magnetoresistance is obtained with applied magnetic fields as small as only a few Gauss. The sensitive response to the magnetic control field offers an interesting principle for the design of a cryoelectronic device family.

As a first experimental approach, we have recently demonstrated the basic concept of a magnetic field effect transistor at liquid-helium temperatures, using single-crystalline p-doped germanium with an acceptor concentration of about 10^{14} cm^{-3} and the typical dimensions of 0.1 x 5 x 6 mm^3 /4,5/. The magnetic control of the filamentary current flow was achieved by a small magnetic field (10 - 100 G), generated by superconducting lines attached to the surface of the semiconductor crystal. In this way, a three-terminal device is obtained, based on a hybrid concept implementing semiconducting and superconducting components. Particularly attractive features of this configuration are its simplicity, the utilization of well-established semiconductor and superconductor technology, and its robust nature, promising a high production yield. Moreover, our device concept may become essential in the development of ultrafast electronic circuits, requiring high packing density and correspondingly lower power dissipation.

In order to evaluate the limitations of the important performance characteristics, we have experimentally performed a miniaturized approach to the cryoelectronic magnetic field effect transistor. Using different sample geometries, the distance between the ohmic contacts was reduced systematically, ending up with a minimum value of 10 µm. As expected from a preliminary estimate /6/, the limitations due to the time of flight per length of the carriers and due to the power losses per distance between the ohmic contacts were found to be about 50 ps/µm and 0.5 µW/µm, respectively. Thus, the miniaturized sample geometry enables operation up to frequencies in the GHz regime.

2 - EXPERIMENTAL METHODS

The samples investigated were prepared from homogeneously p-doped germanium single-crystal slices.(1) The specific resistivity at room temperature ranged from 10 to 20 Ωcm, corresponding to an impurity concentration of (2 - 3) x 10^{14} cm^{-3}. The compensation ratios were definitely smaller than 5 x 10^{-2}. The semiconductor samples (of about 0.1 mm thickness) carried properly arranged ohmic contacts placed on one of the two broad surfaces. The distance between the contacts varied from 0.5 mm to 10 μm. The ohmic contacts were fabricated either by alloying (aluminum) or by ion implantation (boron) /7/. To provide these contacts with an electric field, a voltage bias was applied to a series combination of the sample and an 1 Ω load resistor. The sample current was obtained from the voltage drop at the load resistor. An external magnetic field perpendicular to the broad sample surfaces could be generated either by a superconducting coil surrounding the semiconductor sample or by a superconducting loop attached directly to the surface of the sample. Shaped in the form of a long hairpin, the superconducting niobium lines were fabricated by standard thin-film technology, and electric separation from the underlying semiconductor was achieved by a thin-film insulator. During the experiments, the samples investigated were kept in direct contact with the liquid-helium bath at 4.2 K and were protected against external visible and far-infrared irradiation by an appropriate copper shield.

3 - EXPERIMENTAL FINDINGS AND DISCUSSION

As already mentioned above, the operating principle of the present transistor device is based on the magnetic control of the filamentary current flow during low-temperature avalanche breakdown via the magnetoresistive effect. Since the magnetic field needed for magnetic control is found to be inversely proportional to the mobility of the charge carriers, the required high magnetic field sensitivity of the current flow can only be attained at high carrier mobility. With reference to the experimental verification of extremely high carrier mobilities and drift velocities (of some 10^5 cm^2/Vs and some 10^6 cm/s, respectively)/8/, adequate magnetic control can be achieved already with small magnetic fields in the range of 100 G. In the following, we briefly outline a quantitative evaluation of the most important performance characteristics based on an experimental realization of distinct miniaturized device geometries.

We emphasize that the two essential preconditions of our transistor scheme, namely, first, the appearance of the avalanche breakdown effect and, second, the magnetic control of the current flow in the breakdown regime, have been verified for all samples investigated. For relatively large distances between the ohmic contacts in the range of some 100 μm, differently prepared contacts always led to more or less the same results. But for distances smaller than 100 μm, only samples with ion-implanted contacts showed the expected linear dependence of the breakdown voltage on the contact distance. (Note that such miniaturized geometries are already influenced by the less sharp-edged structure of the heat-treated alloy contacts). By extrapolating the linear voltage versus distance characteristic to zero contact distance, we always found an excess breakdown voltage of about 30 mV. Accordingly, the breakdown electric field strength at small contact distances increases with decreasing distance, ending up with an extrapolated divergence at zero distance. We feel that the existence of an excess breakdown voltage can be directly attributed to the finite ionization energy of the shallow impurities (in the range of 10 meV). Note that avalanche breakdown only occurs if the conduction carriers have gained sufficient kinetic energy from the applied electric field, which at least must be of the order of the impurity ionization energy.

Now we turn to the performance limitations of our transistor scheme due to the time of flight per length of the carriers and due to the power losses per distance between the ohmic contacts. From measured drift velocities of a few 10^6 cm/s in the post-breakdown region /8/ we derive values of about 50 ps/μm for the time of flight per unit length. This estimate agrees reasonably well with the experimentally observed frequency limits of about 10 MHz up to a few GHz for samples with lengths ranging from 0.5 mm to 10 μm, respectively. We have checked that the above high-frequency limits are definitely not influenced by the capacities and inductivities of the electronic circuit configuration, as expected /6/. With the typical electric field of about 5 V/cm for establishing avalanche breakdown and typical current values of about 1 mA in the immediate post-breakdown region, we obtain for the power losses per sample length values of about 0.5 μW/μm, independent of frequency. Taking further into account the increased electric breakdown field at contact distances smaller than 50 μm, the power losses accordingly increase by a factor of five at most. From this we conclude that miniaturization of the transistor device appears to be limited reasonably at the smallest distance of the ohmic contacts chosen (10 μm).

---

(1) Device applications based on extrinsic silicon and gallium arsenide semiconductors are presently also investigated, but not reported here.

Finally, we point out that the functioning of the present transistor scheme can also be confined by noise problems. On one hand, thermal noise is expected to be small due to the operation at liquid-helium temperatures. But, on the other hand, the possible contribution of additional noise generated by the avalanche process /9/ requires further experimental studies and perhaps additional optimization procedures. Crosstalk between the elements of large integrated systems can be minimized by properly arranged superconducting screening lines. Here important concepts can be taken from the well-established Josephson technology. Nevertheless, unwanted stray magnetic und electric coupling sets a fundamental limit to further miniaturization of all presently used circuitry, that is likely to be reached before the limits imposed in the future by modern microfabrication techniques.

4 - CONCLUSIONS

The present work deals with the first experimental realization of a cryoelectronic magnetic field effect transistor based on a novel hybrid configuration of both semiconducting and superconducting components. We have modeled and characterized miniaturized solid state devices using extrinsic germanium semiconductors together with superconducting niobium lines for generating the magnetic control field an for all interconnections. The evaluation of the important performance characteristics of our transistor concept looks very promising. The power losses of such a device are similar to those in a superconducting Josephson junction, representing the high-speed device with the lowest power consumption up to now. We take note of the fact that superconducting wiring can preserve extremely fast pulse rise times required in ultrafast electronic systems. In this way, combining semiconducting device components operating at low temperature with superconducting lines for magnetic control and for all interconnections represents a highly promising new direction which has been neglected so far and which definitely deserves much more consideration.

ACKNOWLEDGMENTS

The authors thank E.E. Haller, W. Prettl, B. Röhricht, and K.M. Mayer for conversations and H. Leier, A. Forchel, and M. Pilkuhn (Microstructure Laboratory of the University of Stuttgart) for providing ion-implanted sample contacts.

REFERENCES

/1/ Mayer, K.M., Gross, R., Parisi, J., Peinke, J., and Huebener, R.P., Solid State Commun. 63 (1987) 55.

/2/ Mayer, K.M., Peinke, J., Röhricht, B., Parisi, J., and Huebener, R.P., Physica Scripta T19 (1987) 505.

/3/ Mayer, K.M., Parisi, J., and Huebener, R.P., Z. Phys. B - Condensed Matter (to be published).

/4/ Mannhart, J., Huebener, R.P., Parisi, J., and Peinke, J., Solid State Commun. 58 (1986) 323.

/5/ Mannhart, J., Parisi, J., Mayer, K.M., and Huebener, R.P., IEEE Trans. Electron Devices ED-34 (1987) 1802.

/6/ Mannhart, J. and Huebener, R.P., J. Appl. Phys. 60 (1986) 1829.

/7/ Peinke, J., Parisi, J., Röhricht, B., Mayer, K.M., Rau, U., Clauß, W., Huebener, R.P., Jungwirt, G., and Prettl, W., Appl. Phys. A (to be published).

/8/ Parisi, J., Rau, U., Peinke, J., and Mayer, K.M., Z. Phys. B - Condensed Matter (to be published).

/9/ Röhricht, B., Wessely, B., Parisi, J., and Peinke, J., Appl. Phys. Lett. 48 (1986) 233.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

FILTER STRUCTURE FOR SUB-MICRON FILTRATION FABRICATED IN SILICON

G. KITTILSLAND and G. STEMME

Department of Solid State Electronics, Chalmers University of Technology, S-412 96 Gothenburg, Sweden

Resume - Design och tillverkningen av ett nytt partikelfilter för filtrering i det sub-mikrona området beskrivs. Filtreringsprincipen är baserad på två, i förhållande till varandra, lateralt förskjutna hålmembran genom vilka fluiden måste passera. Tjockleken hos de stöd av kiseldioxid som separerar hålmembranen definerar storleken på de största partiklar som kan passera genom strukturen. Filterstrukturen tillverkades i en särskild tvåstegs självlinjerande process där ett hålmönster, hård bordopning, anisotrop kiseletsning och kiseldioxidunderetsning utgjorde de viktigaste processtegen. Det faktum att filterstrukturen elektriskt ugör en kondensator möjliggör andra applikationer. Strukturer med membranavstånd på 50 och 200 nm har tillverkats.

Abstract - The design and fabrication of a new particle filter in the sub-micron range are described. The filtration principle is based on two laterally displaced silicon hole membranes, through and between which the fluid has to pass. The thickness of the silicondioxide spacers separating the hole membranes defines the size of the largest particles which can pass through the structure. The filter structure was fabricated using a special two-step self-aligning technique where one single hole pattern, heavy boron diffusion, anisotropic silicon etching and silicondioxide undercut-etching constituted the most important process steps. The fact that electrically the filter structure is a capacitor makes other applications possible. Structures with membrane separation distances of 50 and 200 nm have been made.

INTRODUCTION

There is a great demand for filters and separation techniques in the sub-micron range. One example is the microelectronic industry where ultra-pure gases and liquids are neccesary in order to achieve a high production yield. Other examples are in the field of medicine and biotechnology.

Here a new type of particle filter structure in silicon for sub-micron filtration is presented [1]. Below some of the most important features are listed. Silicon provides a structure which is mechanically very strong and chemically and thermally inert and stable. The filter structure withstands aggressive environments such as acids and solvents and temperatures of more than 1000 °C. These characteristics are important for both use and regeneration of the filter. The special two-step self-aligning technique simplifies the fabrication of the filter structure.

FABRICATION

Fabrication of the structure is based on micromachining of silicon and a special two-step self-aligning technique which makes it possible to define the two membranes and the membrane spacers from a single hole pattern.

A sandwich-structure of 50 to 200 nm silicondioxide, 1 μm polysilicon and 1 μm silicondioxide was created on (100)-silicon wafers using thermal oxidation and LPCVD, as shown in Fig. 3a. A photolithography and a plasma etch step defined a hole pattern in the three layers, as shown in Fig. 3b.

Boron was then diffused into the wafer through the hole pattern at 1200 °C for 4 hours, creating vertical as well as lateral diffusion of boron from each hole, as illustrated in Fig. 3c. After a protective thermal oxidation and a filter frame defining back side lithography the membranes were formed using the anisotropic silicon etch EDP [2], as shown in Fig 3d. Since the EDP silicon etch rate is reduced practically to zero for very high boron doping concentrations ("boron etch-stop") [3], the resulting single-crystal silicon membrane consists of partially overlapping boron doped half-spheres with holes between the silicon half-spheres, as shown in Fig. 2. The hole size and membrane thickness were 5-10 μm.

The procedure where a photolithographically defined polysilicon hole membrane defined the single-crystal hole membrane using boron diffusion constituted the first self-alignment. In the second self-alignment the two silicon hole membranes are used to form the silicondioxide spacers by letting a buffered HF-solution etch away the silicondioxide accessible through the two silicon hole membranes.

The silicondioxide etching was interrupted when the front side and back side under-cut regions of the polysilicon membrane holes met each other, thus creating a free flow path through the structure. The resulting spacer shape is shown in Fig. 2.

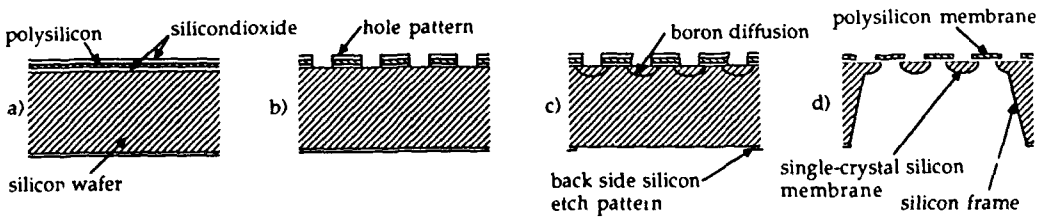


Fig. 3. A cross-sectional view of the fabrication sequence of the structure, not to scale.

a) Layers of polysilicon and silicondioxide on a (100)-silicon wafer.

b) Photolithography produces a hole pattern in the polysilicon and silicondioxide layers.

c) Boron is diffused into the wafer through the hole pattern and a back side lithography defines the silicon etch area.

d) The single-crystal boron doped silicon membrane is formed in the silicon etch, EDP, using the boron etch-stop technique. Finally the silicondioxide between the two membranes is undercut-etched until free flow paths through the hole membranes are obtained.

RESULTS

Filter structures with silicondioxide distances of 50 and 200 nm have successfully been made. The membrane area size is 6.0 x 6.0 mm and the total filter structure size including the supporting frame is 14 x 14 mm.

Preliminary gas flow-pressure measurements have been made. A differential nitrogen gas pressure of 5 psi (35 kPa) across the filters gave flow rates of 48 ml/min and 10 ml/min through the structures with respective membrane separation distances of 200 nm and 50 nm. In both cases the active membrane area was 30 mm^2. A gas pressure of approximately 1 atm could be applied over the filter without membrane breakage.

The two silicon membranes are heavily doped with boron and thus electrically conducting. This means that electrically the filter structure is a plate capacitor where the two silicon membranes constitute the electrodes. They are separated and isolated by the thickness of the silicondioxide spacers. These electrical characteristics can be used in different sensor applications for fluid identification and concentration measurement.

ACKNOWLEDGMENT

This work was supported by a grant from the National Swedish Board for Technical Development.

REFERENCES

[1] G. Stemme, Swedish patent application no. 8800869-3, 1988.

[2] A. Reisman, M. Berkenblit, S. A. Chan, F. B. Kaufman and D. C. Green, J. Electrochem. Soc. 126, 1406 (1979).

[3] H. Guckel, S. Larsen, M. G. Lagally, J. B. Miller and J. D. Wiley, Appl. Phys. Lett. 31, 618 (1977).

Session 6B Room B : Hot carriers

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

GATE CURRENTS AND DEVICE DEGRADATION : CARRIER TRANSPORT IN GATE OXIDES OF MOSFET's

A. SCHWERIN and W. HÄNSCH

Siemens, Corporate Research and Technology, Microelectronics, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

ABSTRACT

A new approach to calculate gate currents in MOS-transistors is presented. We use a nonlocal expression for carrier injection and a modified drift diffusion approximation to describe the flow of charge inside the oxide. Gate currents of n- and p-channel devices are consistently calculated with one set of parameters. The build up of oxide charge is monitored and the shift in device characteristics is described by solving simultaneously trapping rate equations in the 2-D oxide region.

INTRODUCTION

Shrinking device dimensions in modern VLSI circuits leads to increasing electric fields. This causes carrier injection into the oxide of MOS-transistors with subsequent filling of oxide traps and the build up of interface states. In this context gate currents are of some relevancy. They contain information about carrier injection. However, only a fraction of the injected carriers, dependent on the transversal field in the oxide, will reach the gate electrode to contribute to the gate current. Previous analysis of gate currents was unable to investigate the flow of charge in the oxide. For the proper description of the degradation process, however, the knowledge of carrier currents in the oxide is needed. The location of the highest injection of carriers does not necessarily coincide with the location of the damage they cause. A prominent example, for instance, is the LDD-MOSFET with a low n^--dose. There carriers, which are injected by a strong field near the gate edge, move into the spacer oxide where they may get trapped more easily than in the gate oxide.

MODEL FOR THE GATE CURRENT CALCULATION

The injection rate of the carriers, that originate from the Si is calculated with a nonlocal expression for the high energetic carriers, that move perpendicular toward the interface and can surmount the Si/SiO_2 barrier. The injected carriers are assumed to be in nonequilibrium with the oxide field. They relax toward a stationary state with the field over a distance of several nanometers away from the Si/SiO_2 interface [1] (cf. Fig. 1). The flow of the stationary state carriers is described by solving a modified drift diffusion approximation. A detailed description will be given elsewhere [2].

Our model was implemented in the device simulator MINIMOS 4 that allows the calculation of the carrier and field distributions in a realistic MOSFET with doping profiles fed in from 2-D process simulation.

MINIMOS 4 provides a solution scheme for the modified drift diffusion approximation that takes hot carrier transport into account [3]. To solve the continuity equation of both carrier types, electrons and holes,

we used this numeric model and applied it to the oxide configuration of the MOSFET-structure. The mobilities of electrons and holes in the oxide were taken from the literature [4].

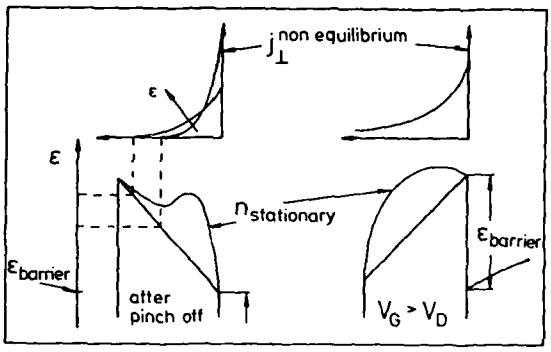


Figure 1: Stationary carrier density versus non-equilibrium current of injected carriers. Note that the stationary carrier density is, symbolically, included in the energy diagram to show its spatial distribution throughout the oxide. This distribution is obtained by solving the generalized drift-diffusion equations in the oxide region. Left side: After pinch off the oxide field opposes electron transport. High energetic electrons penetrate the oxide and relax toward the stationary distribution. Their maximal penetration depth depends on their energy. Right side: The electric field favors electron injection into the oxide. Injected carriers do not have a maximal penetration depth but they will relax toward the stationary density distribution.

RESULTS

(1) Figure 2 shows concentrations of holes and electrons inside the oxide of an n-MOSFET for two different voltage conditions (Fig. 2a: $V_G = V_D/3$, 2b: $V_G = V_D$).

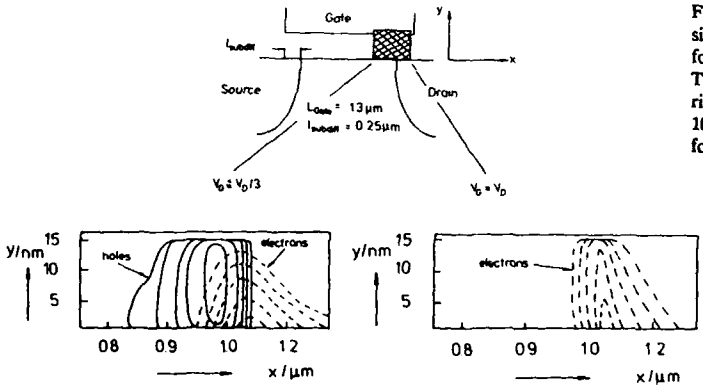


Figure 2: Carrier concentrations inside the oxide: electrons and holes for two different voltage conditions. The levels are equidistant on a logarithmic scale. The maximum level is $10^6 cm^{-3}$ for the holes and $10^7 cm^{-3}$ for the electrons.

(2) The nonlocal character of our injection expression takes heed of the following fact: To surmount the barrier carriers take up energy from the electric field with a high momentum component transversal to the interface. Minority carriers are accelerated by the lateral field, while majority carriers generated by impact ionization are accelerated by the transversal field in the case of low gate voltages. This results in a more effective injection of majority carriers for this voltage condition. Thus, for n-MOS devices the number of holes, that can surmount the barrier of 4.7eV to the oxide is sufficient to give a substantial positive gate current for low gate voltages. For p-channel transistors there is no essential gate current of holes. Figure 3 shows calculated gate currents for an n-MOSFET (3a) and for a p-MOS device (3b). The same injection parameters were used for both channel typs. The well known experimental fact, that hole gate currents are

found in n-MOS but not in p-MOS devices is found as a result of our calculations without additional assumptions. This is an excellent consistency check for the validity of our model.

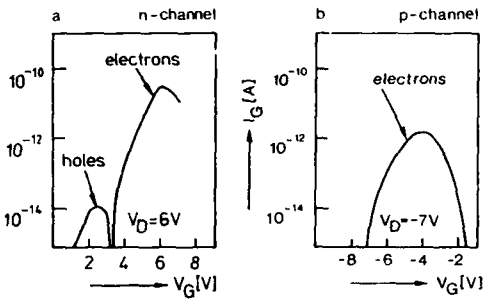


Figure 3: Simulated gate currents:
a) n-MOSFET with t_{ox} = 10nm L_G = 0.9um W = 6um, for V_D = 6V;
b) p-MOSFET with t_{ox} = 17nm L_G = 0.93um W = 10um, for V_D = -7V

(3) We used our modified MINIMOS 4 version to monitor the time evolution of the static stress experiment. To this end we solve trapping rate equations simultaneously with Poisson's equation and the continuity equation in the oxide. The trapped charge enters the right hand side of Poisson's equation. Thus we take into account the feedback of the charge, trapped in the oxide, onto the electric field in the semiconductor, which in turn will modify the number of injected carriers.
We show the results of a model calculation for a p-channel transistor. A p-MOSFET was chosen because it is generally accepted, that the main degradation mechanism there is the filling of oxide traps by injected electrons [5, 6], while the n-channel mechanism is a point of discussion up to now. For each point in the oxide the following trapping rate equation is solved:

$$dN^-/dt = j_{ox}(N^-)\cdot\sigma\cdot(N_0-N^-) \qquad (1)$$

Where N^- is the concentration of filled traps, N_0 the total trap density, σ the trapping cross section and j_{ox} is the carrier current in the oxide, which is calculated as described above. It is very important to take into consideration the effect of the growing negative oxide charge on the elctric field conditions during stress. In a p-channel transistor the build up of negative charge in oxide traps near the drain reduces the electric field peak. Thus the injected current and with it j_{ox} is a function of N^- itself.

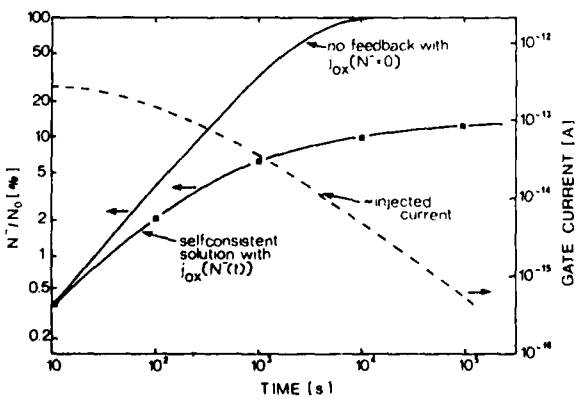


Figure 4: Simulation results of p-channel static stress at V_D = -7V, V_G = -3V. Device of Fig. 3b.
Dotted line - Calculated gate current versus time.
Full squares - Density of filled traps N^- versus time. Data are given for the point of maximum trapping (cp. Fig. 5).
Solid line - Non self-consistent solution of the trapping rate equation (1) (cp. Equ.(2)).

In Figure 4 the decrease of our simulated gate current with time is shown and we compare the time dependence of the maximum charge density in oxide traps, which results from our self-consistent calculation with the solution of equation (1) for constant $j_{ox}(N^- = 0)$:

$$N^- = N_0 \left(1 - \exp(-j_{ox}\sigma t) \right) \qquad (2)$$

The saturation of the charge encountered in the case of this analytic expression is due to the limited number of traps, whereas in our solution a saturation at a much lower level is found, which is due to the decrease in j_{ox}.

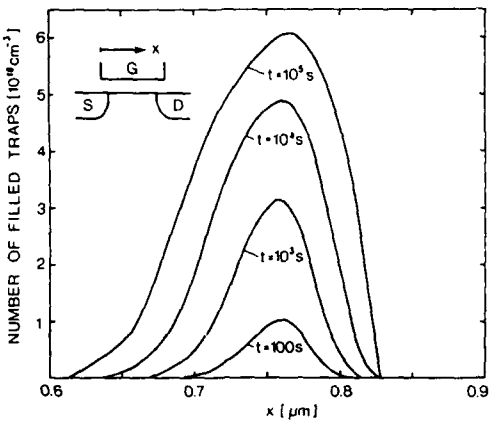


Figure 5: Density distribution of trapped negative charge along the channel calculated for different stress times (same stress as in Fig. 4).

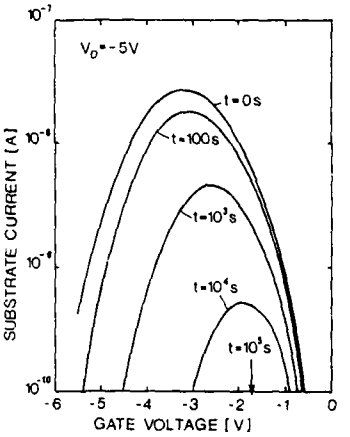


Figure 6: Substrate current characteristics for different stress times at $V_D = -5V$, calculated for the p-channel device of Fig. 4.

Figure 5 shows the density of charged traps along the channel for different stress times. In Figure 6, finally, the substrate current characteristics after different stress times is represented. The experimental known fact of a drastic reduction of substrate current with stress time is reproduced by our calculation. The negative charges trapped near drain decrease the electric field peak, thus diminishing impact ionization. Static stress experiments with p-MOSFETs show the same decrease in substrate current [7].

The example given above was calculated on a VP200 vector processor and took 30 minutes of CPU-time. To simulate the degradation behavior in this way will be of special interest in estimating long time degradation at normal operation voltages starting from short time high voltage stress experiments.

REFERENCES

[1] S.D.Brorson, D.J.DiMaria, M.V.Fischetti, F.L.Pesavento, P.M.Solomon, and D.W.Dong, J. Appl. Phys. Vol. 58, pp. 1302-1313, 1985.
[2] W. Hänsch and A. Schwerin, to be published.
[3] W.Hänsch and S.Selberherr, IEEE Trans. Electron Devices, Vol. ED-34, pp. 1074-1078, 1987.
[4] R.C.Hughes, Solid-State Electronics, Vol. 21, pp. 251-258, 1978.
[5] M.Koyanagi, A.G.Lewis, R.A.Martin, T.-Y.Huang and J.Y.Chen, Ext. Abstracts 18th Conf. on Sol. State Dev. and Mat., Tokyo, pp. 475-478, 1986.
[6] W.Weber and F.Lau, IEEE Electron Device Letters, Vol. EDL-8, pp. 208-210, 1987.
[7] A.Schwerin, W.Hänsch, and W.Weber, IEEE Trans. Electron Devices, Vol. ED-34, pp. 2493-2500, 1987.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988 C4-651

ANALYSIS OF HOT CARRIER DEGRADATION IN AC STRESSED N-CHANNEL MOS TRANSISTORS USING THE CHARGE PUMPING TECHNIQUE

R. BELLENS, P. HEREMANS, G. GROESENEKEN and H.E. MAES

IMEC vzw, Kapeldreef 75, B-3030 Leuven, Belgium

Résumé

La dégradation par porteurs chauds (hot carriers) de transistors nMOS, induite lors de l'application de tensions alternatives (AC-stress) a été évaluée en utilisant la technique Charge Pumping et les résultats ont été comparés à ceux d'une contrainte sous tension continue. En outre d'une composante de dégradation uniquement dépendante du temps de vieillissement, une composante additionelle a été observée qui est proportionelle au nombre d'impulsions appliquées (fréquence * temps). La dégradation s'avère fortement dépendante de la forme de l'impulsion à la grille. En effet, nos expériences prouvent d'une part que le temps de descente de l'impulsion de grille est beaucoup plus important que le temps de montée, et d'autre part que la largeur de l'impulsion de grille détermine le degré de compensation de la charge positive piégée durant la période de tension basse à la grille, par des électrons injectés durant la période de tension haute.

Abstract

Hot carrier degradation induced during AC-stressing of NMOS transistors is evaluated using the charge pumping technique and the results are compared with those from DC-stress. Besides a degradation component that is only dependent on stress time, an additional component is observed that is proportional to the number of applied pulses (frequency * time). A strong dependency of the degradation on the shape of the gate pulse is demonstrated. The falling edge of the gate pulse is shown to be much more important than the rising edge and the width of the gate pulse determines the amount of compensation of the trapped positive charge by injected electrons.

1. Introduction

In the last decade much effort has been spent in the characterization and understanding of hot carrier degradation in DC-stressed n-channel and p-channel transistors /1/. Although there is still no unanimous agreement on the mechanisms causing this DC-degradation, the question already arises whether the results obtained under DC-stress conditions can be simply extrapolated towards AC-stress conditions /2/ or if some additional effects have to be accounted for /3,4/. Therefore, in this study, AC-stressed n-channel transistors are evaluated with the charge pumping measurement technique /5/, which is shown to be a powerful and indispensable tool for evaluation of this kind of degradation /6,7/. The AC-degradation results are compared with those obtained during DC-stress conditions. The influence of the used frequency and of the shape of the gate pulse is studied and a possible explanation of the enhanced degradation is given.

2. Devices and experimental conditions

The devices used in this study had an effective channel length of 2.4 µm and a channel width of 100 µm. The oxide thickness was 27 nm. The devices came from different suppliers with different technologies. In this way, technology-dependent effects could be ruled out.
As illustrated on Fig. 1, during the AC-stressing, a constant voltage V_d was applied at the drain while the gate was pulsed between V_{glow} and $V_{ghigh} \approx V_d$. The other features of the applied pulse are illustrated on Fig. 1. Source and substrate are grounded.

For the charge pumping measurements rectangular pulses were applied at the gate with an amplitude of 5 V and a frequency of 100 kHz. Drain and source were kept at a small reverse bias of 0.1 V while the substrate current was monitored.

3. Expected behaviour based on extrapolation of DC-degradation

By using the charge pumping technique in combination with conventional I-V measurements, the DC-degradation mechanisms have been revealed recently /8,9/. It was found that at medium gate voltages ($V_g \approx V_d/2$) maximum degradation occurs, due to interface trap generation. At low gate voltages ($V_g \approx V_t$), the interface trap generation is masked by positive charge trapping, due to hole injection /9,10/. The latter effect is dominant on the V_t-shift for these conditions. At high gate voltages ($V_g \approx V_d$), interface trap generation is minimal, as is the

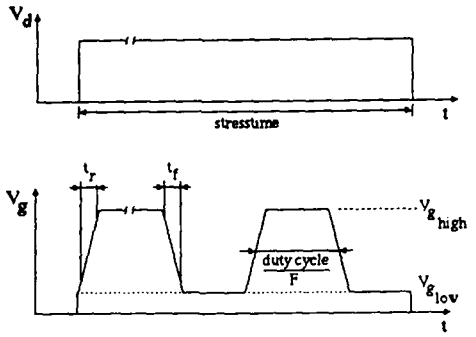


Fig. 1 : Drain and gate voltages applied during AC-stress conditions

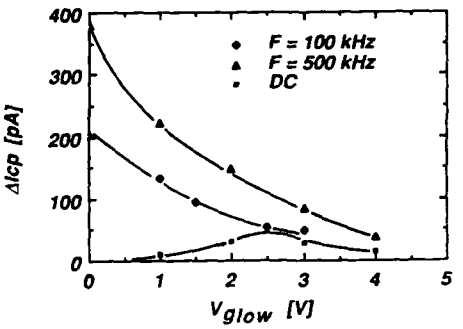


Fig. 2 : ΔI_{cp} as a function of V_{glow} for V_d=7.5V, t_r=t_f=10ns and duty cycle=50%. Stress time was 1000s for AC-stress, 500s for DC-stress. L=2µm, W=100µm

degradation of the I-V characteristic. However, when this condition is applied after a stress at $V_g = V_t$, compensation of the trapped holes occurs, and the interface traps generated at $V_g = V_t$ suddenly influence the I-V characteristics /8,9/. By simply extrapolating the DC-results to pulse shapes as on Fig. 1, one could therefore be inclined to neglect the minor degradation during the part of the gate pulse when $V_g = V_d$ and the degradation would be expected to occur mainly during the $V_g = V_{g_{low}}$ part of the gate pulse. However, during the rise and fall of the gate pulse, a contribution is expected that increases with increasing t_f and t_r. The maximum AC-degradation would however never surmount the worst case of the DC-degradation at the same V_d. Finally, the compensation of the trapped holes mentioned before is expected to be complete during the $V_g = V_d$ part of the pulse, if the trapping time of the electrons is smaller than the pulsewidth.

4. Measurement results

In the same way as for DC-stress conditions, the AC-degradation was studied as a function of stress time for different combinations of the stress parameters. While for DC-stress conditions V_d, V_g and the stress time are the most important parameters, for AC-degradations with a constant V_d, one has to evaluate the possible influence of the frequency F, the rise time t_r and the fall time t_f, the duty cycle, the low value of the gate voltage $V_{g_{low}}$ and the width of the gate pulse. All these parameters will now be studied one by one.

Comparison of AC-degradation with DC-degradation

Fig. 2 shows the increase in charge pumping current ΔI_{cp} as a function of the low value of the gate pulse $V_{g_{low}}$ for a total stress time of 1000 s. In all these experiments $V_{g_{high}}$ was equal to V_d = 7.5 V, $t_f = t_r$ = 10 ns, duty cycle = 50 %. For comparison ΔI_{cp} for DC-stresses during 500 s at V_d = 7.5 V and various V_g was plotted on the same figure. For high $V_{g_{low}}$, i.e. a small difference between $V_{g_{high}}$ and $V_{g_{low}}$, the curve for the AC-degradation approaches that of the DC-stress. For small $V_{g_{low}}$ however, the two curves differ completely and the highest degradation occurs for $V_{g_{low}}$ = 0 V and strongly exceeds the maximum DC-degradation. Moreover, the difference is proportional to the used frequency, which is already indicative of an enhanced degradation during the edges of the gate pulse.

Influence of the frequency of the gate pulse

The stresses at $V_{g_{low}}$ = 0 V were studied extensively in order to clarify the enhanced degradation. Fig. 3 shows the time dependency of the degradation for different frequencies. All degradations were performed with $t_r = t_f$ = 6 ns, duty cycle = 90 %, $V_{g_{high}} = V_d$ = 6.5 V. It follows from the figure that the time dependency of ΔI_{cp} has a slope factor of 0.67 for all frequencies, which is identical to the time dependency during DC-stress. As already illustrated on Fig. 2, one can also observe a dependency on the frequency at a given stress time. This dependency has the same slope factor n = 0.67. In Fig. 3, the time during which $V_g = V_{g_{high}}$ is independent of the frequency for a fixed stress time, since the duty cycle is fixed. Consequently, if the degradation at $V_g = V_{g_{high}}$ would be dominant, the curves for different frequencies would coincide, which is clearly not the case. When one plots the curves of Fig. 3 as a function of the number of applied pulses, i.e. N = F * stress time, one obtains more or less a unique dependency independent of the stress time and the applied frequency, as is shown on Fig. 4. This dependency once more has a slope of 0.67 which again points to a main degradation contribution during the edges of the gate pulse. One can conclude that for the AC-stress case $V_{g_{low}}$ = 0 V, the main degradation occurs only during the falling and/or rising transition edges of the gate pulse, and therefore is proportional to the number of pulses N with a dependency $\Delta I_{cp} = A N^n$ with $n = 0.67$.

Influence of the low part of the gate pulse

In the previous paragraph $V_{g_{low}}$ was always taken equal to 0 V, which means that during the part of the gate pulse when $V_g = V_{g_{low}}$ no degradation occurred while during the high part of the gate pulse ($V_g = V_d$) the DC-degradation is negligible as compared to the degradation during the transition edges of the gate pulses. For conditions where $V_{g_{low}}$ is different from 0 V however one can observe a contribution of

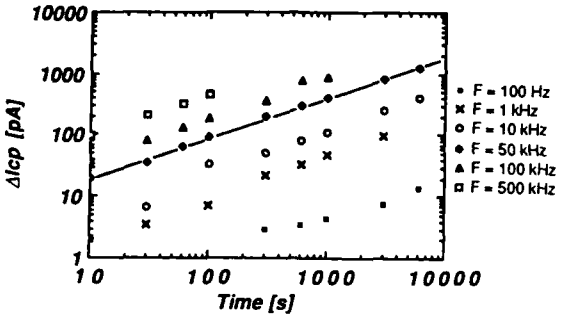


Fig. 3 : ΔI_{cp} as a function of stress time for V_d=6.5V, V_{glow}=0V, t_r=t_f=6ns and duty cycle=90%. L=2µm, W=100µm

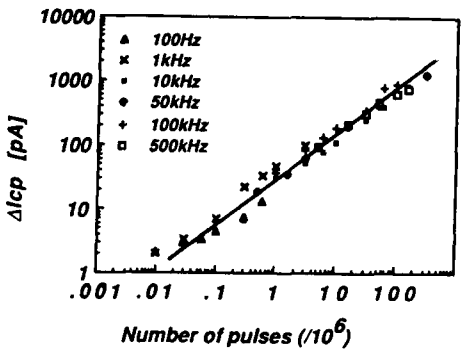


Fig. 4 : ΔI_{cp} as a function of the number of applied pulses for V_d=6.5V, V_{glow}=0V, t_r=t_f=6ns and duty cycle=90%. L=2µm, W=100µm

the DC-degradation during $V_g = V_{g_{low}}$. On Fig. 5, ΔI_{cp} is plotted as a function of the frequency for $V_{g_{low}}$ = 1 V, t_r = t_f = 6 ns, duty cycle = 50 %, $V_{g_{high}} = V_d$ = 7 V, stress time = 1000s. As can be seen ΔI_{cp} does not increase linearly with frequency in this case. At low F, the DC-degradation during the part of the gate pulse when $V_g = V_{g_{low}}$ is much larger than the extra degradation during the transition edges of the gate pulse. The latter shows up on Fig. 5 as the frequency dependent component, while the former is frequency independent, because the duty cycle is kept constant. When one subtracts the part of the degradation occurring at $V_g = V_{g_{low}}$ from the total degradation curve, only the degradation during the transition edges remains. This is shown on Fig. 5 on the curve labeled "AC component". This curve has a slope factor n = 0.7, as is the case for $V_{g_{low}}$ = 0 V.

Influence of the edges of the gate pulse

In order to reveal the cause for the enhanced degradation occurring during the transition edges of the gate pulse, the influence of the edges on ΔI_{cp} for transistors stressed at F = 100 kHz, $V_{g_{low}}$ = 0 V, $V_{g_{high}} = V_d$ = 6.5 V, duty cycle = 50 % and t_f = t_r is shown on Fig. 6. It is obvious that decreasing t_f and t_r increases the degradation. This definitely proves that this degradation is not DC-related, since in this case one would expect a decreasing degradation with decreasing t_f and t_r. In order to find an explanation for the observed behaviour, rising and falling edges were varied independently. It was found that the enhanced degradation almost completely occurs during the fall time, while changing the rise time has little effect. For very large t_r and t_f the degradation again increases because DC-degradation during the transition edges now becomes more important.

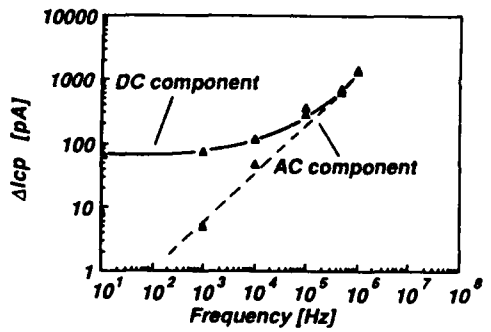


Fig. 5 : ΔI_{cp} as a function of the frequency for stress time=1000s, V_d=7V, V_{glow}= 1V, t_r=t_f=6ns and duty cycle=50%. L=2µm, W=100µm

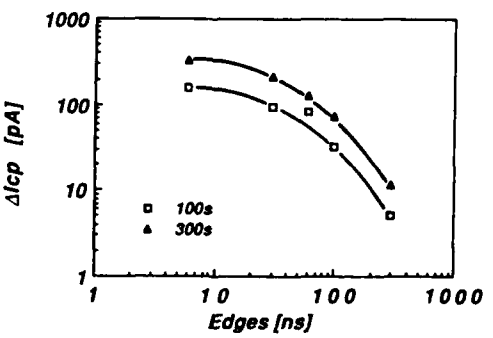


Fig. 6 : ΔI_{cp} as a function of the rise and fall times at 2 stress times for V_d=6.5V, V_{glow}=0V, F=100kHz and duty cycle= 50%. L=2µm, W=100µm

Influence of the duty cycle of the gate pulse

From DC-experiments, it is known that during a stress at $V_g = V_t$ positive charge is trapped in the oxide, that can be compensated by a subsequent stress at $V_g = V_d$ /10/. Since this sequence of stresses is repetitively applied during AC-stress, we were able to study in more detail the hole trapping and subsequent compensation using AC-stress conditions. From charge pumping measurements, it was found that positive charge trapping during $V_g = V_{g_{low}}$ still occurs up to frequencies of 10 MHz. This proves that this hole trapping process has a time constant, smaller than 50 ns. In order to study the compensation effect by electron trapping, stresses at different duty cycles were performed. Fig. 7 shows the influence of the duty cycle on ΔI_{cp} for transistors stressed at F = 100 kHz, t_r = t_f = 6 ns, $V_{g_{low}}$ = 0 V, $V_{g_{high}} = V_d$ =

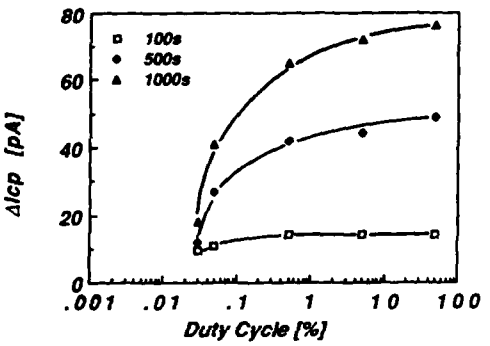


Fig. 7 : ΔI_{cp} as a function of the duty cycle at 3 stress times for V_d=6.5V, V_{glow}=0V, F=10kHz, t_r=t_f=6ns. L=2μm, W=100μm

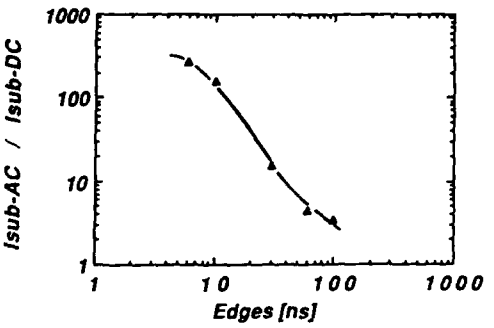


Fig. 8 : AC-component of measured I_{sub} / DC-simulated I_{sub} as a function of the edges for V_d=7.5V, V_{glow}=0V, F=100kHz and duty cycle= 50%. L=2μm, W=100μm

7.5 V and stress time = 200 s . ΔI_{cp} decreases for the lower values of the duty cycle. This behaviour will be explained in the next paragraph.

5. Discussion

From the experiments in the previous paragraph, it can be concluded that the degradation during AC-stress with constant drain voltage is composed of two components : one component is DC-related (mainly occurring at $V_g = V_{glow}$ or during the fall and rise times of the gate pulse); the other component, which is proportional to the number of applied pulses, is not observed during the DC-degradation, and occurs only during the falling edge of the gate pulse. This component has often caused the extrapolation of DC-results towards AC-conditions to fail. In these extrapolations the substrate current is always assumed to be the integral of DC substrate currents. However, for small fall times, this assumption is no longer valid. Indeed, we performed substrate current measurements under identical conditions as during the stress experiments, and compared them with measurement results under DC-conditions. The results are shown on Fig. 8 as a function of the fall time, for F = 100 kHz, V_{glow} = 0 V, V_{ghigh} = V_d = 7.5 V. The substrate current under AC-conditions, for fall times of 10 ns, is seen to be a factor of 250 larger than what is expected from DC-measurements. This enhancement decreases to a factor of 3 for larger fall times. These results are in qualitative agreement with the degradation as a function of the edges, showed on Fig. 6. Moreover, it is observed that this enhancement is not dependent on the rise time, as is also the case for the degradation behaviour. Therefore, this enhanced substrate current can be correlated with the degradation component during the falling edge of the gate pulse. At present, experiments are being performed to clarify the exact origin of this substrate current enhancement. Together with the substrate current increase and accompanying increase in interface trap density for decreasing fall time, a large build-up of positive charge is expected. For the conditions of Fig. 7 with a low duty cycle this positive charge is only partially compensated by the electrons injected during the part of the gate pulse when $V_g = V_d$ /9,10/. Because of the incomplete compensation, a net positive charge remains at the drain which decreases the lateral electric field during the falling edge resulting in a reduced generation of interface traps. For larger duty cycles, the positive charge is however completely compensated within each cycle and the electric fields will be maximal during the fall-off of the gate pulse. From Fig. 7 and from similar additional charge pumping experiments, the time constant for the compensation process (electron trapping on trapped holes) is found to be about 0.5μs.

6. Conclusions

An enhanced degradation has been observed for AC-stress conditions using gate voltage pulses and a constant high drain voltage. This enhancement is dependent on the shape of the gate pulse. Decreasing fall times enhance the effect, while the rising edge has little influence. An enhanced substrate current has been observed for identical stress conditions. At present experiments are being performed to explain this phenomenon. The duty cycle determines the amount of compensation of the positive charge by the injected electrons.

Acknowledgement

The authors wish to thank BTMC Alcatel for providing the devices. P. Heremans is a research assistant of the Belgian National Fund for Scientific Research. Part of this work has been carried out within a IWONL-poject.

References

/1/ C. Hu, S. C. Tam, F. C. Hsu, P. K. Ko, K. W. Terrill " Hot-Electron-Induced MOSFET Degradation-Model, Monitor, and Improvement" IEEE Transactions on Electron Devices, ED-32, pp. 375-385, 1985

/2/ T. Horiuchi, H. Mikoshiba, K. Nakamura, K. Hamano "A Simple Method to Evaluate Device Lifetime Due to Hot-Carrier Effect Under Dynamic Stress" IEEE Electron Device Letters, EDL-7, pp. 337-339, 1986

/3/ M. M. Kwo, K. Seki, P. M. Lee, J. Y. Choi, P. K. Ko, C. Hu "Quasi-Static Simulation of Hot-Electron-Induced MOSFET Degradation Under AC (Pulse) Stress" IEDM Tech. Dig., pp. 47-50, 1987

/4/ J. Y. Choi, P. K. Ko, C. Hu "Hot-Carrier-Induced MOSFET Degradation Under AC Stress" IEEE Electron Device Letters, EDL-8, pp. 333-335, 1987

/5/ G. Groeseneken, H. E. Maes, N. Beltran, R. F. De Keersmaecker "A Reliable Approach to Charge-Pumping Measurements in MOS Transistors" IEEE Transactions on Electron Devices, ED-31, pp. 42-53, 1984

/6/ P. Heremans, H. E. Maes, N. Saks "Evaluation of Hot Carrier Degradation of N-Channel MOSFET's with the Charge Pumping Technique" IEEE Electron Device Letters, EDL-7, pp. 428-430, 1986

/7/ P. Heremans, R. Bellens, G. Groeseneken, H. E.Maes, sent in for publication

/8/ P. Heremans, J. Witters, G. Groeseneken, H. E.Maes, sent in for publication

/9/ R. Bellens, P. Heremans, G. Groeseneken, H. E. Maes "A New Procedure for Lifetime Prediction of n-Channel MOS-Transistors Using the Charge Pumping Technique" PROCEEDINGS IRPS, Monterey, pp. 8-14, 1988

/10/ R. Bellens, P. Heremans, G. Groeseneken, H. E. Maes "Hot-Carrier Effects in n-Channel MOS Transistors Under Alternating Stress Conditions" IEEE Electron Device Letters, EDL-9, pp. 232-234, 1988

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

DYNAMIC HOT-CARRIER DEGRADATION OF FAST-SWITCHING CMOS INVERTERS WITH DIFFERENT DUTY CYCLES

L. RISCH and W. WEBER

Siemens AG, Corporate Research and Development, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

ABSTRACT - The frequency shift of ring oscillators operated at high power supply voltages exhibits hot-carrier degradation similar to the well-known stress effects measured at single transistors. The predicted duty cycles based on substrate currents generated during the fast switching periods yield results which are in good agreement with the degradation data from ring oscillators and externally switched inverters only for short stress times. For long stress times, however,deviations are reported.

1 - INTRODUCTION

Hot-carrier effects in submicron CMOS circuits are considered a major problem when using a power supply voltage of 5V. Recently, dynamic stress experiments performed on transistors with externally switched gate and drain voltages confirmed the assumption of duty cycles based on substrate currents alone /1/ or suggested a modification additionally taking into account the switching current of the transistor /2/.
With externally applied voltages it is difficult to obtain rise and fall times of gate and drain voltages shorter than 3ns. In CMOS circuits such as logic cells or driver stages the rise/fall times of the voltages are much shorter. Consequently degradation experiments using ring oscillators with different load capacitances and gate delay times of 100ps to 3ns are used to investigate dynamic degradation in this high frequency range.

2 - EXPERIMENTAL DETAILS

For an inverter with externally applied gate voltages a test structure was designed which consists of a chain of four CMOS inverters. Each of the inverter outputs can be connected to a pad by a transfer gate for the characterization. The pad capacitance is switched off during the dynamic stress. With this arrangement either each transistor of the inverter chain can be characterized statically or can be stressed dynamically like a ring oscillator with a small load capacitance by an external pulse fed into the first inverter. The n-MOSFET results of the first inverter are shown here. The current degradation was measured in the linear region.
Whereas in static degradation experiments V_t, g_m, or I_D are measured to characterize the degradation of the transistor, here the frequency shift of ring oscillators is used. This quantity can be measured very accurately down to the 0.1% range. Typically the frequency decreases during stress due to the decreasing current of the n-channel transistor caused by hot-carrier injection. Three single-gate and three cascoded (with two n-channel transistors in series) ring oscillators, each with a different number of stages and load

capacitances, were used to investigate the influence of the duty cycle on dynamic hot carrier degradation as listed in Table 1.

The LDD transistors under investigation were fabricated in 4Mbit DRAM technology with a gate length of 1.0um, Wn = 10um, Wp = 20um and tox = 20nm.

fan in/out	cycle time	gate delay
F = 1	10ns	0.18ns
F = 1; 1pf	40ns	1.7 ns
F = 3	60ns	0.35ns

Table 1 Circuit parameters of the ring oscillators

3 - INVERTER CHAIN RESULTS (rise/fall time $\geq$ 3ns)

Fig.1 compares results for different pulse configurations. For a constant rise/fall time of 3ns the pulse period was varied from 40 to 400 and 4000ns. This should reduce the duty cycle by one order of magnitude in each case. The experimental results agree with this as can be observed in Fig. 1.

Then the duty cycle was kept constant and the pulse period was changed. Fig. 1 shows that all three curves agree essentially within the limits of experimental error again demonstrating the validity of the duty cycle assumption.

Comparing these results with static stress experiments by calculating the substrate currents during the switching period, agreement with a calculated factor of 130 is found only at short stress times. At longer stress times the dynamic curves approach the static one and may even cross for long stress times. This finding contradicts the duty cycle assumption and was therefore confirmed for several processes. Furthermore, dynamic stress experiments were performed on single transistors by applying separate pulses at the gate and drain to perform a check on a different experimental setup. The result was again the same.

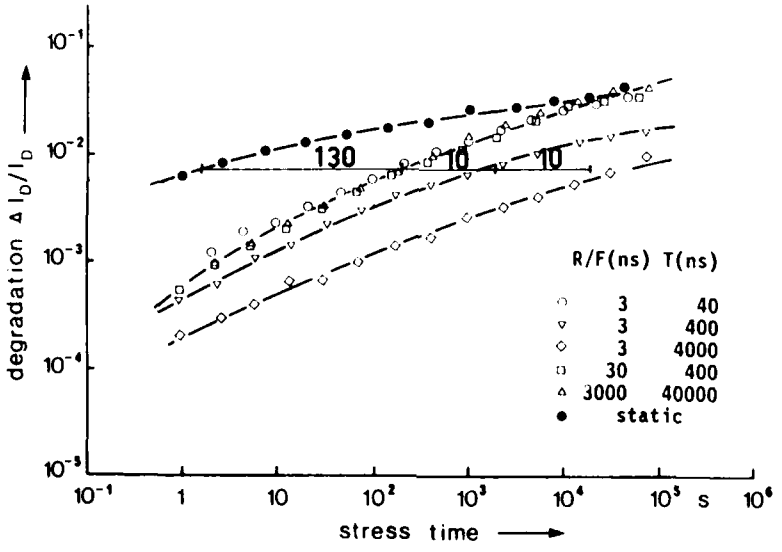


Fig.1 Drain current degradation of a single n-MOSFET in an inverter chain as a function of the dynamic stress time for different input pulse configurations. A static stress result is included included for comparison.

4 - RING OSCILLATOR DEGRADATION (rise/fall time < 3ns)

4.1 Voltage dependence

The frequency shift of ring oscillators due to hot-carrier degradation behaves very similarly to the degradation of the current. Fig.1 shows the frequency shift for a ring oscillator with F = 3 at supply voltages of 5 to 8V as a function of the stress time. After an initial strong shift of the frequency, the degradation of the ring oscillator shows a saturation effect. The frequency shift also increases with power supply voltage and is therefore correlated with the degradation of the transistor parameters.

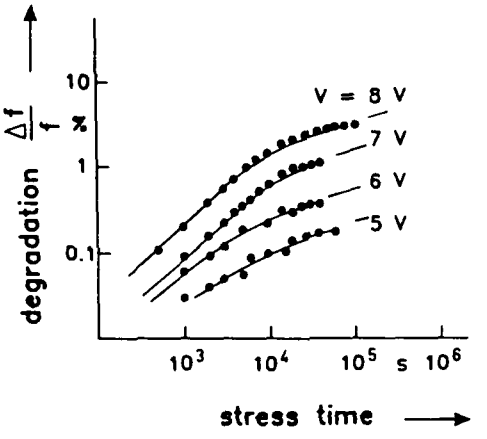


Fig.2 Frequency shift of ring oscillator F = 3 at different voltages

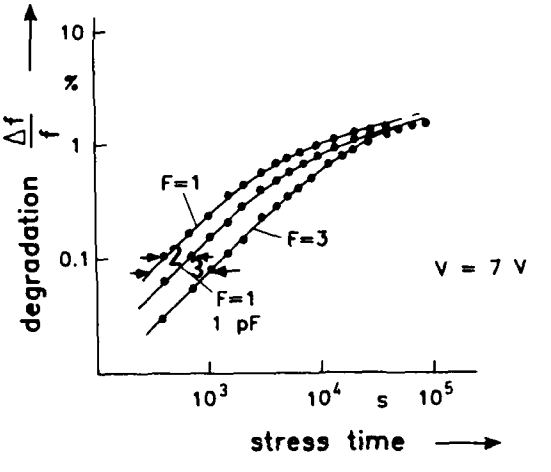


Fig.3 Frequency shift of ring oscillators with different gate delay and cycle times

4.2 Duty cycle dependence

The basis for predicting the dynamic degradation is the ratio of substrate current generated during switching periods to the maximum substrate current of the transistor under static stress conditions. Introducing the substrate current into a circuit simulator such as SPICE with subsequent integration of this current during the cycle time should therefore yield the degradation of the circuit. The drain current dependence (compare /2/) is neglected because all ring oscillators have approximately the same switching current.

The results for three different single-gate inverters are shown in Fig.3. For short stress times, differences can be explained on the basis of the substrate currents (see Fig.4). However, for longer stress times of 10^4-10^5s the three curves level off and lead approximately to the same degradation with the same time dependence as the dynamic results shown in Fig. 1.

Analyzing the duty cycles with SPICE simulations including substrate currents (see Fig.3), the duty cycle theory predicts lifetimes which are 2 and 3 times longer for F = 1, 1pF and F = 3 than for F = 1 respectively. Experimentally this was found only for short stress times while for longer ones the three ring oscillators show approximately the same degradation.

4.3 Substrate current dependence

Cascoded inverter ring oscillators with the same number of stages exhibit greatly reduced substrate currents during switching periods. The peak values are reduced to about 1/7 without changing the duty

cycle. From this an increase in lifetime by a factor of about 120 is predicted. Experimentally, a factor of 50 is found (Fig.5), which is in rough agreement with the calculation. At short stress times the ring oscillator with the fastest gate delay time again exhibits greater degradation. For longer stress times, the same degradation is attained despite different duty cycles. In comparison with Fig.2 the cascoded inverter degrades at 7V like the normal inverter at 5V due to the internal voltage reduction.

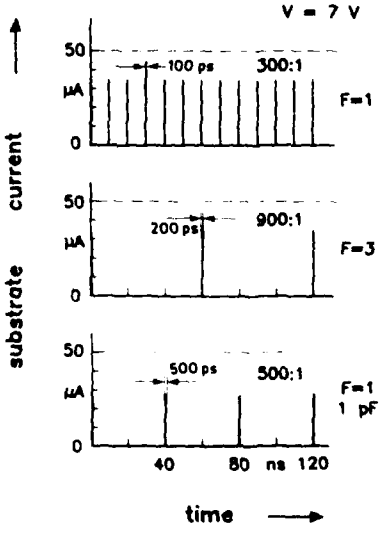


Fig.4 Simulated substrate current peaks for different inverters (shown schematically)

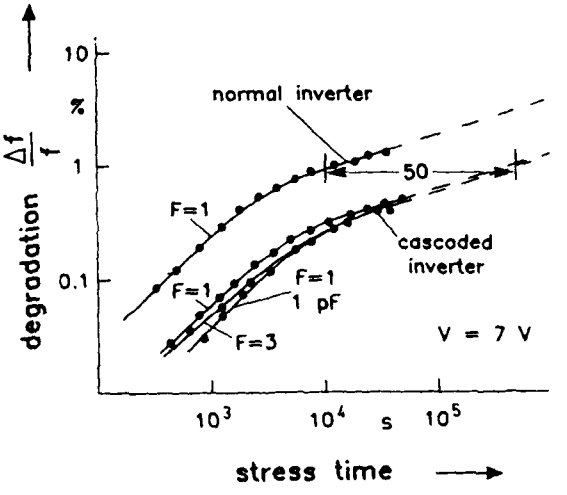


Fig.5 Frequency shift of normal and cascoded ring oscillators with different delay and cycle times

5 - CONCLUSIONS

Comparing the time dependent changes of drain current in inverters with the frequency shift of ring oscillators, a similar behavior can be observed, which shows clearly that both parameters originate from hot-carrier stress and are proportional to one another. Furthermore it is shown that the saturation behavior is less pronounced for long stress times after dynamic stress (Fig. 1). The duty cycle calculation on the basis of substrate currents was verified at short stress times but yields increasing errors for long times (Fig. 1, 3 and 5). Despite different duty cycles, approximately the same degradation is attained by different ring oscillators. The reduced substrate current of a cascoded inverter is not fully correlated with the lifetimes. To extrapolate dynamic lifetimes up to 10 years more accurately further measurements at longer stress times as well as a quantitative refinement of the model will be necessary/3/.

ACKNOWLEDGMENTS

This report is based on a project which was supported by the Minister of the FRG under support-No NT 2696. The authors alone are responsible for the contents.

REFERENCES

/1/ T.Horiuchi, H.Mikoshiba, K.Nakamura, K.Hamano, IEEE Electron Device Lett., vol. EDL-7, pp.337-339, 1986
/2/ W.Weber, C.Werner, A.v.Schwerin, IEDM 86, pp.390-393
/3/ W.Weber, IEEE Trans. Electron Devices, September issue, 1988.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

DEGRADATION OF SHORT-CHANNEL MOS TRANSISTORS STRESSED AT LOW TEMPERATURE

C. NGUYEN-DUC, S. CRISTOLOVEANU, G. REIMBOLD* and J. GAUTIER*

Laboratoire de Physique des Composants à Semiconducteurs, ENSER-INPG, 23, rue des Martyrs, F-38031 Grenoble Cedex, France
D. LETI, Centre d'Etudes Nucléaires de Grenoble, 85X, F-38041 Grenoble Cedex, France

Résumé - Nous étudions le vieillissement à 77 K de transistors MOS de type N soumis à de fortes contraintes électriques: tension de drain V_d = 5,5 V et tension de grille V_g variant de 1,5 à 6,5 V. Nous montrons que les maxima de la dégradation de transconductance et du décalage de la tension de seuil ne se manifestent pas pour les mêmes conditions de contrainte. Les résultats sont expliqués par la nature plus ou moins localisée des défauts créés qui est aussi responsable de la distorsion des courbes de transconductance et de son augmentation temporaire lors du vieillissement. Une augmentation inhabituelle de la transconductance en régime de saturation est également mise en évidence.

Abstract - Hot-carrier stressing was carried out on 1 µm n-type MOSFETs at 77 K with fixed drain voltage V_d = 5.5 V and gate voltage V_g varying from 1.5 to 6.5 V. It was found that the maximum transconductance degradation and threshold voltage shift do not occur at the same V_g. This behavior is explained by the localized nature of induced defects which is also responsible for a distortion of the transconductance curves and even a slight temporary increase in the transconductance during stress. An anomalous increase in the saturation transconductance is also reported.

1 - INTRODUCTION

The hot carrier induced degradation is known to be more accentuated for MOSFET operation at low temperature but most of the experimental work was carried out at 300 K. Recent papers show a lack of concordance concerning the physical mechanisms involved in the device degradation at 300 K and 77 K (generation of interface states /1/ or charge trapping /2/). In relatively long MOSFETs, maximum degradation occurs for a given drain voltage V_d when the gate voltage V_g is approximately $V_d/2$ (maximum substrate current) /3/.

In this work, we study the defect nature and localization in short channel n-type MOSFETs subjected to high-field stressing at 77 K by regarding threshold voltage V_t shift and changes in the transconductance G_m value and shape. Both the linear region and saturation region are reported. Some particularities in the transconductance behaviour of stressed devices are discussed for the first time.

2 - EXPERIMENT AND RESULTS

The devices used were conventional n-channel MOSFETs with effective channel length $L_{eff} \approx 1$ µm, channel width W = 20 µm and oxide thickness t_{ox} = 25 nm fabricated at LETI Laboratories in Grenoble (France). The devices were directly immersed in liquid nitrogen and stressed at a fixed drain voltage V_d = 5.5 V and various gate voltages V_g. The device characterization was performed at 77 K and 300 K, before and after stress, by using HP 4145B Analyzer. We measured the drain current $I_d(V_g)$ and transconductance $G_m(V_g)$ characteristics in the linear (V_d < 100 mV) and saturation region with both forward and reverse modes of operation, the substrate current and the subthreshold slope. V_t was determined at I_{ds} = 0.2 µA/µm of channel width or by extrapolating the linear $I_d(V_g)$ curve.

The transconductance reduction rate $\Delta G_m = (G_{mo}-G_m)/G_{mo}$ and threshold voltage shift $\Delta V_t = V_t - V_{to}$ occuring after one hour stress is presented in Fig. 1. Here G_{mo} and V_{to} denote pre-stress values. We notice that maximum ΔV_t and ΔG_m do not occur at the same V_g : ΔV_t is very small for the $V_g < V_d$ stressing, becomes significant at $V_g \approx V_d$ and then rapidly increases with increasing V_g, while ΔG_m has its maximum in the region of maximum substrate current.

Although for $V_g < V_d$, the degradation rate follows roughly a law of $\Delta G_m \propto t^{0.25}$, the behaviour for $V_g > V_d$ stressing is more surprising ; a temporary increase in G_m is found after 10^4 sec of stress (Fig. 2a). This transconductance "overshoot" is also clearly seen in $G_m(V_g)$ curves (Fig. 2b). Very similar curves have been obtained by 2-D simulation /4/ for the case where generated defects are localized in a region of ΔL = 0.15 µm close to the drain and explained by a two-piece model : the total transconductance of the stressed MOSFET may be dominated by that of the defective region ΔL which has a higher threshold voltage but a much shorter length compared to the defect-free region. Our results offer the first experimental support for the predictions of a transconductance overshoot during the transistor aging.

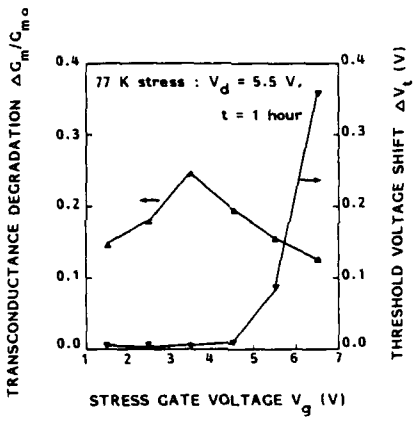


Fig. 1 - Maximum transconductance degradation and threshold voltage shift versus stress gate voltage after 1 hour's stress at 77 K with V_d = 5.5 V.

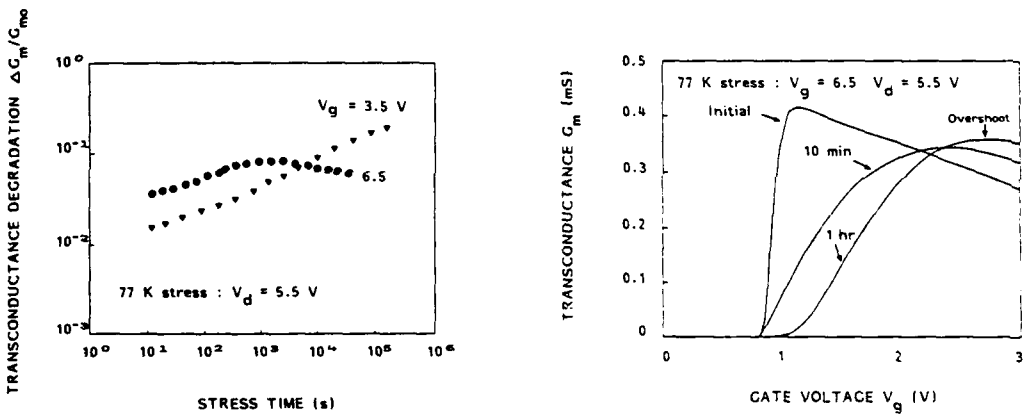


Fig. 2a - Maximum transconductance degradation versus stress time for two typical stress conditions.

Fig. 2b - Transconductance versus gate voltage characteristics of a device stressed at V_d = 5.5 V and V_g = 6.5 V; measurement at 77 K with V_d = 50 mV and stress time as a parameter.

For $V_g < V_d$, we observe first a slight negative shift in the extrapolated V_t followed by a discontinuous transition to positive shifts (Fig. 3a). The negative ΔV_t may be misinterpreted as being due to the generation of positive charge. We propose another explanation based on the fact that V_t depends also on the slope of $I_d(V_g)$ curve /5/ :

$$V_t = V_{gmax} - \frac{G_{dmax}}{G_{mmax}}$$

where G_{dmax} et G_{mmax} are maximum conductance and transconductance, respectively, and V_{gmax} is the gate voltage corresponding to G_{mmax}. It is shown in Fig. 3b that for short stressing periods, the maximum value G_{mmax} decreases while its position V_{gmax} does not change, so that V_t decreases. For longer periods of stress (t ≥ 1 hr) a second G_m maximum emerges which can exceed the first one and leads to an *apparent* increase in the extrapolated V_t. It is, therefore, more reasonable to consider that the important role is played by the generation of defects in a very narrow region ΔL situated above the channel and beside the drain. Indeed, the curves of Fig. 3b are very similar to those obtained by 2-D simulation, for the case where acceptor states or fixed negative charges of density of 5 to 10×10^{11} cm^{-2} are localized in ΔL = 0.05 μm /4/.

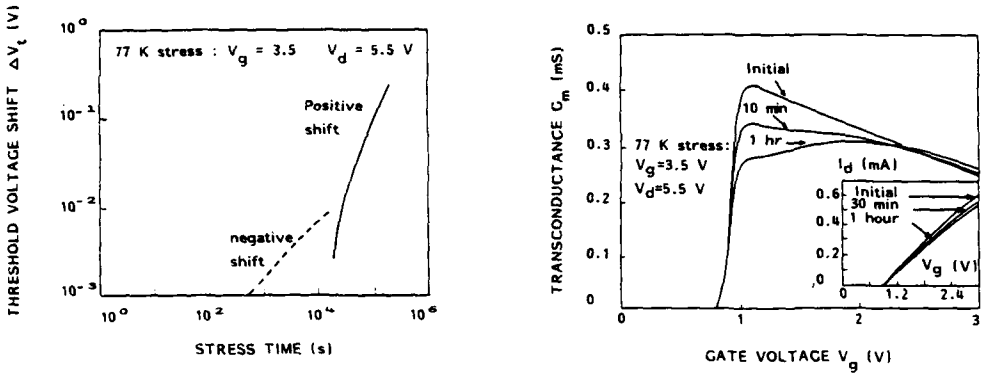


Fig. 3a - Extrapolated threshold voltage shift versus stress time for a device stressed at V_g = 3.5 V and V_d = 5.5 V.

Fig. 3b - Transconductance versus gate voltage characteristics of a device stressed at V_g = 3.5 V and V_d = 5.5 V; measurement at 77 K with V_d = 50 mV and stress time as a parameter.

The relaxation effect at 77 K was found to shift back the curves compared to those taken immediately after stress (curve 3 of Fig. 4); approximately 30 % of the initial degradation is recovered after 60 hours relaxation. This requires the stress-induced degradation to be caused partly by the electron capture in shallow-level traps in the gate oxide as suggested by Ning /6/. Warming up the stressed samples to 300 K shows yet a more beneficial influence; subsequent measurement at 77 K shows a larger decrease (≈ 50 %) in ΔV_t (curve 4). As the effect of damage created by the aging cannot be completely removed we conclude that the stress-induced defects in the MOSFETs studied here consist of (i) reversible electron capture in existing shallow-level traps and (ii) irreversible creation of interface states and/or electron traps in the oxide.

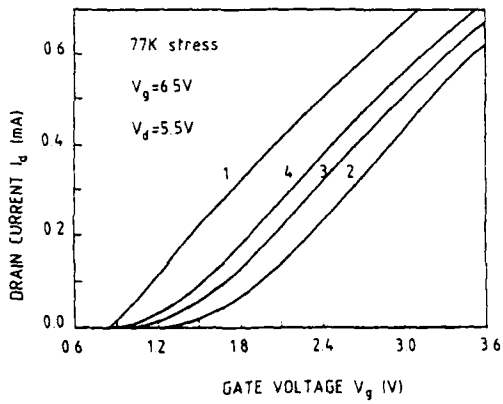


Fig. 4 - $I_d(V_g)$ characteristics showing the relaxation effect; (1) initial, (2) 3 hours stress at V_g = 6.5 V and V_d = 5.5 V, (3) 60 hours pause at 77 K, (4) + 1 hour pause at 300 K.

Another relevant result is observed in the case of $V_g > V_d$ stressing; after stress the saturation transconductance measured for V_d = 5 V in the reverse mode becomes larger than in the forward mode (which remains almost the same as the pre-stress value) (Fig. 5). This improvement in G_{msat} for degraded MOSFETs may be explained by the interaction between the defective and defect-free regions of the channel (as proposed above for the ohmic operation). In the forward-mode operation, G_{msat} of the degraded MOSFET should be the same as the pre-stress value because the defective region is completely screened by the pinch-off region. In contrast, when the source and drain are interchanged, the stressed device can be cosidered as a series combination of two transistors. The total transconductance depends on the interaction of these two regions and can exceed the transconductance of a homogenous MOSFET.

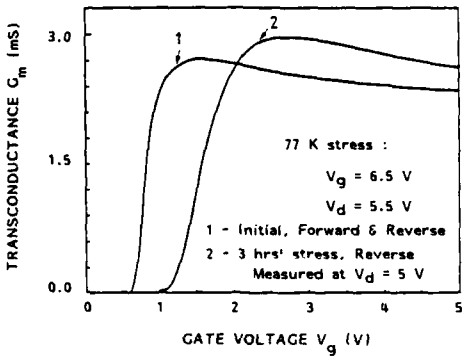


Fig. 5 - Saturation transconductance versus gate voltage of a MOSFET stressed at V_g = 6.5 V and V_d = 5.5 V.

3 - CONCLUSION

Hot carrier degradation has been carried out on 1 μm n-MOSFETs at 77 K and constant drain voltage, with the gate voltage as a variable. It has been found that the extension of the defective region into the channel depends strongly on the V_g/V_d ratio and determines the behavior of the device degradation. For the $V_g < V_d$ stress regime, the degradation of the maximum transconductance is dominant while the threshold voltage hardly changes. In contrast, at $V_g > V_d$ the stress-induced shift of V_t is important and increases with increasing V_g while the change in the transconductance maximum is found to be attenuated and a transconductance overshoot is observed for several periods of stress. The increase in the saturation transconductance of the stressed devices is explained by the two-piece model. The comparison with the results of 2-D numerical calculations allows us to conclude that for a drain voltage as high as 5.5 V, increasing V_g makes the extension of the defective region go deeper into the channel. The partial relaxation at 77 K and 300 K of the created defects suggests that they consist of the electrons captured in shallow-level neutral traps and the acceptor-like interface states and/or negative fixed charges.

REFERENCES

/1/ Von Bruns, S.L. and Anderson, R.L., IEEE Trans. Electron Devices ED-34, (1987), 75.
/2/ Toriumi, A., Iwase, M., Wada, T. and Taniguchi, K., IEDM Tech. Dig. (1986), 382.
/3/ Takeda, E. and Suzuki, N., IEEE Electron Device Lett. EDL-4, (1983), 111.
/4/ Haddara, H. and Cristoloveanu, S., IEEE Trans. Electron Devices, ED-34, (1987), 378.
/5/ Ghibaudo, G., Phys. Stat. Sol. (a) 95, (1986), 372.
/6/ Ning, T., J. Appl. Phys., 49, (1978), 5997.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

HOT ELECTRON RELIABILITY OF DEEP SUBMICRON MOS TRANSISTORS

G. REIMBOLD, F. PAVIET-SALOMON*, H. HADDARA*, G. GUEGAN and S. CRISTOLOVEANU*

LETI, CENG, 85X, F-38041 Grenoble Cedex, France
**Laboratoire de Physique des Composants à Semiconducteurs, ENSERG, F-38031 Grenoble Cedex, France*

Résumé - Nous étudions la dégradation des performances des transistors MOS ultra-courts (0.3 µm - 0.6 µm) engendrée par l'injection de porteurs chauds. Ces dispositifs ont un canal N, une structure conventionnelle (non LDD) et ont été optimisés pour fonctionner à 3 V. Plusieurs types de contraintes ont été analysés. Un suivi systématique des paramètres importants a été réalisé en cours de vieillissement, la dégradation étant ensuite évaluée par des méthodes de caractérisation fine. L'influence des tensions d'alimentation sur la durée de vie des dispositifs est étudiée. Ces résultats sont interprétés en tenant compte de l'extension de la zone de défauts et du taux de génération locale d'états d'interface.

Abstract - The hot electron induced degradation of fully optimized N-channel MOSFET's, having channel lengths in the range 0.3 µm - 0.6 µm, is systematically investigated. The created defects and their influence on the device performance are evaluated with very sensitive techniques and explained using 2D modelling. The device lifetime is analysed as a function of the biasing conditions. These results are interpreted by taking into consideration the extension of the defective region as well as the local generation rate of interface states.

1 - INTRODUCTION

Hot electron reliability of MOS transistors is a key problem for the long term stability of modern VLSI circuits. During the last decade much effort has been devoted to the study of hot carrier-induced degradation in MOSFET's longer than 1 µm. Although many important aspects have been clarified, neither a complete description of the physical mechanisms has been given nor a fully satisfactory technological solution proposed to eliminate aging problems. In addition, an accentuation of the hot carrier deterioration of the Si/SiO_2 interface is expected to occur as the new generation of submicron transistors emerges. Other fundamental questions are related both to the possibility of avoiding LDD structures and the impact of a bias supply limitation at 3 V. However, a very small amount of experimental data was reported on the aging of half-micron transistors.

There are three original aspects in this work :

(1) use of fully optimized deep submicron MOSFET's with 0.3 - 0.6 µm channel lengths,
(2) systematic aging under different stressing conditions and for very long periods of time (6.10^5 seconds),
(3) monitoring of the evolution of the main device parameters during stress, followed by a detailed inspection of the interface damage at the end of the stress using very sensitive characterization techniques (dynamic transconductance and charge pumping) as well as an accurate parameter extraction method .

2 - EXPERIMENTAL CONDITIONS

The transistors used in this study have been fabricated with electron beam lithography. They have channel widths of 20 µm, 15 nm gate oxides, 0.17 µm junction depth and nominal gate lengths from 0.4 to 0.8 µm (the effective channel length being about 0.1 µm shorter). They were optimized for a 3 V bias supply in order to avoid the use of LDD structures [1]. Stress was performed at different gate and drain voltages. The aging effects of low (1.4 V) and higher (3.3 V) gate biases were compared by keeping the drain voltage in the saturation region before the avalanche onset.

Parameter	ΔV_T (V)	$\frac{\Delta Gm}{Gm}$	$\frac{\Delta Isat}{Isat}$	$\frac{\Delta Ilin}{Ilin}$	ΔS (V/dc)
n	0.64	0.42	0.42	0.46	0.71
A	10^{-1}	2.10^{-3}	$1.3\ 10^{-1}$	7.10^{-4}	3.10^{-6}

Table 1. Coefficients of the time dependent degradation rate (At^n) of various parameters in a 0.3 µm MOSFET stressed at V_G = 1.4 V and V_D = 4.4 V.

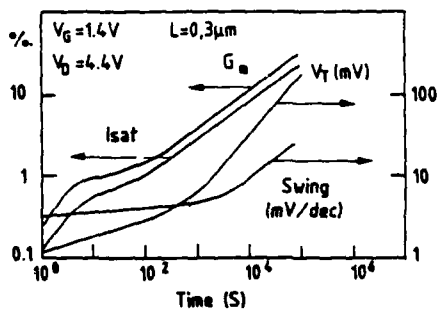
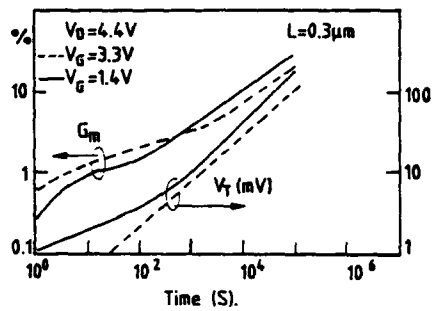


Fig. 1. Degradation of various parameters in a 0.3 µm transistor.

A computerized data acquisition system enabled us to study simultaneously several transistors and monitor the evolution of their parameters. Figure 1 shows the degradation of the main parameters (threshold voltage V_t, maximum of transconductance G_m, saturation current in reversed mode I_{sat}, subthreshold slope) in a 0.3 µm transistor. The time dependent degradation rate of all these parameters follows the classical law At^n (see Table 1). The highest value of n corresponds to the subthreshold swing but no increase in the off-state leakage current is observed. In contrast, the threshold voltage shift is higher than in the micronic transistors, probably because of the greater ratio between the degraded region length and the effective channel length. The usual equivalence between a 10 % degradation in transconductance and a 10 mV shift in V_t does no longer hold in this case (Figs.1 and 2).



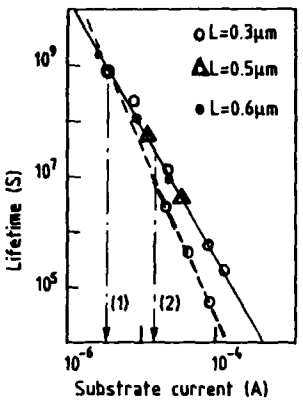


Fig. 2. Transconductance degradation and threshold voltage shift versus time for two stressing conditions.
Fig. 3. Device lifetime corresponding to a 10 % variation in the saturation current (solid line) and to a 10 mV shift in the threshold voltage (dashed line) versus substrate current. The duty ratio is 10 %. (1) and (2) correspond to the case of the shorter transistor (0.3µm) biased at V_D = 3V and 3.5V respectively.

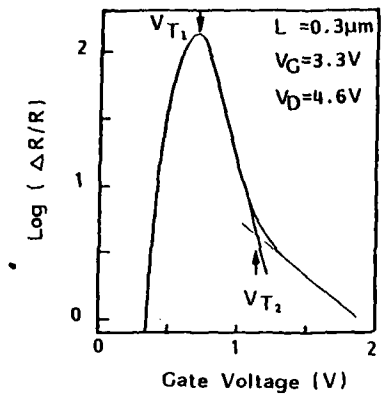
In order to extrapolate the device lifetime, we have chosen two different criteria : (1) 10 % decrease in the reverse mode saturation current (for $V_g = V_d = 3$ V) and (2) 10 mV shift in V_t. The two corresponding lifetimes are plotted in Fig.3. The classical dependence $\tau = C(I_{sub})^{-m}$ is very well fitted by m = 2.5 for the saturation current and m = 3 for the threshold voltage. Assuming a duty ratio of 1/10 for CMOS technology, a 10 years lifetime can be obtained, with both criteria, at 3 V bias supply even for the shortest transistor (0.3 µm). For higher supply voltage the degradation becomes intolerable, the threshold voltage shift being the most limiting factor. This suggests the use of LDD like structures which are believed to reduce the threshold voltage shift and keep a good noise margin for these low voltage transistors. After stress we also noticed an increase by a factor of two in the substrate current due to the local trapping of negative charges that causes the lateral field to increase.

3 - DETAILED CHARACTERIZATION OF THE DEGRADED DEVICES

In order to determine the defective channel length ΔL, the type of the created defects, their density ΔN_q and the carrier mobility, we use a parameter extraction method especially conceived for degraded submicron devices /2/. In this model, the damaged device is presented by the series combination of two transistors with different channel lengths and threshold voltages.

Figure 4 shows an original plot of the degradation rate of the channel resistance versus V_g, which illustrates the interaction between the defective and non defective regions of the channel : for $V_g < V_{t1}$ both regions are in weak inversion, while for $V_g > V_{t2}$ they are in strong inversion. For intermediate gate voltages only the defect free region reaches strong inversion while the degraded one is still in weak inversion. In other words, V_{t1} is the threshold voltage of the virgin device whereas V_{t2} is that of the degraded region. This curve is in very good agreement with previous theoretical calculations /2/. Table 2 gives the values of the threshold voltages shifts $\Delta V_t = V_{t2} - V_{t1}$, mobilities degradation rates $\mu_1 - \mu_2/\mu_1$, degraded region lengths ΔL and densities of created defects ΔN_q. It is found that the defective channel length ΔL becomes more important for the higher gate voltage whereas the density of created defects is weaker.

In order to validate these results, we performed direct measurements of interface states were performed with the dynamic transconductance /3/ and charge pumping /4/ methods. The interface states generated for various conditions of stress are plotted versus drain current (Fig. 5). Indeed for each channel length, a constant drain current, at fixed V_d, is nearly equivalent to a constant inversion charge and a constant surface potential. We clearly observe that the density of created interface states is larger after stress at lower gate voltage (V_g = 1.4 V).



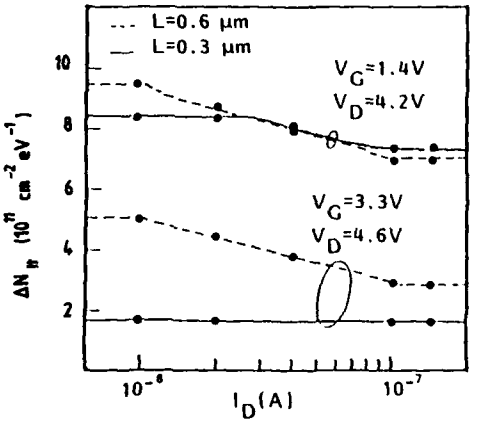
Channel length	L = 0.3 µm		L = 0.6 µm	
Bias conditions	V_G = 3.3V V_D = 4.6V	V_G = 1.4 V V_D = 4.2V	V_G = 3.3V V_D = 4.6V	V_G = 1.4V V_D = 4.2V
ΔL/L	80 %	23 %	10 %	5 %
Δµ/µ	13 %	20 %	16 %	23 %
ΔV_T (V)	0.37	0.45	0.28	0.42
ΔN_q ($10^{11}eV^{-1}cm^{-2}$)	5	6.5	4	6

Fig. 4. Degradation rate of the channel resistance against gate voltage.

Table 2. Parameters of the degraded region of the channel, extracted for various cases. $V_{t1} = 0.78V$.

4 - DISCUSSION

We notice from Table 2 that the ratio ΔL/L can be very large for the higher gate voltage. On the other hand, it is obvious from Table 2 and Fig. 5 that a higher density of interface states are produced at lower gate voltage, as in the case of long MOSFET's. A basic argument is provided by the simulation with MINIMOS of the electric field along the channel (Fig.6). We obseve firstly that the peak of the lateral electric field is almost constant with V_g, unlike the case of longer transistors where this peak is greatly reduced with increasing V_g ; this is experimentally confirmed by the fact that the substrate current measued at fixed V_D does not change significantly with V_G from 1.4 to 3.3 V. Secondly, the vertical field is enhanced by a factor of two when V_g is increased from 1.4 V to 3.3 V, making, therefore, easier the electron injection along a larger part of the channel.



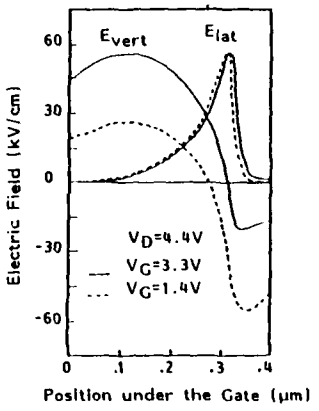


Fig. 5. Density of generated interface states versus drain current after one week stress.

Fig. 6. Profiles of vertical and lateral fields along the channel.

For V_g = 1.4V, not only the electron injection should occur in a narrower region near the drain, but also the vertical field is there inverted, causing the simultaneous injection of hot holes and electrons. As the formation of interface states is usually explained by the injection of both holes and electrons, it is reasonable to measure a higher generation rate of interface states in this case.

To summarize, we think that for small values of the gate voltage a large amount of interface states are created in a very localised region near the drain. For higher gate voltages, the defective region greatly extends towards the source. This region contains trapped negative charges and, at a lower degree, interface states. This difference in the extension of the degraded region is confirmed by $I_D(V_D)$ curves measured in the saturation region : after stress at low V_G the initial drain current value is recovered much earlier (i.e. for lower V_D) than after stress at higher V_G.

5 - CONCLUSION

The degradations of both the maximum transconductance and threshold voltage are greatly influenced by the extension of the defective channel length. The threshold voltage shift seems to be the most limiting factor for the lifetime of ultra short devices. Various degradation tests show that a 10 years *lifetime* can be achieved without LDD even for 0.3 µm transistors operated at 3 V. However, an increase of the supply voltage would need the use of LDD structures.

REFERENCES

/1/ G. Guegan et al., ESSDERC 88.
/2/ H. Haddara and S. Cristoloveanu, SSE, in press.
/3/ H. Haddara and S. Cristoloveanu, SSE, Vol.29, N°8, 1986.
/4/ H. Haddara and G. Ghibaudo, SSE, Vol.31, N°6, 1988.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

INTERFACE STATE GENERATION IN NMOS TRANSISTORS DURING HOT CARRIER STRESS AT *LOW TEMPERATURES*

D. KRISHNA RAO, M.M. HEYNS and R.F. DE KEERSMAECKER

Interuniversity Microelectronics Center (IMEC vzw), Kapeldreef, 75, B-3030 Leuven, Belgium

ABSTRACT - The generation of interface states in NMOS transistors during hot carrier injection is *investigated at temperatures between* 300 and 75 K. It is found that hot carrier stress at low temperatures induces a higher number of interface states than at 300 K. The generation rate of interface states can be fitted with a power law as a function of stressing time with a value of the exponent which is independent of the stressing temperature. A similar dependence on the stressing gate voltage is found for hot carrier stress at 75 K and 300 K. The results suggest a temperature independent mechanism of interface state creation. The important role of oxide hole traps in the interface state generation at low temperatures is demonstrated using two-step experiments.

1-INTRODUCTION

The degradation of MOS-transistors at low temperatures has gained renewed interest due to the advantages of low temperature operation of MOS-devices. However, this is associated with an increased hot carrier degradation due to the higher substrate currents at low temperatures caused by an increased mean free path of the channel electrons. The hot carrier degradation at room temperature in NMOS transistors, resulting from the charge injection and trapping into the gate oxide near the drain region, has been extensively studied in recent years. Although it is generally accepted that the degradation is caused by the injection of high energy carriers into the gate oxide, there is considerable disagreement concerning the physical mechanisms involved during the degradation and, in particular, on the generation mechanism of the hot carrier induced interface states /1-3/. Some disagreement results from the discussion on the interpretation of the measured data obtained from various measurement techniques. Using the charge pumping technique /4/ it has been shown that the number of interface states generated during hot carrier stress can be measured independently from the transistor degradation characteristics /5/. Therefore, in this study this technique is used to provide direct information on the interface state generation during hot carrier stress at various temperatures and stressing conditions.

2-EXPERIMENTAL DETAILS

The devices used in this study are NMOS transistors fabricated by a poly-gate technology. The gate oxide was grown in dry O_2 at a temperature of 950°C to a thickness of 26 nm. The source/drain regions are arsenic doped Transistors having a channel of 20 μm width and channel lengths 2.8 μm and 4.8 μm are subjected to hot carrier stress at various temperatures between 300 and 75 K. The low temperatures are obtained using a Leybold-Heraeus R210 2-stage refrigerator and ROK 10-300 refrigerator-cooled cryostat. The cryogenerator is operated using a closed helium-gas cycle. Samples can be kept at any constant temperature in the range between 10 K and 300 K. The devices are bonded in a package to ensure good electrical contact for low temperature measurements and a pressure of 5×10^{-2} mbar was maintained in the cryostat throughout the measurements. The interface state density (D_{it}, in $eV^{-1}cm^{-2}$) is obtained from the charge pumping current at the various temperatures after introducing the temperature dependent parameters in the charge pumping current equation /4/.

3-RESULTS

A typical result of a charge pumping experiment before and after a hot carrier stress is shown in fig. 1 for a device with a channel length of 2.8 μm which was stressed at 150 K near the maximum substrate current. The maximum of the charge pumping current is a measure for the integrated density of interface states (in cm^{-2}) while the rising and falling flanks of the charge pumping current represent the device threshold and the flatband conditions, repectively. As the sample is cooled down before the stress from 300 K to 150 K the threshold and the flatband voltage are shifted towards higher and lower voltages, respectively. At the same time the charge pumping current is increased, indicating a higher recombination current at the low temperature. This is due to an increase in the emission time constant of the interface states. As a consequence, a larger part of the Si-bandgap contributes to the charge pumping current. D_{it} before the stress is 7.7 10^9 and 9.4 10^9 cm^{-2} eV^{-1} at 300 K and 150 K, respectively. The difference is due to the increasing interface state density close to the Si-band edges. After the stress at 150 K an increase in the interface state density (ΔD_{it}) of 4.3 10^{10} cm^{-2} eV^{-1} is found, evidenced by a strong increase in the maximum of the charge pumping current. The negative shift of the

rising flank of the charge pumping current after the stress and the distortion of the charge pumping curve is an indication of the local positive charge build-up in the oxide near the drain end of the transistor. When the device is warmed to 300 K no change is found in either the interface state density or the trapped positive charge. Three general features follow from the low temperature stress experiments : 1) during the low temperature hot carrier stress the average increase of the interface states inside and outside the energy gap scanned during the charge pumping measurement at 300 K ($E_i \pm 0.38$ eV, where E_i is the middle of the Si-bandgap), 2) no additional interface states are generated during warming-up (further checked by cooling and re-measuring D_{it}) and 3) the positive charge trapped in the oxide at low temperatures does not occupy shallow hole traps since these would have been detrapped during warming up.

The generation rate of the interface states during hot carrier stress was investigated under stress conditions close to the maximum substrate current. To this purpose devices with a channel length of 2.8 μm were subjected to a hot carrier stress at various temperatures between 300 and 75 K during 20,000 s with $V_G = 4$ V and $V_D = 8$ V. The hot carrier stress was interrupted at different times at each temperature and a charge pumping experiment was performed. The result is shown in fig. 2(a). As the temperature is decreased from 300 K the interface state generation is found to increase. At 75 K ΔD_{it} after 20,000 s of stress is about 5 times higher than after an identical stress at 300 K. This increase in interface state generation is partially due to the increase in the substrate current at low temperatures. In fig. 2(b) ΔD_{it} divided by the ratio of the substrate to drain current at the stress temperature is plotted as a function of the stressing temperature. Since the ratio of the substrate to drain current is a measure of the avalanche multiplication in the channel during hot carrier stress, the y-axis in fig. 2(b) indicates the efficiency of the avalanche multiplication in generating interface states. As the temperature is decreased from 300 K to 75 K this efficiency is increased by a factor of 2.5. This suggests that the increase in the interface state generation at lower temperatures is not only due to the increase in hot carrier injection but also due to the increase in the efficiency of the charge injection and/or charge trapping in the oxide at lower temperatures.

In fig. 2(a) the stressing time (t) dependence of interface state generation at various temperatures is found to follow a relation $\Delta Dit = A.\ t^n$ where A and n are constants at a given temperature. For a given device and stress condition the value of n represents the rate of the interface state generation. The experimental data is fitted to the above relation and the constants A and n were obtained for different stress temperatures. The straight lines in fig. 2(a) are graphical representations of the fitting results. The n-values are found to vary only between 0.47 and 0.56 in the stress temperature range between 300 K and 75 K without any systematic dependence, suggesting that n is independent of the stressing temperature. As explained above, at low temperatures the enhanced interface state generation rates are largely due to the increase in the hot carrier injection and trapping in the oxide. Similar results were obtained in the study of the temperature dependence of hot-carrier induced transconductance degradation in NMOS-transistors /6/.

The dependence of the interface state generation on the gate voltage during stressing was investigated at 300 K and 75 K. The stress was performed over a period of 10,000 s with $V_D = 8$ V and V_G ranging from 1 V to 8 V. The channel length of the devices was 2.8 μm and 4.8 μm for the stressing temperature of 300 K and 75 K, respectively. The increase in interface state density normalized to its maximum is plotted as a function of the stressing gate voltage at both temperatures in fig. 3. The interface state generation at both temperatures is found to have a similar stressing gate voltage dependence. The maximum generation both at 300 K and at 75 K is occurring at bias conditions that correspond to the maximum substrate current. Such a behaviour at room temperature was reported before /5/. This again suggests that the interface state generation mechanism is independent of the stressing temperature.

Two-step experiments were performed to study the role of hole traps in the interface state generation at low temperatures. The experiments were carried out at 75 K. In a first experiment a device with a channel length of 4.8μm was first stressed for 1,000 s at $V_G = 1$ V, $V_D = 8$ V. This causes holes to be injected and trapped into the oxide /1/, as was evidenced by a negative shift in the rising flank of the charge pumping current after this stress. Subsequently hot electrons were injected by stressing at $V_G = 8$ V, $V_D = 8$ V /1/ for 1,000 s. This stress caused the rising flank of the charge pumping current to shift back to the original position, indicating the effective neutralization of the trapped holes by the injected electrons. In a second experiment identical stress conditions were used but the order of the stresses was reversed (i.e. first hot electron injection and then hot hole injection). ΔD_{it} after each step of the two experiments is given in Table 1. A large build-up of interface states is found after a hot hole injection while hot electron injection causes only a small increase (compare sample 1 with sample 2). The generation of interface states during the second stress is nearly independent of the first stress performed (compare sample 1 with sample 3 and sample 2 with sample 4). Therefore, no two-step process can be evidenced. The results also clearly demonstrate that even at low temperatures hot hole injection is more efficient in generating interface states than hot electron injection, in agreement with the results found at room temperature using gated diode avalanche injection /8/.

4-DISCUSSION

The generation of interface states during hot carrier stressing at conditions close to the substrate current maximum was observed to obey a power law dependence as a function of stressing time. The exponential factor was found to be independent of the stressing temperature. This suggests that the mechanism of interface state generation is independent of the temperature. This is also supported by the similar dependence of the interface state generation on the stressing gate voltage observed during hot carrier stress at 75 K and 300 K. Under conditions of maximum substrate current the electric field at the drain end of the channel reaches a maximum with the field directing both hot holes and hot electrons into the gate oxide /2/. The observed maximum of the interface state generation under these conditions was, therefore, related to the simultaneous injection of hot holes and hot electrons into the oxide under these conditions /5/. In the two-step experiments it was demonstrated that hot holes are much more efficient in generating interface states than hot electrons. The maximum interface state generation during simultaneous hole and electron injection can, therefore, be explained by the fact that under these conditions the neutralization of the trapped holes by the injected electrons counteracts the field build-up by the charged hole traps which in turn increases the hole injection and trapping probability.

In order to explain the important role of hole traps in the generation of interface states at room temperature a two-step process was proposed according to which the interface states are generated when the trapped holes are neutralized by injected electrons /9/. At low temperatures the generation of interface states in MOS-capacitors was shown to have a different behaviour /10/. The experiments were generally divided in two categories. In the first group of experiments the electrons which pass through the SiO_2 interface are able to immediately generate interface states at low temperature. This is consistent with the observed interface state generation under conditions of hot electron injection reported above. The efficiency of this generation mechanism is, however, very low. In the second type of experiments positive charge trapped in interfacial hole traps less than 2 nm from the Si/SiO_2 interface acts as a precursor for the interface states while positive charge trapped in hole traps further away from this interface (near-interfacial sites between 2 and 7 nm) can be neutralized at low temperature without generating interface states /11/. Holes are first trapped in the near-interfacial sites, and then converted to interfacial trapped holes by thermal energy or the application of high fields. In the two-step experiments on transistor structures, reported above, it was demonstrated that under conditions which favor hole injection, interface states are directly formed at the low temperature at a very high rate. These results suggest that under the experimental conditions used in these transistor experiments the interfacial hole traps are readily filled and neutralized during the stress, thereby generating interface states. This mechanism seems to be in operation both at 300 K and 75 K.

The fact that no two-step mechanism was evidenced in the transistor experiments could result either from the different field conditions during injection or from the fact that the stress conditions employed in the two-step process do not fully reflect the predicted injections (and that e.g. sufficient electrons are available for the two-step process to occur unnoticed). The fact that no interface states are generated when the positive charge generated during hole injection is neutralized by a subsequent electron injection suggests that the positive charge which is left after a hot carrier stress under conditions of hot hole injection is either located in the hole traps further than 2 nm away from the Si/SiO_2 interface or in re-structured interfacial hole traps. The latter suggestion is consistent with earlier findings demonstrating that the trapping of a hole and the generation of an interface state is an irreversible process causing permanent damage in the oxide and making the initial hole trap active as a slow state /12/. These slow states can be charged and discharged without generating fast interface states.

REFERENCES

/1/ Fair, R. B. and Sun, R. C., IEEE Trans Electron. Dev. ED-28 (1981) 83.
/2/ Hofmann,K. R., Weber,W., Werner,C. and Dorda,G., IEEE Trans. Electron. Dev. ED-32 (1985) 691.
/3/ Hu, C., Tam, S. C., Hsu, F. C., Ko, P. K., Chan, T. Y. and Terril, K. W., IEEE Trans. Electron. Dev. ED-32 (1985) 375.
/4/ Groeseneken, G., Maes, H., Beltran, N. and De Keersmaecker, R. F., IEEE Trans. Electron. Dev. ED-31 (1984) 42.
/5/ Heremans, P., Maes, H. E. and Saks, N. C., IEEE Electron. Dev. Lett. EDL-7 (1986) 428.
/6/ Tzou,J. J., Yao,C. C., Cheung,R. and Chan,H., IEEE Trans. Electron. Dev. Lett. EDL-6 (1985) 450.
/7/ Saks, N. S., Heremans, P. L., Van den hove, L., Maes, H. E., De Keersmaecker, R. F., Declerck, G. J., IEEE Trans. Electron. Dev. ED-33 (1986) 1529.
/8/ Heremans, P., Groeseneken, G. and Maes, H. E., presented at IEE "Colloquium on hot carrier degradation in short channel MOS", London, Jan 1987.
/9/ Lai, S.K., J. Appl. Phys. 54 (1983) 2540
/10/ Sung, J. M. and Lyon, S. A., Hu, G., Appl. Phys. Lett. 50 (1987) 1152.
/11/ Wang, S.J., Sung, J.M. and Lyon, S.A., IEDM Technical Digest (1987) 905
/12/ Heyns, M. M. and De Keersmaecker, R. F., presented at "The physics and Technology of Amorphous SiO_2", Les-Arcs, France, June 29 - July 3, 1987.

TABLE 1:Interface state generation during a two-step stress at 75 K.

Sample No	Device (NMOS;Leff=4.8µm;W=20µm) condition before stress	Hot carrier stress at 75K for 1,000s	ΔDit ($cm^{-2}.eV^{-1}$)
1	Virgin	V_G=8V;V_D=8V	$2.2\ 10^8$
2	Virgin	V_G=1V;V_D=8V	$4.4\ 10^9$
3	Pre-stressed with V_G=1V;V_D=8V for 1,000 s	V_G=8V;V_D=8V	$3.0\ 10^9$
4	Pre-stressed with V_G=8V;V_D=8V for 1,000 s	V_G=1V;V_D=8V	$5.0\ 10^9$

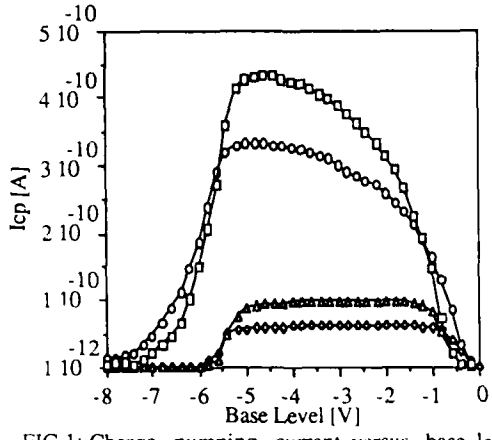


FIG.1: Charge pumping current versus base level voltage before and after an hot carrier stress at 150K. Device:Leff=2.8µm;W=20µm. Stress:V_G=4V; V_D=8V for 20,000 s. —●—: virgin at 300K, —■—: virgin at 150K, —□—: after stress at 150K, —○—: after warming to 300K.

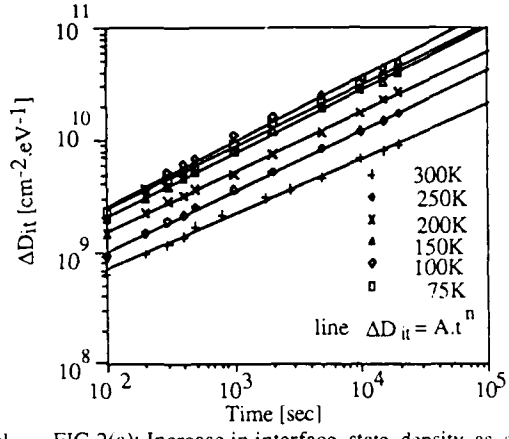


FIG.2(a): Increase in interface state density as a function of stressing time during hot carrier degradation at various temperatures. Device: Leff=2.8µm;W=20µm. Stress:V_G=4V;V_D=8V. Straight lines are fittings.

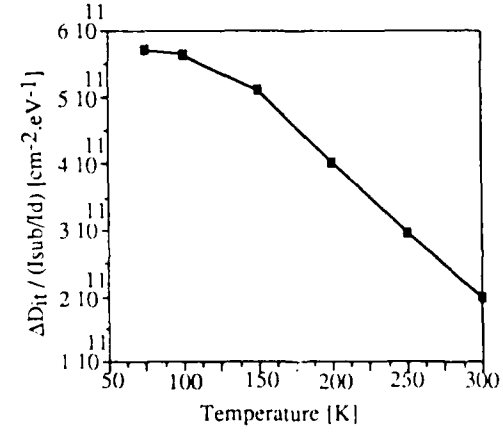


FIG.2(b): Increase in interface state density per unit avalanche multiplication at various temperatures. Stress conditions as mentioned in FIG.2(a).

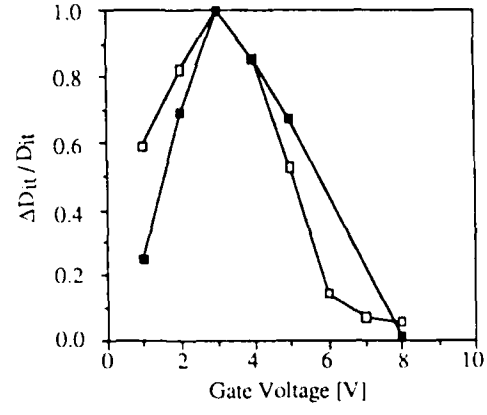


FIG.3: Stressing gate voltage dependence of interface state generation. Stress:V_D=8V;time 10,000 s. □:Device(Leff=2.8µm; W=20µm) at 300K. ■:Device(Leff=4.8µm ; W=20µm) at 75K.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

TWO-DIMENSIONAL COMPUTER SIMULATION OF HOT CARRIER DEGRADATION IN N.MOSFETs

M. GARRIGUES, A. ALEXANDRE, P. ROJO, T. PEDRON*, K. BELHADDAD* and A. PONCET*

Laboratoire d'Electronique, Automatique et Mesures Electriques, CNRS-UA 848, Ecole Centrale de Lyon, BP 163, F-69131 Ecully Cedex, France
**CNET, Centre Norbert Segard, BP 98, F-38243 Meylan Cedex, France*

Résumé : *On présente un simulateur bidimensionnel qui permet de calculer les distributions spatiales des courants d'émission de trous chauds et d'électrons chauds le long de l'interface $Si-SiO_2$ d'un transistor MOS. Les simulations sont comparées aux dégradations expérimentales pour un transistor à canal n et différentes conditions de contrainte. On montre que les effets de l'injection de trous sont seulement importants à faible tension de grille et que la corrélation connue entre la dégradation et le courant de substrat peut être attribuée uniquement aux électrons chauds.*

Abstract : *This paper presents a two-dimensional hot-carrier simulator which calculates the spatial distribution of hot electron and hot hole emission currents along the $Si-SiO_2$ interface of a MOS transistor. The simulation are compared to experimental hot-carrier degradations for a n-channel transistor and different stress bias conditions. It is shown that hole injection effects are only dominant at low gate voltages and that the usual correlation between degradation and substrate current may be attributed only to hot electrons.*

1. INTRODUCTION

Considering the extreme localization of the hot carrier degradations in MOSFET's, its tight dependance on stress bias conditions and technological parameters and its complex relation with both types of carriers, it is obvious that significant progress can only be gained now by a two-dimensional (2-D) simulation of hot carrier current densities for both electrons and holes as first demonstrated in [1].

In this paper, we present INJECT, a 2D specialized hot-carrier post-processor, connected to the previously reported 2D device simulator JUPIN developed in CNET-CNS [2]. An important feature of INJECT is that it includes a comprehensive model for hole heating and injection based on our previous study of uniform hot-hole injection [3].

2. SIMULATION MODEL

INJECT operates on the basic solution including ionization given by the device simulator JUPIN and composed of the 2D distributions of potential, electric field, current and carrier densities.

The model is based on a three step emission process similar to that proposed in [4] (Fig. 1) :

Step (1) : Carrier heating along the current lines by a "lucky-carrier" process.
Step (2) : Redirection toward the interface after an elastic scattering at point (x,y).
Step (3) : Free ballistic path from the scattering point to the interface.

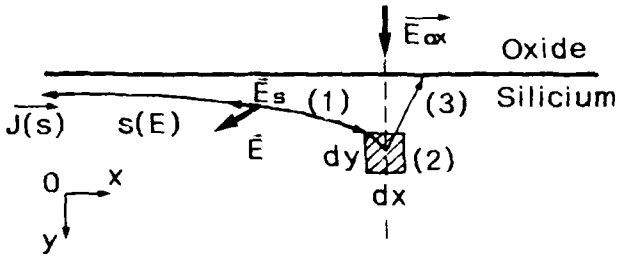


Figure 1 : Schematic of the three step model.

In step (1), the current density of hot carriers with an energy between E and E+dE at the scattering point (x,y) is calculated as :

$$dJ_1(x,y,E) = J(s)\ [\exp - \frac{s(E)}{mfp1}\ ]\ \frac{dE}{mfp1\ \mathcal{E}_s(s)}$$

- s(E) is the distance along the current line the carrier must travel to gain the energy E,
- mfp1 is the optical phonon scattering mean free path of the carrier (electron or hole),
- $\mathcal{E}_s(s)$ is the electrical field component along the current line at abcissa s(E),
- J(s) is the electron or hole density at abcissa s(E).

For step (2), we assume that the probability of an elastic scattering event is ds/mfp2 on a path ds. We note $P_2(E,E_o)$ the probability of a carrier scattered at energy E to have its component of energy directed to the interface greater than E_o.

For step (3), the probability of a free ballistic path from the scattering point (x,y) to the interface is taken as $P_3(y) = \exp[-y/mfp_3]$.

Finally, the current density of hot carriers with a perpendicular component of their kinetic energy greater than E_o at the interface is expressed as :

$$J_h(x,E_o) = \int_y [\ \int_{E_o}^{\infty} dJ_1(x,y,E)\ P_2(E,E_o)\ ]\ P_3(x)\ \frac{dy}{mfp2}$$

The values of the mean free path used in the model are [3],[4],[5] :

mfp1 = mfp3 = 95 Å, mfp2 = 600 Å for electrons ; mfp1 = mfp3 = 45 Å, mfp2 = 600 Å for holes.

$J_h(x,E_o)$ is the hot carrier current density hitting the interface but not necessarily collected by the gate depending on the local value of the vertical electric field. The "external" gate current $[I_G]_{n,p}$ for each type of carrier is calculated by integrating $J_h(x,E_o)$ along the interface for E_o equal to the local $Si-SiO_2$ energy barrier corrected for Schottky effect and for locations x where the vertical electric field has the correct sign for carrier collection (the energy-barrier values we used are : $\Phi_{Bn} = 3.2$ eV for e^- ; $\Phi_{Bp} = 3.8$ eV for h^+).

3. SIMULATION RESULTS

Simulations and experimental measurements have been compared for a typical n-channel transistor made in CNET-CNS using a conventional PolySi-gate CMOS process. The effective channel length, width and gate oxide thickness are respectively : $L_{eff} = 1.3$ µm ; W = 50 µm ; $t_{ox} = 25$ nm. The device description data for JUPIN are directly fed by the 2D process simulator TITAN.

External gate current

Figure 2 shows the result of the simulation for the external gate current I_G as a function of the gate voltage. A satisfactory fit with experimental measurements is found in the region of maximal electron gate current ($V_{DS} = V_{GS}$). However the simulation give a too fast "drop down" of I_G as V_{GS} decreases. It was impossible to make such comparison for the hole external gate current which is predicted by the simulator at low gate voltage, since the current is to much low and probably entirely trapped in the oxyde.

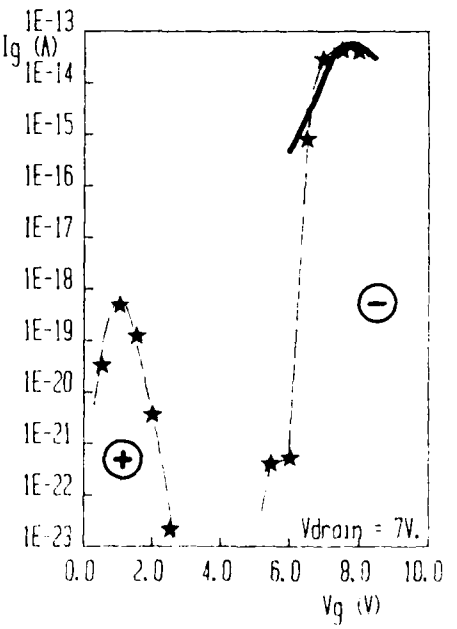


Figure 2 : Comparison between simulated (stars) and measured (solid line) values of the external gate current as a function of gate voltage V_{GS}, for $V_{DS} = 7$ Volts.

Hot-carrier distribution

We have simulated the hot electron and hole current densities $J_{hn}(x,E_n)$ and $J_{hp}(x,E_p)$ for E_n = 3.2 eV and E_p = 3.8 eV (approximate threshold energies for injection into the oxide), in the case of three typical stress conditions (Fig.3.a.b.c) :

(a) V_{DS} = 7 V, V_{GS} = 7 V, (b) V_{DS} = 7 V, V_{GS} = 3.5 V, (c) V_{DS} = 7 V, V_{GS} = 2 V

We notice that the hot electron distribution peak is always located above the drain, within 0.1 µm from the junction while the hot hole peak is located above the channel (also within 0,1 µm from the junction) due to the distance from the neutral drain for hole heating ("dead zone").

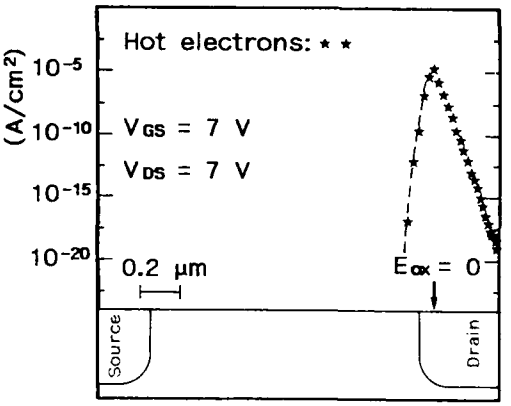
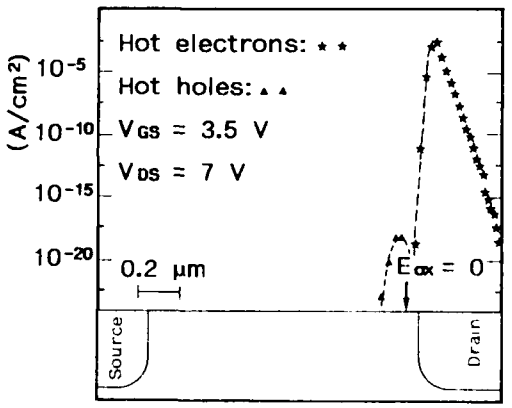
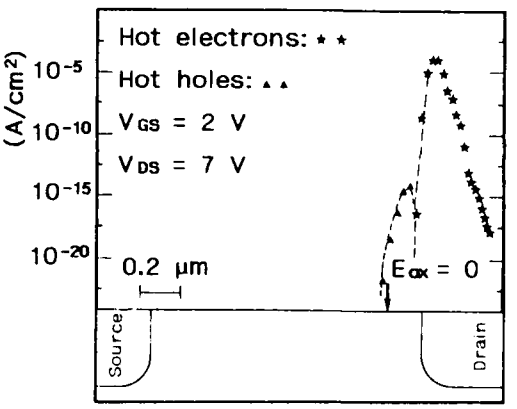


Figure 3 : Simulation results for the hot carrier injected density $J_h(x,E_O)$ for electrons and holes and three different bias conditions, noted on the figures. For V_{DS} = V_{GS} = 7 V, there is no significant hole injection.





The peak values of the hot hole and hot electron distribution are listed in the table below together with the simulated substrate current.

Gate voltage V_{GS}	2 Volts	3,5 Volts	7 Volts
Maximum hot electron current density	1.7 10^{-4} A/cm²	1.2 10^{-3} A/cm²	1.5 10^{-5} A/cm²
Maximum hot hole current density	9.5 10^{-15} A/cm²	8.5 10^{-19} A/cm²	3.2 10^{-28} A/cm²
I_{sub}	- 121 µA	- 265 µA	- 45 µA

The most noticeable result is the very strong increase of the hot hole density as V_{GS} decreases. Particularly the maximum value obtained at V_{GS} = 2 V while the substrate current is two time lower than at V_{GS} = 3.5 V is surprising. Our explanation is that the effect of the enhancement of the maximum lateral electric field E_M strongly overcompensates the decrease of the channel current.

If we consider hot electrons, it can be noticed that their flux is maximum at V_{GS} = 3.5 volts (while I_O is negligible due to the distribution of the vertical electric field).

4. COMPARISON WITH EFFECTIVE DEGRADATIONS

Transistors have been stressed for the simulated DC bias conditions during 10^5 seconds. The following parameters were used to characterize the degradations : V_{TH}, the gate voltage that gives a drain current of 50 µA at V_{DS} = 0.5 V ; g_m, maximum transconductance at V_{DS} = 0.5 V ; I_{SUB} substrate current. The corresponding degradations are listed in the table below :

	V_{GS}	2 V	3.5 V	7 V
Initial stress	ΔV_{TH}	$- 5\ 10^{-3}$ V	$+ 3\ 10^{-2}$ V	$+ 1.5\ 10^{-2}$ V
	$\Delta g_m/g_{mO}$	$+ 2\ 10^{-2}$	$- 6\ 10^{-2}$	$- 4\ 10^{-2}$
	$\Delta I_{SUB}/I_{SUB}$	$- 8\ 10^{-2}$	$- 5\ 10^{-2}$	$+ 5\ 10^{-2}$

The results of the first stress show a "regular" behaviour of the device submitted to DC stress : maximum degradation at maximum substrate current, evidence for electron trapping at V_{GS} = 7 V and for hole trapping at V_{GS} = 2 V.
A second stress was performed to "reveal" trapped holes by injecting only electrons at $V_{DS} = V_{GS}$ = 7 V during 7 hours, just after the initial stress at V_{GS} = 2 V (conditions of enhanced aging as described in [6] or [7]). It clearly reveals that at V_{GS} = 2 V, an appreciable amount of positive charges was trapped which were neutralized by the subsequent electron injection ($\Delta V_{TH} = + 8\ 10^{-2}$ V, $\Delta g_m/g_{mO} = - 5\ 10^{-2}$, $\Delta I_{SUB}/I_{SUB} = + 4\ 10^{-2}$).

Discussion :

i) Considering the relative values of stimulated emission currents in the three bias conditions, hole emission can be made responsible of appreciable degradations only at low gate voltage ;

ii) It implies that the stronger degradation observed at $V_{GS} = V_{DS}/2$ is not due to the effect of combined electrons and holes as sometimes stated but, in fact, to hot electrons alone whose flux is found to be correlated to substrate ionization current (as shown analytically in [4]) ;

iii) The result of sequential hole-electron injection, together with the relatively high hole injection current support the hypothesis of predominant hot hole induced defects at low V_{GS}.

We can further notice that there is not a direct correlation between numerical hot electron densities and corresponding degradations. We think that in the case of a vertical field repelling electrons (as for V_{GS} = 3,5 V), it would be necessary to take also into account the penetration distance of the electrons into the oxide (as in [1]).
Concerning the simulated hot hole emission currents, it appears that the computed values are numerically too much low : assuming a maximal one to one correlation between hot holes and trapped charges (revealed after hot electron injection), the hot hole current density must be about 100 times the simulated value. This result can be corrected by adjusting the mean free path mfpl for holes to a slighly higher value, about 50 ÷ 55 Å.

5. CONCLUSION

We have described a fully functional hot carrier simulator which bring new and useful informations for the understanding of MOS transistors degradation under bias stresses. Our results indicate that hot hole emission can only induce significant degradations at gate voltages near threshold. Hot electrons would be fully responsible of aging at higher gate voltages.
However, quantitative predictions, although promising, are only in infancy, due to remaining ignorances of the defect creation processes.

ACKNOWLEDGMENTS : *This work was partially supported by "Groupement Circuits Intégrés Silicium".*

REFERENCES

[1] K.R. Hofmann, W. Weber, C. Werner and G. Dorda, IEEE Trans.Electron Dev. - ED-32, 691 (1985)
[2] K. Belhaddad, C. Corbex, A. Poncet, NASECOD IV (1985)
[3] Y. Hellouin, F. Chehade, M. Garrigues, J.Appl.Phys. 61, 5342 (1987)
[4] S. Tam, P.K. Ko, C. Hu, IEEE Trans.Electron Devices, ED-31, 1116 (1984)
[5] M. Garrigues and Y. Hellouin, IEEE Trans.Electron Devices, ED-28, 928 (1981)
[6] B.S. Doyle, M. Bourcerie, J.C. Marchetaux and A. Boudou, IEEE Elect.Dev.Lett. - EDL-8, 237 (1987)
[7] R. Bellens, P. Heremans, G. Groesenken, H.E. Maes, IEEE Elect.Dev.Lett. - EDL-9, 232 (1988)

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

SPICE MODEL FOR TRANSIENT ANALYSIS OF EEPROM CELLS

R. BEZ, D. CANTARELLI, P. CAPPELLETTI and F. MAGGIONI

SGS-Thomson Microelectronics, Via C. Olivetti 2, Agrate Brianza (MI), Italy

Abstract. *An equivalent model for the EEPROM cell, described on a SPICE circuit analyzer, is used for the transient analysis of erase(E)/write(W) characteristics with ramp waveform programming pulses. The results of the simulation are compared with the experimental data obtained by an innovative method for measuring E/W curves. The validation of the model is made over different variations of the cell lay-out, different programming voltages and different rise times for the ramp of the programming pulse.*

1 Introduction

The two transistors FLOTOX cell (Fig. 1*A*) is the most widely used among the memory devices proposed for Electrically Erasable Read Only Memories (EEPROM). The first analytical model for this type of EEPROM cell has been proposed by Bhattacharyya in 1984 [1]. A more detailed circuit model has been implemented by Kolodny et al.[2] in order to simulate the shift of the cell threshold voltage after step programming pulses.

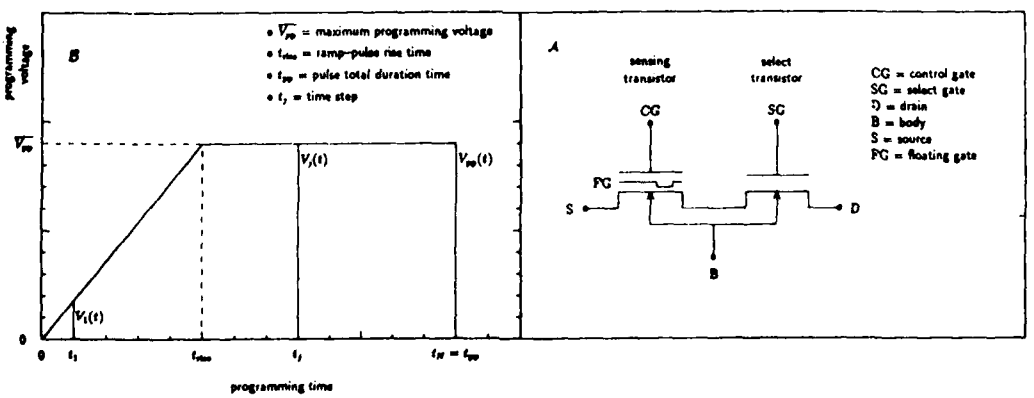


Figure 1. *A*: two transistors FLOTOX cell scheme.
B: a typical ramp waveform programming pulse. $V_{pp}(t)$ is the complete pulse; $V_j(t)$ is the $V_{pp}(t)$ stopped at $t = t_j$

In actual device operation the cell is programmed with a ramp waveform pulse in order to minimize the tunnel current (the peak value of programming current is known to be proportional to the ramp speed): in this case the threshold shift is a function of both the ramp rise time and the pulse width for a fixed pulse amplitude. Therefore an accurate simulation of the transient behavior (*i.e.* threshold shift *vs.* time) of the cell during the programming cycle is desirable.
In this work an equivalent model for transient analysis of EEPROM cell is proposed. The model is described in term of SPICE [3] components in order to utilize a well proven MOSFET capacitance model. A further advantage in using SPICE is the linkability of the cell model to other circuital blocks for global simulation. An innovative experimental method, able to measure transient E/W characteristics, is used for validating and optimizing the model.

2 Experimental

E/W curves are defined as threshold voltage *vs.* time, where, for operative reasons, the threshold voltage (V_{CG}^{T}) is defined as the control-gate voltage measured at fixed drain current (I_D = 10μA with V_D = 2V and V_S = V_B = 0V). Erasing (writing) is defined as the programming operation, which sets the negative (positive) charge state on the floating-gate, *i.e.* the high (low) threshold voltage condition.
E/W curves are measured for programming pulses $V_{pp}(t)$ with a ramp waveform (Fig. 1*B*). A block-scheme of the experimental setup is shown in figure 2*A*. The programming pulses are generated by an *HP 8160A* programmable pulse generator. The threshold detector is a current voltage converter, which forces on the control-gate the voltage needed to sustain a 10μA drain current; this *ad hoc* designed circuit allows a faster measurement of the threshold voltage, which is measured by a *Keithley 192 DMM* voltmeter. All instruments are controlled by an *HP 300* computer.

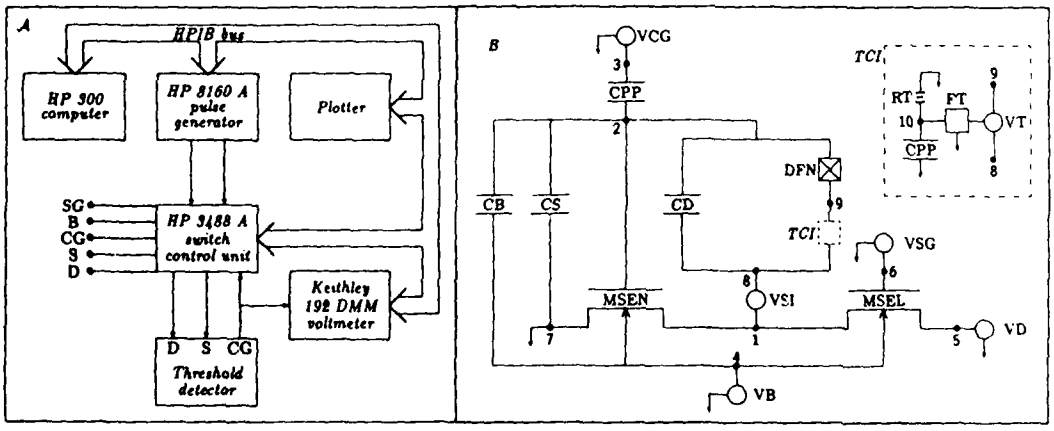


Figure 2. A: block-scheme of experimental setup.
B: SPICE equivalent model for the FLOTOX cell. Element description: MSEN=sensing transistor; MSEL=select transistor; CPP=CG-FG capacitance; CX=FG-X capacitance (X=D,S and B), VY=voltage source (Y=CG,SG,D and B); VSI=voltage source simulating voltage drop in silicon; DFN=Fowler-Nordheim diode; *TCI*=tunnel current integrator. Node description: 1=sensing transistor drain; 2=FG; 3=CG; 4=B; 5=D; 6=SG; 7=S; 8=silicon surface underneath tunnel oxide; 9,10=unreal nodes added for simulation purposes.

In order to clarify how the E/W characteristics are obtained let us refer to the writing curve. The programming pulse timebase is divided into N time steps $t_1..t_j..t_N = t_{pp}$. After cycling the cell few tens of

times in order to stabilize the threshold shift, N values of the threshold voltage ($V_{CG_j}^T$) are measured after stopping the writing pulse at each time step; before each $V_j(t)$ writing pulse, complete (i.e. $V_{pp}(t)$) E/W pulses are applied to reset the cell to the initial erased state. The writing curve is obtained by plotting the N values $V_{CG_j}^T = V_{CG}^T(t_j)$ *vs.* the programming time.

3 Simulation

The equivalent circuit shown in figure 2*B* is described in terms of SPICE components to simulate the EEPROM cell; the description of each component of the circuit is made in the figure caption. The "tunnel diode" (DFN) model can be easily introduced in SPICE using the well known Fowler-Nordheim (FN) equation [4]: external parameters are the pre-exponential and the exponential constants (usually called A and B), the tunnel oxide thickness and the tunnel oxide area. The Ward-Dutton [5] charge model is used for simulating the gate-to-channel capacitance of the sensing transistor. A voltage source (*VSI*) is introduced to take into account the voltage drop in the n^+ depleted region underneath the tunnel oxide during the writing cycle. A Tunnel Current Integrator *TCI* is added to the circuit to determine at the node 10 a voltage value equivalent to the threshold shift.
The DFN parameters are fixed, using theoretical values for A and B ($A = 1.67 \times 10^{-6} AV^{-2}$, $B = 221 MVcm^{-1}$) and the measured ones for the tunnel oxide thickness and area. The capacitance values are chosen as fitting parameters of the model, because they can hardly be determined from experimental data. Anyway the fitting values are within a 15% variation from the lay-out capacitances.
In order to compare simulations and experimental data is useful to put an initial condition on node 10 to obtain the threshold voltage instead of the threshold shift. Moreover an initial condition can be put also on node 2 to simulate the charge on the floating-gate due to a programmed state.

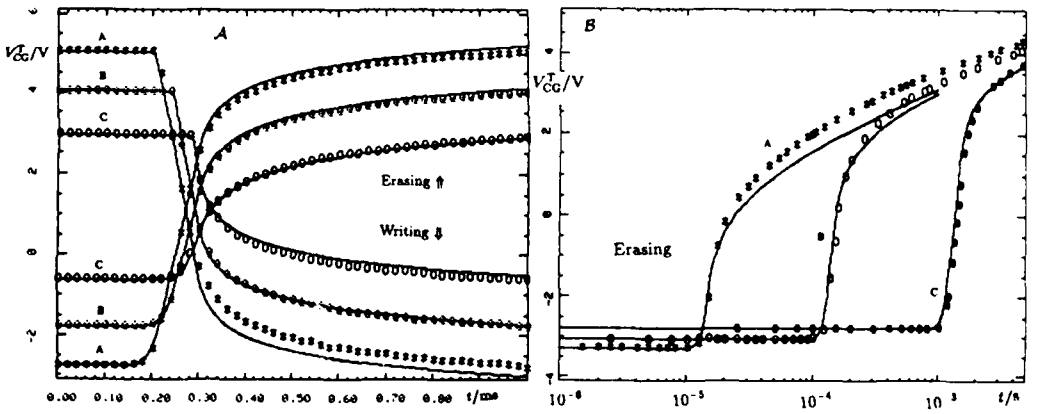


Figure 3. Comparison between simulated (solid line) and measured (symbols) E/W characteristics (i.e. threshold voltage V_{CG}^T *vs.* time t) of the standard cell (SC). *A*: the ramp waveform programming pulse $V_{pp}(t)$ is used with the same t_{rise} (t_{rise} = 300μs) and the same t_{pp} (t_{pp}=1ms) but with different V_{pp} (A: V_{pp}=16V; B: V_{pp}=15V; C: V_{pp}=14V). *B*: the $V_{pp}(t)$ is used with the same V_{pp} (V_{pp}=15V) and the same t_{pp} (t_{pp} = 5ms) but with different t_{rise} (A: t_{rise} = 15μs; B: t_{rise} = 150μs; C: t_{rise} = 1.5ms).

4 Discussion

First the analysis are centralized on a standard cell (SC) operating with $\overline{V_{pp}}$ at different values with the same t_{rise} (Fig. 3$\mathcal{A}$): a good agreement (an error of 3% in the worst case) between simulation and experimental data is found for the complete E/W curve and in particular for the final threshold shift.
Secondly, the model is validated on different cells, each of one having a single lay-out variation with respect to the SC, *e.g.* the tunnel oxide area or the active area are increased; modifying only those physical parameters in the SPICE EEPROM model that are interested in the lay-out variation, the experimental data are again fitted with a good accuracy (Fig. 4).
Finally different t_{rise} and different t_{pp} with the same $\overline{V_{pp}}$ are considered, operating on the SC. In this case the same fitting parameters are not able to give an acceptable fitting, if t_{rise} and/or t_{pp} are changed over wide ranges (Fig. 3$\mathcal{B}$). These discrepancies are attributed to the inaccuracy of the theoretical FN equation to model the actual tunnel current, which is known to deviate from the FN equation at high electric fields [6,7]. In conclusion:

1. the SPICE EEPROM model is able to simulate correctly the transient behavior of the cell (and indeed the physical E/W mechanism), but using fitting parameters;

2. the parameters, found by fitting, are able to take into account cell lay-out variation and/or different $\overline{V_{pp}}$ values, but they are not for different ramp speed;

3. to obtain a self consistent EEPROM model, able to simulate any programming condition, the exact form of the tunnel current in the high electric field range or, at least, different experimental A and B values for different electric field ranges, should be used in the DFN model.

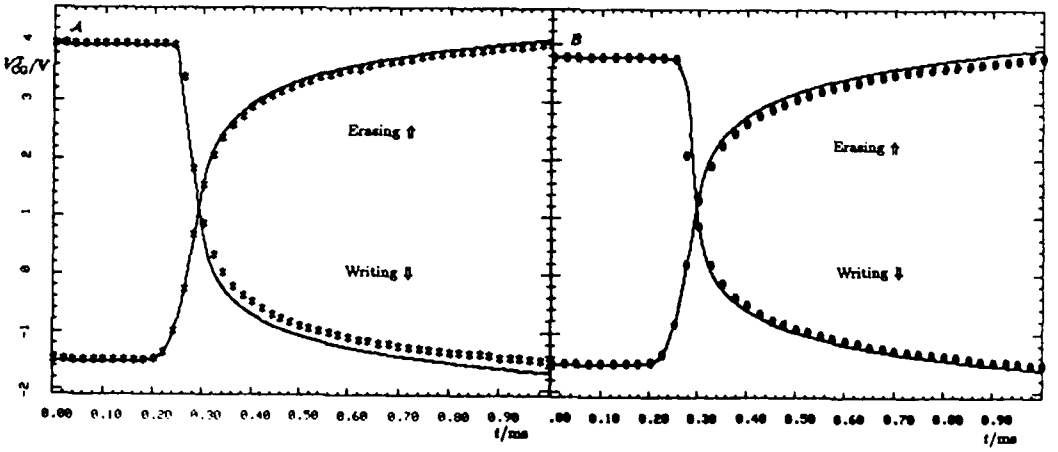


Figure 4. Comparison between simulated (solid line) and measured (symbols) E/W characteristics ($\overline{V_{pp}}$=15V, t_{rise} 300μs and t_{pp}=1ms) of different cells, having a single lay-out variation with respect to the SC. $\mathcal{A}$: tunnel oxide area increased by 40%; $\mathcal{B}$: active area increased by 7%.

[1] A.Bhattacharyya, *Solid State Electronics*, **27**,899,1984
[2] A.Kolodny, *IEEE Trans. Electron Devices*, **6**, 835, 1986
[3] *SGS-Thomson SPICE Version*, 1987
[4] M.Lenzlinger and E.H.Snow, *J. Appl. Phys.*, **4**, 278, 1969
[5] D.E.Ward, *Technical Report G201-11*, Stanford Electrical Laboratories, 1981
[6] M.Itsumi, *J. Appl. Phys.*, **52**, 3491, 1981
[7] M.Shatzkes and M.Av-Ron, *J. Appl. Phys.*, **47**, 3192, 1976

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CAD MOSFET MODEL FOR EPROM CELLS

N. BALLAY and B. BAYLAC

SGS-Thomson Microelectronics, Central R&D/CIS/TDMC, BP 217, F-38019 Grenoble Cedex, France

Résumé - Un modèle CAO de transistor MOS pour la simulation électrique d'une cellule EPROM en mode d'écriture est proposé. Ce modèle est aussi proche que possible de la physique pour permettre au concepteur l'évaluation de plusieurs schémas et variantes de procédé. Il est valable en mode de claquage et prend en compte le courant de grille induit par les électrons chauds.

Abstract - A CAD model of MOS transistor suitable for circuit simulation of an EPROM cell during writing cycle is proposed. This model is as close as possible of physics to allow the designer to evaluate different schemes and changes in the process variables. It is reliable in the breakdown mode too and takes into account the gate current induced by channel hot electrons.

INTRODUCTION

This present CAD model of MOS transistor, including gate current calculation and reliable even in the breakdown mode, should be well suited to simulate EPROM cells during writing cycle when implemented into a SPICE-like program. Such numerical simulator for 3D devices, including snap-back, gate current and transient should not be available in the near future. The assumed equivalent scheme described in chapter 1 includes both MOS and intrinsic bipolar transistors and serie resistors. The sensitivities of snap-back voltage versus different electrical parameters are reported in chapter 2. Its impact on the gate current model based on the well known Lucky Electron model /1/ is then developped in chapter 3. First simulation results, concerning the programming time of an EPROM cell are finally given.

1 - EQUIVALENT ELECTRICAL SCHEME

The device used in the EPROM cell is an MOS transistor whose floating gate potential is induced by coupling between control gate, source and drain. We now only consider the internal MOS device whose gate is the floating one. This device can't be described by a conventionnal MOSFET model because it is operating in or near snap-back regime. When an N-channel MOSFET is operating in saturation region, the electric field near the drain junction can be large enough to cause impact ionization. The generated electrons are then swept into the drain, while the holes move into the substrate. This excess substrate current develops a significant voltage drop across the substrate resistance, and biases the source to bulk junction in the forward direction. Then, the channel current or the avalanche multiplication seed current

increases, causing significant positive feedback leading to snap-back. Moreover, as some of the emitted electrons can be collected by the drain, this positive feedback is enhanced by the gain of the lateral parasitic bipolar transistor /2/3/4/5/.

That is why our electrical equivalent scheme must include an MOS transistor, a bipolar one and a substrate serie resistance. It is also clear that source and drain resistances must also be included as they act like a negative feedback on the contrary. The ionization current induces the breakdown mode and is so the main phenomenon governing breakdown. It is considered in our model like a current source related to the current flowing in the MOS and bipolar transistors and the electric field. The whole circuit is shown in figure 1.

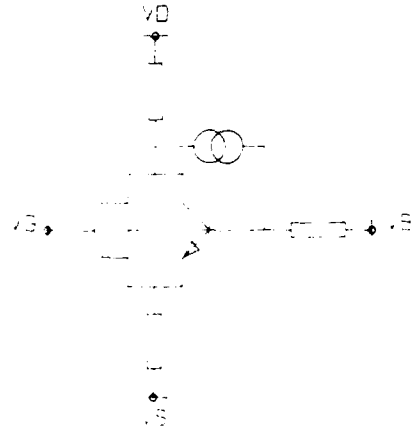


Fig. 1 -Equivalent electric scheme of EPROM device.

2 - SNAP-BACK PREDICTION

We now present some results who give an idea of the ability of the program to fit the reality. Note that the algorithm used to solve the system of equations is a specific one based upon Marquartd's work /6/ which has been modify to fit our problem.

The geometry of the device plays a predominent role as, when the channel length is decreasing :

i) the substrate current is increasing
ii) the effective substrate resistance is higher
iii) the gain of the parasitic BJT is increasing

The channel width also modifies the breakdown characteristics through the isolation implant acting as a shunting resistor which is no more negligible when the width is small. The ratio of the substrate and channel doping levels also plays an important role on the breakdown as the ohmic drop across this resistor is applied either between drain and source (best case) or between source and substrate (worst case).

Figure 2 gives an idea of the agreement between theory and experiment, achieved for a whole set of geometries with the same parameters. On a long and wide MOST one should see larger breakdown voltage and current only flowing between drain and bulk.

- Dotted line : Experiment
- Continuous line : Theory

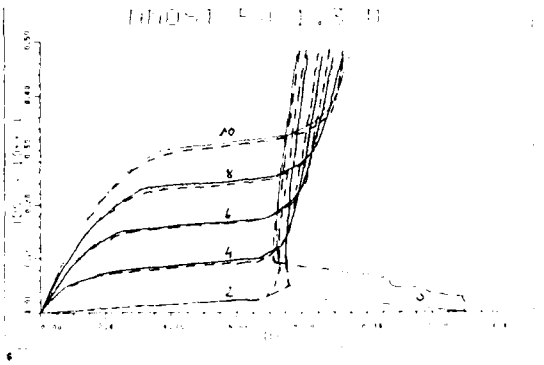


Fig. 2 - Comparison of theory and experiment for the drain current of a short channel MOST.

3 - GATE CURRENT MODELLING

The probability of a carrier to jump the oxide potential barrier is the product of a first probability to gain sufficient energy (more than 3.2 eV for an electron) and of a second one to be able to reach the gate electrode without any collision.

Avalg = AG* EXP (-Phib/(q*Ex*l)) * EXP (-x0/lox)

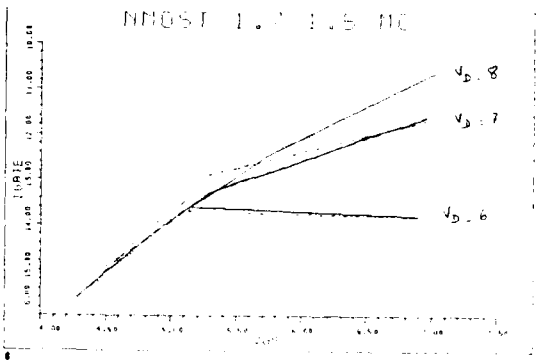
- l = mean free path without any collision in Si
- Phib = potential barrier to jump

The potential barrier may be less than 3.2 eV if the electric field in the oxide is high because of the image force lowering or Schottky effect /7/.
- x0 = SQRT (q/(16*PI*Epsox*Eox)) bounded by hox (oxide thickness).
- lox = mean free path without any collision in SiO2.

Interaction with snap-back is taken into account through the electric field Ex depending on the current density.

The overall probability of injection must now be calculated everywhere in the channel. If the hotest electron are indeed located at the drain because of maximum longitudinal electric field, these electrons might never be collected by the gate (in the case of VGD < 0). That means that the first probability to gain enough energy is increasing when the electron goes in the direction of the drain, but the second one is decreasing in the same time when VGD is negative. The location of the maximum injection point and the values of the two probabilities at this point are very difficult to get analytically. That is why, in our model, we perform a numerical integration of the local injection probability along the channel to get the overall gate current.

On figure 3 is a comparison of theory and experiment for the gate current of a device with W/L = 1.7/1.5. The experimental current has been measured through post-processing of the time dependent threshold voltage of a floating gate transistor.



- Dotted lines : Experiment
- Continuous line : Theory

Fig. 3 - Comparison of theory and experiment for the gate current of the memory cell MOST (1.7/1.5)

4 - CONCLUSION

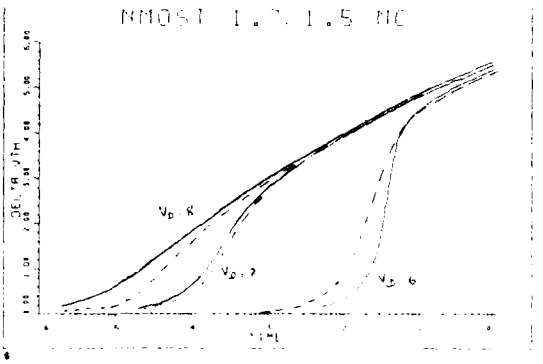
We developped a MOSFET model devoted to EPROM writing simulations. This model includes :

- substrate current
- snap-back (bipolar action)
- gate current

The results we get are very promising when compared to experiment. The major difficulties and uncertainties are :

- evaluation of the effective substrate resistance
- second breakdown behaviour
- (hole) gate injection in snap-back mode

Transient simulations have already been done (see figure 4) but the introduction into a SPICE like simulator would be useful to carry on accurate simulations of the whole EPROM cell, including the load line.



- Dotted line : Experiment
- Continuous line : Theory

Fig. 4 - Threshold voltage versus programming time for an EPROM cell.

REFERENCES

/1/ T.H. Ning, C.M. Osburn, H.N. Yu, Appl. Phys. Let., Vol. 48, 1977

/2/ T. Tobaye, K. Yamaguchi, S. Asai, M.S. Mock, "A Numerical Model of Avalanche Breakdown in MOSFET's", IEEE Trans. Electron Devices, vol. ED-25, no.7, Jul. 1978.

/3/ E. Sun, J. Moll, J. Berger, B. Alders, "Breakdown Mechanism in Short-Channel MOS Transistors", IEDM Tech. Dig., 1978

/4/ W. Muller, L. Risch, A. Schutz, "Short-Channel MOS Transistors in the Avalanche-Multiplication Regime", IEEE Trans. Electron Devices, vol. ED-29, no 11, Nov. 1982

/5/ F.C. Hsu, P.K. Ko, S. Tam, C. Hu, R. Muller, "An Analytical Breakdown Model for Short-Channel MOSFET's", IEEE Trans. Electron Devices, vol. ED-29, no 11, Nov. 1982

/6/ D.W. Marquardt, "An Algorithm for Least-Squares Estimation of Non linear Parameters", J. Soc. Indust. Appl. Math., vol. 11, no 2, June 1963

/7/ G. Barbottin, A. Vapaille, "Instabilities in Silicon Devices", Vol. 1, Part 3, North-Holland, 1986

Session 6C Room C : Materials, epitaxies, HJD

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

VERY RAPID GROWTH OF HIGH QUALITY GaAs, InP AND RELATED III-V COMPOUNDS

M. DESCHLER, K. GRÜTER, A. SCHLEGEL, R. BECCARD, H. JÜRGENSEN and P. BALK

Institute of Semiconductor Electronics, Aachen Technical University, D-5100 Aachen, F.R.G.

Abstract

Application of the hydride growth process at overall pressures below $5x10^3$ Pa yields very high rates of deposition for binary (GaAs, InP) and ternary (GaInAs, GaAsP) III-V semiconductor films. It is proposed that at these conditions the hydrides AsH_3 and PH_3 reach the substrate surface undecomposed and react directly with the group III chloride (GaCl, InCl). The rates can be convenently adjusted over a wide range by varying the reactant pressures. Films of excellent morphology with low background doping, high electron mobilities and well resolved PL spectra were obtained.

Résumé

Des vitesses de croissance pour les semiconducteurs III-V binaires (GaAs, InP) et ternaires (GaAsP, GaInAs) très élevées ont été obtenues en conduisant le procédé de croissance dans le système à hydrure à une pression totale inferieure à 5.10^3 Pa. A ces conditions , les hydrures AsH_3 et PH_3 semblent atteindre le substrat non décomposé et réagissent directement avec les chlorures du groupe III (GaCl, InCl). Les vitesses de croissance peuvent etre ajustées dans un domaine étendu par une variation de la pression partielle des différents composants de la réaction. Les couches possèdent une bonne morphologie de surface, un faible niveau de dopage et une grande mobilité des électrons. De plus, les spectres de photoluminescence obtenus sur ces couches montrent des signaux bien résolus.

Introduction

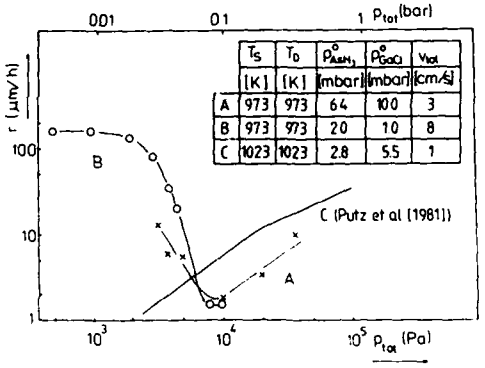
The hydride method has proved to offer a convenient approach for growing high quality epitaxial films for optoelecronic devices like photodetectors and LEDs /1/. For GaAs, operation at reduced pressures (10^4 Pa) is advantageous from the point of view of obtaining improved film uniformity /2/. However, this decrease in pressure also causes the growth rate to drop to rather low values /3/, which is disadvantageous. When exploring growth at still lower pressures we observed to our surprise that the rate increased again, so that in the pressure range below 10^4 Pa extremely fast deposition can be achieved. In this paper we will present first results on this study. We will also show that the phenomena of high growth rates at low pressures is not limited to GaAs; it was also found for InP and for ternary materials (GaInAs, GaAsP) studied so far.

Experimental

The growth experiments were performed in a hydride system /4/ designed for operation at low pressures. Accordingly, the Ga and In sources were protected from backdiffusion of reactants using baffles. A regulating valve permitted adjustment of the pressure in the resistance heated quartz reactor. Semi insulating (100) GaAs and InP wafers, misoriented 2^o towards the (100) direction, were employed as substrates.

Results and discussion

The initial drop in rate upon reducing the overall pressure and the subsequent increase for still lower values are demonstrated for the case of GaAs by the data in fig. 1. The detailed shape and position of the curves depends on the temperature, the linear flow velocity and the partial pressures of the reactants. In this example rates of well over 100 μm/h are attained.



	T_S [K]	T_D [K]	$p^o_{AsH_3}$ [mbar]	p^o_{GaCl} [mbar]	v_{tot} [cm/s]
A	973	973	64	100	3
B	973	973	20	10	8
C	1023	1023	28	55	1

Fig. 1: Dependence of GaAs growth rate on total pressure. Curve a from ref. /3/Fig. 1:

The growth rate curve appears to exhibit four different regions. Just below p ~ 10^5 Pa (curve a) the rate appears to drop roughly with the root of the pressure. This behavior results from the equilibrium

$$GaCl + 1/4\ As_4 + 1/2\ H_2 \rightleftarrows (GaAs)_s + HCl, \qquad (1)$$

where the GaCl is obtained from reaction between injected HCl and an elemental Ga source and the As_4 results from the thermally cracked arsenic source AsH_3. The GaAs activity near the substrate, which is a measure of the growth rate, follows immediately from the above equation:

$$a_{GaAs} = K_1 \times P_2^{1/2} \times P_{GaCl} \times P_{As_4}^{1/4} \times P_{HCl}^{-1} \qquad (2)$$

This equation shows the above mentioned square root dependence on the pressure of the H_2 carrier. Below 2×10^4 Pa the slope of the curves increases (curves a and b) and shows a linear pressure dependence. As proposed in an earlier paper from this laboratory /5/ in this pressure range a kinetic limitation related to the reduction of adsorbed AsGaCl complexes by H_2 molecules occurs. Below approximately 5×10^3 Pa the rate turns up (curves b and c). This phenomenon is the topic of this paper. We suggest that at these relatively low pressures the AsH_3 is not or not completely cracked anymore, so that a substantial density of AsH_3 (or AsH_2 or AsH) reaches the substrate surface. This leads to a change in the growth reaction

$$GaCl + AsH_3 \rightleftarrows (GaAs)_s + HCl + H_2 \qquad (3)$$

which is thermodynamically favored over the process described by eq. (1) because of the positive free enthalpy of formation of this the compound. In the 10^3 Pa region the growth rate saturates. We attribute this effect to the fact that here all GaCl available near the substrate reacts with AsH_3 (or AsH_2 or AsH).

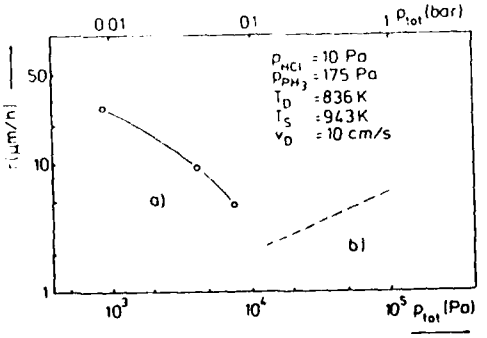


Fig. 2: Dependence of InP growth rate on pressure.
a) this work
b) from /6/

As would be expected, the analogous effect of incomplete dissociation of PH_3 is observed in the low pressure growth of phosphorus compounds. Fig. 2 shows some early data from a study on the growth of InP. In this case growth rates of 35 μm/h were achieved. A further advantage of deposition at low pressure conditions is the large range available for adjusting the rates of deposition of the binaries. These can be varied by almost three orders of magnitude (0.5 - 300 μm/h) for GaAs by varying the HCl and AsH_3 imput pressures. The dependence of the rate on these partial pressures is roughly linear, in contrast to the situation near 10^5 Pa, where the dependence on the reactant pressures is rather weak. For InP a range of 0.3 to 35 μm/h was obtained in the same manner.

Similar high rates (around 60 μm/h) were obtained at low pressures for growth of the ternary GaInAs on InP. Proper adjustment of the lattice constant, i.e. of the Ga/In ratio in the solid, is attained by varying the HCl input flows in the metal source regions or, as shown in fig. 3, by varying the substrate temperature at fixed source temperature. The deposition rate of GaAsP on GaAs at low total pressure is dependent on the ratio of the hydride (AsH_3 and PH_3) pressures and is for this reason a function of the composition of the solid. Around $GaAs_{0.6}P_{0.4}$, a composition typically used for red LEDs, a value as high as 170 μm/h was achieved for this material.

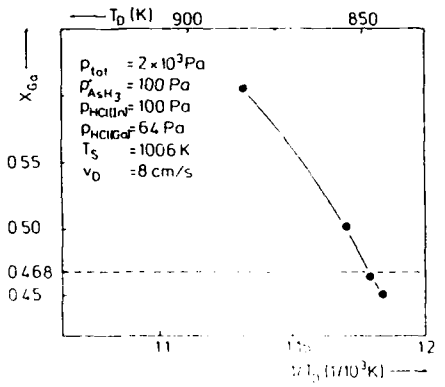


Fig. 3: Dependence of Ga content in GaInAs on deposition temperature for p = 2 x 10^3 Pa.

Besides a perfect surface morphology, layers grown under these low pressure conditions exhibit excellent electrical characteristics, as can be seen from tab.1. For example, 77 K mobilities of 130 000 cm^2/Vs and 65 000 cm^2/Vs were routinely obtained for GaAs and InP, respectively. High resolution 2K PL measurements on GaAs samples yield well resolved exciton peaks with typical linewidth of 0.3 meV. Also, the 300 PL spectrum of a GaAsP structure reveals a sharp FWHM of 15 nm. These results demonstrate the excellent optical properties of the layers.

	GaAs	InP	$In_{0.53}Ga_{0.47}As$	$GaAs_{0.6}P_{0.4}$ Te
n_{77K} (cm^{-3})	1×10^{14}	1×10^{14}	1×10^{15}	—
μ_{77K} (cm^2/Vs)	130,000	65,000	26,000	—
n_{300K} (cm^{-3})	1×10^{14}	1×10^{14}	1×10^{15}	4×10^{16}
μ_{300K} (cm^2/Vs)	7,500	4,900	9,000	2,700

Tab.1: 77 and 300 K electron concentrations and mobilities in GaAs, InP, GaInAs and GaAsP; for the latter compound only data on Te-doped material are available

Conclusions

Our study of growth in the hydride system at low overall pressures ($< 5 \times 10^3$ Pa) reveals the capability of this approach for growing binary (GaAs, InP) and ternary (GaInAs, GaAsP) III-V semiconductor films at very high rates. These rates can be conveniently adjusted over orders of magnitude by varying the reactant concentration in the gas phase. It appears likely that the same technique can also be used to grow quaternary materials. The excellent electrical and optical properties makes this growth method particularly attractive for the preparation of opto-electronic device structures. Also of interest may be the growth of high quality substrates.

References

/1/ M.J. Cardwell
J. Cryst. Growth, 70, 97 (1984)

/2/ N. Pütz, A. Sauerbrey, E. Veuhoff, R. Kirchmann, M. Heyen, H. Lüth and P. Balk

/3/ N. Pütz, E. Veuhoff, K.H. Bachem, P. Balk and H. Lüth
J. Electrochem. Soc., 128, 2202 (1981)
J. Electron. Mat., 14, 645 (1985)

/4/ H. Jürgensen, H.J. Wach, P. Balk and M. Heyen
in: Processing of Electronic materials. Proc. of the first Intern. Conf., Santa Barbara, USA. Vol. IX (American Institute of Chemical Engineering New York, 1987) p.25

/5/ J. Korec and M. Heyen
J. Cryst. Growth, 60, 286 (1982)

/6/ B.R. Singh, H. Jürgensen, M. Heyen and P. Balk
unpublished results

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988 C4-693

IMPROVEMENT OF GaAs EPITAXIAL LAYERS BY INDIUM INCORPORATION[1]

J.P. LAURENTI, P. ROENTGEN*, K. WOLTER**, K. SEIBERT**, H. KURZ** and J. CAMASSEL

Groupe d'Etude des Semiconducteurs, Université des Sciences et Techniques du Languedoc, F-34060 Montpellier Cedex, France
**IBM Research Laboratory, Säumerstr. 4, CH-8803 Rüschlikon, Switzerland*
***Institut für Halbleitertechnik, RWTH Walter-Schottky-Haus, Sommerfeldstr, Seffent-Melaten, D-5100 Aachen, F.R.G.*

RESUME - On étudie l'influence de l'incorporation d'indium (0 - $6,5 \times 10^{19}$ cm^{-3}) dans des couches OMVPE de GaAs, sur les propriétés optiques en bord de gap et les centres profonds. En photoluminescence (PL), les bandes d'excitons liés et de recombinaisons donneur-accepteur glissent vers les basses énergies: il y a formation d'un alliage ternaire GaInAs très dilué. Les intensités de PL augmentent d'un facteur 10. Cette amélioration de la qualité optique des couches est en corrélation avec une réduction de la photoémission impliquant le chrome, dans le même rapport. Par contre, la concentration et le profil de EL2 déterminés par DLTS ne sont pas affectés. Nos résultats, conjointement à ceux de la littérature, semblent indiquer une réduction par l'indium des distorsions de réseau autour des défauts, limitant ainsi la formation de dislocations.

ABSTRACT - The influence of indium incorporation (0 - $6.5 \times 10^{19} cm^{-3}$) in GaAs OMVPE layers, on near-band-edge optical properties and deep trap content has been investigated. The bound exciton and donor-acceptor recombination lines in photoluminescence (PL) spectra shift toward lower energies versus indium content. This indicates the formation of a very dilute GaInAs ternary compound. The corresponding PL intensities increase by a factor of 10. This improvement in the optical quality of the layers correlates with a decrease in a ratio 10 for chromium-related PL intensities. On the opposite, the EL2 concentration and depth profile determined by DLTS are not affected. We suggest that indium relaxes lattice distorsions arround residual defects and limits the formation of dislocations.

In bulk GaAs technology, a large reduction of dislocation density has been noticed by incorporating indium /1/. This resulted in significant improvements of the electrical substrate properties /2/ and the same technics was soon extended to GaAs epitaxial layers. Again a significant improvement in devices properties has been claimed /3/ and, to help clarifying the mechanism of indium incorporation in GaAs, we have investigated a series of samples grown by OMVPE, with calibrated indium concentrations ranging from 0 to 6.5×10^{19} cm^{-3}. This corresponds to about 0.3% in standard notations. We report in this work the results of a cross investigation of both the near band edge optical properties and the deep trap content. All experimental details have been given in separated papers /4, J.P. LAURENTI et al., unpublished/.

In Fig. 1, we show typical luminescence spectra, collected at 2K on a reference sample (indium-free GaAs layer) with different experimental arrangements. At high energy, we used an optical multichannel analyzer with 500 silicon targets and an experimental resolution of 0.1 meV. We resolved, first, a series of excitonic lines, associated with the recombinations of standard bound exciton complexes. Next come two lines, labelled L_1 and L_2, associated with the radiative decay of well identified donor-acceptor pairs. Line L_1 corresponds with both Zn and/or Mg substituted onto the Ga sites and was associated with accidental contamination processes; on the opposite L_2, which is associated with C substituting on the anion sites, was mostly dependent on the growth process; in both cases, weak LO replica have been found.

At lower energy, we have used a Ge-detector, cooled down to liquid nitrogen temperature, and an experimental resolution of typically 0.5 meV. We notice, first, a large luminescence band around 1.23 eV, which corresponds with the reported range of energy for the series of gallium vacancy-substitutional donor complexes first investigated by Williams /5/. Depending on the donor species and the lattice sites where they substitute, the top position varies. However, whatever are the donor atoms, the net concentration (N_d-N_a) must be larger than 10^{16} cm^{-3} in order to get a clear PL signal. This is, at least, one order of magnitude above the residual donor concentration achieved in our samples and makes the residual amount of electrically active impurities hardly involved in the strong luminescent features

[1]Work Partly Supported by the PROCOPE Program

experimentally found. As a matter of fact, we could establish a 1 to 1 correspondence between the magnitude of the donor-vacancy related band and the magnitude of the chromium-related zero phonon line (ZPL). This is shown in Fig. 2 where the intensity of the ($D\text{-}V_{Ga}$) complex is plotted against the intensity of the Cr-ZPL line for all samples and pieces of substrates investigated in this work. Clearly a strong correlation exists, which suggests to identify the previously unknown donor with Cr^{4+}. The sharp luminescence bands, observed at lower energy, correspond with standard features of the Cr^{2+} trigonal center. Their energy positions (ZPL at 0.839 eV) are in excellent agreement with the data of the litterature /6,7/.

Versus indium content, all high energy features shift toward lower energy, in a linear way as illustrated in Fig. 3. The slope parameters are listed in Table I. This indicates that, even at these extremely dilute indium concentrations (~ $6x10^{18}$ to $6x10^{19}$ atoms cm^{-3}), the formation of a ternary compound GaInAs happens and no effet of isoelectronic impurity can be evidenced. This is clear from the slope of the D°X feature obtained in this work (- 16.22 meV per percent). Within experimental uncertainty it is in perfect agreement with the value (- 15.84 meV per percent) reported in Ref. 8 for the free exciton transition. Slightly different values have been obtained for the acceptor-related features and come from microscopic effects /4/ associated with a change in binding energy of the bound holes. These linear relations between the energy shifts and the indium concentration make low temperature PL measurements the most precise and non-destructive tool for quantitative determinations of indium-doping levels. Finally, we display in Fig. 3 the energy position of the Cr-ZPL line versus indium content. Within experimental uncertainty, no finite shift could be evidenced. This small sensitivity to indium content depart from the results of Ref. 9 where bulk (In,Cr)-codoped LEC-grown GaAs samples were investigated. However, in this previous work, higher indium concentrations ($4x10^{18}$ to $1.2x10^{21}$ cm^{-3}) have been used and a shift, towards the low-energy side, of all chromium related features had been noticed versus indium

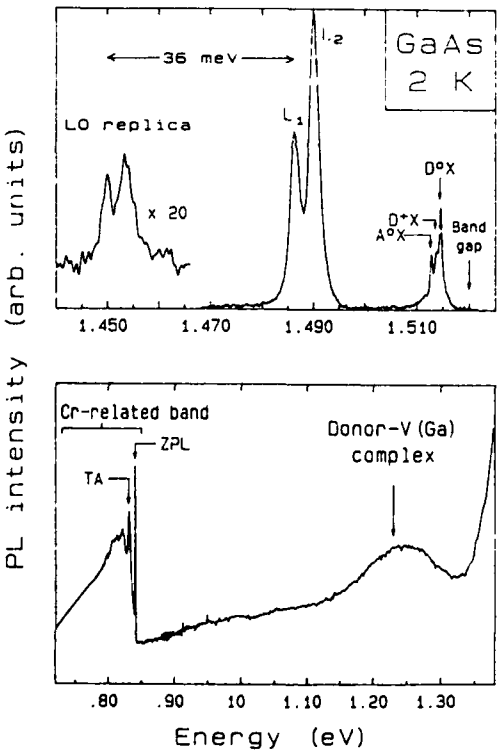


Fig. 1 - Photoluminescence (PL) spectra recorded for a GaAs OMVPE layer at 2K; (i) in the near-band-edge energy range (upper part); (ii) in the range 0.7 - 1.4 eV (lower part). All labelled features are described in the text.

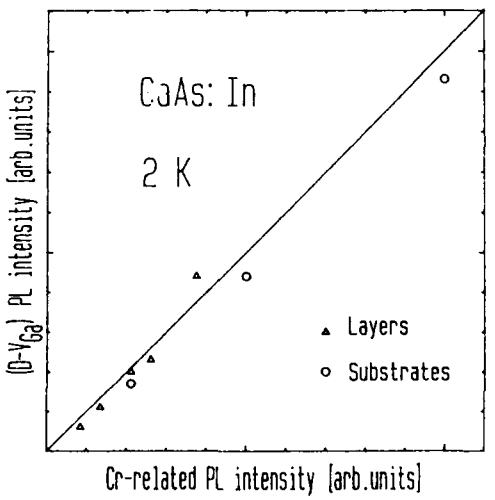


Fig. 2 - Comparison between the PL intensities of the ($D\text{-}V_{Ga}$) complex (1.23 eV), and the chromium-related zero-phonon line (0.839 eV), respectively. A 1 to 1 correspondance between these two PL features is evidenced for all GaAs:In layers and GaAs corresponding substrates.

concentration. Owing to the internal character of the transition, and based on general crystal-field theory, the change in magnitude of the crystal-field splitting was estimated from an inverse fifth power of the lattice constant of the host material /10/:

$$\Delta E \sim E_o \left[\left(a_{GaAs}/a_{GaAs:In}\right)^5 - 1 \right] \qquad (1)$$

With usual values of the crystal parameters /11/, the computed shift is far above our experimental uncertainty. We have found a better agreement using only the Ga-As nearest neighbour distancy in $Ga_{1-x}In_xAs$ /11/. This is shown in Fig. 3.

Concerning now the absolute luminescence intensities, we find the following : (i) all near-band-edge PL intensities increase by a factor of 10 when incorporating indium (Fig. 4); (ii) all defect (chromium) -related transitions decrease by a factor of 10 when incorporating indium (Fig. 5); (iii) there is a one to one correspondance between the changes in the near band edge (NBE) and the chromium related transitions (Fig. 6). On the opposite, the EL2 concentration and depth profile determined by capacitance spectroscopy (DLTS) are not affected.

We believe that most of the improvements concerning bulk and epitaxial materials rely on similar features: there is an interplay of indium and chromium incorporation. Both our results and the work of Ref. 9 show that the concentration of trigonal Cr^{2+} centers decreases; in addition, we show that the concentration of $(Cr^{4+}\text{-}V_{Ga})$ complexes also decreases. The simplest explanation would be that the overall concentration of chromium in the material decreases. However, there are also evidences for a complexation of chromium with indium /9/. This suggests that indium fills in natural cluster surrounding the residual impurities and/or structural defects, and relaxes the lattice distorsions. Lowering the crystal total energy, this would limit the amount of dislocations which can be formed.

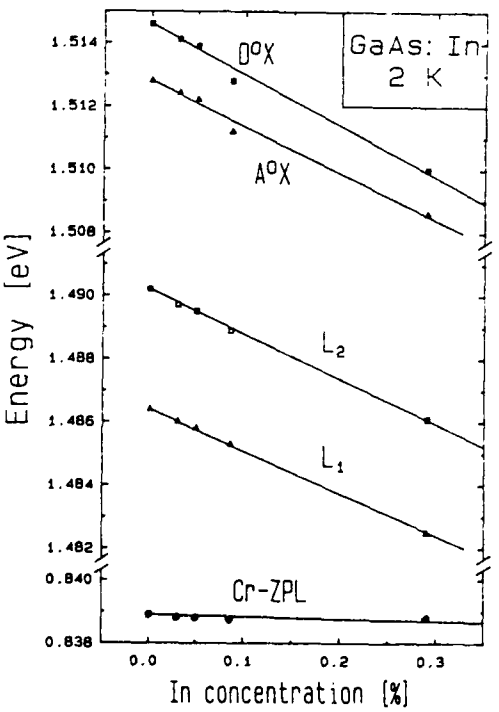


Fig. 3 - Energy positions of the near-gap and Cr-related PL lines, plotted as a function of indium concentration. For the near-gap transitions, the staight lines are drawn with the slope parameters listed in Table I and obtained from a least-mean-square fit. For the Cr-related transition, the line is drawn from Eq. (1) (see text).

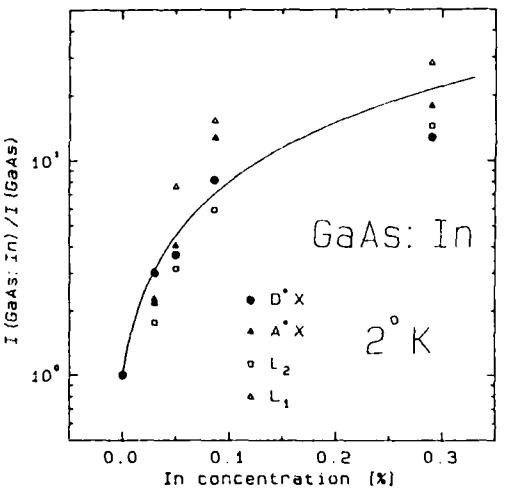


Fig. 4 - Changes in PL intensities as a function of indium concentration for the near-band-edge transitions. All values have been normalized to the In-free reference layer.

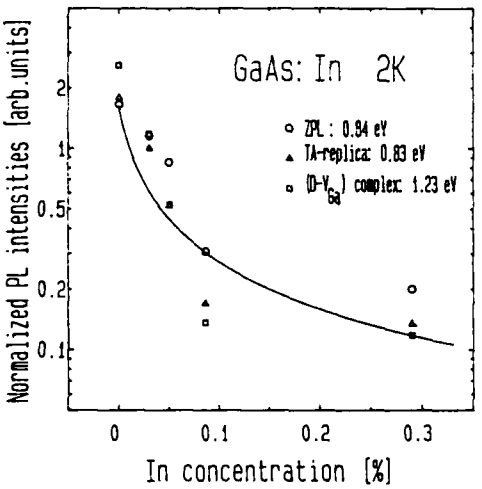


Fig. 5 - Changes in PL intensities as a function of indium concentration for the defect (chromium) -related transitions. For every layer and every transition, all values have been normalized to the PL intensities from the corresponding piece of substrate.

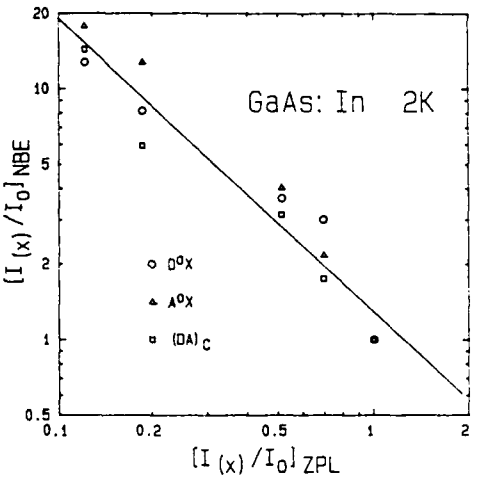


Fig. 6 - Relative intensities of the near-band-edge transitions versus Cr-related zero-phonon line (ZPL). The straight line is a least-mean-square fit on the log-log scale.

Table I : Slope parameters (meV/%) obtained at 2K in this work for the near band edge optical features and comparison with (i) the change in band gap obtained in Ref. 8, and (ii) the initial slope parameters obtained from Eq.(1) for Cr-related ZPL energy (see text).

Transitions	E_g	D°X	A°X	L_2	L_1	Cr^{2+}(ZPL)
This work		- 16.22	- 14.73	- 14.235	- 13.36	- 0.69
Ref.8	- 15.84					
Refs.10,11						- 2.96

REFERENCES

/1/ G. JACOB, M. DUSEAUX, J.P. FARGES, M.M.B. Van Den BOOM and P.J. ROKSNOER, J. Crystal Growth, 61, 417 (1983).
/2/ M. DUSEAUX and S. MARTIN in "Semi-insulating III-V materials" (Shiva, Kahnee-ta, 1984) p 118.
/3/ H. SCHUMACHER, P. NAROZNY, CH. WERRES and H. BENEKING, I.E.E.E. Electronic Devices Letters, 7, 26 (1986).
/4/ J.P. LAURENTI, P. ROENTGEN, K. WOLTER, K. SEIBERT, H. KURZ and J. CAMASSEL, Phys. Rev. B37, 4155 (1988).
/5/ E.W. WILLIAMS, Phys. Rev. 168, 922 (1968).
/6/ Y. FUJIWARA, T.NISHINO and Y. HAMAKAWA, Jap. J. Appl. Phys. 21, L727 (1982).
/7/ J. BARRAU, Do Xuan THANN, M. BROUSSEAU, J.C. BRABANT and F. VOILLOT, Solid State Commun. 44, 395 (1982).
/8/ K.H. GOETZ, D. BIMBERG, H. JÜRGENSEN, J. SELDERS, A.V. SOLOMOV, G.F. GLINSKII and M. RAZEGHI, J. Appl. Phys. 54, 4543 (1983).
/9/ Y. FUJIWARA, Y. KITA, Y. TONAMI, T. NISHIMO and Y. HAMAKAWA, Appl. Phys. Lett. 49, 161 (1986).
/10/ S. SUGAMO, Y. TAMABE and H. KAMIMURA, "Multiplets of transition-metal ions in crystals" Academic Press, N.Y. (1970) p14.
/11/ J.C. MIKKELSEN, Jr and J.B. BOYCE, Phys. Rev. B28, 7130 (1983).

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

ELIMINATION OF TWINNING IN MOLECULAR BEAM EPITAXY OF GaAs/Si and GaAs/INSULATOR

C. FONTAINE, J. CASTAGNE, E. BEDEL and A. MUNOZ-YAGUE

Laboratoire d'Automatique et d'Analyse des Systèmes du CNRS, 7, Av. du Colonel Roche, F-31077 Toulouse Cedex, France

Résumé - Nous présentons une étude de la nucléation de GaAs sur des surfaces (100) de Si et de(Ca,Sr)F2 montrant qu'une croissance 2D peut être obtenue pour des dépôts très fins (15 A°) réalisés à basse température (25-100°C). La stoechiométrie du dépôt est obtenue en augmentant la température jusqu'à 300-400°C ce qui conduit, en même temps, à la cristallisation de la couche. Le diagramme de diffraction électronique met en évidence l'absence de micromacles et l'obtention d'une surface plane au niveau atomique. La procédure proposée ouvre la voie à l'utilisation de techniques basse température pour l'élaboration de structures co-intégrées associant GaAs et Si.

Abstract - We present a study of GaAs nucleation on Si and $(Ca,Sr)F_2$ (100) surfaces which shows that a 2D growth can be achieved for very thin (15 A°) deposits obtained at low temperature (25-100°C). Raising the temperature to 300-400°C leads to a stoichiometric and monocrystalline layer. Electron diffraction shows a diagram without twin spots and a surface devoid of noticeable roughness. The proposed procedure paves the way for the subsequent use of low temperature techniques aimed at the fabrication of co-integrated structures associating GaAs and Si.

1 - INTRODUCTION

Gallium Arsenide epitaxy on Silicon substrates has reveived considerable attention over the last years, and the quality of the layers has been dramatically improved. Although successful demonstration devices have been reported the properties of the GaAs heteroepitaxial layers have still to be improved. This implies the reduction or elimination of the crystalline defects -twins and dislocations- present in these layers, specially near the layer-substrate interface. Even if post-growth annealing (in or ex situ) can be used to overcome this problem, specific growth procedures preventing defect formation during the growth should be preferred.

In addition, low temperature processes have to be developed to reduce the interdiffusion problems at the GaAs-Si interface (namely Si diffusion into GaAs) and to permit fabrication of co-integrated structures.

Among the mechanisms giving rise to the aforementioned crystal defects, the three-dimensional mode (3D-growth) of the heteroepitaxial GaAs nucleation is often invoked. Here, a growth technique is proposed which favors layer-by-layer nucleation (2D-growth) of the GaAs heteroepitaxial layer.

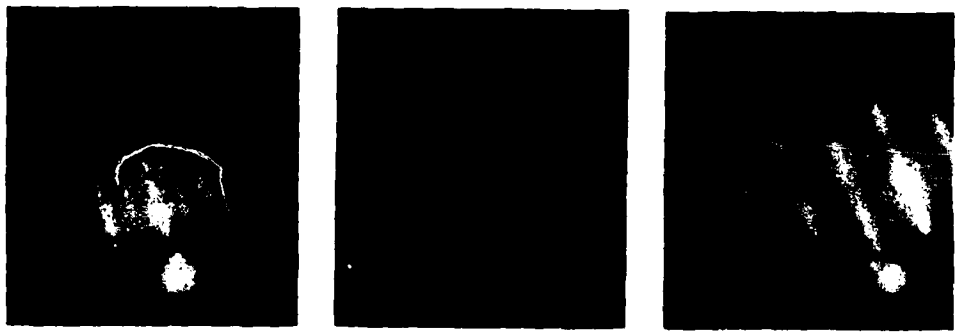
As the technique used for the Si substrate preparation requires heating of the substrate to only 600°C for 10 minutes and as the proposed nucleation procedure works at 400°C, both steps offer promising prospects for low temperature processing in GaAs/Si molecular beam epitaxy (MBE). Finally, the results presented show that extended defects like twins are eliminated in the early stage of growth.

2 - EXPERIMENTAL

An ultra-high vacuum MBE system provided with two main chambers, an Auger electron spectrometer (AES) and an electron diffractometer (RHEED), was used in this study.

The Si substrates were (100), 2° misoriented towards [011]. The preparation of the substrates has been described elsewhere [1] ; it is briefly recalled here as part of the low temperature process : the substrates, with a protective layer of thermally grown SiO_2, were first etched using an HF : ethanol solution to remove this oxide just before loading into the vacuum chamber. AES reveals that, at this point, the surface is devoid of any traces of O or C contamination. RHEED shows a streaky pattern. In the second step, this surface is exposed to low Si atomic flux provided by the effusion cell used for n-type doping of GaAs. This flux is not sufficient to grow a Si buffer layer, but allows the

reconstruction of the Si surface ; indeed, after 10 minutes at 600°C, the RHEED diagram consists of sharp and intense lines with a (2 x 2) reconstruction (figure 1a) that is interpreted as a mixture of (100) terraces with (1 x 2) and (2 x 1) reconstructions. It has been shown that such a Si surface, even with monoatomic steps, allows GaAs to be grown without antiphase defects [2].



a) b) c)

Fig. 1 - *RHEED diagram obtained for the [011] azimuth on a) the cleaned Si surface, b) on the 15A° thick amorphous layer grown at low temperature and c) after annealing of this layer at 400°C.*

3 - RESULTS AND DISCUSSION

The initial step of GaAs/Si MBE epitaxy generally follows the procedure proposed by Wang [3], which consists in the growth of a 200 A° thick layer at moderate temperature (300-400°C) and a growth rate (0.1µm/h). When this procedure is applied, a 3-D nucleation is observed using AES or RHEED : GaAs islands nucleate and grow to finally coalesce for a deposited mean thickness of about 200 A°. The RHEED diagrams are spotty, denoting surface roughness, and frequently reveal twin spots which are eliminated later on, when the growth is continued at higher temperatures (580°C).

To avoid defect creation at the coalescence stage, a 2D nucleation should be preferred. In addition, it would produce a much smoother surface. A 3D nucleation mechanism can be deviated towards a 2D process by lowering the growth temperature : the reduction of the diffusion length of the impinging atoms favors this type of nucleation. The problem, which then arises for GaAs, is that the deposit is no longer monocrystalline and stoichiometric. This is indeed observed experimentally. Figure 1b shows that when a 15 A° thick layer of GaAs is grown on a Si surface at low temperature (between 25 and 100°C) the deposit is amorphous, the RHEED diagram showing a diffuse halo. Auger analysis shows an excess As concentration in the layer due to the increase of the sticking coefficient of the As_4 molecules evaporated from the effusion cell.

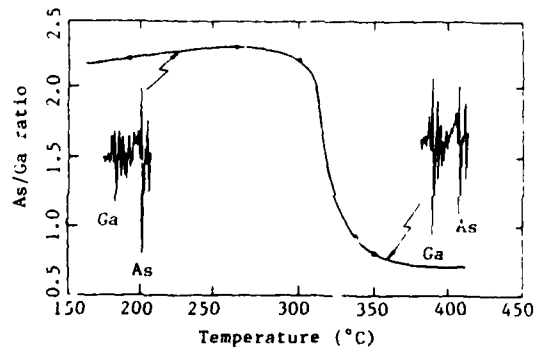


Fig. 2 - Variation of the Auger signal ratio, As (1210 eV)/Ga (1065 eV), for a deposit grown at low temperature when the temperature is raised, evidencing the restoration of the layer stoichiometry during the thermal process.

This can be rapidly improved by a moderate thermal treatment : between 300° and 400°C, the stoichiometry is restored by evaporation of the excess As (figure 2), and simultaneously, a monocrystalline layer is obtained : The RHEED diagram is a streaky pattern, with an incipient reconstruction devoid of twin spots (figure 1c). This result, which has been obtained systematically, demonstrates that the Si substrate is covered up by a very thin GaAs layer with a flat single oriented surface. Usually, such a RHEED diagram is obtained, using standard procedures, after the growth of much thicker (2000 A°) layers.

The proposed procedure also applies to the growth of GaAs/$(Ca,Sr)F_2$/Si structures [4]. The fluoride layer can be used to reduce the lattice mismatch between Si and GaAs, also offering an insulating layer between the semiconductors that could be useful in some applications. Here also, after heating at 400°C the 15 A° thick deposit of GaAs on a $(Ca,Sr)F_2$ layer (10^3 A°) grown on a Si substrate, the resulting layer presents a good crystallinity, without surface roughness (figure 3).

Fig. 3 - <011> RHEED diagram obtained on a 15 A° thick amorphous GaAs deposit on $(Ca,Sr)F_2$/Si (100) after its annealing at 400°C.

The lack of twins in thicker layers grown using this nucleation procedure has also been verified using Raman Spectrometry. It has been shown [5] that, under adequate experimental conditions, the twinned volume in a GaAs (100) layer can be measured from the LO (allowed)/TO (forbidden) phonon ratio. As shown in figure 4 the Raman spectra obtained under those conditions on a GaAs/Si layer exhibit no TO phonon peak, confirming thus that no misoriented twinned regions exist in the GaAs layer.

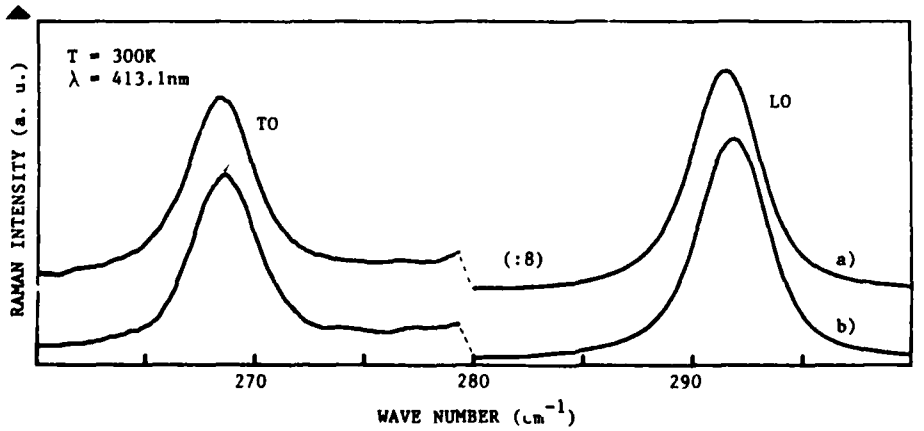


Fig. 4 - Raman spectra obtained on a) a 1 μm thick GaAs layer grown on Si (100) with the proposed procedure and b) a reference homoepitaxial GaAs layer : the peak shifts evidence the presence of thermal stress in the heteroepitaxial layer. The TO/LO amplitude ratio is the same for both samples.

4 - CONCLUSION

We have presented procedures for Si surface preparation and for GaAs heteroepitaxial nucleation using MBE which lead to a good crystalline material and improved surface morphology.

The advantages of employing an initial deposit (15 A° thick) at room temperature, stoichiometry and crystallinity of which are restored by a moderate thermal treatment (400°C), have been demonstrated for GaAs growth on Si and on $(Ca,Sr)F_2$. The proposed techniques, coupled with other means of reducing both the dislocation density and the growth temperature, could be useful in the future growth of co-integrated GaAs-Si structures.

ACKNOWLEDGEMENTS

R. CARLES and G. LANDA are acknowledged for their kind collaboration on Raman spectroscopy.

REFERENCES

/1/ Castagné, J., Fontaine, C., Bedel, E. and Munoz-Yague, A., J. Appl. Phys. to be published in June 1988.
/2/ Kawanami, H., Hatayama, A., Nagai, K. and Hayashi, Y., Jpn. J. Appl. Phys. 26 (1987) L173.
/3/ Wang, W.I., Appl. Phys. Lett. 44 (1984).
/4/ Munoz-Yague, A. and Fontaine, C., Advances in semiconductors and superconductors : Physics and device Applications, S.P.I.E., New-Port Beach, USA, (March 1988).
/5/ Landa, G., Carles, R., Renucci, J.B., Fontaine, C., Bedel, E. and Munoz-Yague, A., J.Appl. Phys. 60 (1986) 1025.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A GaAs ON Si COPLANAR TECHNOLOGY BY EMBEDDED MOLECULAR BEAM EPITAXY

J.B. LIANG, J. DE BOECK, R. MERTENS and G. BORGHS

Interuniversity Micro-Electronics Center vzw, Kapeldreef 75, B-3030 Leuven, Belgium

In recent years, there has been growing interest in the heteroepitaxy of GaAs on Si(1). The motivation for this is the monolithic GaAs/Si (MGS) technology which combines the optoelectronic and high speed properties of GaAs with the high density of Si circuitry. Some succesfull realizations have already been reported on plana: Si wafers. (2-6). When the active regions for the GaAs devices are defined by etching mesa's after growth, exposing the processed Si devices, significant variations in surface topography of several microns are generated. A monolithical GaAs-Si semiconductor combination with a coplanar surface is desirable from a process point of view. In the latter case one can take maximum advantage from the well established high resolution silicon processing. One way to achieve such a coplanar surface is to recess holes several microns deep into the Si substrate and refill them with an epitaxial layer, such that the GaAs surface is even with the Si surface after the GaAs growth. Smooth, device quality GaAs is required in the wells, and in order to obtain this high quality one must avoid the introduction of surface damage or the loss of the intentional misorientation by recessing the Si. Otherwise, the degradation of the morphology and the structural properties of the layer will be extensive and antiphase domains are no longer supressed (7,8). Etching exposes different crystallographic orientations at the sidewalls thus allowing for different nucleation behaviour. Furthermore, the GaAs film is submitted to more severe geometrical constraints causing a different thermal strain field during cool down. The influence of these factors on the formation of dislocations and stacking faults and their propagation through the layer have to be considered. In this paper we present a new technique for the realization of a coplanar GaAs on Si surface, involving a wet chemical etch, a SiO_2 overhang reduction step, embedded MBE growth and removal of the polycrystalline GaAs. Successful metallization crossing the GaAs to Si border without intermediate planarization, after embedded MBE in wet etched recesses in the Si substrate, is reported. Since we have demonstrated recently (9) that the structural properties of the embedded MBE layers are comparable to those of large area GaAs on Si, MGS device applications are realisable using the presented technique.

Light p-doped 10 ohm-cm 3 inch (001) Si wafers oriented 4° off towards [110] were used. The preparation for growth involves forming and patterning a 400nm thick SiO_2 layer and exposing the Si where it is to be etched. The mask consists of $4x4mm^2$ squares containing micro-windows with different widths (20-150μm) and spacings (5-15μm).

Different solutions of HNO_3 and HF diluted in H_2O or CH_3COOH were tested to form a featureless flat bottom and to preserve the intentional misorientation. An isotropic solution of HF:HNO_3 (1:19) was found to meet the above requirements, as discussed below. The reaction can be represented as follows;

an oxidizing step: $Si + 4HNO_3 = SiO_2 + 4NO_2 + 2H_2O$

and an oxide decomposition step: $SiO_2 + 6HF = H_2SiF_6 + 2H_2O$

The etching process tends to be diffusion limited and is carried out at controllable rate. Studies of the recess profile, using a Sloan Dektak® and a JEOL 840 scanning electron microscope, have demonstrated the surface flatness. Figure 1 shows a typical well profile, as measured with the Dektak®, obtained after the isotropic wet etch with the above mentioned solution. Although faster etching occurs near the Si sidewall, the maximum effective angle is less than 0.5°, as demonstrated by the bottom part of Fig.1, so that the intentional misorientation is preserved sufficiently to retain the beneficial influence of the tilt on the heteroepitaxial growth ([7,8]).

After etching the wells, a SiO_2 overhang appears because of the effect of undercutting. MBE growth on such a structure generates voids between the epitaxial GaAS layer and the Si sidewall since the molecular beams are screened from the corner part of the well. In order to allow for reliable metallization in the frame of MGS integration the overhang must be removed. Thus the wafers were etched in buffered HF for more than half the time needed to remove the complete SiO_2 layer. By doing so no voids will occur during growth.

The dielectric is partly stripped from the wafer to permit monitoring the changes in the RHEED pattern during the growth process. Prior to loading, a passivation treatment in $HCL:H_2O_2:H_2O$ (3:1:1:) is used, similar to the last step in the report by Ishizaka et al.([10]), to form a very thin oxide film that is readily removed upon heating in the evaporation chamber. The Si substrate was then loaded into a Riber 2300 MBE system using an indium free mount and was brought to 300°C for 30' to remove moisture and gases. In the evaporation chamber the substrate was heated to 850°C for 15' to decompose the thin oxide layer on the surface. The RHEED showed a very sharp (2x1) reconstruction of the Si surface indicating an atomically clean surface. Thermally strained layers are used, similar to that reported by J.W.Lee ([11]). The substrate temperature is cycled between 650°C and 350°C and subsequently the GaAs film was grown on top of the nucleation layer at a conventional growth temperature of 570°-600°C.

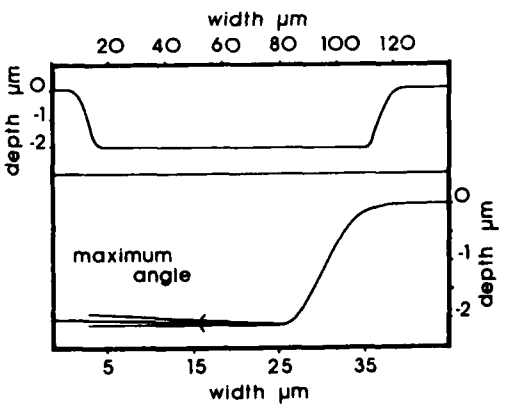


Fig. 1. The profile of a Si recess formed by HNO_3:HF (19:1) solution. The maximum effective deviation angle is about 0.5°.

Fig. 2. A SEM cross section of an embedded GaAs layer. The same nucleation and growth behaviour is found on both the unetched (on the platform) and the recessed Si surfaces.

The morphology of epitaxial GaAs film on Si in the windows is mirrorlike with a typical 'orange peal'-like microtexture. Figure 2 shows a SEM cross section of an embedded GaAs film. The picture is taken from the part of the wafer where the oxide is completely stripped before loading. It shows the smoothness of the recess floor and the typical profile of an isotropically etched surface. In this particular case, an AlGaAs/GaAs periodic heterolayer refills the well. As appreciated from the micrograph, the same nucleation and growth properties are found on both the untreated and the recessed Si. The rough initial GaAs nucleation layer is smoothened out after 0.3μm on both surfaces.

The optical and structural properties of the embedded GaAs layers have been characterized and compared with large area GaAs on Si as discussed in detail elsewhere([9]). All results indicate that the quality of the recessed GaAs on Si is comparable with the large area deposit on the unetched part of the Si substrate.

The GaAs devices are to be fabricated in the embedded single crystal layers in the recess. The realization of the coplanar GaAs on Si surface thus requires precise control of growth rate in order to equalize the thickness of the deposit and the depth of the well. After growth a polycrystalline deposit on the masking dielectric surrounds the recess. A photoresist mask is used to protect the monocrystalline GaAs during the wet etching process to lift off the polycrystalline GaAs. Figure 3 shows a cross sectional SEM photo, taken after the lift off procedure, demonstrating how the GaAs and Si surfaces level. A GaAs surface step of about 0.6μm for a 2μm thick layer appears on the GaAs to Si border. The shape and heigth of the hillock depend strongly on the geometry of the sidewall.

A metallization test, using Au-evaporation, was performed over the border to test the step coverage of the metal line over the GaAs/Si border. The evaporation technique was chosen because it is more sensitive to shadow effects then sputtered deposition and thus constitutes a more critical test for the coplanarity. The test pattern for the metal lift off consisted of lines with varying width and spacing from 10μm to 0.5μm. Figure 4 shows a plane view SEM micrograph of 1.25μm wide Au lines with a thickness 150nm. The wider lines were checked electrically and were found not to be interrupted. SEM-inspection showed that the metallization was succesful down to 1.25μm width

Fig. 3. An SEM photo showing the cross section of the embedded GaAs in the wet etched Si recess after lift off.

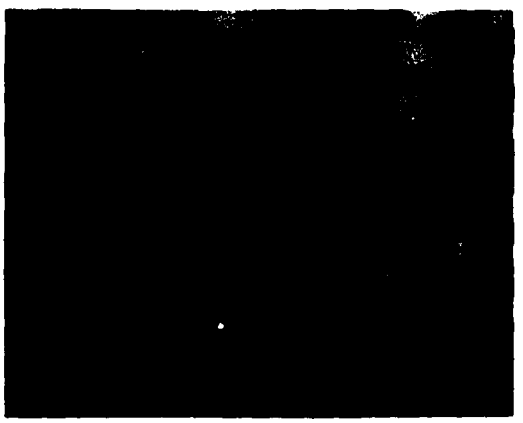


Fig. 4. An SEM photo image showing 150 nm thick with 1.25 μm wide Au-lines crossing the GaAs to Si border.

and spacing, although reliable metal lines should be made wider.

In summary, we reported on GaAs grown embedded in masked wells, which were etched in the Si substrate with a controlled solution of HNO_3 and HF. Coplanarity of the GaAs epilayer and the Si surface is obtained after lifting off the polycristalline GaAs with the masking SiO_2. Successful metallization is performed over the GaAs to Si border. Since the GaAs layer quality is identical to 'normal' GaAs on Si, the presented technique is suitable for MGS applications.

The authors wish to thank D.J. Arent, K. Deneffe, C. Van Hoof for the many fruitful discussions. Further the contribution of P. Roussel, J. Vanhellemont, W. De Raedt, M. Van Hove, M. Lambrechts and W. van De Graaf are kindly acknowledged.

References.

1. "Heteroepitaxy on Si I and II", in *MRS Symposia Proceedings No.67*, edited by J.C.C. Fan and J.M. Poate (1986); and *No.91*, edited by J.C.C. Fan, J.M. Phillips, and B.-Y. Tsaur (1987)(Materials Research Society, Pittsburg, PA)
2. R.J. Fischer, N. Chand, W. Kopp, C.K. Peng, H. Morkoc, K.R. Gleason and D. Scheitlin, IEEE Trans. Electron Devices, ED-33, p. 206, 1986.
3. L.T. Tran, R.J. Matyi, H.T. Yuan, and J.W. Lee, IEEE Electron Dev. Lett. EDL-8, p.50, 1987.
4. H. Shichijo, J.W. Lee, W.V. McLevige, and A.H. Taddiken, IEEE Electron Dev. Lett. EDL-8, p.121, 1987.
5. H.K. Choi, G.W. Turner, and B.-Y. Tsaur, IEEE Electron Dev. Lett. EDL-7, p.241, 1986.
6. H.K. Choi, G.W. Turner, T.H. Windhorn, and B.-Y. Tsaur, IEEE Electron Dev. Lett. EDL-7, p.500, 1986.
7. R. Fischer, D. Neumann, H. Zabel, H. Morkoc, C. Choi, and N. Otsuka, Appl. Phys. Lett. 48, p.1223, 1986.
8. J. S. Harris,Jr , S.M. Koch, S.J. Rosner, in "Heteroepitaxy in Si II", in *MRS Symposia Proceedings No.91*, edited by J.C.C. Fan, J.M. Phillips, and B.-Y. Tsaur (Materials Research Society, Pittsburg, PA,1987)
9. J. De Boeck, J.B. Liang, K. Deneffe, J. Vanhellemont, D.J. Arent, C. Van Hoof, R. Mertens, and G. Borghs, to be published in Appl. Phys. Lett.
10. A. Ishizaka, and Y. Shiraki, J. Electrochem. Soc. 133, p.666, 1986.
11. J.W. Lee, Proceed. 1986 Int. Symp. on GaAs and Related Compounds, *Inst. Phys. Conf. Ser. 83*, p. 111, 1987.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

INVESTIGATION OF THE INFLUENCE OF DX CENTERS ON HEMT OPERATION AT ROOM TEMPERATURE

P. GODTS, E. CONSTANT, J. ZIMMERMANN and D. DEPREEUW

Centre Hyperfréquences et Semiconducteurs, CNRS-UA 287, Université des Sciences et Techniques de Lille Flandres Artois, F-59655 Villeneuve-d'Ascq Cedex, France

RESUME - Une étude expérimentale et théorique des Centres DX dans les couches en AlGaAs est présentée. Il est montré qu'à cause des centres DX, à température ambiante, la transconductance hyperfréquence des HEMTs peut être cosidérablement plus élevée que celle obtenue en régime continu et approche les valeurs qui pourraient être obtenues si les centres DX n'existaient pas. En conséquence, pour des applications à température ambiante, des structures particulières destinées à éliminer les centres DX ne sont pas réellement nécessaires.

ABSTRACT - A study of DX centers in GaAlAs layers is presented, both experimental and theoretical. It is shown that due to DX centers, at room temperature, the microwave transconductance of HEMTs can be considerably higher than that obtained from the DC regime, and approaches the values which could be expected if DX centers did not exist. As a result, for room temperature applications, special structures designed in order to eliminate DX centers are not really necessary.

It is now well known that, due to deep levels associated with DX centers in doped $Al_x Ga_{1-x} As$ [1], the electrical properties of HEMTs may be strongly modified leading to low temperature collapse [2], persistent photoconductivity [3]. Solutions to this have been proposed : a reduction of Al mole fraction x below 0.2; the use of a n^+ GaAs-AlAs superlattice [4]; or even a planar (delta)-doped layer [5]. It is clear that these solutions are necessary for low temperature operation, but are they really useful for room temperature operation? Here,we study the effect of DX centers on DC and microwave properties of HEMTs at room temperature.We use a HEMT simulator which is able to take DX centers into account [6] and we will support our findings by a comparison with a few experimental data.

Our theoretical study is carried out using a method which is somewhat similar to [6]. Since we are interested primarily in the influence of DX centers upon the drain current and device performance, we have to evaluate the importance of electron transfer between the GaAs and GaAlAs layers, requiring a correct evaluation of the gate current through the GaAlAs layer. To do this, we start with some 2 DEG density n_s at some point of

the source-drain channel, from which the exact position of the Fermi level $E_F(n_s)$ with respect to the conduction band is known.

This $E_F(n_s)$ was established from exact results computed with a self-consistent solution of Schrödringer's equation [6]. Using this as the boundary condition for the solution of drift-diffusion and Poisson equations, the Fermi level through the AlGaAs layer is calculated and several assumptions for the effect of DX centers on the electron concentration can be compared. Then, at a given gate voltage, the gate current density can be computed once the correct boundary condition is met at the gate plane. This method is somewhat similar to the one presented by Ponse et al [7]. Thus, we get the n_s and the total charge in the AlGaAs layer as a function of gate voltage. Finally, the drain current and gate-source capacitance are computed using a short channel approximation [8] by solving the following equations:

$$I_{DS} = Z\,(Q_s \cdot v_{GaAs} + Q_{free} \cdot v_{GaAlAs})$$

$$Cgs = Z.L_g \cdot d(Q_s + Q_{free} + Q_{DX}) \,/\, dVgs$$

where $Q_s = q.n_s$, Q_{free} is the free charge in the GaAlAs layer, and Q_{DX} is the charge trapped on DX centers, v_{GaAs} (v_{GaAlAs}) is the peak drift velocity in GaAs (GaAlAs), Z and L_g are respectively the gate length and the gate width.

The key-point here is in the way in which deep DX centers can be taken into account. We used some recent results reported by Theis [9] and made the following assumptions : i) every impurity atom gives rise to a deep donor level (presumably due to a DX center). ii) The deep donor level approximately follows the L-valley minimum as this latter changes with x and is taken to be 140 meV below the L-valley. iii) Shallow donor levels are not taken into account because at high doping concentrations used in HEMTs, the shallow impurity band is resonant with the Γ-valley. As a consequence, each impurity atom gives one electron which either can occupy a conduction band state or can be captured by a(positively) ionized empty deep donor level and the concentrations of these two types of occupied states can be determined since the quasi Fermi level is known.

In order to assess the role of DX centers, the main characteristics (g_m, Cgs) of a typical HEMT structure have been computed both when DX centers, as described above, are taken into account or when they are ignored. This is illustrated in figure 1. DX centers have a strong influence on the transconductance at positive gate voltages . In DC regime a good part (depending on the Fermi level values) of electrons are trapped by DX centers and this limits the upper values of the static drain current and consequently of g_m. When DX centers are ignored, the g_m can be substantially enhanced. Correspondingly , one finds that Cgs is also very sensitive to the effect of DX centers. Static regime can be characterized by times much longer than the emission or capture times on DX centers. Preliminary experiments performed in our laboratory showed that, at 300K, the capture time is of a few μs, although emission time is about 100μs (see figure 2). Consequently, at 500MHz and above,it can be assumed in our model that the charge trapped in DX centers depend on the static bias voltage VgsO and not on the microwave voltage ΔVgs. Taking into account these assumptions, the Q_s, Q_{free}, Q_{DX} values can be determined for Vgs = VgsO and then assuming no variation of the charge trapped in DX centers for Vgs = VgsO + ΔVgs (typical value

used for ΔVgs is 0.1 volt). By using such a method, the microwave g_m including DX centers can be computed and it is found (figure 1) to be closer to the static g_m when DX centers are neglected. Similar results are found for the gate capacitance.

In the rest of the paper we check this idea against experiments. We used two different methods in order to measure the static and microwave g_m of HEMTs with a structure very similar to the one shown in figure 1. In the first method, a pulsed bias voltage is applied to the gate and the corresponding variation of the drain current is recorded (see figure 2 for an example). One then notes that the drain current observed just after the pulse has been applied, can be much higher than the drain current at longer times. This leads, for large positive value of the gate voltage, to microwave g_m 30% higher than the static g_m. As a counter experiment we have checked that this behaviour is never observed in normal GaAs MESFETs where the effect of DX centers is negligable. A second method was also used to confirm these results : on one side the static g_m is measured in the DC regime as a function of Vgs and on the other side, the microwave g_m is obtained from S parameters measurements [10]. The results obtained in this way are presented in figure 3. The fact that at Vgs > 0 the microwave g_m is substantially higher than the static g_m is clearly seen. Moreover the results shown here are in quite good agreement with theoretical predictions of figure 1.

In conclusion, we have shown that, although the influence of the deep level due to DX centers in AlGaAs layers may be strong as concerns static characteristics, the influence on microwave performance is clearly much weaker. Consequently, at room temperature, special structures designed in order to eliminate DX centers formation in HEMTs appear not really necessary.

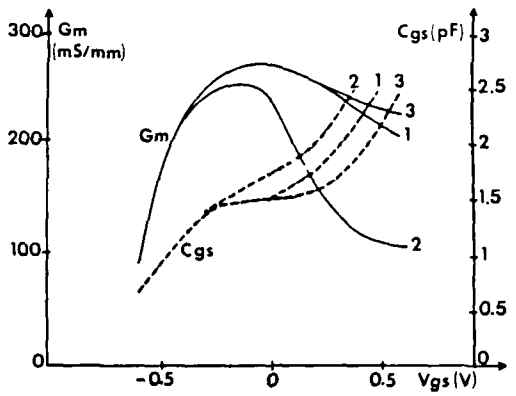


Figure 1 : This figure compares the computed results obtained (1) when DX centers are ignored; (2) when they are taken into account (see text) in DC regime; (3)when they are taken into account in the transient regime which characterizes DX centers behaviour at microwave frequency.

$N_D = 7.10^{17} cm^{-3}$	500Å	X = 0.25	$Al_xGa_{1-x}As$
	10Å	X = 0.25	
	5000Å		GaAs

Simulated structure

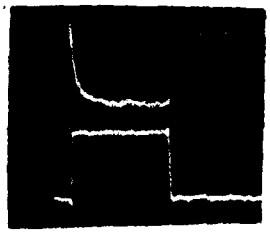


figure 2 :Drain current time response to a pulsed gate voltage. Shorter capture time by DX centers is observed at the beginning of the pulse and longer emission time is observed after extinction of the pulse.

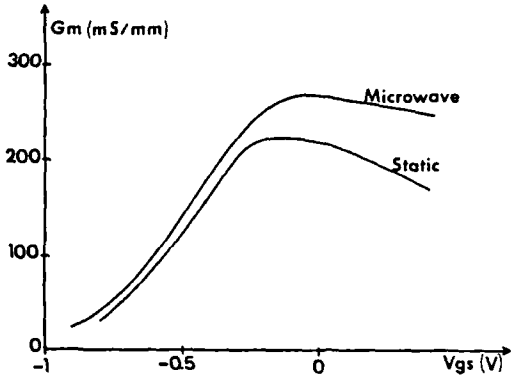


Figure 3 :Experimental transconductance measured on a structure very similar to the one shown in fig.1. Static and microwave (500MHz) data are shown.

REFERENCES

[1] D.V.Lang, R.A.Logan and M.Jaros, Phys.Rev B., 19, 1015 (1979).

[2] J.F.Rochette, P.Delescluse, M.Laviron, D.Delagebeaudeuf, J.Chevrier and N.T.Linh, Inst.Physics Conf.Series., 65, 385 (1982).

[3] M.I.Nathan, Solid State Electron., 29, 167 (1986).

[4] W.Prost, W.Brockerhoff, K.Heime, W.Schlapp and G.Weimann, International Symp. on GaAs and Related Compounds, Heraklion (Crête), sept 1987.(Int. Physics Conf.Series., to appear).

[5] E.F.Schubert, A.Fischer and K.Ploog, IEEE Trans. Electron Dev., ED-33, 625 (1986).

[6] E.Constant, P.Godts, D.Depreeuw and J.Zimmermann, International Symp. on GaAs and Related Compounds, Heraklion (Crête), sept.1987 (Inst.Physics Conf.).

[7] F.Ponse, W.T.Masselink and H.Morkoç, IEEE Trans. Electron Dev., ED-32, 1017 (1985).

[8] A.Cappy, A.Vanoverschelde, M.Schortgen, C.Versnaeyen and G.Salmer, IEEE Trans.Electron.Dev., ED-32, 2787 (1985).

[9] T.N.Theis, International Symp. on GaAs and Related Compounds, Heraklion (Crête), sept. 1987 (Inst.Physics Conf.Series, to appear).

[10] G.Dambrine, A.Cappy, F.Heliodore, E.Playez, IEEE Trans. Microwave Theory and Techniques, to appear in July 1988.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

ON THE MICROWAVE LOW TEMPERATURE ANALYSIS BEHAVIOUR OF HEMTs

A. BELACHE, A. VANOVERSCHELDE, G. DAMBRINE and M. WOLNY*

Centre Hyperfréquences et Semiconducteurs[1], CNRS-UA 287, Université des Sciences et Techniques de Lille-Flandres-Artois, F-59655 Villeneuve-d'Ascq Cedex, France
**Laboratoires d'Electronique et de Physique Appliquée, 3, Av. Descartes, F-94451 Limeil-Brevannes Cedex, France*

Résumé - Le comportement du HEMT en régime statique et hyperfréquence est étudié à basse température. L'influence du piégeage des électrons est analysé à l'aide de procédures spécifiques de mesure de la capacité de grille et la résistance drain-source. L'utilisation d'une structure appropriée permet de supprimer les effets parasites telle que la dégradation des caractéristiques I-V. La transconductance microonde et la fréquence de coupure sont alors améliorées ; ces résultats sont validés par une simulation numérique du composant.

Abstract - Low temperature behaviour of HEMTs is studied under static and microwave conditions. The influence of trapping effects are analysed by means of a specified measurement procedure of the gate capacitance and drain-source resistance. On an appropriate structure that allows to avoid parasitic phenomena such as collapse, it was found that microwave transconductance and cut-off frequency improve, which is confirmed with numerical results.

Two dimensional electron gas field effect transistors (TEGFETs, HEMTs, MODFETs) exhibit very attractive performance for super-low-noise amplifier and ultra-high-speed-logic applications. This is due to superior electron dynamics of electrons spatially separated from their parent atoms and located at the AlGaAs/GaAs interface.

Moreover, cryogenic temperatures give significant improvement of the mobility which increases, for instance, from 8000 $cm^2/v/s$ at 300 K to 1 x 10^5 $cm^2/v/s$ at 77 K [1]. Unfortunately, such an enhancement is not directly reflected to transistor characteristics. In practice, many parasitic phenomena occur at low temperatures, such as :
- the collapse in drain I-V characteristics
- threshold voltage shifts of the FET I-V characteristics
- persistent photoconductivity effects

Many authors [2,3] attributed the origin of these phenomena to the existence of electron trapping in the so-called DX centers [4] in the AlGaAs layer.

Based on experimental approaches, the aim of this work is :
- to progress in the comprehension of the physical mechanisms responsible for these effects
- to define optimal structures to minimize this anomalous behaviour
- to evaluate the potential advantages of modulation doped heterostructures.

Experiments were carried out on several transistors with conventionnal as well as more elaborated structures grown by MBE or MOCVD.

The access source and drain resistances that greatly affect the HEMT parameters at low temperature, are deduced from gate d-c current measurements. These resistances while decreasing with temperature, exhibit for transistors with different cap-layer parameters, a similar minimal value around 0.5 Ωmm at 77 K. The high mobility of electron gas yields an access resistance which is independent of the cap-layer and reflected only by the very low sheet resistance. The limited value then may be attributed to the 2DEG contact resistance. For temperatures below 190 K, these results are greatly affected if drain I-V characteristics were measured before measuring the access resistance. The drain resistance reaches then a considerable value whereas the source one is not modified and vice versa on inverting the role of source and drain (fig. I.1). This is correlated with the $I_D(V_{DS}, V_{GS})$ characteristics degradation resulting from the application of a large drain bias voltage (V_{DS} > 1V).

[1] For the CHS Laboratory, this Work was Supported under a DRET Contact, Number 86/167

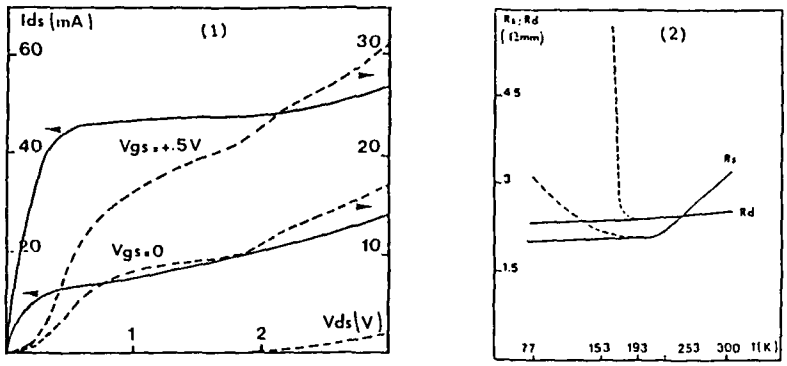


Fig. I : I-V characteristics (fig. I.1) at 77K and access resistances Rs and Rd (fig. I.2) for two HEMTs with (---) and without (—) collapse phenomenon.

However, we observed that the I-V collapse can be suppressed (fig. I.2) by means of :
- replacing the highly doped AlGaAs layer by a NID AlGaAs/N^+ AlGaAs structure.
- reducing the drain ungated recessed area for a standard epitaxial structure. In this case, electrons are directly transfered and collected in the cap-layer of the drain region.

Concerning the influence of technology, MOCVD epitaxy seems to be less sensitive to I-V instabilities even if a large drain voltage is applied to the device.

So, the state of a HEMT being dependent on his history, precisely the manner of reducing the temperature and applying the bias conditions, demands the definition of certain protocols before gate capacitance and drain-source resistance measurements in function of gate voltage.

Three different sequences are successively considered which correspond to electric stresses applied to the transistor during the cool down from room temperature to 77 K. This procedure is made in the very beginning of the experiments in order to separate as possible trapping effects and to make these experiments reproductible. The capacitance measurement is made at 500 MHz and after integrating, the threshold voltage W_{oo} and the total electronic charge quantity under the gate Q_o are deduced. The drain-source measurement gives information about the access resistances $R_s + R_d$ and the channel conductance G_o under the gate region (fig. II).

	Z = 300µm	Q_o (10^{-12} C)	W_{oo} (V)	R_o(Ω)	Rs+Rd(Ω)
1	T = 300K	3,08	2,15	0,68	10,8
2	(T = 77K ; Vg = Vp) Vds = 0	2,92	2,15	0,21	3,9
3	(T = 77K ; Vg = +0,6V) Vds = 0	1,76	1,44	0,10	4,7
4	(T = 77K ; Vg = +0,4V) Vds = 3V	1,59	1,30	collapse	collapse

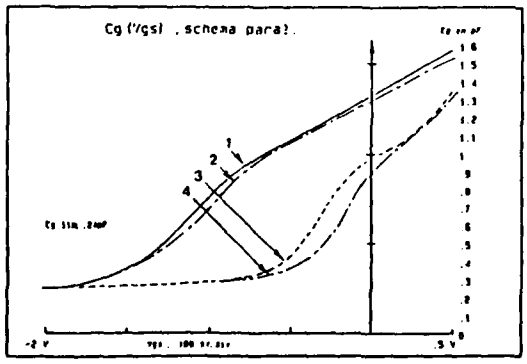


Fig. II : Evolution of gate capacitance for different bias stresses applied on gate and/or on drain contacts (Table).

For the first stress, the whole region under the gate being depleted, the trap level keeps above the Fermi level which allows full ionisation at low temperature. Emission and capture times become then so important that no variation is visible. The capacitance characteristic is close to the room temperature one and the threshold voltage value remains constant. For the second stress, some electrons tied to deep donor level are frozen to these centers. This makes them unavailable for ionization. In the following sequence, a drain source voltage is applied to the sample so that the collapse will be induced. Resistances in the gate drain region and under the gate greatly increase. In the high field region, electrons gain energy and can be captured and frozen in the deep level traps associated with the multivalley structure in AlGaAs. These trapping effects occur at the drain side of the gate where the electric field is the highest, after electrons have jumped the heterojunction barrier. So, if electrons still cold, are immediately collected in the cap layer of the drain region or cross a slightly doped AlGaAs layer, trapping effects such as collapse cannot then be developed. Indeed, no drastic increase of the drain resistance and of the pinch-off voltage shift are observed in case of the third sequence. These conclusions are also verified for transistors realized by MOCVD.

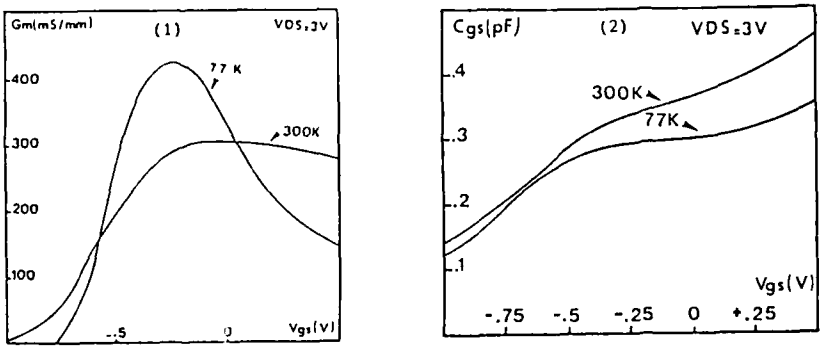


Fig. III : Microwave transconductance (fig. III.1) and gate capacitance (fig. III.2) at room and cryogenic temperatures for typical HEMTs

Due to the importance of defects in controlling electronic properties of HEMTs, it is important to point out their influence on microwave small-signal parameters of the device. An accurate scattering [S] parameters measurement method has been performed at 77 K and the transistor equivalent circuit is deduced under normal bias conditions. At low temperature, the transconductance increases by about 30% and the evolution with the gate voltage shows a selective variation (fig. III.1). This can be explained by a slight mobility improvement of a poor electron density in the AlGaAs layer when the channel is open. Variations of gate-source capacitance remain nearly constant (fig. III.2) and result in an enhancement of the cut-off frequency of more than 50%. This represents the typical HEMT behaviour at low temperature when no collapse phenomena are present.

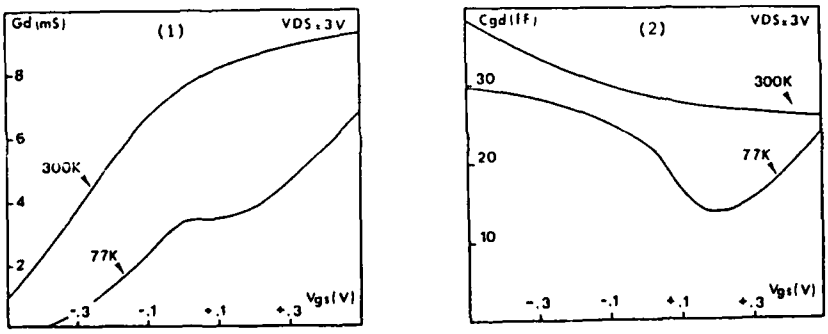


Fig. IV : Microwave output conductance (fig. IV.1) and output capacitance (fig. IV.2) for a device with collapse phenomenon.

The trapping effects imply a large decrease of the microwave transconductance. A pinch off voltage shift around 0.3 V modifies the evolution with gate-source voltage of each transistor parameters. The output conductance diminishes (fig. IV.1) and the capacitance variation ΔCgs is enlarged due to the drain resistance increase when the collapse effects occur. Evolutions of the Cgd (fig. IV.2) vary with different transistors but generally increase for the positive gate voltages. Comparison between microwave and static transconductance values shows for positive gate voltage an R.F. enhancement which is also observed at room temperature. This can be attributed to a less variation of electron quantity in AlGaAs at high frequencies.

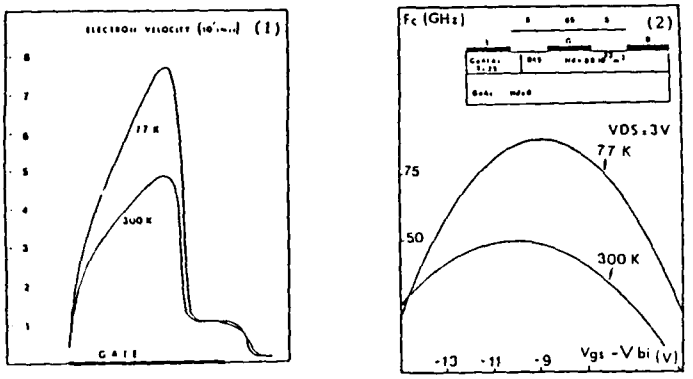


Fig. V : Evolutions of the carrier velocity (fig. V.1) and cut-off frequency (fig. V.2) at room and cryogenic temperatures.

A numerical simulation taking into account the nonstationnary electron dynamics effects for submicrometer gate transistor, allows us to evaluate the potentiality of microwave HEMT performances at 77 K. The small-signal equivalent circuit parameters for each bias point can be determined. Due to the large improvement of electron dynamics [5] with temperature shown in fig.V.1, the transconductance increases as observed by experiments, inducing a large enhancement of the intrinsic cut-off frequency (fig. V.2).

CONCLUSION

At low temperature, HEMT can now provide I-V characteristics without parasitic phenomena such as collapse, but the presence of deep-donor level in AlGaAs layer perturbs the transistor behaviour and limits the low temperature improvement of microwave parameters and consequently cut-off frequency and low noise performances. An optimal structure can be defined in replacing the AlGaAs material by several epitaxial layers, reducing the access source contact resistance and improving electron dynamics by the use of ultra-short gate length.

REFERENCES

[1] S. HIYAMIZU et al, "Improved electron mobility higher than 10^6 $cm^2/v/s$ in selectively doped GaAs/N-AlGaAs heterostructures grown by MBE, Japan J. Appl. Phys, vol. 22, pp. L606-L611, 1983.
[2] R. FISCHER et al, "Instabilities in MODFETs at 77K", Electron. Lett., Vol. EDL-19, pp. 789, 1983.
[3] T.J. DRUMMOND et al, "Bias dependence and light sensitivity of MODFET's at 77K", IEEE, Vol. ED-30, p. 1806, 1983.
[4] D.V. LANG et al, "Large lattice relaxation model for persistent photoconductivity in compound semiconductors", Phys. Rev. B, Vol. 19, p. 1015, 1979.
[5] A. CAPPY et al., "Noise modeling in submicrometer gate two dimensional electron gas field effect transistors", IEEE, Vol. ED-32, p. 2787-2796, 1985.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

MODELING AND PERFORMANCE OF DOUBLE HETEROJUNCTION GaAlAs/GaAs INTEGRATED INJECTION LOGIC. FABRICATION AND DC CHARACTERIZATION OF THE BASIC CELL

A. MARTY, J. JAMAI, J.P. VANNEL and J.P. BAILBE

Laboratoire d'Automatique et d'Analyse des Systèmes du CNRS
7, Av. du Colonel Roche, F-31077 Toulouse Cedex, France

This paper first presents a quantitative estimate of the potentialities of the GaAlAs/GaAs H.I^2L which relies on an accurate modeling related to physical device parameters. The interdependence of the forward and reverse current gains of the DHBTs processed by MBE and Mg-ion implantation which is subsequently analysed provides a verification of the charge-control models used for this evaluation.

I - INTRODUCTION

Single and double heterojunction bipolar transistors have stimulated much interest due to their potentialities as discussed in detail by KROEMER [1] and to the significant progress mode in device fabrication technologies. The emitter/collector interchangeability and the extra degree of freedom in device design resulting from the use of two heterojunctions, provide DHBT's with basic advantages -specially in the case of saturated logic circuits such as Integrated Injection Logic (I^2L)-that may be listed as follows : (i) reduction of the offset and saturation voltages, (ii) suppression of the hole injection from base to emitter and collector, (iii) reduction of the depletion junction capacitances and the base resistance. Both the discrete DHBT's and the H.I^2L circuits in the GaAlAs/GaAs material system, have already shown promising potentialities.

The purpose of this paper is first to provide a quantitative estimate of the potentialities of this technology. The fabrication and characterization of GaAlAs/GaAs DHBTs grown by molecular beam epitaxy with Mg-implanted external base are then presented. DC analysis is particularly concerned with the interpretation of the interdependence of forward and reverse current gains which, in the case of I^2L logic, correspond to the first performance criteria that have to be satisfied.

II - MODELING OF DOUBLE HETEROJUNCTION GaAlAs/GaAs INTEGRATED INJECTION LOGIC AND CIRCUIT PERFORMANCE EVALUATION

The cross-sectional view of the basic gate proposed by KROEMER [1] is shown in Fig.1a. It is realized with a GaAlAs(N)/GaAs(p)/GaAlAs(N) mutilayer structure which defines the double heterojunction transistor and a p^+ implantation which delimits the collector surfaces. Fig. 1b illustrates a typical layout for the rectangular narrow gate geometry which is considered here for discussion purposes. This gate geometry exhibits the most severe high current effects, but is popular because of ease of layout and high density considerations. The partitioning of the NpN base region is shown with dashed lines for a fan-out of two gate. Partitioning in this manner is most convenient for parameter assignment to the macromodel ; section B is the base contact region and sections C_1 and C_2 are the "active" regions (usually identical because of translational symmetry in layout).

The model equations are cast in terms of physically meaningful device parameters and are applicable at all current levels. The trade-off between accuracy and minimization of simulation time has led us to retain the lumped model illustrated in Fig. 2c. A one-dimensional model was suitably applied to represent the intrinsic part of DHBTs because emitter current crowding was expected to be unimportant for its inherently high conductive base layer Both "base contact" region and "external or extrinsic base" regions are also represented by a physical one-dimensional diode model. Series resistance was introduced to account for parasitic elements involved in the merged multicollector structure. In particular, to account for series resistance base distribution while avoiding excessive complexity, the DHBT's intrinsic parts have been "distributed" around the resistances R_{Be} and r_{Be} as shown in Figs 2abc With this layout the one-dimensional models T_i and T'_i correspond respectively to half the intrinsic cell C_i.

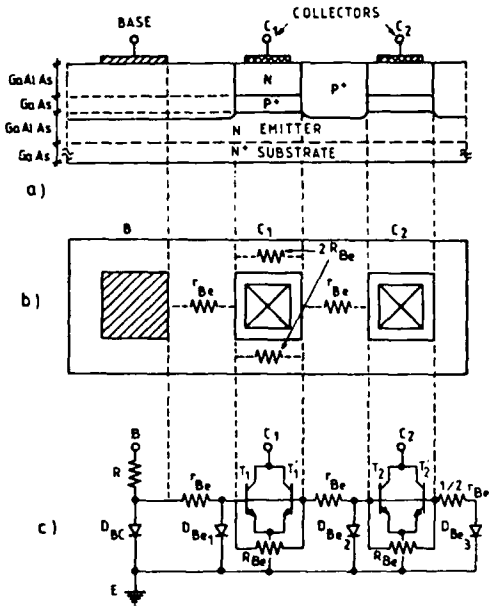


Figure 1 : a) Cross-section of an GaAlAs/GaAs H.I^{2}L inverter with fan-out of two ; b) geometrical layout ; c) lumped macromodel.

Our previously proposed DHBT model is defined by the charge-control formulation that is based on the macroscopic drift-diffusion motion of electrons in the emitter and collector band spikes and the quasi-neutral base region. Our convenient formulation which takes into account both the abrupt and graded cases of GaAlAs/GaAs heterojunctions remains compact and relatively simple. In particular it makes it possible to predict the decrease of the injection efficiency or of the base transport factor and the increase of both the stored charge and the emitter-collector offset voltage when a spike occurs at one of the junctions.

With respect to double heterostructure diodes (DHDs), an approach, in every aspect analogous to the one used to establish the SHBT and DHBT models, allows calculation of the charge residual storage in the low GaAs gap zone corresponding to the confinement of carriers induced by the isotype GaAs(p)/GaAlAs(P) heterojunction as well as the related recombination current.

By using our physical device models that accept as input the material and device fabrication parameters in the ASTEC III [2] circuit simulator, the effects of changes in the processing parameters on the circuit performance are analysed. This design procedure is developped so as to optimize the performance over a wide range of currents. The evolutions is propagation times as a function of the injector current calculated with "1 μm and 2 μm" design rules are given in Fig.2. It can be seen that with a minimum delay time of about 120 ps for the fan-out of one and the finest structure, the calculated power delay product was of the order of 2 fJ for a 1 μm consumption per gate. Note also that a fan-out of 7 gate defined by a "2 μm" design rule still yields an effective current gain of 16 and a noise margin of 72 mV.

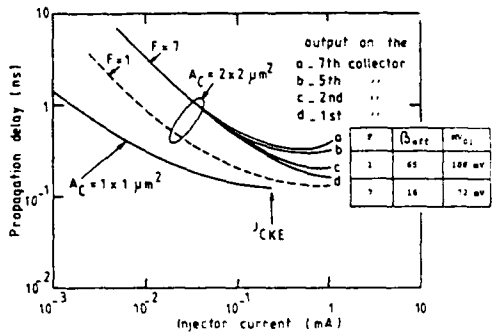


Figure 2 : Propagation delay-injector current characteristics for one and seven collectors optimized gates using 2x2 μm^2 and 1x1 μm^2 collector areas.

III - FABRICATION AND DC CHARACTERIZATION OF THE BASIC CELL

A schematic drawing of the implanted DHBT is shown in Fig. 3. The thirteen-layer structure was grown by MBE on a (100) Si-doped GaAs substrate. The emitter-base and collector-base heterojunction has a graded band configuration to prevent the occurrence of spike barriers which would adversely affect the injection and collection of minority carriers. The Al mole fraction and the linear grading width employed for the emitter and collector regions are 0.25 or 0.4 and 800 Å respectively.A 800 Å GaAs base doped with 1 or $5x10^{18}cm^{-3}$ Be matches a good current gain (base transport factor)-low intrinsic base resistance tradeoff. The undoped GaAs spacer layers of 200 Å were inserted between base and emitter, and between base and collector to allow for Be redistribution during emitter growth. Active emitter and collector are doped with Si to $10^{16}cm^{-3}$. Finally, 2000 Å -

thick GaAs cap layer doped to $2x10^{18}$ cm^{-3} was deposited to facilitate formation of ohmic contact.

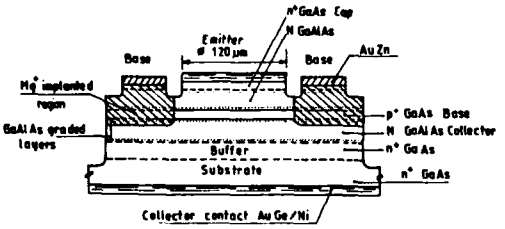


Figure 3 : Schematic cross-section of MBE grown-Mg implanted GaAlAs/GaAs DHBTs.

The P^+ extrinsic base-contact region was formed by selective multi-implantation of Mg selected so as to maintain a flat profile, (~ $10^{19}cm^{-3}$) down to the base layer [3]. Rapid thermal annealing was carried out in order to minimize the base Be diffusion.Au/Ge/Ni and Au/Zn were used as ohmic metals for n and p regions, respectively. The device's active areas were defined by deep chemical etch. The emitter-base and collector-base junction areas of large-size, circular test device are 1.2 10^{-4} and 5.8 $10^{-4}cm^2$ respectively.

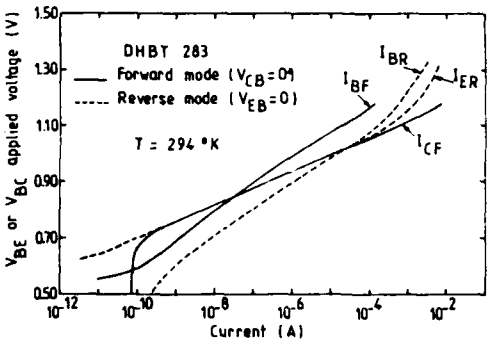


Figure 4 : Typical transfer and input characteristics of the DHBT operating in forward and reverse common-base modes.

Typical Gummel plots of a device operated in both emitter up (forward mode) and emitter down (reverse mode) configurations are show in figure 4. As expected, this device exhibits a perfectly symmetrical behavior in regard to the transfer of minority carriers : the characteristics $I_C(V_{BE})$ and $I_E(V_{BC})$ are fully superposed and exhibit an ideality coefficient of 1 over the whole range of currents for which the influence of the contact resistances remains negligible.

With respect to the base current, both configurations show similar behavior with an exp (qV/2kT) dependence at low levels reflecting the recombination surface effects on the GaAlAs extrinsic base region.

Figure 5 : Output characteristics in both forward and reverse operation.

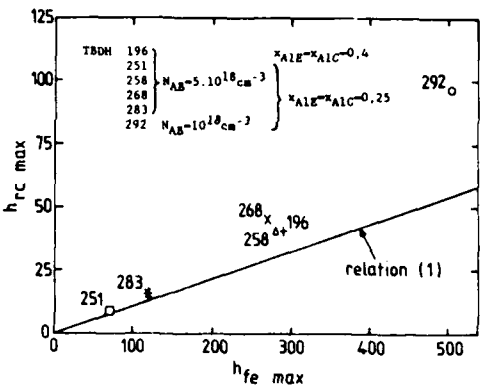


Figure 6 : Interdependence of the forward an reverse current gains of the different devices realized. Comparison with eq. (1) in which it has been considered that $\tau_{nBi} \cdot N_{ABi}/\tau_{nBe} \cdot N_{ABe} = 1$.

Figure 5 shows typical forward and reverse output characteristics. The best current gains ($h_{fe\,max} \approx 500$ and $h_{rc\,max} \approx 100$) and offset voltage ($V_{CEO} \approx 3$ mV) values were obtained for the devices with the lowest base doping level ($N_{AB} = 10^{18}$ cm^{-3}). As shown in figure 6, the correlation observed between the maximal values of the forward and reverse current gains, for the different series realized, constitutes a preliminary experimental verification of the theoretical model which provides for an interdependence relationship of the type :

$$h_{rcmax} = \frac{h_{femax}}{1+2\left[\frac{A_C-A_E}{A_E}\right]\frac{\tau_{nBi}\cdot N_{ABi}}{\tau_{nBe}\cdot N_{ABe}}} \quad (1)$$

in which τ_{nBi} and τ_{nBe} represent the electron lifetime in the intrinsic and extrinsic GaAs base regions. Thus it is possible to confirm the influence of the parasitic conduction of the lateral diode which offsets the symmetry of behaviors of the implanted DHBT.

IV - CONCLUSIONS

With the aid of this modeling directly related to physical device parameters and the results of a study concerning the sensitivity to technological characteristics, we have discussed optimization criteria and evaluated the potentialities of this logic family. The presented results show that for a structure defined by a design rule of "1 μm", a limit propagation time of about a hundred picoseconds is attained for a power consumption of about 300 μW per gate and a merit factor of about 2 fJ for 1 μW.

Preliminary results on double heterojunction GaAlAs/GaAs bipolar transistors processed by MBE and Mg-ion implantation have been presented. The devices making up the basic cell of the H. I^2L exhibited maximum forward and reverse current gains of 500 and 100 respectively and an offset voltage as low as 3 mV. The reported characterization chiefly aimed to analyze the interdependence of the forward and reverse current gains. As to this point, the results demonstrate that while the parasitic conduction induced by the residual storage in the collector-base lateral diode must be taken into account to insure accurate evaluation of the structure potentialities it must not, however, become a major drawback in obtaining an inverse gain permitting correct operation of I^2L (a gain of 100 has been obtained with our test-structures suffering from a very unfavorable A_C/A_E ratio of approximately 5).

ACKNOWLEDGEMENTS

The authors would like to thank B. Descouts, N. Duhamel, C. Dubon-Chevallier (CNET Bagneux) and J. Tasselli for performing ion implantation and RTA.

REFERENCES

[1] H. KROEMER, Proc. IEEE, Vol. 70, pp 13-25, Jan. 1982.

[2] M. HEYDEMANN, Notice d'utilisation ASTEC III, CISI France, 1978.

[3] K. DAOUD-KETATA, C. DUBON-CHEVALLIER, C. BESOMBES, IEEE Electron Device Lett., Vol. EDL-8, pp 205-207, May 1987.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

GaInAs CAMEL DIODES GROWN BY MBE

M. MARSO, A. CHIN*, P. BHATTACHARYA* and H. BENEKING

Institute of Semiconductor Electronics, Sommerfeldstrasse, D-5100 Aachen, F.R.G.
**Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2122, U.S.A.*

Résumé - Cet article présente les caractéristiques d'une diode du type "camel", fabriqué en GaInAs sur substrat InP, avec un facteur d'idéalité de 1,6 à 1,7. La hauteur de barrière de potentiel est 0,44 eV. Des mesures à inversion de courant montrent l'absence de temps d'emmagasinage et un retard à la descente de 660 à 740 ps pour une diode de 50 µm * 50 µm.

Abstract - In this paper we present a camel diode in GaInAs on a InP-substrate with an ideality factor of 1.6 - 1.7. The barrier height is 0.44 eV. Switching measurements show no storage time and a fall time of 660 - 740 ps for a 50 µm * 50 µm device.

1 - INTRODUCTION

Majority carrier devices in GaInAs should have good high frequency performance because of the very low effective electron mass ($m_e^* \approx 0.04 m_e$). Unfortunately the barrier height of metal/n-GaInAs contacts is too low for Schottky contacts.
Another majority carrier diode is the camel diode in which current flow is controlled by a potential barrier in the bulk of the semiconductor. Therefore the electrical properties of this diode are independent of the electrical and metallurgical properties of the interface between the metal and the semiconductor. The camel diode was first proposed by J.M. Shannon /1/ and has been realized so far in Si and GaAs /2,3,4/. It consists of three layers ($n^- p^+ n^+$), where the thickness of the p^+ region is chosen so that in thermal equilibrium this layer is fully depleted of free carriers. Fig. 1 shows the schematic diagram and the potential profile of a camel diode.

Assuming thermionic emission the current-voltage relationship is given by:

$$I_F = A \cdot A^* \cdot T^2 \cdot e^{\frac{-\phi_B}{k \cdot T}} \cdot \left(e^{\frac{V}{n V_T}} - 1 \right)$$

where A is the diode surface, A^* is the Richardson constant, T is the temperature, k is the Boltzmann constant, V_T is the temperature voltage and n is the ideality factor.

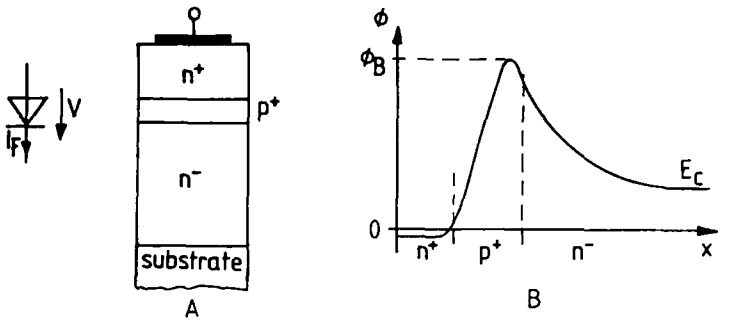


Fig. 1 - Schematic diagram (A) and potential profile (B) of a camel diode

E_C: conduction band edge
ϕ_B: barrier height
V : applied bias voltage
I_F: diode current

2 - DEVICE FABRICATION

The camel layers (n^- p^+ n^+) are grown by MBE on an n^+-InP substrate. The doping concentrations and the layer thicknesses are shown in table 1. A mesa etching is performed to determine the device area. A mixture of H_2SO_4:H_2O_2:H_2O (1:1:10) is used as etching solution. A 0.4 µm thick SiO_2 film is spinned on as isolation layer. The ohmic contact on top of the device is performed with Cr/Au (20nm/200nm). Indium is used for contact on the substrate (Fig. 2).

Table 1 Camel diode data

layer	Thickness/nm	doping concentration/cm^{-3}
n^-	800	$< 10^{16}$
p^+	15	$\approx 5*10^{18}$
n^+	200	$> 10^{19}$

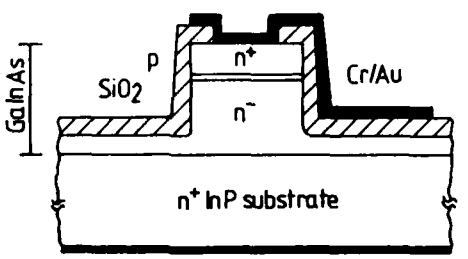


Fig. 2 - schematic diagram of a GaInAs camel diode

3 - MEASUREMENTS

a) I-V measurements

I-V caracteristics were measured on diodes with different areas. The current in the forward direction scales with the diode area, which proves the bulk nature of the current (Fig. 3). The ideality factor n is 1.6 to 1.7. Calculations show that n should be 1.2.

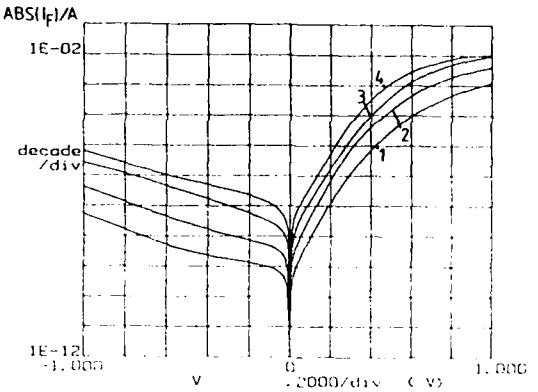


Fig. 3 - I-V caracteristics

diode areas: 1: 400 μm^2 3: 10 000 μm^2

2: 2 500 μm^2 4: 40 000 μm^2

b) I-T measurements

I-T measurements were performed to determine the barrier height. ϕ_B was determined as 0.44 eV at a bias voltage of 0.3 V and 0.5 V.

Using the barrier height, the p-type doping concentration is determined as $N_A=3.5*10^{18}$ cm^{-3} which agrees well with the set value for the p-type doping concentration.

c) Diode switching measurements

The diode was switched from forward to reverse current and the transient current was measured. Fig. 4 shows that there is no storage time. The reverse current reaches a maximum value Ir_{max} and then decreases to its steady state value, following approximately an exponential curve. The fall time T_F, which is determined as the time interval where the current to fall from $0.9*Ir_{max}$ to $0.1*Ir_{max}$, is 660 to 740 ps, dependent on the forward current. Absence of a storage time indicates that there is no minority carrier storage and that the current is in fact due to majority carrier transport. The transient behavior of the current is caused by the reverse of the charge in the space charge capacitance. For smaller geometries faster response is expected.

4 - CONCLUSIONS

Majority carrier camel diodes have been fabricated for the first time in GaInAs, grown by MBE. The ideality factor was 1.6 to 1.7 and the barrier height was 0.44 eV. Switching measurements showed that the current is due to majority carrier transport.

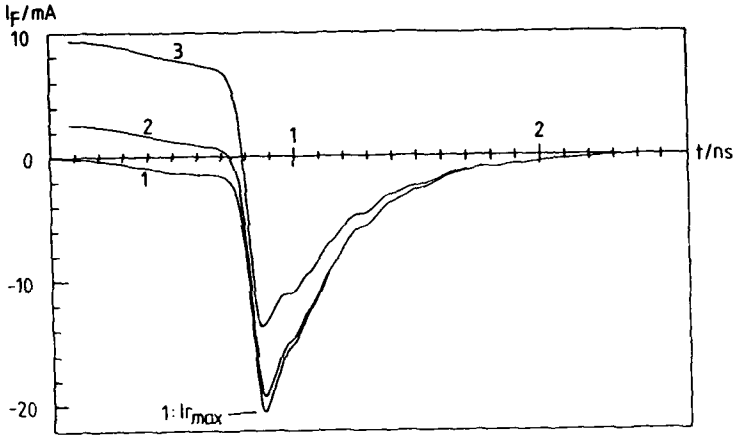


Fig. 4 - Transient current through the diode when switched from forward to reverse bias

diode area: 2500 μm^2

1: I_F = 0 mA ($\hat{=}$ 0 A/cm^2) T_F = 660 ps

2: I_F = 3 mA ($\hat{=}$ 120 A/cm^2) T_F = 700 ps

3: I_F = 10 mA ($\hat{=}$ 400 A/cm^2) T_F = 740 ps

5 - ACKNOWLEDGEMENTS

The authors would like to thank Mr. F. Schulte for helpful discussions and Mr. C. Werres for the switching measurements.

REFERENCES

/1/ J.M. Shannon: A New Majority Carrier Diode - The Camel Diode
Japanese Journal of Applied Physics, 19 (1980)
Supplement 19-1 pp. 301-304

/2/ J.M. Woodcock, J.J. Harris: Bulk Unipolar Diodes in MBE GaAs
Electronics Letters, 19 (1983) pp. 181-183

/3/ L. Vescan and H. Beneking: LPVPE-grown Silicon Camel Diodes
Electronics Letters, 22 (1986) pp. 994-996

/4/ J.M. Shannon: A majority-carrier camel diode
Applied Physics letters, 35(1), (1979) pp. 63-65

SESSION 7

Session 7IP Room A : Invited paper

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

RESIST SCHEMES FOR SUBMICRON OPTICAL LITHOGRAPHY

F. COOPMANS[1]

IMEC vzw, Kapeldreef 75, B-3030 Leuven, Belgium

Résumé - Premièrement on essaie d'expliquer pourquoi la lithographie optique avec des résines positives monocouche est aujourd'hui toujours la technique dominante, même pour une résolution submicronique. On indique comment les matériaux et les procédés ont été améliorés pour répondre aux exigences supérieures posées par les méthodes de fabrications avancées. En suite, nous discutons des limitations imposées par l'équipement et les procédés de transfert pour des dimensions de 0.7µm à 0.3µm. De nouveaux systèmes lithographiques doivent être utilisés pour obtenir une performance maximale, des systèmes bi-couche ou pseudo-monocouche sont les mieux adaptés. On compare leurs mérites et des aspects d'implémentation en fabrication.

Abstract - In this paper we will first try to explain why optical lithography with single layer positive resist is still a dominant technology today, even for submicron resolution. To this end we will indicate how the materials and processes have been improved to cope with the more stringent requirements. Next we will look at the future, at feature sizes under 0.7 µm and down to 0.3µm. For these small sizes the limitations imposed by the imaging equipment and the transfer processes call for new resist systems. The type of systems that are most promising are the pseudo monolayer and bilayer systems. Their relative merits are compared, and issues relating to the implementation in fab environments are discussed.

1 - INTRODUCTION

In recent years several new resist systems have been introduced for advanced lithographic applications. Multilevel systems, antireflection or contrast enhancement layers, dry developable resists, all are the result of intensive research in micron and submicron lithography. The strongly

[1] Presently also with : COBRAIN nv, De Regenboog 11, B-2800 Mechelen, Belgium

increased publication volume in this area supports the increased activity. The evolution of the 'normal' resists has less glamorous aspects but the improvements in performance that are obtained are remarkable. It is exactly because of these improvements that the single layer systems managed to maintain a leading position in volume fabrication and are likely to do so even for submicron processing down to 0.8 μm.
Nevertheless as feature sizes continue to shrink the thickness and flatness requirements can limit the usefulness of single layer systems. Surface imaging and dry transfer will be necessary to reap the benefits from the improving equipment technology.
We will first discuss the very basic lithographic requirements and the trends in single layer positive resists process and material design.
Secondly we will show by refining the criteria where the limitations of the standard processes lay and why surface imaging and dry transfer systems are the best option for advanced optical lithography.

2 - BASIC LITHOGRAPHIC REQUIREMENTS

The formation of a relief image in a photoresist is a result of an imaging action. Obviously, the quality of the relief image will be in the first instance influenced by the quality of the aerial image. The higher the modulation (or the contrast) of the image information, the more relaxed are the requirements for the resist. It is 'straightforward' to print submicron lines in even a very mediocre resist provided the image contrast is high. Therefor one should try to improve the imaging system first, better resist systems can subsequently be used to achieve ultimate performance with a given tool. The resist contrast becomes more important as the information in the optical image degrades. At the resolution limit of the optical projection system the image contrast is inherently low and a high contrast resist is needed to regenerate the image and produce an acceptable relief image. High resolution and steep sidewalls drive the quest for high contrast materials.

High contrast does not mean a good latitude, thresholdlike behaviour will transfer all variability in the optical image. Particularly dose variations due to topography (bulk effect), standing waves and scattering of light will be accurately printed in the resist. It follows that high contrast materials are more useful if these exposure variations can be avoided or compensated.

The exposure tools used present an important capital investment and the sensitivity of the resist should be high enough to allow adequate troughput. In practice this issue is becoming less important as the overhead time in set-up alignment etc. increases. Where sensitivities under $100mJ/cm^2$ were normal in the past, the trend is now to accept values up to $200mJ/cm^2$ provided resolution improvements are obtained. Although the sensitivity seems high this is perhaps the most relaxed criterium, in other imaging applications, e.g. dry film, printing etc., the imaging materials can be 3 to 4 decades more sensitive. The photopolymer also must be compatible with the semiconductor processing in general and the specific application. This poses

stringent criteria in terms of contamination levels (moving into the ppb range). Also the compatibility with the substrates e.g. stress, adhesion must be assured. To be able to apply the same material to a variety of substrates (oxides, metals, semiconductors, polymers) different coupling agents are used. These in turn have to maintain the same contamination levels as the resists.

The relief image is used to transfer the pattern into the active layer by etching or by ion implantation and a certain thickness of the mask is required to perform this process. A minimum thickness is needed at all times to avoid pinhole formation. In addition to this we must include a topography factor as the material thickness on the step is usually smaller than the nominal thickness, depending on the characteristics of the topography and the planarising properties of the polymer. A higher inherent planarisation efficiency requires more resist thickness. To this we must add the erosion of the mask during etch or the penetration of ions into the masking material. For typical cases in todays semiconductor fabrication this results in a minimum resist thickness of about 1 micron. This requirement is very severe considering that the features we want to fabricate are of the same size.

Patterning processes often generate heat. This increase in temperature should not affect the shape of the polymer structures. As the glass transition temperature of novolak-based resist is low, process temperatures above 120° C cause deformations. As the effect is due to plasticising the effective deformation is also dependent on the actual geometry of the structure. The sidewall angle tends to decrease more rapidly for large structures than for small ones. This makes the linewidth variation after e.g. etching dependent on the local geometry. This limitation is caused by the source material itself but process solutions, e.g. hardening techniques or a better temperature control in the patterning processes, can be used advantageously.

3 - POSITIVE OPTICAL RESISTS : A CONTINUING STORY

The first positive optical resist materials used commercially where based on 'natural' novolak. Although this class of materials was the first one to be synthetically reproduced, a huge research work needed to be done to understand and improve the behaviour of the materials in lithographic applications and to produce them consistently. Initially the resists were a simple mixture of sensitiser and base resin in an appropriate solvent, but in the course of the last five to ten years dramatic improvements were made. The impact of this activity is amplified by the improvements in the stepper design. High NA lenses and better focussing systems provide high contrast images on the wafer surface. It is the synergistic development of material and tool that explains why single layer resists maintained a leading position in volume fabrication and succeeded in pushing back the introduction of more performing resist schemes (multilevel) and even new techniques, e.g.

x-ray. We will indicate some of the items that are important in the formulation of new resists.

3.1 FACTORS OF THE RESIN AFFECTING PERFORMANCE

Staying with cresol-novolak systems the most important factors that can be varied to impact the performance are the molecular weight, the isomerisation and the position of the methylene bond, its orientaton in the polymer chain.
Increasing the molecular weight will decrease the overall development rate. This decrease is proportional for the exposed and unexposed regions so that no change in contrast is obtained. It is clear that due to the slower develop rates the sensitivity is degraded.
The isomeric structure in the novolak can be varied by incorporating a given percentage of ortho, meta or para components. The meta-para ratio has an important effect on the development rates. An increase of the para component strongly decreases the development rate of the base resin. This effect is much weaker for the exposed areas and so this ratio influences the contrast of the resist.
The structure of the polymer with respect to the position of the methylene bond can have a similar effect. It was reported [1] that in m-cresol novolak with highly regular ortho-bonding position the dissolution rate of the polymer decreased drastically. This effect was attributed to azocoupling reactions, which are not present in the exposed parts.
In table 1 [1] we give a summary of these trends. Although several other parameters play and the end performance is interrelated, the table presents a possible way of increasing contrast without loosing sensitivity.

Resin factors	Dissolution rate		Resist performance	
	Unexposed	Exposed	Contrast	Sensitivity
Molecular Weight	↓	↓	→	→
meta/para cresol ratio	↑↑	↑	↓	↑
methylene bond position	↓	→	↑	→

Table 1 : Impact of design factors on lithographic performance

3.2 THE PHOTOACTIVE COMPONENT

The photoactive compound (PAC) is usually based on naphtoquinonediazides. An important point in the actual performance is the degree of esterification. If the sensitiser is made polyfunctional one obtains a contrast enhancement

effect. This is illustrated in figure 1 [2]where we show the transformation of the optical modulation function into the modulation of the relevant photoproduct concentration. For a tri-functional photoactive component (q=3) a clear contrast increase can be observed. The first positive resists used typically monofunctional sensitisers, the application of bi- and tri-functional materials has allowed a marked increase in contrast. Also the incorporation of the PAC can be changed from simply mixed with the resin to chemically atatched to the main chain.

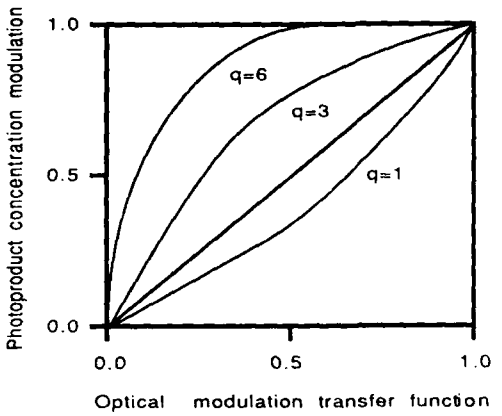


Figure 5 : Theoretical Contrast improvement with multifunctional sensitiser

The sensitiser also acts as a bleachable dye. In this way reflections cannot be completely avoided since the sensitiser must be partly converted (bleached) to function as dissolution rate enhancer. To reduce reflections further one can therefore add a non-bleaching dye. This approach suffers some drawbacks : the side wall angle will decrease and the sensitivity of the system is greatly reduced. Nevertheless, resists using non-bleachable dyes are used and add some process flexibility where reflective notching of patterns is important, typically at metal steps. Attemps to increase the functional speed of dyed systems with low molecular weight speed enhancers are compromised by their reduced thermal stability.

3.3 THE SOLVENT SYSTEM

Safety regulations are imposing changes in the existing solvent systems. The mostly preferred solvent is ECA but reports concerning health hasards have strongly decreased its future importance. Possible substitutions are PMA (propyleneglycolmethyletheracetate) or diglyme, ethyllactate and several others. Apart from the solubility and handling (and cost) , the remaining solvent in the film after coating and prebake is very important. There is appreciable variation among the different solvents. This factor affects the adhesion and the dissolution of the unexposed areas. In addition to the main solvent, a number of co-solvents can be used e.g. xylene,

propylenecarbonate etc. that help to control the evaporation rate which is important for the spincoating uniformity.

3.4 ADDITIVES

A colourful variety of additives has been introduced over the years. We will not give a complete list but simply give the main purpose of the additives and some examples. These materials are obviously part of the black magic that goes together with bringing a photoresist to the market.
Plasticicers are introduced to improve the coating properties. Planarisation is influenced by these additions. Polypropylene, phenoxy resin, polyvinylmethylether and others can be used to this end.
Apart from the actinic dyes we discussed above sometimes also colorants are added that absorb in the red or green. These additions serve to facilitate better alignment precision and more precise measurements with automated optical linewidth systems. They avoid stray reflections at the operating wavelengths of these devices.
Up to 5% adhesion promotors can be added in the resist to enhance the coupling with e.g. doped oxides . Typically materials containing silane groups are used, methyl-disilanemethyl-methacrylate and vinyltrichlorosilane have been used.
Speed enhancers can be added to relatively high (perhaps 20%) weight ratio. These additions serve to increase the development rate of the basic material and are of acidic nature (pitric acid, nitrocinnamic acid....). Although faster resists suffer from reduced contrast and more erosion in unexposed areas, this can cause problems on extensive topography.
Another component that can be used is typically alkylphenoxy-(ethyleneoxy)ethanol as a non ionic surfactant. In addition to this some specific anti-striation agents (a few weight percent) can be added if desired.
The combination of all these products makes the tuning of formulations for specific tasks possible. The difficulty that remains is the design of one single resist that can be used in all applications...

3.5 THE DEVELOPER

To comply with the contamination requirement the older metal-hydroxide based materials must be abandoned and people turned to 'metal ion free' e.g. TMAH type developers. The characteristics of these developers are quite different and new optimisation was necessary. Typically contrast is degraded when using a TMAH-type material while maintaining sensitivity.
A technique for improving resist contrast that has become popular makes use of diluted developer. Because the developer is weaker the sensitivity of the process is reduced. A 20% dilution can produce as much as a 100% decrease in sensitivity. Therefore this approach contains a tradeoff between speed and resolution and is most interesting when one operates at the resolution limit of the exposure system. There is a strong trend in advanced production to sacrifice troughput for resolution and allow resist systems to work with sensitivities up to 200 mJ/cm^2 .

Another approach to improving the developer action is to add surfactants that enhance the selectivity between the exposed and unexposed parts due to the different degree of polarity. Also additions are possible that improve the removal of residues e.g. out of small contact holes. These additions can be complemented by processing refinements such as double step development.

3.6 CONCLUSION

In general the largest gain is obtained from carefully optimising the resist and developer composition together. This requires large amounts of formulation and experimentation. The tuned resist materials and developers play with the equilibrium between the etching and cross-linking or 'reticulation' of the resist by the developer. This results in fast vertical development and passivation of sidewalls to prevent erosion that leads to sloped profiles. Tuned systems therefor minimize the bulkeffect giving excellent profiles and improved process latitude, the interference effect is however not compensated. The combined impact of these improvement makes these resist systems applicable in 1 and 4 megabit type technologies.

4 - LIMITATIONS IMPOSED BY OPTICS AND APPLICATIONS

What holds the future for optical lithography? To answer this question we must investigate more closely the limitations imposed by the projection equipment and the applications that must be supported. We will adress issues as far as they are important for the resist systems used, which means that some important limitations e.g. field size and overlay are not considered here.

4.1 RESOLUTION

In optical projection systems for lithografic applications the minimum linewidth (W) that can be printed is usually given by $W = \kappa\lambda/NA$. Where NA is the numerical aperture of the optics and κ a 'constant' that incorporates a number of imaging and processing effects, e.g. contrast of resist. Obviously higher resolution is obtained by using shorter wavelengths and higher NA optics. Starting initially with g-line exposure, proven equipment is now available for i-line (365 nm) and also deep-uv (248 nm) steppers are emerging. Lenses have been fabricated with a NA above 0,6 for g-line and the same is possible in i-line. For 248 nm lenses with a NA in the 0.5 range are feasible. This implies a basic resolution capacity down to 0,4 μm for i-line and 0,3μm with deep-uv optics, assuming a κ-value of 0.65.
The resist material itself presents no problem at these geometries, its fundamental resolution limits are defined by the chain size and this is still small compared to these feature sizes. Therefore we can conclude that in principle the resolution is limited by the optical information present in the resist material. As maximum resolution is achieved at short wavelengths we

must also assure that the materials are sensitive enough in this range. The novolak systems have a serious drawback with the high absorbance of the resin at these wavelengths.

4.2 THE DEPTH OF FOCUS (DOF)

Two aspects must be discussed and compared, the actual depth of focus of the projection tool and the image depth required by the application.
The depth of focus is in practice controlled by a large number of parameters e.g. λ, NA, the degree of coherence, abberations, stray light, vibration and also the shape and density of the patterns projected. Usually the simplified relationship $DOF = \kappa' \lambda/NA^2$ is employed where κ' is chosen to incorporate all degrading parameters. Clearly the DOF decreases rapidly with high NA optics and shorter wavelengths. For the two maximum resolution lens systems mentioned above (i-line and deep-uv) we find a DOF= 0.7 μm. The actual budget that is required is again a sum of different contributing factors. Resist thickness, wafer topography, wafer and mask flatness, precision of focus setting and tilt. Topography and resist thickness account for the largest part, around 1.5 μm. The other factors contribute about 1 μm in reduction optics. As we discussed in section 2 we allways need a minimum mask thickness of about 1 μm, which means that, even if we reduce all the other factors to zero, the defocus will limit the performance. From this we conclude that we must introduce resist systems that allow image formation in a thin surface layer to push this limit. For a more detailed overview of the DOF issue and its relation to resolution we refer to reference 3.

4.3 THE IMAGE COUPLED INTO THE RESIST

The image as projected is distorted in the resist by standing wave effects, scattering of light and absorption. Even in high contrast situations these effects can degrade the fidelity of the image. The standing wave effect particulary affects the total dose absorbed locally, while the stray reflections introduce faulty image information. The bulk effect can pose problems when the topography on the wafer is large. In addition the absorbtion increases as the wavelength goes down for the common positive resists. This makes it unpractical to fully expose a 1μm thick layer of material in deep uv. To overcome all these effects it is necessary to absorb all radiation in a surface layer, a different technique is than needed to transfer the pattern down into the whole masking polymer.

5. RESIST SCHEMES

In this last section we will discuss how creative thinking has introduced a variety of techniques to overcome the limitations mentioned in the previous section. We do not pretend to present a complete list of possibilities, but we cover what we feel are the important players in the different cathegories. In table 2 we summarise the comparison between all these systems in terms of

performance increase over a standard positive resist. For reference we have included x-ray imaging. Although excellent results are obtained the overall complexity of the technology e.g. mask-making, synchroton source, is such that it cannot favourably compete with optical lithography.

	RESOLUTION	REFLECTION	BULK EFFECT	FOCUS	CONTRAST	LATITUDE	COMPLEXITY	INVESTMENT
DYED RESIST	SMALL	YES	NO	NO	NO	YES	LOW	NO
RESIST + ARC	MEDIUM	YES	NO	NO	NO	WORSE	MEDIUM	SMALL
IMAGE REVERSAL	MEDIUM	YES(DYE)	MEDIUM	SMALL	MEDIUM	NO	MEDIUM	SMALL
TRI-LEVEL	V. HIGH	YES(DYE)	YES	YES	YES	WORSE	V. HIGH	MEDIUM
CEM	HIGH	SMALL	MEDIUM	MEDIUM	YES	YES	MEDIUM	SMALL
PCM(DRY)	HIGH	YES(DYE)	YES	YES	YES	NO	MEDIUM	MEDIUM
DESIRE	V. HIGH	YES(DYE)	YES	YES	YES	YES	MEDIUM	MEDIUM
SABRE	HIGH	YES(DYE)	YES	YES	YES	SOME	HIGH	MEDIUM
NEW POS. RES.	MEDIUM	NO	MEDIUM	NO	MEDIUM	MEDIUM	LOW	NO
X-RAY	ULTRA-HIGH	YES	YES	YES	YES	YES	V.HIGH(MASK)	V. HIGH

Table 2 : Comparison of improvements obtained with different resist systems.

5.1 IMPROVING THE COUPLING

A first group of systems tries to overcome the limitations introduced by interference effects and low image contrast.
Interference can be controlled by adding a dye to the resist. The beneficial effect is however counteracted by an increased bulk effect. Therefor the dye concentration is optimised so that reflective notching is somewhat suppressed but speed and sidewall angle are still acceptable [4,5].
One can avoid the increase in bulk effect by coating a anti-reflective layer (ARC) prior to spinning the resist This implies the complexity of a bi-layer system.
A contrast enhancement material (CEM) can be added on top of the resist layer. Due to its nonlinear bleaching characteristic this layer restores the contrast of the aereal image [6]. This introduces all the benefits from high illumination contrast e.g. improved latitude and high resolution but reflections are not countered.
All these techiques can be combined and have proven usefull for repairing specific situations but none of them tackles the DOF issue.

5.2 REVERSING THE IMAGE

By using a wet developed image reversal resist a number of improvements are seen [7,8]. The process consists of an expose-bake-expose-develop sequence. During the bake and the second exposure the deviations in the first imagewise exposure can be countered. This technique allows excellent control of the bulk effect and gives reasonably high contrast. The chemical

conversion during the reversal bake smooths the reflection effects and notching is somewhat suppressed by the second exposure (if wanted a dye could be added) . This system is relatively simple and shows improvements in nearly every aspect over the positive resists. Unfortunately the DOF criterium for low thickness and topography is not met. However this technology is a strong contender if an intermediate solution is needed.

5.3 GOING DRY

By decoupling the imaging and the masking function we can push the DOF limit. The original systems that were used contained three layers [9]. A planarising botom layer that also functions as the actual mask and as an ARC, a thin top resist to produce a shallow relief image and a buffer layer inbetween to facilitate the transfer of the relief image into the bottom layer. The bulk effect was totally avoided by using anisotropic dry etching as transfer technique. The big drawback of these systems obviously is the complexity. Therefore the quest was to reduce the number of layers and processing steps without losing performance.

Bilayer systems (PCM) using wet transfer were enhanced by using a oxygen etch resistant top resist [10]. These resists are unfortunately difficult to formulate and to work with, although deep-vu sensitivity is no problem. The contrast and compatibility of the organometallic resists used are still a cause for scepticism. Here also the botom layer serves the same purpose as in the tri-level system.

In our opinion the systems that modify the resist after exposure are inherently more promissing because the imaging action is not compromised by the presense of metallic componants. Two types of gas phase functionalised systems are shown in figures 2 and 3 below.

The first one is a bi-layer technique (SABRE- Silicon Added Bilayer Resist) [11]. In the broad sense this structure can work either in positive or negative mode depending on the choise of the materials and behaves like a PCM type system. The difference is that the metallic compound is introduced by diffusion from gas phase after the formation of the relief image in the resist. Some drawbacks of this scheme are in the wet development step that is still needed and in the possible formation of a interfacial layer that causes residue or hinders the silylation process. Nevertheless this technique can work with adapted positive resists as a thin high contrast top layer and should give outstanding performance.

The ultimum form of back integration is presented by pseudo-monolayer systems as in the DESIRE (Diffusion Enhanced Silylating Resist) scheme [12,13,14]. Here a dyed single layer of resist functions as planarising layer, ARC and mask but at the same time a thin surface region of the same layer mimics a thin high contrast resist. After a shallow exposure the material is selectively silylated and the relief image is obtained by anisotropic dry transfer. This system overcomes all the image coupling problems and can be used in deep-uv. Also the DOF limit is not constrained by resist thickness, only the exposure depth (typical about 300nm) is relevant. This makes its performance similar to the multilevel systems while remaining a single layer technique. The materials used can be very close in composition to

common positive resists and this is advantageous for formulation and compatibility.

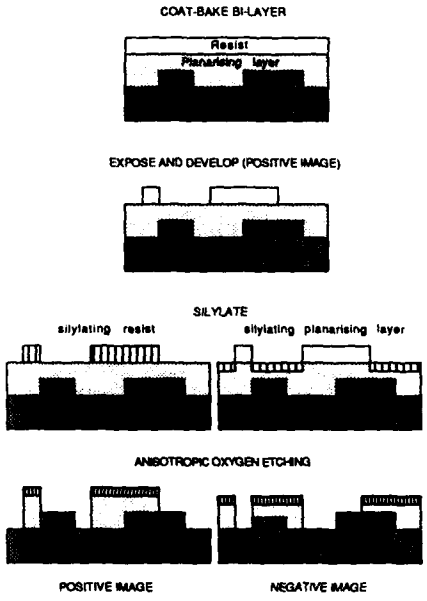


Figure 2 : SABRE process flow

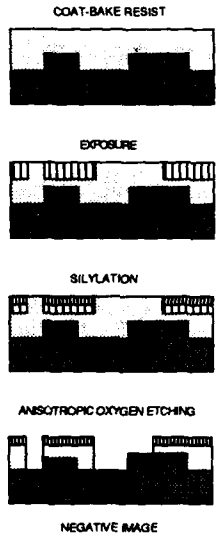


Figure 3 : DESIRE process flow

5.4 IMPLEMENTATION

The implementation of these new resist systems in fabrication lines should produce a minimal distortion of the accepted ways of working. The ARC, CEM and image reversal systems all use wet development and use existing equipment. However, from the discussion above we know that to take full advantage of the possibilities of an advanced stepper we must turn to bilayer or DESIRE-like systems. The introduction of these techniques has been hampered by the availability of suitable production equipment to support the silylation and etching processes. For both steps increasing activity is seen and dedicated tools are being developed for single wafer in-line operation.

6. CONCLUSION

Working resolution in the range from 1-0.7 μm can be achieved by advanced positive resists and additives or by investing in high performance steppers. From 0.5-0.3μm only surface imaging systems can comply with the DOF requirements and they will dominate. In the mid range from 0.7-0.5μm the

battle is on between positive resists, image reversal and surface imaging systems. The selection will be strongly impacted by the will to invest first in resist system technology rather than in steppers (deep-uv). Ultimately both camps must join forces to reach the subhalfmicron regime.

AKNOWLEDGEMENT

The author wishes to thank all the members of the Micropatterning Group at IMEC for their support.

REFERENCES

[1] Hanabata, M., Furuta, A., Uemura, Y., Proc. of SPIE, p. 85-92, (1987),771
[2] Trefonas III, P., Daniels, B.K., Fischer, Jr., R.L., SST, p. 131 (1987) 30, nr. 8
[3] Lin, B. J., Microelectronic engineering 6 , p.31-51,(1987)
[4] Willson, C.G., ACS Symp. series, p.148,(1983),219
[5] Bol, I., Proc. Interface '84, Kodak, Rochester,p.19, (1983)
[6] West, P. R., Griffing, B. F., Proc. of the SPIE, vol. 364, p. 33-38, (1983)
[7] Spak, M. , et al, 7th Int. technical conference on photopolymers, Ellenville, NY
[8] Gijsen, R., et al, Proc. of the SPIE vol. 631, p. 108, (1986)
[9] Havas, J. R., et al, US patent 3,873,361 (1973)
[10] Lin, B.J., ACS Symp. series, p.287-350,(1983),219
[11] McColgin, W.C., Daly, R. C., Jech, J., Brust, T. B., Symposium on VLSI technology Karuizawa,Japan, Technical digest, p.9-10, (1987)
[12] Wolf,E.D., Taylor, G.N., Venkatesan, t., Kretsch, R.T., J.El.Chem.Soc.,p.1664, (1984),131
[13] MacDonald, S.A., Ito, T., Hiroaka, H., Willson, C.G., SPE Conf, Ellenville, p.177,(1985)
[14] Coopmans,F.A., Roland, B., Proc. of SPIE, p.34,(1986),631

Session 7A Room A : MOS transistors

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

HALF-MICROMETER N-MOS TECHNOLOGY FOR GIGABIT LOGIC

G. GUEGAN, M. LERME, J. GAUTIER, M. GUERIN, S. TEDESCO, B. DAL'ZOTTO and G. REIMBOLD

Division LETI, IRDI, DMEL/S.MSC, CENG, 85X, F-38041 Grenoble Cedex, France

Résumé - En vue de situer les performances des technologies N-MOS, un assemblage N-MOS 0.5 µm (Enrichi/Déplété) a été mis au point. Les résultats dynamiques mesurés sur un additionneur trois bits bouclés sont comparés avec les performances d'autres filières. La fréquence de fonctionnement de 1 GHz ainsi obtenue est supérieure à celle mesurée avec une technologie GaAs LPBFL dont la longueur de grille est sensiblement plus élevée pour une puissance consommée similaire et un même nombre de niveaux logiques (10).

Abstract - In order to point out the performances of NMOS technology, a fully-scaled 0.5 µm enhancement/depletion NMOS process was integrated and compared with other technologies, using the same test circuit : a three bit feed-back adder. The average experimental maximum frequency of operation is 1 GHz which is better than results obtained on GaAs LPBFL for about the same power consumption and number of logic layers (10) but lightly shorter gate length.

1 - INTRODUCTION

Many submicrometer MOSFET designs which have been published show a large interest in the fabrication of very small FET structures to explore the performance limits of different technologies. The objective of this study is to report electrical performances of very high speed circuits based on 0.5 µm silicon N-MOS technology. 38 ps propagation delay time and 11 fJ power delay product were achieved from standard 101 stages ring oscillator. A ripple carry 3-bit feed-back adder, which has also been designed, performs a maximum experimental working frequency of 1.02 GHz. This performance is compared with results from the same circuit realized with other technologies (GaAs - STL - Cryogenic NMOS).

2 - DEVICE FABRICATION

The design rules of this NMOS technology was obtained by a factor 2 shrink of a previously optimized 1 µm technology. Geometrical and process characteristics are summarized in table 1.

Table 1 - Geometrical and process characteristics

		1µm	0.5µm
Lithography		optical	E-beam
Supply voltage	V	5	3
Gate length	µm	1	0.5
Channel length	µm	0.8	0.35
Gate oxide	nm	25	15
Junction depth	µm	0.22	0.17
Channel width Enhancement device	µm	9	4.5
Depletion device		5.5	2.75
Hole contact	µm	1.4	0.7

Degradation of the active MOSFET's due to hot carrier generation is one of the key point of submicron design rules. To achieve this requirement a solution used in this work is to scale the power supply voltage down to 3.0 V which has been recognized as a new standard power supply voltage for the near future. This approach simplifies the process by avoiding lightly doped drain or double diffused drain and lead to high drain drivibility. For that channel doping profile and source, drain profiles were optimized to avoid punchtrough current for 3 V supply voltage. Wafers were fully exposed by E-beam direct write on all levels. A trilayer resist system was used for gate, contact and metal levels with either negative (CMS) or positive (WX 214) E-beam resist depending on the level tone. In order to obtain a low gate resistance, tantalum polycide was chosen as gate material with a sheet resistance of 1.8 Ω/□. For scaling consideration, 0.17 µm source-drain junction depth was obtained by Arsenic implant with a measured sheet resistance of 72 Ω/□ . Different doses for depletion channel implant were performed in order to reach either a high speed or a low power dissipation. Process limits were checked by overetching the gate level on given wafers.

3 - RESULTS AND DISCUSSION

Drain current versus drain voltage of a typical enhancement MOS transistor is plotted in Fig. 1. These electrical device characteristics are in good agreement with predictions and optimization criteria : V_T=0.75 V, sub-threshold slope = 100 mV/decade, BVDSS > 4 V, DC transconductance Gm of 122 mS/mm. Short channel effects are minimized and both lithography and etching controls are excellent : the mean value and the standard deviation of the channel length for the typical gate are respectively 0.383 µm and 11 nm (Fig.2). Hot carrier induced degradation tests of the devices show that a 10 years life time is obtained without LDD on a worst case 0.4 µm gate length for 3 V operating voltage. However an increase of the supply voltage would need the use of LDD structure /1/.

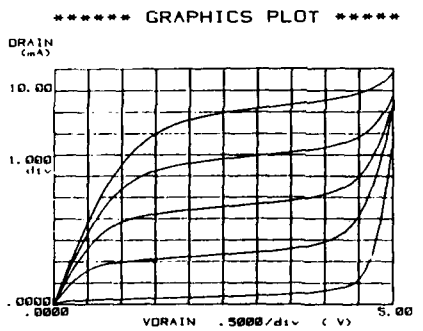


Fig.1 - ID vs VD as a function of VG for an enhancement device with Leff = 0.35 µm, Weff = 19 µm

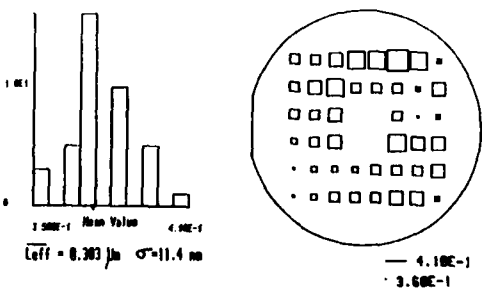


Fig.2 - Statistical results of channel length

Ring oscillators with different loading conditions were processed and tested. Gate delay mean value for FI = FO = 1 is 47 ps (σ= 1.7 ps) and a 33 fJ power delay product for 3 V supply voltage was achieved. Best speed performances are measured on overetched wafers : in this case the channel length is 0.27 µm and a 38 ps propagation delay time was achieved. Ring oscillators work for supply voltage down to 1.8 V and in these conditions very low power delay product was measured : 11 fJ. These results are favorably compared with other N-MOS dynamic performances obtained with different lithographic tools (e-beam /2/, X-ray /3/, ion Beam /4/, optical lithography /5/ /6/ in Fig. 3.

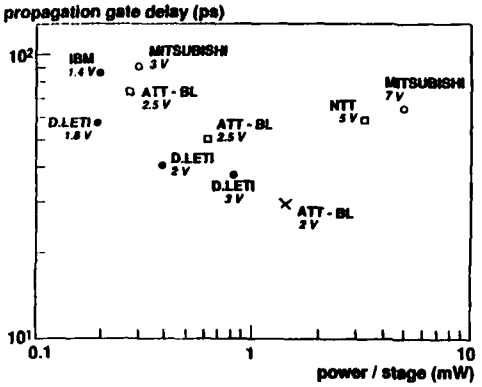


Fig.3 - τ_p - P results for different N-MOS technologies
●E-beam-□ optical-○ ion beam-X xray

Ring oscillators have the advantage of the simplified design but restricted meaning on their dynamic performance. So in order to have a more significant comparison with results obtained with different technologies, a 3-bit feed back adder has been choosen as a better demonstrator circuit because it requires combinatorial and sequential logic (block diagram is shown in Fig. 4). This circuit adds two 3-bit numbers A and B. The sum S_1, S_2, S_3 and carry are injected back to the inputs by four conventionnal 6 NOR gates D flip-flop. The adder speed is the maximum clock frequency for which the adder works. A SEM view of this circuit processed with 1 µm N-MOS and 0.5 µm N-MOS technology is shown in Fig. 5. A maximum measured frequency of 1.02 GHz for a 44 mW power consumption was obtained. Fmax - P chart (Fig. 6) shows that the adder speed with half-µm N-MOS technology is the fastest, even compared to LPBFL GaAs 3 bit adder circuits /7/ (2 µm design rules - 0.6 µm channel length) or cryogenic N-MOS /8/.

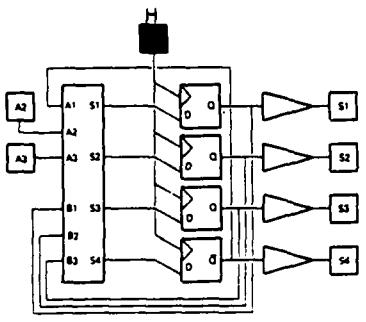


Fig.4 - 3-bit adder block chart

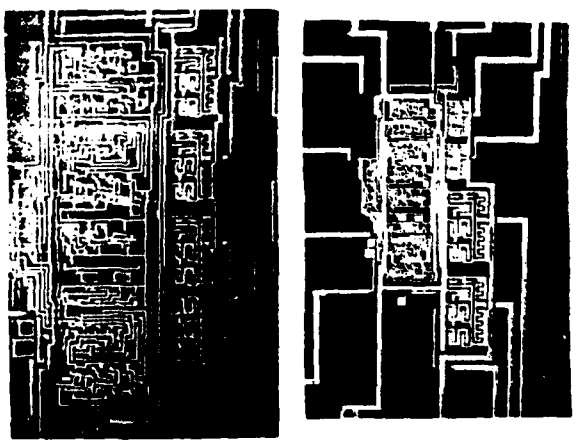


Fig.5 - SEM picture of 1 µm and 0.5 µm 3-bit adder

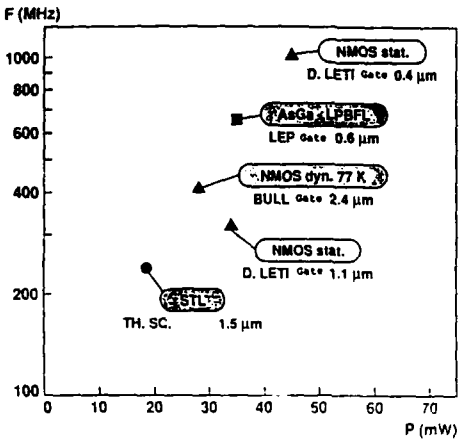


Fig.6 - F_{max} -P results for different technologies

4 - CONCLUSION

A 0.5 µm N-MOS process for high speed circuits has been developed. Process was well controled and a Gigabit logic circuit was realized as a demonstrator. The performances are among the best ones reported (τp = 38 pS, PTp = 11 fJ, 1.02 GHz clock frequency for 3-bit feed back adder) and can even be favorably compared with performances achieved with GaAs technology.

ACKNOWLEDGEMENTS

This work has been supported by DRET contract.
The autors would like to thank the different teams of D.LETI - SMSC for device processing and testing.

REFERENCES

/1/ G. Reimbold et al.
"Hot electron reliability of half-micron MOSFETs"
ESSDERC 88

/2/ M.R. Wordeman et al.
"A Fully Scaled Submicrometer NMOS Technology Using Direct-Write E-Beam Lithography"
IEEE Transactions on Electron Devices ED-32 N° 11 - Nov. 1985

/3/ G.W. Taylor et al.
"A comparison of Si MOSFET and GaAs MESFET Enhancement/Depletion Logic Performance"
IEEE Transactions on Electron Devices ED-32 - N° 9 - Sept. 1985

/4/ H. Morimoto et al.
"Submicrometer - gate MOSFET's by the Use of Focused Ion Beam Exposure and a Dry Development Technique"
IEEE Transactions on Electron Devices, ED 34 - N° 2 - Feb. 1987

/5/ W. Fichtner et al.
"A Submicron NMOS technology suitable for low power High Speed circuits"
IEDM 1985

/6/ H. Inokawa et al.
"A Submicrometer Lifted Diffused-Layer MOSFET"
IEEE Electron Device Letters, Vol. EDL. 8 - N° 3 - March 1987

/7/ A. Mitonneau et al.
"Direct experimental comparison of submicron GaAs and Si NMOS digital IC's"
GaAs IC symposium - BOSTON 1984

/8/ P. Glories et al.
"A High - Performance N-MOS Adder Designed for optimized Cryogenic Operation"
IEEE Journal of Solid-State Circuit - SC 21 - N° 3 - June 1986

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

OPERATION OF MAJORITY AND MINORITY CARRIER MOSFET'S AT LIQUID HELIUM TEMPERATURE

B. DIERICKX, E. SIMOEN, J. VERMEIREN and C. CLAEYS

IMEC, Kapeldreef 75, B-3030 Leuven, Belgium

Résumé - L'opération de transistors MOS "à porteurs majoritairs" ou en "mode d'accumulation" à température de hélium liquide est rapportée et comparée avec MOSFET's normaux (à porteurs minoritaires). Les effets adverses "kink/hysteresis" sont absents, tandis que les niveaux de courant sont très comparables. Cela offre une alternative pour CMOS cryogène: tenant compte de certaines réserves, une technologie CMOS "sans-caisson" basée sur n^+nn^+ et p^+np^+ MOSFET's dans une plaque de type n est une façon d'éviter les anomalies d'un Si-MOSFET à basse température.

Abstract - The operation of "accumulation mode" or "majority carrier" MOS transistors at liquid helium temperature is reported and compared with normal (minority carrier) MOSFET's. The adverse kink/hysteresis effects are absent, while the current levels are highly comparable. This offers an alternative for cryogenic CMOS: under certain reserves, a well-less CMOS technology based on n^+nn^+ and p^+np^+ MOSFET's in a n-type substrate is a way to avoid the Si-MOSFET cryogenic anomalies.

1 - INTRODUCTION

As yet the only technologies suitable for cryogenic operation of Si integrated circuits are MOSFET based /1/. The reason for this is that majority carrier conduction through neutral regions, as in bipolar transistors and JFET's, is greatly reduced by freeze-out effects /2-3/.
Although one expects improved transistor performance from high electron and hole mobility, cold MOSFET's suffer from undesired kink and hysteresis effects /4-10/.

2 - EXPERIMENTS

To avoid sidevall effects the measurements were performed on circular nMOSFET's, processed on <100> material, with a gate oxide thickness of 60 nm. The p-well doping level is about 1e16 /cm^3, and the n-substrate doping is 1.2e15 /cm^3. The transistor width W = 96 µm, its length L = 12 µm. The cryogenic measurements were done in a liquid helium bath cryostat.

3. - "NORMAL" MOSFET CHARACTERISTICS

At 4.2 Kelvin Silicon MOSFET's, especially n-channels, exhibit the kink and hysterisis phenomenon (fig. 1 and 2a).
The observed hysteresis and kink in the I/V characteristics can be attributed to a local bulk effect underneath the transistor channel. A local decrease in depletion layer width or charge density must be compensated by an increase of the inversion layer charge, and thus by increased transistor current. At 4.2 Kelvin the thermal generation/recombination time constants for depletion layer formation approaches infinity /4/, and from this standpoint no important bulk effect on the inversion layer is expected. However it was shown that the depletion layer width and charge density can be modified by a mechanism based on impact ionization and recombination of shallow (dopant) levels in the bulk. This process is induced by a hole current (in a nMOSFET) flowing from the pinch-off region at the drain towards the substrate back contact. The hole

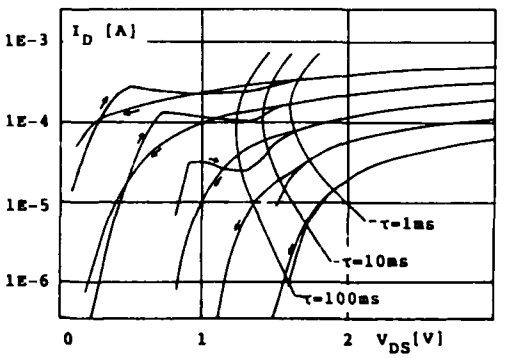


Fig.1 I_D/V_{DS} characteristics of the "normal" n-MOSFET at 4.2 Kelvin on logarithmic scale. Superimposed are lines of constant small-signal time constants.

current originates from the avalanche ionization at the pinch-off end of the channel.

In general these holes interact in 3 ways while travelling through the bulk:
a) the hole simply drifts, as governed by the free hole mobility in the local electric field
b) by impact the hole ionizes a neutral acceptor, thereby freeing an additional hole
c) the hole recombines with an ionized acceptor.

Dependent on the local electric field, and charge densities, an equilibrium space charge will be estabished after some time. The time constant for this "forced depletion layer formation" ("FDLF") is in first instance inversely proportional to the mentioned substrate current /4/, which in turn is proportional to the avalanche ionization rate, being a strong exponentional function of pinch-off voltage (across pinch-off region at the drain), with an onset at about 1.2 Volts (=Si bandgap) /11/.

The time constants for FDLF can be indirectly measured from the transistor small signal transient behaviour. When curves of equal time constants are plotted in the same plane as the transistor I/V sweeps (fig. 1), the origin of the hysteris is identified as time delay in reaching the momentary equilibrium when the transistor's V_{DS} is varied.
In the same view the kink coincides with the closing of the hysterisis loop, i.e. where the time constants get faster than the incremental V_{DS} voltage sweep rate.

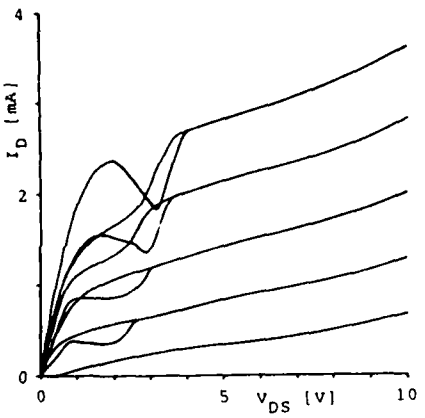


Fig. 2a
nMOSFET at 4.2 Kelvin
I_D/V_{DS} plot: V_{GS} = 1, 1.5, 2, 2.5, 3 V
V_{BS} = 0 V

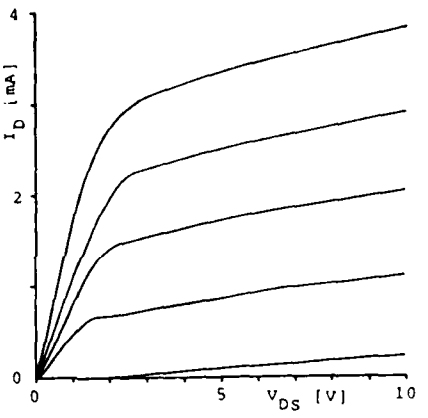


Fig. 2b
majority carrier nMOSFET at 4.2 Kelvin
I_D/V_{DS} plot: V_{GS} = 0, 0.5, 1, 1.5, 2 V
V_{BS} = 0 V

4 - "MAJORITY CARRIER" MOSFET's AT 4.2 KELVIN

The majority carrier equivalent of the above nMOS device is nothing else than exactly the same transistor, on the same wafers, designed in the n-substrate instead of in the p-well.
The majority carrier n-channel MOSFET (fig. 2b) does not feature the kink or hysteresis in its I/V characteristics as it has not the mechanism where the bulk depletion layer (which is absent) can compensate the inversion layer. Conduction takes place through an accumulation layer. The accumulation layer charge obeys the relation

$$Q_{acc} = (V_G - V_T - V_s(y)) \cdot C_{ox} \qquad (1)$$

with V_T the "threshold voltage" for forming an accumulation layer, and $V_s(y)$ the local surface potential.

The current through the device can be calculated as

$$I_D = W \cdot Q_{acc} \cdot \mu_e \cdot (dV_s(y))/(dy) \qquad (2)$$

Substituting (1) in (2) and integrating from source to drain yields the familiar expression:

$$I_D = W/L \cdot \mu_e \cdot C_{ox} \cdot [V_{GS} - V_T - V_{DS}/2] \cdot V_{DS} \qquad (3)$$

Saturation occurs for $V_{DS} = V_{GS} - V_T$, when the accumulation layer at the drain side has vanished. Higher V_{DS} will then be supported by the pinch-off region.
The quadratic behaviour of (3) is reflected in the characteristics of fig. 2b. Besides the absence of kink and hysteresis both transistors behave similar. The small bias mobilities are about 2800 cm^2/V for the normal nMOSFET, and 3200 $cm^2/V.s$ for the majority nMOSFET. At relatively low biases already mobility saturation occurs, so that the curves for consecutive V_{GS} are almost equidistant.

parallel $n^+n^-n^+$ resistor

On the curves of fig. 2b, and more in detail in fig. 3 one observes the onset of conduction of a kind of "parallel resistor" to the transistor. This resistor can be identified with the $n^+n^-n^+$ structure of the transistor. Fig.4 show a schematic cross section both of the transistor in off and on state (i.e. in accumulation).
The conduction at 4.2 Kelvin of such a $n^+n^-n^+$ structure is governed by a similar mechanism as the depletion layer formation mentioned above.

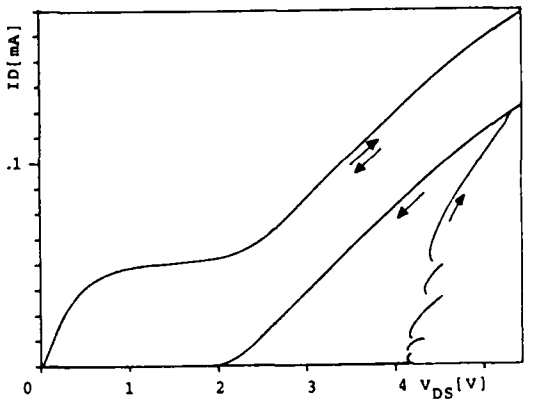


Fig. 3 I_D/V_{DS} of turn-on region of the majority carrier nMOSFET of fig. 2b

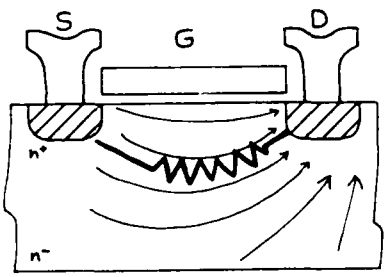


Fig. 4a
Schematic cross-section of the majority carrier nMOSFET, below threshold

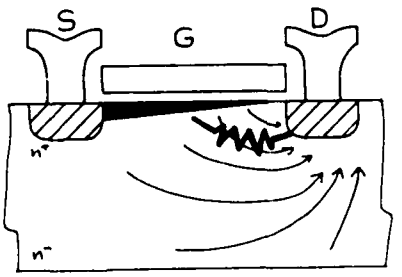


Fig. 4b
The same transistor while conducting current through the accumulation layer

Before conduction the charge distribution in the $n^+n^-n^+$ is homogenous. Any potential step is evenly distributed over the sample. Conduction starts at a certain field by tunneling through the n^+n^- barrier. By shallow impact ionization a charge redistribution builds up across the device, yielding a higher electric field at the injecting contact /12/. The current therefore is space charge limited, and turns off at much lower voltages than it turned on, which explains the bistable behaviour of the transistor in the off-state. In contrast, while the transistor is carrying current the $n^+n^-n^+$ structure is always "on", and the bistability disappears.

According to Dargys and Zuraukas /13/ the current in an $n^+n^-n^+$ resistor at freeze out temperatures obeys:

$$\Delta I = \Delta V \cdot 2.A.v_s.\varepsilon/d^2 \qquad (4)$$

for large enough fields to reach saturation of drift velocity v_s, with d the distance between the contacts (here one can think of some representative distance beween conducting channel and drain) and with A the relevant current-carrying cross-section (approximated by W.d). Substituting $v_s = 1.2E7 cm/s$ and $\varepsilon = 8.85E-13$ F/cm, one obtains a d of 5 µm (compare with a L of 12 µm).

5 - CONCLUSIONS

The combination of the described $n^+n^-n^+$ majority carrier transistor and the regular pMOSFET in the same substrate seems a feasible technological approach for building a cryogenic CMOS process with reduced kink. Indeed the pMOSFET is only slightly affected by kink, and the majority nMOSFET is kink-free except for a turn-on phenomenon. The process is readily obtained by omitting the p-well in the CMOS process sequence.
However, as a certain bulk leakage current will flow from the majority carrier devices, the circuits power dissipation may be a problem at these low temperatures.
The experiments also support the current theory explaining the 4.2 Kelvin anomalous behaviour in normal silicon MOSFET's.

references

/1/ B.Lengeler, "Semiconductor devices suitable for use in cryogenic environments", Cryogenics 14, 439 (1974).
/2/ W.P.Dumke, "The effect of base doping on the performance of Si bipolar transistors at low temperatures", IEEE Trans. Electron Devices ED-28, 418 (1981).
/3/ A.K.Kapoor and T.S.Yayadev, "Operation of Poly bipolar transistors near liquid helium temperatures (9K)", IEEE Electron Devices Letters EDL-9, 177 (1988).
/4/ B.Dierickx, L.Warmerdam, E.Simoen, J.Vermeiren and C.Claeys, "Model for hysteresis and kink behaviour of MOS transistors operating at 4.2 K", to be published in IEEE Trans. Electron devices, July 1988.
/5/ S.B.Broadbent, "Operation of CMOS devices at 10 Kelvin", Extended Abstracts of the Electrochemical Society, Vol 87-2, 486 (n° 344) (1987).
/6/ R.L.Maddox, "p-MOSFET parameters at cryogenic temperatures", IEEE Trans. Electron Devices, ED-23, 16 (1976).
/7/ H.Hanamura, M.Aoki, T.Masuhara, O.Minato, Y.Sakai and T.Hayashida, "Operation of bulk CMOS devices at very low temperatures", IEEE Journal Solid-State Circuits SC-21, 484 (1986).
/8/ S.B.Stetson, D.B.Reynolds, M.G.Stapelbroek and R.L.Stermer, "Design and Performance of Blocked-Impurity-Band detector focal plane arrays", Proceedings of SPIE, Vol 686, 48 (1986).
/9/ E.A.Hendricks, R.J.J.Zijlstra and J.Middelhoek, "Current Noise in <111> n-channel Si-MOSFET's at T=4.2K", Physica 147B, 297 (1987).
/10/ F.Balestra, L.Audaire and C.Lucas, "Influence of substrate freeze-out on the characteristics of MOS transistors at very low temperatures", Solid-State Electron. 30, 321 (1987).
/11/ B.Eitan, D.Frohman-Bentchkowsky and J.Shappir, "Impact ionization at very low voltages in silicon" J.Appl.Phys 53, 1244 (1982)
/12/ D.D.Coon and S.D.Gunopala, "New injection mode infrared detector", J.Appl.Phys. 57, 5525 (1985)
/13/ A.Dargys and S.Zurauskas, "Observation of the transition from impact-ionisation-dominated to field-ionization-dominated impurity breakdown in silicon", Solid State Commun. 52, 139 (1984).

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A NOVEL METHOD FOR DIMENSIONAL LOSS CHARACTERIZATION

P. CAPRARA, C. BERGONZONI and T. CAVIONI

SGS-Thomson Microelectronics, Via C. Olivetti 2, Agrate Brianza (MI), Italy

Abstract. *The present tendency toward devices scaling in VLSI technology makes more and more difficult the electrical characterization of channel dimensions.*
The narrow channel effects have had a great consideration for what concern device modeling, but a minor attention about dimensional loss problems. The aim of this work is to propose an effective width characterization method based on transconductance and not affected by the typical problems related to narrow channel devices.
Moreover, a compared analysis of this method to an other one previously proposed [1] is shown pointing out the phenomenological differences.

1 Introduction

In 1 μm processes development the active areas pitch scaling is not necessarily associated to an adequate scaling of the isolation process.
For EPROM devices the dimensional losses induce poor current and speed performances and the possibility of a marginal verification of the memory cell.
For this reason it is useful the implementation of a reliable electrical evaluation of the effective channel width of a MOST to control the process; moreover the evaluation method must be independent from the process variations.

2 Experimental and results

The measures were carried out on two sets of n-channel transistors with 280Å gate oxide thickness, polysilicon gates and ion implantation process.
The first set was constitued by native and enhancement mode devices with a channel stop implant of 6.5×10^{12} cm^{-2} and has been treated at a temperature of 920 °C (low T); the other set was in the same mode with a channel stop implant equal to 7.5×10^{12} cm^{-2} and treated at 1100 °C (high T).
The threshold shift implant was equal to 7×10^{11} cm^{-2} for both sets and the substrate concentration was 7×10^{15} cm^{-3}.
The channel length was 10μm for all the samples and the width varied between 2.6μm and 1.0μm with a W_{step} equal to 0.2μm.

We are interested in the evaluation of the MOST effective geometrical width W_{eff}, that is by definition the width of the active area below the gate oxide and delimitated by the bird's beak.
The selective oxidation by Si_3N_4 pattern permits the field oxide penetration under the Si_3N_4 itself; for this reason the final active area dimension will be less than the one on mask.
Now we can consider the behavior of drain current of a MOST operating in the linear region:

$$I_d = \frac{W_{eff}}{L} \mu C (V_g - V_t) V_d$$

where $W_{eff} = W_{nom} - \Delta W$
W_{eff} is the active area width covered by the gate oxide and limited by the bird's beak, ΔW is the dimensional loss and W_{nom} is the dimension on mask.
By plotting the drain current values as a function of W_{nom} for a set of constant length MOS devices and choosing the best straight line, the value of ΔW is determined as intersection with the W_{nom} abscissa.
In order to avoid the narrow channel effects on threshold voltage it is advisable to apply a $V_g \gg V_{th}$ [1].
In figure 1 we show the application of the method with V_{gs} as a parameter; it is evident the strong dependence of ΔW by the gate voltage value.
On the other hand the dependence of the I_d method by temperature treatments is shown in figures $2(a, b)$ where the SEM analysis results of the dimensional loss are reported as a comparison.
This behavior is explainable with different diffusion profiles of channel stop implant; in fact we can consider the real MOST as a parallel of an ideal transistor with $W = W_{eff}$ and a "parasitic" one whose gate oxide is the bird's beak and the isolation implant is the threshold implant.
To understand the behavior of the narrow channel MOST it is useful to plot the I_d and the trasconductance (beta) against gate voltage. In figure 3 the curves of low T n-channel native transistor are shown; it can be noted the beta canonic shape.
Viceversa when we get an high T treatment of the isolation implant the drain current grows almost linearly with V_{gs} as showed by the constant beta value also for high gate voltage values (fig.s $4.a, b$); this effect is stronger for the enhancement mode device as reported in fig.$(4.b)$.

3 Transconductance method

By considering the previous results we can summarize the following considerations:

- the enhancement transistor presents a more evident drain current contribution of the bird's beak parasitic than the native one; this is due to the similarity of the two "parallel" devices threshold voltages (fig.s $4.a, b$) and is enhanced by an higher T treatment on channel stop implant;
- the transconductance *plateau* is induced by the"parasitic" contribution compensating the mobility loss;
- in figure 5 the mobility loss gets over the parasitic when V_{gs} is higher than $4 \div 5$ Volts.

Avoiding the effect described above we implemented the transconductance measure in the same way as we treated the drain current, that is by plotting the maximum beta versus W_{nom}.
The implementation of this method is shown in figures 6 and 7; in both cases we measured native and enhancement transistors where in fig.6 the samples were.treated at low T, while in fig.7 at high T.
As we can see in the insert the agreement with SEM analysis is good.
This method is more precise than the I_d one because it permits to identify the same channel status (the strong inversion beginning) irrespectively of the threshold voltage. Because we check the beta maximum, the problem connected with the non-linear effect of V_{gs} on mobility is no more a constraint.

4 Conclusions

The method presented and based on transconductance seems to be quite reliable because it overcomes the problems related to mobility saturation [2], narrow channel effects, threshold voltage and isolation process.
We have outlined the problems connected with the drain current method pointing out the behavior of the narrow channel devices.

References

[1] Ying-Ren Ma et al.(IEEE TRANS. ON ELECTRON DEV.,VOL.ED-29,NO.12, DEC.1982).
[2] P.Wang (IEEE TRANS. ON ELECTRON DEV.,VOL.ED-25,NO.7,JULY 1978).
We thank A. Maggis of S.T. for the useful discussions and suggestions.

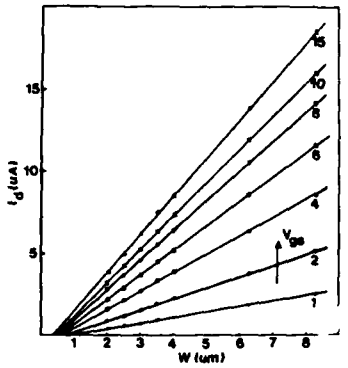


Fig.1 - I_d vs. W_{drawn} with V_{gs} as a parameter for n- channel enhancement transistor.
[Gate length 10μm]

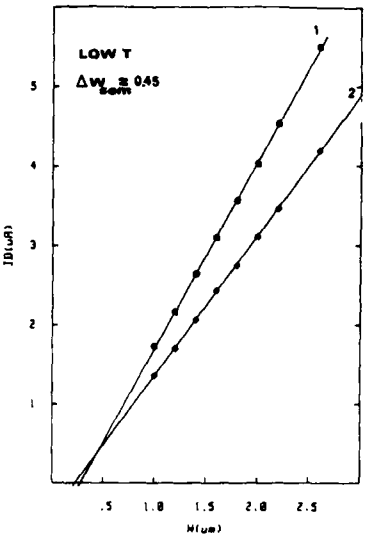


Fig. 2.a - I_d vs W_{drawn} on native (a) and enhancement (b) transistors for low Temperature treatment (920 C).
[V_{gs}= 3 V native - V_{gs}= 5 V enhanc.]
[Gate length 10μm]

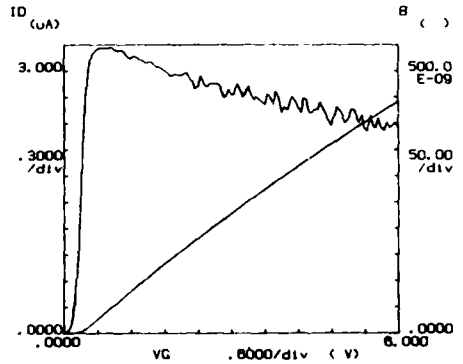


Fig. 3 - Threshold voltage and transconductance characteristics against gate voltage for native n-channel transistor at low T.
[W_{drawn} 1μm, L 10μm]

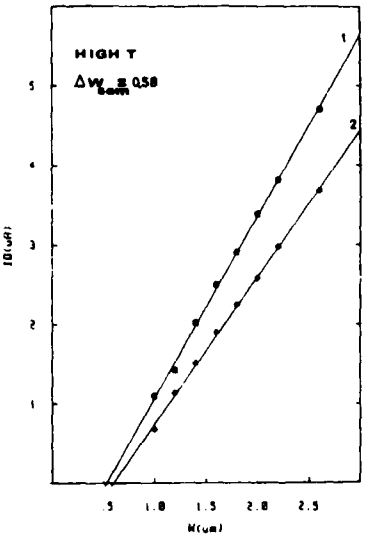


Fig. 2.b - I_d vs W_{drawn} on native (a) and enhancement (b) transistors for high Temperature treatment (1100 C).
[V_{gs} = 3 V native - V_{gs} = 5 V enhanc.]
[Gate length 10μm]

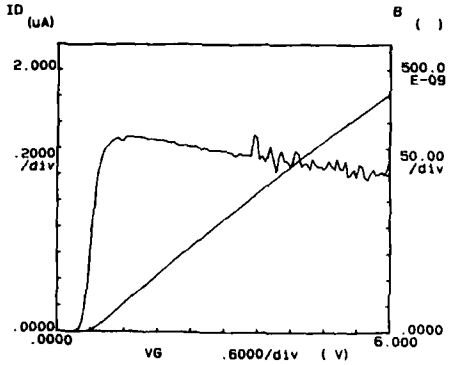


Fig. 4.a - Threshold voltage and transconductance characteristics against gate voltage on n-channel native transistor at high T.
[W_{drawn} = 1μm, L = 10μm]

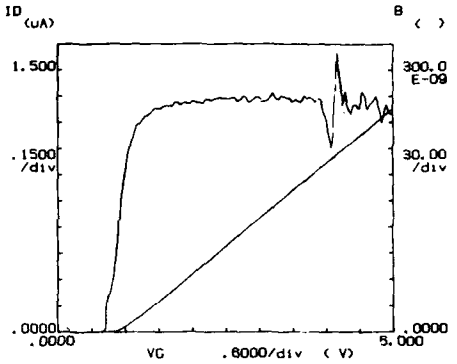


Fig. 4.b - Threshold voltage and transconductance characteristics against gate voltage on n channel enhancement transistor at high T.
[W_{drawn} = 1μm, L = 10μm]

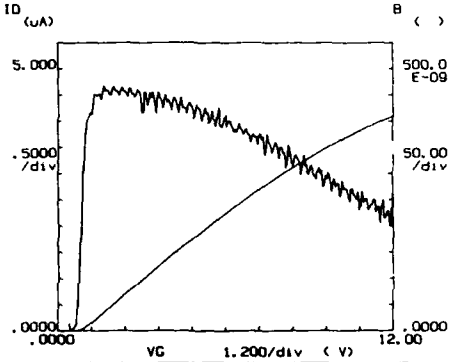


Fig. 5 - Threshold voltage and transconductance characteristics on n-channel transistor till high V_{gs} values (high T treatment).
[W_{drawn} = 1μm, L = 10μm]

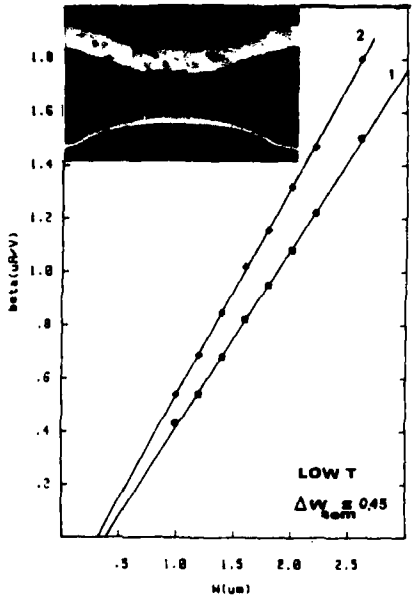


Fig. 6 - Transconductance versus W_{drawn} for n-channel native (1) and enhancement (2) transistors.
The SEM cross section is reported in the insert [21K magnification].

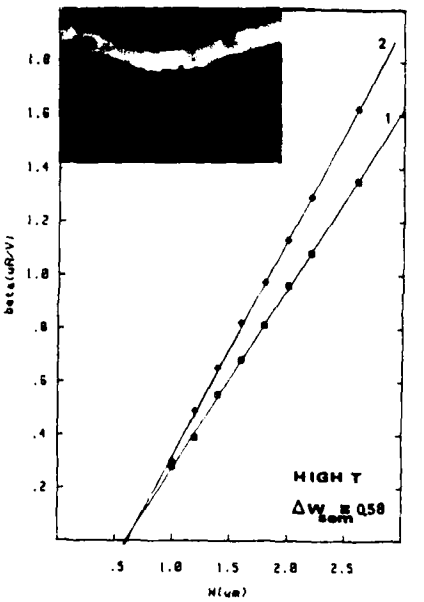


Fig. 7 - Transconductance versus W_{drawn} for n-channel native (1) and enhancement (2) transistors.
The SEM cross section is reported in the insert [21K magnification].

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

NARROW CHANNEL EFFECT ON n- AND p-CHANNEL DEVICES FABRICATED WITH THE SILO AND BOX ISOLATION TECHNIQUES

J.L. COPPEE[1], E. FIGUERAS*, B. GOFFIN, D. GLOESENER and F. VAN DE WIELE

Université Catholique de Louvain, Laboratoire de Microélectronique, Place du Levant 3, B-1348 Louvain-la-Neuve, Belgium
**Centro Nacional de Microelectrónica, CSIC-UAB, Universitat Autònoma de Barcelona, SP-08193 Bellaterra, Barcelona, Spain*

Résumé - Nous avons étudié l'effet canal étroit sur des dispositifs actifs fabriqués avec les techniques d'isolation SILO et BOX. L'importance de la diffusion latérale du bore dans la zone active est démontré, ainsi que l'avantage de la technique BOX, qui utilise un oxide déposé au lieu d'un oxide thermique dans le procédé SILO. En utilisant une dose d'implantation faible pour réduire l'effet canal étroit, un effet de double tension de seuil dans la caractéristique en faible inversion est inévitable.

Abstract - We have studied the narrow channel effect (NCE) on active devices fabricated with the SILO and BOX isolation techniques. The importance of the boron encroachment in the active region is shown, as well as the advantage of the BOX technique which uses a deposited oxide instead of a thermal oxide in the SILO process. With a low field implantation dose used to reduce the NCE, a double threshold effect in the subthreshold characteristic is inevitable.

I. Introduction

Recently several new isolation technologies have been proposed as alternatives to the conventional LOCOS process for VLSI application. These new technological processes are characterized by very different fabrication steps and various profiles of the isolation oxide which influence drastically the electrical behaviour of MOS devices.

The present paper is devoted to the study of the narrow channel effect (NCE) on active devices,which originates from the merging of the active device region with the field oxide region and results in an increase of the threshold voltage as the channel width decreases. This effect is very dependent on the isolation technology. In this study, active devices were fabricated using the SILO technique /1,2/ and a modified BOX technique /3,4/. The SILO process uses a thermal field oxide, contrary to the BOX process which employs a deposited oxide. This difference is essential whith respect to the NCE.

II. Process description

The starting material was <100> p-type silicon wafers with a resistivity of 18 Ωcm. In the SILO process, the isolation mask was a sandwich structure of LPCVD Si3N4 (10 nm)/CVD SiO2 (30nm) / LPCVD Si3N4 (150nm). After the etching of the three layers, a KOH solution was used to etch the silicon in the case of the fully-recessed structures. For the NMOS devices, a boron field implantation was performed at 25 keV with different doses. The field oxide was grown in a wet ambient at 950°C for 180 min up to a thickness of 500 nm. In the BOX process, the field oxide was performed from a CVD oxide layer by using a polysilicon counter-mask and a wet etching of the oxide /3/. Fig.1 shows SEM cross-sections of fully recessed SILO and BOX field oxides.

III. Results and discussion

In Fig.2, we compare the NCE of enhancement NMOS transistors fabricated with semi-recessed (SR) or fully-recessed (FR) field oxide, and different field implantation doses (1E13-2E13-4E13 at/cm2). The variations are chiefly due to the lateral diffusion of the channel-stop implantation into the active region. Indeed, the NCE increases when the lateral boron diffusion increases, i.e. with a semi-recessed field oxide or with a high field implantation dose.

Fig.3 shows the NCE of enhancement NMOS transistors fabricated with the BOX process and with different field implantation doses.

[1]supported by the Institut pour l'encouragement de la Recherche Scientifique dans l'Industrie et l'Agriculture - IRSIA.

Fig.1. SEM cross-sections of fully recessed SILO and BOX field oxides.

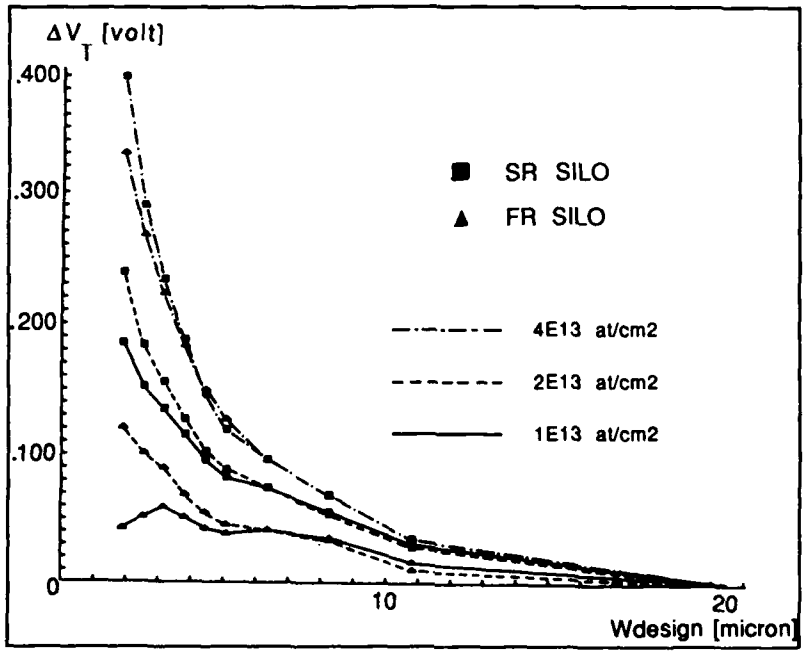


Fig.2. NCE of enhancement NMOS transistors fabricated with different SILO processes.

The comparison between the BOX and SILO process is illustrated in Fig.4 for NMOS transistors without channel implantation. The field implantation dose was 2E12 and 2E13 at/cm2 respectively, and in both cases the field transistor threshold was the same. The better behaviour of the device isolated with BOX is explained by the absence of thermal field oxidation which permits to use a lower field implantation dose and also to reduce the boron encroachment into the active region.

As shown in Figs.2 and 3, a way to reduce the NCE is to lower the channel-stop implantation. In this case, however a double threshold effect, caused by a parasitic sidewall transistor at the edges of the active region, can appear in the subthreshold characteristic of MOS devices isolated with a fully-recessed SILO (Fig.5) or a BOX field oxide (FIG.6). This effect is avoided by using a semi-recessed structure with the SILO process, or a higher field implantation dose with both isolation techniques. In each cases, the increased boron concentration at the isolated edges of the transistors raises the parasitic sidewall transistor threshold. As shown in Fig.7, the double threshold effect is undesirable because :

1. if the channel width is large, it introduces excess leakage current when the gate bias is below the threshold voltage.
2. if the channel width is narrow, it may cause a drop in the threshold voltage, the so-called inverse narrow channel effect.

We have found another way to reduce the NCE by using a field pre-implantation oxide, which has a more important thickness on the sloped walls of the silicon grooves etched in a KOH solution. This reduces the implanted boron dose on the edges of active region and so the boron encroachment. Fig.8 shows the influence of this oxide for a SILO structure.

The NCE has also been measured on PMOS transistors fabricated with the SILO process. The difference between the threshold voltage of transistors with a width of 2 microns and 20 microns is 170 mV. In this case, there is no field implantation, but the accumulation of phosphorus impurities below the field oxide during the thermal oxidation acts as the lateral boron diffusion in NMOS transistors.

IV. Conclusion

With the SILO and BOX isolation techniques, the NCE of NMOS transistors proceeds from the boron encroachment in the active region. Nevertheless, the BOX process, which uses a deposited field oxide instead of a thermal oxide, creates a lower NCE. The reduction of the NCE by using a low channel-stop implantation produces an undesirable double threshold effect in the subthreshold characterisric. But we have shown the capability of a field pre-implantation oxide to reduce the NCE of NMOS transistors fabricated with a fully-recessed structure. In the case of PMOS transistors fabricated with the SILO process and without field implantation, the NCE is due to the accumulation of phosphorus implantation below the field oxide.

References

[1] J.L. Coppée, F. Van de Wiele, Solid-State Electronics, Vol.31, n°5, pp. 887-991, 1988.

[2] J. Hui,T.Y. Chiu , S. Wong , W.G. Oldham, IEDM Technical Digest,pp. 220-223, Dec. 1982.

[3] E. Figueras, J.L. Coppée, F. Van de Wiele, Proceedings of the 17th European Solid State Device Research Conference (ESSDERC), 1987, pp 473-476.

[4] T. Shibata, R. Nakayama, K. Kurosawa, S. Onga, M. Konaka, H. Iizuka, IEEE IEDM Tech. Dig., paper 2.3,1983

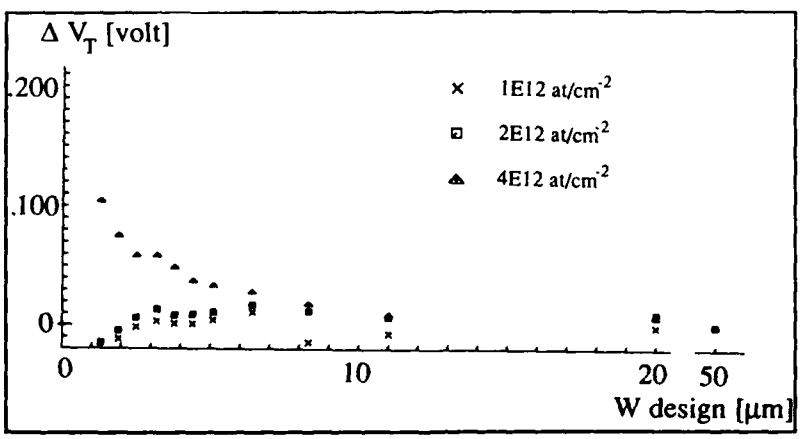


Fig.3. NCE of enhancement NMOS transistors fabricated with different BOX processes.

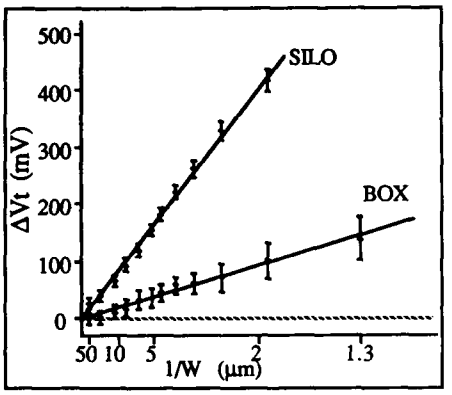


Fig.4. NCE of NMOS transistors without channel implantation.

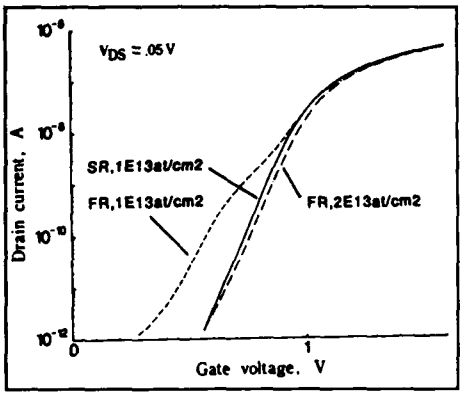


Fig.5. Subthreshold characteristics of enhancement transistors fabricated with different SILO processes.

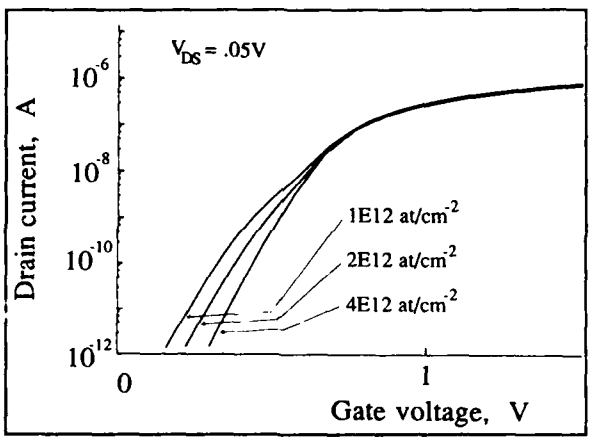


Fig.6. Subthreshold characteristics of enhancement transistors fabricated with different BOX processes.

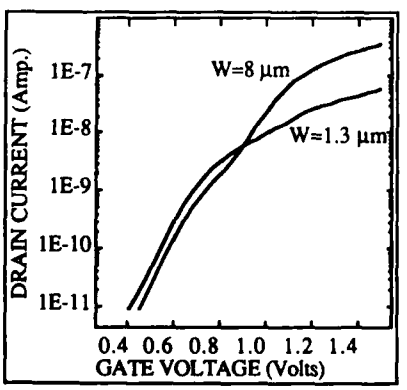


Fig.7. Dependence of the channel width on the subthreshold characteristics.

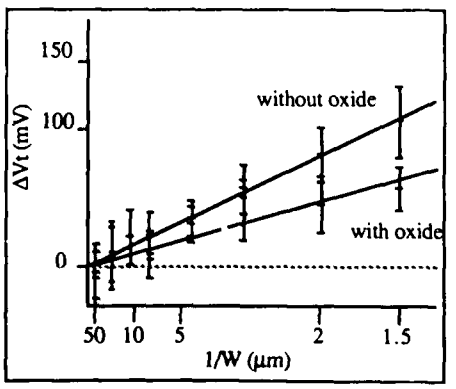


Fig.8. Influence of the pre-implantation oxide on the NCE of enhancement NMOS transistors.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

HIGH VOLTAGE SWITCHING P-CHANNEL STRUCTURE FOR CMOS ARCHITECTURES

C. MORI, G. CRISENZA and P. PICCO

SGS-Thomson Microelectronics, Via C. Olivetti 2, I-20041 Agrate (MI), Italy

RÉSUMÉ

On presents ici l'analyse d'une structure à canal p à haut voltage complétement compatible avec toutes les architectures CMOS. Elle met en évidence le rôle critique joué par la couche du "drain extension" sur la prestation électrique du dispositif et par l'épaisseur de l'oxyde de champ périphérique sur la fiabilité du dispositif. Ces résultat sont utiles pour dessiner des dispositifs à haut voltage avec un procés compatible dans une technologie CMOS avancée.

ABSTRACT

An analysis of a P-ch High Voltage structure fully compatible with all CMOS architectures is presented. It points out the critical role played by drain extension layer on electrical device performance, and by the field thick oxide periphery on the device reliability. These results are useful in designing HV process compatible devices in advanced CMOS technology.

1 Introduction

The electronic evolution towards submicron dimensions finds an important constraint in the compatibility with existing systems. Many of these must be supplied with voltages greater than the maximum voltages allowed for scaled circuits, but must often be directly interfaced to them. Hence, the problem of designing and characterizing High Voltage MOS devices compatible with submicron technologies must find an increasing interest [1,2].
Luckily, the general trend from NMOS to CMOS technologies introduces some degrees of freedom to build-up new, not standard structures, which may be useful for high voltages.
But unluckily, the well known technological tricks to produce low-threshold MOS devices weaken the structure with respect to HV effects, and especially with respect to hot electron effects.
In this paper we examine a particular, self-aligned P-ch structure compatible with all CMOS processes. We consider its evolution on different process generations, and after technological and electrical simulations, we identify the critical points and some design criteria for such a structure in the advanced processes.

2 P-channel structure

Three main effects limit the HV perfomances in a MOS structure: the thickness of the gate oxide over the drain junction, the drain dopant concentration and the curvature of the drain junction.
The P-ch structure shown in Fig.1 is an optimization of these three points with respect to an usual P-ch structure. The oxide over the drain junction is the field oxide, the thickest one available in the process. The transistor drain extension is made by P-isolation implant, at least two order of magnitude less doped than normal P^+ drain. The junction curvature is softed by the long diffusion necessary for field oxide (LOCOS) growth.
The interest of this structure resides in the fact that it is the only structure compatible with both P- and N-well CMOS technologies, since it is based on P-isolation, i.e. the P-type channel stopper necessary in both types of architectures. In the shown section an N-well technology has been chosen, but the only difference with P-well technology is that in the latter the transistor is embedded in the N-type substrate instead of in the N-well.

In any case, the P-iso layer determines the transistor performance. To improve it, the P-iso sheet resistance ought to be reduced by increasing its doping density. This induces a (process compatible) increase in N-ch parasitic threshold, but might equally induce an excessive lowering of N^+/P-iso junction breakdown on N-ch devices [3].
At the first glance, the device optimization looks to be simply to find the best compromise of opposite requirements. But other structural parameters must be taken into account.

3 Experimental

The P-channel structure has been simulated by using the program *TITAN* [4] and performed on silicon in two different processes, according to the scheme shown in Tab.1, to verify the suitability to be scaled down in different process generations and to verify the full compatibility with the assigned process architecture (no extra mask, ion implant or thermal treatments). Both these processes are CMOS: Process 1 is P-well based and its typical gate length is 3μm; Process 2 is N-well based and its typical gate length is 1.5μm.
In Fig.2 a comparison is made between the output characteristics of HV P-channel structure and the output characteristics of a standard P-ch transistor of same length achieved by the same process. But since reliability is a major concern when strong electric fields are involved, a main parameter to be carefully taken into account is the drain/channel overlap [5,6], i.e. the encroachment between the P-iso and the thin oxide region, simulated in Fig.3 according to the process schemes given in Tab.1.
Fig.4 shows the substrate parasitic currents for transistors produced by Process 1 and Process 2. It is worthwhile to note that the bell-shaped curve is associated with the structure that in Fig.3 results well overlapped (Process 1), while the diverging current is related to the not overlapped structure (Process 2).
After considering these currents, it is straightforward to consider the degradation of the Process 2 transistor. The Fig.5 shows that no degradation occurs on the not overlapped transistor.

4 Discussion

In the general trend of CMOS processes towards thinner oxides and tighter design rules, the described experimental data show the critical role played in the discussed HV device not only by P-iso layer, but also by the *LOCOS* bird's beak shape.
P-iso sheet resistance is critical in determining the transistor gain and current drive capability [7]. Its effects are apparent in the output curves of Fig.2: the decrease of current drive capability is heavier for Process 2 (10 kΩ/$\square$ P-iso layer) than for Process 1 (3 kΩ/$\square$ P-iso layer), while the device reliability more strongly feels the shape of LOCOS beak.
From the beak height/width ratio F the drain/channel encroachment strongly depends. For the considered Process 1 and Process 2 HV transistors, the different Nitride/Pad oxide ratio, and the different field oxide growth conditions, produce different LOCOS beak shapes. The different shape produces the different P-iso/channel overlap of Fig.3, and this produces the different substrate currents shown in Fig.4.
The unexpected and important effect found in this work, i.e. the absence of degradation on the non overlapped transistor, is just explained by considering the LOCOS beak shape and the underlying region. If the beak is very steep as in Process 2 (F=3, in contrast to Process 1, where the shape is very soft (F=1)), an almost vertical gap exists between the channel region and the lateral P-iso diffusion. This gap is not controlled by the gate voltage (then substrate current diverges, Fig.4). the maximum electric field is not close to the thin gate oxide but it is deeper in the semiconductor. This inhibits any charge injection into the gate oxide, then inhibiting device degradation too.
The correctness of this interpretation is confirmed by considering the Fig.6, that shows the degradation occurring on a normal P-ch transistor achieved by Process 2, after a stress comparable to that performed on the HV transistor. For this normal P-ch transistor, the gate oxide thickness is the same as for the HV transistor, body concentration is the same, and there is evidence of a

positive drain/channel overlap (see Fig.7), but a 15% degradation occurs. The reason is that the maximum electric field is close to gate oxide, so that charge injection may occur.

5 Conclusions

A HV P-ch structure has been shown and discussed on different CMOS processes. Its reliability has been related to the local physical characteristics of and under the LOCOS periphery. This structural element is of increasing importance in the general process trend towards thinner oxides, for which reliability is a major concern.

REFERENCES

1. Z.Parpia,C.A.T.Salama,R.Hadaway, IEEE Trans.El.Dev.,31,2335(1987)
2. V.Rummenik,D.L.Heald, IEDM 1982, 77
3. S.Ochi,T.Okabe,I.Yoshida,K.Yamaguchi,M.Nagata, IEEE Trans.El.Dev., 27,399 (1980)
4. A.Gerodolle,S.Martin, TITAN 4 2D PROCESS SIMULATION, CNET GRENOBLE (1987)
5. P.K.Ko,T.Y.Chan,A.T.Wu,C.Hu, IEDM 1986,292
6. T.Mizuno,Y.Matsumoto,S.Sawada,S.Shinozaki,O.Ozawa, IEDM 1985,250
7. J.G.Mena,C.A.T.Salama, Sol.St.El.,29,647(1986)

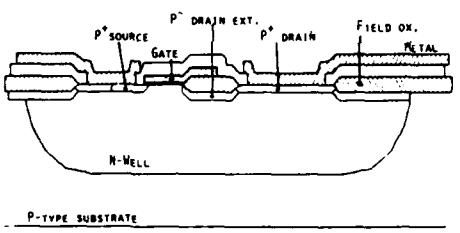


Figure 1: HV switching P-ch structure for N-well CMOS architecture.

Tab 1

	Process 1	Process 2
Transistor body concentration	8 10^{15} cm $^{-3}$	1 6 10^{16} cm $^{-3}$
Nitride thickness	3000 Å	1900Å
Pad oxide thickness	900 Å	100Å
P-iso ion dose	7 10^{13} cm $^{-2}$	2 10^{13} cm $^{-2}$
P-iso energy	40KeV	25 KeV
Field oxide growth	1000°C 260'wet	1000°C 160'wet
Field oxide thickness	9500 Å	7000Å
Gate oxide growth	875°C 48'wet	800 C 100'wet
Field oxide thickness	700 Å	330Å
Threshold shift dose	1 10^{11} B cm $^{-2}$	5 10^{11} BF_2 cm $^{-2}$
Threshold shift energy	40KeV	60 KeV

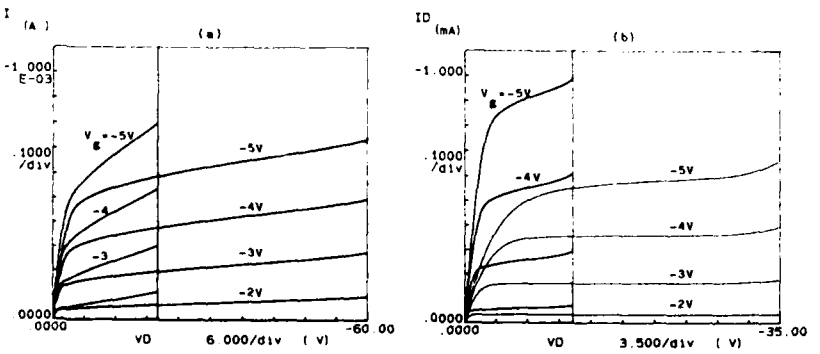


Figure 2: Output curves for Process 1 (a) and Process 2 (b) HV transistors.

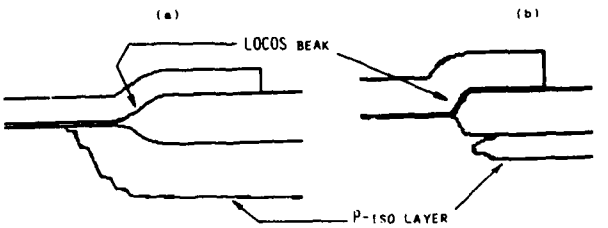


Figure 3: Drain-channel overlap in Process 1 (a) and Process 2 (b) HV transistor.
Note, in (b), the almost vertical gap between channel and P-iso layer.

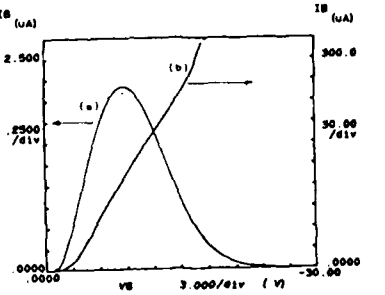


Figure 4: Substrate current vs gate voltage for Process 1 (a) and Process 2(b) HV transistors.
For (a) $L_{eff} = 5\ \mu m$
For (b) $L_{eff} = 3\ \mu m$

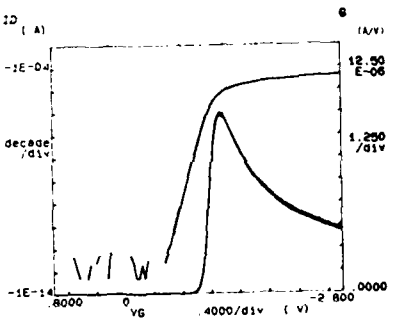


Figure 5. Process 2 HV transistor.
Comparison of I_D current (also in subthreshold region) and transconductance before and after voltage stress: *no difference is appreciable.*

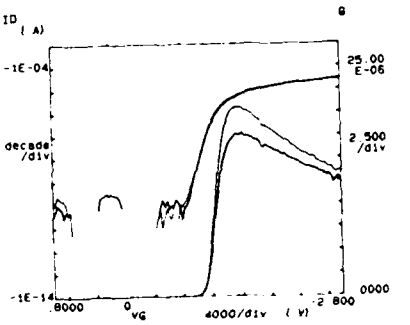


Figure 6. Process 2 std not HV transistor.
Comparison of I_D current (also in subthreshold region) and transconductance before and after a voltage stress comparable to that performed on H.V. transistor.
The gain (G) degradation is apparent.

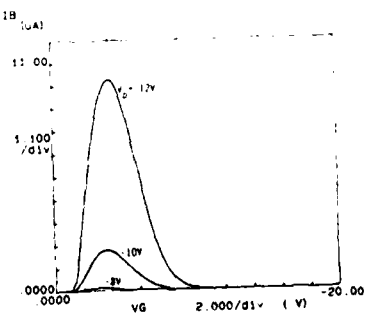


Figure 7. Substrate current of Process 2 std not HV transistor

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

INFLUENCE OF HYDROGEN RELATED DEFECTS ON THE Q_{ot}, D_{it} BUILD-UP DUE TO STRESS AND ANNEALING IN THE MOS SYSTEM

F. WULF and D. BRÄUNIG

Hahn-Meitner-Institute Berlin GmbH, Department of Data Processing and Electronics, Glienicker Str. 100, D-1000 Berlin 39, F.R.G.

Résumé: La génération des charges d'oxyde Qot et des états d'interface Dit par "bias temperature stress" (BT-stress) et par irradiation ionisante ainsi que le comportement de rétablissement après le stress ont été étudiés. Le processus de rétablissement après des stress respectifs ainsi que le processus de génération lors du BT-stress est décrit par la réaction chimique de l'hydrogène dans le Sio2. Le modèle présenté permet d'établir une corrélation entre le BT stress et la sensibilité à l'irradiation. En outre, ce modèle explique le rétablissement inverse, l'effet de l'intensité de la dose et le comportement de rétablissement des Dit en relation à la température.

Abstract: The build-up of oxide charges, Q_{ot}, and interface states, D_{it}, during stress -irradiation or bias temperatures (BT)- as well as the annealing of these defects has been investigated. The post-stress annealing and the degradation process due to BT-stress can be explained by the chemical reaction of the hydrogen in the SiO_2. With this new model it becomes possible to describe the correlation of irradiation and BT-stress and moreover the sensitivity of MOS-devices against carrier injection as well as reverse annealing, dose rate effects and the temperature dependent D_{it}-annealing.

1 - INTRODUCTION

During the last two decades, countless work has been devoted to understand the build-up and annealing of oxide charges and interface states in the SiO_2/Si-system. Based on these studies, we found that the hydrogen related defects, Si-H and SiOH, which represents extrinic defects of donor and acceptor character resp., are the main source for the charge trapping process. This process causes bond breaking and therefore a chemical reaction with the SiO_2-network. The new model describes the build-up of Q_{ot} and D_{it} due to BT-stress by a reaction and diffusion controlled process. The annealing behaviour after BT-stress and irradiation can be explained by the same mechanisms. It can be shown that the Si-H defect is responsible for the generation of the positive charged Q_{ot} and D_{it} in the case of BT-stress and irradiation. This result makes it possible to correlate the BT-stress and irradiation sensitivity of a given MOS-technology.

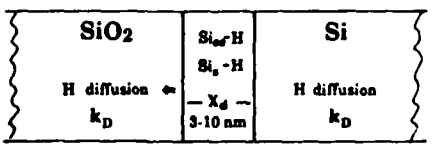
2 - EXPERIMENTAL

The experiments are performed on a variety of MOS devices:
MOS structures (varactor, transistor), CMOS (commercial and radiation hardened) and HCMOS. The following stress modes are applied:
1) BT-stress at temperatures between 100 and 200°C and applied biases of both polarities ($|E_{ox}| \leq 5$ MV/cm). During heat-up and cool-down the respective bias is maintained.
2) Irradiation at room-temperature and applied biases of both polarities up to an accumulated dose of 60 krad (Si). Radiation sources are CO-60 Source, X-ray-tube ($E_{max} = 150$ KeV) and 2.5 MeV-electron-van de Graaff accelerator.
Devices from the same lot are used for both stress modes to facilitate a comparison. The separation of Q_{ot} and D_{it} from C(V) or $I_D(V_{GS})$- characteristics, which is essential to the analysis of the degradation process is accomplished by taking the midgap voltage shift for Q_{ot} - determination. At this particular value it is assumed that the interface charge is neutral and any shift reflects the action of Q_{ot} alone (1-4). The interface state density is then obtained from low-high frequency C(V) measurements [5-7] or just from high frequency C(V) measurements [8] in the case of varactors. For transistors the Q_{ot} are deduced from the subthreshold voltage shift V_{Th} [9], whereas the D_{it} are evaluated from the slope of the I_D (V_{GS}) characteristic in the subthreshold region [10].

3 - GENERATION OF OXIDE CHARGE AND INTERFACE STATES

3.1- BIAS-TEMPERATURE-STRESS

The polarity of the oxide field, E_{ox}, is crucial for the effect of the BT-stress. Negative oxide field (BT⁻-stress) generates Q_{ot} and D_{it} and a positive field (BT⁺-stress) anneals these active defects. There are many emperical [11-18] and physical [19-25] models which describe the degradation mechanisms, but up to now a coherent understanding has not been developed. Fig. 1 shows the schematic structure of the MOS-system, with the small region X_d in which the concentration of the Si-H defect is higher than in the bulk of the SiO_2 [26].



Si_S: trivalent Si-atom bounded to the Si-network at the SiO_2/Si-interface
Si_{OS}: trivalent Si-atom bounded to three oxygen atom in the area of X_d

Fig.1 Schematic structure of the SiO_2/Si-System

Under the influence of the elevated temperature and the electric field, the Si_S-H bond is broken and the released hydrogen diffuses out of the strained region near the interface. Simultaneously the presence of the hydrogen causes a chemical

reaction with the SiO_2-network and the result can be described by the right-hand side of eq. 1:

$$\underbrace{\equiv Si_s - H}_{\text{defect}} + \equiv Si_{os} - O - Si_{os} \equiv \xrightarrow[E_{ox}<0]{T} \underbrace{\equiv Si_s \cdot}_{\text{Dit}} + \underbrace{\equiv Si_{os}^{+}}_{\text{Qot}} + \equiv Si_{os} - OH + e\uparrow^{Si} \qquad (1)[27]$$

The trivalent $\equiv Si_S \cdot$ acts as an interface state.

The energetic level in the Si band gap depends on the local structure of the defect [28]. The released hydrogen reacts with the SiO_2-network and a breakage of the Si-O-bond gives a positive charged Si-atom. Eq. 2 illustrates some possible reactions of the hydrogen with the SiO_2-network.

a)	Si-H + H	⇆ Si· + H_2	d)	Si-OH	⇆ Si-O· + H
b)	H_2	⇆ 2H	e)	Si-O-Si + H	⇆ Si^+ + Si-OH + e
c)	Si-H	⇆ Si· + H	f)	Si-OH + H-Si	⇆ Si-O-Si + H_2

(2)[27,29]

Expressed as a set of coupled differential equations, the eq. 2 can only be solved by numerial calculation. This very complex process, can be approximated by a reaction of pseudo-nth-order.

$$\frac{dH}{dt} = -k_H \cdot H^{n+1} \quad 1 \le n+1 \le 3 \qquad (3)$$

H	:	concentration of hydrogen
H_0	:	initial concentration of hydrogen
k_H	:	reaction rate
n + 1	:	order of the reaction

With the initial conditions $H = H_0$ at $t = 0$ the following solution is obtained:

$$H(t) = H_0 (1 + n\, k_H H_0^n t)^{\frac{-1}{n}} \qquad (4)$$

The diffusion of the released hydrogen can be described by a simplified expression given by Wait [30].

$$k_D = 4\pi\, r_0 \cdot D_0 \cdot \exp\left(-\frac{E_{aD}}{kT}\right) \qquad \text{with: } |r_{Si} - r_H| > r_0 \qquad (5)[30]$$

E_{aD} is the activation energy of the diffusion process, D_0 and r_0 are the diffusion coefficient and reaction radius, respectively.

Therefore the generation of electrically activated defects N_{ot} is described by the following equation

$$\frac{d\Delta N_{ot}}{dt} = (N_{Def} - \Delta N_{ot}) \cdot k_D \cdot H(t) \qquad (6)$$

N_{Def}	:	initial concentration of defects
ΔN_{ot}	:	change of oxide charge density
k_D	:	diffusibility of hydrogen
H(t)	:	hydrogen concentration as a function of time

The solution of eq. 6 using eq. 4 and 5 is

$$\Delta N = N_{Def} \cdot \left\{1 - \exp\left[\frac{k_D \cdot H_0^{1-n}}{k_H \cdot (n-1)} \left(1 - \left(1 + n \cdot k_H \cdot H_0^n \cdot t\right)^{\frac{n-1}{n}}\right)\right]\right\} \qquad (7)$$

Fig. 2 shows the increase of N_{ot} as a function of time for static fields of ± 5MV/cm and different stress temperatures. The experimental data are fitted by eq. 7 (solid line). The change of N_{ot} in the case of positive field ($E_{0X} > 0$) is more than three orders of magnitude less than for $E_{0X} < 0$.

Eq. 7 explains the degradation process in the SiO_2 caused by BT-stress and combines the emperical description by Reynolds [11] and the analytical approach by Jeppson et. al.[25]. If we plot ΔN_{ot} as a function of the mth-root of the time (Fig. 3, with m = n/n-1) we get a nearly straight line. This explains why other authors [11,14,16,18] find that ΔN_{ot} is proportional to $t^{1/m}$.

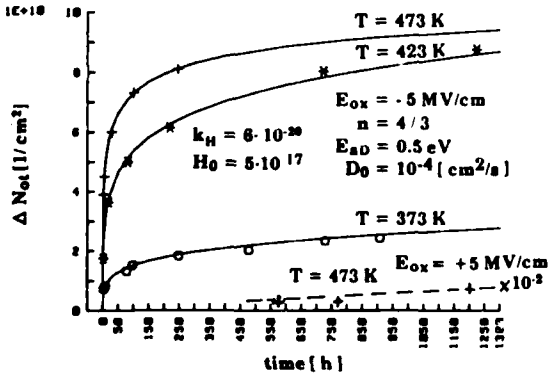


Fig.2 Generation of ΔN_{ot} after BT-stress vs time with electric field and temperature as a parameter

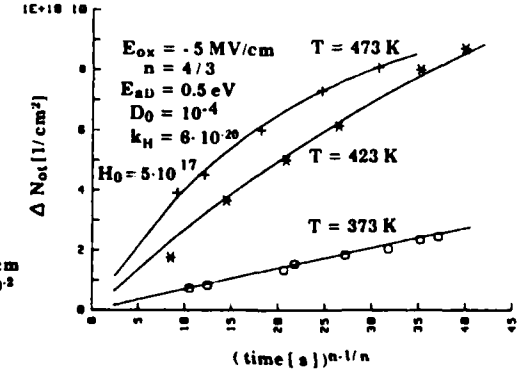


Fig. 3 Generation of ΔN_{ot} after BT-stress as a function of $t(s)^{n-1/n}$

3.2 - IRRADIATION STRESS

As shown in fig. 4, the amount of generated oxide charges is directly proportional to the amount of accumulated dose and can be described by eq. 8.

$$Q_{ot} = q \cdot K \cdot f_H \cdot d_{ox} \cdot A \cdot D \qquad [25,26] \qquad (8)$$

With:

- q = elementary charge
- K = Yield of electron/hole pairs per cm² and rad (Si) in SiO_2
- f_H = fractional yield. Ratio of seperated to generated carriers
- d_{ox} = oxide thickness
- A = hole trapping factor. Ratio of captured to separated holes
- D = accumulated dose in rad (Si)

The generation of Q_{ot} is strongly correlated to the generation of D_{it}, as verified in fig. 5.

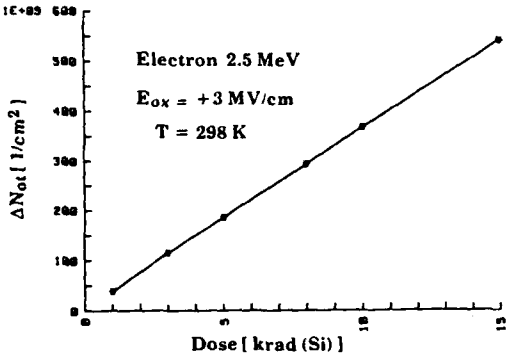


Fig. 4 Generated oxide states vs. accumulated dose.

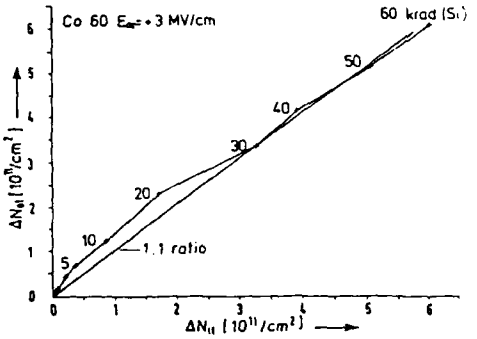


Fig. 5 Oxide state vs interface state change for irradiation at room temperature.

Revesz [27] has shown that the presence of the irradiation induced hole is sufficient to brake the Si-H bond. The released hydrogen acts in the same manner as described above for the BT-stress.

$$\equiv Si_s - H + \equiv Si_{os} - O - Si_{os} \equiv + h \rightarrow \equiv Si_s \cdot + \equiv Si_{os}^{+} + \equiv Si_{os} - OH \qquad (9)$$

3.3 - HOT ELECTRON STRESS

After outlining the interdependence of BT- and irradiation stress, a possible link to the hot electron problem is presented. The hydrogen in the SiO_2 forms a Si - H bond as well as a Si - OH bond. This latter defect acts as an acceptor and becomes negatively charged via an electron capture process according to eq. (10).

$$Si - OH + e \rightarrow Si - O^{-} + H \qquad (10)$$

The ratio of Si - H and Si - OH bonds in the SiO_2 depends on the oxidation process parameters [26]. By this dependency, the effect of irradiation, BT-stress and the hot electron sensitivity is interrelated to the process optimization [33].

4 - ANNEALING

Besides other possible annealing mechanisms after stress [34] (thermal or optical detrapping e.g.) we focus on chemical detrapping. The oxide charges and the interface states can be neutralized by hydrogen. After stress there is an enhanced defect concentration in the region X_d and the corresponding capture of hydrogen in this region causes a diffusion gradient of the released hydrogen from the bulk of the SiO_2 (Si) towards the interface.

The process of free hydrogen generation can be similarly described as outlined by eq. 2-5. Eq. 11 shows the chemical reactions of the hydrogen at the SiO_2/Si- interface

$$\equiv Si_{os}^{+} + H \rightarrow \equiv Si_{os} - H + h\uparrow_{Si} \qquad \equiv Si_s \cdot + H \rightarrow \equiv Si_{os} - H \qquad (11)$$

Only under the influence of a positive oxide field the trapped hole can move into the silicon substrate. This explains the field dependence of the annealing process.

$$\frac{d N_{ot}}{d t} = - N_{ot} \cdot k_D \cdot H(t) \qquad (12)$$

The annealing of N_{ot} can be described by eq. 12. With the initial condition $N_{ot} = N0_{ot}$ for $t = 0$ the solution of eq. 12 using eq. 4 and 5 is:

$$N_{ot}(t) = N0_{ot} \cdot exp\left| \frac{k_D \cdot H_0^{1-n}}{k_H \cdot (n-1)} \left(1 - (n \cdot k_H \cdot H_0^{n} \cdot t + 1)^{\frac{n-1}{n}}\right)\right| \qquad (13)$$

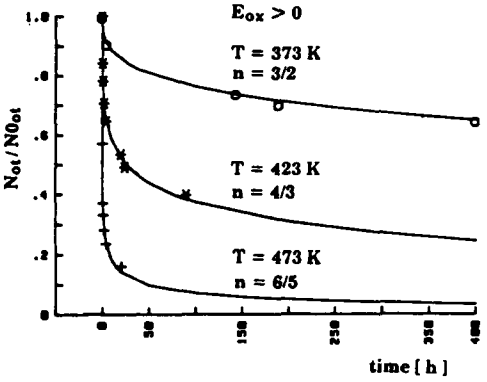


Fig. 6 Decrease of N_{ot} during annealing

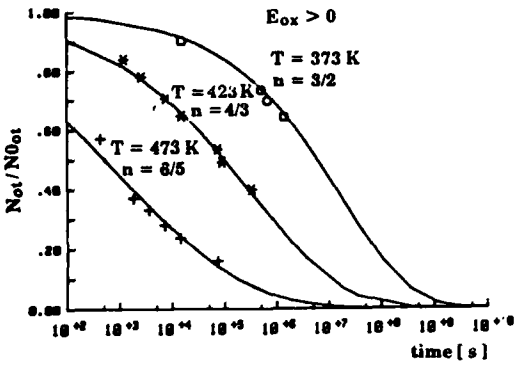


Fig. 7 Annealing behaviour: Comparison of measured data with model over a large time scale

5 - CONCLUSION

The presented unified model of charge build up and annealing takes into account that the same defects in the SiO_2 are responsible for the degradation of the MOS-system in both BT-stress and irradation. The generation of Q_{ot} and D_{it} is explained by the thermal or hole activated breakage of the ≡Si-H bond. The time dependence of the degredation and annealing process is determined by reaction and diffusion controlled mechanisms of hydrogen in the SiO_2. The influence of temperature and field polarity can be explained quantitativly for the BT-stress and can be correlated to the irradiation stress. The hot electron sensitivity, the delayed generation of the D_{it} after irradiation and the dose rate effects can also be described by this model. In addition, the feasibility of using irradiation as a time accelerating diagnostic tool for long term stability prediction has been shown [35].

6 - ACKNOWLEDGEMENT

For helpful discussions and performing measurements thanks are due to Dipl. Ings. H. Lippold, B. Schwarz, K.H. Helling. The preparation of MOS structures and part of the irradiation has been done by K.J. Tiedt and I. Siegel, resp.. Thanks to them, too.

REFERENCES

[1] Lai S.K. Young D.R; INFOS 81, Springer Verlag, Berlin, Heidelberg, New York 1981, p 118
[2] Ma T.P. Scoggan G. and Leone R.; Appl. Phys. Lett. 27, 12, 1975, p 61
[3] Hughes G. W. J.; Appl. Phys.48 , 12 , 1977 , p 5357
[4] Knoll M.; Dissertation, TU-Berlin, FB 19, 1983
[5] Kuhn M.; Solid State Elec. 13 , 1970 , p 873
[6] Berglund C.N.; IEEE Trans. on Electron. Dev.13 , 1968 , p 701
[7] Castagné R. Vapaille A.; Measurements Surface, Sci.28 , 1971 , p 157
[8] Terman L.M.; Solid State Elec. 5, 1962, p. 285
[9] van Overstraeten R.J.; Declerck G.J., Muls P.A. IEEE Trans. Elec. Dev., ED-22 , 5 , 1975, p. 283
[10] Winokur P.S. Schwank J.R., IEEE Trans. on Nucl. Sci.,NS-3 , 6 , 1984 , p 1453
[11] Reynolds F.H.; Proc. 9th Ann Rel, Phys. Symp. 1971, p 46
[12] Reynolds F.H.; Proc. IEEE 119, 12, 1972, p 1683
[13] Reynolds F.H.; Proc. 11th IEEE Rel. Phy. Symp. 1973, p 170
[14] Sinha A.K. and Smith T.E.; J. Electrochem. Soc. 125, 5, 1978, p 743
[15] Sinha A.K.; Levinstein H.J., Adda L.P. Fuls E.N. and Povilouis E.I.,;Solid State Electronics 21, 1978, p 531
[16] Noboru S. and Tokehisia Y.; Japanese Journal of Applied Physics 18,6,1979, p 1087
[17] Shiono N,Hashimoto C.;IEEE Trans. Electron Devices ED-29,3,1982, p 361
[18] Haller G.,Knoll M.,Bräunig D.,Wulf F.,Fahrner W.R.; J. Appl. Physics 56, 1984, p 1844
[19] Hofstein S.R.; Solid-State Electron 1010, 1967, p 657
[20] Breed D.J.; Solid-State Electron 17,1974, p 1229
[21] Breed D.J.;Appl. Phys. Lett. 26, 3, 1975, p 116
[22] Fischer F. and Fellinger J.; Siemens Forsch- und Entwickl.-Ber. 5, 1976, p 25
[23] Hubner K.; Phys. of Solid Surfaces 1984, p 115
[24] Hübner K.; Akademie Verlag Berlin 1984
[25] Jeppson K.O. and Svensson C.M.; J. Appl. Phys. 48. 5. 1977., p 2004
[26] Beckmann K.H. and Korrick N.J.; J.Electrochem. Soc. Solid State Science 118, 1971, p 614
[27] Revesz. A.G.; IEEE Trans. Nuc. Science NS-18,6, 1971, p113
[28] Sakurai T. and Sugano T.; J. Appl. Phys. 52, 4, 1981, p 2889
[29] Revesz A.G.; IEEE Trans. Nuc. Sci., NS-24161, 1977, p 2102
[30] Waite T.R.; Phys. Review Vol. 107, 2, 1957, p 471
[31] Freeman R. and Holmes-Siedle A.; IEEE Trans. on Nucl. Sci NS-25, 1978, p 1216
[32] Weinberg Z.A.,Johnson W.C. and Lampert M.A.; J. Appl. Phys. 47, 1, 1976, p 248
[33] Mc Brayer I.D. Fleetwood D.M., Pastorek R.A., Jones R.V.; IEEE Trans. Nucl. Sci.NS-32 , 1985 , p 3935
[34] Di Maria D.J.; The physics of SiO_2 and its interface Yorktown Heights, New York , 1978, p 160
[35] Wulf F.,Bräunig D., Nickel W.; Proc. 26th, Reliability Physics, 1988, p 150-157

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A NEW PUNCHTHROUGH MODEL FOR SHORT CHANNEL MOSFET

D. CHEN and Z. LI

Institute of Microelectronics, Tsinghua University, Beijing, China

Abstract -After the source-drain punchthrough was considered detailly, a new set of analytical models for punchthrough voltage V_P were suggested, which are suitable to NMOSFET with uniformly doped substrate or channel ion-implanted substrate and PMOSFET with buried channel, as well as a high speed numerical simulation method was developed for autosearching of V_P. Excellent agreements were shown between the results of numerical simulation and analytical models.

The source-drain punchthrough is one of the main limit factors to the short channel MOSFET. Based on a rigorous physical consideration of this effect, a new analytical model and a high speed numerical simulation method for the punchthrough voltage V_P were suggested.

During punchthrough the drain current mostly flows along a minimum potential path (for NMOS) in the bulk silicon. The control factor of this current is the potential Ψ_P at the ridge point (as shown in Fig.1), corresponding to the electron barrier height on the current path. At this point, we have $\partial\Psi/\partial z = \partial\Psi/\partial y = 0$, and the current can be expressed as $I_P = I_O exp(q\Psi/kT)$. For short channel devices Ψ_P is mainly dependent on drain voltage V_{DS} due to DIBL, therefore we can define V_P as the value of V_{DS} at which

$$\Psi_P = \Psi_{pt} = 2\Phi_F - n\frac{kT}{q}$$

where n is determined by the ratio of defined punchthrough current value I_{pt} and gate controlled channel current I_{ch}. For example, when

$$\frac{I_{ch}}{I_{pt}} = 10^3, \qquad n = 8.$$

By this definition, an autosearching function of V_P (through autosearching of Ψ_{pt}) was implemented in the 3- dimensional device simulation program TDMOS[1] with a simple algorithm. The

excellent agreement of the calculated values with the experimental results[2] are shown in Fig.2 and Fig.3. Generally no more than 100 seconds CPU time is required for each searching in TDMOS running on VAX11/750.

On the basis of above physical consideration, a set of analytical expressions for V_P of MOSFET in different cases (NMOS with uniformly doped substrate or channel ion-implanted substrate and PMOS with buried-channel) had also been derived out. The derivation is based on the solution of Poisson's equation by quasi-two dimensional way similarly to that had been done for V_T by Toyabe *et al*[3], but with different approximated $\Psi(y)$ in different cases.

Following are the expressions of V_P obtained for different cases:

(a) NMOSFET with uniformly doped substrate:

$$V_P = \frac{C_1(\Psi_P + D/A)^2}{\Psi_{bi} + D/A} e^{\sqrt{A}L/C_2} - \Psi_{bi} - D/A$$

where

$$A = \frac{1}{W^2}\frac{4 - 6K_m}{K_m(1-K_m)^2} = \frac{K_A}{W^2}, \qquad K_m = \frac{2}{3 + \sqrt{9 - \frac{12V_{BS}}{\Psi_P - V_{BS}}}}$$

$$D = \frac{q}{\epsilon_s}N_b - \frac{1}{W^2}(K_A - \frac{4(1-K_m)}{K_m})V_{BS}$$

L is the effective length of channel, V_{BS} is the substrate bias, N_b is the substrate doping level, Ψ_{bi} is the built-in voltage which is determined by

$$\Psi_{bi} = \frac{kT}{q} ln(\frac{N_b N_S}{n_i^2})$$

where N_S is the donor concentration in the surface of source region. n_i is the intrinsic concentration, W is the effective thickness of the depleted layer which can be expressed approximately by

$$W = x_j + \frac{\epsilon_s}{\epsilon_{ox}} t_{ox} + \sqrt{\frac{2\epsilon_s}{qN_b}(\Psi_{bi} - V_{BS})}$$

x_j is the source and drain diffusion depth, t_{ox} the gate oxide thickness, ϵ_s and ϵ_{ox} are permittivities of silicon and oxide respectively.

The constants are $C_1 = 1.5$, $C_2 = 1.8$.

(b) NMOSFET with ion-implanted channel region:

$$V_P = \frac{C_3(\Psi_P + D/A)^2}{\Psi_{bi} + D/A} e^{\sqrt{A}L_{eff}/C_4} - \Psi_{bi} - D/A$$

where A and D are expressed by the same formulas as that in the case (a), but W is determined by

$$W = x_j + \sqrt{\frac{2\epsilon_s}{qN_b}(\Psi_{bi} - V_{BS})} - d$$

the constants $C_3 = 0.48$, $C_4 = 1.7$.

According to the fact that most of the punchthrough current flows in the depleted layer (where impurity concentration is lower) far from the surface, the effective length of current path L_{eff} is used instead of the gate effective length L. From a simple geometrical consideration, we have

$$L_{eff} = L + 2(\sqrt{(d + y_m)^2 + x_j^2} - x_j)$$

where d is the thickness of channel ion-implanted layer, y_m is expressed by $y_m = K_m W$.

(c) PMOSFET with buried-channel:

$$V_P = \frac{C_5(\Psi_P + D/A)^2}{\Psi_{bi} - D/A} e^{\sqrt{A}L/C_6} - \Psi_{bi} + D/A$$

where

$$A = \frac{2(2t'_{ox} + W_d + d)}{W_d d(2t'_{ox} + y_m)}, \qquad D = \frac{q}{\epsilon_s}(N_A - N_D) - \frac{2V_{BS}}{dW_d}$$

N_A and N_D are acceptor and donor concentration respectively, the effective oxide thickness

$$t'_{ox} = \frac{\epsilon_s}{\epsilon_{ox}} t_{ox}$$

d is the thickness of the surface acceptor layer, y_m is the depth of the point of potential maximum which is determined by

$$y_m = \frac{N_A d - N_D W_d}{N_A - N_D}$$

the thickness of depleted layer $W_d = \sqrt{(N_D t'^2_{ox} - 2V_{BS} + N_A d(2t'_{ox} + d))/N_D} - t'_{ox}$

the constants $C_5 = 1.7$, $C_6 = 1.8$.

Fig.4 and Fig.5 are examples of agreement between the numerical results and the analytical results for case of channel ion-implanted NMOSFET and taking the given parameters as variables. The similar excellent agreement is observed in all other cases.

References:

[1] D.T.Chen *et al*, Math. and Comp. in Simulation, vol.29, 461 (1987).

[2] F.C.Hsu *et al*, IEEE Trans. Electron Dev., ED-30, 1354 (1983).

[3] T.Toyabe *et al*, IEEE Trans. Electron Dev., ED-26, 453 (1979).

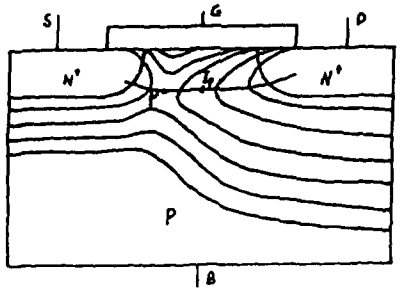


Fig.1. 2-D distribution of potential at punchthrough

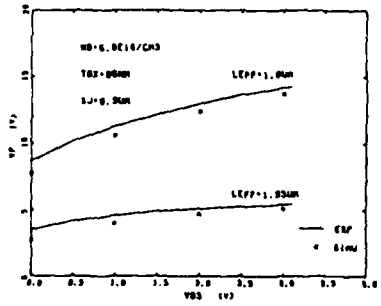


Fig.2. the relationship of V_P and V_{bs}

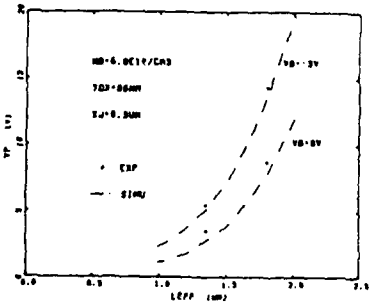


Fig.3. the relationship of V_P and L_{eff}

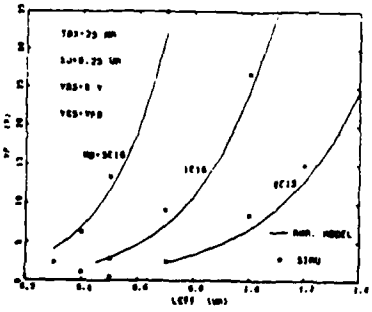


Fig.4. V_p vs L_{eff} and N_b for ion implanted NMOS

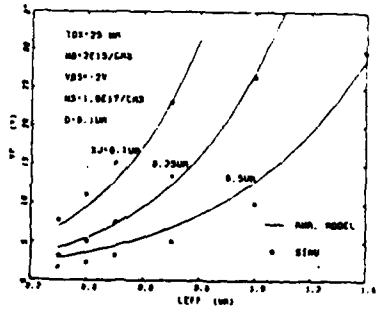


Fig.5. V_P vs L_{eff} and X_j for ion implanted NMOS

Session 7B Room B : Hot carriers

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

HOT CARRIER SENSITIVITY OF MOSFET's EXPOSED TO SYNCHROTRON-LIGHT

G. PRZYREMBEL, R. MAHNKOPF and H.G. WAGEMANN

Institut für Werkstoffe der Elektrotechnik der Technischen Universität Berlin, Jebensstr. 1, D-1000 Berlin 12, F.R.G.

Abstract - The influence of synchrotron-light irradiation for p- and n-channel MOSFET's on their sensitivity to hot carrier degradation was investigated. The radiation induces additional interface states and a positive oxide charge. Annealing at 450°C reduces the interface state density to its initial value but not the oxide charge. A hot carrier stress can compensate this remaining charge by trapping electrons. This effect produces an enhanced shift of the threshold voltage compared to non-irradiated devices. After compensating all of the charge due to the irradiation the devices have a degradation behavior comparable to the non irradiated ones.

1 - INTRODUCTION

X-ray lithography is expected to become important for the realization of MOSFET's with submicron channel lengths. X-rays however are damaging the oxide by the generation of interface states and positive oxide charges. An complete annealing is possible only at elevated temperatures. Yet after the metallization of the wafer with aluminum the highest applicable temperature is 450°C. Using X-ray lithography for structuring the aluminum, an annealing of the irradiation at a subsequent temperature step of less than 450°C is often incomplete. The amount of the remaining positive oxide charge proves different for various technologies.

A highly detrimental behavior of short channel MOSFET's is the effect of degradation by hot carriers, which has been proven to be enhanced for devices exposed to ionizing radiation. This correlation was recently shown for n-channel MOSFET's, which were irradiated by a 10 KeV-X-ray source /1/.

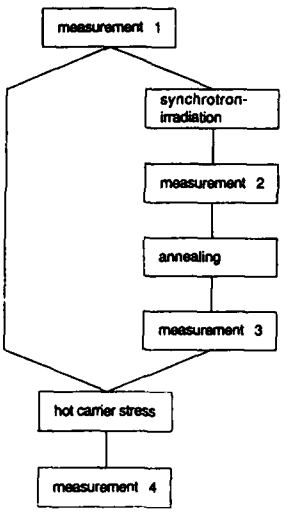
In our experiments the degradation of hot carrier stress for p- and n-channel MOSFET's exposed to synchrotron-light was investigated. The wavelengths of the radiation facility at the BESSY laboratory IMT in Berlin are ranging between 0.3 and 1.5 nm. The temperature at the subsequent annealing process was 450°C, the highest permitted value for a fully processed wafer. The stress behavior of these MOSFET's was compared to devices, which were not subjected to the radiation and annealing procedure.

2 - DEVICE PREPARATION AND MEASUREMENTS

An overview of the whole experimental procedure is given in Fig. 1. The test devices are available on a 4"-wafer. The interface trap density was determined by the charge pumping technique. The single steps of irradiation, annealing and hot carrier stress are described in the next sections:

1. Irradiation

The conditions at the radiation process were chosen to be comparable to a usual exposure step for X-ray resists (e.g. power, time, geometry).



MOSFET devices:

poly-gate, d_{ox} = 38 nm
p-channel: L_{eff} ≈ 2.2 µm, W = 500 µm
n-channel: L_{eff} ≈ 1.4 µm, W = 100 µm

measurement:

V_T threshold voltage (transfer characteristic)
D_{it} interface traps (charge pumping)

radiation procedure:

synchrotron-light (BESSY)
dose = 0.37 J/cm²
(power ~ 1.5 mW/cm², time ~ 240 sec)

annealing:

450°C, 15 min, N_2, H_2 atmosphere

hot carrier stress:

p-channel: V_G = - 2.5 V, I_D = - 40 mA
n-channel: V_G = + 4.0 V, I_D = + 40 mA

Fig. 1: Flowchart and parameter of the experimental procedure

In Fig. 2 the transfer characteristics of the p- and n-channel devices before and after irradiation with synchrotron-light are given. The shift of the threshold voltage is the same for both p- and n-channel devices. The increase of the interface state density can be deduced from the change of the subthreshold slope.

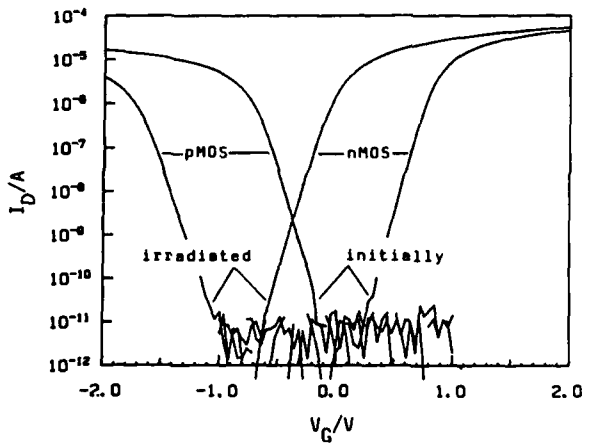


Fig. 2: Transfer characteristics of p- and n-channel MOSFET's before and after irradiation with synchrotron light

ii. Annealing
The annealing of the radiation damage was carried out at 450°C in N_2 ambient for 15 min. The density of the interface traps reaches the initial value. For a lot of the devices the positive oxide charge was not completely reduced.

iii. Hot carrier stress
The stress condition for the p-channel MOSFET was V_G = - 2.5V at a drain current of I_D = - 40mA. The n-channel device was stressed at V_G = + 4V and I_D = 40mA. The duration of the stress was 1 min for both MOSFET-types.

3 - RESULTS AND DISCUSSION

In table I the values for the threshold-voltages V_T and interface trap densities D_{it} are listed.

n	V_T	X	0.8	0	0.6	0.75
		0	0.8			0.81
	D_{it}	X	$7 \cdot 10^9$	$5 \cdot 10^{10}$	$7 \cdot 10^9$	$1.4\ 10^{10}$
		0	$7 \cdot 10^9$			$1.1\ 10^{10}$
p	V_T	X	- 1	- 1.8	- 1.25	- 1.05
		0	- 1			- 0.99
	D_{it}	X	$2 \cdot 10^{10}$	$1.35\ 10^{11}$	$2 \cdot 10^{10}$	$1.75\ 10^{11}$
		0	$2 \cdot 10^{10}$			$1.12\ 10^{11}$
			1	2	3	4
			1 initially 2 after irradiation 3 after annealing 4 after hot-carrier stress			
		X	procedure with radiation step			
		0	procedure without radiation step			
	V_T/V , $D_{it}/eV^{-1}\ cm^{-2}$					

Table I: Average values of threshold voltage V_T and interface traps density D_{it} after different treatment (referring to one wafer each).

Fig. 3 schematically displays the course of the threshold voltage V_T during the experimental procedure. The lower curve of both devices applies to hot-carrier-stress with preceding irradiation/annealing treatment, the upper ones apply to hot carrier stress only.

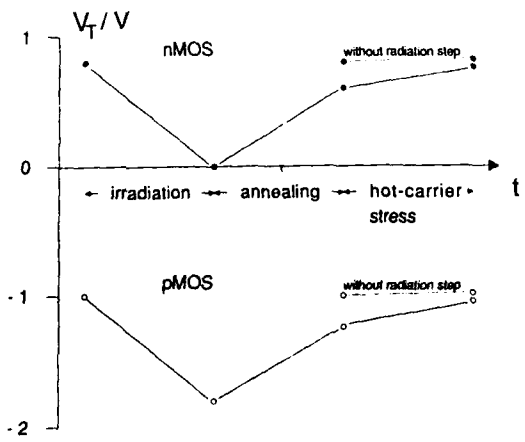


Fig. 3: Threshold voltage after different treatment

After annealing the threshold voltage tends towards the preradiation value. The amount of this change depends on the time and temperature of the annealing step. Often there is no complete reduction of the positive oxide charge producing a threshold voltage more negative than the initial value. A subsequent hot carrier stress gives rise to a neutralisation of the residual charge. Therefore the threshold voltage shift is enhanced within the initial phase of stressing. Because of the different threshold voltages at the beginning of the stress the effective gate voltage was adjusted according to equivalent hot electron stress conditions.

Fig. 4 displays the shift $\Delta V_{T\ stress}$ after a hot electron stress of 1 min as a function of $\Delta V_{T\ anneal}$, which is the difference between the threshold voltage after annealing and the pre-irradiation value and an indication for the quality of the annealing process.

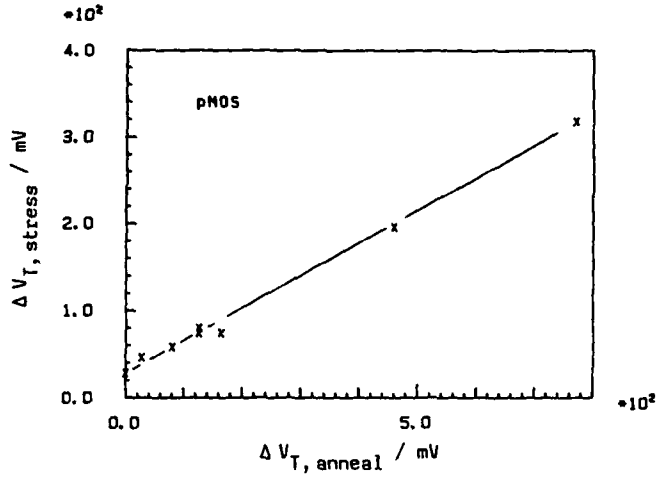


Fig. 4:
Threshold voltage shift $\Delta V_{T,\,stress}$ after a hot electron stress versus the difference $\Delta V_{T,\,anneal}$ between the threshold voltage after annealing and the pre-irradiation value. $\Delta V_{T,\,anneal}$ results from the remaining positive charge after irradiation and annealing.

From Fig. 4, it is apparent that devices with a higher amount of residual positive charge have a higher shift in threshold voltage for the same stress conditions. After neutralisation of this charge the MOSFET will show a stress behavior equal to a non-irradiated MOSFET. This fact can be taken from the time dependence of threshold voltage during hot electron stress given in Fig. 5.

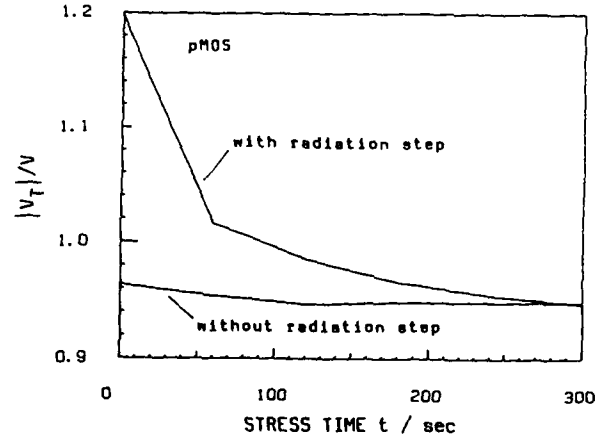


Fig. 5: Threshold voltage as a function of time during hot electron stress.

4 - CONCLUSION

An enhanced sensitivity to hot electron stress for devices priorly exposed to synchrotron light was observed. This effect is caused by a positive charge generated by irradiation and remaining in the device after an incomplete annealing process at 450°C. This charge will be compensated by the hot electrons injected into the oxide resulting in a significantly higher threshold shift at the starting phase of stressing. This phenomenon was observed for both p- and n-channel MOSFET's. A complete annealing procedure at temperatures applicable to fully processed wafers can overcome this enhanced stress effect.

ACKNOWLEDGEMENTS For device preparation we appreciate the aid of Dr. D. Haack, Fa. R. Bosch, Reutlingen. Furthermore we thank Dr. W. Seifert for technological support. This work has been completed with the aid of the FhG-Insitute IMT Berlin and with financial support of the DFG Bonn (Deutsche Forschungsmeinschaft).

REFERENCES

/1/ Reich, R.K., et al.: "Radiation-Dependent Hot-Carrier Effects", IEEE El.Dev.Lett., Vol. EDL-7, No. 4, April '86, pp. 235-237

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

ANNEALING OF HOT-CARRIER-INDUCED MOSFET DEGRADATION

R. MAHNKOPF, G. PRZYREMBEL and H.G. WAGEMANN

Institut für Werkstoffe der Elektrotechnik der Technischen Universität Berlin, Jebensstr. 1, D-1000 Berlin 12, F.R.G.

Abstract - The annealing of fixed oxide charge and interface states generated by hot-carrier stress is investigated in the temperature range of 100°C - 450°C. First order rate equations are given, which approximately describe two subsequent processes involved in the annealing and ending at neutralization. The related activation energies are determined. For comparison the annealing of synchrotron light induced damage is examined.

1 - INTRODUCTION

Little is known about the annealing of hot-carrier-induced damage. The analysis of the temperature dependence can lead to a better understanding of degradation mechanism in MOSFET's /1/. In this study the results of our experiments on p-channel MOSFET's will be shown. The changes of the densities of fixed oxide charge N_{ox} and interface states D_{it} due to the hot carrier stress and the annealing procedure are evaluated. In addition to the influence of the temperature (100°C - 450°C) on the annealing step the time dependence was investigated. The characteristic behavior is approximated by two subsequent first order reaction steps.

From plotting the data in an Arrhenius-diagram the activation energies for the annealing processes of the oxide charges and interface states are determined. To clarify the influence of a different prior treatment on the annealing procedure, the same temperature steps were carried out for devices irradiated with synchrotron light (at the IMT Bessy facility Berlin). The changes of N_{ox} and D_{it} were evaluated and are shown for comparison.

2 - EXPERIMENTS

The devices used were polysilicon gated MOSFET's with gate oxide of 38 nm, effective channel length of 2.2 µm and channel width of 500 µm. The MOSFET's were stressed for one minute with a gate voltage V_G = - 3 V and a drain voltage V_D = - 16 V. Before and after stress HF-capacitance (CV) and charge-pumping (CP) characteristics were measured. From this the density of the stress generated negative oxide charge $N_{ox,\,stress}$ was determined to $4.7 \cdot 10^{11} cm^{-2}$. The increase of interface state density $D_{it,\,stress}$ amounts to $1.5 \cdot 10^{10} cm^{-2} eV^{-1}$. After hot carrier stress the devices were annealed in a N_2-ambient at various temperatures ranging between 100°C and 450°C.

In contrast to the method proposed in /1/ each step of annealing temperature was performed with another device stressed at the same condition. By taking the same device over all steps a different behavior for the annealing progress was observed. After annealing HF-CV and CP-characteristics were taken again to monitor the change of N_{ox} and D_{it}. Additionally for two devices the time of thermal annealing at 200°C and 300°C was varied subsequently from 1 to 100 min.

In comparison the whole procedure was carried out with devices damaged by ionizing radiation instead of hot carrier stress. The irradiation was done by synchrotron light with a dose of 0.37 J/cm². The homogeneously distributed increase of charge densities in the channel region was determined to $N_{ox,\,stress} = 4.5 \cdot 10^{11} cm^{-2}$ and $D_{it,\,stress} = 2.5 \cdot 10^{11} cm^{-2} eV^{-1}$. It can be confirmed that in contrast to hot carrier stress the irradiation induced oxide charge has a

positive sign and the generated interface state charge is more than a decade larger.

3 - RESULTS AND DISCUSSIONS

Fig.1a shows the effect of the temperature treatment for the hot carrier stressed devices by the decrease in the density of fixed oxide charge ΔN_{ox} and the increase in interface state density ΔD_{it}.

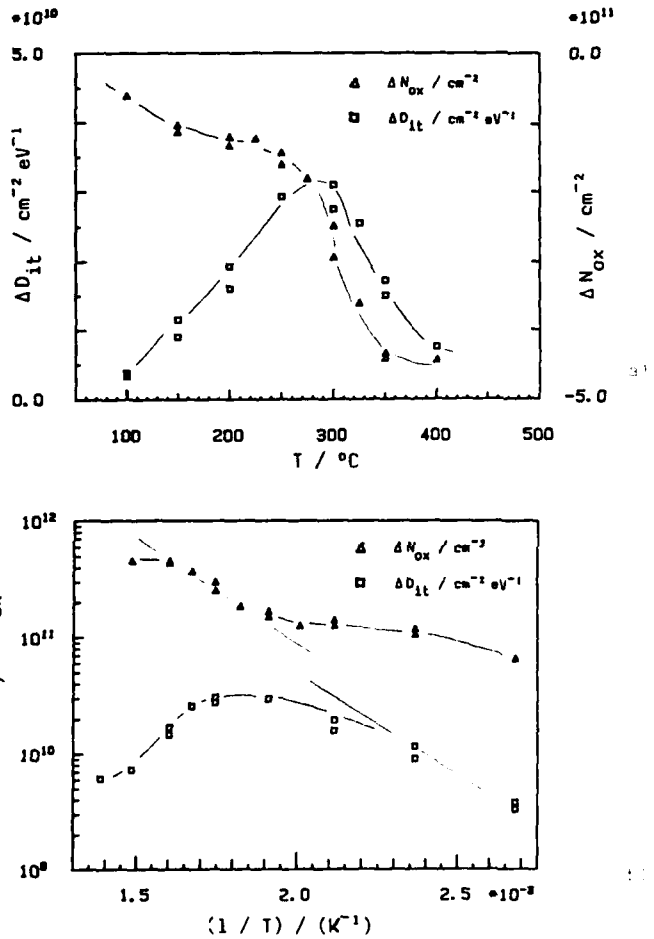


Fig. 1: Change of the oxide charge density ΔN_{ox} and interface state density ΔD_{it} as a function of annealing temperature for hot carrier stressed devices. The same data are plotted vs. T (a) and 1/T (b).

The stress induced density of fixed oxide charge decreases continuously with increasing temperature. Above 400°C no further decrease occurs because the stress induced charge is completely annealed. The interface state density however shows an increase with a maximum at 300°C. This behavior can be understood assuming only two processes:

process 1) a change from fixed oxide charges Q_{ox} ($Q_{ox} = q \cdot N_{ox}$) to interface states D_{it} followed by

process 2) an annealing of interface states D_{it}, ending at neutralized charge:

$$N_{ox} \xrightarrow{k_1(T)} D_{it} \xrightarrow{k_2(T)} \text{neutralization}$$

The densities of the charges are given by the rate equations, where $k_1(T)$ and $k_2(T)$ describe the rate coefficients:

$$-\frac{d\,N_{ox}}{dt} = k_1(T) \cdot N_{ox}$$

$$+\frac{d\,D_{it}}{dt} = k_1(T) \cdot N_{ox} - k_2(T) \cdot D_{it}$$

$$\text{with } k_{1,2} = k_{1,2\,0} \cdot \exp\left(-\frac{W_{1,2}}{kT}\right)$$

In Fig. 1b the results are shown in an Arrhenius-plot. From the slope of the decrease of fixed oxide charge density as well as from the slope of the increase of interface state density an activation energy of 0.35 eV for the transformation process 1 can be evaluated. This indicates that hydrogen may be involved in the process of annealing, which has been suggested by several authors /2,3/.

Fig. 2 shows the time dependence of annealing for the temperatures T_1 = 200°C and T_2 = 300°C. At 200°C the interface state density is still increasing with increasing time, at 300°C however the process of interface state annealing (process 2) begins to predominate the process of interface state generation from fixed charge (process 1) after [illegible] min. So the time dependence of the annealing agrees well with the proposed model.

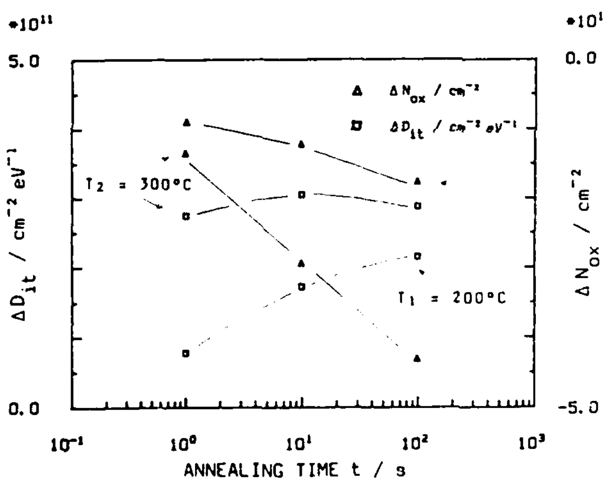


Fig. 2: Time dependence of annealing for T_1 = [illegible] and T_2 = [illegible]

In Fig. 3 the effect of the temperature treatment for the hot carrier-irradiated devices is shown. The induced densities of positive fixed oxide charge as well as interface state charge are decreasing [illegible] with increasing temperature.

Consequently both fixed oxide charges, the positive one of irradiated devices and the negative one of hot carrier stressed devices are decreasing with temperature. The interface state charge however is decreasing for irradiated and increasing for hot carrier stressed devices. This behaviour can be explained by the different relation between interface state charge and fixed oxide charge at the beginning of the temperature treatment.

The Arrhenius-plot for the irradiated devices is shown in Fig. 3b. From the slopes an activation energy of W = [illegible] eV can be evaluated for the decrease of both densities. According to that the same physical mechanism may be supposed for the annealing of damage induced by hot carriers and by [illegible] irradiation.

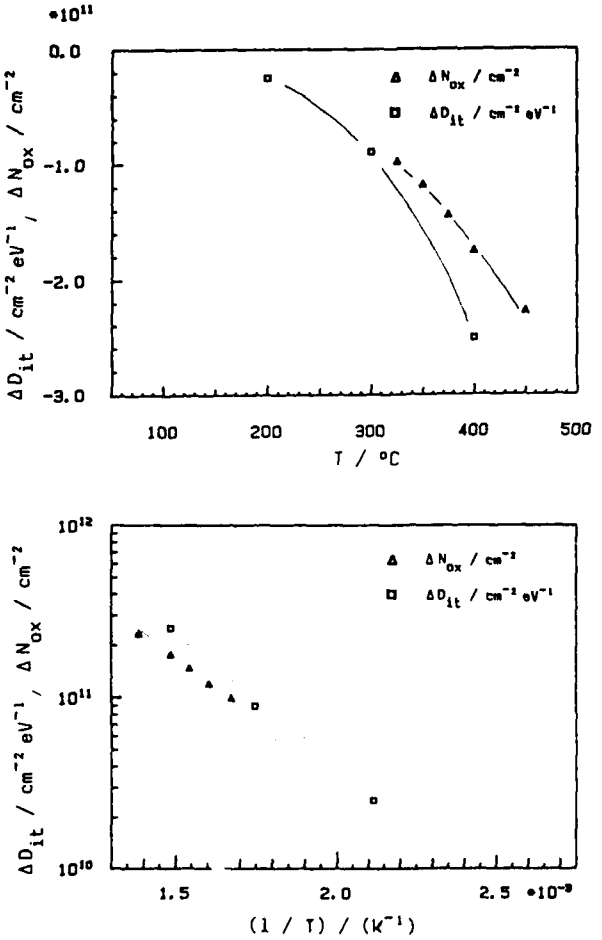


Fig. 3:
Change of the oxide charge density ΔN_{ox} and interface state density ΔD_{it} as a function of annealing temperature for irradiated devices. The same data are plotted vs. T a and 1/T b).

4 - CONCLUSION

The annealing of damage generated by hot carriers within MOSFET devices was investigated. At p-channel devices the density of negative oxide charge N_{ox} is decreasing in the whole range of temperature. On the other hand the interface state density D_{it} is increasing at temperatures below 400°C indicating a transformation process for the fixed oxide charge. This behavior can be approximated by two subsequent first order reactions. The activation energies of the generation of the intermediate product D_{it} and of the decrease of the primary product N_{ox} were determined to ~ 0.35eV. This value suggests the participation of hydrogen atoms in the annealing process. Measurements on devices treated with ionizing irradiation instead of hot carrier stress provide the same activation energies. These devices however are showing a decrease of interface state density even at low temperatures.

ACKNOWLEDGEMENT - For device preparation we appreciate the aid of Dr. D. Haack, Fa. R. Bosch, Reutlingen. Furthermore we thank Dr. W. Seifert for technological support. This work was completed with the aid of the FhG-Institute IMT Berlin and with financial support of the DFG Bonn (Deutsche Forschungsgemeinschaft).

REFERENCES

/1/ Saitoh, M., et al.: IEEE Trans. ED-34, p. 2584, 1987
/2/ Svensson, C.M.: in "The physics of SiO_2 and its interfaces", S.T. Pantelides, Ed. New York: Pergamon, 1978 pp. 328-332
/4/ Wulf, F.: Dissertation, TU Berlin, 1988

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A NEW METHOD FOR THE DETERMINATION OF THE SPATIAL DISTRIBUTION OF HOT CARRIER DAMAGE

R. MAHNKOPF, G. PRZYREMBEL and H.G. WAGEMANN

Institut für Werkstoffe der Elektrotechnik der Technischen Universität

Berlin, Jebensstr. 1, D-1000 Berlin 12, F.R.G.

Abstract - A new method is presented for the experimental evaluation of the spatial distribution of interface states and fixed oxide charges generated by hot carriers. This method is fully based on HF-capacitance (CV) and charge-pumping (CP) measurements. A general discussion of the method is followed by an application to MOSFET's with different channel lengths stressed with various bias conditions.

1 - INTRODUCTION

Hot carrier damage in MOSFET's is inhomogenously distributed over the channel region. Maximum damage has been observed at the drain side, where the hot carrier injection is predominant due to the high field strength /1-5/. In this work a method is presented, which allows the determination of the spatial distribution of hot carrier generated interface states as well as fixed oxide charge. HF-capacitance (CV) and charge-pumping (CP) characteristics taken before and after stress are sufficient for achieving all information.

For the determination of the local position in the channel our method neither needs the one-dimensional equation for the depletion width /1-3/ nor a 2-dimensional simulation program (e.g. MINIMOS) /4,5/. The first one does not truly describe the surface conditions, the second one depends on the evaluation of several model parameters.

The next section will give a detailed explanation of the measurement and evaluation procedure. Then the results of systematically varied stress conditions will be described.

2 - EVALUATION METHOD

The MOSFET's under test are p-MOS and n-MOS devices with a poly-gate and an oxide thickness of 38 nm. The channel width W is 100 µm, the channel length L_{eff} 1.1 µm or 3.0 µm for p-MOS and 2.5 µm for n-MOS devices respectively.

1) p-MOS

First the procedure for the determination of the spatial distribution of hot carrier induced damage is demonstrated. A p-MOSFET with an effective channel length of 3.0 µm is stressed for 1 min with a gate voltage V_G = - 4 V and a drain voltage V_D = - 16 V. Under this condition avalanche hot electrons are injected into the oxide.

HF-CV characteristics of the gate-substrate capacitance are measured before and after stress with the source and drain terminals grounded (Fig. 1a). After stress the capacitance approaches the accumulation level asymptotically at a higher level of gate voltage than before, indicating local regions with a different density of fixed negative charge. The ratio between the capacitance before and after stress at a given gate voltage is regarded as a measure for the channel area under consideration.

The spatial distribution of the interface states is deduced from charge pumping measurements, also taken before and after hot carrier stress (Fig. 1b). The low level of the gate-pulse is kept constant while varying the high level. So after stress the regions of enhanced negative charge can be scanned with increasing the high-level of gate voltage.

For moderate high levels the pumpcurrent is lower than initially; parts of the channel area near the drain junction can't reach accumulation due to the induced negative fixed oxide charge. Then at a specific value of high level voltage the pumpcurrent I_{CP} rises sharply because of the additional interface states in the stressed region now contributing to the pump-signal. The correlation between the gate-level and the position in the channel is taken from the CV-measurements.

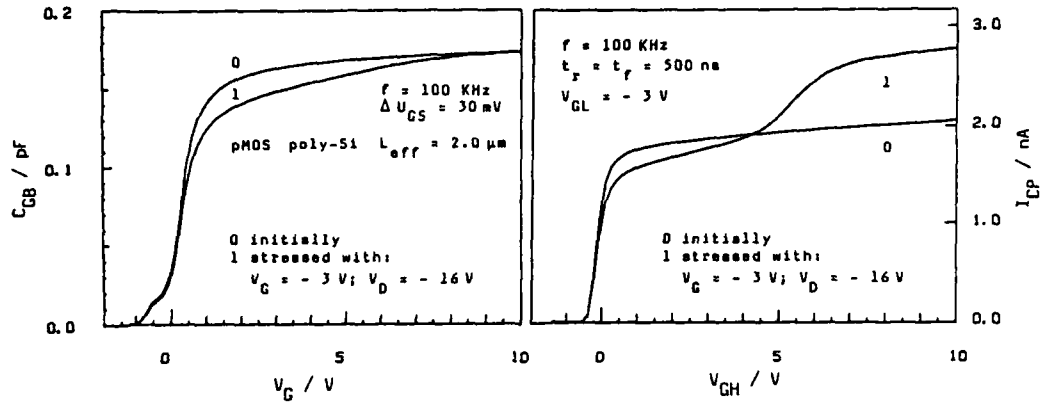


Fig. 1: Gate-substrate capacitance C_{GB} vs. gate voltage V_G (a) and charge-pumping current I_{CP} vs. gate-high-level V_{GH} (b) for a p-MOSFET before and after hot carrier stress

In Fig. 2 distributions of the generated densities for different channel lengths and different drain voltages for stress are shown. For higher drain voltages an enhanced maximum of the charge densities and an extension into the channel can be observed.

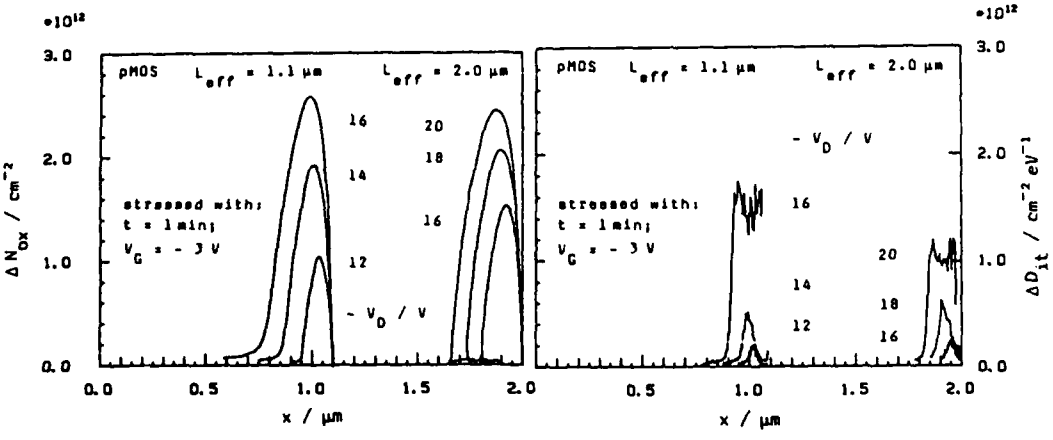


Fig. 2: Generated densities of interface states ΔD_{it} and fixed oxide charge ΔN_{ox} as a function of position after hot carrier stress for p-MOSFET's with different channel lengths and different drain stress voltages.

The stress induced interface state densities are smaller than the densities of fixed oxide charge. This relation is mostly seen at p-MOS devices and allows to extract the distribution of fixed oxide charge from CV-measurements.

The spatial distributions of the stress induced charge densities ΔN_{ox} and ΔD_{it} are given by the expressions

$$\Delta N_{ox}(x) = \frac{\Delta V_G(V_G(x)) \cdot C_{ox}}{q \cdot W \cdot L_{eff}}, \qquad \Delta D_{it}(x) = \frac{1}{f \cdot q \cdot W} \cdot \frac{\Delta I_{CP}}{\Delta x}\Big|_x$$

$$\text{with } \Delta I_{CP} = I_{after} - I_{before} \cdot \frac{C_{after}}{C_{before}}.$$

$$x(V_G) = L_{eff} \cdot \frac{C_{after}}{C_{before}}\Big|_{V_G}, \quad x = \text{border between the accumulated and the inverted channel area.}$$

In Fig. 3 the total amount of the stress-induced oxide charge ΔQ_{ox} is shown as a function of the injected charge Q_G flowing across the oxide during stress. It can be seen that for the same amount of charge Q_G an equal build-up of trapped charge ΔQ_{ox} exists for both MOSFET's with different channel lengths (see Fig. 2a: L_{eff} = 1.1 µm, V_D = - 16 V ↔ L_{eff} = 2.0 µm, V_D = - 20 V).

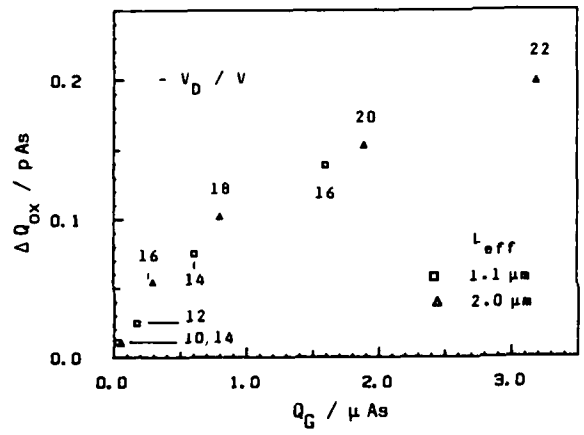


Fig. 3:
Stress induced oxide charge ΔQ_{ox} as a function of the injected charge Q_G for different channel lengths and different drain stress voltages

11) n-MOS

In Fig. 4 the measured characteristics for n-MOSFET's with an effective channel length of 2.5 µm before and after hot carrier stress are shown. Two stress conditions with different gate voltages were chosen. It has been verified that for the gate voltage of 1.5 V hot holes, and for the gate voltage of 8.0 V hot electrons are predominantly injected into the oxide.

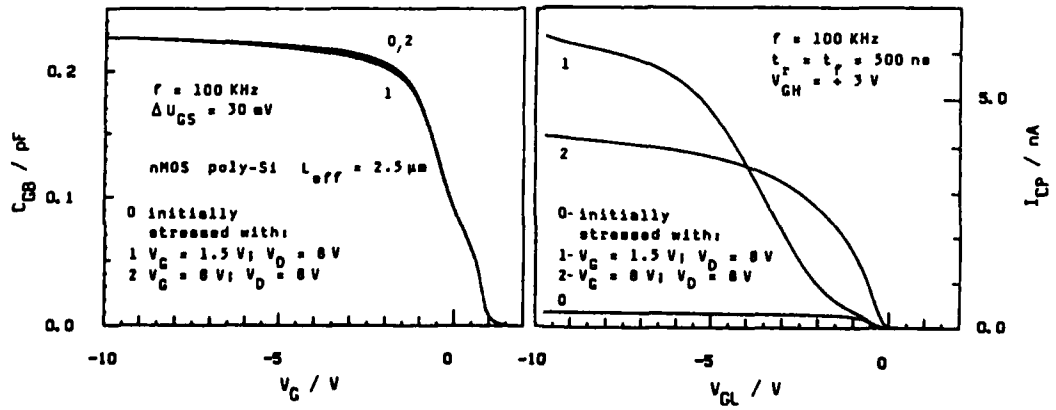


Fig. 4: Gate-substrate capacitance C_{GB} vs. gate voltage V_G (a) and charge-pumping current I_{CP} vs. gate-low-level V_{GL} (b) for a n-MOSFET before and after hot carrier stress

From the HF-CV-characteristics a build-up of positive oxide charge for the stress at lower gate voltage can be seen. The stronger electron injection at higher gate voltage causes no significant change of the characteristic.

From the CP-measurements the positive oxide charge generated at 1.5 V is indicated by a continuous increase of the pumpcurrent over a range of nearly 6 V. For the stress carried out at V_G = 8 V the CP-characteristics show no remarkable oxide charge. This is in accordance with the statement drawn from the CV-measurement and can be explained by a low efficiency of the damage process induced by injected electrons. This contrasts to the behavior of the p-MOS devices. Furthermore it can be seen from the CP characteristics that the generation of interface states is more pronounced for stressing by the lower gate voltage.

In Fig. 5 the spatial distribution of generated oxide charge for a n-MOSFET stressed at V_G = 1.5 V is shown. The evaluation was performed with the CV-characteristics given in Fig. 4a. The induced charge is located in a very narrow region ($\lesssim$ 50 nm) at the drain end of the channel. This value is commonly assumed for the extension of damage to n-MOSFET's.

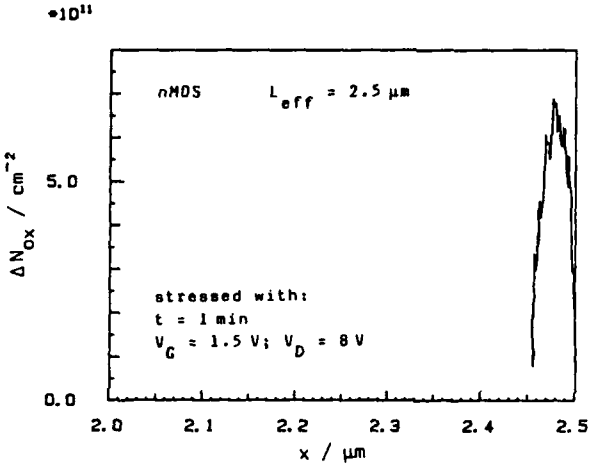


Fig. 5: Generated density of fixed oxide charge ΔN_{ox} as a function of position after hot carrier stress for a n-MOSFET.

3 - CONCLUSION

With the presented method the spatial distributions of hot carrier induced interface traps and fixed oxide charges can be determined seperately. The influence of various stress conditions on the distribution of the damage along the channel has been investigated. Examples for different channel lengths and bias conditions have been given in this work. Comparative studies of p- and n-MOS devices demonstrate the different stress behavior of both types. A further insight into the mechanism of degradation can be expected.

ACKNOWLEDGEMENTS

We wish to thank Dr. D. Haack, Fa. R. Bosch, Reutlingen, for technological support and the DFG/Deutsche Forschungsgemeinschaft Bonn as well as the FhG-Institut IMT Berlin for financial aid.

REFERENCES

/1/ Schwerin, A.; et al., IEEE Trans. Electron Devices, vol. ED-34, p. 2493 (1987)
/2/ Poorter, T.; Zoestbergen, P., IEDM Techn. Dig., p. 100 (1984)
/3/ Takeda, E.; et al., IEEE Electron Dev. Letts., vol. EDL-4, p. 329 (1983)
/4/ Maes, H.E.; et al., Electronics Letts., vol. 18, p. 372 (1982)
/5/ Schmitt, D.; Dorda, S., Electronics Letts., vol. 17, p. 761 (1981)

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

HOT CARRIER STRESS INDUCED CHANGES IN MOST TRANSCONDUCTANCE STRUCTURE

C. BERGONZONI, R. BENECCHI* and P. CAPRARA

SGS-Thomson Microelectronics, Via C. Olivetti 2,
I-20041 Agrate Brianza (MI), Italy
**Istituto di Fisica del Politecnico, p.zza Leonardo da Vinci 32,*
I-20133 Milano, Italy

ABSTRACT-*An analysis of MOS transistor hot carrier degradation is presented, based upon experimental, simulated and analytical study. The transconductance curve study is shown to be a good tool for the correct characterization of the stress induced damage, and experimentally observed transconductance degradations are reproduced by simulations and by means of a simple analytic model. The usual classification of damage in terms of threshold voltage shift and maximum transconductance degradation is shown to fail under general conditions, where structural alterations of electrical characteristics take place.*

1 Introduction

Due to the strong increase of electric fields inside short devices, hot carrier degradation is one of the most outstanding limits to transistor dimensions scaling down for VLSI. This work presents an approach to this problem devoted to a correct characterization of the effect of hot carrier aging on MOS devices, based upon the analysis of transconductance curve alterations induced by stress, to achieve the physical interpretation of the electrically measured effects. We assume in this work that electric charge be present, after stress, in some region of the gate oxide and/or of the oxide/ silicon interface, following a general agreement on this topic. Experimental results are discussed together with device 2-D simulations, and an analytic two-piece model is presented which reproduces and explains the observed transconductance degradations.

2 Experimental

The experimental measurements were performed on N-channel conventional devices:the device of figure 1 had 1.5μm gate length and 1.4μm effective channel length and the device of figure 2 had 1.2μm gate length and 0.95μm effective channel length . The characteristics of the native device were compared to the after stress curves using a HP 4145 semiconductor parameter analyser. Fig.1 reports the degradation of transconductance and drain current, each curve being taken after a stress cycle composed by 1000 sec. at Vd=9V,Vg=1V and 10 sec. at Vd=8.5V, Vg=8.5V.

[1]This work has been partially supported by ESPRIT project 554-SPECTRE

This kind of stress cycle is used in order to inject holes, thus creating electron traps in the oxide, and to fill then those traps with the second stress step [1]; it should then represent a kind of worst case for aging, also simulating a dynamic operation of the device. Fig.2 reports the degradation of transconductance and drain current, after 100 sec at Vd=8V,Vg=1V and 20 sec at Vd=8V, Vg=8V. It can be seen that in fig.1 an hump in the transconductance curve takes place after some stress cycles, modifying the usual structure of the characteristic, while in fig.2 the maximum transconductance peak is higher after the stress cycle than in the native device (overshoot). It can be also seen that these effects are not easily observed in the Id vs. Vg curves.

3 Simulations

Simulation results for aging effects on current characteristics have been discussed in [2]; the experimentally onserved transconductance degradations have been here reproduced by means of 2-D device simulations, using the Poisson and continuity solver HFIELDS [3]. The results of simulations are shown in fig.3 and 4. The simulation conditions were:

- damage extension: -100 nm; interface fixed negative charge concentration: $2 \cdot 10^{12} cm^{-2}$ (fig.3)
- damage extension: 40 nm; interface fixed negative charge concentration: $2 \cdot 10^{12} cm^{-2}$ (fig.4)

The minus sign in the first item damage extension means that charge is localized, starting from the metallurgical junction, towards the bulk of the drain junction. In the second simulation charge is localized above the channel. A qualitative agreement between experimental results and simulations can be readily seen.

4 Analytic modeling

The fundamental hypothesis of this section is that an hot carrier aged transistor can be modeled as the series of two devices: a first one, with the unchanged properties of the original transistor, and a second one, with some alterations in electrical parameters caused by the hot carrier injection. To reproduce analytically this conditions, an empirical function has been chosen, reproducing the qualitative behavior of a MOS transistor. Expression (1) shows the dependence of transconductance on gate voltage as it has been chosen in this work:

$$g(x) = \frac{A}{2}[1 + tanh(k(x - \alpha))] \tag{1}$$

The function (1) gives a good fit of linear MOST transconductance under a *constant mobility* assumption. The constant mobility approximations has been taken in order to emphasize the effects of aging of transconductance structure, avoiding to mix them with the usual transconductance decrease due to mobility degradation. The primitive function of (1) thus represents the conductance of the device in the linear region:

$$G(x) = \frac{A}{2k}[log(cosh(k(x - \alpha))) + k(x - \alpha) + log(2)] \tag{2}$$

It can be seen that both these functions depend on three parameters, α, k and A, respectively representing, as it can be easily verified, the threshold voltage, the subthreshold slope and the linear gain of the transistor. The model of the aged transistor is thus achieved by calculating the transconductance of a series of two devices, both of them being represented by expressions like (1) and (2) for g(x) and G(x), but with generally different values for α, k and A. Fig. 5 shows the result of the application of this model with different values for *all* the parameters involved in the analytic expressions, while fig. 6 reports the transconductance calculated when only A and α assume different values.

5 Discussion and conclusions

The comparison of the experimental, simulated and analytical results presented in the previous sections shows that:

- the experimental effects can be reproduced by 2-D device simulations under the simple hypothesis that trapped negative charge is present in the gate oxide or at the interface. Different charge localizations lead to structurally different effects.
- a simple two-piece analytic model reproduces again both the experimental and simulated data; in the analytic case, qualitatively different transconductance degradations are calculated, depending on the number of parameters on which the degradation is supposed to act.

The alteration of the parameters A and α in (1) and (2) gives a result consistent with the output of the 2-D simulation of a device when the damaged region is extended above the channel; this shows that the main effect of this kind of damage is the local alteration of the threshold voltage of a region of the transistor, thus leading to an aged device operating like a series of two transistors, with different lengths and thresholds; the overshoot in the transconductance curve taking place in this situation as an effect of control exchange between the two transistors of the series was discussed in [4]. The situation becomes more intricate when an additional hump takes place in the aged device transconductance; we have shown here that this effect is reproduced by device simulation if the damaged region is extended in the gate/drain overlap region, and by analytic modeling when also the subthreshold slope is supposed to be affected by aging. These results can be explained by observing that, when the trapped charge is localized in the overlap region, it exists in the channel a quickly decreasing potential, due to the charge itself, which gives a spatially variable threshold to the terminal region of the channel. This region will so show a generally different subthreshold slope, in agreement with the analytical result. An other interpretation of the subthreshold alteration can be given by the presence of aging induced fast interface states in a localized region of the channel, which could be detected e.g. with the charge pumping technique.

The results presented in the previous sections show that the analysis of transconductance is needed for the characterization of aging effects when significant structural modifications take place in the electrical characteristics of a device. As device dimensions scale down, the structural alterations induced by hot carrier aging become stronger and stronger; we have suggested a scheme of analysis that can help in the identification and classification of the observed effects in terms of damage intensity and spatial localization, together with the conventional characterization techniques. A complete hot carrier degradation picture should give the correlation between the operation conditions and the electrical device degradations; the here proposed model provides the link between the physical damage and the observed degradation, giving information about the stressed device which could be hardly deduced from the experimental or simulated analysis of current characteristics.

References

[1] M. Bourcerie, B.S.doyle, J.C.Marchetaux and A.Boudou,"The creation of electron traps by hole injection in hot carrier stressing of NMOS transistors", private communication

[2] C.Bergonzoni and B.S. Doyle,"Simulations of aging effects in MOS transistors", Proceedings of ESSDERC '87, Bologna, p.721, 1987

[3] G. Baccarani, R.Guerrieri, P. Ciampolini and M.Rudan,"HFIELDS: a highly-flexible 2-D semiconductor-device analysis program", Proceedings of the fourth International Conference on the Numerical Analysis of Semiconductor Devices and Integrated Circuits, J.J.H.Miller ed., Dublin: Boole Press, pp.3- 12, 1985

[4] H.Haddara and S.Cristoloveanu,"Two-dimensional modeling of locally damaged MOSFET's operating in the linear region",IEEE Trans. Electron Devices, vol. ED-34,p.378, 1987

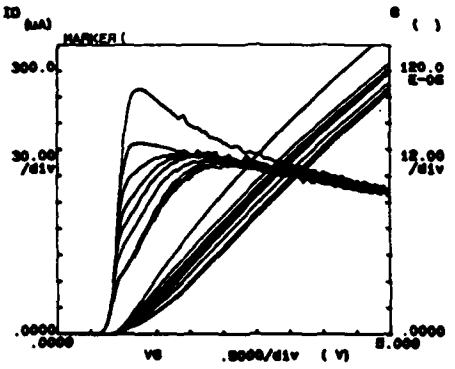


Fig.1-Experimental transconductance hump after hot carrier stress (stress condition reported in the text)

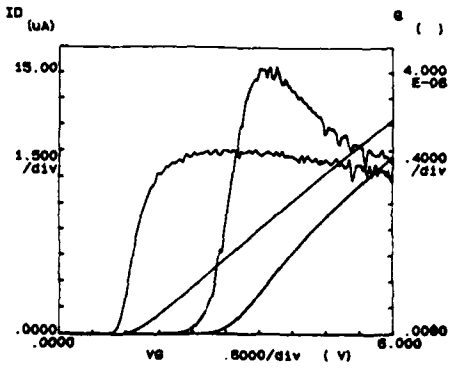


Fig.2-Experimental transconductance overshoot after hot carrier stress (stress conditions reported in the text)

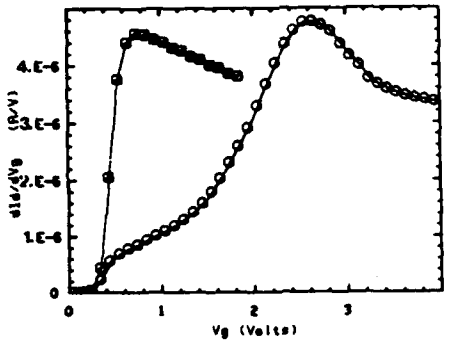


Fig.3-Simulation of transconductance hump; negative charge in the gate/drain overlap region (100nm, $2 \cdot 10^{12} cm^{-2}$)

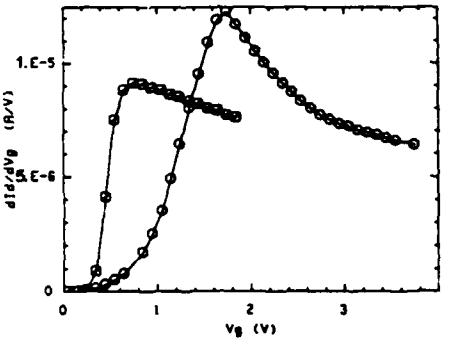


Fig.4-Simulation of transconductance overshoot; negative charge above the channel (30nm, $2 \cdot 10^{12} cm^{-2}$)

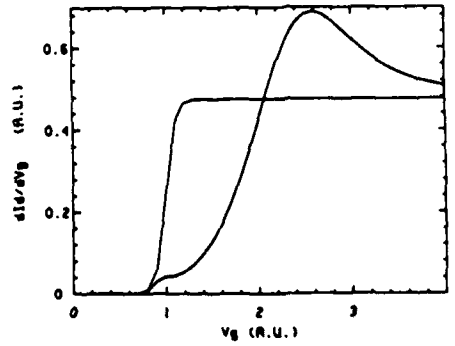


Fig.5-Analytically calculated transconductance hump; degradation of α, A and k; constant mobility

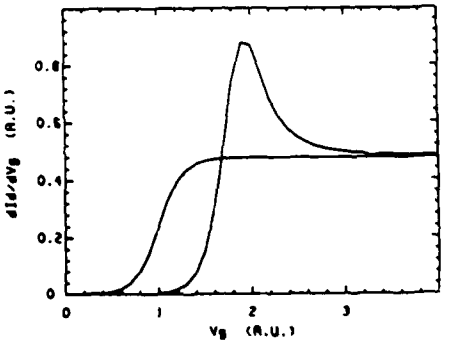


Fig.6-Analytically calculated transconductance overshoot; degradation of α and A; constant mobility

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

OXIDE DEGRADATION AND BREAKDOWN IN STRESSED MOS CAPACITORS

I. PLACENCIA, J. SUÑÉ, N. BARNIOL, E. FARRÉS and X. AYMERICH

Centro Nacional de Microelectrónica, Departamento de Física, Universidad Autónoma de Barcelona, SP-08193 Bellaterra, Barcelona, Spain

Résumé-Un nouveau model qui relatione les phénomènes de degradation et de rupture de rupture de couches minces de SiO_2 dans des dispositifs MOS a été presenté. Le point principal de ce model c'est que la vitesse de génération de pièges d'électrons est proportionelle a l'énergie dissipée dans la couche d'oxyde. Dans ce scheme, l'evolution experimentalle de la tension apliquée en expériences de stress a courant constant et la rélation quasi-linearie entre log(temps de rupture) et log(densité de courant) ont été prédites.

Abstract-A new model that relates the degradation and breakdown phenomena of SiO_2 films has been presented. The crux of the model is that the trap generation rate is proportional to the energy lost by the electrons by interaction with the SiO_2 lattice. In this framework, the experimental evolution of the applied voltage in constant-current stress experiments and the quasi-linear relationship between log(time-to-breakdown) and log(current density) have been predicted.

1 - INTRODUCTION

Intrinsic dielectric breakdown of thin SiO_2 films has been received considerable attention in the recent literature. A deeper understanding of this phenomenon is needed when dealing with reliability problems of high-quality thin oxide films suitable for VLSI devices. Nowadays, it is widely accepted that the intrinsic breakdown is an history-dependent process. Two stages have to be considered: a pre-breakdown stage during which the injection damages the SiO_2 lattice and a very fast runaway process that produces a low-resistance ohmic path between the electrodes. In this paper, we focus our attention on the pre-breakdown process and in determining the critical conditions that trigger the final runaway.

Two main models to describe the pre-breakdown physics have been presented in the past few years. The first one, developped by Harari /1/, postulates that the breakdown is due to the progressive accumulation of negative charge in the oxide bulk that increases the anode field until a critical value is reached. The second one, presented by Chen and co-workers /2/, proposes a positive feedback mechanism of local enhacement of the cathode field due to hole trapping near the cathode interface. Even though some correlations have been found between the generation of positive charge and the breakdown /3,4/, the generation of positive charge reaches most of its amount after times which are much shorter than the normal times-to-breakdown. While, the generation of electron traps does not saturate.

Hence, in this paper, we will consider that the degradation of the SiO_2 structure is reflected in the generation of new electron traps.

In the presented model, it has been assumed that a fixed degree of degradation must be achieved or, what is the same, a critical density of electron traps must be reached to trigger another mechanism that brings the oxide to breakdown after a very fast runaway process. The breakdown is always triggered locally because the creation of defects is not homogeneous over the whole structure area. For this reason, the breakdown is a random phenomenon and it has to be treated statistically. In the presented model, all the breakdown parameters (time-to-breakdown, charge-to-breakdown, defect density at breakdown, etc.) must be understood as mean values.

2 - DEGRADATION AND BREAKDOWN IN MOS STRUCTURES

In stress experiments, part of the kinetic energy of the injected electrons is dissipated by interaction with the SiO_2 lattice. Only a very small fraction of this energy is converted into the creation of deep electron traps. Some of the generated traps are filled with electrons and the resulting negative charge distribution modifies the electric field in the oxide. Thus, the stress conditions have to be changed to mantain the stress

specifications.The breakdown is triggered by the presence of a critical density of created defects, N_{bd}, generated after dissipation of a critical energy per unit of volume, $E_{bd}=N_{bd}\Delta$, where Δ is the average energy to create one electron trap. Because of the statistical nature of breakdown, N_{bd} would depend on the structure area: the smaller capacitor area, the greater N_{bd}. On the other hand, N_{bd} will depend on the oxide thickness in accordance with the thickness dependence of the mechanism that controls the final runaway.

It is possible to define a time-dependent degradation degree as $D(t)=N(t)/N_{bd}$, where N(t) is the density of generated defects until time t. The degradation, D, varies between 0, non-degradated oxide, and 1, breakdown threshold. The fraction of dissipated energy that is used to create the electron traps is defined as $dE_a=\beta\, dE$, where dE_a is the differential energy density used to create the traps. β will obviously depend on the temperature at which the stress experiment is performed and on the particular oxide structure that varies with the technological process used in the fabrication of the MOS structure.

The total energy density that is dissipated per unit of time in the SiO_2 conduction band is given by the Joule law:

$$\frac{dE(t)}{dt} = J(t)F(t) \qquad (1)$$

F(t) being the average electric field in the oxide and J(t) the flowing current density. For the high fields used in the stress experiences, the current will be given by the Fowler-Nordheim expression. To take into account the effects of the negative trapped charge, we will assume, as a first approximation, that the F-N expression remains valid if the average field, F(t), is changed by the field at the injecting electrode, F´(t), that for the case of a homogeneus distribution of trapped charge is given by:

$$F'(t) = F(t) + \frac{\rho(t)\, tox}{2\,\epsilon ox} \qquad (2)$$

being t_{ox} the insulator thickness, $\rho(t)$ the trapped charge density at time t, and ϵ_{ox} the SiO_2 permitivity constant. We have assumed that $\rho(t)=-qN(t)$ that means that all the generated defects are filled up with electrons. This fact is not strictly the case, because of the actual trapping-detrapping mechanisms /5/. *In any case, to be consistent,* one can give a different physical meaning to β and define it as the efficiency of creation of the traps that are filled in the steady state of the trapping-detrapping processes. *It has been experimentally shown* /5/ that the trap generation rate depends exponentially on the oxide field. However, having redefined β in terms of the occupied traps, to consider β independent of the applied field is a very good first approximation because the increase in the generation of traps at high fields is compensated to some extent by the easier detrapping at those fields.

In constant-current stresses, the current density must be fixed at a value $J=J_o$, and hence, the applied voltage has to be increased to compensate for the influence of the increasing number of trapped electrons /1-4/. Derivating equation (2) and since $\rho(t)=-qE_a(t)/\Delta$, it follows that:

$$\frac{d\rho(t)}{dt} = \frac{-q\beta}{\Delta}\frac{dE(t)}{dt} \qquad (3)$$

Then, using equation (1) for dE(t)/dt, with the condition F´(t)=constant (this condition is a direct consequence of the constancy of the current), the differential equation that describes the evolution of $V(t)=F(t)t_{ox}$ is obtained and its solution is:

$$V(t) = V(0)\exp\left(\frac{t}{\tau_v}\right) \qquad (4)$$

where $\tau_v = 2\epsilon_{ox}\Delta/q\beta t_{ox}J_o$, being V(0) the oxide voltage at t=0 if the oxide is initially non-degradated. The evolution of the degradation, $D(t)=E(t)/E_{bd}$, can be easily calculated substituting equation (4) in equation (1) and integrating. Then, applying the breakdown condition $D(t_{bd})=1$, one obtains:

$$t_{bd} = \tau_v \,\mathrm{Ln}\left\{1 + \frac{F_{oo}}{F(0)}\right\} \qquad (5)$$

where $F_{oo} = (t_{ox}qN_{bd})/(2\epsilon_{ox})$. The average charge injected to breakdown, Q_{bd}, is given by

the product $J_o t_{bd}$. Doing so, it can be concluded that higher currents have associated lower values of Q_{bd}.

3 - DISCUSSION AND RESULTS

In order to compare our theoretical predictions with the experimental data, we have analyzed the degradation characteristics corresponding to four VLSI PolySi-SiO_2-Si(n) samples ($t_{ox} \approx 100$ Å) taken out from the literature /1,6,7/. We have used the following procedure: fitting the data in the linear regime of the V(t) characteristic (since the condition $t_{bd} \ll \tau_v$ is always satisfied, equation (4) can be approximated by a straight line), we obtain the value of τ_v. From the value of τ_v we can directly obtain β/Δ. On the other hand, from the experimental t_{bd} values, we have calculated N_{bd} for each sample. All these results are presented in table I. Also in this table, one can see that N_{bd} is of the same order of magnitude for all the samples. The small variations can be attributed to differences in the structure areas and oxide thicknesses. In any case, the electron trap density is so high that the final runaway can be reasonably attributed to ressonant tunneling as proposed by Riccó et. al. /8/.

SAMPLE	t_{ox} (Å)	Area ($10^{-6}cm^2$)	J_o (A/cm^2)	τ_v (sec)	β/Δ ($10^{-9}eV^{-1}$)	t_{BD} (sec)	N_{BD} ($10^{18}cm^{-3}$)
S1	63	2.3	0.435	385	6.6	23	5.82
S2	100	6.4	1	303	2.3	21	4.31
S3	103	58	0.01	40400	1.7	4500	5.86
S4	94	2.3	0.435	266	6.0	47	9.5

Table I - Degradation and breakdown parameters for four MOS structures subjected to constant current stress.

In figure 1 we have plotted the predicted time-to-breakdown as a function of the applied current. To calculate t_{bd} as a function of J_o we have assumed a unique value of N_{bd} for all the samples, $N_{bd} = 6.5\ 10^{18}\ cm^{-3}$, which is approximately the average of the particular results shown in table I. The times-to-breakdown as a function of J_o have been calculated by using equation (5). As seen in figure 1, the validity of expression (5) is confirmed by the experimental results. In addition to the data that correspond to samples S1-S4 whose β/Δ values have been obtained from their respective V(t) characteristics, in figure 1 we have plotted the experimental data corresponding to two other samples (S5 and S6) of similar thickness. The theoretical predictions for these samples have not been done because we do not have their V(t) characteristics and hence we can not calculate their corresponding β/Δ values.

In the literature is has been experimentally shown that the dependence of $\log(t_{bd})$ on $\log(J_o)$ is quite well described by a straight line with slope $(d\log(t_{bd})/d\log(J_o))=-1-s$ with $s > 0$ /1,6,10,11/. This behaviour can be also observed in figure 1. Although equation (5) does not predict a linear dependence between $\log(t_{bd})$ and $\log(J_o)$, it can be demonstrated that $d\log(t_{bd})/d\log(J_o)$ is constant with only 5% of error, when the current density varies over three orders of magnitude. For the case of the analyzed samples, equation (5) predicts that the above mentioned slope is approximately -1.06. The experimental values of s are slightly higher but always positive as predicted by equation (5). Note that s=0 is equivalent to Q_{bd}=constant whereas $s > 0$ means that Q_{bd} decreases

with the applied current.

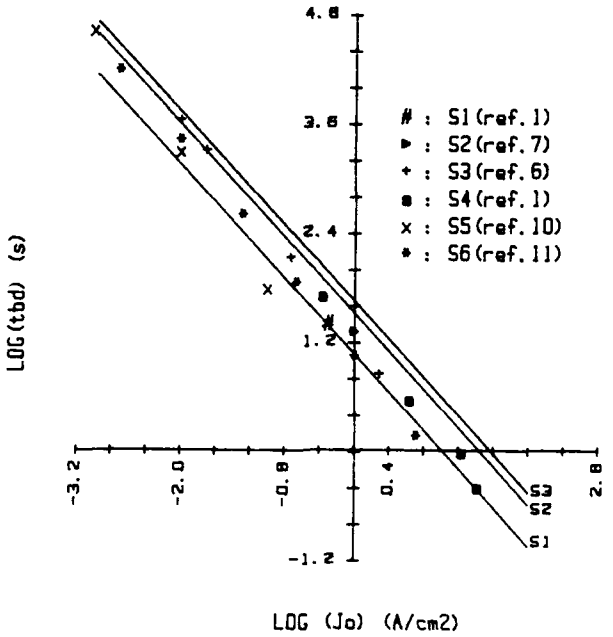


Fig 1 - Experimental and theoretical mean time-to-breakdown (t_{bd}) versus applied current density (Jo) in constant current breakdown tests.

As a conclusion, the presented model is able to explain the time-to-breakdown dependence on the applied current. From the experimental evolution of the degradation, we have been able to predict the quasi-linear relationship between $\log(t_{bd})$ and $\log(J_o)$ providing a good fit of the experimental breakdown data for more than three decades of applied current.

REFERENCES

/1/ E. Harari, J. Appl. Phys., 49 (1978) 2478.
/2/ I. C. Chen, S. Holland and C. Hu, IEEE Trans. Elect. Dev., ED-32, (1985) 413.
/3/ Z. A. Weinberg and T. N. Nguyen, J. Appl. Phys., 61 (1987) 1947.
/4/ S. Holland and C. Hu, J. Electrochem. Soc., 133 (1986) 1705.
/5/ Y. Nissan-Cohen, J. Shappir and D. Frohman-Bentchkowsky, J. Appl. Phys., 60 (1986) 2024
/6/ P. Olivo, T. N. Nguyen and B. Riccò, Proc. of The Physics and Technology of a-Si:F_x Conference, Les Arcs (1987), Ed. Plenum (in press).
/7/ P. Olivo, B. Riccò, T. N. Nguyen, T. S. Kuan and S. J. Jeng, Appl. Phys. Lett., 51 (1987) 2245.
/8/ B. Riccò, M. Ya Azbel and M. H. Brodsky, Phys. Rev. Lett., 51 (1983) 1795.
/9/ M. Dutoit, P. Fazan, A. Benjellou, M. Ilegems and J. M. Moret, Appl. Surf. Sci., 30 (1987) 333.
/10/ T. Kusaka, Y. Ohji and K. Mukai, IEEE Elec. Dev. Lett., EDL-8, (1987) 160.
/11/ P. Fazan, M. Dutoit, J. Mantley, M. Ilegems and J. M. Moret, ECS Fall Meeting, San Diego, CA (1986) 604.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

THE IMPACT OF DIFFERENT HOT-CARRIER-DEGRADATION COMPONENTS ON THE OPTIMIZATION OF SUBMICRON n-CHANNEL LDD TRANSISTORS

P.T.J. BIERMANS, T. POORTER and H.J.H. MERKS-EPPINGBROEK

Philips Research Laboratories, PO Box 80.000, NL-5600 JA Eindhoven, The Netherlands

Resumé

Mesurer $\Delta I_{ds}/I_{ds}$ durant uniquement une courte période de temps en vue de caractériser la dégradation dûe aux porteurs chauds, des transistors submicroniques au drain légèrement dopé, peut mener à des prédictions erronées de la durée de vie de ces transistors. L'évaluation des composants de dégradation ΔR_s,$-\Delta\beta/\beta$ et ΔV_T est proposée. Une explanation physique de l'orgine de ces composants est avancée conjointement avec leur dépendance de la dose n^- de la région du drain légèrement dopé.

Abstract

Measuring $\Delta I_{ds}/I_{ds}$ for only a short time period to characterize the hot-carrier degradation of submicron LDD transistors can lead to erroneous predictions of the transistor lifetime. Evaluation of the degradation components ΔR_s,$-\Delta\beta/\beta$, and ΔV_T is proposed. A physical explanation for the origin of these components is provided together with their dependence on the n^- dose of the LDD region.

1 Introduction

Hot-carrier degradation of submicron n-channel Lightly Doped Drain (LDD) transistors is often characterized as a relative decrease of the drain current, i.e. $\Delta I_{ds}/I_{ds}$ [1]. The slope of $-\Delta I_{ds}/I_{ds}$ versus time on a double-logarithmic scale shows a saturation behaviour. Since this slope is much smaller than for conventional devices a very long lifetime for LDD transistors is predicted. If, however, the electrical stress is continued an increase in the slope of $-\Delta I_{ds}/I_{ds}$ as a function of time is observed. This is shown in fig. 1: after a stress time of $2 \times 10^1 s$ the slope increases and after $2 \times 10^3 s$ a second increase in the slope is observed.

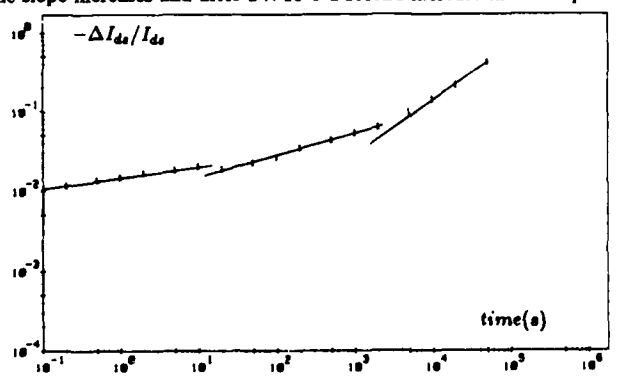


Figure 1: Relative current decrease $-\Delta I_{ds}/I_{ds}$ as a function of time after stressing at $V_{ds} = 9.0V$ and $V_{gs} = 4.0V$, measured at $V_{ds} = 0.1V$ and $V_{gs} = 1.5V$. The n^- dose is $2 \times 10^{13}/cm^2$.

In this paper the different components of the hot-carrier degradation occurring in the n-channel LDD transistors are analysed. A new method to characterize these components is described. Furthermore, the relative importance of these degradation components is investigated: experimental and device-simulation results are shown for LDD transistors with different n^- doses. Finally, a general guideline for the optimization of LDD transistors is given based on the obtained insight.

2 The degradation components

The n-channel MOS transistors used in the experiments have an effective channel length $L_{eff} = 0.6\mu m$ and a gate-oxide thickness $x_{ox} = 17.5nm$. The spacer width is $0.2\mu m$ and the phosphorous n^- dose is varied from $1 \times 10^{13}/cm^2$ to $8 \times 10^{13}/cm^2$.

In fig. 2 the linear characteristic of an LDD transistor with an n^- dose of $2 \times 10^{13}/cm^2$ is shown before the stress (1) and after a number of stress intervals (2-5).

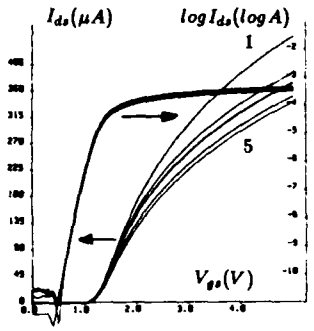


Figure 2: $I_{ds} - V_{gs}$ characteristics measured at $V_{ds} = 0.1V$ before the stress (1) and after different stress times (2 : $100s$; 3 : $1000s$; 4 : $10000s$; 5 : $20000s$) at $V_{ds} = 9.0V$ and $V_{gs} = 4.0V$. The n^- dose is $2 \times 10^{13}/cm^2$.

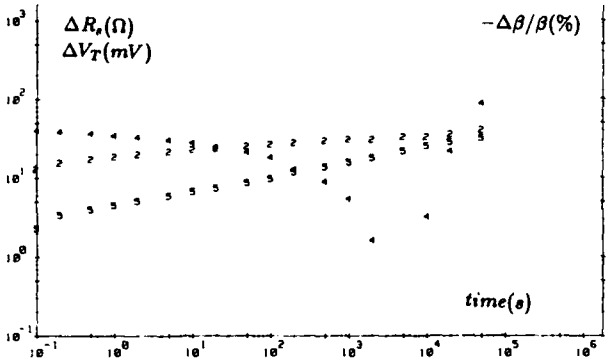


Figure 3: $\Delta R_s(2), -\Delta\beta/\beta(5)$ and $\Delta V_T(4)$ as a function of time during a stress at $V_{ds} = 9.0V$ and $V_{gs} = 4.0V$. The n^- dose is $2 \times 10^{13}/cm^2$.

Initially, the curvature of the characteristic increases significantly indicating a large series-resistance increase. Next, the slope of the curve decreases caused by a reduction in the transconductance. For longer stress times a small threshold-voltage shift starts to develop. This indicates that the degradation of the $I_{ds} - V_{gs}$ curve can be described by the superposition of three effects, namely a series-resistance modification ΔR_s, a relative change of the transconductance $\Delta\beta/\beta$ and a threshold-voltage shift ΔV_T. This is based on the fact that for low V_{ds} values the drain current I_{ds} can be approximated by:

$$I_{ds} = \beta \frac{(V_{gs} - V_T)}{1 + \beta R_s (V_{gs} - V_T)} V_{ds} \tag{1}$$

The change of the inversion-layer mobility due to stressing is incorporated in a change of β. From this expression the relative change of the drain current, i.e. $\Delta I_{ds}/I_{ds}$, is calculated as a function of ΔR_s, $\Delta\beta/\beta$ and ΔV_T:

$$\frac{\Delta I_{ds}}{I_{ds}} = \frac{\frac{\Delta\beta}{\beta} - \frac{\Delta V_T}{V_{gs} - V_T} - \beta(V_{gs} - V_T)\Delta R_s}{1 + \beta R_s (V_{gs} - V_T)} \tag{2}$$

In our method for characterizing hot-carrier stress the components ΔR_s, $\Delta\beta/\beta$ and ΔV_T are measured after each stress interval. In fig. 3 ΔR_s,$-\Delta\beta/\beta$, and ΔV_T are plotted as a function of time for the same set of curves as in figs. 1 and 2.

From fig. 3 the change in R_s is already obvious after a stress of $0.1s$. R_s increases proportionally with time, however, at a very small rate. Initially $-\Delta\beta/\beta$ is quite significant and increases also proportionally with time. The threshold-voltage shift develops much later and increases proportionally with time, similar to the degradation behaviour of conventional devices [2]. For small stress times a negative threshold-voltage shift is obtained. This is an artefact of the parameter-determination method: no initial threshold-voltage shift is observed in the subthreshold curves. The proportionallity constant is much higher for this ΔV_T component than for $-\Delta\beta/\beta$ indicating that ΔV_T eventually becomes the dominating degradation component.

3 The relative importance of the degradation components versus the n^- LDD dose

The relative importance of the different degradation components depends on the actual construction of the LDD transistor. The n^- dose is one of the key parameters in the design of reliable LDD transistors [3]. Fig. 4 shows measured results of the degradation components for an n^- dose of $4 \times 10^{13}/cm^2$ and $8 \times 10^{13}/cm^2$.

The series-resistance increase is more significant for lower n^- doses. The explanation for this behaviour lies in the exact position of the lateral peak electric field with respect to the gate edge [4]. In fig. 5 simulation results obtained with CURRY [5] are shown for the electron-hole-pair generation rate due to impact ionization using the impact-ionization rates given in [6] for an n^- dose of $2 \times 10^{13}/cm^2$ and $4 \times 10^{13}/cm^2$.

Reduction of the n^- dose leads to a shift of the peak electric field towards the spacer, resulting in a higher generation rate under the spacer. This indicates that the effective electron temperature is higher leading to a higher injection rate of electrons in the spacer oxide. These electrons are trapped, causing an increase in the series resistance R_s after very short stress times. For higher n^- doses the field under the spacer is lower causing less injection of electrons. Furthermore, it is more difficult to deplete the n^- region because of its higher

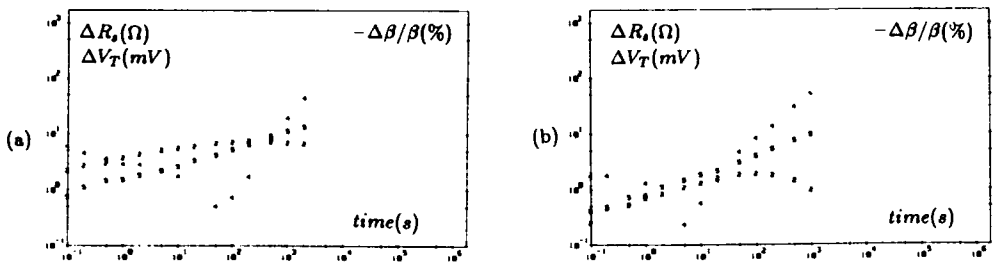


Figure 4: $\Delta R_s(2), \Delta\beta/\beta(5)$ and $\Delta V_T(4)$ as a function of time during a stress at $V_{ds} = 9.0V$ and $V_{gs} = 4.0V$ for (a) an n^- dose of $4 \times 10^{13}/cm^2$ and (b) $8 \times 10^{13}/cm^2$.

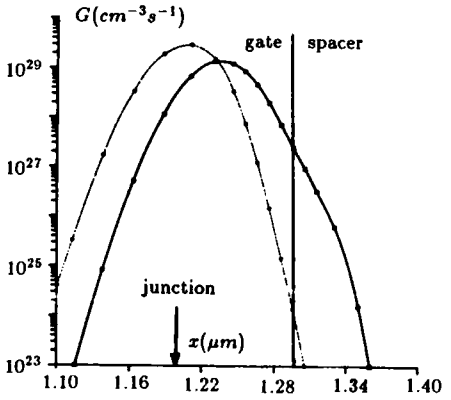


Figure 5: Generation rate of electron-hole pairs G as a function of the position x for an n^- dose of $2 \times 10^{13}/cm^2$ (□) and $4 \times 10^{13}/cm^2$ (○). $V_{ds} = 9.0V$ and $V_{gs} = 4.0V$.

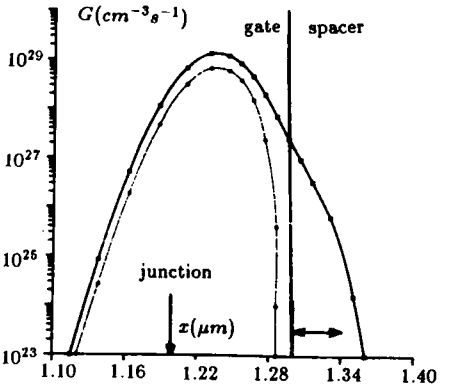


Figure 6: Generation rate of electron-hole pairs G as a function of the position x for an n^- dose of $2 \times 10^{13}/cm^2$. $V_{ds} = 9.0V$ and $V_{gs} = 4.0V$.(□): no oxide charge. (○): negative oxide charge.
The extension of the charge layer is indicated by the arrow.

concentration. This series-resistance increase is a self-limiting process: the negative charge trapped in the spacer reduces the injection rate by separating the current path from the peak electric field region. This is indicated in fig. 6 where the generation rate of electron-hole pairs due to impact ionization at the $Si - SiO_2$ interface of the transistor is plotted as a function of the position with and without negative charge in the oxide for an n^- dose of $2 \times 10^{13}/cm^2$. A significant reduction of the generation rate under the spacer is observed.

Another observation which can be made is that the relative transconductance reduction $-\Delta\beta/\beta$ is higher for lower n^- doses. One of the reasons is that the lateral electric field at the surface of the transistor above the drain junction is higher, leading to more injection of hot electrons in the gate oxide (fig. 5). This, together with a lower dope concentration in the high electric field region, results in a larger effective channel length after stressing.

The threshold-voltage shift ΔV_T is more severe for higher n^- doses. This is caused by the higher electron temperature in the channel region near the drain junction, as indicated by fig. 5. The injection of hot carriers in the gate oxide also causes the inversion-layer mobility to decrease due to the creation of interface states and fixed oxide charge at the interface above the channel [7]. As a result the transconductance decreases with increasing stress time, however at a lower rate than the threshold-voltage increase.

The behaviour of the substrate current I_{sb} as a function of the stress time can also be explained with these degradation components. In fig. 7 the measured substrate current is plotted for different stress times.

Initially, I_{sb} decreases due to the series-resistance increase and the accompanying decrease of the generation rate of electron-hole pairs. Next, an incease takes place due to an increase of the peak electric field: the negative charge present in the gate oxide above the n^- region causes a higher voltage drop across the drain-channel depletion layer. When the threshold-voltage shift becomes significant the peak in I_{sb} will shift to higher V_{gs} values. Eventually, I_{sb} will decrease due to the decrease in I_{ds}.

In fig. 8 the simulated substrate current is plotted for different charge distributions, corresponding to different stress times. The qualitative agreement between measured and simulated curves is clearly observed.

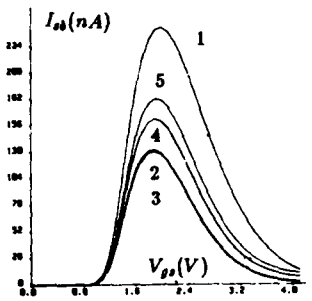


Figure 7: Measured I_{sb} as a function of V_{gs} before the stress (1) and after different stress times (2 : 100s; 3 : 1000s; 4 : 10000s; 5 : 20000s) at $V_{ds} = 9.0V$ and $V_{gs} = 4.0V$. The n^- dose is $2 \times 10^{13}/cm^2$.

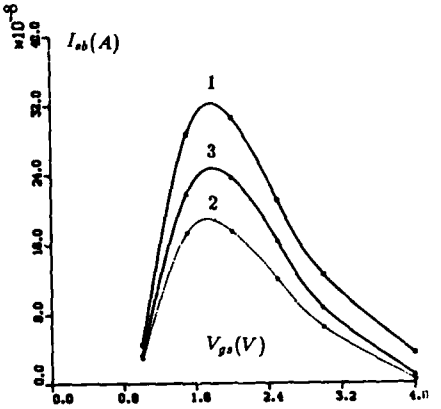


Figure 8: Simulated I_{sb} as a function of V_{gs} for different charge distributions: 1: no charge; 2: spacer oxide charge; 3: gate and spacer oxide charge. The n^- dose is $2 \times 10^{13}/cm^2$.

4 Optimization

From the abovementioned results it is clear that the optimum n^- dose depends on the applied lifetime criterion. Since the $\Delta I_{ds}/I_{ds}$ versus time characteristic does not have a single slope, ΔR_s, $-\Delta\beta/\beta$, and ΔV_T are analyzed as a function of the stress time. The optimum n^- dose is obtained in this case when all three degradation components reach the maximum acceptable change after approximately the same stress time. For an n^- dose of $4 \times 10^{13}/cm^2$ this situation is almost met assuming that a 10% modification of the parameters V_T, β and R_s determines the lifetime of the transistor. Since V_T of the analyzed transistors is approximately $1V$ and R_s 100Ω, this implies that a ΔV_T of $100mV$ and a ΔR_s of 10Ω are acceptable values.

However, the maximum allowed change in the parameters V_T, β and R_s influences the optimum n^- dose. For example, for a 1% modification of the parameters the increase in R_s will limit the transistor lifetime for almost any value of the n^- dose. In this case an n^- dose of $8 \times 10^{13}/cm^2$ is necessary. For most circuit applications however this lifetime criterion is far too stringent. Even a 10% current reduction will in practice not be noticeable in a circuit since the processing spread has to be taken into account in the design of a circuit. This means that ΔV_T and $-\Delta\beta/\beta$ are the limiting degradation mechanisms if the initial ΔR_s has been adjusted to a sufficiently low level.

5 Conclusions

Evaluation of the degradation components ΔR_s, $-\Delta\beta/\beta$ and ΔV_T leads to a better insight in the degradation behaviour of LDD transistors. Moreover, these components characterize the hot-carrier degradation unambiguously.

References

[1] Y. Toyoshima et al., 1984 IEDM Techn. Dig., p.786

[2] T. Poorter et al., 1984 IEDM Techn. Dig., p.100

[3] P. Biermans et al., 1986 ESSDERC conf. abstr.,p.183

[4] P. Biermans, 1985 ESSDERC conf. abstr., p.312

[5] S. Polak et al., Int. Journal for Numerical Methods and Eng., Vol.24

[6] J. Slotboom et al., 1987 IEDM Techn. Dig., p.494

[7] S. Sun et al., IEEE Journal of Solid-State Circuits, Vol. SC-15, No.4, p.562, 1980

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

MOSFET GATE CURRENT MODELLING USING MONTE-CARLO METHOD

J. VOVES and J. VESELY

The Czech Technical University of Prague, Faculty of Electrical Engineering, Suchbátarova 2, 166 27 Praha 6, Czechoslovakia

Abstract - The new technique for determining the probability of hot-electron travel through the gate oxide is presented. The technique is based on the Monte Carlo method and is used in MOSFET gate current modelling. The calculated values of gate current are compared with experimental results from direct measurements on MOSFET test chips.

1 - INTRODUCTION

Injection of hot electrons from MOSFET channel into gate is an important short channel effect which causes time-dependent degradation of transistor threshold voltage and channel conductance . The probability of electron emission into gate electrode is usually determined as a product of three probabilities: 1/ probability that electron has enough energy to surmount Si-oxide barrier, 2/ probability of free travel from depth of semiconductor to interface and 3/ probability of free travel through the oxide. Previous gate current models /1/,/2/ have used the probability of electron travel through the oxide in exponential form with argument x_o/λ_{ox}, where x_o is a distance of the maximum of potential barrier from Si-oxide interface for positive oxide field and x_o is equal to oxide thickness for negative oxide electric field. λ_{ox} is mean-free-path of electrons in the oxide. Another model /3/ supposes injection into the gate only for positive oxide field. Present model uses the Monte Carlo method for describing the collisions of electrons in oxide and determines more precisely the probability of electron travel through the oxide in dependence on the oxide electric field.

2 - THEORETICAL MODEL

Our study concentrates on the electron path through the gate oxide. No electron trapping in the oxide is supposed, if electron energy E is greater than bottom of oxide conduction band E_c in all points of electron path. The random path of electron in the oxide is started by Monte Carlo method. The collision number probabilities P_n are determined in dependence on the oxide thickness t_{ox}. They are calculated from 10^5 electron paths. But it is sufficient to perform the calculation for certain oxide thickness only once. Next simulations use those probabilities only as input data. If the average energy loss by the collision ΔE is known, then the probability of electron travel through the oxide P_{ox} can be determined in dependence on the oxide electric field $\mathcal{E}_{ox}$:

$$P_{ox}(\mathcal{E}_{ox}) = \sum_{n=0}^{n_o(\mathcal{E}_{ox})} P_n \qquad (1)$$

where n_o is a maximum of possible collisions without trapping of electron in

the gate oxide :

$$n_o(\varepsilon_{ox}) = \text{int}(q \varepsilon_{ox} t_{ox} / \Delta E) \qquad (2)$$

The algorithm checks condition $E > E_c$ in all points after collisions. The probabilities calculated for different ratios t_{ox}/λ_{ox} are shown in Fig. 1.

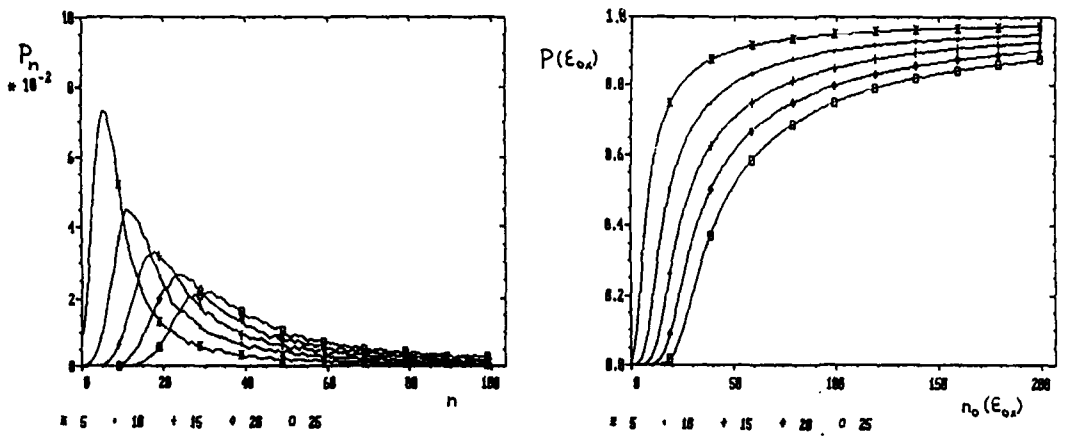


Fig. 1 Probability P_n of n collisions in oxide and dependence P_{ox} on ε_{ox} for t_{ox}/λ_{ox} equal to 5, 10, 15, 20, and 25.

We have used this technique in connection with the originally one dimensional "lucky electron model" /4/. This model has been extended to two dimensions. Electric field, potential and electron concentrtion in modelled transistors have been obtained with the help of MOSFET 2-D simulator MINIMOS. Program SUPREM II has been used for the calculation of impurity profiles in source, drain and channel regions.

3 - DEVICE CHARACTERISTICS

Direct measurements of gate current have been realised on MOSFET test chips prepared by different technologies. Minimum gate length is equal to 1μm. So, transistors with submicrometer effective channel length have been measured. We have chosen four types of transistors. All are based on standard NMOS technology and differs only in the preparing of source and drain regions. All types are described in tab.1.

Tab. 1 - Technologies of measured transistors

technology	enregy	dose	diffusion	L_{eff}
As drain	110 keV	$5.10^{15}cm^{-2}$	20'in O_2 20'in N_2	0.25μm
P drain	110 keV	$7.10^{15}cm^{-2}$	20'in O_2 20'in N_2	0.50μm
DD drain (P) (As)	110 keV 110 keV	$1.10^{14}cm^{-2}$ $5.10^{15}cm^{-2}$	20'in O_2 40'in N_2 after As implant	0.22μm
DD drain separate drive in	110 keV 110 keV	$1.10^{14}cm^{-2}$ $5.10^{15}cm^{-2}$	20'in O_2 40'in N_2 after P implant	0.26μm

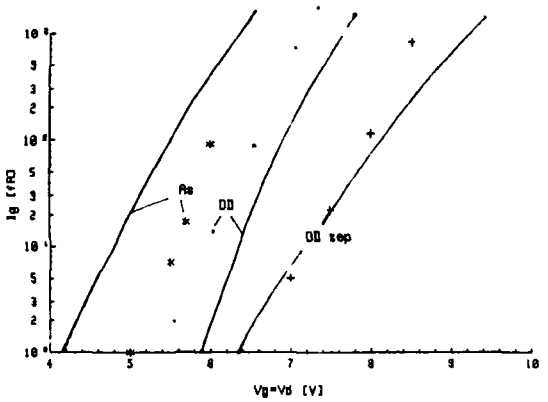


Fig. 2 Calculated and measured gate current for transistors with 1μm gate length

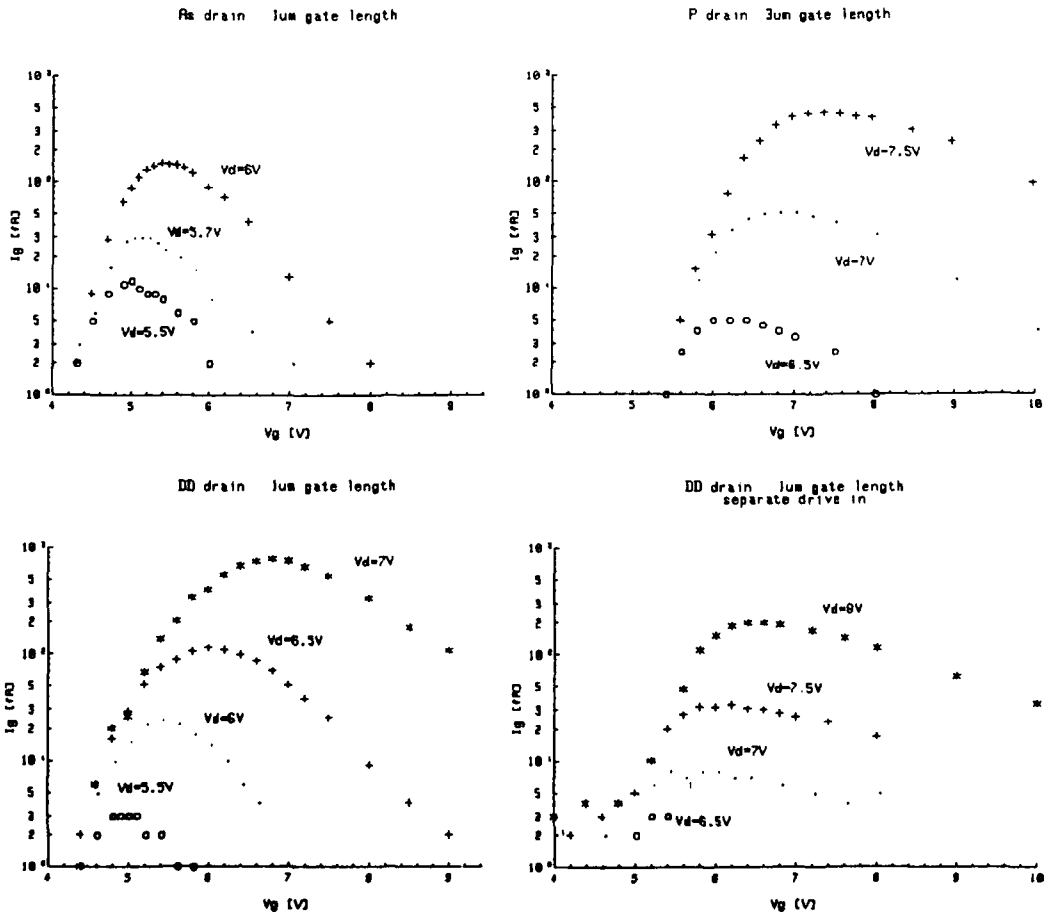


Fig. 3 Measured characteristics of gate current vs. gate voltage

Effective channel lengths have been determined from I-V characteristics of transistors with different gate lengths. In tab. 1 is presented L_{eff} of transistors with 1μm gate lenth. Only in P drain technology there is the smallest working transistor with 3μm gate. Oxide thickness is equal to 44 nm.

4 - RESULTS AND DISCUSSION

We have performed the gate curent modelling for all transistors with minimum gate lentgs. The average energy loss ΔE by the collision in the oxide is determined from dependence of electron velocity on oxide electric field published in /5/. If λ_{ox} = 3.2nm is assumed then ΔE is equal to 0.05eV. The mean-free-paths of electrons in silicon were taken from /1/ (λ = 9.2nm, λ_p = 61.6nm). In Fig. 2 comparission of model with experimental results is presented. Modelling and measurements have been realised for gate voltage equal to drain-source bias. The substrate bias have been chosen equal to-2V. The solid lines represents the modeled gate currents. Measured values are signed with points. Measurements have been realised for constant drain voltage as well. These characteristics are showm for all types of transistors in Fig. 3. To determine time-dependet degradation in correlation with gate current the dependece of transistor threshold voltage on the stress time has been measured. An example with stress conditions V_g = 3.5V, V_d = 7V is in Fig.4. Degradation of transistor with As drain is almost three times faster.

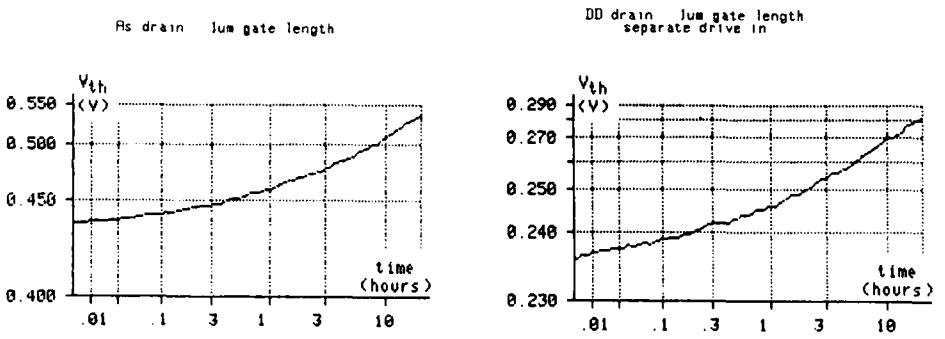


Fig. 4 Threshold voltage of As drain and DD sep. drain transistors in dependence on the stress time.

Neither fitting parameter is used in present model. Correlation with experiments is qualitatively good and gives the possibility to disscuse the applicability of different modifications of unipolar IC technologies for minimizing hot-electron induced aging. DD drain structure with "separate drive in" have good quality from this point of view. Our MMC based "oxide travel" model is applicable to different 2-D gate current simulators. In connection with more precise emission model than "lucky electron model" will give better results.

5 - ACKNOWLEDGEMENTS

The authors would like to express their appreciation to P.Kalina from Tesla Research Institute for his help in test chips fabrication.

REFERENCES

/1/ Tam, S., Ko, P. and Hu, Ch., IEEE Trans. El. Dev. ED-31 (1984) 1116
/2/ Hofmann ,K.R., Werner, Ch., Weber, W. and Dorda, G., IEEE Trans. El. Dev ED-32 (1985) 691
/3/ Watanabe, D.S. and Slamet, S., IEEE Trans. El. Dev. ED-30 (1983) 1042
/4/ Ning, T.H., Osburn, C.M. and Yu, H.N. J. Appl. Phys. 42 (1977) 286
/5/ Hughes, R.C., Solid-St. Electron. 21 (1978) 251

Session 7C Room C : Characterization

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

ELECTROLUMINESCENCE FROM SILICON DEVICES -A TOOL FOR DEVICE AND MATERIAL CHARACTERIZATION-

K. PENNER

Siemens AG, Corporate Research and Technology, Microelectronics, Otto-Hahn-Ring 6, D-8000 München 83, F.R.G.

Résumé-La lumière géneré à l'interieur d'un composant de silicium donne à la mème fois des informations sur le fonctionnement du composant ainsi que sur les materiaux utilisé. On a installé un système d'enregistrement photographique amélioré avec une sensibilité spectrale du proche UV à proche IR (300-1200nm) qui permit l'enregistrement de phénomen d'électroluminescence diverses, avec une resolution temporelle et spatiale très haut.

Abstract-The light distribution generated by electroluminescence phenomena inside silicon devices yields information on the functioning of the devices as well as on the materials used. An improved camera system has been installed, with a spectral sensitivity from near UV to near IR (300-1200nm) allowing the registration of various electrolumiscence phenomena, with high temporal and spatial resolution.

1-INTRODUCTION

The detection of the infrared radiation due to recombination of excess carriers in silicon is a widely used technique which originally has been developed for the characterization of power devices /1/. This non-destructive measurement technique is unique as it yields information about the carrier distribution within the bulk of power devices in the on state but also during turn on and turn off. The light intensity Φ is determined by the carrier densities n,p. In a first order approximation one usually assumes:
$\Phi = B*n*p$.
In the on state the low doped base regions are swamped with excess carriers and due to charge neutrality the electron and hole densities are equal. The wavelength of about 1.14 µm emitted from the radiative recombination of electrons and holes corresponds to the bandgap in silicon.
It is our intention to show that this sensitive method can also be applied to a large number of other semiconductor devices and materials. Beside the radiative recombination of excess carriers in silicon there are other electroluminescence phenomena in the interesting spectral range caused by hot electrons, for example the Drain Avalanche in MOS-FET's /2/, avalanche-breakdown in blocking pn-junctions and oxide currents /3/. The hot electrons are caused by high electric fields and have energies of a few eV above the conduction band minimum. When they return radiatively to their ground state they emit light with the corresponding wavelengths which can extend to the visible range.

2-EXPERIMENTAL

We have installed a new system with improved features which allows the detection of the radiation with spatial and temporal resolution. The core of this system is a high-speed, highly sensitive shutter camera /4/ with a selected S1-cathode as image converter, a microchannel plate (MCP) and a silicon intensified target (SIT)-Vidicon to attain high gain, and a digital video frame memory for the quantitative evaluation of the measured intensity distribution. (Fig. 1).

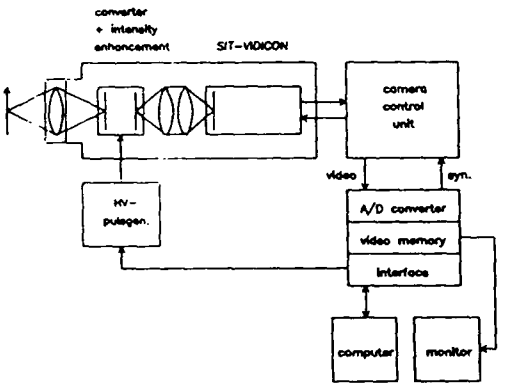


Fig. 1: Infrared highspeed camera system

The spatial resolution is determined by the optics used, the wavelength of the emission, the focus depth and the device geometry. The temporal resolution is obtained by the application of short voltage pulses to the MCP. The MCP works like a shutter with a system dependent minimum time window of 100ns.
The equipment is controlled by a host computer which has been programmed to permit noise reduction and background subtraction. Very low light levels can be detected by a stroboscopic mode operation provided the device under test can be driven in a repetitive manner. The range of physical dimensions for devices to be tested extends from a few microns to a few centimeters in diameter. Ideally the samples should be transparent to the radiation generated. Generally however some metallization is required to drive the device. In this case one can detect the straylight beneath the metallization or one can etch a pattern of small holes which does not disturb the electrical behaviour and which allows light penetration. In order to obtain spatial resolution in axial direction of vertical structures one often needs special preparation techniques, like cutting and polishing slices of the device preserving the relevant optical and electrical characteristics.

3-EXAMPLES

The method will be emphasized in the following examples:

- Analysis of the quality and electrical homogeneity of Si solar cells. They are driven as diodes in forward bias in order to swamp the active regions with carriers. Beneath the contacting metal stripes one can recognize the single cristallites and the active area of the solar cells (Fig. 2)
- Observation of the plasma spreading during turn- on of thyristors , GTO's and other laterally extended devices, which is crucial for the switching behaviour.
- Detection of local latch-up in C-MOS devices. Latch-up is correlated with a thyristor-like structure, so the photon emission mechanisme and the detectibility is the same, although it is only a local phenomena.
- Visualization of oxide currents: After tunneling through an oxide barrier between two silicon layers the carriers recombine from the upper band of the oxide to the conduction or valence band of silicon emitting faint light in the visible range. Localized oxide defects can be seen very well.
- Localization of current distribution in power semiconductor modules to recognize possible inhomogeneties due to thermal or contact problems during turn-on or in the on-state.
- Drain avalanche in Si-MOSFETs, which appears in the pinch off region when $U_{DS} > U_{sat}$ and the carriers are accelerated by a high electric field (Fig. 3).

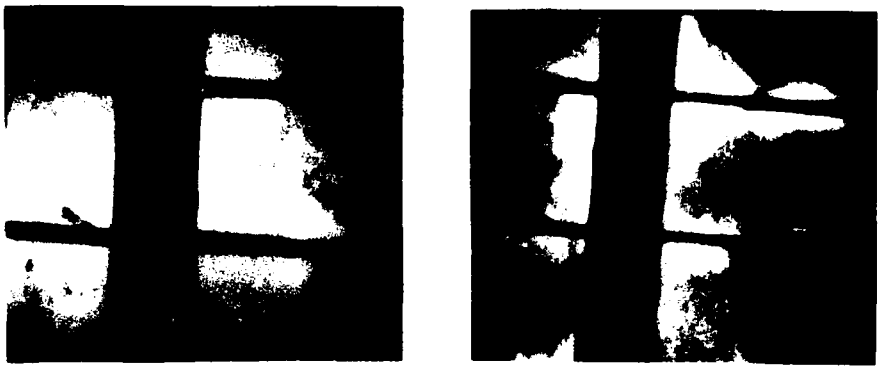


Fig. 2: Multi crystalline solar cell. Left: Top view
Right: Radiation distribution indicating enhanced carrier density along specific grains (the solar cell was operated in a forward biased diode mode). The black stripes are the contact pads.

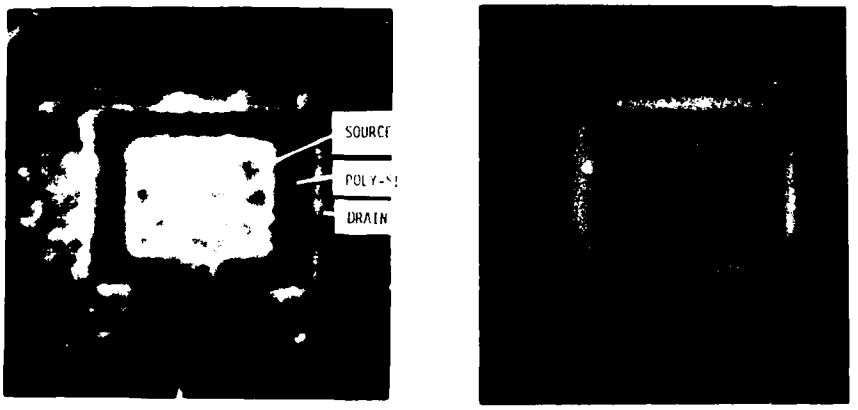


Fig. 3: MOS-FET testpattern. Left: Top view
Right: Light emission due to drain avalanche.

- The measurement of the axial carrier distribution n,p in PIN-diodes which delivers data on the diffusion length L
$n=C*\cosh(x/L)$
and so on the carrier lifetime τ
$\tau=D^2/L$.
The decrease of the carrier density at the edges of the device is an indication for the surface recombination velocity. Theoretically expected and simulated carrier respectively light distributions can be fitted to the experimental data using the appropriate parameters recursively (Fig. 4).
- Dynamical avalanche injection in Si-devices /5/: Of particular interest is the first observation of current squeezing and a subsequent generation of current filaments during critical turn-off of a MOS-GTO /6/. These filaments are caused by dynamic avalanche due to the high electric field and the high carrier density at the blocking junction during turn-off of the device. The generation of such filaments is localized at critical points correlated with the design of the device, but once created they can can move in lateral direction across the device.

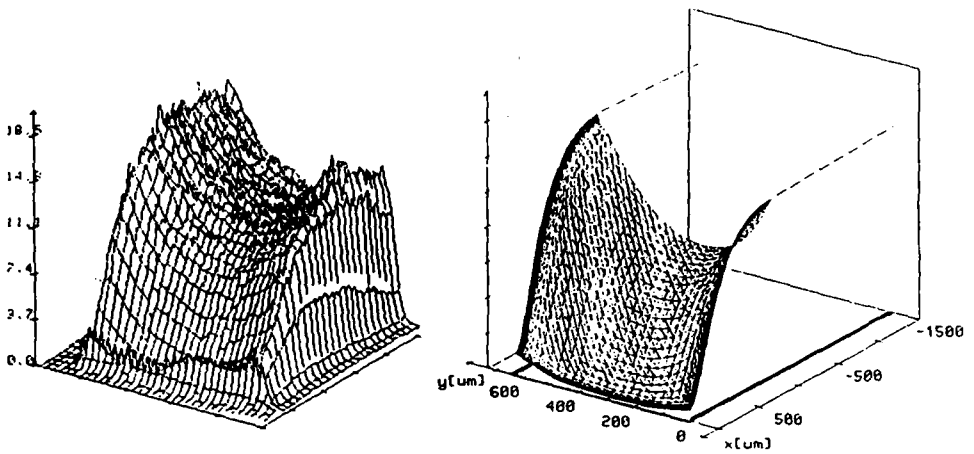


Fig. 4: Radiation profile from a PIN-diode.
left: measured right: simulated

4-CONCLUSIONS

The method of time- and space-resolved observation of the IR electroluminescence is not only a useful tool for the analysis of large area power-devices but also for a variety of other devices. The unique specification of our system is an improved sensitivity due to computer aided signal management. This makes accurate quantitative evaluation possible and thus allows for comparison of the results with theoretical curves or with the simulated device behaviour.

This work has been supported by the Federal Departement of Research and Technology of the Federal Republic of Germany. The author alone is responsible for the contents.

REFERENCES

/1/P. Voss, Proc. VDE-Meeting, Dynamische Probleme der Thyristortechnik, (Edit. by M. Depenbrock) Achen, p. 351 (1971)
/2/A.G. Chynoweth and K.G. McKay, Phys. Rev.102,no.2 (1956) 369
/3/N. Khurana and C.-L. Chiang, IEEE/IRPS (1987) 72
/4/Hamamatsu S1-II Shutter Camera
/5/A. Nakagawa et al., IEEE IEDM Tech.Dig., 150 (1985)
/6/M. Stoisiek, H. Strack, IEEE IEDM Tech.Dig. 158 (1985)

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

SOME ASPECTS OF THE SCANNING ACOUSTIC MICROSCOPE CONTRIBUTIONS IN THE EVALUATION OF DEVICE RELIABILITY

A. SAIED, C. AMAUDRIC DU CHAFFAUT, J.M. SAUREL and J. ATTAL

Laboratoire d'Automatique et de Microélectronique, Université des Sciences et Techniques du Languedoc, F-34060 Montpellier Cedex, France

RESUME

La microscopie acoustique, une technique non destructive d'investigation des circuits microélectroniques est presentée dans cet article. Des applications spécifiques ont été effectuées sur des structures typiquement rencontrées dans la plupart des circuits à semiconducteurs.

ABSTRACT

The scanning acoustic microscope, a technique suited to non-destructive investigations of microelectronic devices is presented in this paper. Specific applications are carried out on structures typically encountered in most of semiconductor components.

INTRODUCTION

The reliability of electronic systems is strongly affected by the quality of the components used. These components increase in complexity and in order to ensure their performance, a steady improvement of methods for detecting, locating and characterizing flaws is required. The scanning acoustic microscope (SAM) is a high resolution system and one of the most potential methods of defects investigations. It has proved to be a non destructive technique of evaluation suited to the characterization of material surfaces and to the inspection of the internal structures of optically opaque devices /1 - 2/. This paper outlines the principles of the SAM. Then, it presents through typical examples, the progress achieved by this technique for the detection of subsurface defects, located especially in bonding interfaces of semiconductor devices. Quantitative studies at high frsquency of layered structures encountered in integrated circuits will be discussed.

PRINCIPLES AND APPLICATIONS OF THE TECHNIQUE

The physical phenomenon underlying the SAM is the interaction of an ultrasonic wave with the mechanical properties (density, velocity, viscosity) of the sample to be examined. This acoustic wave is generated by a piezoelectric transducer, typically ZnO or $LiNbO_3$ which transforms a high frequency electrical signal into an ultrasound wave propagating at the same frequency. By means of a spherical sapphire lens a converging acoustic beam wave is focused on the sample surface. A coupling fluid is necessary for transmitting the acoustic signal from the lens to the sample. Water is commonly used, but the instrument performances are increased when liquid metals such as mercury are employed (this will be discussed below). In the reflection mode which covers a wide range of applications, the SAM operates with a single lens for transmitting and receiving the acoustic signal. The image is obtained by moving the focused beam on the sample. This is carried out by scanning the sample in two perpendicular directions relative to the lens. Focusing is obtained by displacing the lens towards the object. The resolution is determined by the focal spot size and is limited by the diffraction to about one wavelength. In the high frequency range (above 1 GHz), this resolution is comparable to that achieved by optical microscopes.

The main reasons that make the SAM a powerful technique employed where others are impractical or provide insufficient results are : first, acoustic waves are sensitive to the elastic properties of materials; second, they appear in different forms, i.e., longitudinal, transverse and surface modes. This latter particularity allows two possibilities of imaging by the instrument which are applied to subsurface testing and surface or near surface imaging and characterization.

1 - Subsurface testing

One of the most promising applications of the SAM is to probe deeply into materials. This ability arises from the fact that there is practically no attenuation of the acoustic energy in most solids. Moreover, the large change in sound velocity occuring at the couplant liquid-sample interface leads the acoustic waves to be refracted inside the sample, obeying

the same laws of refraction as in optics. Penetrating into material, the bulk waves (longitudinal and transverse) are reflected where they encounter discontinuities in the elastic properties, i.e, where acoustic impedance (density x sound velocity) changes. Thus, they give information on the internal features of the material and detect defects and inhomogeneities at several hundreds of microns below the surface. Due to the acoustic wave attenuation occuring mostly in the coupling fluid, for deep investigations the SAM employs low frequencies ranging from 100 to 500 MHz and giving within a specimen a resolution between respectively 50 and 10 microns, depending on the nature of the sample.

In the microelectronic field, among numerous applications for testing, the SAM has been used to control non-destructively integrated circuits throughout the thickness of their substrate, in order to evaluate for example, the wires bonding on the circuit metallization pads /3/. In the same way, different processes like welding, eutectic bonding, etc.. have been distinguished and inspected too. This kind of application has been carried out on power semiconductor modules. Indeed, for electrical and thermal conductivity, the different levels of these components are made of various materials and contain therefore different bonding areas. Many types of assembling techniques are used and the results obtained are of a great importance in the device functioning. The SAM has been employed to check joints in order to evaluate the circuit reliability and to eventually know the cause of failures, which may allow the improvement of the assembling process. In addition to many structures tested such as Cu/Al203/Cu, Si/BeO etc../4/, other typical examples will be discussed here. The experiments were achieved at 200 MHz which seems to be a good frequency for the resolution of details expected. Mercury has been used as a coupling fluid since it is more adequate than water for subsurface imaging. This arises from the fact that mercury impedance is similar to that of most solids which, reduces the impedance discontinuity at liquid-sample interface and improves the energy penetrating the sample. Moreover, being four times less absorbent than water, mercury allows one to double the operating frequency and to increase consequently the resolving power by a factor of about 2. However, it can attack metals such as gold, copper, silver etc.. so, to avoid damages, a protective thin layer deposited on the sample surface is required.

Figure 1 shows an acoustic image of an epoxy bond layer between a silicon chip and its copper heatsink.

Fig.(1): Si/Cu interface

Fig.(2): Si/Mo interface

Fig.(3): Mo/Cu interface

Bonding interfaces revealed by SAM at 200 MHz and with mercury as coupling fluid.

The bonding is visualized through 250 microns thick copper, and the resolution is of about 20 microns. The image reveals inhomogeneities in the epoxy layer marked by vertical dark traces. In addition, many voids of different sizes appear in bright circular areas. Since they stop the acoustic waves, they are highly reflective. Besides, the chip edges are not entirely visible which indicates that the epoxy layer thickness is not uniform.

Brazing is an another example of soldering used for materials assemblies and is successfully controlled by the SAM. Images on figures 2 and 3 are respectively of Si/Mo and Mo/Cu interfaces and show the type of defects occuring in brazed materials. The Si/Mo assembly is a part of a power transistor. This type of structure is usually encountered since molybdenum and silicon have nearly the same thermal expansion and elasticity modulus. The image is taken through 250 microns thick Mo and reveals defects at the interface marked by bright traces indicating that the Mo has been laminated. It seems that disbonds occur at irregularities caused by the lamination.

In figure 3, we see another type of defect at the interface between Mo and copper constituting the anode of a thyristor. The defects are detected through the Mo with a resolution of about 25 microns. In addition to the solder grain structures, the examination reveals dark regions denoting that bond is maintained, while bright areas show lack of adhesion. Using image processing, it is possible to evaluate the defect areas rate and therefore to estimate the device reliability.

2 - Surface imaging and characterization

The second important application is surface imaging and the determination of the elastic properties of microstructures located at less than one wavelength under the sample surface. High resolution is therefore required and the operating frequency extends beyond the GHz range. In that case, the acoustic signal arising from the sample is the result of an interference effect between two kinds of waves propagated in different directions: a bulk wave (longitudinal) reflected at the normal direction and a leaky surface acoustic wave (mostly Rayleigh wave) which is characteristic of the material, generated at the liquid-sample interface and detected at a critical angle from the surface (figure 4). This phenomenon has no counterpart in electromagnetic systems and is responsible for the contrast variations occuring in the surface imaging. This variation of contrast is more displayed when the lens-sample spacing is decreased. Furthermore, in the non-scanning system, the variation of the reflected acoustic signal recorded as a function of the sample defocus is a valuable method for measuring quantitatively, on a microscopic scale, the acoustic properties of the sample, and for explaining the contrast observed in surface images. Several authors have studied this technique which is called V(z) where V is the voltage received on the transducer connected to the reflectance function and z the sample defocusing distance. V(z) curves are periodic and are treated as a signature of the material since they exhibit the Rayleigh wave disturbance by the mechanical parameters of the surface. Thus, by measuring the curve periodicity Δz, the Rayleigh wave velocity can be determined, and hence, we can identify the material under test and its cristallographic orientation. For bulk specimens, the Rayleigh velocity is frequency independent but when a thin layer is deposited on the material, the propagation velocity of the surface wave is altered and becomes a function of both the frequency and the layer thickness. The V(z) technique is then of a great importance since it is possible to characterize a thin deposited film and to monitor point by point with a high lateral resolution the film thickness variations. In that manner, lack of adhesion of deposited layers and inhomogeneities in localized regions can be detected. Moreover, damages caused by implanted areas and defect structures such as cracks, voids etc.. change the Rayleigh wave velocity and can be therefore observed /5-6/.

Figure 5 shows an acoustic image of an integrated circuit which contains various layers made up of different materials. The contrast appears with different grey levels providing information on the nature of each region of the circuit. The image is taken at 1 GHz and the resolution is around 0.75 microns. As we have mentioned above, in order to understand the contrast mechanism and to interpret the acoustic images, characterization studies in relation to V(z) curves have been performed for materials such as silicon and aluminium used in integrated circuits. As an example, we show in figure 6 a V(z) curve obtained for a (111)Si. The experiments has been carried out at 600 MHz with a spherical lens having a half opening angle of 50° and water as a coupling fluid. To increase the signal- to- noise ratio, the curves have been recorded at 52°C since high temperatures decrease the ultrasound absorption in water. The periodicity Δz which depends on the frequency and the velocity in water, is estimated to be of 24.5 microns. Upon using the following equation, the Rayleigh wave velocity Vr in (111)Si is determined with an accuracy of 0.3% : Vr = 4779 m/s.

$$Vr = Vw \,/\, (1-(1 - Vw/2F\Delta z)^2)^{\frac{1}{2}},$$ Vw is the velocity in water and F is the frequency /7/.

The second curve in figure 6 shows the V(z) obtained for a 2500 A°thin layer of aluminium deposited on (111)Si.

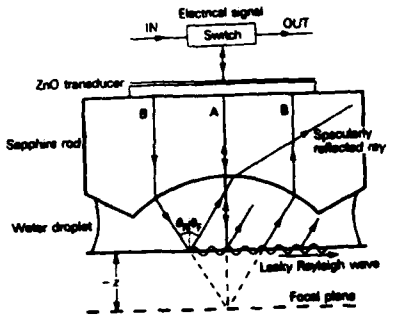


Fig.(4): Schematic diagram showing Rayleigh wave propagation.

Fig.(5): Acoustic image of a CMOS integrated circuit taken at 1 GHz.

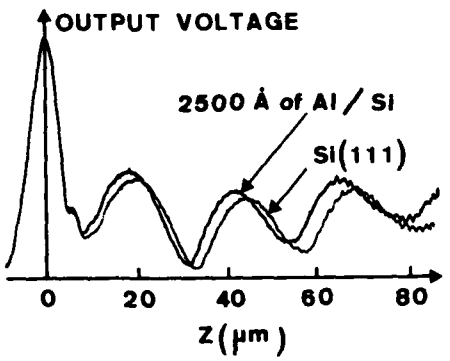


Fig.(6):Acoustic signature of Si and Al/Si-600MHz.

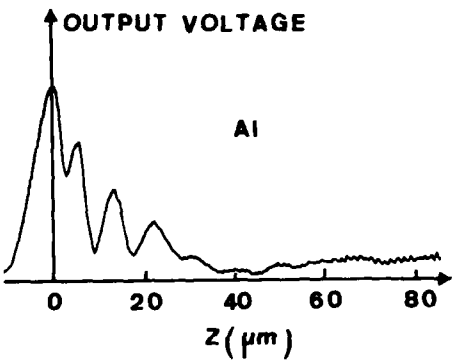


Fig.(7):Acoustic signature of Al-600MHz

Since the layer thickness is much less than the aluminium Rayleigh wavelength λr (which is of about 4.9 microns at 600 MHz), the V(z) curve remains similar to that of the substrate. However, we note that the effect of the aluminium layer is to lower of about 4% the Rayleigh velocity in Si. This is shown by a shift in the priodicity: Δz = 22.55 microns which gives Vr = 4590 m/s. This variation will affect the contrast of an acoustic image. If the aluminium layer thickness increases and extends λr, the Rayleigh velocity will decrease and reach the Rayleigh velocity in aluminium. Figure 7 shows the V(z) recorded in the same experimental conditions for a bulk aluminium. The curve is quite different since Vr=2980m/s.

CONCLUSION

The results presented here are significant. They indicate that the SAM contributions in the microelectronic field should lead to the use of this instrument in industrial environments for the control of components and research into developing processes. However, further work is needed to visualize simultaneously, with the highest resolution, multiple interfaces of a component and to interpret the information provided.

References

/1/A.J.MILLER: Scanning microscopy in electronic research.IEEE sonics and ultrasonics. Vol. SU 32 n°2 March 1985, pp 320-324.

/2/A.J.MILLER: Applications of acoustic microscopy in the semiconductor industry. Proceedings Acoustical imaging - vol.12 July 1982, pp 67-77.

/3/J.ATTAL, A.SAIED, J.M.SAUREL, C.C.LY: Acoustic microscopy:deep focusing inside materials at gigahertz frequencies.Acoustical imaging, Vol.18, June 1988 (in press).

/4/J.ATTAL, J.M.SAUREL, A.SAIED, J.C.HUBERT : Failure testing in materials for microelectronics by acoustic microscopy - 3rd international conference quality in electronic components. Bordeaux March 1987, pp 148-154.

/5/P.J.BURNETT, G.A.D.BRIGGS: The elastic properties of implanted silicon. Journal of materials science 21 - 1986, pp 1828-1836.

/6/J.KUSHIBIKI, N.CHUBACHI: Material characterization by line focus beam acoustic microscope. IEEE sonics and ultrasonics, Vol. SU32, n°2, March 1985, pp 189-212.

/7/R.D.WEGLEIN: Acoustic micro metrology.IEEE sonics and ultrasonics,March 1985,pp 225-234.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

ELECTRON BEAM WRITING ERASURE SWITCHES

P. GIRARD, B. PISTOULET and M. VALENZA

Laboratoire d'Automatique et de Microélectronique de Montpellier (UA 371), Université des Sciences et Techniques du Languedoc, Place Eugène Bataillon, F-34060 Montpellier Cedex, France

Résumé - Dans cet article, nous montrons pour la première fois, qu'il est possible de faire basculer un transistor MOS à grille flottante de l'état passant à l'état bloqué et réciproquement à partir d'une irradiation par faisceau d'électrons dans un microscope électronique à balayage. Nous déterminons l'influence du champ d'extraction sur les rendements de charge positive, et celle de l'état électrique initial de la grille flottante sur le potentiel positif atteint. Ces résultats ouvrent la voie au test et à la reconfiguration de circuits intégrés par le seul faisceau d'électrons.

Abstract - In this paper, it is shown for the first time that an electron beam is able to switch an oxide embedded floating gate MOS transistor from the on to the off state and inversely. These experiments are achieved in a classical Scanning Electron Microscope. We derive the influence of : i) the extraction voltage on the positive charging yield, ii) the initial floating gate voltage on its final positive voltage. These preliminary results open the way to future e-beam testing and reconfiguration method of integrated circuits.

I - INTRODUCTION

Today the electron beam testing of integrated circuits is coming of age and the equipments now available are consistent with the performances of integrated circuits /1,2/. However the interaction of the electrons with sensitive parts of the circuits may lead to new electron beam testing possibilities. For example the activation of switches with the primary electrons is possible without any extra electrical wire /3/. It opens the way to partitioning or reconfiguration of circuits in connection with testing. A first possibility consists in short circuiting a reversely biased junction by the excess electron hole pairs located in the vicinity, this is the EBIC effect. Unfortunately, in order to cross the various oxide sheets above the silicon substrate high energy electrons are required. This may lead to damages on adjacent devices /4/ and makes difficult primary energy changes. A second way consists to deposit charges on or above an oxide embedded floating gate MOS transistor. Up to now efficient negative charge deposition has been observed at primary beam energy of 5 /3/ and, more recently of 3 keV /5/. Ultra violet irradiation is required in order to erase the deposited charge.

In this paper, we report experimental results showing that writing-erasure cycles can be achieved only by means of electron beam deposition into the top oxide covering the floating gate of first a negative and secondly a positive charge. The results and the main parameters governing the positive electron beam charging are discussed below.

II -EXPERIMENTS AND DISCUSSIONS

The experiments have been achieved on n channel depletion mode floating gate transistors embedded in a 700 nm top oxide. The polysilicon gate dimensions are 25 x 25 μm^2 and the oxide thickness around the gate merges around 2 µm. The devices are irradiated inside an ISI SS 40 Scanning Electron Microscope. The evolution of the electrical parameters is established using a Hewlett Packard HP 4145 transistor parameter analyser. The charge deposition is obtained in the scanned mode, the surface exposed to the electrons is much larger than the gate of the device.
In a first step the floating gate normally on transistors are shifted in the

off state by a 3 keV primary beam negative charging /5/. Then positive charging is achieved by irradiating with primary electrons of 1.4keV and using a 0.28 kV/mm extraction field, this switches the device to the on state (fig. 1).

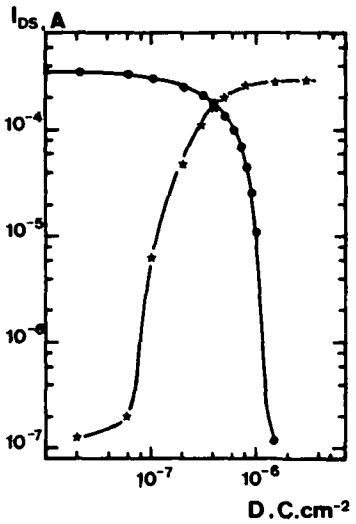


Fig. 1 - Example of the drain source currents of the floating gate MOS transistor, at 5 V drain source polarisation versus the incident dose. Negative and positive charging correspond respectively to (•) and (x).

In both cases, a 3.10^{-6} C cm^{-2} incident charge is sufficient to obtain a current variation of three orders of magnitude. The parameters driving the negative charging are well established /6/ but, up to now, this is not the case in the positive region. The floating gate voltage variations are deduced /6/ by comparing the electrical characteristics of same size floating gate and normal MOS transistor processed on the same chip. The experimental results are reported on fig. 2 at various extraction fields.

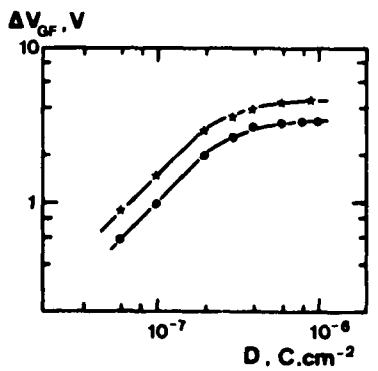


Fig. 2 - Evolution of the positive floating gate voltage variation versus the incident dose. The extraction fields used are respectively 1 kV/mm (*) and 0.28 kV/mm (•).

These variations correspond to the difference between the initial -3 V gate voltage and successive steps when positive charging occurs. The deposited charge first increases and then saturates when a final positive voltage is obtained. The limit voltage increases with the extraction voltage.
The charging yield Y of the top oxide,i.e. the ratio between deposited and incident charges, depends on the backscattered and on the secondary yields η

and $\delta(Ep, Xj)$ respectively /7/. This last yield depends on the primary beam energy Ep and also on other parameters Xj considered below.
The bound voltage is obtained when $Y = \eta + \delta(Ep, Xj) - 1$ (1) tends towards zero. The yield variation is deduced at various extraction voltages (fig.3) from the evaluation of the charging slope in the linear zone (see fig.2).

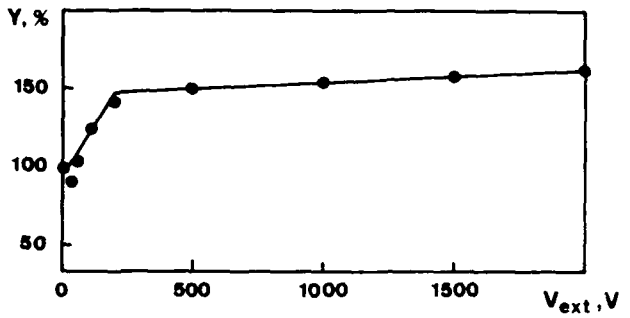


Fig. 3 - Charging yield variations versus extraction voltage.

Since a large planar extracting grid about 1 cm diameter is placed above the specimen, in the objective lens bore, the homogeneous extraction field strongly affects the electron trajectories. In the first zone the major part of the reemitted electrons is directed out of the insulator surface, while in the second part only the particles with highest angles of incidence are driven by the field.
The positive saturation voltage reached increases with the extraction voltage (see fig.2), but depends only slowly of the initial voltage of the floating gate within the -3 V, + 0.75 V range (fig.4).

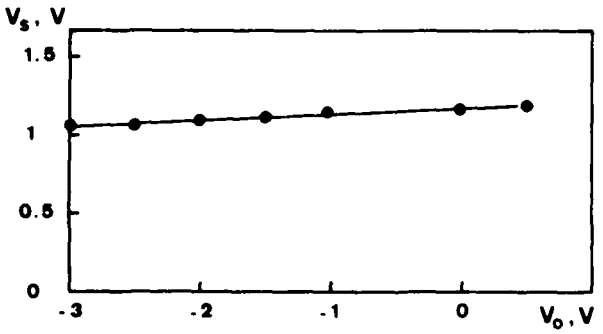


Fig. 4 - Evolution of the saturation voltage at various initial states.

On one hand the limit obtained is largely sufficient to switch on again the transistor. On the other hand it remains higher than the capacitively induced surface voltage, i.e. $Vsc \simeq (Cvacuum / Coxide) Vext.$, which is equal to 0.15 V on the device.
Let us now come back to the secondary yield which is strongly related to the saturation voltage. It depends on : i) the primary beam energy as currently admitted /7/, ii) the extraction field as established here, iii) the positive surface voltage.
When the voltage is increased, low energy electrons are driven back to it, it gives then a low energy cut off filter. Let us call this function F(Vs), where Vs is the surface potential, it is written

$$F(Vs) = \int_{qV_s}^{50eV} N(E)\, dE \;/\; \int_{0}^{50eV} N(E)\, dE \qquad (2)$$

with $N(E) = 1,5\ E \exp(2 - (8/3\ E)^{1/2})$ (3)

is the secondaries energy distribution as established by Kollath /8/. Consequently if we suppose that $\delta o(Ep)$ represents the normal secondary yield, i.e. without any surface effect on the electrons trajectories, F(Vs) the filter function, and G (Vextr.) the influence of the extraction field, the effective secondary yield may be written

$$\delta(Ep , Vs , Vextr.) = \delta o(Ep) \cdot F(Vs) \cdot G(Vextr.) \quad (4)$$

Then the final attainable voltage corresponds to :

$$F(Vs) = (1 - \eta) / (\delta o(Ep) \cdot G(Vextr.)) \quad (5)$$

It is clear that the higher the extraction voltages and the normal yields, the lower the filtering function and the higher the positive surface allowable voltage are. This simple approach gives a qualitative agreement with experiment. In fact, the charging mechanism in presence of an extraction field is a complex phenomenon /9,10,11/ requiring more further developments and studies.

III-CONCLUSION

We have demonstrated for the first time, that both negative and positive charging of silicon dioxide embedded floating gates MOS transistors are achievable in a Scanning Electron Microscope. The activation of such a switch has been actually demonstrated, and the main parameters for positive charging have been established. It is shown that the extraction field may appreciably change the charging yield and consequently the positive voltage attained by the gate voltage. This parameter is only slowly dependent on the initial gate voltage, so a MOS transistor can be switched from the off to the on state with a current ratio of three orders of magnitude. This work shows that the concept of e-beam testing and reconfiguration using the same tool does not appear unrealistic.

AKNOWLEDGEMENTS

We acknowledge the EEC for the initial support in the early stages of this work within the ESPRIT 824 project (Wafer Scale Integration).

REFERENCES

/1/ Menzel E., Buchanan R., Elect. Lett., 20, 10 (1985) 408.
/2/ Frosien J., Plies E., Microelectronic Engin.7 (1987) 163,Wolfgang E., Courtois B. ed., Elsevier Science Pub. (North Holland),Amsterdam.
/3/ Shaver D.C., J. Vac. Sc. and Techn. 19 (4) (1981) 1010.
/4/ Girard P., Scann. Microsc., 2, 1 (1988) 151.
/5/ Girard P., Pistoulet B. Valenza M., Lorival R., IFIP Workshop on Wafer Scale Integration, to be published, North Holland.
/6/ Girard P., to be published in Scanning Microscopy .
/7/ Goldstein J.L., Yakowitz H., Practical Scanning Electron Microscopy, Plenum Press, New York (1975).
/8/ Kollath R., Hanbuch der Physik, 21, 241 (1956) Springer Verlag,Germany.
/9/ Brunner M., Menzel E., J. Vac. Sci. Technology B1(4) (1983) 1344.
/10/ Nye P., Dinnis A., Scanning, 7 (1985) 117.
/11/ Reiners W., Herrmann K.D., Kubalek E.,Scann. Microsc.,2, 1 (1988) 161.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

CHARACTERIZATION OF ANOMALOUS LATCH-UP EFFECTS BY MEANS OF INFRARED MICROSCOPY AND SPICE SIMULATION(1)

C. CANALI, F. CORSI*, M. MUSCHITIELLO**, M. STUCCHI** and E. ZANONI*

Dipartimento di Elettronica ed Informatica, Universita' di Padova, Via Gradenigo 6a, I-35131 Padova, Italy
**Dipartimento di Elettrotecnica ed Elettronica, Universita'di Bari, Via Re David 200, I-70125 Bari, Italy*
***Tecnopolis CSATA, Str. Prov. per Casamassima km 3, I-70010 Valenzano (Bari), Italy*

ABSTRACT
Anomalous effects have been evidentiated during pulsed I/O overvoltage tests, such as "window effects", i.e. disappearing of the latch-up condition for high I/O injected current. Infrared microscopy observation reveals that anomalous effects are due to the dynamic redistribution of supply current between different latch-up paths. This analysis is confirmed by the SPICE simulation of the lumped equivalent circuit of a CMOS output comprising two coupled pnpn parasitic structures.

1. INTRODUCTION

Pulsed I/O overvoltage tests allow to evaluate eletrical sensitivity to latch-up of finished CMOS integrated circuits, reducing device heating and consequent damage, and providing a more realistic simulation of transients [1].

However, due to the three-dimensional current distribution within a single or between different pnpn structures on the same chip, anomalous effects may arise, such as : decrease of the latch-up steady-state current at increasing the current injected into I/O, Fig. 1, and "window effects", i.e. the existence of a well defined interval of pulsed injected currents, with lower and upper limits, for latch-up to occur, Fig. 2.

These effects may lead to neglect potentially dangerous latch-up paths and can induce serious errors in the evaluation and screening of CMOS integrated circuits [2-4].

2. ELECTRICAL CHARACTERIZATION AND ANOMALOUS EFFECTS

We tested a wide number of commercial devices with different technologies, layout and logic functions by applying pulsed overvoltages to I/O pins. Supply current Idd was monitored to verify the occurrence of latch-up. Several anomalous effects were observed, which can be better described by reporting the steady-state value of the supply current Idd in latch-up condition as a function of the negative pulse current applied to the output, Iout, Figs. 1 and 2.

Figs. 1 and 2 refer to a specific output of a 4 bit full-adder manufactured with a standard 3 μm CMOS p-well technology, for a pulse duration τd = 20 μs and a supply voltage Vdd = 11 V. In this circuit latch-up can be initiated by applying a negative voltage pulse (τd = 20 μs) to a specific output, thus forward biasing the n^+/p drain junction of the NMOS output transistor.

Two anomalous effects were identified during measurements; the first one is observed at high values of Vdd, 12.2 V in Fig. 1. As can be seen, at Iout = 24 mA latch-up is initiated and Idd reaches the value of 120 mA. At increasing Iout, the steady-state latch-up current changes, suggesting the existence of more than one latch-up path. In fact, four different latch-up conditions can be observed in Fig. 1, identified as a), b), c) and d). Moreover, Idd is not an increasing function of Iout as can be expected if an increasing number of latch-up paths had turned on.

The boundaries between conditions a), b) c) and d) were found to be markedly dependent on Vdd and τd and, at decreasing Vdd, the interval of existence of some conditions of latch-up is reduced until they disappear.

When Vdd is decreased below 12 V a second anomalous effect can be observed, which consists of the appearance of a "latch-up window", i.e. of an interval of Iout values, with lower and upper limits, for latch-up to occur.

(1) Work partially supported by CNR, Progetto Finalizzato Materiali e Dispositivi per l'Elettronica a Stato Solido

Figure 2 reports the steady-state latch-up current Idd as a function of Iout for Vdd = 10 V; as can be seen latch-up can only be triggered for 24 mA < Iout < 172 mA and Iout > 272 mA. Finally, for Vdd < 8 V, latch-up can not be triggered anymore.

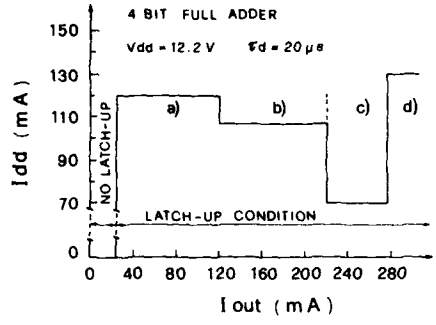


Fig. 1 Steady-state latch-up current Idd vs Iout for Vdd = 12.2 V and τd = 20 μs in the 4 bit full adder. Four different latch-up conditions are identified as a), b) c) and d).

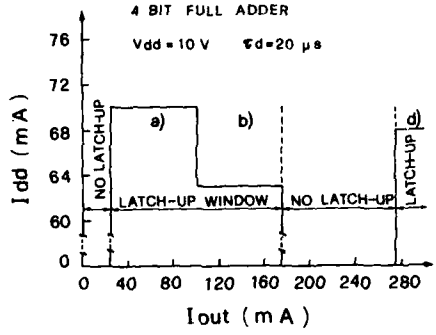


Fig. 2 Steady-state latch-up current Idd vs Iout for Vdd = 10 V and τd = 20 μs in the 4 bit full adder, showing the presence of the latch-up window effect.

3. INFRARED MICROSCOPY ANALYSIS

To locate active latch-up paths and explain anomalous effects we made recourse to IR microscopy [5]. This technique allows to identify latch-up current flow within the parasitic pnpn structures by detecting the near IR radiation (λ = 1.1 μm) emitted due to electron-hole recombination in Si crossed by large current density. A special IR system (Infrapol by Reichert-Jung) was adopted; this system is equipped with a PbS vidicon camera and with objectives modified and treated to minimize aberrations and glare and to optimize contrast and image quality in the IR spectrum. The response covers the range between 450 nm and 1800 nm but the visible part of the spectrum is usually filtered to avoid saturating the detector. Because blackbody radiation of a sample at T < 300 °C has negligible intensity for wavelenghts lower than 2 μm, the map of the IR radiation emitted by the device under test will show the actual distribution of recombination current and not the areas of maximum heat dissipation. Moreover, because Si is transparent to IR radiation, the emission due to latch-up can be observed from the back, through the Si substrate, even when latch-up sites are covered on top by metal lines. This greatly favour IR microscopy with respect to other analytical techniques for latch-up analysis and layout correction of VLSI circuits with tight design rules.

Fig. 3 shows the IR images (in latch-up condition) of the output of a quad and-or select gate manufactured with a standard 3 μm CMOS p-well technology. At increasing the Iout current, five different levels of steady-state latch-up current absorption can be measured. When the device is observed by IR microscopy in the same electrical conditions, the presence of five different combinations of active latch-up paths are observed, Figs. 3 a) - e). Areas of maximum current density appear as white spots in Fig. 3.

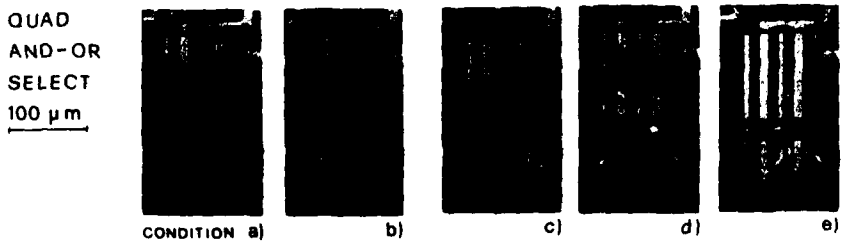


Fig. 3 IR micrographs (from the top) with topography superimposed of the output region of the quad and-or-select in five different conditions of latch-up current absorption.

The 4 bit full adder whose electrical characteristics were described in Figs. 1 and 2 also exhibits the presence of four different latch-up current paths, corresponding to the a), b), c) and d) latch-up conditions identified in Fig. 1. At a lower Vdd voltage, Fig. 2, IR

microscopy reveals the turning on of the paths a), b) and d), while in correspondence of the Iout range where latch-up condition c) was previously observed, Fig. 1, no latch-up is observed, so that a "latch-up window" is created, Fig. 2.

In all the circuits which exhibited latch-up windows we verified, by means of IR microscopy, the presence of different latch-up paths. Similar anomalous effects were observed in a large number of CMOS devices, comprising VLSI technologies with very small geometries.

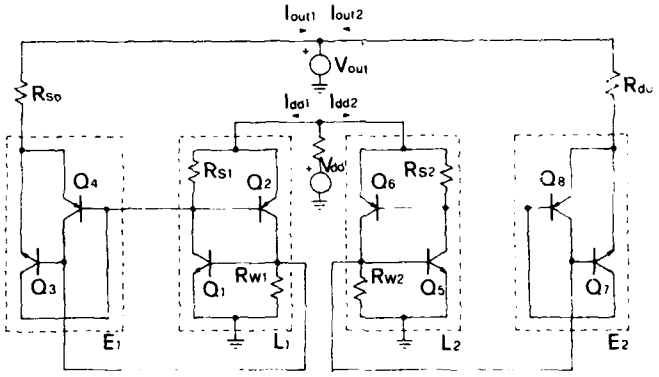


Fig. 4 Lumped equivalent circuit model of the output of a CMOS IC's comprising two parasitic pnpn paths in shunt connection, L1 and L2, and two output circuits E1 and E2. The two structures E1-L1 and E2-L2 are simmetric, except for the two output resistances Rso and Rdo. Rso is lower than Rdo.

4. SPICE SIMULATION

Results of electrical measurements and IR microscopy suggest that both anomalous effects depicted in Figs. 1 and 2 are due to dynamic competition and current redistribution between different latch-up paths. To confirm this hypothesis, we made recourse to the SPICE simulation of a lumped-parameter circuit which comprises two parasitic structures in shunt connection, Fig. 4. In Fig. 4, bipolar transistors Q1-Q2 and Q6-Q7 correspond to actual SCR structures, associated with sources of MOS transistors and identified as L1 and L2 in Fig. 4; Q3-Q4 and Q7-Q8 are instead due to MOS drains and form the "output" structures, E1-E2 in Fig. 4, driven by the same voltage source Vout. The two structures E1-L1 and E2-L2 differ only in the value of the output series resistances Rso and Rdo, so that different currents Iout1 and Iout2 are injected into E1 and E2. In particular, Rso is lower than Rdo.

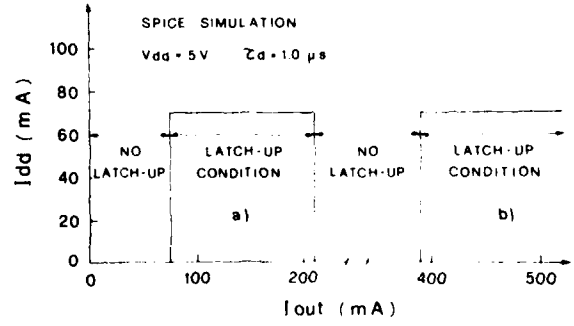


Fig. 5 Steady-state latch-up current Idd vs Iout for Vdd = 5 V and τd = 1 μs as obtained by the SPICE simulation of the circuit of Fig. 4. Latch-up conditions a) and b) correspond to the turning on of structure L1 and L2 respectively. The presence of the latch-up window effect is also evident.

Results of the simulations are summarized in the Idd-Iout plot of Fig. 5. Two latch-up windows, separated by an interval of no-latch-up condition were obtained. In fact, latch-up

can only be triggered for 75 mA < Iout < 210 mA, when the path L1 is in latch, and for Iout > 390 mA, when the latch-up current flows through L2. Owing to the assumption of two identical latch-up structures in the model, the value of the latch-up current Idd in steady-state conditions is the same for the two latch-up conditions identified as a) and b) in Fig. 5. For small negative excitations the L1 path is activated thanks to the current injected into the base of Q2, which is greater than the corresponding current injected into the base of Q6. For high negative excitation voltages transistors Q4 and Q8 operate in the far reverse active region during the Vout pulse. Because of the longer recovery time of the output circuit E2, when the external pulse is removed, transistor Q6 is still holded in the forward active region, while Q2 has already turned off, so that circuit L2 latches.

Competition between E1-L1 and E2-L2 circuits is essentially dynamic and "window" behaviour can be obtained also when switching times of the two branches of the circuit are made different by changing the value of emitter-base shunt resistances and/or of forward and reverse transistor transit times.

5. CONCLUSIONS

Infrared microscopy analyses of anomalous effects confirm that discontinuities in the value of Idd are due to current switching between different parasitic structures in shunt connection. The presence of latch-up windows appears to be also due to the same reason; in fact, window effects are always coupled with the existence of more than one parasitic path within the same chip.

Despite the well-known limitations of the lumped equivalent circuit model of latch-up, and the difficulty of measuring correct model parameters for parasitic bipolar devices [6], the SPICE simulation of two parasitic SCR structures in shunt connection is able to reproduce anomalous effects. Dynamic current redistribution between the two structures, which can be due to many possible difference in circuit parameters, causes the switching of latch-up from one structure to the other, and at low Vdd values, the presence of latch-up windows.

Our results are in agreement with those described by other authors [3,4] who studied the latch-up window effects as a function of the dose rate of ionizing radiations. In [4] Johnston and Baze used a Q-switched Nd:glass laser (λ = 1.06 μm) to induce latch-up in tested devices and found out that competition between different latch-up regions was the cause of window behaviour in CMOS devices. However, beam diameter of the scanning laser was 150 μm, so that lateral resolution was too poor to follow the behaviour of single distributed pnpn structures. Our observations demonstrate that latch-up windows with respect to I/O injected current can be also due to three-dimensional current redistribution within the same pnpn, Fig. 3. Similar three-dimensional effects were found to cause hysteresis effects in static I-V characteristics of large pnpn structures [2].

Finally, IR microscopy demonstrated to be a powerful tool to easily identify latch-up sites and to correlate latch-up current distribution with electrical characteristics. Its capability of observing latch-up paths through the Si substrate also when parasitic structures are covered on top by metal lines represents a fundamental advantage for the observation of CMOS VLSI circuits with tight design rules.

REFERENCES

[1] E. Zanoni, M. Giannini, A. Senin, G. Simeone, C. Canali, "Observation of latch-up time evolution in CMOS IC's by means of stroboscopic voltage contrast techniques", IEEE Journal of Solid-State Circuits, Vol. SC-22 (1), 1987, pp. 65-70.

[2] L. Selmi, F. Venturi, E. Sangiorgi and B. Ricco', "Three dimensional distribution of latch-up current in scaled CMOS structures", in P. U. Calzolari, G. Soncini (Eds.), Proc. of ESSDERC'87, Bologna 1987, pp. 783-786.

[3] F. M. Coppage, D. J. Allen, P. V. Dressendorfer, A. Ochoa, J. Rauchfuss, T. F. Wrobel, "Seeing through the latch-up window", IEEE Trans. on Nucl. Science Vol. NS-30, 1983, pp. 4122-4126.

[4] A. H. Johnston and M. P. Baze, "Mechanisms for the latch-up window effect in integrated circuits", IEEE Trans. on Nucl. Science, Vol. NS-32 (6), 1985, pp. 4018-4025.

[5] M. Khurana, "Pulsed Infra-red Microscopy for debugging latch-up in CMOS products", IEEE 22nd Ann. Proc. Int. Reliability Phys. Symp., 1984, pp. 122-127.

[6] R. R. Troutman, "Latch-up in CMOS technologies", Kluwer, Boston, 1986.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988 C4-813

ELECTRICAL PERFORMANCES COMPARISON OF SEMI AND FULLY RECESSED ISOLATION STRUCTURES

E. DUBOIS[1], J.-L. COPPEE*,[2], B. BACCUS and D. COLLARD

ISEN, Institut Supérieur d'Electronique du Nord, CNRS-LA 253, 41, Bd Vauban, F-59046 Lille Cedex, France
**Université Catholique de Louvain, Laboratoire de Microélectronique, Place du Levant 3, B-1348 Louvain-la-Neuve, Belgium*

Résumé - Les caractéristiques de structures d'isolation LOCOS semi-enterré et LOCOS enterré ont été analysées. En particuliers, les simulations bidimensionnelles couplées de technologie et de dispositif, réalisées avec le logiciel IMPACT, ont démontré l'influence de l'interface Si/SiO_2 sur les performances d'isolation. Un excellent accord simulation-expérimentation a été trouvé pour les structures étudiées.

Abstract - Electrical characteristics of semi-recessed and fully-recessed LOCOS isolation techniques are analyzed. Two-dimensionnal process/device simulations, performed with the IMPACT package, have demonstrated the effect of the Si/SiO_2 interface shape on isolation efficiency. An excellent agreement has been found between simulations and measurements for the studied structures.

I - INTRODUCTION

Device isolation presents critical aspects for circuit packing density in VLSI. LOCal Oxidation of Silicon (LOCOS) is classically limited by the bird's beak extension and lateral diffusion of the channel stop region into active transistor area. Several technological alternatives were proposed to control lateral growth of oxide, such as SILO technique (Sealed Interface Local Oxidation) by using successive silicon nitride and deposited oxide layers /1,2/. In order to evaluate a new isolation process, numerous technological and device physics aspects have to be investigated to clarify important two dimensional (2D) parasitic effects /3/. So the use of coupled process/device simulations is of prime necessity. This paper proposes a comparison of electrical performances for both semi and fully-recessed isolation structures by means of process-device simulations and experimental results. A special attention has been devoted to the bird's beak shape description to demonstrate the electrical implications induced by geometrical effects. An advanced version of the IMPACT3 device simulator extended for isolation structure modeling linked to the IMPACT2 /4/ process simulator was extensively used for two dimensional electrical analysis.

II - PROCESS DESCRIPTION

Table I outlines the main characteristics of the isolation stuctures that has been investigated. SEM cross-sectional views of both structures are shown in figure 1 . For each of them, two versions have been fabricated:
- an aluminium gate version for which a 0.5 micron poly-silicate-glass (PSG) layer was deposited prior the aluminium level. (figure 2)
- a polysilicon gate version. In this case, the n⁺ diffusion regions are self-aligned to the polysilicon. A 550 A° thick oxide layer of about 3 microns long exists between the field oxide mask and the edge of the poly mask. (figure 3)

III - ISOLATION PROPERTIES

The major difference between the semi-rox and the full-rox structure is the bird's beak shape. Because the junctions are shallower than the planar part of the

[1] Subsidized by Centre National de la Recherche Scientifique - CNRS
[2] Supported by l'Institut pour l'encouragement de la recherche scientifique dans l'industrie et l'agriculture - IRSIA

Si/SiO_2 interface, the oxide corners are incorporated into the channel region. Figures 4 and 5 show the potential variations along the interface for both structures with a polysilicon gate. Differents barrier deformations are observed when the gate bias affects sufficiently the interface potential distribution. Indeed, the potential barrier of the fully-recessed structure is then located under each bird's beak corner, but remains spread over the whole field oxide interface in the case of the semi-recessed structure. Because the current in the overall structure is controlled by electron emission over the potential barrier at the source side, the fully-recessed device shows better isolation properties than the semi-recessed one. This point is illustrated by the ID-VG characteristics for both aluminium and polysilicon gates (resp. figures 6 and 7). Notice that an excellent agreement is found between measured and calculated drain currents that validates the simulation approach.

Further simulations have been carried out on 2 and 7 microns long aluminium gate devices without the PSG layer, for a 5 volts drain bias condition. Characteristics plotted in figure 8 demonstrates the sensitivity of the semi-recessed isolation structure to the n^+ to n^+ spacing (L). By opposition, the fully-recessed structure does not reveal this typical 1/L dependence in the subthreshold mode of operation. Indeed, the effective channel length is in this case reduced to the curved part of the field oxide interface and makes the whole parasitic transistor dominated by the two corner devices.

IV - SENSIBILITY TO DIBL EFFECT

The simulations performed on a 2 microns long device have allowed to analyze the effect of Drain Induced Barrier Lowering (DIBL) combined to the corner effect. The drain bias modulates the potential distribution around the n^+ diffusion and first affects the potential barrier in the corresponding region. As the curved part of the interface located on the drain side requires more band bending to invert because of the influence of the drain controlled depletion region, the channel is firstly formed in the source region (figure 9). Figure 10 illustrates the potential distribution along the interface for gate bias varying from 0 to 15 volts. This result shows, once again, how the newly located potential barrier leads to a better immunity against the DIBL effect, for the full-rox isolation structure.

CONCLUSION

The presented results obtained from coupled process-device analysis have outlined the sensitivity of isolation structure to the field oxide shape. As the corner effect provides better characteristics for threshold voltage, subthreshold slope and drain voltage dependence, buried oxide isolation structures with sharpened corners appear as a valuable solution for micron and submicron technologies. Moreover, the experimental characteristics of the studied structures have revealed to be in good agreement with two dimensional numerical modeling, and have proved the capabilities of the IMPACT simulation tool to investigate physics of new isolation structures.

REFERENCES

/1/: J. HUI et al., IEDM Technical Digest, pp 220-223, Dec. 1982
/2/: S. SAWADA et al., IEEE Trans. Electron. Devices, vol ED-32, 1985
/3/: J.L. COPPEE and P. VAN DE WIELE, Solid-State Electronics, vol 31, 1988
/4/: D COLLARD and K. TANIGUCHI, IEEE Trans. Electron. Devices, vol. ED-33, 1986
/5/: S. GOODWIN and J. PLUMMER, IEEE Trans. Electron. Devices, vol. ED-31, 1984

FIGURE 1 : SEM CROSS-SECTIONAL VIEWS OF THE FIELD OXIDE EDGE
SEMI-RECESSED STRUCTURE FULLY-RECESSED STRUCURE

TABLE I

	Semi-Rox Structure	Full-Rox Structure
Field Implant	2 10^{13} cm^{-2} at 25 Kev	2 10^{13} cm^{-2} at 25 Kev
Field Oxidation	1000° C 220 mn	950° C 370 mn
Oxide thickness	716 nm	735 nm
S/D implant	5 10^{15} cm^{-2} at 140 Kev	5 10^{15} cm^{-2} at 140 Kev
Annealing	950° C 45 mn	950° C 45 mn

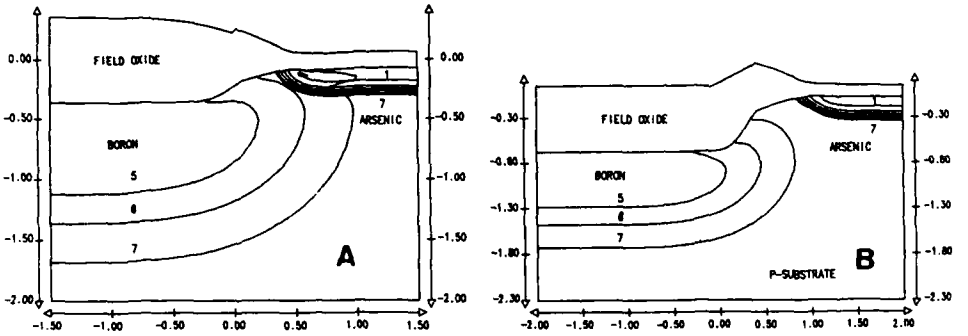


FIGURE 2 : ALUMINIUM GATE STRUTURES 2-D IMPURITIES EQUI-CONTOURS LINES
(A) SEMI-RECESSED STRUCTURE (B) FULLY-RECESSED STRUCTURE

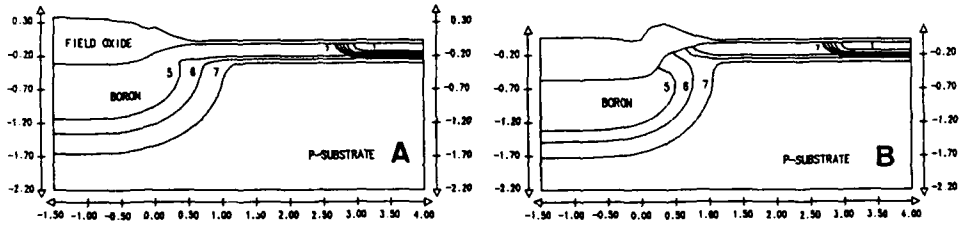


FIGURE 3 : POLYSILICON GATE STRUCTURES 2-D IMPURITIES EQUI-CONTOURS LINES
(A) SEMI-RECESSED STRUCTURE (B) FULLY-RECESSED STRUCTURE

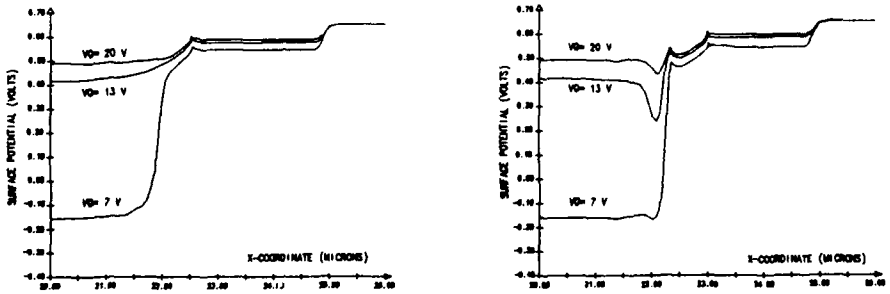


FIGURE 4 : SEMI-ROX STRUCTURE FIGURE 5 : FULL-ROX STRUCTURE
INTERFACE POTENTIAL DISTRIBUTION FOR POLYSILICON GATE DEVICES ON DRAIN SIDE
V_{Drain} = 50 mVolts , V_{Gate} = 0 Volt , V_{Source} = 0 Volt , W/L= 100/20

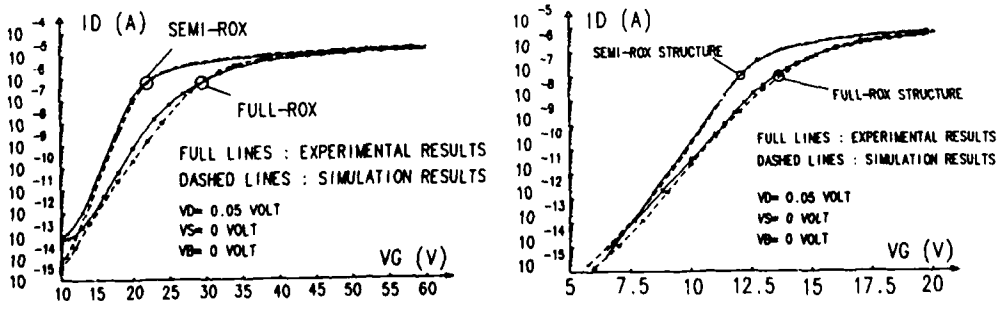


FIGURE 6 : ALUMINIUM GATE VERSION FIGURE 7 : POLYSILICON GATE VERSION
SIMULATED AND MEASURED ID-VG CHARACTERISTICS
W/L= 100/20

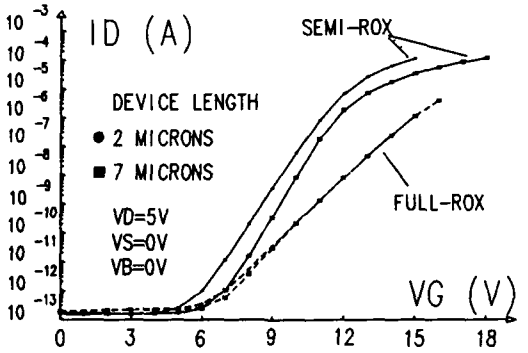


FIGURE 8 : CALCULATED ID-VG CURVES FOR DIFFERENT DEVICE LENGHTS
V_{source} = 0 Volt , V_{bulk} = 0 Volt , V_{drain} = 5 Volts
W = 100 Microns

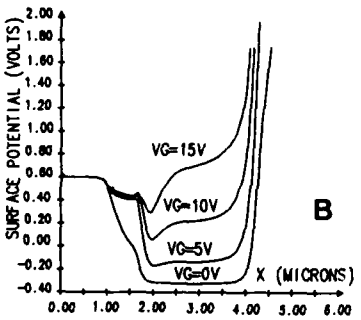


FIGURE 9 :
POTENTIAL DISTRIBUTION
ALONG THE INTERFACE

V_{source} = 0 VOLT

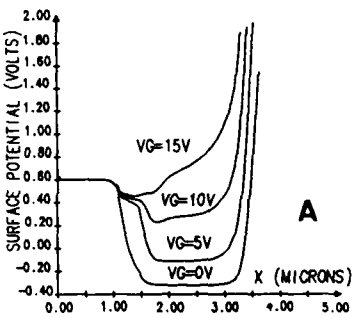
V_{bulk} = 0 volt

V_{drain} = 5 volts

V_{gate} = 0,5,10,15 volts

L= 2 microns

(A) SEMI-ROX STRUCTURE
(B) FULL-ROX STRUCTURE



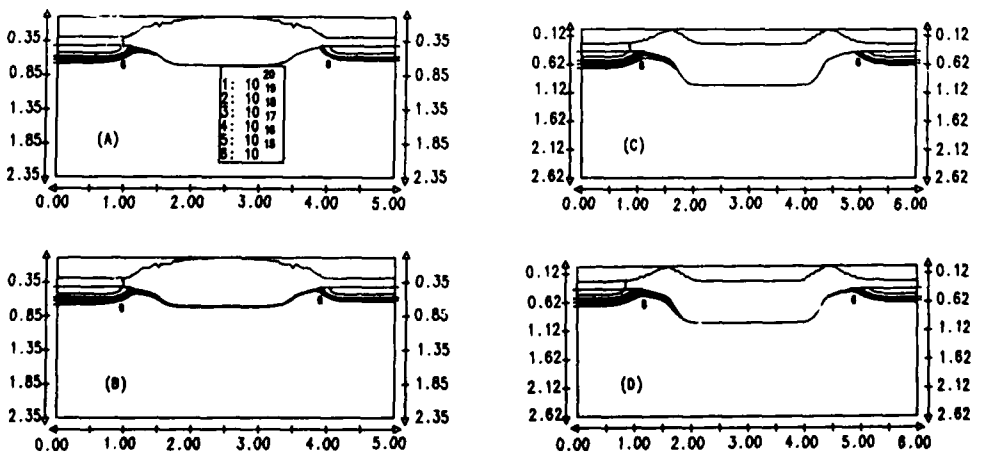


FIGURE 10: ELECTRON DENSITY DISTRIBUTION DURING CHANNEL FORMATION
V_{source} = 0 volt , V_{bulk} = 0 volt , V_{drain} = 5 volts

SEMI-ROX STRUCTURE : (A) V_{gate} = 5 volts , (B) V_{gate} = 15 volts
FULL-ROX STRUCTURE : (C) V_{gate} = 5 volts , (D) V_{gate} = 15 volts

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

A NEW METHOD FOR THE EXTRACTION OF MOSFET PARAMETERS AT AMBIENT AND LIQUID HELIUM TEMPERATURES

F. BALESTRA, I. HAFEZ and G. GHIBAUDO

Laboratoire de Physique des Composants à Semiconducteurs (CNRS-UA), ENSERG/INPG, 23, Av. des Martyrs, F-38031 Grenoble Cedex, France

Résumé - Une méthode originale pour l'extraction des paramètres des Transistors MOS est présentée en fonction de la température. Cette méthode, qui repose sur une combinaison des caractéristiques de transfert donnant le courant de drain et la transconductance, permet d'obtenir des valeurs satisfaisantes pour les paramètres de tension de seuil et de mobilité.

Abstract - An original method for MOSFET parameter extraction is presented as a function of temperature. This method, which relies on combining drain current and transconductance transfer characteristics, enables reliable values of the threshold voltage and mobility parameters.

1 - INTRODUCTION

At ambient temperature, different methods for the extraction of the MOSFET parameters have been proposed [1-3]. These methods are based primarily on the premise that the threshold voltage V_t can be obtained from extrapolating the linear region of the transfer characteristics $I_d(V_g)$ to zero drain current, or by taking the low field mobility μ_o equal to the maximum field effect mobility. At very low temperature, no reliable extraction method based on satisfactory mobility and current modelling has been proposed up till now. The aim of our method is to present a new parameter extraction which avoids the effect of mobility reduction with gate voltage on the determination of both the threshold voltage and low field mobility parameter. The extraction is conducted, as is usual, within the strong inversion regime of the MOSFET linear region and relies on the drain current I_d and normalized transconductance g_m ($g_m = (1/V_d).dI_d/dV_g$) expressions [4,5].

2 - EXTRACTION METHOD AT AMBIENT TEMPERATURE

a) Theory

The basic idea of our method consists of constructing a function which eliminates the influence of the mobility reduction with gate voltage (At 300K the effective mobility is given by : $\mu_{eff} = \mu_o / [1 + \theta (V_g - V_t)]$). This is achieved by the following function :

$$I_d/(g_m)^{1/2} = ((W/L)\, C_{ox}\, \mu_o\, V_d^2)^{1/2}.(V_g - V_t) \ ,$$

where W and L are the effective channel width and length respectively, C_{ox} is the gate oxide capacitance, θ is the mobility reduction coefficient, μ_o is the low field mobility, V_t is the charge threshold voltage, V_g and V_d are the gate and drain voltages respectively.

Since this quantity ($I_d/g_m^{1/2}$) is linear with the gate voltage, we can easily extract the real values of charge threshold voltage (V_t) and low field mobility (μ_o) from the intercept and slope respectively. It is worth emphasizing that the present method allows a separate determination of V_t and μ_o which is not the case in the previous works [1-3]. Moreover, it is worth noting that the charge threshold voltage must not be confused with the extrapolated threshold voltage V_{text} of the linear region of $I_d(V_g)$ to zero drain current. Indeed, when the mobility reduction factor θ is not equal to zero, V_{text} is smaller than the charge threshold voltage V_t [4]. V_{text} is strictly equal to V_t only when the mobility is gate voltage independent.

The mobility reduction factor θ can be determined by the following expression, which is constant in strong inversion :

$$\theta = [\ I_d/(\ g_m\, V_d\ \ (V_g - V_t)\ ) - 1\ ]/(V_g - V_t) \ .$$

b) Experiment

Measurements have been carried out on n-channel MOSFETs fabricated at the LETI/LIR with a technology adapted for low temperature operation, with a surface channel doping of 1.5×10^{15} cm^{-3}, a gate oxide thickness of 120 nm, and different geometries. Therefore, the threshold voltages are very low at ambient temperature (around 0V) and become higher at low temperature. Fig. 1 (a) and (b) show typical plots of the drain current $I_d(V_g)$ and the field effect mobility $\mu_{FE}(V_g)$ (which is obtained from the normalized transconductance by : $\mu_{FE} = g_m/((Z/L).C_{ox})$. Note the sub-linear dependence of I_d above threshold due to the decrease of the mobility at high gate voltages. Fig. 1 (c) presents the corresponding plot of the $I_d/(g_m)^{1/2}$ characteristics. As expected, a linear dependence with gate voltage is obtained. The previous method allows the determination of the ambient temperature MOSFET parameters : V_t = -0.035 V, μ_o = 1236 $cm^2/V.s$, and θ = 0.039 V^{-1}. Moreover, we can observe that V_t is higher than the extrapolated threshold voltage (V_{text} = -0.065 V), and μ_o is greater than the maximum field effect mobility (μ_{FEmax} = 1214 $cm^2/V.s$). This method has also been successfully applied at liquid nitrogen temperature, demonstrating the validity of the previous relations for this temperature range.

3 - EXTRACTION METHOD AT LIQUID HELIUM TEMPERATURE

a) Theory

At very low temperature (4-40K), the conventional relations are no longer valid, and thus the previous extraction method cannot be applied. Indeed, the effective mobility

of an inversion layer at very low temperature has a bell-shaped behaviour [5] with inversion charge Q_i ($\mu_{eff} = 2\ \mu_m / (Q_i/Q_m + Q_m/Q_i)$, where μ_m is the maximum mobility and Q_m is the corresponding inversion charge). Therefore, the effective mobility in strong inversion can be expressed by using the classical relation of the inversion charge ($Q_i = C_{ox}\ (V_g-V_t)$) as :

$\mu_{eff} = 2\ \mu_m\ \theta'\ (V_g-V_t)/[\ 1 + \theta'^2\ (V_g-V_t)^2\ ]$, where $\theta' = C_{ox}/Q_m$ is the mobility parameter which characterizes the gate voltage dependence of μ_{eff} at low temperature.

As for ambient temperature, the basic idea is to eliminate the effect of the effective mobility dependence on gate voltage. This is achieved by using the following function, obtained by combining the drain current and normalized transconductance which are expressed using the low temperature mobility law :

$$(I_d)^{2/3}/(g_m)^{1/3} = (\ (W/L)\ C_{ox}\ \mu_m\ \theta'\ V_d^2\ )^{1/3}.(V_g-V_t)\ .$$

This quantity provides both the charge threshold voltage V_t and the maximum mobility μ_m, since θ' can be determined from the maximum of g_m given by :

$$V_{gmax} = V_t + 1/(\ \sqrt{3}\ \theta')\ .$$

b) Experiment

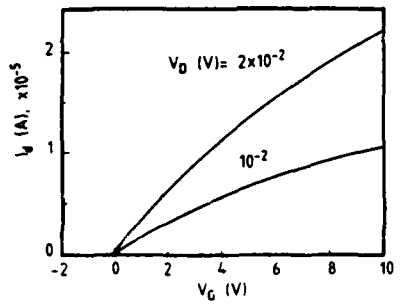
Fig. 2 shows the transfer characteristics, the field effect mobility and the $I_d^{2/3}/g_m^{1/3}$ function. We can observe the expected linear dependence of this function on the gate voltage which confirms the validity of the model used in the parameter extraction. The previous method enables the determination of the low temperature MOSFET parameters : V_t = 0.171 V, θ' = 0.115 V^{-1}, and μ_m = 4589 $cm^2/V.s$. Two interesting points stand out : V_t is lower than the extrapolated threshold voltage (V_{text} = 1.78 V), and μ_m is lower than the maximum field effect mobility (μ_{FEmax} = 5600 $cm^2/V.s$), which is not the case for ambient temperature ($V_t > V_{text}$ and $\mu_o > \mu_{FEmax}$).

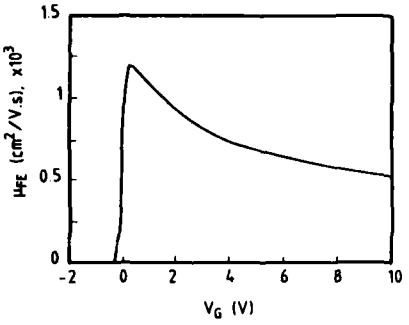
4 - CONCLUSION

An original method of MOSFET parameter extraction has been proposed to avoid the effects of mobility reduction with gate voltage on the determination of both the threshold voltage and the mobility parameters. Different methods have been successfully tested for ambient, nitrogen and helium temperature ranges.

REFERENCES

[1] De la MONEDA F.H. et al, IEEE Electron Devices Letters, EDL-3, p.10 (1982)
[2] CHU HAO et al, Solid-State Electronics, 28, p. 1025 (1985)
[3] KRUTSICK T.J. et al, IEEE Trans. Electron Devices, ED-34, p. 1676 (1987)
[4] GHIBAUDO G., Phys. Status Solidi (a), 95, p. 323 (1986)
[5] GHIBAUDO G. and F. BALESTRA, Solid-State Electronics, 31, p. 105 (1988)





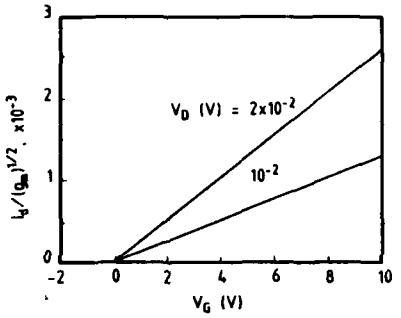
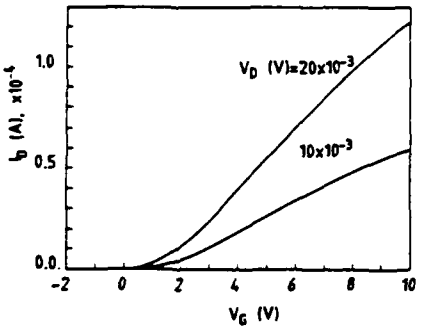
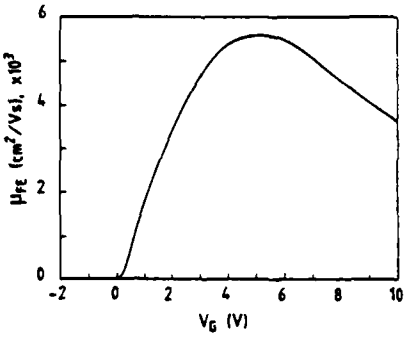


Fig. 1 : Ambient temperature
(a) Transfer characteristics,
(b) field effect mobility, and
(c) $I_d/(g_m^{1/2})$ function.





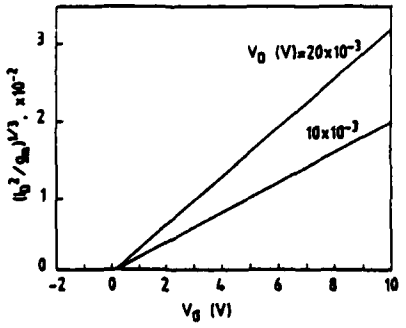


Fig. 2 : Liquid helium temperature
(a) Transfer characteristics,
(b) field effect mobility, and
(c) $(I_d)^{2/3}/(g_m)^{1/3}$ function.

JOURNAL DE PHYSIQUE
Colloque C4, supplément au n°9, Tome 49, septembre 1988

SUBSTRATE AND DEFECT INFLUENCES ON THE POSITION-RESPONSE LINEARITY OF POSITION-SENSITIVE DETECTORS

Y.Z. XING

Faculty of Electrical Engineering, Delft University of Technology, PO Box 5031, NL-2600 GA Delft, The Netherlands

Abstract — The performance of position—sensitive detectors (PSD's) with an epilayer was investigated. Defects which are responsible for a great position—response nonlinearity in the p—type resistive layer were fond. The influence of the substrate potential, the photocurrent generation in the junction between the epilayer and the substrate, and the dependency of the resistive layer thickness on the local potential were analyzed. By connecting the substrate to the same potential as the p—type layer, distortions in the position—response grid are reduced to a certain degree. A simple method was used to discover this kind of defect in an early stage.

1. — Introduction

In recent years, the commercially developed two—dimensional position—sensitive detectors (PSD's) have been fabricated with a diffusion or an ion—implanted layer on a substrate. In 1979, Noorlag /1/ presented a two—dimensional PSD which uses an epilayer as the second resistive layer. PSD's based on this structure have the virtues of providing better position—response linearity because of the better homogeneity of the epilayer, perfect electrical isolation along the edges through deep p diffusion and avoidance of process steps on the back of the wafer so that the standard process can be implemented. But introducing an epilayer into the PSD also causes new problems, as defects in the epilayer in the form of stacking fault were discovered. The effect on the performance of the PSD was discussed in an early paper /2/. In the investigations following it was found that the photocurrent generation in the lower junction and the substrate potential have a great influence on the PSD. The dependency of the resistive layer thickness on the local potential is discussed. The analysis results are also applicable to one—dimensional PSD's with on—chip signal processing circuit /3/.

2. — Variation of the resistive layer thickness

In 1960, Lucovsky /4/ derived a differential equation describing the lateral photoeffect in the PSD, which lately has been adopted by many in their theoretical studies on PSD's with different configurations /1,5/. However, when the dependency of the thickness of the resistive layer on the lateral photovoltage is taken into account, the basic differential equation can be derived as:

$$\nabla(\sigma_s \nabla\phi) = J_{pn} \qquad \text{(p—type layer)}$$

$$\nabla(\sigma_s \nabla\phi) = -J_{pn} \qquad \text{(n—type layer)}$$

where J_{pn} is the total current density through the junction and σ_s the sheet conductivity of the resistive layer. If σ_s is constant, Lukovsky's equations in a steady—state are obtained.

The influence of the resistive—layer thickness variation is strong in one—dimensional PSD's which are designed with a large length to width ratio and operated with voltage outputs/3/. A computer simulation result is shown in Fig.1. A 0.05 % nonlinearity in the position response will occur for a p—type layer with a sheet resistivity of 3 kΩ and a length to width ratio of 20. From Fig.1 we can also

see that the nonlinearity is an independent function of the light beam movement in the direction parallel with the electrodes, so that a simplified one—dimensional model can be used for the nonlinearity calculation. The maximum measurement error Δx obtained by the one—dimensional model is given by:

$$\Delta x = \frac{w_d V_{max}^2}{w_0 I_{ph} R_p V_{pn}} l$$

in which w_d is the thicknesses of the depletion layer in the corresponding resistive layer, w_0 the resistive layer thickness, V_{max} the maximum voltage, I_{ph} the photocurrent, R_p the resistance of the resistive layer, V_{pn} the revers—bias voltage, l the length of the device.

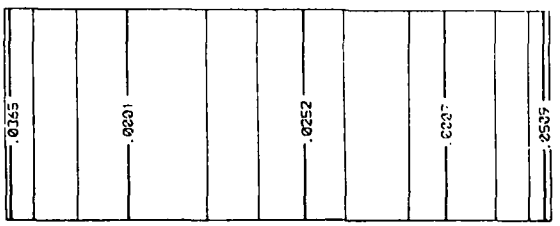


Fig.1 — Simulation results for a one—dimensional PSD (figure is not to scale). The lines in the grid represents the loci of the light beam with equal nonlinearity in percentage.

3. — PSD's with one defect

In the analysis it is assumed that the size of the defect is small so that it can be considered as an extra current generating or absorbing source with a fixed position. For a PSD with one defect the light—spot position x_m measured by the detector becomes then:

$$x_m = \frac{I_2 - I_1}{I_2 + I_1} l = \frac{I_{ph}(2x_g/l - 1) + I_{def}(2x_d/l - 1)}{I_{ph} + I_{def}} l$$

where I_{ph} is the photocurrent generated in the upper junction, I_{def} the current through the defect, x_g and x_d are the positions of the light beam and the defect on the x—axis respectively. The current I_{def} is defined as positive when it flows from the substrate to the p—type layer. The effective length of the resistive layer l_{eff} and the shift of the zero point Δx now are obtained:

$$l_{eff} = \frac{I_{ph}}{I_{ph} + I_{def}} l \qquad \Delta x = \frac{I_{def}(2x_d/l - 1)}{I_{ph} + I_{def}} l$$

which shows that the measured effective layer length and the zero point of the PSD are changed.

For PSD's with one defect and a floating substrate, as shown in Fig.2(a), the value of I_{def} is always positive since the photocurrent generated in the lower junction can only flow to the resistive layer through the defect. Therefore the measured effective layer length is smaller than the length of the real resistive layer. When the electrode on the substrate is held at the same potential as the electrodes on the p—type layer, as shown in Fig.2(b), the photocurrent generated in the lower junction flows directly to the electrode on the substrate. Because the local potential in the p—type layer is always higher than that in the substrate due to the lateral photoeffect, I_{def} is negative. This means a prolonged effective layer length. The measurement results of a PSD with one defect is shown in Fig.3. The reduction of the measured effective layer length is 30 %, which agrees with the measurement showing that the same percent of photocurrent is coming from the lower junction. From Fig.3(b) we can see that by connecting the substrate to the same potential as the upper electrodes the effective layer length is longer than when the substrate is floating, but it is still smaller than the real layer length. This might be due to contact resistance of electrode on the substrate, so that there is still a portion of photocurrent generated in the lower junction flowing to the p—type resistive layer.

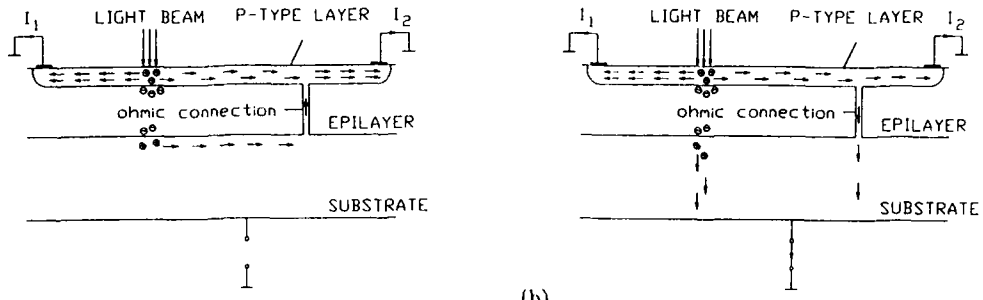


Fig.2 An illustration of a PSD with one defect: a) the substrate is floating and b) the substrate is held at the same potential as the p—type layer.

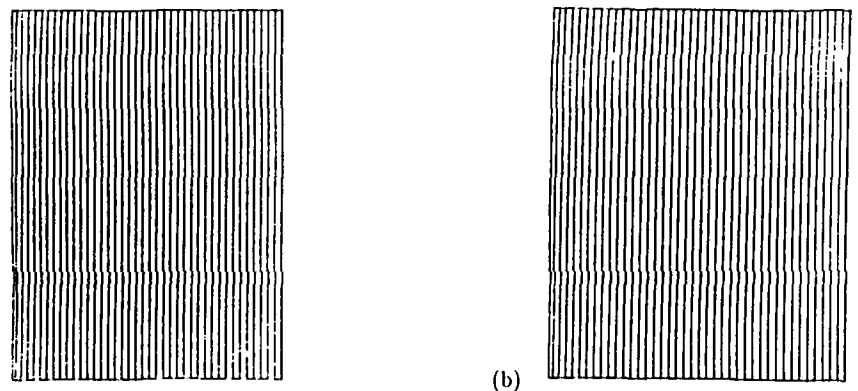


Fig.3 Measurement results. (a) the substrate is floating; (b) the substrate is held at the same potential as the p—type layer.

4. — PSD's with two defects

Figure 4 depicts a PSD with two defects. In the case of a floating substrate the defects, together with the substrate, form a current path. When the light beam is moved close to one of the defects, say at a position on the right—hand side of the left defect, as shown in Fig.4(a), the photocurrent, which was destined for the electrode on the left, will be diverted to the substrate through that defect, flow back to the resistive layer through the other and be finally collected by the electrode on the right. This causes a shift in the measured light—spot position to the right. The influence of this current diversion on the position—response grid is smaller when the light beam is moved to the left of the left defect. Because in that case, although taking a different path, the photocurrent which is diverted to the substrate by the defects still flows to the electrode it ought to, *i.e.* the electrode on the right. This explains the directional deformation of lines in the measured position—response grid. A computer simulation result is shown in Fig.5(a). When the substrate is put on a fixed potential, as shown in Fig.4(b), the diverted current does not flow back to the resistive layer. Therefore, the directional deformation should disappear, as shown in Fig.5(b). In practice the measurement result shows only a slight decrease in the deformation, which may be caused by the contact resistance on the substrate so that there is still a portion of photocurrent which could flow to the opposite electrode.

5. — Measurement results

Measurements were carried out on a large number of PSD's (total 80) and PSD's fabricated without epilayer (total 8). It was verified that the existence of defects in the PSD's can be revealed by carrying out electrical measurement instead of optical as this kind of defect causes a large resistance deviation (7.6%) of p—type layers over the whole wafer, and an ohmic characteristic in the IV—curve measured between the p—type layer and the substrate. The extent of distortion is closely related to the value of the resistance which varies from several kilo ohms to hundreds of kilo ohms.

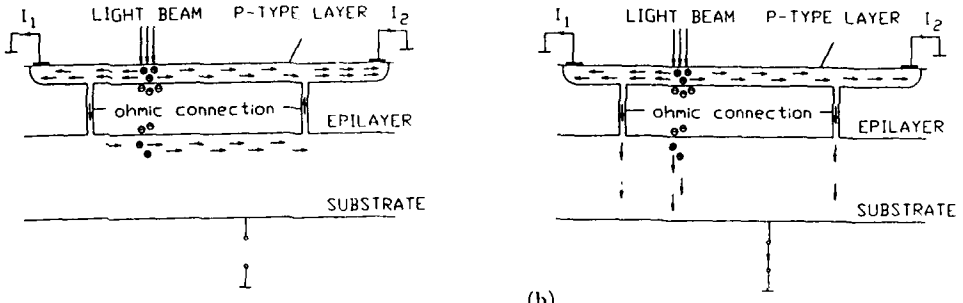


Fig.4 A schematic illustration of a PSD with two defects: a) the substrate is floating and b) the substrate is held at the same potential as the p—type layer.

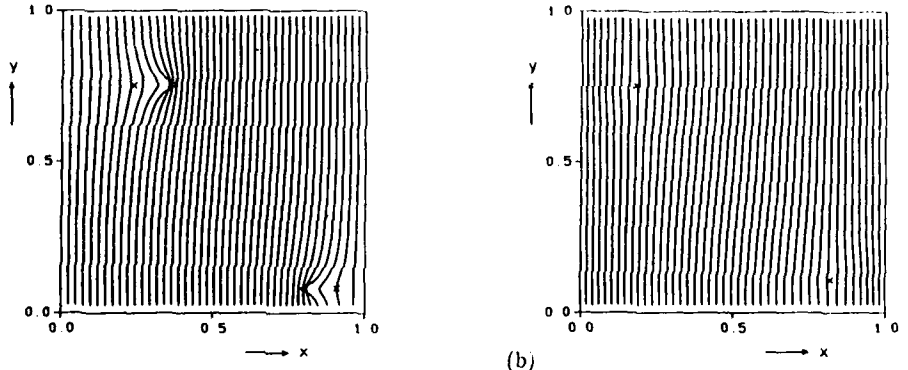


Fig.5 Simulation results of the PSD with two defects. (a) the substrate is floating, (b) the substrate is held at the same potential as the p—type layer.

6. — Conclusions

Variations in the resistive layer thickness affects the position—response linearity of the PSD when the PSD is operated with a voltage output. The influence is proportional to the length—to—width ratio of the device. When defects in the form of stacking fault are present in the epilayer, photocurrent generation in the lower junction reduces the effective length of the resistive layer. By holding the substrate at the same potential as the p—type resistive layer this influence is reduced. Therefore it is recommended to operate PSD's with the lower junction reverse biased. Besides measurements on the leakage current and resistance distribution of resistive layers, the current—voltage characteristic between the p—type layer and the substrate can be used as a criteria for qualifying PSD's. Defects of the kind described above can be discovered in an early stage when the current—voltage characteristic shows an ohmic character. This makes the usually time—consuming optical measurement unnecessary.

REFERENCES

/1/ D.J.W. Noorlag and S. Middelhoek, "Two—dimensional position—sensitive photodetector with high linearity made with standard I.C.—technology", IEE J. Solid—state and Electron Device, vol.3, pp.75—82, 1979.

/2/ H. Yashiro, Y.Z. Xing and S. Middelhoek, "Influence of surface defects on the performance of Si position—sensitive detectors", Proc. of 4th International Conference on Sensors and Actuators, pp.423—426, Tokyo, 1987.

/3/ Y.Z. Xing, "A simple light—spot position detector with a light—intensity variation—insensitive output", Proc. EUROSENSORS I, pp. 227—228, Cambridge, 22—24, September 1987.

/4/ G. Lucovsky, "Photoeffects in nonuniformly irradiated p—n junctions", J. Appl. Phys. vol.31, pp.1088—1095, 1960.

/5/ Y.Z. Xing and C.P.W. Boeder, "A new angular—position detector utilizing the lateral photoeffect in Si", Sensors and Actuators, Vol.7, pp.153—166, 1985.

AUTHOR INDEX