MICROELECTRONIC ENGINEERING

An International Journal of Semiconductor Manufacturing Technology

Including Nanoelectronic Engineering

∧nfos'99

Proceedings of the
11th biennial conference on
Insulating Films on Semiconductors

June 16-19, 1999
Kloster Banz, Germany

19991123 118

Edited by

Max Schulz

Rolf Brendel

COMPLETE VOLUME

MICROELECTRONIC ENGINEERING

Aims and Scope

The aim of *Microelectronic Engineering* is to bring together in one publication the results of European, American and Japanese work in the rapidly expanding field of integrated microelectronic, optic and optoelectronic devices. *Microelectronic Engineering* is an archival, peer-reviewed journal. It publishes full research papers, short communications, and review articles. The scope of this Journal includes materials, methods and designs for microfabrication, processing and inspection for microelectronic and optoelectronic elements from centimeters to nanometers.

The wide range of topics covered by the journal include:

1. *Lithography*
- optical lithography (submicron dimensions, deep u.v. lithography)
- electron optical methods and systems
- X-ray optical methods and systems
- resists
- limits

2. *Pattern Transfer*
- ion technology
- plasma etching
- self-organisation

3. *Materials*
- metallization
- microstructuring growth processes
- silicon on insulators
- compound semiconductors
- dielectrics (low K and high K)
- interconnect architectures

4. *Inspection and Testing*
- electron beam testers
- laser probes
- signal and image processing
- metrology

5. *Advanced Processing*
- process integration
- three dimensional integration
- defect free processes
- manufacturing science
- rapid thermal processing
- process modelling
- equipment modelling
- laser assisted processing

6. *Micromechanical Structures*
- fabrication
- design

7. *Advanced Devices*
- nanometer structures
- dimension-sensitive devices properties
- effect of small scaling

Journal home page: www.elsevier.nl/locate/mee
This allows access to manuscripts status system for authors.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 1999	3. REPORT TYPE AND DATES COVERED Conference Proceedings

4. TITLE AND SUBTITLE

Insulating Films on Semiconductors

5. FUNDING NUMBERS

F61775-99-WF057

6. AUTHOR(S)

Conference Committee

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Erlangen
Staudtstrasse 7
Erlangen D-91058
Germany

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

EOARD
PSC 802 BOX 14
FPO 09499-0200

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

CSP 99-5057

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

A

13. ABSTRACT (Maximum 200 words)

The Final Proceedings for Insulating Films on Semiconductors, 16 June 1999 - 19 June 1999

This is an interdisciplinary conference. Topics include dielectrics, semiconductor/insulator interfaces, and the physics and technology of field effect (MIS, MOS) semiconductor devices, atomic scale processing, characterization and measurement techniques, radiation and hot carrier reliability, novel materials, and processing technologies. Special emphasis will be given to ultrathin oxides on silicon.

14. SUBJECT TERMS

EOARD, Semiconductors, Semiconductor Growth

15. NUMBER OF PAGES

447

16. PRICE CODE

N/A

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

MICROELECTRONIC ENGINEERING

Proceedings of the
11th biennial conference on
Insulating Films on Semiconductors

June 16-19, 1999
Kloster Banz, Germany

Edited by

Max Schulz
Rolf Brendel

AQF00-02-0494

Infos'99

Proceedings of the
11th biennial conference on
Insulating Films on Semiconductors

June 16-19, 1999
Kloster Banz, Germany

Edited by

Max Schulz
Rolf Brendel

Volume 48/1–4
1999

ELSEVIER

Abstracted/indexed in:

Cambridge Scientific Abstracts
Current Contents
Engineering Index
INSPEC
Science Citation Index

Contents

Traps and degradation

Measurements

Technology

Silicon carbide

Alternative dielectrics

Silicon on insulator

Preface

The INFOS'99 conference on Insulating Films on Semiconductors was organized by the University of Erlangen-Nürnberg at the Kloster Banz Conference Center near Lichtenfels, approximately 80 km north of Erlangen in Germany. This conference was the 11th sequel of the series started 1979 in Durham England. INFOS'99 attracted 130 participants from industry, research institutes, and universities. Attendants were registered from 20 different nations. The biennial European conference series specializing in insulating films on semiconductors will be continued by INFOS'01 to be held in Udine, Italy in the year 2001.

Since INFOS'99 was the last conference in this century of this series specializing on insulating films on semiconductors, this proceedings volume presents a review of the achievements of the MOS technology made in the past forty years and gives an outlook into the future of the next century by two invited talks. The present status of our understanding the insulating properties of films and their interface to semiconductors is surveyed by further 11 invited and 80 contributed papers. Special emphasis was given at this conference to the understanding of the growth, the properties, and the limits of ultrathin gate oxides on silicon; however, sessions devoted to oxide traps, alternative dielectrics, degradation and cleaning of the interface were also included among others. Oxides on silicon carbide were discussed at this conference for the first time. Non-volatile memory and silicon-on-insulator (SOI) properties were discussed in special workshops.

The organizational load of the conference was distributed on many collaborators. The service of the program committee, of the session chair persons, and of many conference attendees who reviewed all the publication manuscripts during the conference, was very much appreciated and made the quick publication of these proceedings possible. Special thanks also go to the coworkers from the Institute of Applied Physics at the University of Erlangen and from the Center of Applied Energy Research (ZAE Bayern) who helped to run the conference. The administration work of the registration and the finances by Mrs. G. Loy is also gratefully acknowledged. The staff of the Hanns Seidel Foundation made the stay at Kloster Banz during the conference very pleasant; this service to the conference is also gratefully acknowledged.

Erlangen, July 1999   M. Schulz   R. Brendel

Conference Chairman:
M. Schulz, University of Erlangen , Institute of Applied Physics,
Staudtstr. 7 , D-91058 Erlangen/Germany

Workshop Organizers:
"Advanced Silicon on Insulator Materials and Devices"
S. Cristoloveanu, LPCS Grenoble, France

"The Physics and Technology of Non-Volatile Memory Devices"
E. Sangiorgi, Universita di Udine, Italy

Program Committee:
V.V. Afanasev, St. Petersburg, Russia
P. Balk, DIMES, Delft, The Netherlands
S. Cristoloveanu, LOCS, Grenoble France
R. Degraeve, IMEC, Leuven, Belgium
W. Eccleston, University of Liverpool, UK
O. Engström, Chalmers University, Sweden
W. Hönlein, Siemens AG, München, Germany
F. Koch, Technical University, München, Germany
F. Kuper, Philips Semiconductors, Nijmegen, The Netherlands
V.G. Litovchenko, Academy of Science, Kiev, Ukraine
J. Morante, Universitat Barcelona, Spain
J. Moreland, Wacker Siltronic, Burghausen, Germany
H. Przewlocki, Institute of Technology, Warsaw, Poland
E. Sangiorgi, Universita di Udine, Italy
M. Schulz, University of Erlangen-Nürnberg, Erlangen, Germany
A. Stesmans, Universiteit Leuven, Belgium
M. Uren, DERA, Malvern, UK
D. Vuillaume, CNRS Institut d'Electronique, Lille, France

International Advisors:
H. Iwai, Toshiba Kawasaki, Japan
D.J. DiMaria, IBM Yorktown Heights, NY USA
H.Z. Massoud, Duke University, Durham NC USA

Local Committee:
M. Schulz, Conference Chair
R. Brendel, Proceedings Editor
G. Loy, Secretary

REVIEW AND OUTLOOK AT THE TURN OF THE CENTURY

40 years MOS technology - from empiricism to science

Pieter Balk

DIMES, Delft University of Technology,
2600 GB Delft, The Netherlands

This paper summarizes the development of the MOS field effect technology from the early beginnings. It will deal with its empirical basis and present some examples of the way the technology progressed: the search for the ideal semiconductor-dielectric system, the continuous battle against impurities and the tenacity with which the problem of the interaction of hydrogen with Si/SiO_2 interfacial defects has defied complete interpretation. A few highlights from the period under discussion will conclude the review.

1. INTRODUCTION

It took 35 years after Lilienfeld's first 1925 patent on a field effect device to turn his idea into a usable structure where a substantial modulation of the conduction could be attained. To reach this point fundamental insight into the characteristics of metals, semiconductors and insulators and into the behavior of contacts between these materials had to be acquired and basic knowledge of their technology had to be obtained. Only on this foundation could the MOS field effect technique live up to its potential and develop over the next 40 years -from 1960 till today- into a technology of great economic importance for the fabrication of ever faster devices of ever smaller dimensions at an ever lower price.

2. EMPIRICAL BASIS

Field effect devices, as described in two patents of Lilienfeld [1,2] and in a later patent by Heil [3] would only become attractive when implemented in the silicon technology. However, in the forties and fifties the development of this technology was directed towards bipolar devices, whereas field effect devices received hardly any attention.

Nevertheless, the bipolar work produced some important insights for our topic of interest. One was the possibility to reduce the high density of interface states (broken bonds) [4], as found underneath the gate electrode on "real" (etched) surfaces, by thermal oxidation [5[. Equally important was separating the conduction of the surface inversion layer from the bulk conduction by using a source and a drain diode. In this manner a hundred percent modulation of this parameter in field effect devices with MOS gate system could be easily obtained. Last, but not least, thermally oxidized silicon had found an important application for local masking in the planar process [6], which is essential for the fabrication of small devices and integrated circuits.

3. FURTHER DEVELOPMENT

The first properly operating MOS field effect devices became available around 1960 [7]. They utilized a 200 nm gate oxide and an Al gate. These devices had stable characteristics and constituted a good starting point for further development of the technique up to the present time.

This development demanded new and more sensitive methods for electrical characterization, as charge pumping [8], to explore the distribution of interface states in the gate area of devices, and RTS (random telepraph signal) studies, which allow electrical characterisation of single interface traps in small devices [9,10]. For atomic identification of defects ESR (electron spin resonance) [11-13] has proven to be a powerful approach.

In the field of material science questions to be dealt with were : Is there an alternative for the Si/SiO$_2$ system? What parameters determine the carrier mobility in the channel? What is the mechanism of carrier transport and trapping in the dielectric in relation to its defect structure? In which ways can this structure be optimized? What is the role of impurities and the micromorphology of the interface in determining the properties of the MOS system? In what manner can these parameters be affected?

In the following we will try to characterize the developments in the MOS field during the past 40 years using three examples : the search for alternatives for the Si/SiO$_2$ system; controlling impurities in SiO$_2$; the interaction between hydrogen and Si/SiO$_2$ interfacial defects. It will not be our intent to give an exact description of our present understanding. We will rather indicate how our insight developed and what complications were encountered along the way.

4. THE CHOICE FOR Si/SiO$_2$

Notwithstanding the potential advantages of other semiconductors, silicon has one very attractive feature : the open flexible structure of its glassy oxide SiO$_2$, thermally grown on crystalline Si, which provides a relatively easy fit to this substrate and thus a low density of interface states [14]. Moreover, its properties are conveniently modified by incorporation of small amounts of dopants like N [15], F [16] or Cl [17]. Only in recent years the rather low value of the dielectric constant along with the small available area in advanced DRAM capacitors has created renewed demand for alternative materials with higher dielectric constant. Unfortunately, with the exception of Si$_3$N$_4$, the alternative dielectrics that may be of interest are polycrystalline, which tends to cause gate leakage. By optimization of the technology and the use of very thin SiO$_2$ interlayers one has learned to live with this problem. Thus, high ε dielectrics like Ta$_2$O$_5$ and TiO$_2$, and ferro-electrics like BaSr-titanate appear to be making the grade [18].

Since the sixties there has been an interest in producing MOS devices on non-Si substrates, like Ge, GeSi, SiC or some III-V compound. The aim would be , for example, obtaining a higher channel mobility or a higher operating temperature (because of a larger bandgap). This requires finding a dielectric which would provide a similarly good fit to the substrate as SiO$_2$ does to Si. For Ge the search for a suitable dielectric has been unsuccessful. For the binary semiconductors one has an extra element to contend with and also in this case break-throughs have not been reported.

5. CONTROLLING IMPURITIES

Metallic impurities generally have a deleterious effect on the properties of MOS structures. The drift in device characteristics caused by Na in the gate oxide may be strongly reduced by trapping this impurity in a thin phosphosilicate glass film grown on top of the thermal oxide, originally (and accidentally) obtained as a "byproduct" of doping the source and drain of n-channel devices in the planar technology. Compromises between polarisibility of the film (which negatively affects its stability) and the maximum protection against Na drift can be made very precisely since the SiO$_2$-P$_2$O$_5$ system has been well characterized [19]. Using this approach, the maximum tolerable Na concentration in production equipment could be allowed to increase to a level which was considered easy to control. Unfortunately, the development towards thinner gate oxides forced to the use of ever thinner and less concentrated phosphosilicate films which ultimately caused the complete loss of protection against Na.

For this reason a second method, the use of a gaseous Cl compound (for example TCE, trichloroethene , C$_2$HCl$_3$) became dominant in the seventies [20]. The chlorine compound not only cleans the furnace tube, but, by injecting it as an additive to the oxidant, is instrumental in substantially reducing the Na concentration in the gate oxide. An added advantage of the Cl-method was that it is simpler to use.

Present day ultra-clean processing techniques, which were introduced around 1990, constitute an integral approach to device fabrication. They are not only concerned with metallic impurities, but also consider the roughening of the surface, which effectively causes the film thickness to vary over the surface, and the deposition of particulate matter and organics. Gaseous Cl compounds are still part of the cleaning process. However, also liquid etching

appears to be an unavoidable step in the ultra-clean technology [21].

6. HYDROGEN AND INTERFACIAL DEFECTS

The interaction between hydrogen and interface states has occupied the attention of the MOS community throughout its history. Only at the present time the problem appears to approach a solution. Part of the difficulty is the omnipresence of hydrogen in MOS structures and the difficulty of determining its concentration and state of bonding [22].

Passivation of silicon dangling bond interfacial traps by reaction with molecular hydrogen [23,24] or atomic hydrogen (formed at the Al-SiO$_2$ interface) [25] at temperatures around 400 ^{0} C was early recognized as an effective way to enhance the transconductance in field effect devices. Heating in forming gas (a mixture of H$_2$ and N$_2$) was originally performed to improve the bond between gate electrode and gate dielectric, and to provide improved electrical contact to source and drain. In later years indications were obtained that under irradiation in MOS systems containing hydrogen also the reverse process, the depassivation of interface traps, could take place. In this case hydrogen would be transported through the oxide and react at the interface as OH$^-$ radicals [26], H atoms [27] or, as a later insight, protons [28].

Experiments, in which samples were exposed to atomic H at different temperatures, indicated the existence of an equilibrium concentration of passivated interface states [29]. It was also concluded from experiments with atomic H that not all interface states generated are of the dangling bond type [30] but that some may have a rather different structure. One possibility is the interfacial defect which has been suggested to occur in UV-irradiated samples, where H would be bonded to a lone electron pair at the bridging O-atom in the Si-O-Si network at the interface [31].The latter structure has been also proposed to be the long known oxidation-induced fixed charge center, which would be located at or near the interface [32]. A complicating factor in this work is that the generation of interface states appears to depend on the detailed structure of the MOS system (thermally grown SiO$_2$; various types of SOI).

7. HIGHLIGHTS

The results of the past 40 years in our field have shown the development of a considerably deeper insight in the details of increasingly complex systems of strongly reduced dimensions (by a factor of 50 for the thickness of gate oxide) during this period [33]. New phenomena were particularly reported during the first 20 years. We will only mention here the effects of quantisation in inversion layers [34] and carrier trapping in dielectrics [35]. The rapidly growing MOS technology provided high quality samples for physical experimentation, which in turn contributed strongly to the development of the field. An excellent demonstration of this development over the second 20 years may be found in the proceedings of the international biennial conference on insulating films on semiconductors (Infos), published since 1979.

A period of 40 years equals approximately the duration of a professional's life; it is about the maximum time span one can report on from direct experience. Many scientists, presently working in the MOS area, have not personally known the workers of the first hour , who made such decisive contributions in shaping the field. As typical representatives of this group I would like you to remember Sir Neville F.Mott (Cambridge) : growth of SiO$_2$ [36] ; Nic Klein (Haifa) : dielectric breakdown [37] ; Ed Nicollian (Murray Hill, N.J. /Charlotte N.C. : hydrogen in SiO$_2$ [38] ; electrical measurement techniques [39]. They passed away, but their work had a strong impact on our field. Their names stand for scientific excellence, and are also typical for the international character of the MOS community.

REFERENCES

1. J.E.Lilienfeld, U.S.Patent 1745175 (1925).
2. J.E.Lilienfeld, U.S.Patent 1900018 (1928).
3. O.Heil, U.K.Patent 439457 (1935).
4. W.Shockley and G.L.Pearson, Phys.Rev. **74**, 232 (1935).
5. M.M.Atalla, E.Tannenbaum and E.J.Scheibner, Bell Syst.Tech.J. **38**, 749 (1959).
6. J.A.Hoerni, IRE Electron Devices Meet., Washington D.C., (1960).
7. D.Kahng, U.S.Patent 3102230 (1960).

8. G.Groeseneken, N.Beltran and H.E.Maes, IEEE Trans. Educ. **31**, 42 (1984).

9. M.J.Kirton and M.J.Uren, Adv. in Phys. **38**, 367 (1988).

10. H.H.Mueller and M. Schulz, J. Appl. Phys. **83**, 1734 (1998).

11. E.H.Poindexter and P.J.Caplan, in Insulating Films on Semiconductors, eds M. Schulz and P.Pensl, (Springer, Berlin, Heidelberg, New York, 1981) p. 150.

12. P.M.Lenahan, Microelectronic Eng. **22**, 129 (1993) (Proc. Infos 1993).

13. A.Stesmans and V.V.Afanas'ev, Appl. Phys. Lett. **72**, 2271 (1998).

14. A.G.Revesz, J. Non-Crystalline Solids **11**, 309 (1973).

15. Y.Okada, P.J.Tobin, P.Rushbrook and W.L.Dettart, IEEE Trans. Electron Dev. **ED-41**, 191 (1994).

16. V.V.Afanas'ev, J.M.M.de Nijs and P.Balk, Proc.ESSDERC '93 (1993) p.415.

17. J.R.Monkowski, R.E.Tressler and J.Stach, J. Electrochem. Soc. **125**, 1867 (1978).

18. P.Balk, J. Non-Crystalline Solids **187**, 1 (1995).

19. P.Balk and J.M.Eldridge, Proc. IEEE **57**, 1558 (1969).

20. G.J.Declerck, Proc. ESSDERC '79 (1979) p.133.

21. M.Meuris et al, Microelectronic Eng. **22**, 21 (1993) (Proc. Infos 1993).

22. A.Revesz, J. Electrochem. Soc. **126**, 122 (1979).

23. E.Kooi, Philips Res. Rep. **20**, 528 (1965).

24. P.Balk, Extended Abstract Electronics Division No.109, E.C.S.Meeting, San Fransisco Ca. May 1965, (The Electrochem. Soc., Pennington N.J. 1965) p.238.

25. P.Balk, Extended Abstract Electronics Division No.111, E.C.S.Meeting, Buffalo N.Y., Oct.1945 (The Electrochem Soc., Pennington N.J. 1965) p.29.

26. C.T.Sah, IEEE Trans. Nucl. Sci. **NS-23**, 1563 (1976).

27. C.M.Svensson, in Proc.Intl.Conf. on Phys. of SiO_2 and its Interfaces, ed. S.T.Pantelides, (Pergamon, New York, 1978) p.328.

28. F.B.McLean, IEEE Trans. Nucl. Sci. **NS-27**, 1651 (1980).

29. L.Do Thanh and P.Balk, J. Electrochem. Soc. **135**, 1797 (1988).

30. E.Cartier, J.H.Stathis and D.A.Buchanan, Appl. Phys. Lett. **63**, 1510 (1993).

31. de Nijs, K.G.Druyf, V.V.Afanas'ev, E. van der Drift and P.Balk, Appl. Phys. Lett. **65**, 2428 (1994).

32. V.V.Afanas'ev and A.Stesmans, Appl. Phys Lett. **72**, 79 (1998).

33. The National Technology Roadmap for Semiconductors : Technology Needs (Semiconductor Industry Association, 1997) p.46.

34. T.Ando, A.B.Fowler and F.Stern, Rev. Mod. Phys. **54**, 437 (1982).

35. D.J.DiMaria, in Proc. Intl. Conf. on Phys. of SiO_2 and its Interfaces, ed. S.T.Pantelides, (Pergamon, New York, 1978) p.160.

36. N.F.Mott, in Insulating Films on Semiconductors, eds. G.G.Roberts and M.J.Morant (The Inst. of Phys., Bristol and London, 1979) p.12.

37. N.Klein and P.Solomon, J. Appl. Phys. **47**, 4364 (1976).

38. E.H.Nicollian, C.N.Berglund, P.F.Schmidt and J.M.Andrews, J. Appl. Phys. **42**, 5654 (1971).

39. E.H.Nicollian and J.R.Brews, MOS Physics and Technology (Wiley, New York, 1982).

OUTLOOK OF MOS DEVICES INTO NEXT CENTURY

Hiroshi Iwai
Tokyo Institute of Technology
4259 Nagatsuta, Midori-ku, Yokohama, 226-8502, Japan
Tel: +81-45-924-5471,   Fax: 81-45-924-5487,   E-mail: h.iwai@ieee.org

Silicon MOSFETs have achieved a great success as the primary component of LSIs for 30 years by the downsizing. However, it is predicted that the downsizing will encounter certain limit in the early period of next century.   This paper describes about the outlook of future of MOS transistors and LSIs in the next century in terms of its predicted limit and possibility of further evolution.

1.   INTRODUCTION

Since D. Kahng and M. Atalla made the first Si-MOSFET in 1960 [1], MOS devices started a great progress.   In early 70's, MOSFETs were integrated in LSI and continued to make remarkable progress for almost 30 years.   The first generation of LSIs, which appeared as products such as a 1Kbit DRAM and a 750 kHz MPU (Micro Processor Unit), has evolved up to 64Mbit DRAMs and 600 MHz MPUs, as shown in Table I.   This evolution has seen the critical dimension (CD) decreases to 1/50, the number of memory bit increase by 256,000 times, and the MPU clock frequency increase by 800 times.   Recent revolutionary progress in the information and communication fields, as represented by the Internet, and the popularization of intelligent mobile devices owes much to this remarkable development of LSI technologies.

Such a tremendous growth of LSIs illustrated in Table II has been achieved by the downsizing of components according to Moore's law [2], which says that CD decreases to 2/3, chip size increases to 3/2, and number of components in a chip quadruples every 3 years, or at every new generation.   Regarding the future, this historical trend is expected to continue according to the SIA (Semiconductor Industry Association) National Technology Roadmap for Semiconductors [3] and its updated version – International Technology Roadmap[4] –, as shown in Table III, and continuous LSI growth is predicted until the 0.035 µm generation with gate length of 0.025 µm in the year of 2014, though with a comment of no known solution to some of the requirements for the future generations.

If the same trend is assumed up until 2059, the CD of LSI products will be less than the atomic distance in the silicon crystal.   This would be the ultimate limit of conventional LSI.   Before reaching this limit, however, we can expect certain other limitations to arise the reasons.   In particular, the 0.1 µm generation expected to start in 2005 is already thought to be a critical stage because several limitations coincide.

In spite of the expected limit in the downsizing, the downsizing of the MOSFETs has been accelerated year by year by year due to severe development competition among the chip suppliers which are aiming to release the next generation technology before the roadmap schedule, resulting in the revision for updating the roadmap as shown in Figs. 1 and 2.

This paper describes on a personal view of the perspective of MOSFETs and MOS LSIs into the 21st century, considering their possible limits in the early next century.

Table I   Evolution of LSI from 1970 to 1999

Year	CD	ratio	DRAM bit	ratio	MPU clock	ratio
1970 /72	10 μm	1	1K	1	750 KHz	1
1998	0.18 μm	1/50	256M	256,000	600 MHz	800

CD: Critical Dimension

Table II   Historical trend of CD and number of DRAM bits

Year	1970	1974	1977	1980	1983	1986	1989	1989	1995	1998
CD (μm)	10.0	6.0	4.0	3.0	2.0	1.2	0.8	0.5	0.35	0.25
Reduction	--	0.6	0.67	0.75	0.67	0.6	0.67	0.63	0.70	0.71
DRAM	1K	4K	16K	64K	256K	1M	4M	16M	64M	256M

Table III   Future trend predicted by International Technology Roadmap for Semiconductors updated 1998

Year	1997	1999	2002	2005	2008	2011	2014
Dense lines: Half pitch (μm)	0.25	0.18	0.13	0.10	0.07	0.05	0.035
Reduction	--	0.72	0.72	0.77	0.7	0.71	0.70
Isolated lines: MPU gate (μm)	0.20	0.14	0.10	0.07	0.05	0.035	0.025
Reduction	--	0.70	0.71	0.70	0.71	0.70	0.71
DRAM@ samples/introduction	256M	1G	4G	16G	64G	256G	1T
DRAM @production ramp	64M	256M	1G	4G	16G	64G	256G
MPU Clock frequency (MHz)							
On-chip, local, high performance	750	1250	2100	3500	6000	10000	16903
On-chip, across the chip, high performance	375	1200	1600	2000	2500	3000	3674
On-chip, across the chip, cost performance	400	600	800	1100	1400	1800	2303
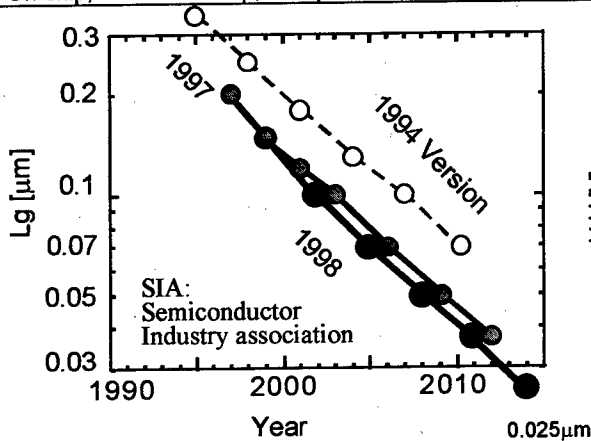
Fig.1   SIA Roadmap for isolated gate length
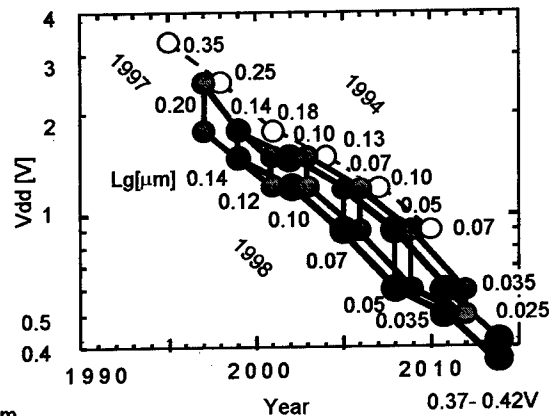
Fig.2   SIA Roadmap for minimum supply voltage

2. DOWNSIZING BY SCALING METHOD

In digital circuit applications a MOSFET functions as a switch. Thus, complete cutoff of leakage current in the "off" state, and low resistance or high current drive in the "on" state are required. In addition small capacitances are required for the switch to rapidly turn on and off.

When making the gate length small, even in the "off" state, the space charge region near the drain – the high potential region near the drain – touches the source in a deeper place where the gate bias cannot control the potential, resulting in a leakage electron current from source to drain via the space charge region as shown in Fig.3. This is the well-known short-channel effect of MOSFETs. The short-channel effect is often measured as the threshold voltage reduction of MOSFETs when it is not severe.
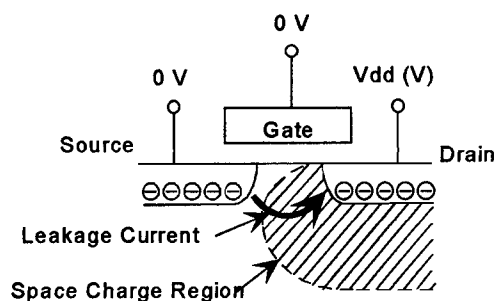
Fig.3 Short-channel effect

In order for a MOSFET to work as a component of an LSI, the capability of switching-off or the suppression of the short-channel effects is the first priority in the designing of the MOSFETs. In other words, the suppression of the short-channel effects limits the downsizing of MOSFETs.

In the "on" state, reduction of the gate length is desirable because it decreases the channel re-resistance of MOSFETs. However, when the channel resistance becomes as small as source and drain resistance, further improvement in the drain current or the MOSFET performance cannot be expected. Moreover, in the short-channel MOSFET design, the source and drain resistance often tends to even increase in order to suppress the short-channel effects. Thus, it is important to consider ways for reducing the total resistance of MOSFETs with keeping the suppression of the short-channel effects.

The capacitances of MOSFETs usually decreases with the downsizing, but care should be taken when the fringing portion is dominant or when impurity concentration of the substrate is large in the short-channel transistor design.

Thus, for the MOSFET downsizing, the suppression of the short-channel effects with the improvement of the total resistance and capacitances are required. In other words, without the improvements of the MOSFET performance, the downsizing becomes almost meaningless even if the short-channel effect is completely suppressed.

To suppress the short-channel effects and, thus, to secure good switching-off characteristics of MOSFETs, the scaling method was proposed by R. Dennard et al. [5], where the parameters of MOSFETs are shrunk or increased by the same factor K as

shown in Fig. 4, resulting in the reduction of the space charge region by the same factor K and suppression of the short-channel effects.
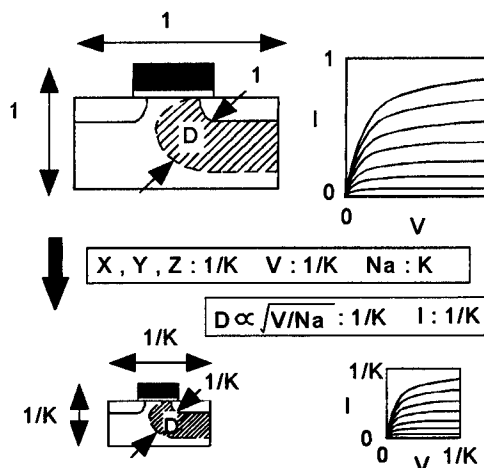
Fig.4 Scaling method

So far, great success has been achieved with the scaling method in miniaturizing MOSFETs down to gate-lengths of 0.14 μm at the LSI product although the actual scaling of the parameters has been somewhat different from that originally proposed as the ideal scaling, as also shown in Fig.5.
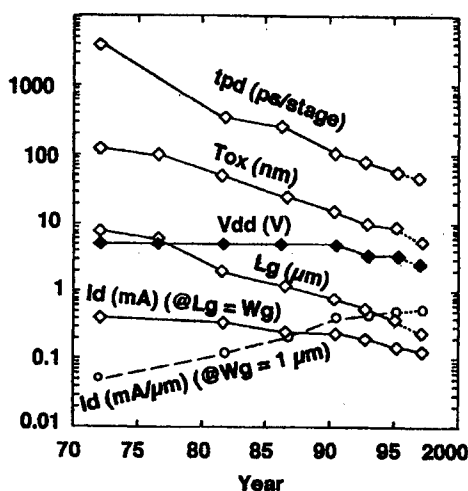
Fig.5 Trend of actual scaling for parameters

The major difference is the supply voltage reduction. The supply voltage was not reduced in the early phase of LSI generations in order to keep a compatibility with the supply voltage of conventional

systems and also in order to obtain higher operation speed under higher electric field. However, now, it is not easy to reduce the supply voltage because of difficulties in reducing the threshold voltage of the MOSFETs. Too small a threshold voltage leads to significantly large subthreshold leakage current even at the gate voltage of 0 V.

3. LIMITS EXPECTED AT CRITICAL GENERATIONS

If we could assume the same trend of the road-

map shown in Table III to the year of 2059, CD of LSI products would become less than atomic distance in silicon crystal, and 64E (Exa:10^{18}) bit DRAM would be produced with a supply voltage of 2 mV as shown in Table IV. This would be the a) ultimate limit of LSI determined by the atomic distance. The gate electrode could be produced by atomic manipulation such by STM, but it is hardly believable that this works as a component of LSIs. Thus, anyway, the progress of the LSI by the component downsizing will encounter the hard limit by the middle of the next century at the longest.

Table IV Simple extrapolation of the roadmap of Table III to future and expected downsizing limiting factors

Year (Product)	2005	2017	2026	2047	2059
CD	0.1 µm (= 100 nm)	0.025 µm (= 25 nm)	0.01 µm (= 10 nm)	0.001 µm (= 1 nm)	0.00025 µm (= 2.5 A)
Min Lg	0.07 µm	0.018 µm	0.07 µm	0.0007 µm	0.00018 µm
tox	1.5 – 2 nm	0.4 – 0.5 nm	0.15 – 0.2 nm (= 1.5 – 2 A)	0.015 – 0.02 nm (= 0. 15 – 0.2 A)	0.004 – 0.005 nm (= 0.04 – 0.05 A)
Min Vd	0.9 – 1.2 V	0.18 – 0.25 V	0.09 – 0.12 V	0.009 – 0.012 V	0.0018 – 0.002 V (= 1.8 – 2 mV)
DRAM	16G bit (Giga : 10^9)	256G bit	16T bit (Tera: 10^{12})	256P bit (Peta : 10^{15})	64E bit (Exa : 10^{18})
Downsizing Limiting factor	e) Scaling parameter f) Performance? g) Lithography? h) Interconnects? i) Economy?	d) Short-channel effects MOSFET switch-off?	b) Thermal noise? c) Uncertainty principle?		a) Atomic distance

Before this limit, however, there are expected certain levels of limitation by several reasons as shown in the table. In the year of 2047, the device feature size would reach 1 nm, and it has been told that b) thermal noise and c) uncertainty principle would limit the transistor and LSI functions. In the year of 2017, we could reach the 25 nm generation paying a huge effort, but the MOSFETs would not switch off due to the so-called d) short channel effects [6].

Especially, 0.1 µm generation is already thought to be a critical stage due to 5 downsizing limiting factors which relate each other; e) scaling parameter, f) performance, g) lithography, h) interconnects, and i) costs. Each factor is explained in the following sections.

4. LIMITS BY SCALING PARAMETERS AND DEVICE PERFORMANCE

In order to realize sub-0.1 µm MOSFETs, significant modification of the scaling method is required because some of the parameters have already reached their scaling limitation in the 0.1 µm generation as shown in Table V. In the 0.1 µm generation, the gate oxide thickness has already reached the direct-tunneling leakage limit of 3 nm. The substrate impurity concentration (or the channel impurity concentration) has already reached 10^{18}cm^{-3}. If the concentration is further increased, the source-substrate and drain-substrate junctions become highly doped pn junctions and act as tunnel diodes. Thus, the isolation of source and drains with substrate can-

not be maintained. The threshold voltage has already decreased to 0.2 - 0.4 V and further reduction causes significantly large subthreshold leakage current. Further reduction of the threshold voltage and thus, the further reduction of the supply voltage are difficult.

Table V Scaling parameter limitation

Scaling parameter	Value at 0.1μm generation	limiting factor
Lg	0.07 μm	cost of lithography
xj	20 - 40 nm	diffusion layer resistance
tox	1.5 – 2 nm	direct-tunneling leakage
Nsub	1E18/cm3	direct-tunneling leakage
Vd	0.9 – 1.2 V	lower limit of Vth
Vth	0.2 – 0.4	subthreshold leakage

As obvious from Table IV, if we would continue the simple scaling for future generations, tox, Vd, and Vth will soon become the unrealistically small values. Thus, in 1993, by using somewhat irregular scaling scheme as shown in Table VI – no further scaling of the tox, Nsub and Vd values, but instead, aggressive downsizing of Lg and xj values –, successful operation of 40 nm gate length MOSFETs was without the short-channel effects was confirmed as shown in Fig. 6 [7].

Table VI Scaling scheme used for 0.1 μm to 40 nm. Parameter values of 0.1 μm MOSFETs in 1993 was slightly different from those of Table IV.

	0.1 μm	→	40 nm
Lg	0.1 μm	2/5	40 nm
tox	3 nm	1	3 nm
xj	40 nm	1/4	10 nm
Nsub	10^{18}cm^{-3}	1	10^{18}cm^{-3}
Vd	1.5 V	1	1.5 V

However, it was found that the deep sub-0.1 μm MOSFETs – including the 40 nm Lg MOSFETs – made without downsizing tox do not show any current drivability increase as shown in Fig. 7. By the experiment of 1.5 nm direct-tunneling oxide MOSFETs [8], it was confirmed that gate oxide should go beyond the direct-tunneling limit – 3 nm –, and the SIA roadmap was changed aggressively as shown in Fig. 8.
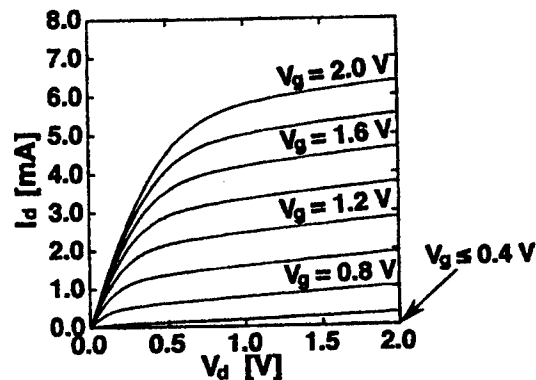
Fig.6 40 nm gate length n-MOSFET operation
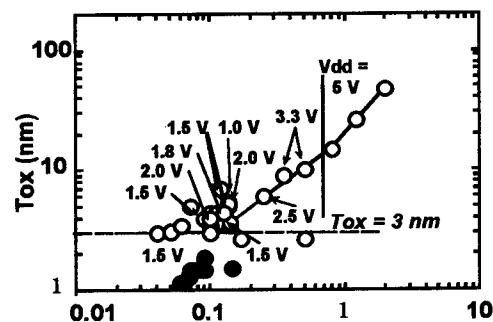
(a) gate oxide thickness
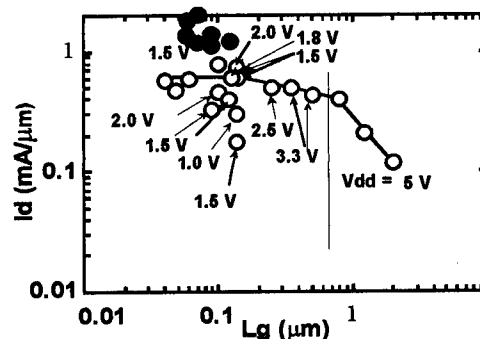
(b) Drain current

Fig.7 Correlation between the oxide thickness (a) and current derivability (b) plotted from published data.
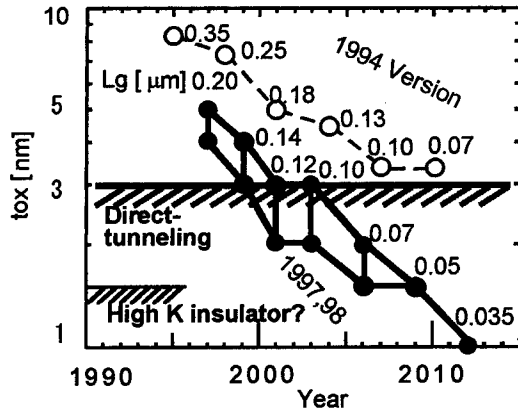
Fig.8. SIA Roadmap change for tox

Progress of the LSI has been achieved by the downsizing of components, mainly because the ca-

pacitance values decreases and thus switching time of the circuit decreases. This does not seem to be necessarily true when the CD decreases to 0.1 μm and below. The problems related with the interconnects are explained in the interconnects section. Regarding the MOSFET downsizing, also performance improvement cannot be guaranteed in the generations of 0.1 μm and below.

Table VII shows the current status of the gate lengths for each level of front-end research and products. As described in the above, the minimum Lg achieved for transistor level (Lg = 0.04 μm) does not give the highest transistor performance (At this moment, it is given at Lg = 0.06 - 0.07 μm), and that the minimum Lg achieved for CMOS ring oscillator (Lg = 0.075 μm) does not give the highest circuit performance (It is given at Lg = 0.1 μm).

Table VII　Gate lengths which give the highest performances

Digital Logic	Tr Level	Min. Lg (Simulation)	n-MOS: 0.025 μm	p-MOS: 0.025 μm
		Min. Lg (Experiment)	n-MOS: 0.04 μm	p-MOS: 0.05 μm
		Lg for max Id (Experiment)	n-MOS: 0.06 μm ;　1.8mA/μm @1.5 V	
		Lg for max gm (Experiment)	n-MOS: 0.06 μm: 1120mS/mm @1.5 V	
			n-MOS: 0.07 μm: >1100mS/mm @1.5V	
			n-MOS: 0.10 μm: 1210mS/mm @2.5V	
	Circuit Level	Min. Lg　(Experiment)	CMOS Ring Osc.: 0.075 μm ; 22ps @1.5V, 16ps @3.0V	
		Lg for min tpd (Experiment)	CMOS Ring Osc: 0.1 μm ; 8.0 ps @2.5V	
			SOI CMOS Ring Osc: ~0.1 μm? (Leff=0.07μm); 7.9ps@2.1V	
	LSI Level	Min Lg for MPU (Products)	0.18 μm :　550 – 600 MHz	

Thus, at this moment, the minimum Lg for MOSFET operation does not guarantee the maximum transistor, circuit or LSI performance. Large resistance caused by ultra-shallow S/D junction will harm the transistor and circuit performance, for example. For the logic devices, increase of the power consumption with increase in the clock frequency is another major concern. The reduction of the supply voltage is the most effective method to reduce the power with the sacrifice of the clock speed.

It should be noted that no circuit performance improvement has been realized below 0.1 μm generation

even in the research level, even though deep sub-0.1 μm gate length MOSFETs were realized. This would lead to the downsizing limit in terms of the LSI performance. We hope this is solved by future development of downsizing technology.

5. DOWNSIZING LIMIT BY LITHOGRAPHY

Lithography is the most fundamental process to realize the downsizing. The progress of the lithography resolution has been achieved by the reduction of the wavelength used for the light for the exposure of the lithography. Now, 248 nm (or 0.248 μm) wave length light emitted from a KrF excimer laser is used

for 0.18 µm devices. It should be noted that the KrF is for the lithography which size is much shorter than the wavelength. In order to realize this, several kinds of resolution enhancement technique (RET) has been/are going to be developed. Furthermore, it is expected that 193 nm wavelength ArF light with combination of the RET will realize down to 0.13 µm lithography in near future.

So far, 0.13 µm generation would be handled by ArF stepper. One of the most serious problem, however, is that there is no light source commonly accepted in the industry for the next generation to ArF. There are several kinds of lithography techniques proposed, but none of them have been confirmed at this moment as the cost-effective candidate for further 0.1 µm and below. We hope this will be solved in near future.

6. LIMITS BY INTERFCONNECTS

Interconnects are another big concern because the length of the interconnects and hence the number of interconnect layers continues to increase with generations. Thus, the number of the backend process steps increases, resulting in the significant production cost increase. Because the interconnect line becomes longer, narrower and denser for every generation, its resistance and capacitance increases dramatically. These RC delay and production cost for the interconnects would certainly work as a limiting factor to the downsizing in the regime of 0,1 µm and below.

Of course, the introduction of the new techniques such as Cu-damascene and low k interlayer material as shown would be very effective to extend the limit. However, the limitation of the interconnect multi-layer and downsizing will arrive in not far future.

8. OUTLOOK INTO THE NEXT CENTURY

As described in the previous sections, the next century will featured by the i) end of downsizing, ii) end of gate oxide thinning, iii) end of the increase in integration iv) end of progress in clock frequency, in the middle range. However, it does not mean the end of the R & D in semiconductor industry and its progress. LSIs are now indispensable products for human society, similar as automobile, aircraft manufacturing or steel industries. Even though, the progress in the speed or size of the airplane saturate, there is always a huge market and R &D are indispensable to win the

competition. Furthermore, as the function of LSI evolves to higher level for handling more sophisticated operations, new application and market will be explosively increase, which would grow the semiconductor industry further. In the middle to short range, beginning of the next century will be an exciting period for the researchers, because now the research for new materials which would replace the gate oxide or poly Si gate electrode, etc. are really demanded and there is a high possibility that these new material somehow extend the downsizing limit of MOSFETs or LSIs.

What would be a possibility of new device structures? In order to reduce the power consumption under considerably high speed operation in the mobile environment , higher drive current under lower bias is desirable. Single electron transistors does not seem to meet the requirement at this moment because of poor current drive or signal propagation capability. In the early period of the next century, telecommunication technology will make a huge progress. Under such an environment, high speed computations would be done by the base station and requirement to the high current drive would become less aggressive.

What will be the future direction of technology development ? Will it leave the beaten path? Will there be any major breakthrough? Table VIII shows some of the major discovery and invention in electronics for done from the end of last century.

Table VIII Discovery and invention in electronics
1897 Discovery of electrons by Thomson
1904 Invention of vacuum tube (Diode) by Flemimg
1907 Invention of vacuum tube (Triode) by De Forest
1925 Invention of concept of MOSFET by Lilinefeld
1946 Realization of computer by Eckert and Mauchly
1947 Invention of bipolar Tr by Bardeen, Bratten, Shockley
1959,61 Invention of IC by Kilby and Noyce
1960 Realization of MOSFET by Kahng and Atalla
1970, 71 Fabrication of 1kDRAM and 4bit MPU by Intel
1990's Internet, mobile computation / telecommunication

Considering such a tremendous progress in this century, it is impossible at all to predict the technology progress of next century. However, some personal view for the outlook of MOSFETs and LSIs in the next century is shown in Table XI, because this is the main objective of this paper. No one can make predictions about future breakthroughs, but it seems that an immediate revolution on the hardware side is unlikely.

Table IX Personal view: Outlook of the progress of MOSFET and LSI technologies

	21st century		
	Early period	Middle period	Late period
Hardware aspect	↓ Hardware performance limit? ↓ Downsizing limit? ↓ Interconnect problem? ↓ Lithography tool Problem? ↑ Progress in new material introduction?	↑ Suitable process and structure for new application or algorithm?	↑ Breakthrough for process or device???
Software aspect	↑ Software performance progress? ↑ New application more and more intelligent? more and more information &communication? more and more personal and mobile?	↑ Breakthrough in algorithm or architecture like biological system??	
Economy Market aspect	↓ Production cost increase ↓ Economy crisis in Asia, Russia, etc. ↓ Economy panic in the world.?? ↑ Increase in market size rapid growth in Chain, India etc.? new application described in the above?	↑ New market corresponding to the breakthrough shown in the above??	

Rather than hardware, there is a better chance of a breakthrough, or at least of innovation, on the software side, such as new application, algorithm and system architecture side. 'Internet' would be one of the examples. It is one of the major forces to expedite the realization of the information community and thus to develop new application fields of LSIs. Popularization of the intelligent mobile devices is another example.

In particular, it has been noted that today's computer architecture is much more inefficient than biological structures. Thus, there is room for revolutionary improvements in the efficiency of our computer structures at some point in the future. Once a breakthrough or innovation occurs on the software side, hardware will follow. New hardware demands imposed by new systems will become a strong motive force. In fact, the past development of digital LSIs has strongly depended on the conventional computer architectures. The discussion of the necessity of the downsizing in the roadmap has been made also on the premise of today's architecture. It is greatly expected that new algorithm and architecture of computer totally different from those of today will promote the development of a completely new style of LSIs.

Even now, prototypes of future hardware device components might already be in existence simply waiting for suitable applications.

Research on conventional path is still most important. It is likely that research on the introduction of many kinds of new materials – whether inorganic, organic, or even biological – to semiconductor devices will be the most productive; certainly this is an exciting area with potential for extending some limitation of LSIs. From the process research point of view, we look forward to new approaches that will dramatically reduce the production costs. Also, environmental issues must be considered of utmost importance in the development of any new techniques.

REFERENCES

1. D. Kahng et al, IRE Solid-StateDev. Res. Conf., (1960).
2. G. E. Moore, IEDM Tech. Dig., p. 11, December, (1975)
3. Semiconductor Industry Association, 'The national technology roadmap for semiconductors,' (1997)
4. Semiconductor Industry Association, 'The international technology roadmap for semiconductors: (1998)
5. R. H. Dennard et al. , IEEE J. Solid-State Circuits, SC-9, p.256 (1974)
6. H. Iwai, IEEE J.Solid-State Circuits, vol.34, p.357 (1999)
7. M. Ono et al., IEDM Tech. Dig., p. 119, (1993).
8. H. S. Momose et al., IEDM Tech. p. 593 (1994)

ULTRATHIN DIELECTRICS

The initial growth steps of ultrathin gate oxides

Takeo Hattori, Hiroshi Nohira, and Kensuke Takahashi

Department of Electrical Engineering, Musashi Institute of Technology
Fax: +81-3-5707-2173, Tel: +81-3-3703-3111, E-mail: hattori@ipc.musashi-tech.ac.jp

The studies on the atomic-scale surface roughness, interface structures and interface-state-distribution in silicon bandgap at the initial growth steps of ultrathin oxides formed on Si(100) are reviewed in comparison with ultrathin oxides formed on Si(111). Interface-state-density distribution in silicon band gap was found to change periodically with progress of oxidation in accordance with layer-by-layer oxidation. Therefore, the oxide film thickness must be adjusted with the accuracy of less than 0.1 nm in order to minimize the interface-state-densities. The structural difference between 1-nm-thick structural transition layer and bulk silicon dioxide was detected from the measurement of O1s photoelectron spectra. The effect of elastic scattering on Si 2p photoelectrons in silicon oxide, which is important in the accurate structural analysis of ultrathin gate oxides, is also discussed.

1. INTRODUCTION

Since MOSFET with 1.5-nm-thick gate oxide was shown to operate at room temperature,[1] extensive studies have been performed on the reliability of ultrathin gate oxides. Because 1.5 nm is close to the thickness of structural transition layer consisting of SiO_2 and the amount of electronic defects in the structural transition layer must be larger than those in bulk SiO_2, the reliability of ultrathin gate oxides must be strongly influenced by the chemical structure of structural transition layer. In the present paper surface and interface structures and correlation between interface states and structures at the initial growth steps of ultrathin gate oxides in addition to the effect of elastic scattering of Si 2p photoelectrons in silicon oxides, which is important in the accurate structural analysis of ultrathin gate oxides, will be discussed.

The layer-by-layer oxidation process on silicon surface was first discovered using transmission electron microscope[2] and was later confirmed on an atomic-scale in lateral direction using scanning reflection electron microscope.[3,4] This layer-by-layer oxidation reaction was shown to occur at the interface from X-ray photoelectron spectroscopy studies on the interface structures. Namely, on Si(111) surface the periodic changes in interface structures appear as a result of bonding nature of Si crystal at the interface.[5] The abrupt composi-tional transition takes place on Si(111) and (100) surfaces and is weakly affected by the atomic steps on the initial surface.[5,6]

In order to relax oxidation-induced stress near the interface the structural transition layers are formed on both sides of the interface. The thickness of this structural transition layer on the oxide-side was determined to be about 1 nm from the analysis of infrared absorbance and X-ray reflectance.[7, 8] The valence band discontinuity of 0.2 eV was found at the boundary between this structural transition layer and bulk SiO_2 layer for oxides formed on Si(111) and Si(100)[9, 10] and was shown to result from the distortion of silicon oxides in the structural transition layer.[11] Furthermore, the structural transition layer was found be different from bulk SiO_2 layer from the measurement of O1s photoelectron energy loss spectra probably because there is a large difference in defect structure between these two layers. This can be expected from the recent medium-energy ion-scattering studies that the oxidation reaction occurs not only at the interface but also near the interface.[12]

It was found from the studies on oxide surfaces using non-contact mode atomic force microscope[13] that there is a clear difference in surface morphologies of oxides formed on Si(111) and Si(100). With the progress of oxidation from the thickness of 1.0 to1.7 nm the size of protrusions on the surface of oxides formed on Si(111) increases,

while the size of protrusions on Si(100) does not change significantly.

2. EXPERIMENTAL DETAILS

Because the formation of native oxide can be suppressed by terminating Si surface with hydrogen atoms,[14,15] atomically flat hydrogen-terminated Si surfaces (abbreviated as H-Si surfaces hereafter) were used as initial surfaces before the oxidation. Here, an atomically flat H-Si(111)-1×1 surface was obtained by the treatment[16] in 40% NH_4F solution, while an atomically flat H-Si(100)-2×1 surface can be obtained by the surface reconstruction[17] of Si(100) surface annealed at high temperature (>1100°C) in a H_2 atmosphere at 1 bar.

The extremely uniform oxide films studied in the present paper were prepared as follows. In order to preserve the flatness of the initial Si surface during the oxidation at high temperature, the nearly 0.5 nm thick preoxides were formed at 300°C[18] in 1-4 Torr dry oxygen by oxidizing atomically flat hydrogen-terminated Si(111)-1×1 and Si(100)-2×1 surfaces without breaking Si-H bonds. Through these preoxides the oxidations at 600-900°C in 1 Torr dry oxygen were performed. In order to form device-grade oxide the effect of impurities in the oxidizing atmosphere on the oxidation process was minimized by forming oxides in flowing condition of high-purity oxygen gas with a pressure of 1 Torr. The amount of water vapor in oxygen gas used was below 37 ppb. In order to heat Si wafers in oxygen gas under high pressure and at high temperature Si wafers were only heated optically.

Interface structures of ultrathin silicon oxides were studied from the measurement of photoelectron spectra excited by monochromatic AlK α radiation with an acceptance angle of 3.3 degrees, using ESCA-300 manufactured by Scienta Instrument AB,[19] while surface structures of ultrathin silicon oxides were studied from the observation of noncontact-mode atomic force microscope (NC-AFM) images with a force constant of 39 N/m and resonant frequency of about 300 kHz, and a single-crystalline silicon probe, using instrument manufactured by OMICRON Vakuum Physik GmbH.[20]

The interface state densities were measured using the method developed by Lau et al.[21] First, an organic molecules of 2-propanol were absorbed on oxide film. Second, oxide is charged by electron beam irradiation with electron kinetic energy of 2 eV to produce voltage drop across the oxide film. Third, this charging-induced changes in electrical potentials at surface and interface, which are necessary for the determination of interface state densities, were obtained from the measurements of chemical shifts in C 1s level of organic molecules and Si 2p core level of Si substrate. The organic molecules were decomposed during the oxidation. Other experimental details and analytical procedure of Si 2p photoelectron spectra were described elsewhere.[22]

3. OXIDATION-INDUCED SURFACE ROUGHNESS

Recently, it was found that the surface roughness changes in accordance with the changes in interface structure.[13] Figure 1 shows non-contact mode atomic force microscope images of surface of oxides formed on Si(111) and Si(100) measured over an area of 200 nm × 200 nm for oxide film thicknesses of. 1.0 and 1.7 nm. According to this figure the surface morphologies are quite different for these two oxide surfaces. The height distribution on the surface of oxides formed on Si(111) and Si(100) are shown in Figs. 2 and 3 and can be approximated by the Gaussian function. As shown in these figures, full-width at half-maximum (FWHM) of this Gaussian function is about two atomic step heights on Si(111), while on Si(100) FWHM is within single atomic step height for the thickness smaller than the thickness of structural transition layer and increases with increasing thickness. In other words, FWHM for oxides formed on Si(100) is affected strongly by the oxidation-induced stress near the interface.

Figure 4 shows averaged values of root mean square(rms) surface roughness of oxide films formed on Si(111) and Si(100), respectively, measured at three positions on the oxide surfaces, as a function of thickness. Oscillations in surface roughness appear on both surfaces and can be attributed to the layer-by-layer oxidation. The oscillations in surface roughness observed on Si(111) can be correlated with the periodic changes in the amount of Si^{1+} as shown in Fig. 4(a) and can be explained as follows. In the case of forming
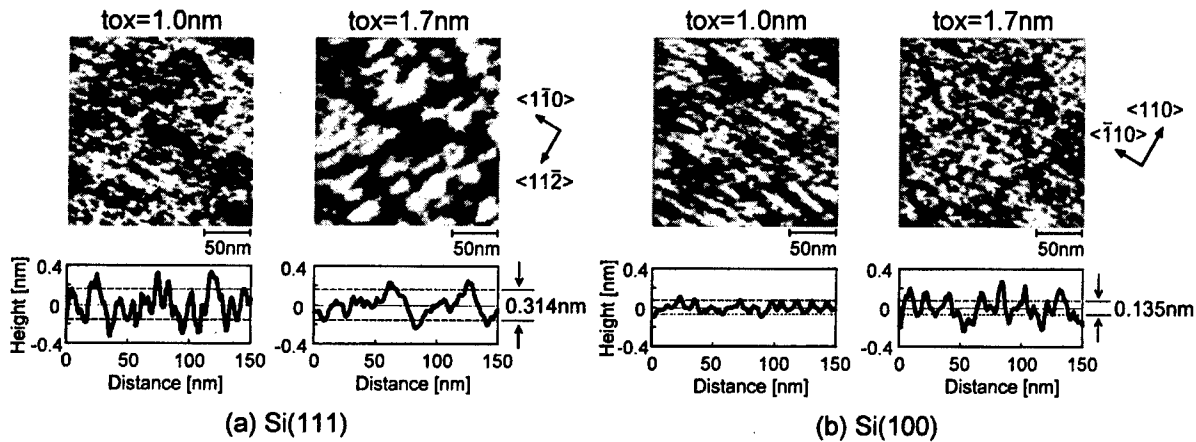
(a) Si(111)   **(b) Si(100)**

Fig. 1   Non-contact mode atomic force microscope images of surface of oxides formed on (a) Si(111) and (b) Si(100) for two thicknesses.
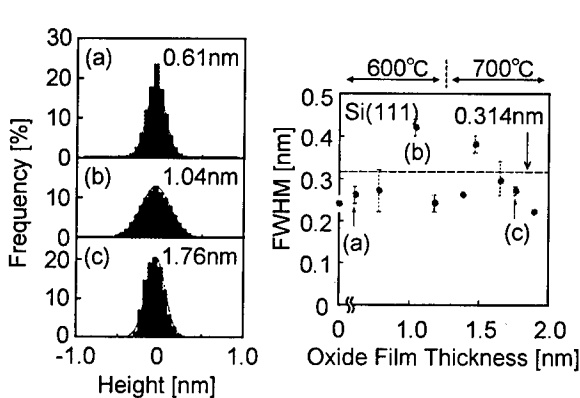
Fig. 2   Height distributions on the surface of oxides formed on Si(111) for three oxide film thicknesses and FWHM of height distribution as a function of oxide film thickness are shown.

Fig. 3   Height distributions on the surface of oxides formed on Si(100) for three oxide film thicknesses and FWHM of height distribution as a function of oxide film thickness are shown.

Si^{1+} the insertion of an oxygen atom between two Si atoms consisting of a Si-Si bond oriented along the <111> direction at the interface expands the oxide network only along the <111> direction, while in the case of forming Si^{3+} the insertion of an oxygen atom between two Si atoms at the interface expands the oxide network mostly along the direction perpendicular to <111>. Therefore, the formation of Si^{1+} at the interface results in the increase in surface roughness caused by the formation of protrusions on the oxide surface.

4. INTERFACE STATES CORRELATED WITH INTERFACE STRUCTURES

Figure 5 shows the changes in interface structures with progress of oxidation of Si(111), while Fig. 6 shows the changes in interface state distribution in Si bandgap with oxide film thickness as a parameter.[23]   According to these figures, if the amount of Si^{1+} takes its maximum value, the drastic decrease in interface state densities near the midgap of silicon appear.   This drastic decrease in interface state densities appears periodically in accordance with periodic changes in interface structures.   As
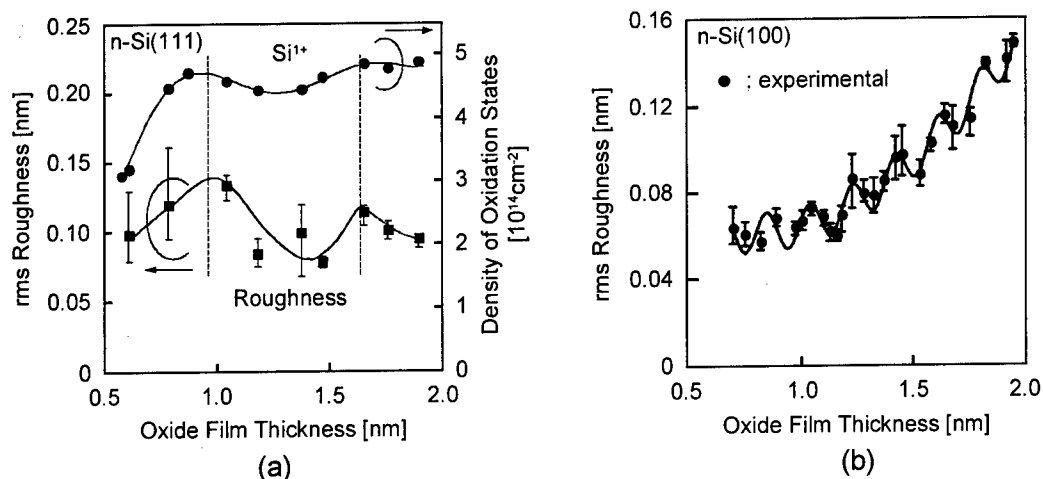
Fig. 4 Surface roughness for oxide films formed on (a) Si(111) and (b) Si(100) as a function of thickness.
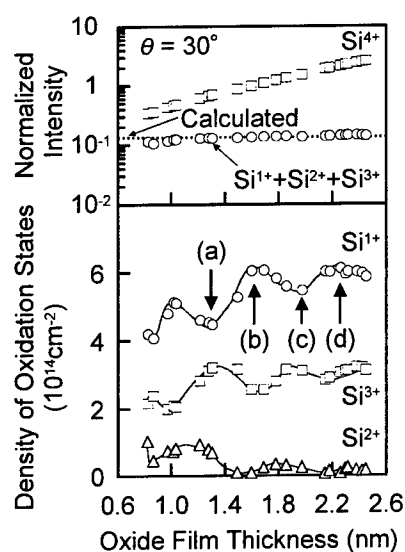
Fig. 5 Changes in interface structures with progress of oxidation of Si(111): upper part shows the dependence of normalized spectral intensity of Si^{4+} and normalized total intensities for all intermediate oxidation states on the oxide film thickness, while lower part shows the dependence of areal densities of Si^{1+}, Si^{2+} and Si^{3+} on the thickness.

Fig. 6 Changes in interface state distribution in silicon bandgap with oxide film thickness as a parameter.

described in section 3 the bond breaking probability of Si^{3+}-Si bond must be the largest at the thickness where the amount of Si^{3+} takes its maximum value, while that must be the smallest at the thickness where the amount of Si^{3+} is takes its minimum value.

The difference in dangling bond densities for these two thicknesses can be the possible origin for the difference in the interface state densities near the midgap of silicon for two thicknesses.

Figure 7 shows the changes in interface structures with progress of oxidation of Si(111), while Fig. 8
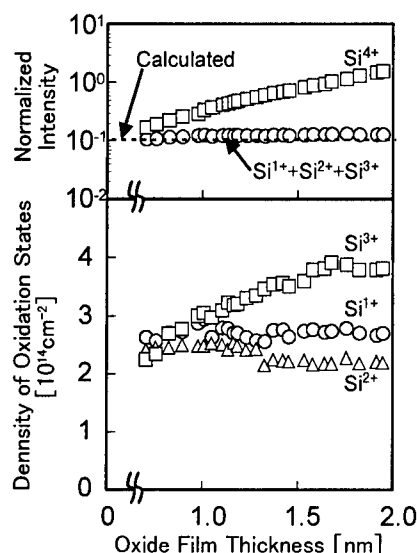
Fig. 7 Changes in interface structures with progress of oxidation of Si(100): upper part shows the dependence of normalized spectral intensity of Si^{4+} and normalized total intensities for all intermediate oxidation states on the oxide film thickness, while lower part shows the dependence of areal densities of Si^{1+}, Si^{2+} and Si^{3+} on the thickness.
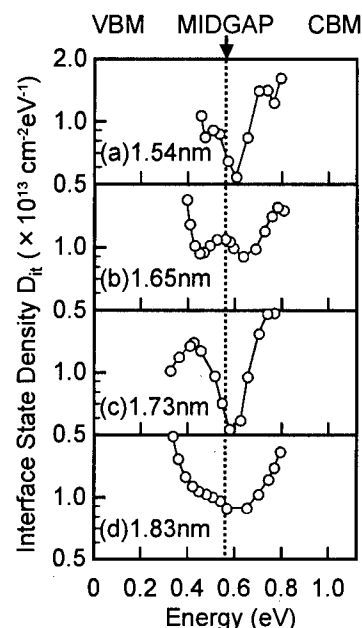
Fig. 8 Changes in interface state distribution in Si bandgap with oxide film thickness as a parameter.

shows interface state densities at and near the midgap of Si with oxide film thickness as a parameter for oxide films formed on Si(100) surface.[24] There is no clear correlation between interface structures and interface state distribution. However, there is a close correlation between atomic-scale surface roughness and interface state distribution in silicon bandgap. Namely, for oxide film thicknesses of 1.65 and 1.83 nm, where the rms roughness takes its maximum value, the interface state densities at and near the midgap are almost equal to 1×10^{13} cm^{-2}eV^{-1}, while for oxide film thicknesses of 1.54 and 1.73 nm, where the rms roughness takes its minimum value, the interface state density near the midgap decreases drastically down to 5×10^{12} cm^{-2}eV^{-1}. Therefore, the drastic decrease in interface state density near the midgap appears periodically with progress of oxidation and is closely correlated with the atomic-scale smoothness of oxide surface, which must reflect the atomic-scale smoothness of the interface. However, the size of atomically smooth interface must be on

the order of 5 nm deduced from the atomic force microscope images in Fig. 1(b). Therefore, the atomic step at the interface must be the structural origin of the interface states. This result is contrasted with the observation for oxide films formed on Si(111) surface that the increase in surface roughness results in the decrease in interface state density in the middle of silicon bandgap.

5. O1s PHOTOELECTRON ENERGY LOSS IN STRUCTURAL TRANSITION LAYER

Figure 9 shows O1s photoelectron energy loss spectra normalized by the total number of elastically scattered and non-scattered O1s photoelectrons with oxide film thickness as a parameter and spectra expanded by 15 times.[25] According to this figure the energy loss arising from bandgap ionization in bulk SiO$_2$ starts to appear above the thickness of 1 nm.[26,27] In Fig. 10 the numbers of inelastically scattered O1s photoelectrons normalized by the total numbers of elastically scattered and non-scattered O1s photoelectrons is shown as a function of the suboxide spectral intensity normalized by the oxide
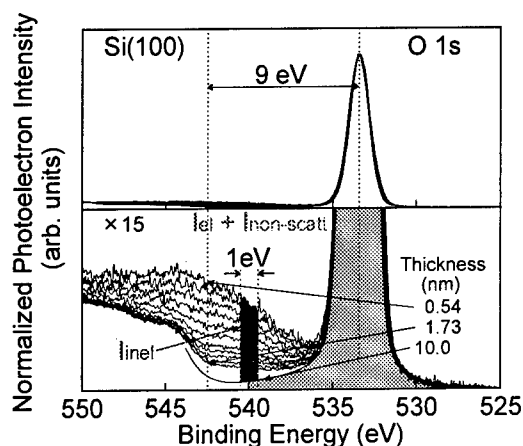
Fig. 9   O1s photoelectron energy loss spectra normalized by total numbers of elastically scattered and non-scattered photoelectrons with oxide film thickness as a parameter.
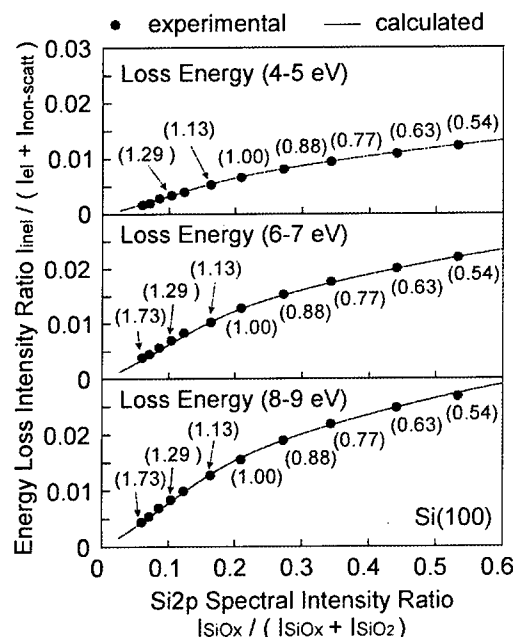
Fig. 10   Numbers of inelastically scattered O1s photoelectrons normalized by the total numbers of elastically scattered and non-scattered photoelectrons as a function of the suboxide spectral intensity normalized by the oxide spectral intensity for the loss energy in the range from 6 to 7 eV.
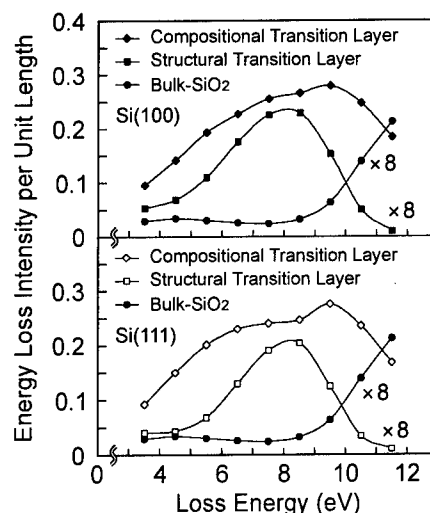
Fig. 11   O1s photoelectron energy loss spectra in compositional and structural transition layers and those in bulk SiO_2 for oxides formed on Si(100) and Si(111).

spectral intensity for three loss energy ranges. Sublinear dependence is obtained for these loss energy ranges.   If the energy loss occurs only in the compositional transition layer at the interface, a line almost close to the straight line is obtained.   If the energy loss arises not only from the compositional transition layer but also from the structural transition layer, sublinear curve is obtained.   Therefore, the experimental results can be explained by considering the energy loss in compositional and structural transition layers.   From these analyses the energy loss spectra in compositional and structural transition layer and the energy loss spectrum in bulk silicon dioxide shown in Fig. 11 were obtained for oxides formed on Si(100) and Si(111).[25,28] According to this figure, the energy loss mostly arises from compositional transition layer and the energy loss spectra for structural transition layer is essentially different from that in bulk silicon dioxide.

6. ELASTIC SCATTERING OF Si 2p PHOTO-ELECTRONS IN SILICON OXIDES

Figure 12 shows the photoelectron diffraction pattern, that is, the dependence of Si 2p photoelectron spectral intensity arising from Si substrate on photoelectron take-off angles with oxide
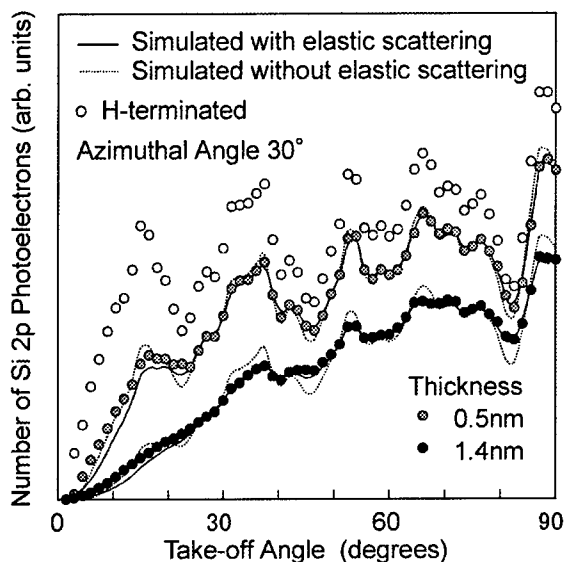
Fig. 12   Dependence of Si 2p photoelectron spectral intensity arising from silicon substrate on photoelectron take-off angles with oxide film thickness as a parameter.

film thickness as a parameter, which results from the interference between the primary photoelectron wave and elastically scattered photoelectron wave in single crystalline silicon.[29]   The oxidation-produced attenuation of photoelectron diffraction pattern can be simulated by considering elastic and inelastic scattering of Si 2p photoelectrons in silicon oxides.   In the simulation the Monte Carlo calculation of path of elastically and inelastically scattered Si 2p photoelectrons were performed by considering the following: 1) the structure of hydrogen-terminated silicon surface does not change by the oxidation, 2) a reported value of electron escape depth in Si,[30] 3) all inelastic scattering of Si 2p photoelectrons in silicon oxide, 4) anisotropic elastic scattering based on Wentzel model for atomic potential.   It was found from this simulation that the elastic and inelastic scattering cross-sections in silicon oxides are 1.4×10^{-20} m^2 and 1.5×10^{-20} m^2, respectively.   According to Fig. 12, the oxide film thickness of 1.4 nm is not thin enough to neglect the effect of elastic scattering of Si 2p photoelectrons in silicon oxides on the photoelectron diffraction. It should be noted that the concept of electron escape depth can be used only for special angles where the effect of elastic scattering can be neglected effectively as can be seen in Fig. 12.   Other

experimental details were described elsewhere.[29]

7. SUMMARY

The surface roughness and interface state distribution in silicon bandgap measured for oxide films formed on Si(111) and Si(100) exhibit periodic changes with progress of oxidation and can be correlated with layer-by-layer oxidation.   It is found from the measurement of O1s energy loss spectra that the defect structure in structural transition layer must be quite different from that in bulk silicon dioxide layer, although the structural transition layer exhibits almost the same Si 2p photoelectron spectra as those measured for silicon dioxide.   Elastic and inelastic scattering cross-sections of Si 2p photoelectrons in silicon oxides were determined to be 1.4×10^{-20} m^2 and 1.5×10^{-20} m^2, respectively.

REFERENCES

[1] H. Sasaki Mmose,   M. Ono,   T. Yoshitomi, T. Ohguro, S. Nakamura, M. Saito,   and H. Iwai, (1994), IEDM 94-593.

[2] F. M. Ross and J. M. Gibson,   Phys. Rev. Lett. **68**, 1782(1992).

[3] H. Watanabe,   K. Kato,   T. Uda, K. Fujita, M. Ichikawa, T. Kawamura, and K. Terakura, Phys. Rev. Lett. **80**, 345 (1998).

[4] N. Miyata,   H. Watanabe,   and M. Ichikawa, Phys. Rev. B **58**, 13670 (1998).

[5] K. Ohishi and T. Hattori, Jpn. J. Appl. Phys. **33**, L675 (1994).

[6] T. Aiba, K. Yamauchi, Y. Shimizu, N. Tate, M. Katayama, and T. Hattori,   Jpn. J. Appl. Phys. **34**, 707 (1995).

[7] K. Ishikawa,   H. Ogawa,   S. Oshida,   K.

Suzuki, and S. Fujimura, Ext. Abstr. of Int. Conf. on Solid State Devices and Materials (Osaka, 1995) p. 500.

[8] Y. Sugita, N. Awaji and S. Watanabe, Ext. Abstr. of Int. Conf. on Solid State Devices and Materials (Yokohama, 1996) p. 380.

[9] H. Nohira and T. Hattori, Appl. Surf. Sci. **117/118**, 119 (1997).

[10] H. Nohira, A. Omura, M. Katayama, and T. Hattori, Appl. Surf. Sci. **123/124** (1998) 546.

[11] H. Hirose, H. Nohira, T. Koike, K. Sakano,and T. Hattori, Phys. Rev. B**59**, 5617 (1999).

[12] E. P. Gusev, H. C. Lu, T. Gustafsson, and E. Garfunkel, Phys. Rev. B **52**, 1759 (1995).

[13] M. Ohashi and T. Hattori, Jpn. J. Appl. Phys. **36** (1997) L397.

[14] T. Takahagi, I. Nagai, A. Ishitani and H. Kuroda, J. Appl. Phys. **64** (1988) 3516.

[15] M. Sakuraba, J. Murota and S. Ono, J. Appl. Phys. **75** (1994) 3701.

[16] G. S. Higashi, R. S. Becker, Y. J. Chabal and A. J. Becker, Appl. Phys. Lett. **58** (1991) 1656.

[17] H. Bender, S. Verhaverbeke, M. Caymax, O. Vatel and M. M. Hynes, J. Appl. Phys. **75** (1994) 1207.

[18] T. Ohmi, M. Morita, A. Teramoto, K. Makihara and K. S. Tseng, Appl. Phys. Lett. **60** (1992) 2126.

[19] U. Gelius, B. Wannberg, P. Baltzer, H. Fellner-Feldegg, G. Carlsson, C. -G. Johansson, J. Larsson, P. Munger and G.

Vergerfos, J. Electron Spectrosc. & Relat. Phenom. **52** (1990) 747.

[20] P. Guthner, J. Vac. Sci. & Technol. B**14** (1996) 2428.

[21] W. M. Lau and X. -W. Wu, Surf. Sci. **245**, 345 (1991).

[22] H. Nohira, Y. Tamura, H. Ogawa, and T. Hattori, IEICE Trans. Electron. E75-C, 757 (1992).

[23] N. Watanabe, A. Omura, H. Nohira and T. Hattori, Ext. Abstr. of Intern. Conf. on Solid State Devices and Materials, Hiroshima, 1998, p. 130.

[24] N. Watanabe, A. Omura, H. Nohira, a nd T. Hattori, unpublished.

[25] H. Nohira, K. Takahashi, and T. Hattori, Proc. 4th Int. Symp. UCPSS'98, Ostend, 1998, pp. 241-244.

[26] F. G. Bell and L. Ley, Phys. Rev. B**37**, 8383 (1988).

[27] S. Miyazaki, H. Nishimura, M. Fukuda, L. Ley, and J. Ristein, Appl. Surf. Sci. **113/114**, 585 (1997).

[28] H. Nohira, K. Takahashi, and T. Hattori, reported at 14th Int. Vacuum Congress, Birmingham, 1998.

[29] T. Hattori, K. Hirose, H. Nohira, K. Takahashi, T. Yagi, Appl. Surf. Sci. **144/145**, 297 (1999).

[30] Z. H. Lu and J. P. McCaffrey, B. Brar, G. D. Wilk, R. M. Wallance, L. C. Feldman, and S. P. Tay, Appl. Phys. Lett. **71**, 2764 (1998).

Understanding the Limits of Ultrathin SiO$_2$ and Si-O-N Gate Dielectrics for Sub-50 nm CMOS

M. L. Green, T. W. Sorsch, G. L. Timp, D. A. Muller, B. E. Weir, P. J. Silverman, S. V. Moccio and Y. O. Kim

Bell Laboratories/Lucent Technologies
600 Mountain Avenue, Murray Hill, New Jersey 07974, USA

In spite of its many attributes such as nativity to silicon, low interfacial defect density, high melting point, large energy gap, high resistivity, and good dielectric strength, SiO$_2$ suffers from one disadvantage, low dielectric constant (K=3.9). Thus, ultrathin SiO$_2$ gate dielectric layers are required to generate the high capacitance and drive current required of sub-50 nm transistors. The silicon industry roadmap dictates 4 nm SiO$_2$ gate dielectrics for 0.25 μm technology today, and calls for <1 nm equivalent SiO$_2$ thickness for 0.05 μm technology in 2012. SiO$_2$ layers in this thickness range may suffer from boron penetration, reduced drive current, reliability degradation, and high gate leakage current. We will argue that none of these problems are limitations for thicknesses greater than about 1.3 nm. Below that thickness, the fundamental problems of high tunneling current and reduced current drive will prevent further scaling, unless alternate gate dielectrics are introduced.

1. INTRODUCTION

The continuous scaling of integrated silicon microelectronic devices has required a continuous decrease in the thickness of the SiO$_2$ gate dielectric [1], as is shown in Fig. 1. As the SiO$_2$ gate dielectric becomes thinner, the Si/SiO$_2$ and SiO$_2$/gate interfacial regions play a more influential role in the properties of the device. A 1.3 nm SiO$_2$ layer contains about five silicon atom layers, at least two of which reside at the interfaces. Transistors with gate lengths smaller than 50 nm have already been produced, Fig. 2, utilizing SiO$_2$ gate dielectrics as thin as 1.0 nm [2].

Nature has endowed the silicon microelectronics industry with a wonderful material, SiO$_2$. It is native to silicon, and forms a low defect density interface. It also has a high melting point (1713°C), large energy gap (9 eV), high resistivity ($\geq 10^{15}$ Ω-cm), and very good dielectric strength (10^7 V/cm). However,
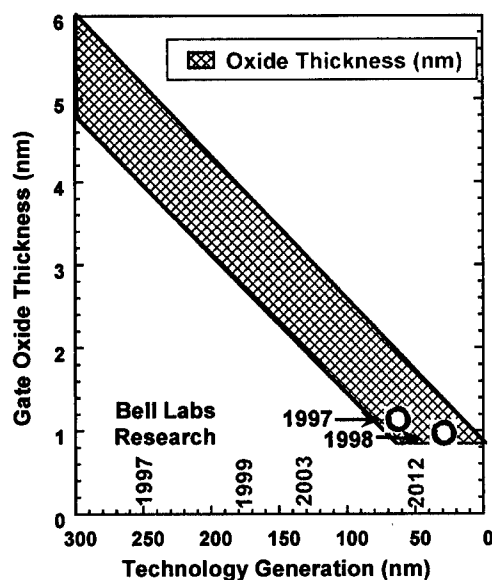
Figure 1. Decrease in SiO$_2$ gate dielectric thickness with scaling of silicon device dimensions.
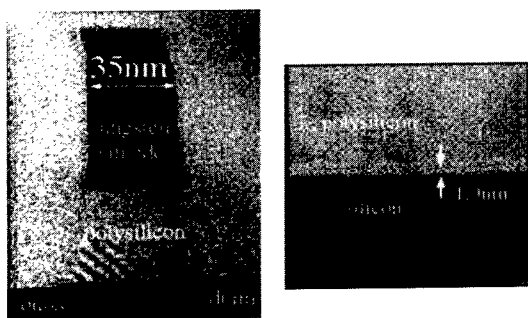
Figure 2. Cross-section transmission electron photomicrographs of a 35 nm transistor, and a detailed view of the 1.0 nm SiO_2 gate dielectric.

SiO_2 suffers from one disadvantage, a low dielectric constant (K=3.9). Since high gate dielectric capacitance is necessary to produce the required drive currents for sub-50 nm devices, the SiO_2 layers must be ultrathin. This gives rise to a number of problems, including the need to grow ultrathin and uniform SiO_2 layers, boron penetration [3], reduced drive current [4], reliability degradation [5], and high gate leakage current [2]. In this paper, we will address these issues as a function of decreasing oxide thickness, and attempt to understand the critical limits on oxide performance imposed by the thickness decreases. We will come to the conclusion that all of these problems can be overcome as long as the SiO_2 thickness is greater than about 1.3 nm. Below that thickness, excessively high leakage current and reduced drive current become fundamental limitations to further scaling.

Of course, one solution to the scaling limitation is the use of alternate, higher dielectric constant (K) gate dielectrics, and much research is currently underway in this field [6]. None of this work will be covered here. The introduction of alternate gate dielectrics will not be simple, since SiO_2 has only one disadvantage, low K, and most alternate gate dielectrics have only one advantage, high K.

2. ISSUES POTENTIALLY LIMITING THE USE OF ULTRATHIN SiO_2 LAYERS

2.1. Ultrathin and uniform film growth

The silicon industry roadmap [1] dictates 4 nm SiO_2 gate dielectrics for 0.25 µm technology today, and calls for <1 nm equivalent SiO_2 thickness for 0.05 µm technology in 2012. SiO_2 layers in this thickness range can be grown by either furnace or rapid thermal oxidation (RTO) techniques. Due to their large thermal mass, furnaces must be operated at either low temperature (600-800°C) or low pressure (1-100 Torr) to achieve thickness control. On the other hand, since RTO systems can visit high temperatures for short periods of time, growth can occur at high temperatures (900-1100°C), which may be preferable. Fig. 3 depicts a state-of-the-art RTO cluster tool for thin oxide growth. A further advantage of single wafer RTO processing is the integration of the clean before oxidation, and the polysilicon gate deposition after oxidation. In particular the in-situ vapor phase clean has been used to great advantage in preparing ultrasmooth and clean Si surfaces, which results in enhanced device performance [2,4,7].

Integrated (Cluster-Tool) Gate Dielectric Processing
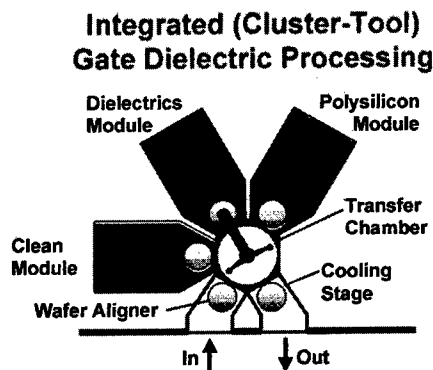
Figure 3. Schematic depiction of an integrated cluster tool for clean/gate dielectric/polysilicon processing.

Fig. 4 is a plot of RTO oxide thickness as a function of oxidation pressure, at 1000°C, demonstrating the ability to grow 1.0 nm oxides. Excellent film thickness uniformity ($3\sigma \leq 4\%$) has been achieved on 200 mm wafers by both furnace and RTO techniques.
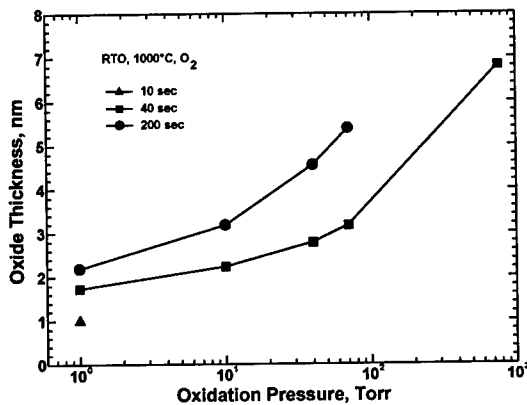
Figure 4. Oxide thickness as a function of growth pressure for RTO oxides grown at 1000°C.
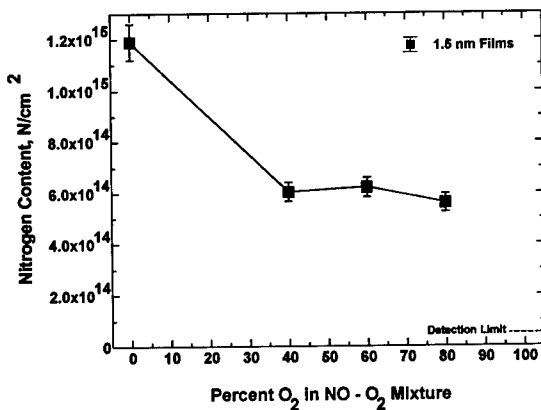
Figure 5. Nitrogen content of oxynitrides grown in O_2-NO mixtures by RTO at 1000°C.

2.2 Boron penetration

It is well known that boron diffuses easily through SiO_2, and can cause threshold voltage shifts when present in the channel

region [8]. Careful control of thermal budget can prevent boron from diffusing significantly [2]. However, even if it only diffuses into the gate dielectric, boron may degrade oxide reliability. Nitrogen doping of the SiO_2 may be necessary to stop boron from diffusing out of the gate polysilicon. In this respect, it is easy to incorporate significant amounts of nitrogen into ultrathin SiO_2 by RTO in O_2-NO mixtures, as is shown in Fig. 5. It can be seen that up to 1.2×10^{15} N/cm^2 (the equivalent of about two monolayers) can be incorporated into 2.0 nm SiO_2 films. This is enough nitrogen to prevent boron penetration for typical thermal budgets [3].

2.3 Reduced drive current

Reduced drive current has been reported in small transistors with ultrathin gate
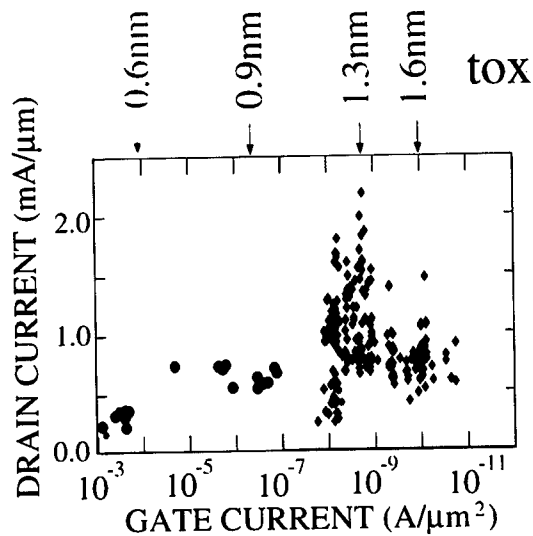
Figure 6. DC transistor performance (drain current) as a function of gate leakage current, for various ultrathin gate dielectric thicknesses (typical values indicated on top).

dielectrics [4]. Fig. 6 shows that as gate leakage current increases (as is expected for decreasing SiO_2 thickness), drive current does not increase, but rather decreases slightly. Thus, for SiO_2 layers thinner than about 1.3 nm, there is no

advantage in performance (drive current) for incurring the burden of an ever-increasing gate leakage current. This would suggest that SiO_2 layers thinner than 1.3 nm no longer deliver any performance advantage.

The cause of the decreased drive current is not well understood now. One possibility is an additional scattering component from the upper (SiO_2/polysilicon gate) interface. Alternatively, it could be a universal mobility curve effect, i.e., lowered mobility due to enhanced scattering because of extreme carrier confinement in the inversion layer of the ultrathin oxide.

2.4 Gate oxide reliability

Recently, an analysis of SiO_2 reliability data came to the conclusion that films thinner than about 2.2 nm would not have the reliability required by the industry roadmap [5]. Therefore, it was predicted that scaling might stop at that point. Our recent data, co-plotted with the aforementioned data in Fig. 7, indicates that for at least one oxide thickness, 2.8 nm, higher reliability is possible. If future
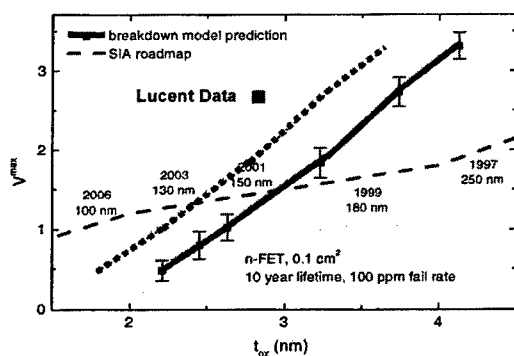
Figure 7. Oxide reliability prediction, adapted from [5]. Note our data point that indicates higher than predicted reliability for a 2.8 nm SiO_2 film.

results on our thinner oxides show the same trend, then SiO_2 might be reliable to

thicknesses of about 1.5 nm. We attribute the higher predicted reliability of our SiO_2 film to a larger Weibull slope due to enhanced film uniformity, and to higher film quality as evidenced by longer time to breakdown.
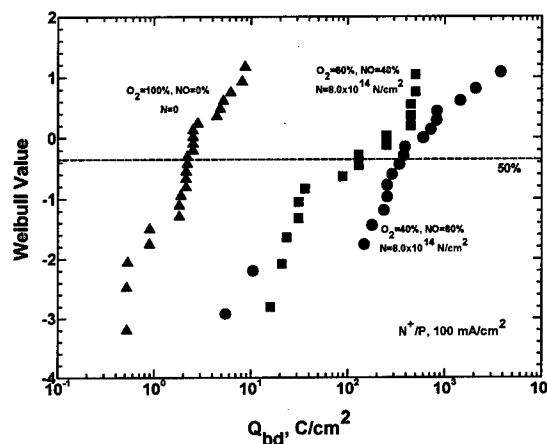
Figure 8. Q_{bd} of various O_2 and O_2-NO films, all 2.0 nm thick, grown at 1000°C by RTO.

We have also observed that reliability can be improved through the incorporation of nitrogen into thin SiO_2. Fig. 8 shows that up to a one hundred-fold increase in Q_{bd} is possible in 2.0 nm oxynitrides grown in O_2-NO mixtures, as opposed to pure O_2.

2.5 Gate leakage current

We have extensively studied the extent to which gate leakage current can be controlled by processing parameters such as starting wafer roughness and pre-oxidation cleaning. Atomic force microscopy (AFM) [7] and scanning tunneling microscopy (STM) [9] were extensively used to characterize wafer surfaces and interfaces. Typical starting wafers had RMS roughness of 0.05 nm measured by AFM, and 0.12 nm measured by STM. STM is more sensitive than AFM because it measures both mechanical and electrical roughness, whereas AFM only measures the former.
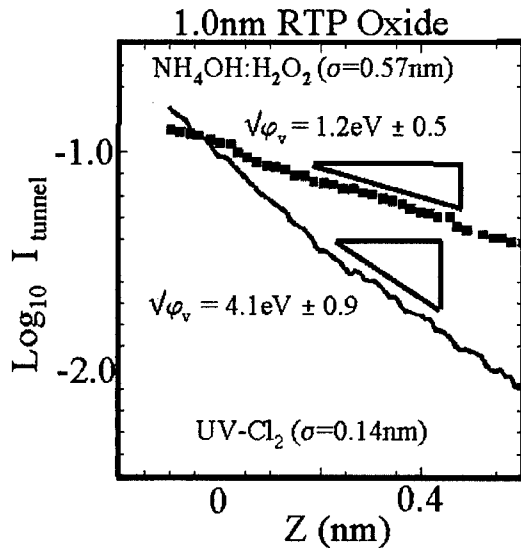
Figure 9. Tunneling current as a function of distance between probe tip and wafer surface, for smooth (UV-Cl₂ cleaned) and rough (NH₄OH:H₂O₂ cleaned) wafers.

A variety of cleans were used to study the effects of cleaning on surface roughness. Fig. 9 is a plot of tunneling current as a function of distance between the STM tip and the surface of the wafer. The slope of the data is proportional to the effective tunneling barrier height. It can be seen that the smoother surface, produced by a UV-Cl₂ vapor phase clean, has a much higher barrier for tunneling than does the rougher surface produced by the NH₄OH/H₂O₂ wet clean. Thus, the smoothest surface is required for the best tunneling behavior, and therefore the lowest gate leakage current.

Taking all of the proper processing precautions into account, sub-50 nm transistors were produced, and their gate leakage currents measured, as is shown in Fig. 10. The leakage current is seen to increase exponentially with decreasing thickness, as expected. These leakage currents represent state-of-the-art values. Assuming a maximum allowable gate current density of 1 A/cm² for desktop applications, and 10⁻³ A/cm² for portable
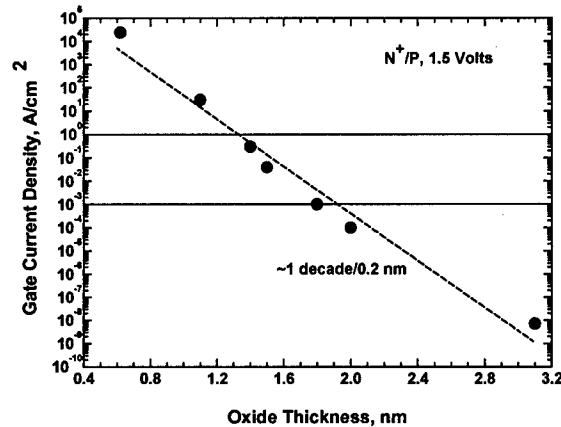
Figure 10. Gate leakage current as a function of oxide thickness for 35 nm transistor structures.

applications, minimum acceptable SiO₂ thicknesses would be 1.3 and 1.9 nm, respectively.

Recently, electron energy loss spectroscopy (EELS) in a scanning transmission electron microscope (STEM) was carried out on ultrathin SiO₂/Si interfaces [10]. Oxygen profiles across the interface were obtained, and are shown in Fig. 11 for SiO₂ films of 1.0 and 1.8 nm thickness. Based on this analysis, it has been calculated that the minimum oxide thickness, before leakage current becomes overwhelming, is 1.2 nm. This comes from the fact that a satisfactory SiO₂ tunnel barrier is formed when it is equal in thickness to six times the decay length of the interfacial states, about 0.7 nm, plus 0.5 nm contribution from interfacial roughness.

3. CONCLUSIONS

We have systematically examined the repercussions of decreasing SiO₂ gate dielectric thickness on several physical and electrical parameters: ultrathin and uniform film growth, boron penetration, reduced drive current, gate oxide reliability and gate leakage current. CMOS applications of SiO₂ or Si-O-N gate dielectrics are neither limited by the ability to
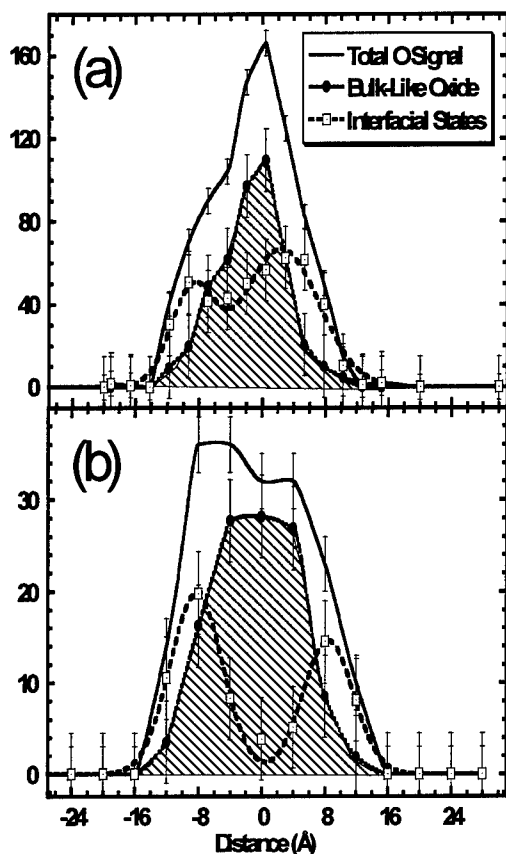
Figure 11. Oxygen bonding profiles measured by EELS. The Si substrate is on the left, and the gate polycrystalline silicon is on the right. a) 1.0 nm (ellipsometric) oxide, annealed 1050°C/10 sec. The bulk-like O signal yields a FWHM of 0.85 nm, whereas the total O signal yields a FWHM of 1.3 nm. The overlap of the two interfacial regions is correlated with a very high gate leakage current, 10^2 A/cm^2. b) a thicker (1.8 nm ellipsometric) oxide, also annealed. The interfacial regions no longer overlap, and the gate leakage current is 10^{-2} A/cm^2.

grow the films commercially, nor to suppress the diffusion of boron. Reliability of SiO$_2$ gate dielectrics will probably be acceptable to thicknesses of 1.5 nm, and can be enhanced by nitrogen incorporation. The fundamental limits imposed on SiO$_2$ are reduced drive current and excessively high gate leakage current. Both of these parameters impose a limit of 1.2-1.3 nm as the thinnest SiO$_2$ acceptable. With potential

improvements in interfacial roughness, the leakage limit may be pushed to about 1.0 nm.

In any event, technology generations that require <1.0 nm equivalent SiO$_2$ thickness will have to rely on alternate, high K gate dielectrics. This transition will not be simple to implement, because SiO$_2$ is the reason that silicon technology has been so successfully integrated thus far. SiO$_2$ has only one disadvantage (low K), and most alternate gate dielectrics have only one advantage, (high K).

4. ACKNOWLEDGMENTS

REFERENCES

1. Semiconductor Industry Association, *The National Technology Roadmap for Semiconductors*, Sematech, Austin, 1997.
2. G. Timp et al., IEDM Technical Digest, (1998) 615.
3. K. S. Krisch, M. L. Green, F. H. Baumann, D. Brasen, L. C. Feldman and L. Manchanda, IEEE Trans. Elec. Dev., 43 (1996) 982.
4. G. Timp et al., IEDM Technical Digest, (1997) 930.
5. J. H. Stathis and D. J. DiMaria, IEDM Technical Digest, (1998) 167.
6. K. J. Hubbard and D. G. Schlom, J. Mater. Res., 11 (1996) 2757.
7. J. Sapjeta, T. Boone, J. M. Rosamilia, P. J. Silverman, T. W. Sorsch, G. Timp and B. E. Weir, Mat. Res. Soc. Symp. Proc. 477 (1997) 203.
8. J. M. Sung, C. Y. Lu, M. L. Chen, S. J. Hillenius, W. S. Lindenberger, L. Manchanda, T. E. Smith and S. J. Wang, IEDM Technical Digest, (1989) 447.
9. T. Sorsch et al., VLSI Technical Symposium, (1998).
10. D. A. Muller, T. Sorsch, S. Moccio, F. H. Baumann, K. Evans-Lutterodt and G. Timp, Nature, 399 (1999) 758.

The role of native traps on the tunneling characteristics of ultra-thin (1.5-3 nm) oxides

A. Ghetti[a], E. Sangiorgi[b], T.W. Sorsch[a] and I. Kizilyalli[a]

[a]Lucent Technologies - Bell Labs,
600 Mountain Ave., Murray Hill, NJ 07974, USA

[b]DIEGM - University of Udine
Via delle Scienze 208, 33100 Udine, Italy

In this paper we report experiments and simulations on thin oxide (1.5-3 nm) MOS devices, showing that native traps can play a dominant role in the tunneling characteristics of such oxides. The number of traps, and thus their role on the tunneling current, can be minimized by careful and simple processing: in this case traps affect the tunneling current in relatively thick oxides while their role vanishes at oxide thicknesses of 2 nm and below.

1. INTRODUCTION

MOS devices with gate oxide thicknesses below 2 nm have been recently fabricated [1,2], where a significant gate tunneling current is present at normal operating conditions. As a consequence the gate current impact on device performance and stand-by power consumption can no longer be neglected [3] and its magnitude may represent one of the most critical parameters for future MOS technologies [4]. Thus, for ultra-thin oxide devices to be feasible, it is important to minimize the gate leakage current for a given oxide thickness, and to characterize the different contributions to the tunneling current.

2. SIMULATION MODEL

The simulation model used in this work [5] includes the self-consistent calculation of Poisson and Schrödinger equations [6] and it is used for both the determination of the physical oxide thickness (t_{ox}), by fitting the CV curves, and the calculation of the tunneling current from bound and free electron states. The model computes both the "unassisted", i.e. direct component (DT), and the "trap-assisted" component (TAT) of the tunneling current. TAT is modeled according to a two-step model [7]. For a known struc-

ture, i.e. a given potential profile, DT computation depends only on the oxide effective mass and barrier height. Since there is a substantial agreement in the literature on the numerical values of these quantities, (respectively $m_{ox} = 0.5m_0$ and $\Phi_B = 3.1eV$), the DT calculation does not rely on device dependent parameters. TAT calculation, instead, depends on the product between the trap capture cross section σ and the trap density distribution $N_{tr}(E,x)$.

3. NATIVE TRAPS

Figs. 1, 2 report the CV and IV characteristics of two virgin N-MOS capacitors (labeled DEVA and DEVB), respectively. They were fabricated over two different but equally lightly doped ($10^{15}cm^{-3}$) p-type epitaxial silicon wafers. DEVA oxide was grown by RTO at $900°C$ on a wafer with initial AFM roughness of $0.0883nm$ RMS, while DEVB oxide was grown by RTO at $1000°C$ on a wafer with initial AFM roughness of $0.117nm$ RMS followed by a 30s anneal in N_2 at $930°C$. Then a 1000$\mathring{A}$ thick amorphous polysilicon layer was deposited. It was implanted with $3 \times 10^{15}cm^{-2}$ arsenic dose at $25keV$ and activated at $1050°C$ for 5s. Finally the wafers were metallized depositing aluminum and patterned using only a single level lithography.
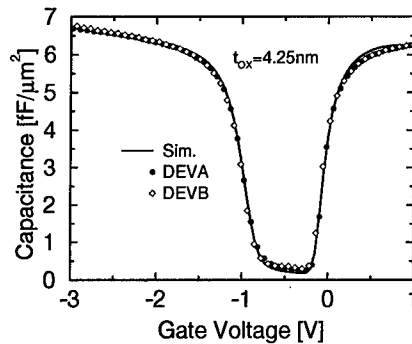
Figure 1. Measured (symbols) and simulated (line) low frequency capacitance as a function of the gate voltage for two N-MOS capacitors (labeled DEVA and DEVB). Oxide thickness extracted by fitting CV measurements using quantum mechanical simulations is $t_{ox} = 4.25nm$.
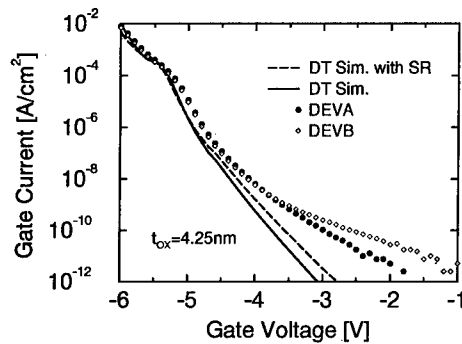
Figure 2. Measured (symbols) and simulated (lines) tunneling IV characteristics of the same devices of Fig. 1. Solid line: DT simulations; dashed line: DT simulations with surface roughness (SR). The extra current with respect to the DT component is due to native traps.

DEVA and DEVB exhibit the same CV characteristics (Fig. 1), the same tunneling current for high enough voltage, but their IV characteristics differ in the low voltage regime (Fig. 2). In addition, DT simulations featuring the t_{ox} extracted by fitting CV data with quantum mechanical simulations (solid line) well reproduce experiments at high voltage, while experiments feature a larger current and a smaller slope with respect to DT simulations at low voltage.

Among the possible explanations of these discrepancies our data allow us to rule out the following possibilities: a) incorrect t_{ox}, because the two devices have the same CV, hence the

same t_{ox}; b) surface roughness (SR), because, although simulation accounting for SR feature a slightly larger current at low voltage (dashed line in Fig. 2), the simulated current is still lower than experimental data. Therefore we conclude that the difference between the experimental gate current of the two devices and the theoretical limit of DT in the low voltage regime is due to an additional tunneling current component assisted by native (intrinsic) traps, present in a different amount and/or distribution in the two devices.
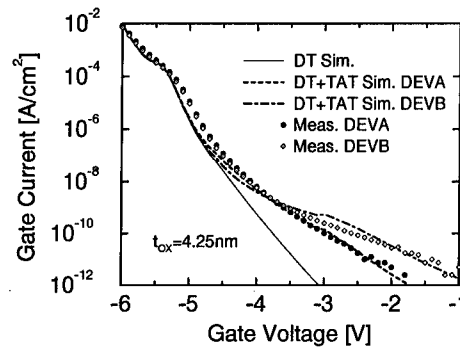
Figure 3. Measurements (symbols) and simulations including native trap assisted tunneling (lines) of the tunneling IV characteristics of the same devices of Figs. 1, 2. Different native trap distributions are needed for the two devices.

Fig. 3 shows the results of simulations performed including TAT component. Here the trap distribution is used as fitting parameter, while, for the sake of simplicity, σ is kept constant at $10^{-16}cm^{-2}$. Lacking reliable data on the oxide trap distribution we assumed a distribution constant in space and variable in energy $(N_{tr}(E))$. Fig. 4 reports the trap distributions adopted for the simulations in Fig. 3. Notice that the falling edge of DEVB's N_{tr} allows us to fit the smoother slope featured by the IV measurements of this device. However, because of the limited amount of data available, we cannot exclude that others $N_{tr}(E, x)$ may fit the data as well.

4. OXIDE THICKNESS DEPENDENCE

To better understand the role of the native traps, we have investigated different (1.5-3 nm)
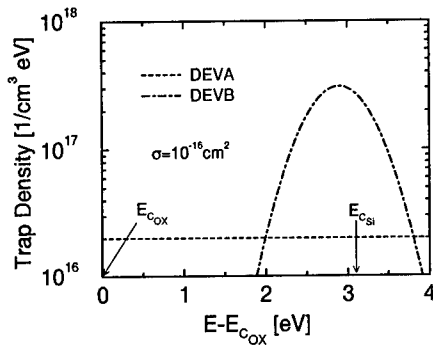
Figure 4. Energy trap profile adopted for the simulation of Fig. 3. The energy is measured downward from the oxide conduction band edge ($E_{C_{ox}}$).

oxide thickness large area ($200 \times 200 \mu m$) N-MOS capacitor devices (CAP), subjected to a simple and careful process. The epitaxial substrate was either left lightly doped ($10^{15} cm^{-3}$), or implanted with $10^{13} cm^{-2}$ boron dose at $16 keV$, resulting in a peak substrate doping concentration of $1.07 \times 10^{18} cm^{-3}$. Different thickness oxides were subsequently grown by RTO at $1000°C$ in pure oxygen. Then, a $200 nm$ thick amorphous polysilicon layer was deposited. The polysilicon layer was implanted with $5 \times 10^{15} cm^{-2}$ arsenic dose at $60 keV$ and annealed at $1050°C$ for $5 s$. Finally the wafers were metallized depositing aluminum and patterned using only a single level lithography.
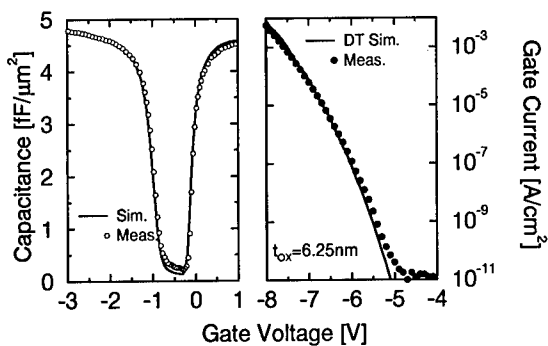
Figure 5. Measured (symbols) and simulated (line) low frequency capacitance (left) and gate tunneling current (right) for a CAP device. Oxide thickness extracted by fitting CV measurements using quantum mechanical simulations is $t_{ox} = 6.25 nm$.

Fig. 5 reports the CV and IV characteristics of a CAP device with $t_{ox} = 6.25 nm$. In this case, DT simulations featuring the t_{ox} extracted by fitting CV data with quantum mechanical simulations (solid line) is very close to the experimental data, showing that the TAT component is not relevant in this case. This observation can be explained by the fact that, in such thick oxide, TAT is below the noise level of our setup or by the reduced number of traps due to a larger distance from the oxide/silicon interface.
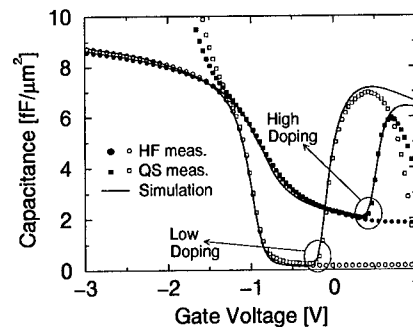
Figure 6. Measured high-frequency (HF) and quasi-static (QS) and simulated (lines) capacitance for CAP devices with $t_{ox} = 3 nm$ and with high (closed symbols) and low (open symbols) doped substrate. QS measurements become inaccurate at high voltage because the large leakage current overcomes the precision limits of our setup.
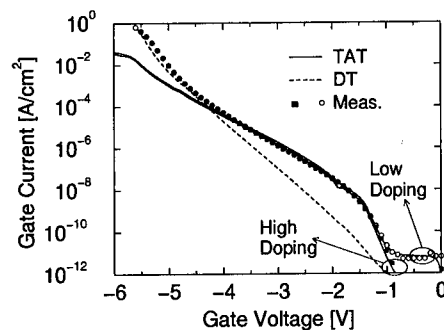
Figure 7. Measured (symbols) and simulated (lines) gate tunneling current for the same devices of Fig. 6. TAT dominates at low voltage while DT prevails at high voltage. The two measurements coincide for $V_G < -1V$.

Figs. 6, 7 report the measured and simulated CV and IV characteristics of two capacitors with

$t_{ox} = 3nm$ and different substrate doping. In these devices the simulated DT is lower than the measured value, indicating that, for the same oxide quality, the TAT becomes more important in relatively thick (3-5nm) oxides. In order to fit the measured IV, a simple trap distribution, uniform in space and energy throughout the SiO_2 forbidden gap, was assumed. The best fitting was found with $\sigma N_{tr} = 30cm^{-1}$, a value comparable with the one of DEVA and DEVB.

Fig. 7 reports the results of the simulations and shows the relative importance of the DT and TAT components. TAT dominates at low applied bias, where there are enough traps to sustain the current and the DT component is still very low. Eventually the DT prevails and the TAT almost saturates due to the limited number of traps.
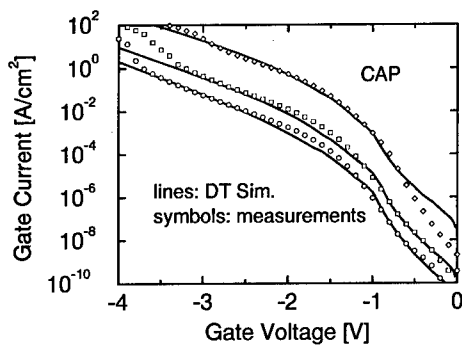
Figure 8. Measured (symbols) and simulated (lines) gate tunneling current for CAP devices with implanted substrate. The oxide thicknesses are (from top to bottom): 1.5 nm, 1.8 nm, 2 nm. Here only DT is needed to fit the experiments. TAT computed using the same parameters as in Fig. 7 is negligible.

In Fig. 8 the measured and simulated IV characteristics of thinner (2-1.5 nm) oxide CAP devices with implanted substrate are shown. In all these cases the DT current alone fits the experiments and the TAT, computed using the same parameters as in Fig. 7, is negligible. Similar results were obtained on CAP devices featuring the same oxide thicknesses but undoped substrate.

This decreasing importance of TAT with decreasing t_{ox} is schematically illustrated in Fig. 9, and it is due to the different functional dependence on t_{ox} of the two components.
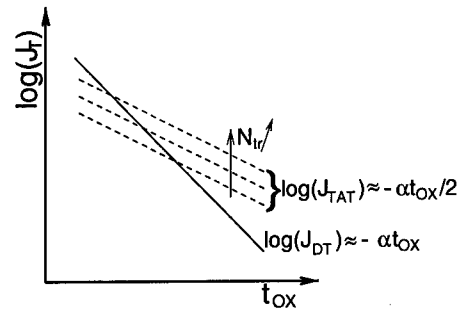
Figure 9. Schematic representation of the tunneling current (J_T) dependence on t_{ox} at a fixed bias. The oxide thickness at which DT = TAT depends on σN_{tr}.

5. CONCLUSIONS

In conclusions we have shown that in ultra-thin oxide MOS devices the leakage current can largely exceed the theoretical value. We propose that the extra component is due to the presence of native traps probably induced by back-end processing. If the number of native traps is kept low by careful and simple processing, the TAT current still shows up in relatively thick oxides but disappears at 2 nm thickness and below.

ACKNOWLEDGEMENTS

REFERENCES

1. H. Momose et al., IEEE Trans. on Electron Devices, vol. 43, p. 1233, 1996.
2. G. Timp et al., in IEDM Technical Digest, p. 930, 1997.
3. S.-L. Lo et al., IEEE Electron Device Letters, vol. 18, p. 209, 1997.
4. Y. Taur et al., IBM Journal of Research and Development, vol. 39, p. 245, 1995.
5. A. Ghetti et al., Submitted to IEEE Transaction on Electron Devices, .
6. F. Stern et al., Phys. Rev., vol. 163, no. 3, p. 816, 1967.
7. E. Suzuki et al., Journal of Applied Physics, vol. 60, no. 10, p. 3616, 1986.

Novel mechanisms in nm-thick gate SiO$_2$ growth at low temperatures utilizing activated oxygen

T.Fuyuki[a], T.Futatsuyama[b], Y.Ueda[b], K.Moriizumi[b], and H.Matsunami[b]

[a]Material Science, Nara Institute of Science and Technology,
8916-5 Takayama, Ikoma, Nara   630-0101   JAPAN

[b]Department of Electronic Science and Engineering, Kyoto University,
Sakyo, Kyoto   606-8501   JAPAN

We propose a novel technique of direct oxidation of Si surfaces using activated oxygen species which can react with Si surfaces to form SiO$_2$ at low temperatures (< 500 ℃). Specific oxidation mechanisms have been revealed, and the atomic level planarization at surfaces and interfaces could be achieved for the first time. Electronic properties of grown films and SiO$_2$/Si interfaces were characterized in detail to show the feasibility for the application to MOSFET devices.

1. INTRODUCTION

For the increasing integration in advanced ULSIs, the formation of nanometer-thick SiO$_2$ gate insulators becomes a crucial process to realize high reliablity and performance. Accurate control of film thickness and uniformity should be strongly required. In addition, low temperature process is also strongly pursued to avoid impurity redistribution and thermal distortion. We have proposed a novel technique of direct oxidation of Si surfaces using activated oxygen [1]. Chemically active oxygen species can react with Si surfaces to form SiO$_2$ at low temperatures (< 500 ℃). Specific oxidation mechanisms have been revealed [1], and the electronic properties of grown films and SiO$_2$/Si interfaces were characterized [2] to show the feasibility for the application to MOSFET devices. In this study, interface structures were studied using the atomic force microprobe (AFM) method, and the atomic level planarization at interfaces and surfaces could be achieved for the first time. Obtained performances of fabricated MOSFETs were also discussed in detail relating with interface electronic properties.
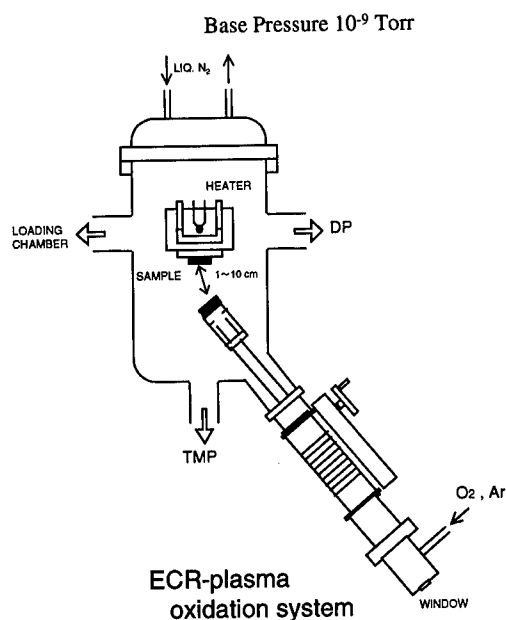
Figure 1. Experimental apparatus.

2. EXPERIMENTS and GROWTH

Atomic-level controlled oxidation could be realized using activated oxygen at as low as 300-500 ℃ [1,2]. An experimental apparatus in Fig.1 was explained in detail elsewhere [1]. The main chamber was evacuated down to a back pressure of 10^{-9} Torr. Activated oxygen species were generated in Ar based ECR plasma with typical conditions of 100W and 10^{-3} Torr. The distance between a substrate and the orifice of the plasma source was about 30 - 100 mm. Substrates were Si (001) wafers, and a conventional RCA cleaning process was
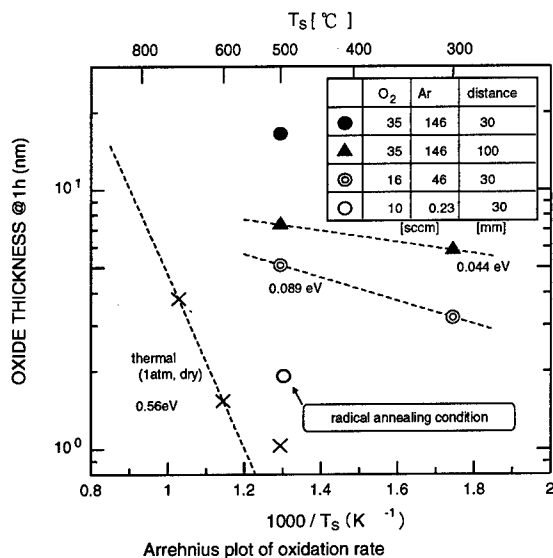
Figure 2. Growth rate as a function of substrate temperature.

used before oxidation. Typical growth rates are shown in Fig.2 as a function of substrate temperature under various oxidation conditions. Optical emission spectroscopy of 777 nm emission related to the excited states of oxygen atom revealed that the amount of activated oxygen species could be controlled by changing Ar/O_2 flow ratio. The oxidation rate was varied in a range of 3 - 15 nm /hour, which improves the feasibility in ultra-thin SiO_2 formation. It should be noted that in a range of 300 - 500 ℃, the oxidation rate has found to be almost independent on the substrate temperature (see ▲ (activation

energy : 0.044 eV) and ◎ (activation energy : 0.089 eV) in Fig.2,) at a fixed Ar/O_2 ratio and also found to be much higher than the thermal oxidation case (✕). Activated oxygen species (atomic oxygen at excited and ground states and/or excited molecular oxygen) play a dominant role in oxidation.
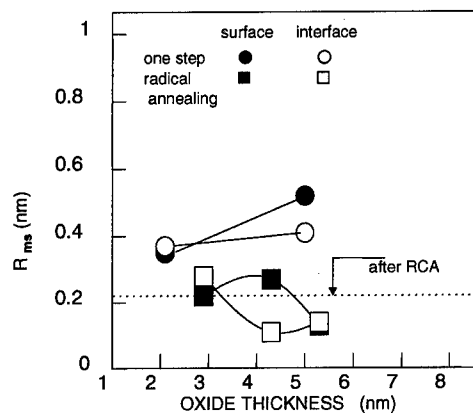
Figure 3. Surface and interface roughness.

We have found a novel phenomenon of planarization at interfaces and surfaces by activated oxygen treatment. After oxide growth at a fixed thickness a sample was irradiated by activated oxygen species for 10 min at the condition shown by ○ in Fig.2. A total film thickness was not increased at all (less than 0.2 nm), by this "radical annealing", but roughness at surfaces and interfaces was improved drastically. The scanning area in AFM measurement was 0.6 $\mu m^{\square}$, and the lateral and depth resolutions were 0.05 and 0.01 nm, respectively. Figure 3 shows the values of R_{ms} (root mean square) st surfaces and interfaces as a function of film thickness. The roughness at interfaces was measured after removal of SiO_2 by diluted HF. After an oxidation using activated oxygen, roughness increased monotonously with increasing thickness as shown by ●, ○, on the contrary, by introducing the radical annealing after oxidation roughness decreased down to the as-prepared value or less shown by ■, □.

The minimum value was about 0.15 nm, which was equivalent to a single atomic step. Detail mechanisms of planarization are now investigated using AFM observation under various oxidation conditions.

3. ELECTRONIC CHARACTERIZATION
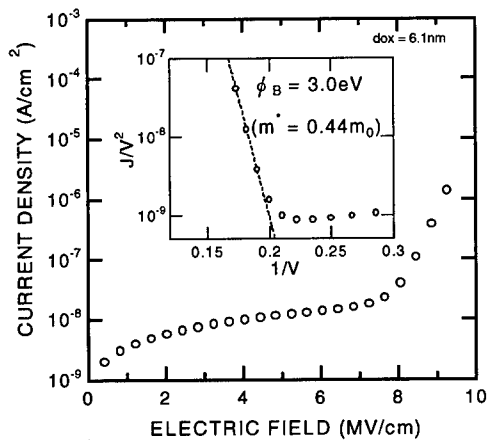
Figure 4. Typical current/voltage characteristics.

Electronic properties of grown SiO_2 films and interfaces were characterized by current / voltage and capacitance / voltage measurements using metal / oxide / semiconductor (MOS) structures fabricated on n-type Si surfaces. Without "radical annealing" grown films showed rather poor qualities partly due to rough interfaces as mentioned above. We did not observe the clear relationship between roughness and electronic properties such as resistivity and/or interface state density quantitatively, but we could achieve high-quality electronic properties reproducibly by "radical annealing". In Fig.4, a typical current / voltage characteristic is shown for a SiO_2 thickness of 6.1 nm. A high resistivity of 10^{14} - $10^{15}\,\Omega\,cm$ was achieved at a low electric field strength.
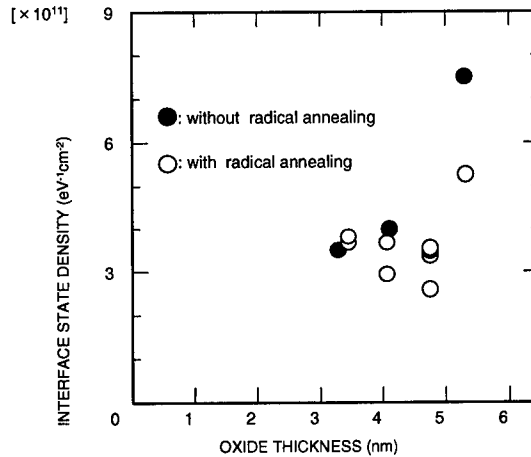
Figure 5. Interface state density.

The steep increase in a range higher than 8 MV/cm was caused by a Fowler-Nordheim (F-N) tunneling current. A F-N tunneling plot shown in the insertion gave a barrier height (Si/SiO_2) of 3.0 eV with an electron effective mass of 0.44. The derived barrier height is a little bit smaller than the reported value (3.2 eV) in the case of thermally grown one. But, the electronic band diagram at the interface formed by activated oxygen was considered to be almost identical to that made by a thermal process at high temperatures [3] if we take into account of the measured values distributed in a range of 3.0 - 3.2 eV in our experiments. Capacitance / voltage characteristics showed no hysteresis, and the estimated minimum interface state density was as low as $3\times10^{11}\,cm^{-2}eV^{-1}$ as shown in Fig.5.

MOSFETs with a channel length/width of 30/250 μ m were fabricated on p-Si (001) surfaces. A maximum effective electron mobility of 130 - 300 cm^2/Vs was realized, which reflects the excellent interface properties as discussed above. To authors' knowledge, the achieved mobility is one of the high values reported up to now for low-temperature processed gate insulators [4], but

is about 20 - 40 % in the case of thermally grown SiO_2. Figure 6 shows a typical example of the effective electron mobility as a function of measurement temperature. A rather high value (10^{12} - 10^{13} cm^{-2}) of fixed charge density was deduced by comparing the experimental values with reported ones [5] in the range of coulomb scattering limited region. The fixed charge was incorporated during FET fabrication processes since rather low fixed charge density of 10^{11} cm^{-2} could be realized in the case of a simple MOS structure process. Further improvement of FET fabrication processes including interface modification by hydrogen annealing and/or ozone treatment is required.

4. CONCLUSIONS

Novel growth mechanisms in nanometer - thick SiO_2 growth using activated oxygen were presented. Precise thickness control in oxidation could be achieved, and interface planarization in an atomic level was demonstrated for the first time. Tunneling transport through ultra-thin SiO_2 showed high-quality electronic properties as a thermal process at high temperatures. Fabricated FETs were studied in detail, using mobility / temperature dependence, and perspectives for further improvement was discussed.
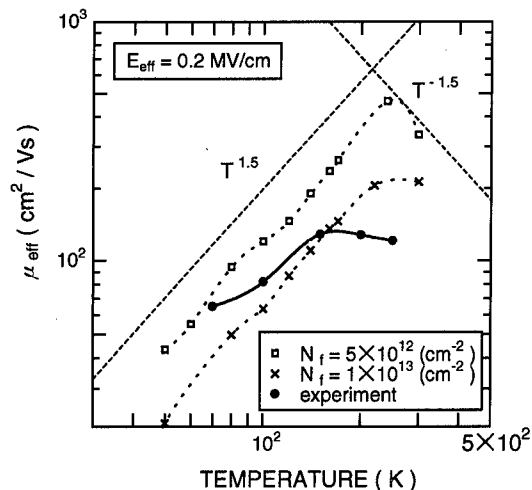
Figure 6. Effective electron mobility.

REFERENCES

1. T.Fuyuki, S.Muranaka, and H.Matsunami; Applied Surface Science 117/118 (1997) 123-126.
2. T. Fuyuki, T. Futatsuyama, Y.Ueda, K. Moriizumi and H. Matsunami;Proc. Intn'l Conf. Solid State Devices and Materials, Hiroshima, Japan Sep. 1998, pp.104-105.
3. T.Yoshida, S.Miyazaki, and M.Hirose; Proc. Intn'l Conf. Solid State Devices and Materials, Yokohama, Japan, 1996, pp.539-541.
4. J.Tasiak, J.Batey, E.Tierney, and J.Li;IEEE Elect. Dev. Lett. 10 (1989) 245.
5. C.T.Sah, T.H.Ning, and L.L.Tschopp, Surface Science, 32 (1972) 561-575.

The Effects of Ge Content in Poly-$Si_{1-x}Ge_x$ Gate Material on the Tunneling Barrier in PMOS Devices*

A. Shanware,[1] H. Z. Massoud,[1] A. Acker,[2] V. Z.-Q. Li,[2] M. R. Mirabedini,[2] K. Henson,[2] J. R. Hauser,[2] and J. J. Wortman[2]

[1]Semiconductor Research Laboratory, Department of Electrical and Computer Engineering, Duke University, Durham, North Carolina 27708–0291, U.S.A.

[2]Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27695, U.S.A.

The use of SiGe gates in MOSFET technology has promise as a single-gate material for both n- and p-channel MOSFETs. The Ge content in the gate, however, affects the gate energy band diagram. While Ge in the SiGe gate does not affect the conduction-band energy level, it is found to raise the valence-band energy level and reduce the gate bandgap. This change results in an increase in the gate current resulting mainly from the tunneling of electrons from the valence band of the gate in PMOSFETs. This paper reports on the effects of Ge content in SiGe gates on the tunneling characteristics of PMOSFETs.

1. INTRODUCTION

The presence of Ge in the poly-Si gate of PMOS devices is reported to change their flatband voltage [1–4]. It has been found that an increase in the Ge content decreases the flatband voltage, thus improving the current drive and transconductance. Such a technology promises future single-gate materials for both p- and n-channel devices. The purpose of this work is to study the effects of Ge content in SiGe gates on the gate tunneling current of PMOS structures biased for gate injection and to determine tunneling parameters such as the barrier height and the electron effective mass. The gate current in p^+-gate PMOSFETs biased with $V_{GS} < 0$ is attributed to the tunneling of electrons from the valence band of the SiGe gate, as illustrated in Fig. 1. In this work, it is found that the tunneling barrier height is reduced by as much as 0.395 eV when the percentage of Ge is increased from 0 to 60% in the gate. As expected, the electron effective mass in the oxide is not affected.

Tunneling Mechanisms in PMOSFETs

p^+ Poly Gate SiO_2 n–type Substrate
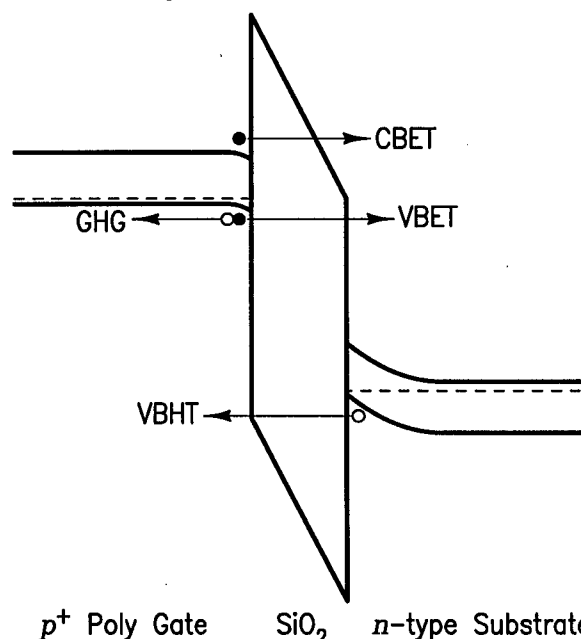
Fig 1. Tunneling mechanisms in PMOSFETs. Electron tunneling from the conduction (CBET) is negligible due to the small concentration of electrons in the p^+-gate.

The presence of Ge in the gate is found not only to change the flatband voltage but also the barrier height of electrons tunneling from the valence band. This observation was recently reported for Poly-Si$_{0.7}$Ge$_{0.3}$ gates [5]. Our present work studies the effect of the Ge percentage on the tunneling barrier height. We report that the barrier height decreases with the increase in Ge content and that the reduction can be as much as 0.395 eV when the Ge percentage reaches 60%.

2. EXPERIMENTAL RESULTS

PMOS devices with oxide thicknesses ranging from 35 to 70Å were fabricated. In-situ boron-doped poly-Si$_{1-x}$Ge$_x$ gates were deposited in a cold-wall RTCVD system, at 560-580°C and 4 Torr, using a mixture of Si$_2$H$_6$, GeH$_4$, and B$_2$H$_6$ gas. Poly-Si$_{1-x}$Ge$_x$ gates with different Ge content were obtained by changing the ratio of GeH$_4$ and Si$_2$H$_6$. The film composition was determined by energy-dispersive X-ray spectrometry (EDX). The Ge content in the SiGe gate was changed from 0 to 60%.

The oxide thickness and flatband voltage were determined by fitting the $C(V)$ characteristics. The experimental $C(V)$ data and model fits are shown in Fig. 2.

Figure 3 shows the experimentally observed dependence of the gate current density on the oxide electric field for devices with different Ge percentages. It is found that the Ge presence in the gate does not change the barrier height of the electrons tunneling from the substrate. Figure 4 shows the experimental results for electrons tunneling from the gate. The trend in this figure indicates that for the same electric field in the oxide, the gate current density for gate injection increases when the percentage of Ge in the gate is increased. We attribute this increase in J_G to a reduction in the barrier height for electrons tunneling from the valence band of the gate.

Extraction of the electron effective mass in the oxide and the tunneling barrier height is carried out with the help of Fowler-Nordheim plots where we fit $\ln(J_G/\mathcal{E}_{OX}^2)$ vs $1/\mathcal{E}_{OX}$, where $\mathcal{E}_{OX}$ is the oxide electric field, as is shown in Fig. 5. The barrier height was found to be 4.3 eV in the case of PMOS devices with no Ge in the polysilicon gate material.

Fig 2. The capacitance-voltage $C_{gb}(V_{GB})$ characteristics of three MOSCAPS with SiGe gates containing 0, 40, and 64% Ge.
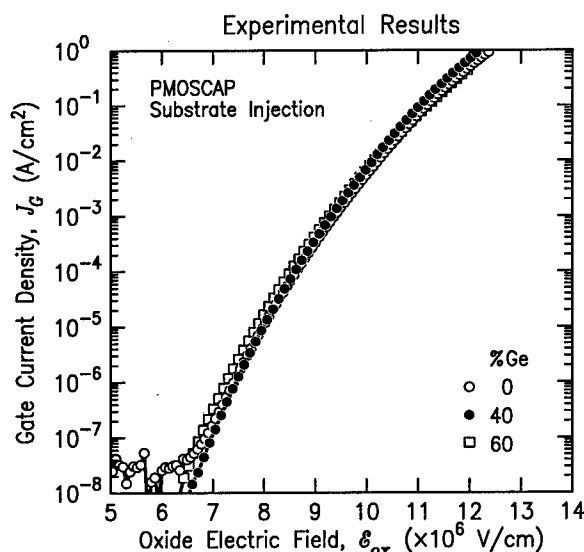
Fig 3. Dependence of the DC current-voltage $J_G(\mathcal{E}_{OX})$ characteristics on Ge percentage for substrate injection.

Experimental Results
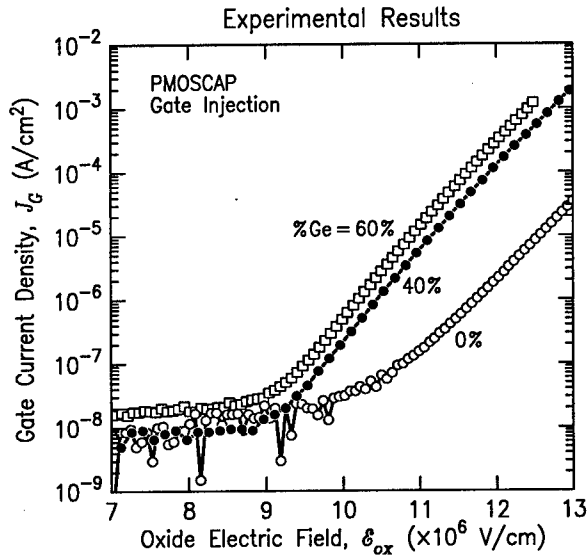
Fig 4. Dependence of the DC current-voltage $J_G(\mathcal{E}_{OX})$ characteristics on Ge percentage for gate injection.

Fit Using the Fowler–Nordheim Equation
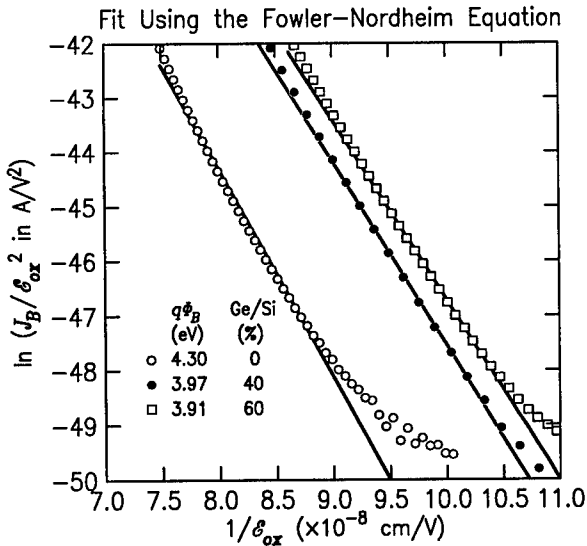
Fig 5. Influence of the percentage of Ge in the SiGe gate on the Fowler-Nordheim plots for gate injection. The solid lines are the straight-line FN fits to the data.

Effect of % Ge on the Barrier Height
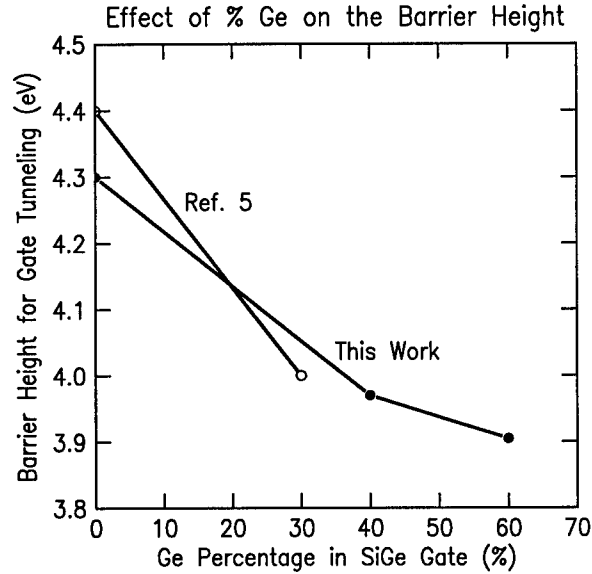
Fig 6. Influence of the percentage of Ge in the SiGe gate on the barrier height of electron tunneling from the gate.

The electron effective mass in the oxide was found to be $0.39m_0$ for all cases. It is not expected that this tunneling parameter would be affected by the presence of Ge in the gate.

3. DISCUSSION

Figure 6 shows our results for the effect of Ge content in SiGe gate on the decrease in the barrier height for electron tunneling from the valence band in comparison with the results obtained by Salm et al. [5]. Lee et al. [4] and Salm et al. reported that the change in the workfunction of NMOSFETs with n^+-poly SiGe gates was negligible and made similar conclusions that the conduction-band energy level is hardly affected by the presence of Ge in SiGe gates. From their results and the results obtained in our study, we can conclude that, as the Ge percentage is increased in SiGe gates, the conduction-band energy level is hardly changed and that the valence-band energy level is strongly affected in a direction that results in the reduction of the SiGe band gap as schematically shown in Fig. 7.
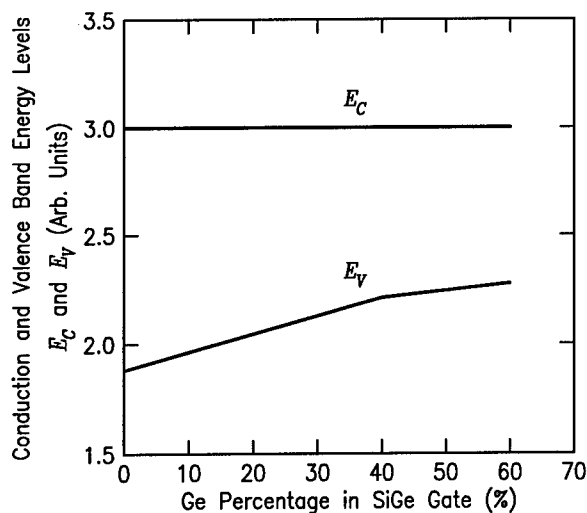
Fig 7. Dependence of the conduction-band energy level E_C and valence-band energy level E_V on the Ge percentage in poly-SiGe gates.
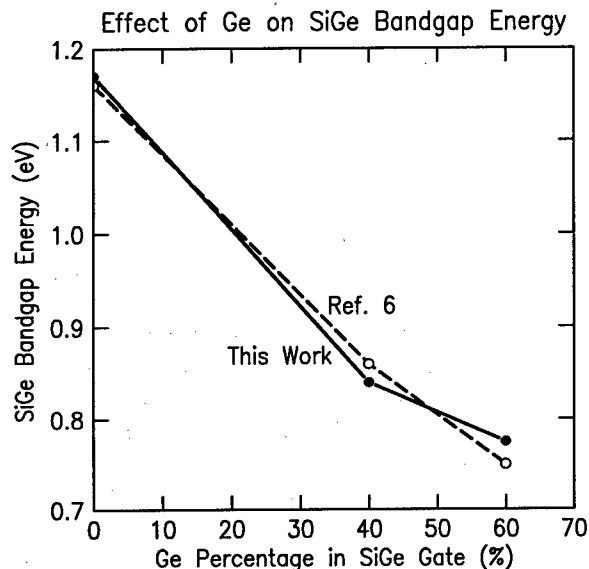
The results reported by Salm *et al.* [5] and compared to our results only extend to Ge content of up to 30%. For Ge percentages exceeding this value, we compare our results with those obtained by Lang *et al.* [6] using optical photocurrent spectroscopy to measure the optical absorption spectra of coherently strained layers of SiGe. It can be easily seen in Fig. 8 that the decrease in the tunneling barrier of electrons obtained in our study closely matches the decrease in the bandgap of SiGe material measured using optical methods.

We conclude that the use of poly-SiGe as the gate material for dual-gate CMOS transistors has to take into account the increase in the gate current consisting of electrons tunneling from the valence band of PMOSFET gates due to the presence of Ge. The germanium content in the gate should be optimized to limit the gate tunneling current in PMOSFETs especially at gate biases larger than 1 V. This effect should be considered a process integration issue in the optimization of ULSI devices.

4. CONCLUSIONS

The presence of Ge in the p^+ gate of PMOS devices reduces the barrier height of electrons tunneling from the valence band of the gate. This result poses a limitation on the use of SiGe gates in devices with ultrathin dielectrics.

REFERENCES

1. V. Z.-Q. Li, *et al.*, pp. 833, IEDM Digest 1997.
2. T.-J. King, *et al.*, pp. 253, IEDM Digest 1990.
3. Y. V. Ponomarev, *et al.*, pp 829, IEDM Digest 1997.
4. W.-C. Lee, *et al.*, IEEE Elec. Dev. Lett., 19, 247 (1998).
5. C. Salm, *et al.*, IEEE Elec. Dev. Lett., 19, 213 (1998).
6. D. V. Lang *et al.*, Appl. Phys. Lett., 47, 1333 (1985).

Fig 8. The reduction in the bandgap of SiGe gate due to increasing percentage of Ge in gate. Experimental results obtained in this work are compared with direct measurements of the bandgap obtained by optical methods [6].

X-ray irradiation effect on the reliability of ultra-thin gate oxides and oxynitrides

M. Houssa [a], S. De Gendt [a], P. de Bokx [b], P.W. Mertens [a] and M.M. Heyns [a]

[a]IMEC, 75 Kapeldreef, B-3001 Leuven, Belgium

[b]Philips Research Laboratories, 4 Prof. Holstlaan, 5656 AA Eindhoven, The Netherlands

During X-ray irradiation of MOS devices, interface states as well as bulk electron and hole traps are generated at the Si/SiO_2 interface and in the SiO_2 layer respectively. In this paper, we investigate the effect of X-ray irradiation on the electrical characteristics of 4.2 nm oxide and 4.4 nm oxynitride layers. It is shown that the bulk electron traps generated during irradiation lead to an increase of the gate current in the low gate voltage region, behaviour which can be attributed to trap assisted tunneling. From the analysis of time-dependent dielectric breakdown characteristics, it is demonstrated that the irradiation-induced bulk traps participate in the formation of the breakdown (percolation) path in the gate oxide layer.

1. INTRODUCTION

The effect of X-ray irradiation induced damage on the electrical properties of MOS (metal-oxide-semiconductor) devices has been widely investigated [1-4]. Ionizing radiations generate interface states at the SiO_2/Si interface as well as bulk electron and hole traps in the gate oxide layer. The generation of these traps results in the degradation of the device properties. While most of the reported works on X-ray irradiation damage of thin gate oxides have focused on the investigation of interface states, not much work has been performed on the study of bulk oxide traps [5]. In this paper, we report the effect of X-ray irradiation on the electrical properties and reliability of ultra-thin gate oxides and oxynitrides. It is shown that the bulk oxide traps generated during the X-ray irradiation behave quite similarly to the neutral electron traps induced by electrical stress of the gate oxide. It is also demonstrated that the irradiation-induced traps participate in the formation of the breakdown path in the SiO_2 layer.

2. EXPERIMENTAL

MOS capacitor structures with 4.2 nm oxide and 4.4 nm oxynitride layers and with P-doped poly-Si gate electrodes were patterned on n-type (100) Si wafers. The wafers were irradiated at 5 different spots (3 cm in diameter) during 120 s. The X-ray beam was applied perpendicular to the wafer surface. The X-ray tube (Rh anode) was set to 24 kV and 125 mA. The estimated irradiation dose was estimated to be 2.5 J/cm² [Si].

3. RESULTS

The current-voltage characteristics of 1.96×10^{-5} cm² area capacitors with a 4.4 nm oxynitride layer are shown in Fig.1. One observes that the leakage current is increased in the low gate voltage region after X-ray irradiation [6], a behaviour quite similar to the stress-induced leakage current (SILC) observed in devices stressed under a constant gate current or gate voltage [7,8].
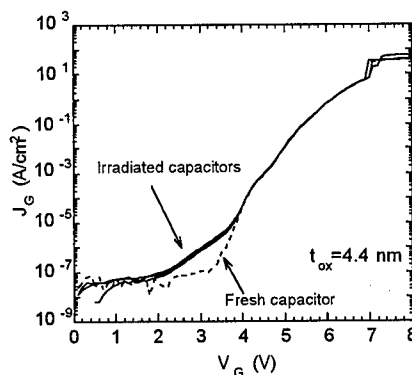
Fig.1. Current-voltage characteristics of MOS capacitors with a 4.4 nm oxynitride layer before and after X-ray irradiation.

The radiation-induced leakage current (RILC) has been analysed within a trap assisted tunneling model [9]. The trap-assisted tunneling current can be approximated by the following expressions [9]

$$J_{TAT}^{DIR} = C \exp\left[-\frac{2\sqrt{2m^*}\left(\phi^{3/2} - (\phi - V_{ox})^{3/2}\right)}{3\,e\,\hbar\,E_{ox}}\right] \quad (1)$$

$$J_{TAT}^{FN} = C \exp\left[-\frac{2\sqrt{2m^*}\,\phi^{3/2}}{3\,e\,\hbar\,E_{ox}}\right] \quad (2)$$

in the direct (DIR) and Fowler-Nordheim (FN) tunneling regime respectively. In the above equations, $C = e\,t_{ox}\,D_{ot}/\tau_o$, where e is the electron charge, t_{ox} the oxide thickness, D_{ot} the density of bulk neutral traps, τ_o the de-trapping time and E_{ox} the oxide field.
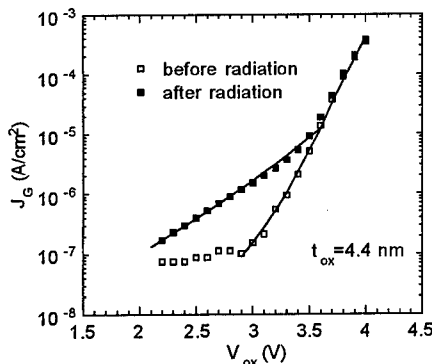
Fig.2. J_G-V_{ox} characteristics of a 4.4 nm gate oxide layer before and after X-ray irradiation. Solid lines are fits to the data obtained with a trap assisted tunneling model [9].

We have used Eq. (1) and (2) in order to fit the current-voltage characteristics of MOS capacitors before and after X-ray irradiation, see Fig.2. The theoretical curves shown in Figs. 2 were obtained by fixing the following values of the physical parameters: $m^* = 0.4\,m_o$, $\tau_o = 2\times10^{-12}$ s, $\phi = 3.2$ eV for tunneling from the conduction band (fresh capacitors) and $\phi = 4.2$ eV for tunneling from the valence band (irradiated capacitors) [9]. The excellent agreement between the experimental results and the theoretical curves suggests that the radiation-induced leakage current can be attributed to the trap assisted tunneling mechanism, like the stress-induced leakage current [8,9]. From these fits, the density of bulk electron traps D_{ot} induced by the X-ray irradiation

can be extracted, see Table 1. One can see that the density of bulk traps increases by about 2 orders of magnitude after irradiation.

The time-to-breakdown t_{BD} distributions before and after irradiation of MOS capacitors with a 4.4 nm oxynitride layer are presented in Fig.3. These distributions were obtained by stressing 20 devices (1.26×10^{-3} cm^2 area) under a constant gate voltage stress $V_G = 5.6$ V (substrate injection). One observes that the t_{BD} values are lower after X-ray irradiation and that the slope of the Weibull distribution β is smaller after irradiation.
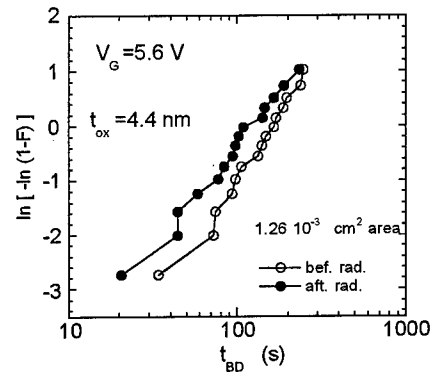
Fig.3. Time-to-breakdown distributions of MOS capacitors with a 4.4 nm oxynitride layer.

In Fig. 4, the intrinsic (50%) t_{BD} values are presented as a function of the capacitor area for the 4.2 nm oxide and 4.4 nm oxynitride layer. One observes that the 50% t_{BD} values are lower after X-ray irradiation, and that this relative decrease becomes more important as the capacitor area increases. From the data shown in Fig.4, the slope of the time-to-breakdown distributions can be extracted [10,11] before and after X-ray irradiation according to the expression

$$t_{BD}^{A2} = \left(A_1 / A_2\right)^{1/\beta} t_{BD}^{A1} \quad (3)$$

where t_{BD}^{A2} (t_{BD}^{A1}) is the intrinsic time-to-breakdown corresponding to area A_2 (A_1) and β is the Weibull slope of the time-to-breakdown distribution.

Table 1. Density of bulk electron traps D_{ot} obtained from RILC analysis, Weibull slope β and corresponding critical trap density $D_{ot,crit}$ as obtained from the analysis of t_{BD} data. $\Delta D_{ot,crit}$ is the difference between the critical density of electrically induced traps before and after irradiation, i.e. the traps induced during irradiation.

Sample		D_{ot} (cm^{-3}) (RILC)	β	$D_{ot,crit}$ (cm^{-3}) (t_{BD})	$\Delta D_{ot,crit}$ (cm^{-3})
4.2 nm	bef. rad.	6.8×10^{16}	3.05 ± 0.16	1.5×10^{19}	9.7×10^{18}
	aft. rad.	1.2×10^{19}	2.22 ± 0.15	5.3×10^{18}	
4.4 nm	bef. rad.	1.6×10^{17}	2.36 ± 0.12	8.1×10^{18}	6.9×10^{18}
	aft. rad.	7.6×10^{18}	1.86 ± 0.14	1.2×10^{18}	

The values of β extracted from the data shown in Fig.4 are given in Table 1. One can see that β decreases after irradiation, i.e. a wider distribution of breakdown times is observed after irradiation.
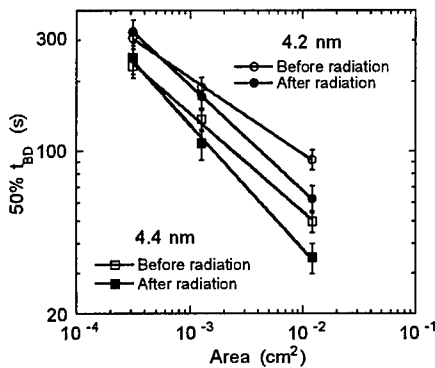
Fig.4. Area dependence of the intrinsic time-to-breakdown values of 4.2 nm and 4.4 nm gate oxide layers before and after X-ray irradiation.

It has been shown [12] that the slope of the Weibull distribution is related to the critical density of neutral

electron traps $D_{ot,crit}$ generated during the electrical stress and needed to trigger breakdown (or soft breakdown) of the gate oxide layer. The Weibull slope β is plotted versus $D_{ot,crit}$ in Fig.5. This curve was obtained from Monte-Carlo simulations, assuming that electron traps are generated randomly in the SiO_2 layer and form a percolation path between the cathode and anode of the capacitor at breakdown or soft breakdown [12,13].
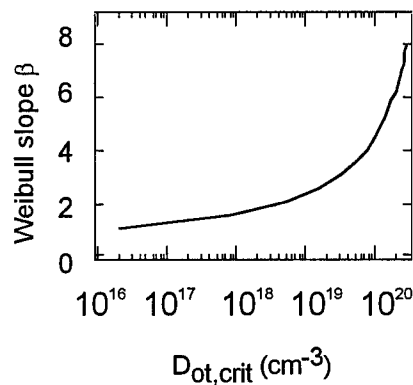
Fig.5. Weibull slope β as a function of the critical density of bulk oxide traps $D_{ot,crit}$.

This percolation model predicts that $D_{ot,crit}$ is lowered as β decreases. From the data shown in Fig.4 and the results summarized in Table 1, it can be deduced that the critical density of *electrically-induced traps* at breakdown (or soft breakdown) is lowered in irradiated capacitors, namely *X-ray irradiation-induced traps are participating in the formation of the breakdown path*. By substracting the $D_{ot,crit}$ values before and after irradiation, the bulk electron trap density $D_{ot} = \Delta D_{ot,crit}$ *induced by irradiation* can be estimated, see. Table 1. These latter D_{ot} values are quite close to those obtained from the analysis of the irradiation-induced leakage current. Furthermore, the probability of observing soft breakdown is increased from about 20% before irradiation to about 60% after irradiation. The bulk electron traps generated during X-ray irradiation thus favor the occurrence of soft breakdown in the oxide, i.e. these traps activate the formation of the percolation path responsible for the time-dependent dielectric breakdown.

4. CONCLUSIONS

The effect of X-ray irradiation on the electrical characteristics of ultra-thin gate oxides and oxynitrides has been investigated. It has been shown that the tunneling current is increased in the low gate voltage region after the irradiation of MOS capacitors with 4.2 nm oxide and 4.4 nm oxynitride. This phenomenon is closely related to the stress-induced leakage current observed in electrically stressed MOS devices and is due to the generation of X-ray irradiation-induced bulk electron traps in the SiO_2 band gap, traps which increase the probability of electrons to tunnel from the Si substate to the poly-Si gate. The current-voltage characteristics of irradiated capacitors have been quite well reproduced by using an approximate expression for the trap assisted tunneling current. From these fits, realistic values of the X-ray induced bulk oxide traps D_{ot} have been obtained. Next, we have shown that the time-to-breakdown values obtained by stressing MOS capacitors with a constant gate voltage are decreased in irradiated capacitors, the more so in the largest tested capacitor area. By plotting the intrinsic t_{BD} values as a function of the capacitor area, the slope β of the Weibull distribution has been extracted before and after X-ray irradiation. It has been shown that β decreases after irradiation, which implies that the critical density of neutral electron traps needed to

trigger breakdown (or soft breakdown) in the gate oxide is decreased after irradiation. By substracting the values of the critical trap density before and after X-ray irradiation, the bulk electron trap density generated by the irradiation could be also extracted. It has been shown that consistent results were obtained from the analysis of the radiation-induced leakage current and the time-to-breakdown characteristics. These results strongly suggest that the X-ray induced bulk electron traps participate in the formation of the percolation path and which triggers the breakdown or soft breakdown of the gate oxide layer.

REFERENCES

1. P.U. Kenkare ans S.A. Lyon, Appl. Phys. Lett. **55**, 2328 (1989).
2. B.J. Mrstik and R.W. Rendell, IEEE Trans. Nucl. Sci. **38**, 1101 (1991).
3. P.M. Lenahan and J.F. Conley, Appl. Phys. Lett. **71**, 3126 (1997).
4. G.S. Ristic, M.M. Pejovic, and A. Jaksic, J. Appl. Phys. **83**, 2994 (1998).
5. V.V. Afanas'ev, J.M.M. de Nijs, A. Stesmans and P. Balk, Microelec. Eng. **28**, 43 (1995).
6. A. Scarpa et al., IEEE Trans. Nucl. Sci. **44**, 1818 (1997).
7. N.K. Patel and A. Toriumi, Appl. Phys. Lett. **64**, 1809 (1994).
8. J. De Blauwe et al., IEEE Trans. Electron Devices **45**, 1745 (1998).
9. B. Ricco, G. Gozzi, and M. Lanzoni, IEEE Trans. Electron Devices **45**, 1554 (1998).
10. T. Nigam et al., in *Proc. of the 1998 IRPS* (Reno, USA), p.62.
11. J. Sthatis and D. DiMaria, in *IEDM Tech. Dig.* (IEEE, 1998), p. 167.
12. R. Degraeve et al., in *IEDM Tech. Dig.* (IEEE, 1995), p. 863; R. Degraeve et al., IEEE Trans. Electron Devices **45**, 904 (1998).
13. M. Houssa et al., J. Appl. Phys. **84**, 4351 (1998); M. Houssa et al., in *IEDM Tech. Dig.* (IEEE, 1998), p.909.

Investigation of temperature acceleration of thin oxide time-to-breakdown

B. Kaczer, R. Degraeve, N. Pangon, T. Nigam, and G. Groeseneken

IMEC, Kapeldreef 75, B-3001 Leuven, Belgium

We study the effect of elevated temperature on the thin oxide breakdown. We find that the Arrhenius law does not describe the process well—in contrast to that we observe the log of the time-to-breakdown to scale linearly with temperature for a wide range of oxide thicknesses. We further observe that both the hole fluence and the neutral electron trap density at breakdown decrease at increased temperatures and do not reach constant critical values. We also show how elevated temperature affects the reliability of oxide films and predict insufficient reliability for oxides below 2.5 – 2.8 nm.

1. INTRODUCTION

As the downscaling trend continues, the operating temperature of logic devices is likely to rise. Integrated circuits based on more conservative technologies used in certain specialized mission-critical applications, such as telephone exchanges, automobile and jet engines, etc., are also required to operate at elevated temperatures. Moreover, already during fabrication, devices are typically exposed to increased-temperature stress during various plasma-processing steps. Thorough understanding of the degradation phenomena caused by electrical stress at elevated temperatures is therefore critical for proper assessment of the reliable device lifetime.
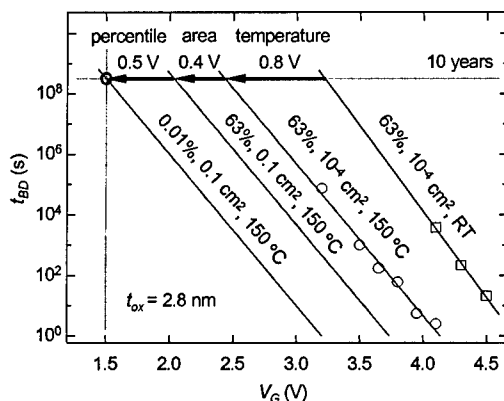
Figure 1. Prediction of the maximum operating voltage for the 2.8 nm oxide by considering T-acceleration, area scaling, percentile recalculation, and a linear projection of the gate voltage V_G.

2. CONSEQUENCES FOR OXIDE RELIABILITY

The temperature acceleration of oxide breakdown constitutes an additional constraint for further device downscaling. For example, the effect of elevated temperature T on the lifetime of 2.8 nm oxide—typical for the 0.15 μm CMOS technology—is demonstrated in Fig. 1 in combination with area scaling and percentile recalculation (these two effects stem from the breakdown statistics [1]). Typically specified maximum device failure of 0.01% after 10 years of continuous operation at 150°C and effective gate area of 0.1 cm^2 is achievable only at ~1.5 V—the expected operating voltage for the 0.15 μm technology.

Furthermore, this narrow reliability margin will completely disappear for oxide thicknesses below ~2.5 – 2.8 nm (depending on the exact operating conditions), which corresponds to the 0.13 – 0.15 μm technologies [2]. Consequently, for technologies below 0.13 μm, the insufficient T-accelerated gate oxide reliability might constitute the end of the SiO_2-based gate dielectrics era.

We, however, have to point out that the above predictions assume a linear V_G projection to operating voltages, which cannot be reasonably tested experimentally and may not be correct. Also, the exact T-dependence of the Weibull slope β, which affects the percentile and area scalings [1], is not completely clear. A certain hope also remains that β could be positively influenced (i.e., increased) by, for example, particular processing of the gate oxide.

3. TEMPERATURE DEPENDENCE OF TIME-TO-BREAKDOWN

T-acceleration of time-to-breakdown t_{BD} is typically assumed to be governed by the Arrhenius law. However, a wide range of activation energies is reported in the literature, while the activation energies frequently depend on the oxide thickness, applied electric field, or even the temperature range (see e.g., Ref. [3]). That already suggests that this law might not be well suited for describing this acceleration process. From our own measurements, carried out over a wide range of oxide thicknesses between the room temperature (RT) and 200°C we conclude that the Arrhenius law does not acceptably describe the observed T-dependence of t_{BD} (Fig. 2).

In Fig. 3(a) we present t_{BD} data measured as a function of T and the stress current J for the 7.3 nm oxide. Figure 3 demonstrates that the log of the measured t_{BD}'s can be very well linearized with respect to T [Figs. 3(b) and (c)] and ln J (this dependence agrees with the $1/E_{ox}$ model). When the

Figure 3. (a) Measured dependence of t_{BD} on T and J is fitted with the relationship ln $t_{BD} = a(T) + b(T)$ ln J. (b)(c) The dependence of the coefficients a and b on T is linear.

data originally given in Fig. 2(a) are replotted, as shown in Fig. 4, a *linear* dependence of ln t_{BD} on T is apparent in the full range of the measured oxide thicknesses.

The observed T-dependence strongly contrasts with the Arrhenius law. If t_{BD} values are extrapolated from data measured at lower then the operating T, using this dependence may have a significant impact on the reliability prediction. On the other hand, if the t_{BD} data are measured directly at the

Figure 2. Arrhenius plot (a) for different oxide thicknesses and (b) for different stress currents for one oxide thickness. Data do not lie on straight lines and no single activation energy can be determined.
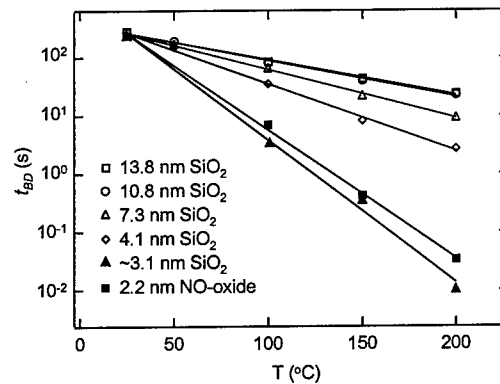
Figure 4. Data for each oxide thickness lie on a straight line, suggesting that oxide breakdown scales linearly with T. With decreasing oxide thickness the T-acceleration of t_{BD} increases.
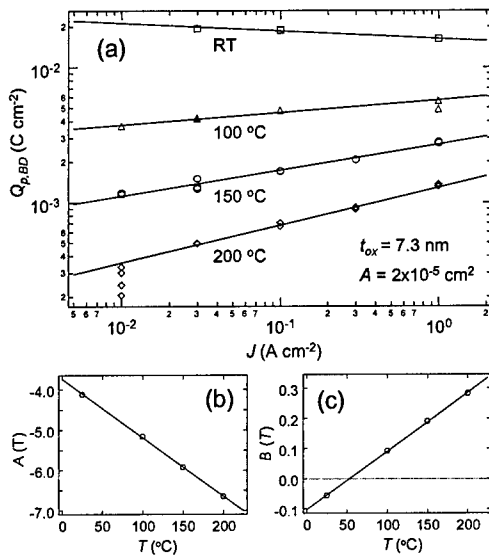
Figure 5. (a) Measured dependence of $Q_{p,BD}$ on T and J is fitted with the relationship $\ln t_{BD} = A(T) + B(T) \ln J$. (b)(c) The dependence of the coefficients A and B on T is linear. $Q_{p,BD}$ at higher T is not constant.

operating T or over a wide T range and *interpolated* to the operating T, the exact choice of the t_{BD} T-scaling law is not critical [2].

4. DISCUSSION

The possible mechanisms governing the observed T-dependence of t_{BD} were more closely studied on the 7.3 nm oxide. Some findings are now discussed.

4.1. Effect of the barrier lowering

In order to explain the observed log-linear decrease of t_{BD} with T, we have considered the effect of the Si/SiO$_2$ barrier height, which is seen to decrease *linearly* with T. This effect was found to be responsible for the log-linear dependence of FLASH memory retention times on T [4]. We observe a linear lowering of the barrier with T ($\sim$5 – 7 $\times$ 10^{-4} eV K^{-1}, assuming that both substrate-Si/SiO$_2$ and SiO$_2$/poly-gate barriers lower at the same rate).

However, the only direct effect of the decreased barrier with increased T is an increased current flowing through the oxide in the Fowler-Nordheim regime [4]. In the constant-current stress used on the 7.3 nm oxide, this results in a *decrease* of the electric field in the oxide, which by itself should

increase t_{BD}. We therefore conclude that this effect cannot be directly responsible for the observed T-dependence of t_{BD}. This is also apparent from the observations discussed below.

4.2. Hole fluence

The role of holes in the T-accelerated breakdown is discussed in Fig. 5. The hole fluence at breakdown $Q_{p,BD}$ at RT was previously observed to reach a critical value [5], which was an argument supporting the anode hole injection model. We find $Q_{p,BD}$, obtained by integration of the substrate current to t_{BD}, decreasing at higher T, and becoming a function of J [Fig. 5(a)]. The fitting curves can again be well linearized [Figs. 5(b) and (c)].

We note that an increase in $Q_{p,BD}$ for sub-RT was observed in Ref. [6]. This observation is in agreement with and complements our result. Considering that non-constant $Q_{p,BD}$ was also observed on wet oxides at RT [7], we suggest that the "constant" critical value of $Q_{p,BD}$ at RT might be merely accidental. Among plausible explanations we list the possibility that the substrate current is not a true measure of the number of holes passing through the oxide, and could be, for example, partially mediated by photons, as some experiments suggest [8].

4.3 Trap density at breakdown

Finally, we investigate the influence of T on the oxide trap density D_{ot} at breakdown ($D_{ot,BD}$). A constant critical value of D_{ot} has been previously linked with breakdown through the percolation model [1]. We extract this quantity by measuring
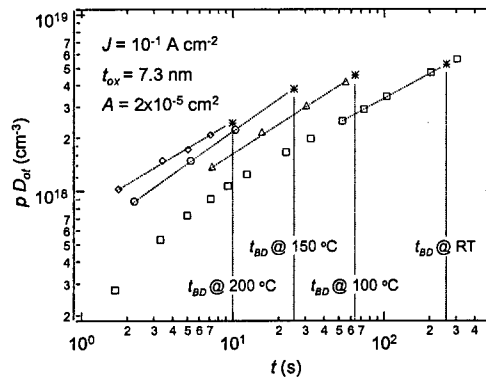
Figure 6. The dependence of pD_{ot} vs. injection time t at different T's is used in combination with the t_{BD} data (Fig. 3) to extract $D_{ot,BD}$. The filling fraction p is assumed to be constant for all measurements.
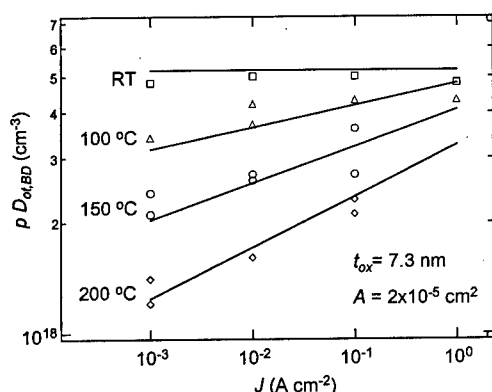
Figure 7. D_{ot} at breakdown decreases with increasing T and decreasing J. The lines assume Arrhenius-like increase in the trap generation rate.

D_{ot} (at RT [1]) as a function of the electron fluence Q_{inj} injected at different temperatures and extrapolating the dependence to breakdown (Fig. 6). From this figure it is also apparent that the trap generation rate increases with T.

We find that the $D_{ot,BD}$ extracted from our experiment and shown in Fig. 7 decreases with T. This effectively means that fewer traps are needed to cause breakdown at temperatures above RT. Somewhat surprising is also the J-dependence of $D_{ot,BD}$ observable at elevated T. Consequently, modifications to the very successful percolation model [1] appear to be necessary to incorporate the decrease in $D_{ot,BD}$ [9].

From Figs. 6 and 7 it is evident that the T-acceleration of oxide breakdown is affected by a combination of two factors—the trap generation rate and $D_{ot,BD}$—which both depend on T. Consequently, even though trap generation rate is apparently governed by the Arrhenius law [9], the T-dependence of $D_{ot,BD}$ is clearly non-Arrhenius, which results in the observed non-Arrhenius dependence of t_{BD} [10]. On the other hand, the observed log-linear T-dependence of t_{BD} seems to be obtainable only if $D_{ot,BD}$ assumes a certain special dependence on T so that it fully *compensates* the Arrhenius behavior of trap generation. We therefore cannot rule out the possibility that the observed log-linear T-dependence of t_{BD} is just coincidental and might disappear if, for example, data are obtained over a wider T range.

We also note the apparently similar dependencies of $Q_{p,BD}$ and $D_{ot,BD}$ on T and J (Figs. 5 and 7). In the

light of the above observations it is, however, difficult to establish the exact connection between holes and the neutral electron traps—conceivable dependencies between the two quantities range from causal to merely correlational.

5. SUMMARY

We observe that the Arrhenius law does not well describe the T-acceleration of oxide time-to-breakdown. We find that ln t_{BD} decreases *linearly* with T for a wide range of oxide thicknesses. Furthermore, we find that Q_p and D_{ot} do not reach a constant critical value at breakdown at elevated temperatures and the values of both these quantities at breakdown decrease with increasing T. We also show how elevated T affects the reliability of oxide films and predict that reliability will be insufficient for thicknesses below 2.5 – 2.8 nm.

ACKNOWLEDGEMENTS

REFERENCES

1. R. Degraeve *et al.*, IEEE T. Electron Dev. 45 (1998) 904.
2. B. Kaczer, R. Degraeve, N. Pangon, T. Nigam, and G. Groeseneken, accepted to ESSDERC'99.
3. R. Moazzami, J. C. Lee, and C. Hu, IEEE T. Electron Dev. 36 (1989) 2462.
4. B. De Salvo, G. Ghibaudo, G. Pananakakis, B. Guillaumot, and P. Candelier, ESSDERC Proc. (1998) 228.
5. I. C. Chen, S. Holland, K. K. Young, C. Chang, and C. Hu, Appl. Phys. Lett. 49 (1986) 669.
6. H. Satake and A. Toriumi, IEDM Tech. Digest (1993) 337.
7. K. Sakakibara, N. Ajika, and H. Miyoshi, IEEE T. Electron Dev. 45 (1998) 1336.
8. E. Cartier, J. C. Tsang, M. V. Fischetti, and D. A. Buchanan, Microelectron. Eng. 36 (1997) 103; M. Rasras *et al.*, submitted to IEDM'99.
9. N. Pangon, R. Degraeve, P. J. Roussel, G. Groeseneken, and H. E. Maes, presented at SISC'98.
10. D. J. DiMaria and J. H. Stathis, Appl. Phys. Lett. 74 (1999) 1752.

Reliability of Ultra Thin Oxide and Nitride Films in the 1 nm to 2 nm Range

B. Yuwono[a], T. Schloesser[a], A. Gschwandtner[a], G. Innertsberger[a], A. Grassl[a], A. Olbrich[a] and W.H. Krautschneider[b]

[a]Infineon Technologies,
Otto-Hahn-Ring 6, 81739 Munich, Germany
e-mail: budy_yuwono@hotmail.com, till.schloesser@infineon.com

[b]TU Hamburg-Harburg, Technische Elektronik & Bildverarbeitung,
Eissendorfer Str. 38, 21073 Hamburg, Germany
e-mail: krautschneider@tu-harburg.de

The scaled-down MOSFETs of the 16 Gbit generation and beyond will require a gate dielectric thickness less than 2 nm. Reliability of ultra thin dielectrics were investigated in DC and AC stress at room and elevated temperatures.

1. INTRODUCTION

The scaled-down MOSFETs of the 16 Gbit generation and beyond will require a gate dielectric thickness less than 2 nm [1]. In this paper, results will be presented about the reliability of ultra thin dielectric films (< 2nm) in terms of breakdown behavior, their homogeneity and reproducibility and their stability when stressed by DC and AC voltages at room and elevated temperatures. It will be shown that ultra thin oxide, nitrided oxide, oxynitride and nitride films may be suited for high Gigabit applications.

2. SAMPLE FABRICATION

For our experiments we used MOS capacitors processed on a n-type substrate doped with antimony (20 mOhm-cm).

A HF vapor preclean, the processing of the ultra thin dielectric films and the deposition of the polysilicon layer (thickness: 400 nm, phosphorus doped 5.10^{20} cm^{-3}) were performed in a hot cluster tool in order to avoid exposure to room air. This procedure precludes any growth of native oxide, so that the total thickness of the gate dielectric can be exactly controlled. Four different types of dielectric films were produced by rapid thermal processing (800°C to 1000°C, 5 sec. to 10 sec.): Thermal oxide, nitrided oxide, oxynitride and nitride. The dielectric film thicknesses were in the range from 1.5 nm to 2 nm. After deposition of an interlevel dielectric (thickness 1750 nm), the wafers were annealed for 10 sec. at 1072°C.

3. BREAKDOWN BEHAVIOUR

DC stressing caused a soft breakdown with a modest current increase (Fig. 1). Extrapolation of the breakdown time T_{BD} (V_{DC}) data shows for a voltage less than 2 V a

lifetime exceeding ten years (Fig. 2). This will provide a very good margin at the expected maximum operation voltage of 1.2 V for the 16 Gbit generation [1]. After soft breakdown, the current level has increased by a factor of about five to ten and exhibits a significant current noise (Fig. 1). Our samples did not show a major change of the general shape of the current voltage characteristics after stress and soft breakdown (Fig. 4). The breakdown charge Q_{BD} of the ultra thin dielectrics exponentially increases with decreasing stress voltage (Fig. 3).

Further stress experiments were carried out using an AC stress voltage (duty cycle 1:1, $V_{MAX} = \pm 4$ V). AC stress was not more detrimental than DC stress. The change of the current voltage behavior of the MOS capacitors after DC and after AC stress was similar (Fig. 5).

4. REPRODUCIBILITY

The current voltage characteristics of the MOS capacitors with ultra thin oxide films showed a significant spread across a wafer (Fig. 6). The gate area of the samples was 10 μm^2. TEM cross sections show some thickness variations of the oxide film that may be generated in the beginning of the oxidation [2]. But, these variations cannot sufficiently explain the data spread. For further analysis, Conducting Atomic Force Microscopy [3] was used to measure the surface topography and local current distribution (Fig. 7). Local areas of significantly greater current densities were found that may control the current across the dielectric films.

We observed that our nitrided oxide (Fig. 8a), oxynitride (Fig. 8b) and nitride (Fig. 8c) films showed a significantly smaller data spread corresponding with a better uniformity of the film on a microscopic scale. This may be caused by the deposition of the nitride with a low pressure in the hot cluster tool.

Stressing at T = 140°C increases only very slightly the impact of stress on the current voltage behavior (Fig. 9).

5. CONCLUSION

Because of the exponential increase of the breakdown charge Q_{BD} at reduced gate voltages, it can be expected that MOS transistors of the 16 Gbit generation with ultra thin dielectric gate films will meet the 10 years reliability specifications.

A problem may arise for the reproducibility of oxide because of relatively large variations of the tunneling current. Nitrided oxide, oxynitride and nitride films showed a much narrower distribution of the tunneling current.

DC and AC stress affected the current voltage characteristics in a similar manner. After soft breakdown, the current increased by about one decade. Stressing at 140°C accelerated and intesified the stress only very slightly.

In general, ultra thin dielectric films (thickness < 2 nm) may be usable for MOS transistors of the 16 Gbit generation and beyond.

REFERENCES

[1] Semiconductor Industry Association, Roadmap, 1997.
[2] H.S. Momose et al., IEEE Trans. Electr. Dev., vol. 45, p. 691, 1998.
[3] A. Olbrich et al., Proc. IRPS, p. 163, 1998.
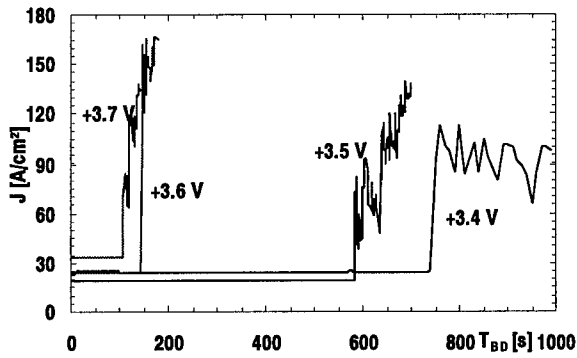
Fig. 1 Current density of MOS capacitors during DC stress until soft breakdown occurs. Dielectric: oxide, T_{ox}=1.5 nm.
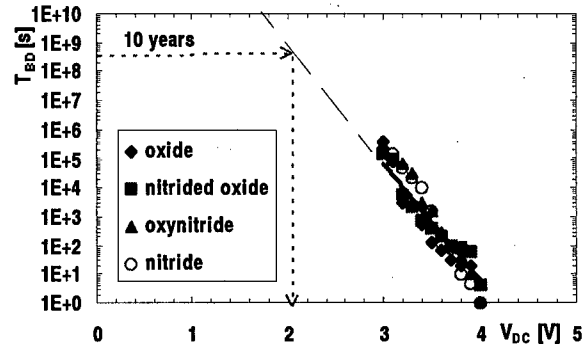
Fig. 2 Time to breakdown T_{BD} at different DC stress voltages for oxide and nitride (1.5 nm), nitrided oxide and oxynitride (2 nm).
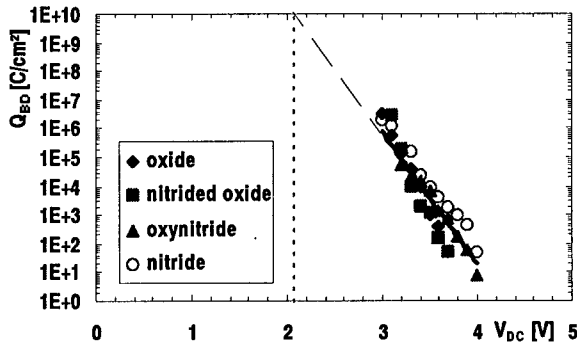
Fig. 3 Breakdown charge in dependence of the DC stress voltage. Dielectrics: s. Fig. 2.
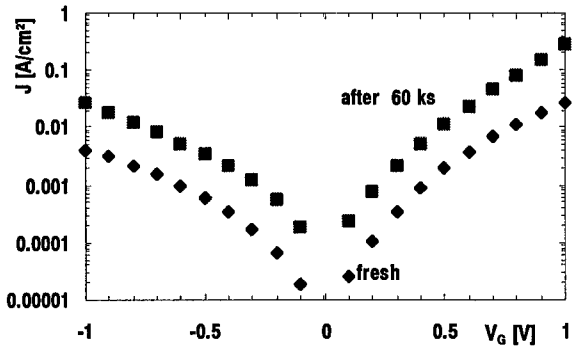
Fig. 4 I-V characteristics of MOS capacitor with oxide (T_{ox}=1.5 nm) before and after DC stress with V_{stress}=+3.5 V for 60 000 sec.
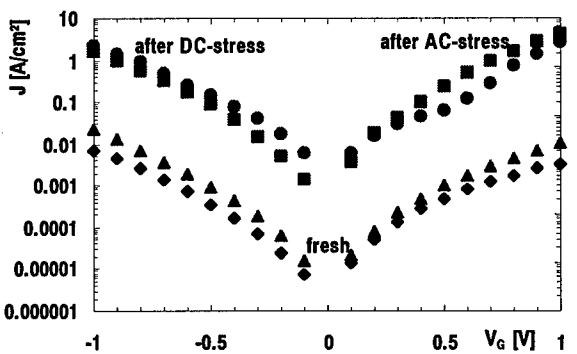
Fig. 5 Comparison of impact of DC and AC stress on the I-V characteristics of oxide (T_{ox}=1.5 nm).
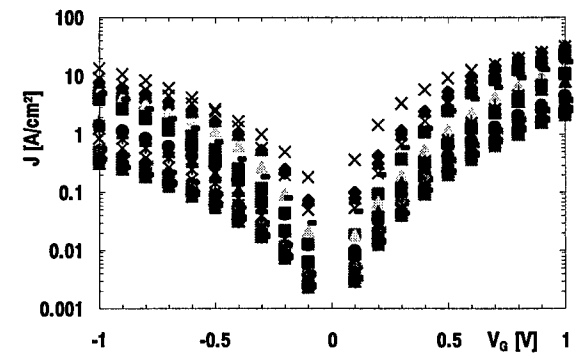
Fig. 6 Thin oxide films show a significant chip to chip variation of the I-V characteristics, T_{ox}=1.5 nm.
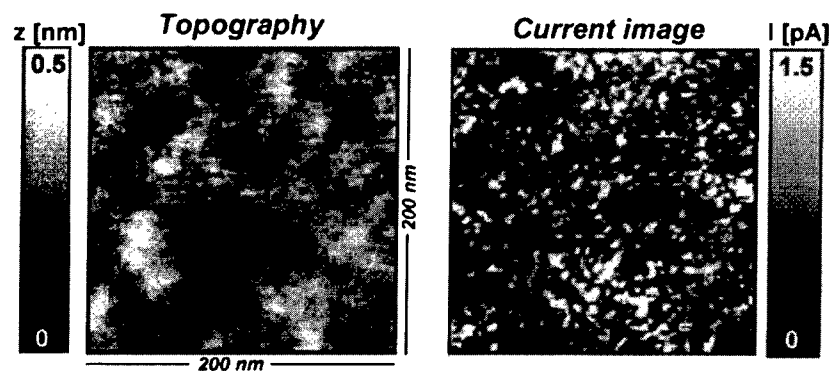
Fig. 7 Topography and local current distribution of oxide (T_{ox} = 1.8 nm)

nitrided oxide (NO): 1.8 nm
oxide: 800°C, 5s
nitride: 900°C, 10s
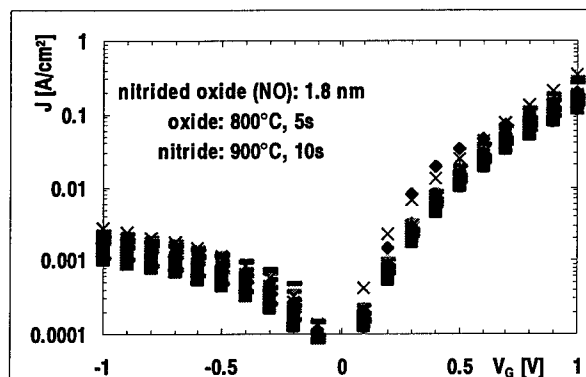
Fig. 8a Small data spread is observed for the I-V characteristics of thin NO films, T_{NO} = 1.8 nm.
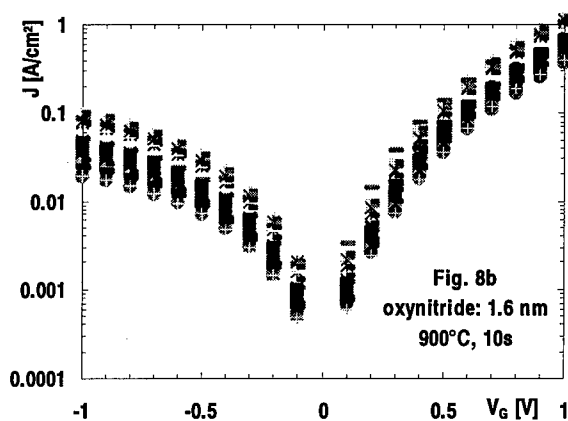
Fig. 8b
oxynitride: 1.6 nm
900°C, 10s
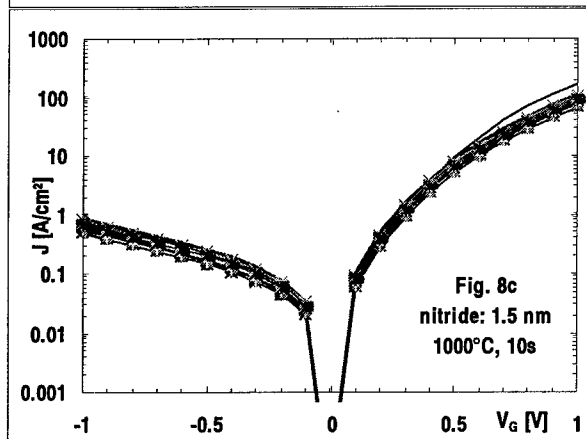
Fig. 8c
nitride: 1.5 nm
1000°C, 10s

140°C + AC

140°C

25°C
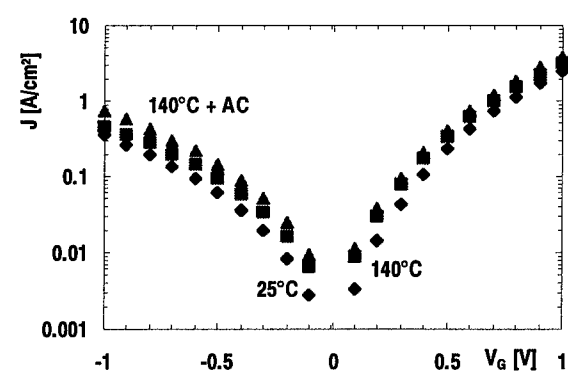
Fig. 9 Stressing at T=140°C (DC and AC) only slightly influences the I-V characteristics of oxide films (T_{ox}=1.5 nm).

Ultra Thin High Quality Stack Nitride/Oxide Gate Dielectrics Prepared by *in-situ* Rapid Thermal N$_2$O Oxidation of NH$_3$-nitrided Si

S. C. Song, H. F. Luan, C. H. Lee, A. Y. Mao, S. J. Lee, *J. Gelpey, *S. Marcus, and D. L. Kwong
Microelectronics Research Center, Department of Electrical and Computer Engineering
The University of Texas, Austin, TX 78758, (512) 471-5992, Fax: (512) 471-4345, dlkwong@mail.utexas.edu
*Steag AST Electronik, USA. Tempe, AZ 85284

Abstract

In this paper, we report ultra thin high quality nitride/oxide gate dielectrics prepared by rapid thermal NH$_3$ nitridation of Si followed by *in-situ* N$_2$O oxidation (NH$_3$+N$_2$O process). These films show excellent interface properties, significant lower leakage current (~10^2X), enhanced reliability, and superior boron diffusion barrier properties compared with SiO$_2$ of identical equivalent oxide thickness (Tox,eq).

Introduction

Formation of oxynitride gate dielectrics using nitrogen-containing gas (i.e., N$_2$O, NO) has been widely studied, and improved reliability and superior resistance to boron diffusion have been demonstrated [1]. However, as we continue to scale down oxynitride thickness, the maximum nitrogen concentration incorporated through above processes becomes less for a given thermal budget process. In addition, N located at the SiO$_2$/Si interface in these processes increases fixed charge and interface trap densities, and reduces peak channel mobility [2]. In this paper, we report a novel method to enhance nitrogen incorporation with N located away from SiO$_2$/Si interface. This not only blocks boron diffusion effectively, but also increases dielectric constant of the resulting film to achieve thinner equivalent oxide thickness without degrading SiO$_2$/Si interface properties

Experiment

Both n+ poly NMOS and p+ poly PMOS devices are fabricated. NH$_3$-nitridation of Si substrate is performed in RTP at 700~800°C for 10 seconds to grow ultra thin Si$_3$N$_4$ layer, followed by *in-situ* N$_2$O oxidation at 800~950°C for 15~30 seconds. N$_2$O oxidation temperature plays an important role in controlling final film thickness as shown in Fig. 1. Control oxide films were also grown in pure O$_2$ in the same RTP system. POCl$_3$ doping is used for n+-poly gate NMOS and BF$_2$ ions are implanted at 50 keV to a dose of 5x10^{15} cm^{-2} for p+-poly PMOS. The amount of boron penetration is varied by changing the RTA drive-
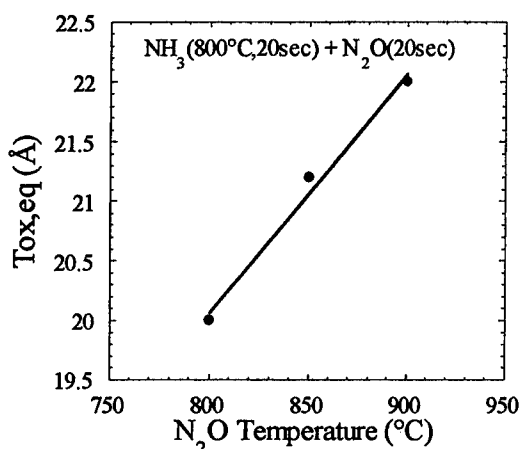
Fig.1 Effects of N$_2$O oxidation temperature on final NH$_3$+N$_2$O film thickness. N$_2$O temperature is important for final thickness

in conditions for p+-poly PMOS capacitors. Equivalent oxide thickness (Tox,eq) is determined by high frequency CV curve in strong accumulation region without considering quantum mechanical effects.

Results and Discussion

A. Interface properties

Fig. 2 shows high frequency and quasi-static C-V
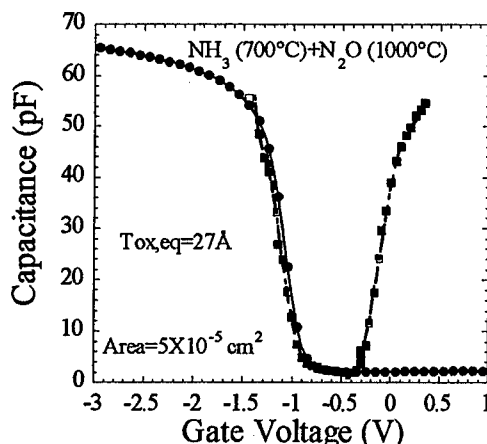
Fig.2 High and quasi-static C-V curves of n+-poly NMOS capacitor with NH$_3$+N$_2$O film. Identical C$_{min}$ value indicate extremely low interface state density
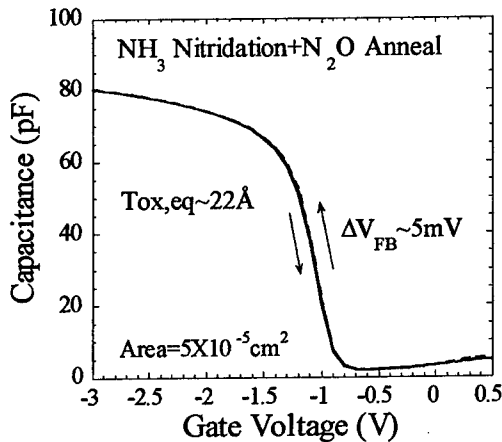
Fig. 3 Hysteresis characteristics of NH_3+N_2O film. Stronger oxidizing capability of N_2O anneal eliminates bulk traps effectively, resulting in negligible amount of hysteresis.
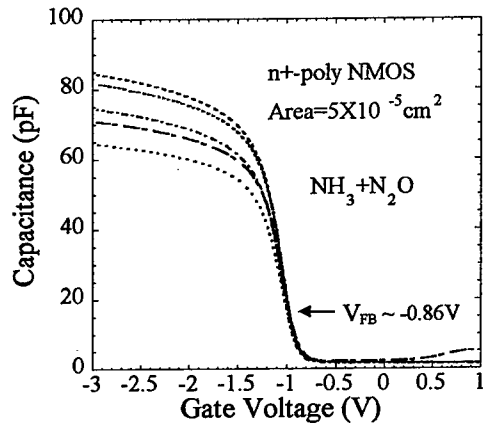
Fig. 4 High frequency C-V curves of NH_3+N_2O films of different thickness. Due to effectively removed H and interfacial oxide grown during N_2O oxidation, NH_3+N_2O films show extremely low positive fixed charge, which results in identical V_{FB} (~0.86V) with different C_{ox}.

Fig.5 High frequency C-V curves of NH_3+N_2O and SiO_2 of identical equivalent oxide thickness. V_{FB} difference is as small as 50mV.
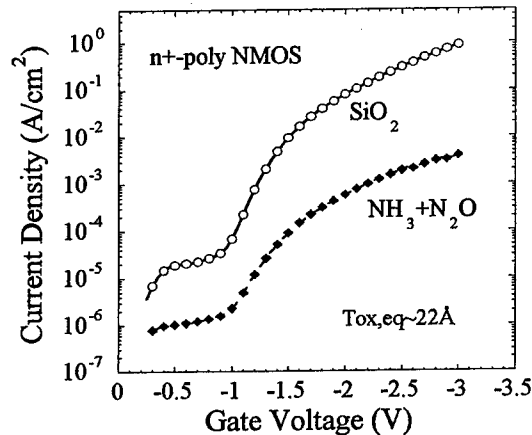
Fig.6 Jg-Vg curves of NH_3+N_2O and SiO_2 of identical equivalent oxide thickness. NH_3+N_2O exhibits significantly low leakage current probably due to thicker physical thickness

curves of NH_3+N_2O film. Contrary to conventional Si_3N_4 film which contains high interface states, smooth transition in depletion region and identical C_{min} value of NH_3+N_2O film indicate extremely low interface state density. Angle-resolved XPS data of NH_3+N_2O films indicated an ultra thin (<10Å) pure SiO_2 layer was formed between Si_3N_4 and Si after N_2O reoxidation [3], suggesting N is removed away from SiO_2/Si interface during N_2O reoxidation. This result is very significant in that this is the first time that a nitride/oxide stack layer is realized by low thermal budget thermal growth process.

Negligible amount of hysteresis (ΔV_{FB} ~ few mV) is observed in NH_3+N_2O film as shown in Fig. 3 due to eliminated bulk traps by N_2O oxidation. Since N_2O has stronger oxidizing capability compared to O_2 [4], N_2O

anneal is more effective to remove defects in NH3 grown Si_3N_4 film. Fig. 4 shows high frequency C-V curves of NH_3+N_2O films of different thickness. Identical V_{FB} values (V_{FB}~0.86V) with different C_{ox} imply negligible amount of fixed charge in these films. Since positive fixed charges in conventional Si_3N_4 film is due to N-H bonds at the interface, N_2O oxidation of NH_3 grown Si_3N_4, which removes H effectively and grows ultra thin interfacial oxide layer, allows negligible amount of positive fixed charge in NH_3+N_2O film.

B. Low leakage current

High frequency C-V and Jg-Vg curves of both SiO_2 and NH_3+N_2O films of identical Tox,eq are shown in Figs. 5, 6. Significantly lower leakage currents are
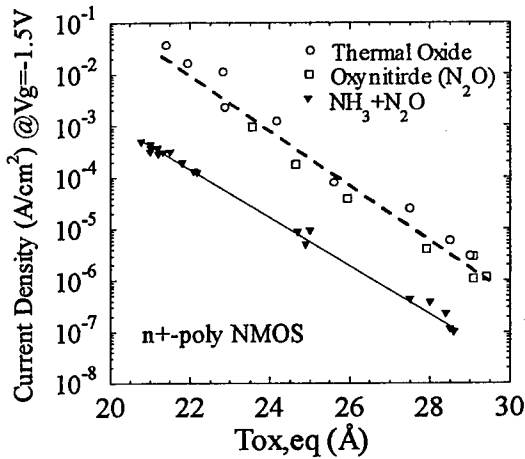
Fig.7 Current density @Vg=-1.5V of SiO$_2$, N$_2$O oxynitride and NH$_3$+N$_2$O devices with respect to Tox,eq. NH$_3$+N$_2$O films show more than 15X lower leakage current compared to SiO$_2$ or N$_2$O oxynitride

Fig.8 Comparison of boron diffusion barrier properties. Suppression of boron penetration is observed in NH$_3$+N$_2$O films even up harsh drive in condition of 1050°C,30sec.

obtained in entire gate bias range from NH$_3$+N$_2$O films, while the flat band voltage difference between NH$_3$+N$_2$O and control SiO$_2$ is as small as 50mV. Effective nitrogen incorporation from NH$_3$-grown Si$_3$N$_4$ layer increases dielectric constant, which allows physically thicker film with identical Tox,eq compared to SiO$_2$ or N$_2$O oxynitride. In Fig. 7, leakage current density (@Vg=-1.5V) is plotted as a function of equivalent oxide thickness for SiO$_2$, N$_2$O oxynitride, and NH$_3$+N$_2$O. More than 10^{2}X lower leakage current is seen for NH$_3$+N$_2$O films compared to SiO$_2$ or N$_2$O oxynitride of identical equivalent oxide thickness.

C.   Suppression of boron penetration.

Fig. 8 shows boron penetration characteristics of NH$_3$+N$_2$O and SiO$_2$. NH$_3$+N$_2$O film shows virtually no V$_{FB}$ shift even up to 1050°C, 30 seconds of implantation anneal, while significant V$_{FB}$ shift is observed in SiO$_2$ film due to born penetration. Since boron penetration becomes harder as N concentration of gate dielectric increases [5], complete suppression of boron penetration in NH$_3$+N$_2$O film is attributed to large amount of N incorporation from thermally grown nitride.

Large amount of boron penetration enhanced by fluorine generates traps in the bulk of gate oxide [6], which increases leakage current at low field regime as well as degrades reliability. Boron penetration-induced low field leakage current is also suppressed effectively in NH$_3$+N$_2$O films, as shown in Fig. 9, implying that boron penetration is blocked from the top surface.

D.   SILC and TDDB

In Fig. 10, stress induced leakage current (SILC)
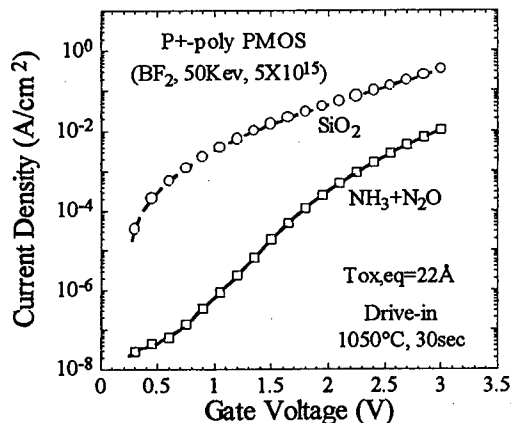
Fig.9 Boron penetration induced low field leakage current is also suppressed effectively in NH$_3$+N$_2$O films due to less trap generation during boron penetration

characteristics of NH$_3$+N$_2$O and control oxide are shown. Gate injection stress of Vg=-3.5V was applied for 1000sec and leakage current increment is monitored at Vg=-1.5V. As shown in the figure, stress induced leakage current after stress is found to be much higher for SiO$_2$ compared to NH$_3$+N$_2$O device, and SiO$_2$ films tend to show soft breakdown characterisitcs during stress, which is also known as B-mode SILC [7].

NH$_3$+N$_2$O has higher breakdown field with tighter distribution as shown in Fig. 11. A 10X longer time-to-breakdown is seen in n+ poly NMOS capacitor with NH$_3$+N$_2$O at a given electric field compared to SiO$_2$ (Fig. 13). Recently, gate oxide failure is reported to be limiting factor for scaling of oxide thickness, since time-to-

S. Song et al. / Microelectronic Engineering 48 (1999) 55–58
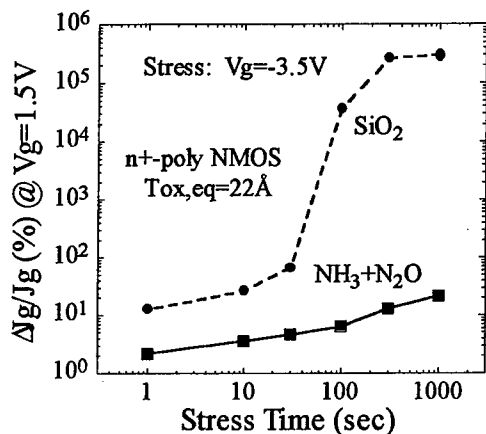
Fig.10 Stress induced leakage current as a function of stress time for NH_3+N_2O and SiO_2. SiO_2 films tend to show soft breakdown characteristics.

breakdown decreases exponentially with increasing gate leakage current [8]. Due to significantly lower leakage current, NH_3+N_2O film can extend scaling limit of SiO_2 in terms of dielectric reliability as well as stand-by power consumption. Projected electric field for 10 years lifetime of NH_3+N_2O films is -12.5MV/cm. For substrate injection condition (positive gate bias) on p+ poly PMOS capacitor with NH_3+N_2O films, even higher electric field (13.2MV/cm) at 10 years lifetime is obtained in devices with severe drive-in condition (1050°C, 30seconds), as shown in Fig. 12.

Conclusion

In this paper, we report ultra thin high quality nitride/oxide gate dielectrics prepared by rapid thermal NH_3 nitridation of Si followed by in-situ N_2O oxidation (NH_3+N_2O process). These films show excellent interface
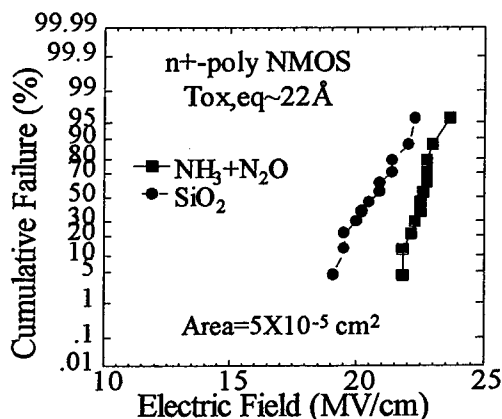
Fig.11 Breakdown field distribution of NH_3+N_2O and SiO_2. NH_3+N_2O has higher and tighter breakdown field distribution.
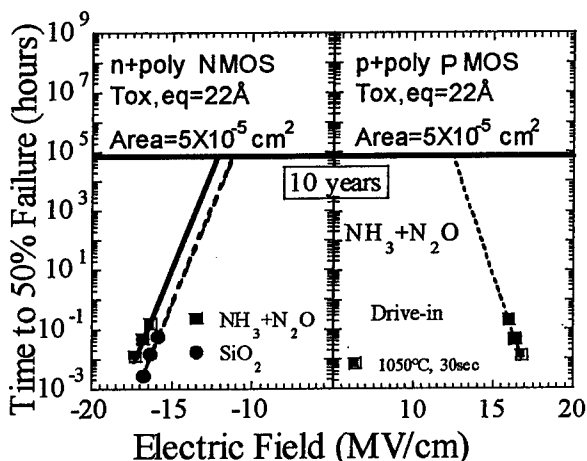
Fig.12 10 years lifetime projection of n+-poly NMOS capacitors with NH_3+N_2O and SiO_2 and NH_3+N_2O p+-poly PMOS for two drive conditions. 10X longer T_{BD} is seen from NH_3+N_2O than SiO_2.

properties, significant lower leakage current ($\sim10^2X$), enhanced reliability, and superior boron diffusion barrier properties compared with SiO_2 of identical thickness.

References

[1] H. Hwang, et. al., "Electrical and reliability characteristics of ultrathin oxynitride gate dielectric prepared by rapid thermal processing in N_2O", IEDM Tech. Dig., p. 421, 1990

[2] T. Hori., "Inversion layer mobility under high normal field in nitrided-oxide MOSFETs", IEEE Trans. Electron. Devices, vol. 37, no. 9, p.2058, 1990

[3] A. Y. Mao, et. al., "N_2O oxidation kinetics of ultra thin thermally grown silicon nitride", MRS Proceeding, spring, 1999. to be published.

[4] W. S. Lau, et. al., "Evidence that N_2O is a stronger oxidizing agent than O_2 for post deposition annealing of Ta_2O_5 on Si capacitors", Jpn. J. Appl. Phys. Vol. 36, pp. 661–666, 1997

[5] R. B. Fair, "Physical model of boron diffusion in ultra thin gate oxides", J. Electrochem. Soc, vol. 144, no. 2, p. 708, 1997

[6] J. R. Pfiester, et. al., "The effects of boron penetration on p+ polysilicon gated PMOS devices", IEEE Trans. Electron Devices, vol. 37, p. 1842, 1990

[7] K. Okada, et. al., "A new prediction method for oxide lifetime and its application to study dielectric breakdown mechanism", Symp. on VLSI Tech., p.158, 1998

[8] J. H. Stathis, et. al., "Reliability projection for ultra thin oxides at low voltage", IEDM Tech. Dig., p. 167, 1998

Reliability of ultra-thin gate oxides grown in low-pressure N_2O ambient or on nitrogen-implanted silicon

A.J. Bauer[a], M. Beichele[a], M. Herden[b], H. Ryssel[a,b]

[a]Fraunhofer-Institut für Integrierte Schaltungen - Bauelementetechnologie (IIS-B),
Schottkystr. 10, D-91058 Erlangen, Germany

[b]Lehrstuhl für Elektronische Bauelemente, Universität Erlangen-Nürnberg,
Cauerstr. 6, D-91058 Erlangen, Germany

The influence of nitrogen incorporated at the gate oxide interface on the reliability and electrical behavior of MOS capacitors was investigated. Incorporation of nitrogen was accomplished either by either rapid thermal oxidation (RTO) in N_2O or N_2O/O_2 atmosphere at pressures between 75 to 500 mbar or by implantation of nitrogen at 20 keV and doses between 10^{13} to 10^{15} cm^{-2} into the substrate prior to RTO. Characterization of MOS capacitors with oxide thicknesses between 2.9 to 3.9 nm was achieved by reliability measurements at constant current or constant voltage stress, respectively. These measurements were correlated to the nitrogen dose at the gate interface which was determined by SIMS analysis. Stress reliability showed a sensitive dependence on the processing conditions. An accurate adjustment of the nitrogen dose at the gate interface is mandatory to attain improved reliability of nitrided oxides.

1. INTRODUCTION

Nitrogen at the gate oxide interface is supposed to harden the gate oxide in order to increase hot carrier immunity and to suppress diffusion of boron from a p-type gate electrode through the gate oxide [1, 2]. This becomes of great importance for gate oxide thicknesses lower than 4 nm which are used in sub 0.25 μm device generations and are investigated in this work. Rapid thermal oxidation of silicon under low pressure in an environment containing N_2O is known as an alternative and promising method to grow high-quality ultra-thin gate oxides [3]. Using low pressure, thickness homogeneities can be improved and due to longer process times, better process control can be achieved. Moreover, nitrogen implantation into the substrate prior to gate oxidation offers the possibility to obtain different oxide thicknesses of high quality within the same manufacturing process e.g. for fabrication of embedded memories [4].

2. EXPERIMENTAL

Gate oxidation in N_2O containing ambient at low pressure was carried out in a rapid thermal processing system using O_2, N_2O or 1:1 N_2O/O_2 gas at pressures of 500, 150 and 75 mbar on p-type silicon (100) with a resistivity of 4-6 Ωcm. The oxidation temperature was 1050 °C and the oxidation times were adapted to the individual pressures to yield an oxide thickness of 3.6 nm.

For oxides grown on nitrogen implanted silicon, nitrogen was implanted into the silicon substrate of PMOS devices through a 24 nm thick screen oxide at an energy of 20 keV and at doses from 10^{13} cm^{-2} to 10^{15} cm^{-2}. After removing the screen oxide, gate oxides were grown at temperatures of 1050 or 1000 °C, respectively, at a pressure of 500 mbar in dry oxygen ambient by RTO. Gate electrodes were formed by deposition of a 500 nm thick polysilicon layer at 620 °C and additional $POCl_3$ diffusion at 950 °C for n-type electrodes. P-type electrodes were doped by implantation of BF_2 (50 keV, $2 \cdot 10^{15}$ cm^{-2})

and annealing (1070 °C, 5 sec). The MOS capacitors were formed using lithography and etching.

Gate oxide reliability was determined by measuring time-dependent dielectric breakdown (TDDB) using constant current stress (CCS) or constant voltage stress (CVS), respectively. CCS measurements were performed on capacitors with an area of 0.1 mm² at a current density of 100 mA/cm². CVS measurements were conducted on capacitors with an area of 0.02 mm^2 at field strengths from 12 up to 17 MV/cm. During the CVS measurements the occurrence of soft-breakdowns was checked by regularly switching to a lower test voltage (2 V).

The nitrogen concentration at the gate interface was investigated by SIMS analysis using CsN^+ ions to minimize matrix effects in secondary ion yield. The oxide thicknesses were measured by ellipsometry using a refractive index of 1.46.

3. RESULTS

For gate injection, the Q_{bd}-distributions determined by CCS measurements show almost parallel spreads in their Q_{bd}-distributions. As a strong polarity gap occurs, for substrate injection Q_{bd}-values higher than 100 C/cm^2 and unacceptably long measurement cycles occured.
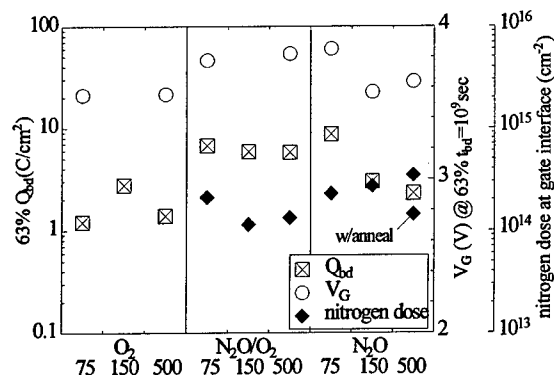
Figure 1. Oxides grown in nitrogen containing ambient. 63 % Q_{bd}-values measured by CCS, gate voltages for 30 years lifetime extracted from CVS measurements and integrated nitrogen doses at the gate interfaces.

The 63 % values of Q_{bd} for the oxides grown in atmospheres containing N_2O are depicted in figure 1. The reference oxides grown in O_2 show the lowest 63 % Q_{bd}-values. A slight improvement in Q_{bd} is found when the O_2-oxides were grown at 150 mbar. This value is comparable to N_2O-oxides grown at 150 and 500 mbar. Regardless of pressure the largest Q_{bd}-values were found for all N_2O/O_2-oxides and for the 75 mbar N_2O-oxides. The latter show the best results with respect to Q_{bd}.

Like CCS, the CVS measurements yield linear Weibull distributions for the t_{bd}-values (Fig. 2). Soft-breakdown occurred a decent time before hard-breakdown. Compared to the O_2-sample and the 500 mbar N_2O-sample the Weibull-plots of the N_2O/O_2 and the N_2O-oxides are significantly shifted to higher t_{bd}-values.
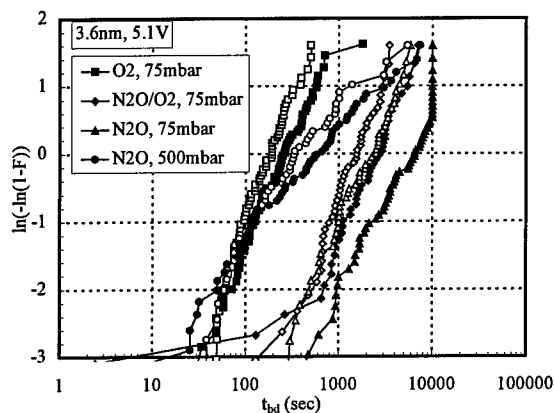
Figure 2. t_{bd}-distributions determined by CVS measurements for gate oxides grown in O_2, N_2O or N_2O/O_2 atmosphere at low pressure. (Open symbols for soft-breakdown, filled symbols for hard-breakdown).

Figure 3 illustrates the 63 % values of t_{bd} for hard-breakdown that were found by CVS measurements at different gate voltages. In order to compare the individual samples, the data set of each sample was extrapolated to a lifetime of 30 years. A unique extrapolation slope was used for all samples to account for statistical variations. The corresponding voltages for 30 years lifetime were used to compare the oxides' reliabilities. These voltage values are also depicted in figure 1. The measured data sets can be divided into two groups. The first group, which shows significantly shorter operating voltages

contains the O_2-samples along with the N_2O-samples that were grown at 150 and 500 mbar. The N_2O/O_2-samples together with the N_2O-sample grown at 75 mbar are shifted towards higher operating voltages.
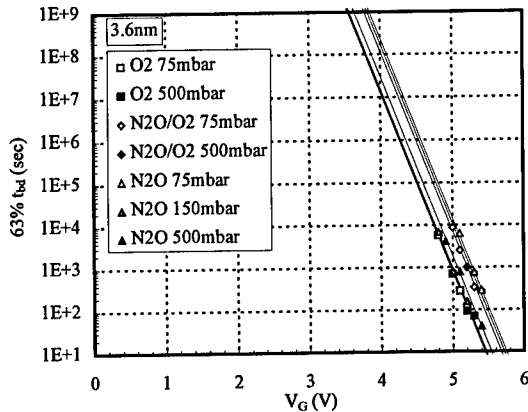
Figure 3. 63 % t_{bd}-values for hard-breakdown from CVS measurements for different gate voltages of oxides grown in O_2, N_2O or N_2O/O_2 atmosphere at 1050 °C by RTO at low pressure.

When oxides are grown in an environment containing N_2O nitrogen piles up at the SiO_2/Si-interface. The SIMS measurements of the nitrogen concentrations were carried out on 4.0 nm oxides. Figure 1 shows that the nitrogen content is strongly influenced by oxidation pressure. For the oxides grown in pure N_2O the nitrogen dose increases with pressure, even though at higher pressures shorter oxidation times are required for the same thickness. When oxides are grown in an N_2O/O_2 ambient the nitrogen dose is hardly affected in the pressure range from 500 down to 150 mbar whereas a further decrease of pressure down to 75 mbar leads to a rise of the nitrogen dose by almost a factor of 2. In general, the nitrogen dose incorporated into the N_2O/O_2-oxides is significantly smaller than that of the N_2O-oxides. However, the concentrations of the N_2O/O_2 and the N_2O-oxide grown at 75 mbar nearly coincide. An in-situ N_2 anneal immediately after oxidation in an N_2O ambient at 500 mbar leads to a significant drop of the nitrogen content to almost $1/3^{rd}$ of the initial dose.

The CCS measurements of MOS capacitors on nitrogen implanted silicon also yield a linear

dependence of the Q_{bd}-distributions in a Weibull-plot. The 63 % values of Q_{bd} are illustrated in figure 4 with respect to different processing conditions. No unique correlation between Q_{bd} and the processing parameters is observable. The largest Q_{bd}-values are obtained from devices with a nitrogen implantation dose of 10^{13} and 10^{15} cm^{-2}, whereas for a dose of $5 \cdot 10^{13}$ cm^{-2} a significant decrease in the 63 % value of Q_{bd} occurs. Thicker oxides show higher Q_{bd}-values for the same nitrogen implantation dose. An anneal after nitrogen implantation and before gate oxidation did not influence the breakdown behavior.
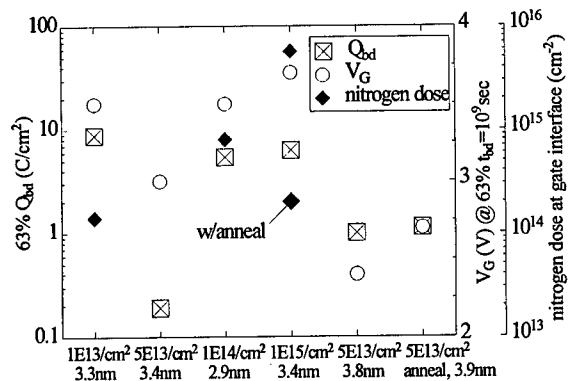
Figure 4. Oxidation of nitrogen implanted devices. 63 % Q_{bd}-values measured at CCS, gate voltages for 30 years lifetime extracted from CVS measurements and integrated nitrogen doses at the gate interface.

To confirm the CCS measurements, CVS measurements on nitrogen implanted oxides are conducted. In figure 5, the 63 % values of t_{bd} are shown for different applied stress voltages. The values of each sample were extrapolated to a corresponding voltage for 30 years lifetime. The relation between the extrapolated voltages and the processing conditions is illustrated in figure 4. It was found that capacitors with an implantation dose of 10^{15} cm^{-2} reach the highest lifetime values. Whereas, as was found by CCS measurements, capacitors with an implantation dose of $5 \cdot 10^{13}$ cm^{-2} show the lowest lifetime. Here, insufficient incorporation of the nitrogen at the gate interface and thus formation of traps occurs. The thicker oxides show a smaller slope of the extrapolation line and thus the extrapolated operating voltages are

lower. Annealing before the gate oxidation slightly increases the operating voltage for 30 years lifetime.
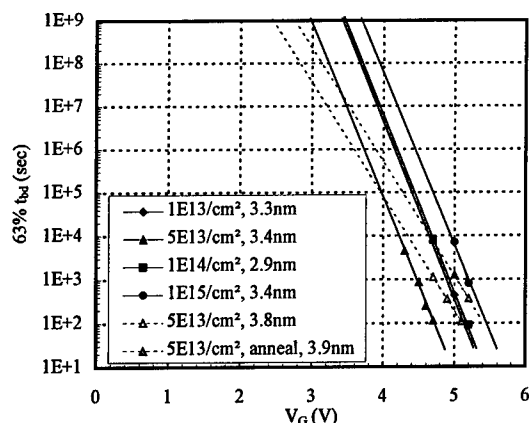
Figure 5. 63 % t_{bd}-values of MOS capacitors on nitrogen implanted silicon for different gate voltages. The values were determined by CVS.

From SIMS measurements, a pile-up of nitrogen at the SiO_2/Si-interface is observed. The nitrogen concentration increases with the implantation dose. Performing an annealing step after the implantation lowers the nitrogen concentration significantly. Due to segregation mechanisms, the detected nitrogen dose at the gate interface is larger than the implanted one, as is illustrated in figure 4. For a nitrogen implantation dose of 10^{15} cm^{-2} about one monomolecular nitrided layer is formed in which also silicon nitride precipitates were observed previously [5]. At lower doses, a stoichiometric silicon nitride layer cannot be formed so that the incorporated nitrogen can generate traps which decrease the oxide reliability. For a implantation dose of 10^{13} cm^{-2} and for the annealed sample with an implantation dose of 10^{15} cm^{-2} the nitrogen concentration at the SiO_2/Si-interface is comparable to that of the N_2O oxides. In case of annealing before the gate oxidation a portion of the nitrogen is stripped off together with the screen oxide and the remaining nitrogen is bound on lattice sites where also crystal tensions can be formed [5]. Therefore, the process parameters have to be adjusted properly to minimize these effects and to maximize the benefits of nitrogen for gate oxide reliability.

4. CONCLUSION

Nitrogen implantation before gate oxidation or N_2O-oxidation are both suitable to reach high oxide reliability if the amount or incorporated nitrogen is adjusted properly. For an implantation dose of 10^{15} cm^{-2} even a slightly higher reliability as for oxides grown in dry oxygen ambient can be achieved at which the oxidation time is ten times longer for the implanted oxides. Pressure during N_2O-oxidation was also a suitable parameter to improve the gate oxide reliability, where the highest reliability values were obtained form N_2O/O_2-oxidation or N_2O-oxidation at 75 mbar. The experiments show that the amount of nitrogen near the gate interface is not the only parameter that determines the oxide quality. CCS and CVS measurements show comparable results when the thicknesses of the compared oxides are equal and higher than 3.5 nm. For thicknesses below 3.5 nm only the CVS measurement yields a reliable statement. Both methods are eligible to improve gate oxide reliability and to achieve better process control with respect to RTP.

REFERENCES

1. D.M. Fleetwood, N.S. Saks, J. Appl. Phys. 79(3), 1996, p. 1583

2. G.D. Leonarduzzi, D.-L. Kwong, Semiconductor International, July 1998, p. 225 or C.T. Liu, Y. mA, H. Luftman, S.J. Hillenius, IEEE Electron Device Letters, Vol. 18, No. 5, 1997, p. 212

3. R. McIntosh, C. Galewski, J. Grant, Mat. Res. Soc. Symp. Proc., Vol. 342, 1994, p. 209

4. C.T. Liu, Y. Ma, et. all., International Electron Device Meeting, San Fransisco, CA, Technical Digest, 1998, p. 589

5. A.J. Bauer, P. Mayer, L. Frey, V. Häublein, H. Ryssel, Ion Implantation Technology, Kyoto, Japan, Technical Digest, 1998

PHOTOELECTRON YIELD SPECTROSCOPY OF ELECTRONIC STATES AT ULTRATHIN SiO₂/Si INTERFACES

S. Miyazaki, T. Maruyama, A. Kohno and M. Hirose
Department of Electrical Engineering, Hiroshima University
Kagamiyama 1-4-1, Higasi-Hiroshima 739-8527, Japan

Total photoelectron yield spectroscopy (PYS) has been applied to evaluate the energy distribution of electronic defect states for ultrathin SiO_2($\leq$4.5nm)/Si(100) system. The novelty and usefulness of this technique have been demonstrated through the observation of the reduction in defect states with progressive dilute-HF treatment for such ultrathin SiO_2/Si(100) and the comparison of the PYS and the quasi-static C-V measurements for 4.5nm-thick SiO_2/Si(100).

1. INTRODUCTION

Controlling electrically-active defect states in SiO_2 and near the SiO_2/Si interface is a matter of great importance to achieve a good and highly-reliable performance for metal-oxide-semiconductor (MOS) devices. Conventionally, the spatial and energy distribution of electronic trapping centers of SiO_2-Si system has been evaluated from the analysis of capacitance-voltage (C-V) characteristics for MOS structures [1]. However, for gate oxides thinner than 3nm which will be required for MOS memories in gigabit generations, difficulties in accurately observing the C-V characteristics emerge because of high leakage current due to the direct tunneling. To overcome the difficulties, a modified C-V method with a non-contact electrode [2] or an electronic circuit for tunnel-current compensation [3] has been developed, and also an alternative technique using x-ray photoelectron spectroscopy (XPS) [4-6] in which surface band bending due to electron trapping is measured from the binding energy shit for core level photoemission has been demonstrated. Even such techniques are only applicable to the evaluation of defect states more than $10^{11}cm^{-2}eV^{-1}$ located within $\pm$0.3eV from midgap for ultrathin ($\leq$3nm) SiO_2/Si. Recently, we have reported a novel approach to the study of the energy distribution of surface states for chemically-cleaned Si wafers by using total photoelectron yield spectroscopy (PYS) and demonstrated that the PYS technique enables us to evaluate the surface state density as low as $10^9cm^{-2}eV^{-1}$[7].

In this work, for ultrathin SiO_2 ($\leq$4.5nm)/Si interfaces, we show the results on application of the PYS technique for determining the energy distribution of the defect state density and report the influence of oxide thinning in a dilute HF solution on the defect states.

2. EXPERIMENTAL

Clean oxides with thicknesses of 1.8 - 4.5nm were grown on Si(100) wafers with a boron concentration of 0.5-1.0x10¹⁵cm⁻³ or a phosphorous concentration of 0.3-1.0x10¹⁹cm⁻³ in 2% O_2 diluted with N_2 at 900-1000°C or in pure O_2 at 700-800°C. The total photoelectron yields for the SiO_2/Si(100) samples so prepared were measured as a function of incident photon energy in the range from 4 to 6eV with an energy resolution of ~20meV to evaluate electronic defect states in the Si bandgap near the interface. Some of the SiO_2 layers grown at 1000°C were etched back stepwise in a 0.1%HF solution and followed by pure water rising. At each SiO_2-thinning step, very homogeneous SiO_2 etching was confirmed by AFM measurements and then total photoelectron yield measurements were carried out to examine the change in defect states by the dilute HF treatment. The details of the experimental technique have been given elsewhere[7, 8].

3. RESULTS AND DISCUSSIONS

Typical PYS spectra for as-grown SiO_2/n⁺Si(100) are compared with that for H-terminated n⁺Si(100) in Fig. 1. For SiO_2/n⁺Si (100), the valence electrons of the SiO_2 layer can not contribute to the yield in this energy region because of the SiO_2 energy bandgap of ~8.9eV[9] and a yield reduction due to the scattering of photoelectrons escaping through the SiO_2 layer should be taken into account. Under flat band condition, since the photoemission from the Si valence band starts at an incident photon energy of 5.15eV, which corresponds to the energy separation between the valence band edge and the vacuum level, observed yields for photons below ~5eV are attributable to

photoemission from filled gap states at the interface and the surface. Notice that no significant difference in the spectral shape between the yields from 2.5nm-thick $SiO_2/n^+Si(100)$ and 4.5nm-thick $SiO_2/n^+Si(100)$ is observed except the absolute intensity, which indicates that the contribution from the defect states in SiO_2 at least 2.5nm away from the interface is negligible. In order to differentiate between photoemissions from the Si valence band and gap states, the photoemission yield from the valence band were characterized by power law dependence which is theoretically predicted [10] and empirically well-accepted[11, 12]. As shown in Fig.2, the threshold energies for direct and indirect excitations from the Si valence band with nearly parabolic density of states to the vacuum level were determined from the linear plot and the cubic root plot for the yield spectrum, respectively. The direct and indirect threshold energies for the SiO_2/n^+Si sample shift toward the higher photon energy side by 50meV from those for H-terminated Si. This could be interpreted in terms of the negative charge-up of the SiO_2 layer and/or a number of interface states which result in the upward surface band bending. Since the indirect excitation threshold energy corresponds to the Si valence band maximum E_V

measured from the vacuum level, the photoemission yield below the indirect excitation threshold reflects the gap state distribution at H-terminated n^+Si surface or the SiO_2/Si interface in an integrated form. Although the origin of the observed surface states for H-terminated n^+Si remains unclear, a small amount of oxidation as confirmed from weak O_{2p} signals at a binding energy of 7eV in the UPS valence band spectrum, which could occur during air exposure and pure water rinse, might be responsible for this surface states. Considering the energy shift in E_V and the oxide thickness, the escape depth for electrons excited by 5.3-6eV photons is determined to be 2.2nm. To derive the energy distribution of occupied states from the observed PYS spectrum, we assumed the following: (1) For the electron emissions induced by 4-6 eV photons, the electron escape depths for SiO_2 and Si are kept constant at 2.2nm and 2.5nm [12], respectively. (2) The energy dependence of ionization cross-section of localized defect states can be neglected. These assumptions seem to be valid taking into the fact that a quite similar yield spectrum is measured irrespective of the SiO_2 thickness. Furthermore, to a first approximation, XPS valence band data of c-Si reflects the real density-of-states
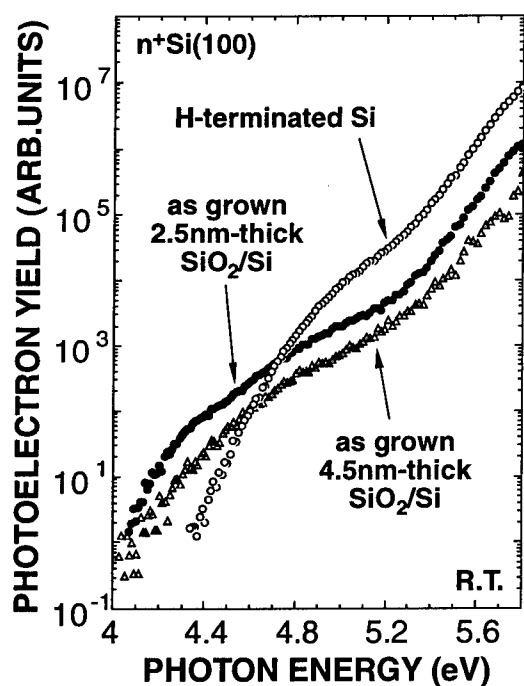
Fig. 1 PYS spectra for as-grown SiO_2 (2.5nm or 4.5 nm, 1000°C)/$n^+Si(100)$ and H-terminated $n^+Si(100)$ which was obtained by a removal of as-grown 2.5nm-thick SiO_2 layer with a 4.5% HF solution.
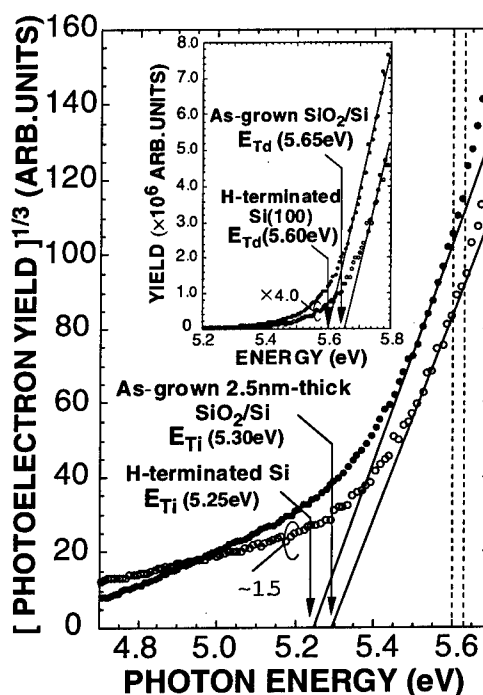
Fig. 2 Cubic root plots of the PYS spectra as-grown 2.5nm-thick $SiO_2/n^+Si(100)$ and H-terminated $n^+Si(100)$ represented in Fig. 1. Inset shows linear plots in a higher energy region.

distribution in the valence band and thus the photoelectron yield Y_s at an energy E_s (typically 5.7eV) above the direct transition threshold E_{Td} can be crudely converted into the density of states by

$$Y_s = \rho_s \lambda_s (N_s/N_t) \exp(-t_{ox}/\lambda_o)$$

where ρ_s is the density of valence electrons in c-Si $(2 \times 10^{23} cm^{-3})$, λ_s and λ_o are the photoelectron escape depths for c-Si and SiO_2, N_s/N_t is the ratio of the states integrated within E_s from the valence band maximum E_v to the sum total of the valence states in the XPS valence band spectrum and t_{ox} is the oxide thickness. And then, from the first derivative of the yield spectrum with respect to photon energy [13], the energy distribution of occupied states existing above E_v of Si from the observed PYS spectrum was quantified . Figure 3 shows the energy distribution of occupied states so derived from each PYS spectrum. In this figure, the influence of oxidation temperature on the defect state distribution was also represented for 2.5nm-thick $SiO_2/$ $n^+Si(100)$. Provided that the density of states at the donor level is equal to the donor concentration in bulk Si, unoccupied state density near the conduction band edge is also deduced from considering the surface Fermi level position as indicated with narrow lines in the figure. Even for 1000°C as-grown SiO_2/n^+Si, the defect state density near midgap is as high as $5 \times 10^{11} cm^{-2} eV^{-1}$ independent of the oxide thickness. And obviously, the lower oxidation temperature, the
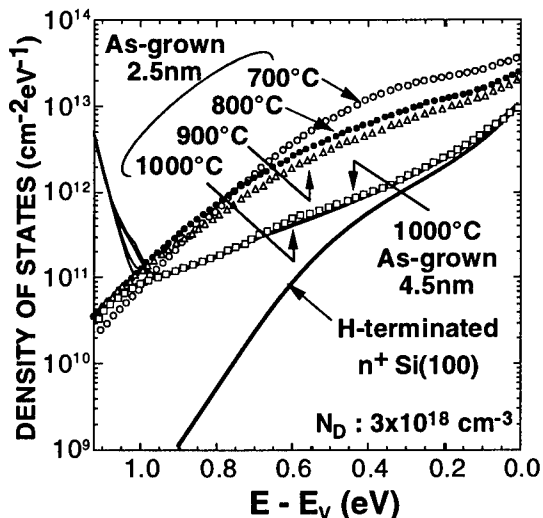
higher defect state density in the Si band gap as reported from C-V characteristics of MOS capacitors with thick SiO_2 [14]. Figure 4 shows the energy distribution of occupied states obtained at each SiO_2-thinning step. It is evident that the defect states above midgap are reduced with progressive SiO_2-thinning. A similar result was obtained for the SiO_2/p-Si system as represented in Fig. 5. This suggests that hydrogen atoms diffuse into the SiO_2 layer and passivate the interface states during the HF etching. Consistently, the fact that the photoemission yield decreased by the oxide thinning is partly recovered by 800°C annealing for 10min in N_2 can be explained by the thermal dissociation of hydrogen-induced defect passivation. Figure 6 shows comparison of the energy distribution of the defect state density determined from PYS measurements and that from quasi-static C-V measurements for as-grown 4.5nm-thick SiO_2/p-Si(100). Because the PYS technique enables us to measure not only filled gap states at the SiO_2/Si interface but also electron trapping states in SiO_2 near the interface whose response time is too slow to be detected with the quasi-static C-V method, that tends to provide higher density of defect states than the value determined by the C-V method. In fact, the midgap
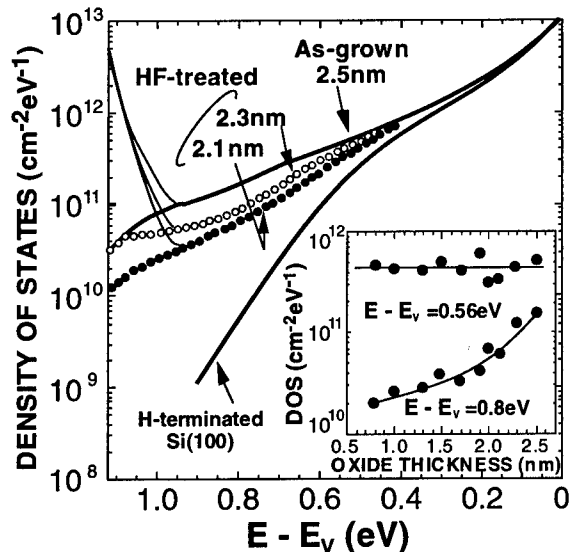
Fig. 4 Energy distributions of density of states determined from PYS spectra for as-grown SiO_2(2.5nm, 1000°C)/n^+Si(100), HF-treated SiO_2 (2.3 or 2.1nm, 1000°C)/n^+Si(100) and H-terminated n^+Si(100). Bold curves, circles and dots denote occupied states. The sum of occupied and unoccupied states for each sample is indicated by a narrow curve. Defect density at 0.8eV from the valence band edge Ev and midgap state density as a function of the oxide thickness are also shown in the inset.

Fig. 3 Energy distributions of density of states determined from PYS spectra for as-grown SiO_2(2.5nm or 4.5nm)/n^+Si(100), and H-terminated n^+Si(100). Bold curves, circles, triangles, squares and dots denote occupied states. The sum of occupied and unoccupied states for each sample is indicated by a narrow curve.
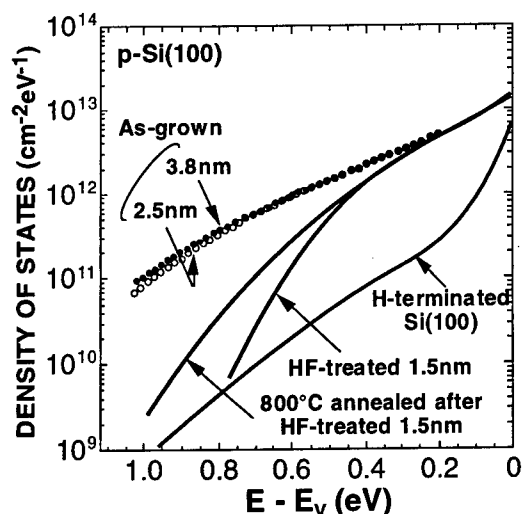
Fig. 5 Energy distributions of occupied state densities for HF-treated SiO_2 (1.5nm, 1000°C)/p-Si(100) before and after 800°C annealing for 10min in N_2. Those for as-grown SiO_2(2.5 or 3.8nm, 1000°C)/p-Si(100) and H-terminated p-Si(100) are also shown as references.
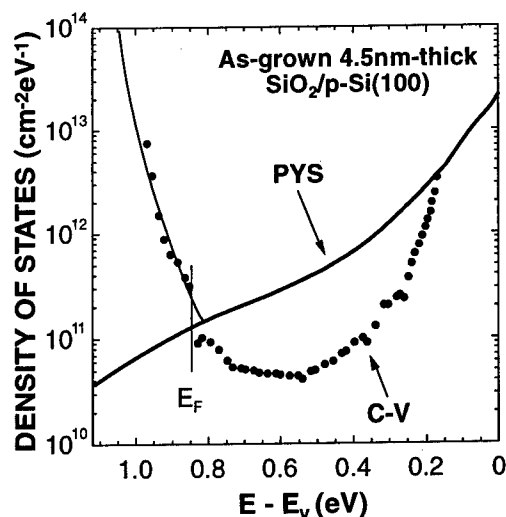
Fig. 6 Energy distributions of gap state densities determined from PYS and quasi-static C-V measurements for as-grown SiO_2 (4.5nm, 1000°C)/p-Si(100).

density determined by the PYS technique is about one order of magnitude higher than that determined by the quasi-static C-V method. It is interesting to note that almost the same amount of the gap state density is obtained at 0.2 eV from the valence band edge by both methods and if we assume the quasi E_F position for electrons during PYS measurements is located at 0.85 eV from the valence band, the gap state distribution in the region above 0.85eV is in a good agreement with the result of the C-V method. These may support the validity of assumptions made in deriving the gap state distribution from the PYS spectrum.

4. CONCLUSIONS

Total photoelectron yield spectroscopy for ultrathin SiO_2/Si enables us to quantify the energy distribution of electronic defects states over nearly entire bandgap of Si with a high enough sensitivity for detect states as low as $10^{10}cm^{-2}eV^{-1}$. As-grown defect states in ultrathin SiO_2/Si are passivated significantly by hydrogen in dilute HF etching as confirmed by highly sensitive PYS measurements.

ACKNOWLEDGMENTS

REFERENCES

[1]   For example, E. H. Nicollian and J. R. Brews, eds., MOS Physics and Technology (John Wiley

& Sons, 1982).

[2]   T. Yoshida, H. Hasegawa and T. Sakai, Jpn. J. Appl. Phys. 38 (1999) 2349.

[3]   H. Song. E. Don, X. Q. Sun and K. P. Farmer, Proc. of Intern. Conf. on Characterization and Metrology for ULSI Technology (Gaithersburg, Maryland, 1998) p. 231.

[4]   W. M. Lau and X. -W. Wu. Surf. Sci. 245 (1991) 345.

[5]   N. Watanabe, A. Omura, H. Nohira and T. Hattori, Extended Abstracts of the 1998 Intern. Conf. on Solid State Devices and Materials (Hiroshima, 1998) p.130.

[6]   T. Yuasa, K. Yamanaka, K. Yoneda, Y. Todokoro and H. Kobayashi, ibid. p.128.

[7]   S. Miyazaki, T. Tamura, T. Maruyama, H. Murakami, A. Kohno and M. Hirose, Mat. Res Soc. Symp. Proc. Vol.500 (1998) 81.

[8]   S. Miyazaki, T. Maruyama, A. Kohno and M. Hirose, Mat. Sci. Semicon. Process (1999) to be published.

[9]   S. Miyazaki, N. Nishimura, M. Fukuda, L. Ley and J. Restain, Appl. Surf. Sci. (1997) 585.

[10] E. O. Kane, Phys. Rev. 127 (1962) 131.

[11] G.W.Gobeli and F. G. Allen, in Semiconductor and Semimetals Vol. 2, eds R. K. Willardson and A. C. Beer (Academic Press, 1966) Chapter 11.

[12] G.W.Gobeli and F. G. Allen, Phys. Rev. 127 (1962) 141.

[13] K. Winer, I. Hirabayashi, L. Ley, Phys. Rev. Lett. 60 (1988) 2697.

[14] B. E. Deal, M. Sklar, A. S. Grove and E. H. Snow, J. Electrochem. Soc. 114 (1967) 266.

Ultrathin nitrided gate dielectrics by plasma-assisted processing

E.P. Gusev, D.A. Buchanan, P. Jamison, T.H. Zabel and M. Copel

IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA
and
IBM Advanced Semiconductor Technology Center, Hopewell Junction, NY 12533, USA

Ultrathin (<3nm) gate dielectrics made by plasma nitridation of SiO_2 films have been studied by a combination of physical (ellipsometry, Nuclear Reaction Analysis, Medium Energy Ion Scattering, and Atomic Force Microscopy) and electrical (C-V, I-V, and constant voltage stress) methods. The main observation we report here is a reduction of leakage current in the nitrided oxides at the expense of reduced (peak) mobility, flatband voltage shift (for high concentration of incorporated nitrogen) and lower breakdown strength.

1. INTRODUCTION

As the thickness of gate dielectrics in ULSI devices shrinks below 3 nm, boron penetration from poly-Si gate, direct tunneling currents, and oxide reliability become significant limiting further CMOS gate oxide scaling. Recent work [1] has emphasized that dielectric reliability may limit thin oxides for logic applications with the limit for the (quantum-mechanical [2]) thickness of SiO_2 films being approximately ~2.4 nm. This work also showed that orders of magnitude improvement in dielectric reliability is required in order to meet the SIA roadmap requirements.[1] Leakage current exceeds the limit of ~1A/cm² for SiO_2 films thinner than ~ 1.7 nm [2]. Nitrided SiO_2 films are considered as a short-term replacement for pure SiO_2. Thermal (oxy)nitridation in N_2O [3] or NO [4] helps to reduce boron penetration. However, the concentration of nitrogen incorporated into the film during thermal NO or N_2O processes is typically below 1x10¹⁵ N/cm² (<15%) [4], which is not enough to increase the dielectric constant of the film and, consequently, reduce the leakage currents. In addition, these processes do not offer significant (i.e. orders of magnitude) reliability improvements.

The goal of this work is to demonstrate the electrical and physical characterization methodology required to assess novel dielectric materials with lower leakage current and reliability as the key benchmarks. For this study, we focussed on understanding a variety of plasma processed nitride films with nitrogen concentrations higher >1x10¹⁵ N/cm².

2. EXPERIMENTAL

Ultrathin 2-2.5 nm SiO_2 oxides were thermally grown on 8" Si p-type wafers after wet-chemical clean. This step was followed by nitridation of the SiO_2 films in N_2 or NH_3 plasmas [5-8]. We explore the nitridation as a function of plasma power (in the 100-1000 W range), pressure, and time. After nitridation, the films (on patterned wafers) were subjected to rapid thermal anneal in nitrogen to reduce plasma-induced damage. (RTA anneal was not performed on unpatterned wafers used for materials analysis). After defining gate stack in the above way (and poly Si deposition), patterned wafers were sent to conventional CMOS processing to complete MOSCAP and MOSFET structures.

After device fabrication, the films were characterized for materials and electrical properties. Materials analysis included ellipsometry (both single wavelength and spectroscopic) for optical thickness measurements, nitrogen nuclear reaction analysis (NRA) to measure integrated amount of nitrogen in the film, medium energy ion scattering (MEIS) for depth profiling, and AFM to monitor surface roughness. Electrical measurements were performed on capacitor and transistor structures (with poly-Si gate) of different sizes using high-frequency C-V, ramp I-V and constant voltage stress techniques. As TBD measurements alone are not sufficient, both Q_{BD} and leakage currents were measured independently in order to fully evaluate the quality of a dielectric film.

3. RESULTS AND DISCUSSION

We found that plasma nitridation is an efficient way of nitrogen incorporation into initial SiO_2 films. We were able to control (integrated) nitrogen concentration in the films in the range of 5×10^{14} - 5×10^{15} N/cm^2 by varying plasma power, pressure, nitridation time and nitrogen source (N_2 and NH_3). These amounts are much higher than that of typical nitrogen concentration in ultrathin thermally grown oxynitrides.
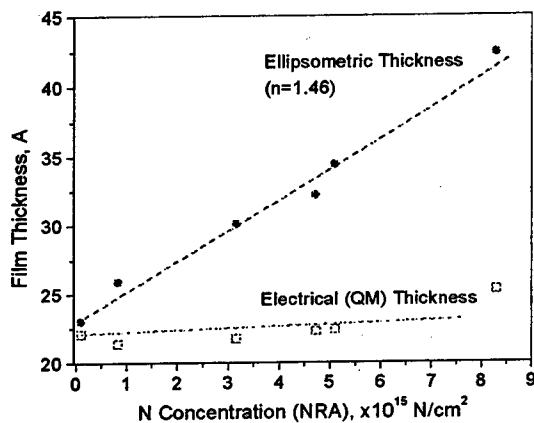
Fig. 1 Electrical (quantum-mechanical thickness defined in the accumulation at -2.1 V) and optical thickness (measured with the refractive index of n=1.46) as a function of nitridation.

Ellipsometric thickness increases as function of nitrogen concentration, whereas electrical thickness of the dielectric almost doesn't change after oxide nitridation (Fig. 1) These observations can be explained by an increase in the dielectric constant of the films (which results in thinner equivalent electrical oxide thickness) and higher value of the refractive index (i.e. thicker optical thickness, respectively) [9].

Fig.2 shows the difference between the electrical and optical (at fixed n=1.46) thickness deduced from the data presented on Fig. 1 as well as other results for NO and N_2O based thermal oxynitrides. One should point out here that the difference can be used as a crude method(calibration curve) to estimate nitrogen concentration based on optical and electrical electrical measurements.
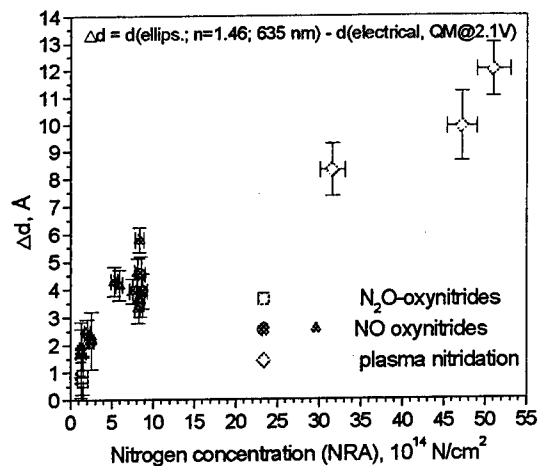
Fig. 2 Open squares correspond to thermally grown N_2O oxynitrides, solid circles and triangles to NO based films, and open diamonds to films after plasma nitridation.

One could think about plasma nitridation as about an energetic process which may change (i.e roughen) oxide film morphology. Non-contact AFM analysis demonstrated that the RMS value of the surface roughness after nitridation was ~ 0.11 nm, similar to that of the initial film (~0.15 nm).

The very good thickness uniformity has also been observed in C-V and I-V measurements (Fig. 3). On both plots and for both processes, data from 16 devices over 8" wafers were taken and analyzed. The analysis resulted in a very good 1σ variation of approximately 0.02 nm.

One can clear see from Fig. 3 that gate leakage current is lower for the gate dielectrics nitrided by plasma compared to that of control SiO_2 of similar thickness. (The similarity of the thicknesses is evidenced from very similar capacitance for the two processes (see upper part of the Fig. 3). The leakage current reduction (normalized to flatband voltage shift for high nitrogen concentrations – see (Fig. 4 below) can be as much as approximately one order of magnitude depending on the concentration of incorporated nitrogen.
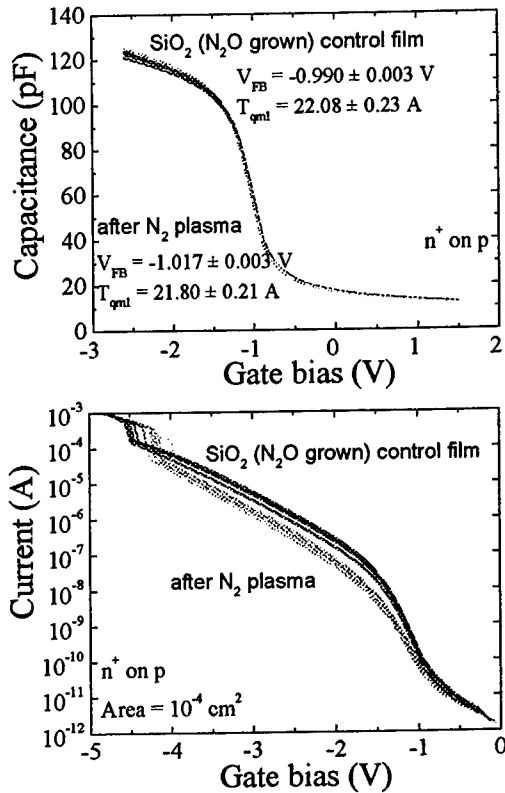
Fig. 3 Capacitance-voltage and gate current-voltage characteristics of N₂O-grown reference film (solid lines) and film after plasma nitridation (dot line).

One can see from the analysis of the C-V curves on Fig. 3, there is a shift in flatband voltage for nitrided films. On Fig. 4, we plot flatband voltage as a function of nitrogen concentration. One can see a linear dependence of the flatband voltage on nitrogen concentration. For high concentrations, the shift may be too large to be compensated by ion implant engineering. In transistor structures, this effect reflects in threshold voltage shift we also observed. We and others believe nitrogen atoms are the source of positive charge in the film. This could be an explanation of the large shift for higher concentration of incorporated nitrogen.

Another detrimental effect of nitrogen located close to the gate is degradation of channel mobility (Fig. 5). NMOSFET peak

(low-field) mobility (G$_{m\_peak}$) was observed to decrease proportionally to the amount of incorporated nitrogen. One should mention however that high-field mobility may still be comparable with SiO₂ what was demonstrated for other nitrided systems.
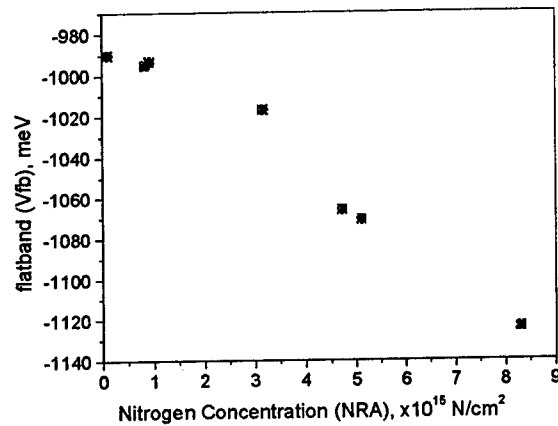
Fig. 4 Flatband voltage for NMOS capacitors.
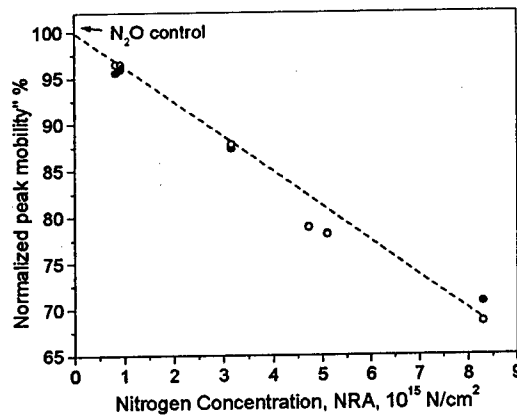
Fig. 5 Peak mobility (normalized with respect to control SiO₂ film).

The reason for reduced channel mobility may be same as that of flatband voltage shift, namely nitrogen located too close to the SiO₂/Si interface. MEIS analysis was performed to address the nitrogen localization issue. Modeling of the MEIS spectrum shown on Fig. 6 resulted in the following result. The thickness of the oxynitride layer for this particular film is ~ 1.5 nm and only ~0.6 nm of the interfacial oxide is nitrogen-free.
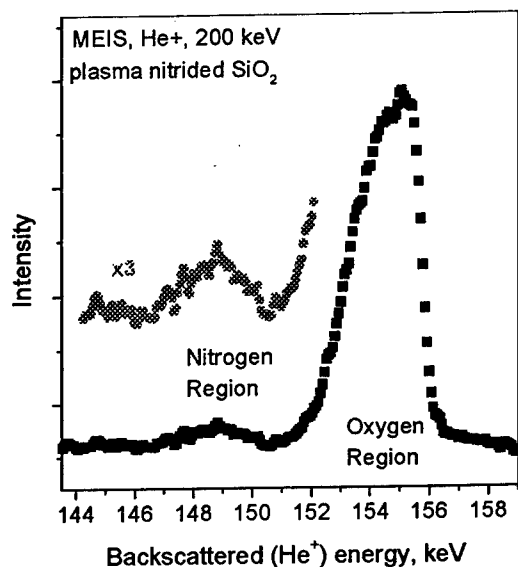
Fig. 6 MEIS spectrum for plasma nitrided film.

Finally, we note that plasma induced damage may be an issue as it may be reflected in lower dielectric strength (lower Q_{bd}) .

4. SUMMARY

Plasma nitridation of ultrathin SiO_2 oxides has been evaluated by electrical and physical methods. We found this manufacturable process offers good thickness uniformity and reduced leakage current. However, there is a flatband voltage shift and reduced peak mobility. Both of the detrimental effects may be large at high nitrogen concentrations. In addition plasma induced damage may effect oxide reliability properties.

ACKNOWLEDGEMENTS

REFERENCES:

1. J.H. Stathis and D.J. DiMaria, Proceedings IEDM 1998, p. 7.2.1
2. Y. Taur, D. Buchanan, W. Chen, D.J. Frank, K.I. Ismail, S.-H. Lo, G.A. Sai-Halasz, R.G. Viswanathan, H.-J.C. Wann, S.J. Wind, and H.-S. Wong, Proc. IEEE 85, 486 (1997)
3. K. A. Ellis and R. A. Buhrman, J. Electrochem. Soc 145, 2068 (1997).
4. M. L. Green, D. Brasen, L. Feldman, E. Garfunkel, E. P. Gusev, T. Gustafsson, W. N. Lennard, H. C. Lu, and T. Sorsch, in Fundamental Aspects of Ultrathin Dielectrics on Si-based Devices, edited by E. Garfunkel, E. P. Gusev, and A. Y. Vul' (Kluwer Academic Publishers, Dordrecht/Boston/London, 1998), p. 181.
5. R. Kraft, T.P. Schneider, W.W. Dostalik, and S. Hattangady, J. Vac.Sci. Technol. B15, 967 (1997)
6. D. Kapila, S. Hattangady, M. Douglas, R. Kraft, M. Gribelyuk, J. Eelectrochem. Soc. 146, 1111 (1999).
7. G. Lucovsky, in Fundamental Aspects of Ultrathin Dielectrics on Si-based Devices, edited by E. Garfunkel, E. P. Gusev, and A. Y. Vul' (Kluwer Academic Publishers, Dordrecht/Boston/London, 1998), p. 147.
8. I.J.R. Baumvol, C. Krug, F.C. Stedile, M.L. Green, D.C. Jacobson, D. Englisham, J.D. Bernstein, J. Shao, A.S. Denholm, P.L. Kellerman, Appl. Phys. Lett. 74, 806 (1999).
9. D.M. Brown et.al J. Electrochem. Soc., 115, 311 (1968).

A Study of Atomically-Flat SiO₂/Si Interface Formation Mechanism, Based on the Radical Oxidation Kinetics

H. Itoh, M. Nagamine, H. Satake, and A. Toriumi

Advanced LSI Technology Laboratory,
Corporate Research & Development Center, Toshiba Corporation
8, Shinsugita-cho, Isogo-ku, Yokohama 235-8522, Japan

We have reported both the experimental results of the flat interface formation and its mechanism based upon the theoretical analysis of the oxygen radical transport in the grown SiO_2. We obtained the logarithmic dependence of the oxide film thickness on the oxidation time on the assumption that the deactivation of the oxygen radicals is proportional to the concentration. The characteristic length "a" of the oxygen radicals plays an important role in forming the flat interface.

1. Introduction

The downscaling of ULSIs requires the ultra-thin gate oxides with higher electrical reliability. It is important that the flatness of the initial surface of Si wafer is maintained in order to obtaine the flat SiO₂/Si interface [1]. However there are many points with the local roughness at the practical Si surface, just before the gate oxidation in the ULSI process. The interface of SiO₂/Si plays an important role in the device reliability. Recently, we have found the oxygen radical oxidation can produce tougher SiO_2 and the atomically-flat interface, compared with the conventional dry oxidation process [2].

In this paper, we investigate the mechanism of the atomically-flat SiO₂/Si interface formation, based on the radical oxidation kinetics.

2. Experimental Results and Discussion

Details of the experimental apparatus and procedure have been presented in another paper [2]. In order to analyze the kinetics of the radical oxidation, experiments on radical and dry oxidation were carried out under the conditions of oxidation temperatures in the range 700 °C to 900 °C, and oxidation pressures of 5 Torr and 760 Torr.

2.1 Formation of Atomically-Flat Interface

Figure 1 shows a cross-sectional transmission-electron-microscopy (TEM) photograph of the SiO₂/Si interface obtained by the radical oxidation at the 10 nm oxide thickness, compared with the conventional dry oxidation.
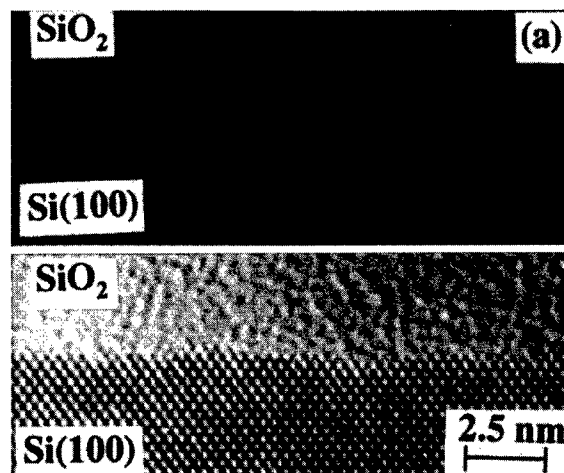
Fig.1 Transmission-Electron-Microscopy (TEM) photograph of the SiO₂/Si(100) interface produced by the radical oxidation at 900 °C at 10-nm oxide thickness (a), compared with the conventional dry oxidation (b) [2].

The atomically-flat interface can be clearly recognized in the radical oxidation. Such a flat interface cannot be obtained in the dry oxidation, where there are many silicon micro-clusters appeared at the interface at the same oxide thickness.. We defined the height of silicon clusters appeared at the SiO_2/Si interface as the interface roughness. It is the peak-to-peak hight in the TEM photograph, where the distance between the silicon lattice images in [100] is 0.14 nm.

Figure 2 shows the dependence of the interface roughness on the oxide thickness in the radical and dry oxidation at 900-1000 °C. The interface roughness decreases monotonously with increase in the oxide thickness in both the radical and dry oxidation [1], while solid and doted lines are for guidance only. The interface roughness of the radical oxide decreases more rapidly than that of the dry oxide does [2]. We think that the dependence of the interface roughness on the oxide film thickness reflects the kinetics of the oxide growth in the radical oxidation.
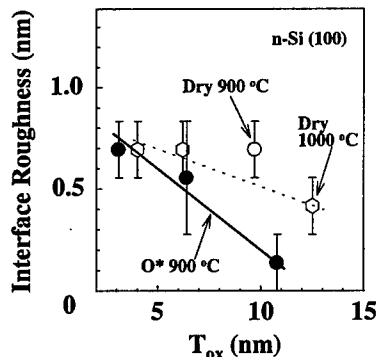
Fig.2 Dependence of the interface roughness on the oxide thickness in the radical and dry oxidation [2]. Solid and doted lines are for guidance only.

2.2 Radical Oxidation Kinetics

In order to clarify the mechanism of the flat interface formation, we analyzed the growth kinetics of the radical oxidation. Figure 3 shows the oxidation time dependence of the SiO_2 film thickness at 900 °C in the radical and dry oxidation.

The radical oxidation proceeds rapidly, compared with the dry oxidation at the same oxidation pressure of 5 Torr, which shows the high reactivity of the oxygen radicals with Si atoms. Furthermore, the radical oxidation proceeds rapidly at the initial stage of the oxidation and slows down quickly with increase in the film thickness, as shown in Fig.4.
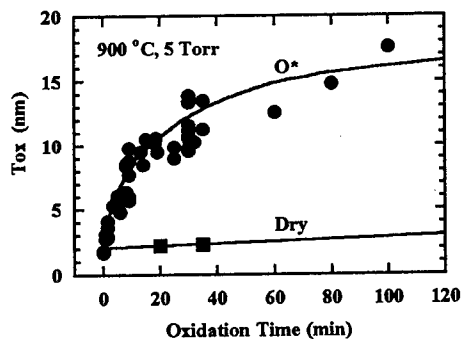
Fig.3 Dependence of the SiO_2 film thickness on the oxidation time at 900 °C in the radical and dry oxidation (5 Torr). Solid lines are for guidance only.
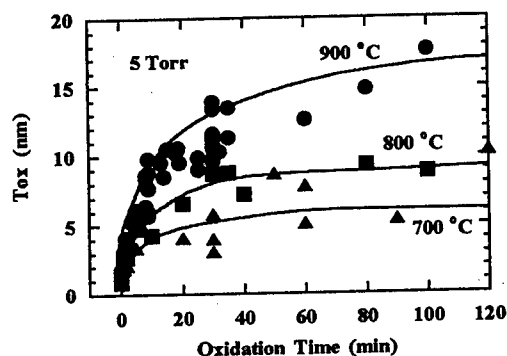
Fig.4 Dependence of the SiO_2 film thickness on the oxidation time as a function of the oxidation temperature in the radical oxidation (5 Torr). Solid lines are for guidance only.

In order to clarify the rapid oxidation of the initial stage in the radical oxidation, we evaluated the initial oxidation rate of the radical oxidation. Figure 5 shows the temperature dependence of the reaction rate R of the oxygen radical with Si atoms, compared with the linear coefficient B/A of the Deal-Grove model in the dry oxidation. R has very large values and is insensitive to the ambient temperature.

It is important that the initial oxidation proceeds rapidly and that the oxide growth is significantly suppressed with increase in the oxide film thickness. This fact strongly suggests that the diffusion of oxygen radicals to the SiO_2/Si interface is suppressed as the oxide layer grows thicker. We think that the deactivation of the oxygen radicals occurs partially during the diffusion through the SiO_2.
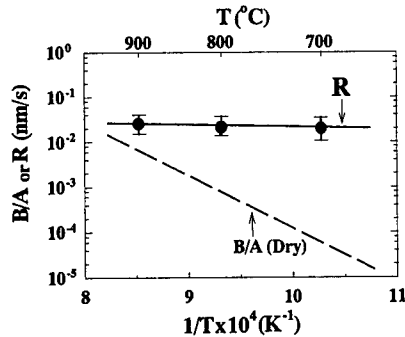
Fig.5 Arrhenius plot of the initial oxidation rate R in the radical oxidation and the B/A coefficient of the Deal-Grove model in the dry oxidation. R is the reaction rate of the oxygen radicals with Si atoms, and shows little temperature dependence.

In order to explain the effect of the deactivation, we considered the mass transfer equation with the effect of the deactivation of the oxygen radicals as schematically described in Fig.6.
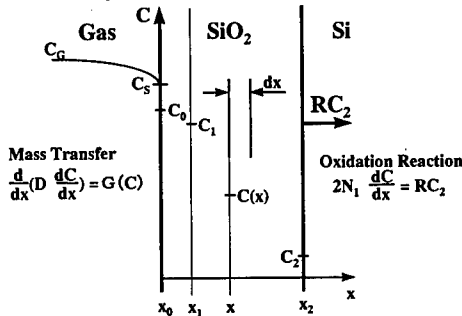
Fig.6 Models of the deactivation of the oxygen radicals during the diffusion through the growing oxide. D is the diffusion coefficient of the oxygen radicals, and C_G, C_S, C_0, C_1, C_2, and $C(x)$ are concentrations of radicals in the bulk of the gas, at the surface, at the outer surface of the oxide, at the initial oxide thickness, at the SiO_2/Si interface, at a distance x, respectively. R is the reaction rate of the oxygen radicals to Si, and N_1 is the number of oxygen atoms incorporated into a unit volume of oxide (2.3×10^{22} cm^{-3}). See reference 3.

Under the assumption of a steady state condition, the mass transfer equation is as follows,

$$\frac{d}{dx}\left(D\frac{dC}{dx}\right) = G(C), \qquad (1)$$

where x is the oxide film thickness, C is the concentration of the oxygen radicals, D is the diffusion constant of the oxygen radicals in SiO_2

(assumed to be constant in the present case), and G(C) is the function representing the deactivation effect of the oxygen radicals.

For the most simple case, we assume that the deactivation of the oxygen radicals is proportional to their concentration, that is, $G(C)=C/\tau$ (τ is the life time of the oxygen radicals), and then solved equation (1). In this case, we can solve equation (1) analytically, and there is logarithmic dependence of the oxide film thickness x on the oxidation time t (direct logarithmic relation),

$$x = a \ln(t+c) + b, \qquad (2)$$
$$C(x) = C_0 \exp(-x/a) \qquad (3)$$

where a, b and c are constants, and C_0 is the concentration of the oxygen radicals at the outer surface of the oxide.

The experimental results are well fitted to the theoretical ones with Eq. (2) as shown in Fig.7, in which solid lines denote results fitted with the parameters listed in Table 1 based on Equation (2).

We assumed that the diffusion constant of the oxygen radicals are constant in our theorital treatment. We think that our model is not only one. If we assumed the variable diffusion constant of the oxygen radicals without the deactivation of the oxyen radicals, we should obtaine another oxidation scheme. In this case, however, the physical meaning of the oxygen radicals is not so clear as that in our present model.
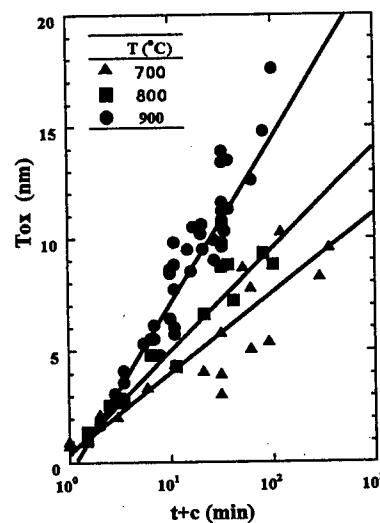
Fig.7 The dependence of oxide thickness on the oxidation time. Solid lines denote results fitted with parameters listed in Table 1, based on Eq.(2). The slope of each line means the characteristic length a.

Table 1　List of parameters a, b, and c

T (°C)	a(nm)	b (nm)	c (min)
700	1.6	0.5	1.0
800	2.0	0.18	1.5
900	3.5	0.0	2.0

2.3 Flat-Interface Formation Mechanism

The parameter "a", which is $(D\tau)^{0.5}$, represents the characteristic length of the oxygen radicals. Oxygen radicals move this distance decreasing exponentially (Eq.3). As shown in Fig.8, the diffusion of the oxygen radicals is strongly suppressed in the direct logarithmic scheme, compared with the conventional Deal-Grove model, in which the oxidant species are directly proportional to the gradient of the oxidant concentration. Then, the concentration of the oxygen radicals at the SiO_2/Si interface becomes extremely low with increase in the oxide film thickness, as shown in Fig. 8, which is important in the formation of the atomically-flat interface of the SiO_2/Si.
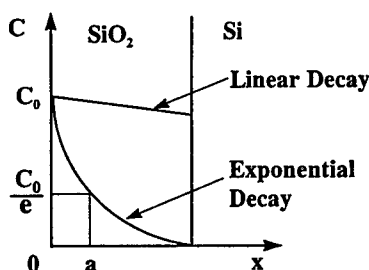
Fig.8　Schematic view describing lower concentration of the oxygen radicals in the direct logarithmic relation, compared with a case of the linear decay of the concentration. "a" is the characteristic length of the oxygen radicals.

The reaction of oxygen radicals with Si atoms is very fast, as shown in Fig.5. Thus, under the small concentration of the oxygen radicals, the radicals mainly react with Si atoms in the region with the length "a", while they cannot interact with Si atoms outside of the "a" length. In the other point of view, Si atoms constructing Si micro-clusters appeared at the SiO_2/Si interface interact only the oxygen radicals in the region with the "a" length. In the case of larger value of "a", the oxidation of Si atoms in the micro-clusters proceeds more rapidly. "a" has a temperature dependence as shown in Table 1, which means that the flatness is obtained relatively faster in the higher temperature oxygen radical process.

It is difficult to achieve the flat interface in the

reaction-limited region [4], which is same as that in the Deal-Grove process [3]. Thus, it is inferred that the diffusion of oxygen radical in grown SiO_2 might be responsible for the flat interface. The parameter "a" has a physical meaning of the characteristic length in which the diffusion of oxygen radical can influence the oxidation process. In that sense, "a" can be called as "the interface roughness screening length" in the radical oxidation. This is schematically described in Fig.9.
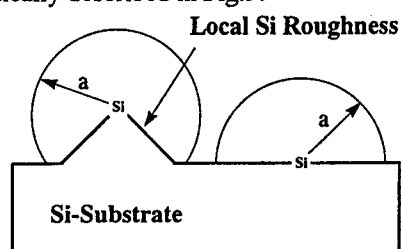
Fig.9　Schematic view describing the atomically flat interface formation. Oxygen atoms in the region of "a" react to the local Si roughness at the interface. In the case of large "a" value, the interface becomes flat faster more easily.

3. Summary

We report both the experimental results of the flat interface formation and its mechanism based upon the theoretical analysis of the oxygen radical transport in the grown SiO_2. These results demonstrate the kinetics of the radical oxidation of Si on the assumption of oxygen-radical deactivation. We obtained the logarithmic dependence of the oxide film thickness on the oxidation time on the assumption that the deactivation of the oxygen radicals is proportional to the concentration. The characteristic length "a" of the oxygen radicals plays an important role in forming the flat interface. We propose the radical oxidation is a very promising process to control the SiO_2 uniformity and its reliability.

References

[1] M. Niwa, et al., *Jpn. J. Appl. Phys.*, **33**, p.388 (1994).

[2] N. Nagamine, et al., in IEDM Tech. Dig., p.593 (1998).

[3] A. S. Grove, "Physics and Technology of Semiconductor Devices", (John Wiley &Sons, Inc.) pp.23-31 (1967).

[4] A. H. Carim and R. Sinclair, *J. Electro. Chem. Soc.*, **134**, p. 741 (1987).

Effect of local surface structure on electronic properties of hydrogenated silicon surfaces

Th. Dittrich [a], V. Yu. Timoshenko [a,c], M. Schwartzkopff [a], E. Hartmann [a],
J. Rappich [b], P. K. Kashkarov [c], and F. Koch [a]

[a]Technische Universität München, Physik Department E16
D-85748 Garching, Germany

[b]Hahn-Meitner-Institut, Abt. Photovoltaik
Rudower Chaussee 5, D-12489 Berlin, Germany

[c]M.V. Lomonosov Moscow State University, Faculty of Physics
119899 Moscow, Russia

Morphology and electronic properties of hydrogenated Si surfaces are studied by scanning tunneling microscopy and by methods of surface photovoltage and photoluminescence (PL), respectively. The microscopic roughness on the nm scale is more important than the macroscopic roughness of a hydrogenated Si surface for the densities of nonradiative surface defects and hole traps. The adsorption of water or oxygen molecules on a hydrogenated Si surface leads to an increase of the intensity of interband PL of c-Si which is controlled by surface nonradiative recombination. This effect is interpreted as a decrease of the recombination cross section of a nonradiative surface defect due to charge trapped at the adsorbed molecules.

1. INTRODUCTION

Hydrogenated Si surfaces are important not only for ultraclean processing of semiconductor devices [1]. Reactive surface sites and/or adsorbed molecules at hydrogenated Si surfaces can be used, for example, for the modification of Si surfaces and interfaces at low temperatures by grafting [2]. This is of interest for the connection of Si based electronic devices with organic materials. However, the control of the density and distribution of reactive surface sites such as dangling bonds (DBs) and of adsorbed molecules and their correlation with the morphology of hydrogenated Si surfaces has not been well studied up to now.

The density of the interface states (D_{it}^{min}) and of the concentration of nonradiative (nr) surface defects at Si surfaces can be measured ex-situ and in-situ by the pulsed surface photovoltage (SPV) [3,4] and photoluminescence techniques (PL) [5,6], respectively. Recently we showed by quantitative analysis the sensitivity of the pulsed PL technique

to the nr defect concentration in very broad range of concentrations [7].

In the present work we describe effects of the local surface structure on electronic properties of hydrogenated Si surfaces. In the first two parts we summarize known facts concerning D_{it}^{min} and positive charging of hydrogenated Si surfaces. New experimental results are presented in the next two parts. Methods of scanning tunneling microscopy (STM), SPV and PL are used. Experimental details are skipped for shortness (the respective references are given).

2. SURFACE MORPHOLOGY AND D_{it}^{min}

Different types of hydrogenated Si surfaces can be distinguished by their surface morphology. Si surfaces treated in concentrated HF are hydrogenated [8] but not atomically flat [9]. Atomically flat and ideally hydrogenated Si surfaces are obtained on Si(111) facets after treatment in NH$_4$F solution [9, 10]. The values of D_{it}^{min} are in the order of 10^{12}

[4,11] and below 10^{11} eV^{-1}cm^{-2} [12] for HF and NH$_4$F treated Si(111) surfaces, respectively. The surfaces of as-prepared porous Si are also hydrogenated [13] and well passivated ($D_{it}^{min} < 10^{11}$ eV^{-1}cm^{-2} [14]). The lowest values for D_{it}^{min} ($D_{it}^{min} \approx 10^{10}$ eV^{-1}cm^{-2} [15]) were reached on electrochemically hydrogenated Si surfaces. The most important step for the electrochemical hydrogenation is the anodic oxidation in the oscillating regime [16] which produces a macroscopically rough surface [17] (see also this work).

3. POSITIVE CHARGING

Hydrogenated Si surfaces are in general positively charged (see, for example, [17] and refs. therein). The values of "fixed" positive charge (Q_f), hysteresis and D_{it}^{min} measured by SPV depend sensitively on the conditions of the measurement [17]. The hole traps at hydrogenated Si surfaces are relatively slow. A correlation between Q_f, hysteresis and D_{it}^{min} was observed [17]. The largest values of D_{it}^{min} are found for hydrogenated Si surfaces with the largest hysteresis and Q_f.
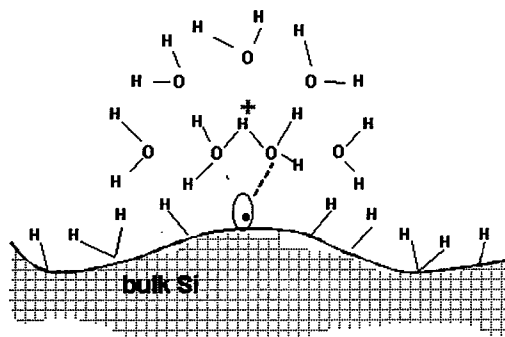
Figure 1. Configuration for a Si dangling bond - [H$_5$O$_2$]$^+$ complex surrounded by physisorbed water molecules.

The origin of hole traps at hydrogenated Si surfaces was discussed from the point of view of chemi- and physisorbed water molecules [17]. Wa-

ter molecules are of the donor type [18] and they form charged complexes as [H$_5$O$_2$]$^+$. The hole traps are localized, in our opinion, near the defects which are responsible for the quenching of the PL intensity since the anticorrelation between D_{it}^{min} and the PL intensity is well. Figure 1 shows an example for a respective configuration. A DB serves in this case as the initial adsorption site for a water molecule.

4. ROLE OF MICROROUGHNESS

STM micrographs of Si(111) surfaces hydrogenated electrochemically after anodic oxidation in the oscillating regime for 2 and 30 min are compared in figure 2. The electrochemically treated Si surfaces contain pore like structures on the scale of some of 10 nm. The depth of these structures is less than 5 nm. The pore like structures are more pronounced and small protrusions diminished for longer time of anodic oxidation in the oscillating regime. We remark that the surface structure of the electrochemically hydrogenated Si surfaces weakly depends on the orientation of the crystal.

The formation of a hydrogenated Si surface can be monitored in-situ by the anodic current transient and by the stroboscopically probed PL intensity. Examples are presented in figure 3. The chemical etching of the oxide and Si is currentless. The formation of the hydrogenated Si surface is characterized by an anodic current transient [16] and by a strong increase of the PL intensity due to the passivation of reactive surface sites. The PL intensity reaches the maximum when the anodic current transient levels out. The maximum of the PL intensity is higher after longer anodic oxidation in the oscillating regime, i.e. for the surface with the lower micro-roughness.

Nonradiative surface defects are located at steps, kinks etc. which are responsible for the micro-roughness. This is confirmed also by spectroscopic ellipsometry and SPV at BHF treated Si(111) surfaces [19]. The probability of the formation of DBs by oxidation of Si backbonds is relatively high at surface steps or corners. Microscopically smooth hydrogenated Si surfaces are stabilized and therefore less reactive [20].
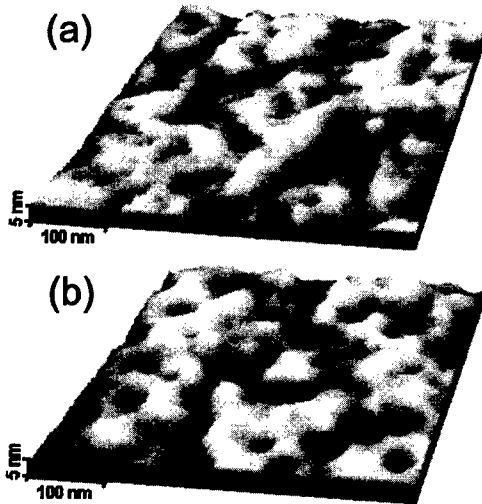
Figure 2. STM micrographs of Si(111) surfaces hydrogenated electrochemically after anodic oxidation in the oscillating regime for 2 (a) and 30 (b) min. See for the experimental details [17].
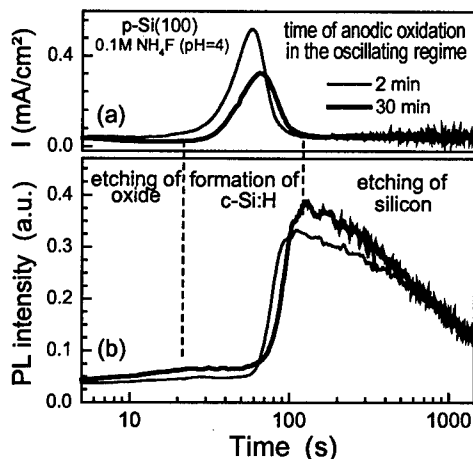
Figure 3. Anodic current transient (a) and time dependent PL intensity (b) during the formation of a hydrogenated Si surface after anodic oxidation in the oscillating regime for 2 and 30 min. See for the experimental details [6] and [15].

5. EFFECT OF MOLECULE ADSORPTION

Figure 4 shows an example for the influence of adsorbed water and oxygen molecules at a hydrogenated Si surface on the photovoltage (U_{ph}) and on the PL intensity in the case of HF treated p-Si(111). The photovoltage and PL signals were excited with a laser diode (wavelength 902 nm, pulse duration 150 ns) and with a N_2 laser (wavelength 337 nm, pulse duration 0.5 ns), respectively. The signal of interband PL (1.1 μm) of c-Si was used as a probe of surface nonradiative defect density. The experimental details of PL and SPV measurements are described in [5-7] and [15], respectively. The U_{ph} and PL measurements were carried out in different high vacuum chambers (oil free vacuum, basic pressure 10^{-6} mbar). The sample was moved away from the mica spacer during the adsorption procedure of the SPV experiment. The reversibility of the U_{ph} and PL signals was very good during the adsorption/pumping cycles.

The value of $|U_{ph}|$ increased for adsorption of water molecules and decreased for adsorption of oxygen molecules. This is not surprising since water and oxygen molecules are of the donor type or acceptor type, respectively [18].

The PL intensity increased strongly after adsorption of both water and oxygen molecules, while the increase was stronger for the adsorption of water molecules. The relative increase of the PL intensity was independent of the absolute value of the PL intensity, i.e. on the initial concentration of nr surface defects. We remark that the PL intensity was independent of illumination with bias light while $|U_{ph}|$ dropped under bias light.

The stationary values of $|U_{ph}|$ and PL intensity were reached some minutes after switching from high vacuum to the H_2O or O_2 atmospheres. This shows that some charge transfer to adsorbed water or oxygen molecules is necessary to influence $|U_{ph}|$ and PL intensity. We point out that the features demonstrated in figure 4 are characteristic for all hydrogenated Si surfaces independent of type of hydrogenation, doping level, type of doping and crystallographic orientation.
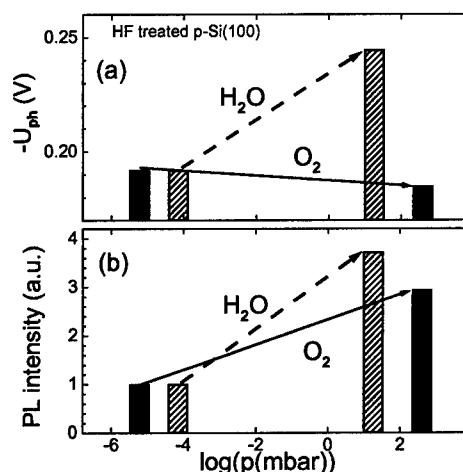
Figure 4. Photovoltage (a) and PL intensity (b) for a hydrogenated Si surface in high vacuum and in water or oxygen atmospheres.

The dependence of U_{ph} and PL intensity on the adsorption of water and oxygen molecules gives evidence that the increase of the PL intensity is not induced by changes of the surface band bending. The excellent reversibility suggests that the number of nr surface defects is constant during the adsorption experiments. Therefore, the adsorbed molecules directly influence the capture cross sections for electrons and holes of the nr surface defects.

6. CONCLUSIONS

The electronic properties of hydrogenated Si surfaces are determined by complexes consisting of defects like DBs and adsorbed molecules as shown in figure 1. The surface defects are located at microscopic imperfections on the hydrogenated Si surface and serve as initial adsorption sites for water molecules. The capture cross sections of the defects depend strongly on adsorbed molecules.

ACKNOWLEDGEMENTS

REFERENCES

1. See, for example, 1st int. Symp. On Ultraclean Processing of Silicon Surfaces. UCPSS'92, Sept.17-19, Leuven (Belgium) 1992.
2. P. Allongue, Phys. Rev. Lett., 77 (1996) 1986.
3. Y. W. Lam, J. Phys. D: Appl. Phys., 4 (1971) 1370.
4. K. Heilig, H. Flietner, J. Reineke, J. Phys. D: Appl. Phys., 12 (1978) 1370.
5. V. Yu. Timoshenko, J. Rappich, Th. Dittrich, Jpn. J. Appl. Phys., 36 (1997) L58.
6. J. Rappich, V. Yu. Timoshenko, Th. Dittrich, J. Electrochem. Soc., 144 (1997) 493.
7. V. Yu. Timoshenko, A. B. Petrenko, M. N. Stolyarov, Th. Dittrich, W. Füssel, J. Rappich, J. Appl. Phys., 85 (1999) 4171.
8. E. Yablonovich, D. L. Allara, C. C. Chang, T. Gmitter, T. B. Bright, Phys. Rev. Lett., 57 (1986) 249.
9. K. Itaya, R. Sudawara, Y. Morita, H. Tokumoto, Appl. Phys. Lett., 60 (1992) 2534.
10. G. S. Higashi, Y. J. Chabal, G. W. Trucks, K. Raghavashari, Appl. Phys. Lett., 56 (1990) 656.
11. Th. Dittrich, H. Angermann, W. Füssel, H. Flietner, phys. stat. Sol. (a), 140 (1993) 436.
12. H. Angermann, K. Kliefoth, W. Füssel, H. Flietner, Microel. Engen., 28 (1995) 51.
13. A. Venkateswara Rao, F. Ozanam. J.-N. Chazalviel, J. Electrochem. Soc., 138 (1991) 153.
14. Th. Dittrich, K. Kliefoth, I. Sieber, J. Rappich, S. Rauscher, V. Yu. Timoshenko, Thin Solid Films, 276 (1996) 183.
15. S. Rauscher, Th. Dittrich, M. Aggour, J. Rappich, H. Flietner, H. J. Lewerenz, Appl. Phys. Lett., 66 (1995) 3018.
16. H. Gerischer, M. Lübke, Ber. Bunsenges. Phys. Chem., 92 (1988) 573.
17. Th. Dittrich, M. Schwartzkopff, E. Hartmann, J. Rappich, Surface Science (in press).
18. V. F. Kiselev, O. V. Krylov, Electronic Phenomena in Adsorption and Catalysis on Semiconductors and Dielectrics, Springer, Berlin, 1987.
19. H. Angermann, W. Henrion, M. Rebien, J.-T. Zettler, A. Röseler, Surf. Sci., 388 (1997) 15.
20. Th. Dittrich, V. Yu. Timoshenko, J. Rappich, Appl. Phys. Lett., 72 (1998) 1635.

Al/SiO$_2$(2.0-2.5 nm)/p-Si tunnel junction as a light emitter

N. Asli[a]*, S.V. Gastev[b], I.V. Grekhov[b]*, P. Seegebrecht[a],
A.F. Shulekin[b]*, S.E. Tyaginov[b]*, M.I. Vexler[a]* and H. Zimmermann[a]

[a]Lehrstuhl für Halbleitertechnik, Christian-Albrechts-Universität,
Kaiserstraße 2, D-24143 Kiel, Germany

[b]A.F.Ioffe Physicotechnical Institute,
Polytechnicheskaya Str. 26, 194021 St.-Petersburg, Russia

Room-temperature hot electron luminescence of MOS tunnel diodes has been experimentally studied. Complete radiation spectra are presented for the first time for the direct-tunneling regime. Essential parameters are: oxide thickness 2.0...2.5 nm, doping level $2 \cdot 10^{18}$ cm^{-3}, current density $\sim 10^2$ A/cm^2, and photon energy 0.73...3.2 eV.

1. INTRODUCTION

Today, there is no necessity to motivate the actuality of a study of MOS structures with an ultrathin ($d_{ox} \sim 2.0$ nm) oxide layer, considered as basic components of future MOSFETs [1].

Hot-electron-induced light emission belongs to those aspects of their operation, which have so far received relatively little attention. Just several papers on the luminescence of MOS diodes with a sub-3 nm insulator layer can be cited [2,3].

The results of investigation of light generation in such structures would naturally find their place as extension of a series of previous works on optical properties of MOS devices with different oxide thickness. By varying d_{ox}, one can "shift" the range of electron and photon energies where convenient and reliable measurements may be performed. For any given d_{ox}, the apparent restriction at high electron energies is the damage of structures and at low energies is the impossibility of detecting light (too low currents J).

2. CONCEPTS OF LUMINESCENCE

In the diodes on p-Si, monoenergetic hot electrons (HE) are injected into the semiconductor under the accumulation (forward bias) condition. In this case, their initial energy E can be unequiv-

ocally defined as the difference of the Fermi level of the metal and the conduction band edge within the quasi-neutral Si (Fig.1). Such a definition is justified since the electron mean free path λ substantially exceeds the accumulation layer width [3]. The value of E depends on the applied bias
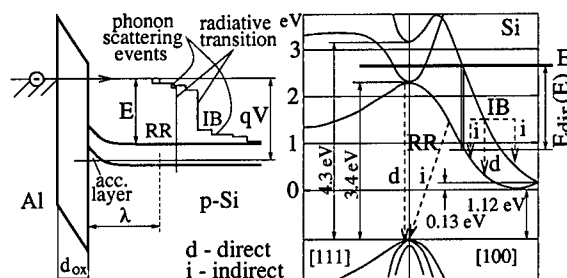
Figure 1. Radiative transitions in the Al/SiO$_2$/p-Si diode (an example).

and is suggested to govern the luminescence spectrum of a tunnel MOS structure.

Injected HE are involved in radiative intraband (IB) transitions inside the conduction band(s) of Si as well as in radiative recombination (RR) with holes. Both these types of transitions may be direct or indirect in the k-space, with corresponding

limitations for the photon energies $\hbar\omega$ introduced by E and by the band structure of Si (Fig.1). For example, the spectral range of *direct* IB transitions is $\hbar\omega = [0...E_{dir}(E)]$ (Fig.1).

A light quantum is usually emitted, if any, after the multiple phonon scattering events. No peak in the HE radiation spectra associated with the initial electron energy should therefore appear.

These spectra could be estimated as

$$\frac{dI}{d(\hbar\omega)} \approx \frac{J}{2q} \frac{\hbar\omega}{\hbar\omega_0} \int_0^E \frac{1}{\tau_{ph}^{-1}(\tilde{E})} \frac{d\tau_{le}^{-1}(\tilde{E}, \hbar\omega)}{d(\hbar\omega)} \, d\tilde{E} \quad (1)$$

In this equation, $d\tau_{le}^{-1}/d(\hbar\omega)$ is the rate of photon emission in Si in the spectral range of $[\hbar\omega ... \hbar\omega + d(\hbar\omega)]$, τ_{ph}^{-1} is a phonon scattering rate and $\hbar\omega_0$ is optical phonon energy. Practically, however, it is not easy to perform the calculations, because of the lack of reliable data on $d\tau_{le}^{-1}/d(\hbar\omega)$.

Light emission in the near-E_g range is additionally contributed to by the radiative recombination of thermalized electrons. The intensity of this component is proportional to the electron lifetime and behaves roughly as $\exp(-\hbar\omega/kT)$, being superimposed to the spectrum given by Eq.(1).

3. TECHNICAL DETAILS

3.1. Sample fabrication

We have fabricated and investigated the tunnel MOS diodes Al/SiO$_2$/p-Si (111) with acceptor concentration N_A of $2 \cdot 10^{18}$ cm^{-3}. Thin oxide films ($d_{ox} = 2.0...2.5$ nm) were grown by oxidation in dry O$_2$ at 700 °C. Non-transparent circular Al electrodes (area $S = 1.26 \cdot 10^{-3}$ cm^2) were deposited at 200 °C.

3.2. Characterization

Forward-bias current-voltage characteristics of our samples are shown on Fig.2. They remained unchanged during the measurements. The electron energy was practically estimated as

$$E \approx qV - E_g \quad (2)$$

where V is to be understood as the applied bias corrected, if necessary, for the parasitic contact and ohmic voltage drops.

The radiation emitted along the border of an Al electrode was transmitted (via a glass waveguide
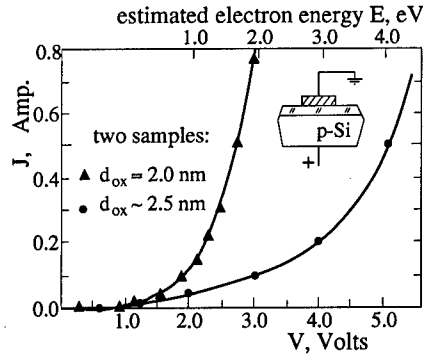
Figure 2. Current-voltage characteristics of studied MOS tunnel structures.

with 6 mm diameter) either to the monochromator and further to the photomultiplier ($\hbar\omega > 1.4$ eV) or to the appropriate optical filters and an InGaAs-detector ($\hbar\omega > 0.73$ eV).

The measurements were performed at 300 K.

4. RESULTS OF OPTICAL MEASUREMENTS AND THEIR DISCUSSION

4.1. Measurement of intensity

For a forward-biased diode we measured the detector-limited light intensity I in the spectral ranges of $\hbar\omega = [0.73...1.00$ eV] and $[1.4...3.5$ eV] using adequate filters for selecting the range of $[\hbar\omega_{min}...\hbar\omega_{max}]$. Each dependency of the ratio of intensity I to the diode current J on the bias V (Fig.3) is characterized by a certain threshold voltage V_{thr}. However, for $\hbar\omega = [1.4...3.5$ eV] there is some intensity even in the sub-threshold region ($V < V_{thr}$). This is a quite natural result. Indeed, the photons with $\hbar\omega = [0.73...1.00$ eV] cannot be produced by recombination, and there is no light in this range up to $E \approx \hbar\omega_{min} = 0.73$ eV (i.e. up to $V_{thr} = 1/q \cdot (E + E_g) \approx 1.8$ V, Fig.3). Conversely, in the spectral range of $\hbar\omega = [1.4...3.5$ eV] there is some recombination-related intensity at lower electron energies E than 1.4 eV.

Similar measurements were performed in [2] using MOS structures with $d_{ox} \sim 3$ nm. In [2], the values of $\hbar\omega_{min}$ were rather high (> 1.7 eV)
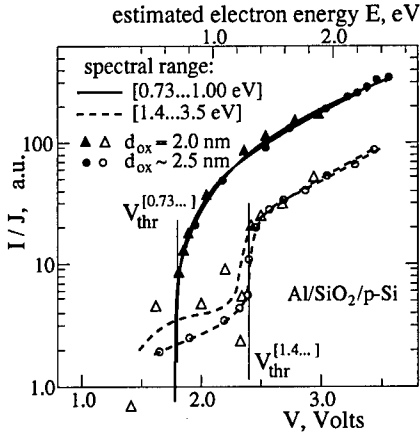
Figure 3. Light intensity in two different spectral ranges, depending on the applied bias.

Figure 4. Radiation spectra of MOS tunnel structure with $d_{ox} \sim 2.5$ nm. Along the vertical axis, the values of $\ln(\mathrm{d}I/\mathrm{d}[\hbar\omega])$ in a.u. are presented.

and the threshold voltages V_{thr} have been referred to the condition $\hbar\omega_{min} = E_{dir}$, not to $\hbar\omega_{min} = E$. However, this does not contradict our results, since in our case, i.e. for $E < 1.5$ eV, $E_{dir} \approx E - (0.7 - \sqrt{E})^2 \sim E$ (E is in eV).

4.2. Measurement of spectra

The spectra were recorded for the Al/SiO$_2$/p-Si diodes with two oxide thicknesses (Figs.4,5).

As it should have been expected, the enhancement of V leads to the shift of a high-energy border of the HE-IB-spectrum to the right (Fig.4). Similar effect could be deduced from an inspection of Fig.3. But in Fig.4 we are dealing with relatively high E, and the borders of the IB-spectra are seen to correspond to $\hbar\omega \approx E_{dir}$ rather than to $\hbar\omega \approx E$ (e.g., for $V = 4.1$ V, we have $E = 3.0$ eV and $E_{dir} \sim 2.5$ eV [2], see also Eq.(2)). This observation supports the statement of [2] about the dominant role of *direct* IB-processes.

For the sample with $d_{ox} = 2.0$ nm, the electron energy E did not exceed 1.4 eV ($V < 2.5$ V, higher voltages were not tolerated), so that the contribution of intraband processes remained outside Fig.5, plotted only for $\hbar\omega > 1.4$ eV.

At $\hbar\omega$ higher than the high-energy border of the direct IB luminescence (E_{dir}) one can see a

Figure 5. Spectra of a diode with $d_{ox} = 2.0$ nm.

weak signal (Fig.4) arising from the indirect IB-transitions ($\hbar\omega < E$) and RR ($E_g < \hbar\omega < E + E_g$). In Fig.5, the radiation throughout the entire spectral range is due to recombination.

These data complete the results obtained in [2] for thicker MOS structures. Namely, the smallness of d_{ox} in our devices enabled us to perform exact measurements in the range of lower electron and/or photon energies E, $\hbar\omega$ than in [2] as well as to observe a HE-RR component, whose relative role reduces with increasing E [4].

Note also that at lowest photon energies falling

into Figs.4,5 (< 1.5 eV) there is a substantial increase of intensity originating probably from a recombination of already *thermalized* electrons.

Possible features of spectra owing to the direct recombination processes ($\hbar\omega \sim 3.4$ eV, ~ 4.3 eV) need to be studied separately.

5. ADDITIONAL REMARKS

5.1. On the oxide degradation

Our preliminary results suggest that light emission from the tunnel MOS diodes on p-Si should not be strongly affected by the oxide degradation. An evidence for this statement is provided in Fig.6 where the same dependencies as in Fig.3 (i.e. I/J) are shown for a structure before and after the stress. The stress-induced enhancement of
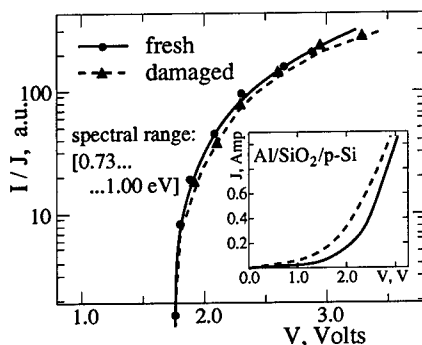
Figure 6. The ratio I/J for a diode measured before and after the oxide damage. *Inset*: corresponding current-voltage characteristics.

J is large, but the values of I/J are not changed as dramatically, i.e. the radiation intensity has increased roughly proportional to the current. Possible explanation assumes that the degradation may eventually be formalized as a local "thinning" of the oxide film. The electron energy E is set by V (Eq.(2)) and is independent on d_{ox}. Therefore, the native or created non-uniformities of oxide thickness should be of minor importance for optical measurements.

5.2. On the photon emission rates

Luminescence measurements on the MOS tunnel diodes can serve for extraction of photon emission rates $d\tau_{le}^{-1}/d[\hbar\omega](E,\hbar\omega)$. Moving all the multipliers of Eq.(1), except the integral, into the left side and then numerically taking a partial derivative over E from both sides, we may find the photon emission rate. For this purpose, we need a series of spectra $dI/d[\hbar\omega]$ to be measured for different J, E, and the dependency $\tau_{ph}^{-1}(E)$ to be independently known from the literature [5]. The data on photon emission rate would be very useful for the analysis of HE luminescence in any silicon device. The realization of such a method remains, however, beyond the scope of this work.

6. CONCLUSION

In this work, an experimental study of hot electron luminescence in the MOS tunnel structures Al/SiO$_2$/p-Si with a 2.0...2.5 nm oxide layer has been performed. Due to the small insulator thickness, the current density was high enough even at low electron energies. Reliable optical measurements covered the broad spectral range: from 0.73 to 3.2 eV. Measured spectra indicated the dominant role of direct intraband optical transitions and revealed also the recombination component, including the contribution of thermalized electrons. The position of the high-energy border related with the intraband radiation depends on the energy of injected electrons which is set by the applied forward bias.

REFERENCES

1. H.S. Momose, M. Ono, T. Yoshitomi, T. Ohguro, S. Nakamura, M. Saito and H. Iwai IEEE Trans. ED-43 (1996) 1233.
2. E. Cartier, J.C. Tsang, M.V. Fischetti and D.A. Buchanan, Microelectron. Eng. 36 (1997) 103.
3. I.V. Grekhov, A.F. Shulekin and M.I. Vexler Proc. MIEL-21, Niš, Yugoslavia (1997), 153.
4. J. Bude, N. Sano and A. Yoshii, Phys. Rev. B45 (1992) 5848.
5. Y. Wang and K.F. Brennan, J. Appl. Phys. 75 (1994) 313.

Characteristics of ultrathin [4-7 nm] gate oxides for SiGe quantum well MOS structures

S. Kar[a] and P. Zaumseil[b]

[a]Department of Electrical Engineering, Indian Institute of Technology, Kanpur - 208016, India

[b]Institute for Semiconductor Physics, Walter-Korsing-Strasse 2, 15230 Frankfurt (Oder), Germany

SiGe quantum well metal-oxide-semiconductor (MOS) structures were realized by the preparation of the undoped Si buffer layer, the undoped SiGe quantum well layer, and the undoped Si cap layer by molecular-beam epitaxy or low pressure chemical vapor deposition, and low thermal budget rapid thermal oxidation of the Si cap layer in dry oxygen at 1100 °C for 30-65 s. Comprehensive electrical characterization by steady state admittance spectroscopy was carried out. The relaxation of the strained SiGe layer was monitored by x-ray diffraction. The effects of the composition of the SiGe layer [Ge content, and minute C content] on the trap density and the device characteristics were investigated. Also examined were the effects of the oxidation temperature, the post-oxidation annealing [temperature and time] in nitrogen, and the post-metallization annealing [temperature and time] in forming gas, on the trap density and the device admittance characteristics.

1. INTRODUCTION

The SiGe quantum well MOS structures [back contact/Si substrate/epitaxial Si buffer layer/epitaxial SiGe quantum well layer/epitaxial Si cap layer/SiO$_2$ layer/gate metal] have received attention on account of their promise in realizing p-channel transistors with high hole mobility, relative immunity to mobility degradation by the Si-SiO$_2$ interface roughness and traps, and reduction in hot carrier injection into the gate oxide. One of the main problems, that has come in the way of progress in this area, concerns growth of a high quality gate oxide on substrates containing a strained SiGe quantum well layer. To avoid the relaxation of the strained SiGe layer by the generation of misfit dislocations and Ge diffusion, investigators have tended to carry out the gate oxide growth or deposition with low thermal budget, often, employing unconventional processes, such as wet oxidation or low-pressure or plasma-enhanced chemical vapor deposition [1-4]. Our approach to obtain a superior quality gate SiO$_2$ layer, without relaxing the strained SiGe layer, has been to employ the thinnest possible gate oxides, and to carry out the oxidation in dry O$_2$ at high temperatures for the shortest possible times. This has enabled us to grow ultrathin (4-7 nm), gate oxides with very low trap densities, without causing any relaxation of the strained SiGe quantum well layer.

2. EXPERIMENTAL

The MOS samples for this work were fabricated in the following manner. After surface preparation of the silicon substrate (B-doped) with 10-20 Ω.cm resistivity and (100) orientation, the Si buffer layer [ca. 30 nm thick], the SiGe or SiGeC well layer [ca. 30 nm thick with 21 % Ge], and the Si cap layer [ca. 15 nm thick] were grown by molecular beam epitaxy [MBE] or chemical vapor deposition [CVD] with no intentional doping, at 500 or 690 °C, respectively. The cap layer was oxidized thermally in dry oxygen at atmospheric pressure at 1100 °C for times between 30 and 65s, to obtain an oxide thickness, t$_{ox}$, in the range of 4-7 nm. Preceding oxidation, the wafers were cleaned ultrasonically in trichloroethylene, acetone, and methanol, and etched in HF, followed by decanting in ultrapure water with 18.2 MΩ.cm resistivity and total organics < 5 ppb. No post-oxidation annealing [POA] was carried out. Subsequently, Al front metal dots (0.1 cm diameter) and the Au back contact were deposited by filament evaporation in an ultra high vacuum, ion-pumped system [Varian VT-112], at pressures below 5x10^{-7} torr. Post-metallization annealing [PMA] was carried out at 400 °C in forming gas for 15 min. Figure 1 illustrates the physical structure of the fabricated SiGe quantum well MOS structure. Control MOS samples were fabricated

exactly the same way, except that the SiGe layer was replaced by a Si layer of the same thickness.

Gate Metal [Al]

Gate Oxide: 4-7 nm
---
i-Si Cap Layer: 14 nm originally

Undoped Si0.8Ge0.2 Quantum Well: 30 nm
---
i-Si Buffer: 30 nm
---
p-Si (100) 10-20 W.cm Substrate

Back Contact [Au}

Figure 1. Schematic of the fabricated SiGe quantum well MOS structure.

Comprehensive admittance-voltage-frequency characteristics were measured using the HP 4192A LF impedance analyzer, the HP 4042B quasistatic C-V and I-V meter, the Keithley 595 static C-V meter, the HP 310M controller, and a data acquisition and data analysis software package developed by us [5]. The Capacitance-voltage, C-V, and the conductance-voltage, G-V, characteristics were measured at different frequencies between 100 Hz and 3 MHz. The capacitance-frequency, C-f, and the conductance-frequency, G-f, characteristics were measured at different values of the bias. The quasi-static C-V characteristic was measured at a ramp rate of 0.001 V/s. The static C-V characteristic was measured with 0.01 V step and a step delay of 20 s. The current-voltage, I-V, characteristic was measured, using a step voltage of 0.01 V, and a step delay of 60 s.

3. RESULTS AND DISCUSSIONS

Figure 2 presents the capacitance-voltage [C-V] characteristics of the sample SiGe-44, with CVD-grown epitaxial layers and a 6.5 nm thick gate oxide, measured using the sinusoidal small signal of 100 Hz, 100 kHz, and 300 kHz, and the quasi-static C-V characteristics, measured with the ramp rate of 1 mV/s. The dispersion of the accumulation capacitance, C_{acc}, is due to a high series resistance. The value for this resistance, R_s, obtained directly from the measured 300 kHz impedance, [R_s + $j/\omega C_{ox}$], in strong accumulation [- 2.5 V], was 76 Ω, while

its value calculated from the C_{acc} dispersion was 92.5 Ω. The value of the spreading resistance, calculated using the relation r/2d, where r is the Si substrate resistivity [10-20 Ω.cm], and d the diameter of the gate metal dot, was 50-100 Ω. This indicates that the main contribution to R_s is from the spreading resistance. The most remarkable feature of the C-V characteristics of Fig. 2 is the near absence of dispersion around flat-band [at – 0.80 V], in depletion [between – 0.80 and – 0.47V], and in weak inversion[between – 0.47 and - 0.12V]. The well in the quasi-static C-V characteristic is deep because of the high Si resistivity and the ultrathin gate oxide. Such deep wells greatly enhance the reliability and the sensitivity of direct trap detection. This means that even a very low trap density would be easily visible as a dispersion between the quasi-static C-V and the 300 kHz C-V characteristics. The quasi-static C-V curve looks ideal in almost every respect, including the flat-band voltage of - 0.80 V.
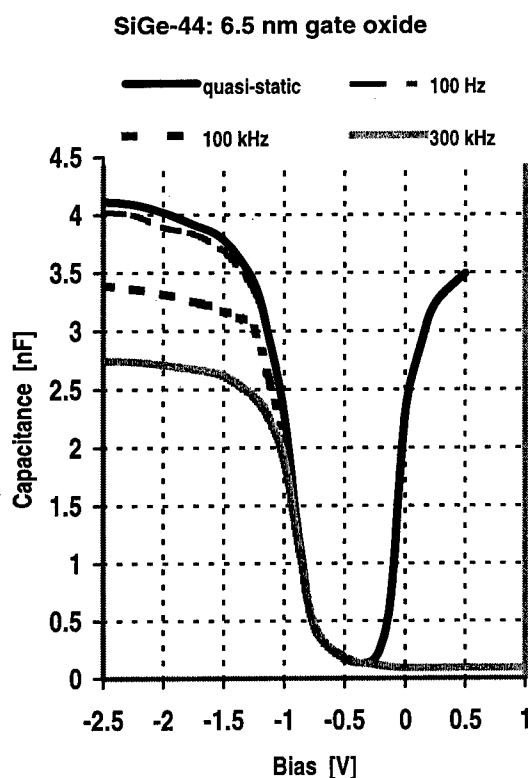
Figure 2. Measured capacitance-voltage characteristics of sample SiGe-44.

Figure 3 displays the conductance-voltage [G-V] characteristics of sample SiGe-44, measured at 30 kHz,

100 kHz, and 300 kHz. The entire measured conductance appears to be due to the series resistance, and the characteristic peaks generated by traps are conspicuously absent in these characteristics. This directly confirms the strong indications of Fig. 2 of a very low trap density.

Experimental values of G_p/ω were obtained from the measured MOS conductance-frequency and capacitance-frequency curves and the experimental values of C_{ox} and the series resistance R_s [5-6], for different bias values corresponding to depletion and weak inversion. No peak could be detected in any of these characteristics. The experimental points were random and scattered. The experimental values of G_p/ω were below 5.0 pF for all the bias values in depletion and weak inversion, which corresponds to a D_{it} value of about $8 \times 10^9/\text{cm}^2\text{V}$, for single level states.

Sample SiGe-44: 6.5 nm gate oxide
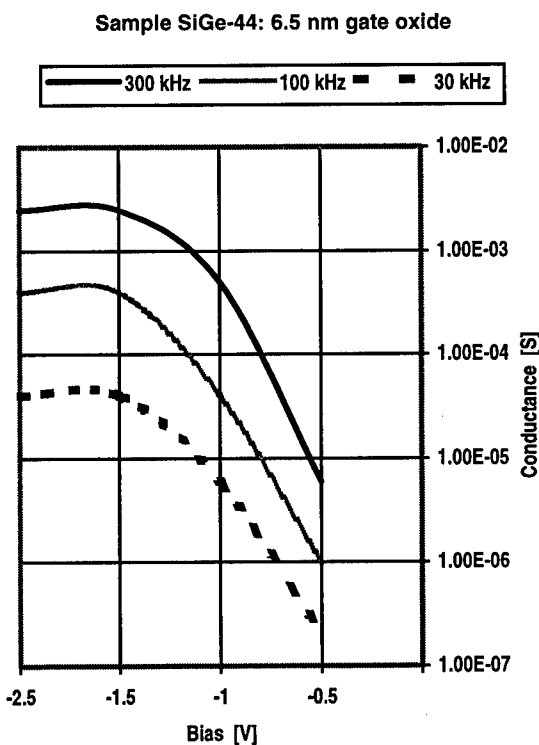
Figure 3. Measured conductance-voltage characteristics of sample SiGe-44.

The combined experimental data suggest nearly ideal and nearly trap-free admittance characteristics, except for the effect of the series resistance, which will be inevitably present in the case of high resistivity silicon bulk. There are a number of problems in estimating the trap density from these characteristics. Since, no peaks

are obtained in the G_p/ω versus f curves, a value for the trap density cannot be determined from the conductance data, except a limiting maximum value, as indicated above. Extraction of D_{it} from the capacitance data, using the high-low approach, presents the following problems: [1] The difference between the quasi-static [low-frequency] and the high frequency capacitance in depletion and weak inversion, cf. Fig. 2, is in the range of the measurement error. [2] The high frequency capacitance does not reflect the space charge capacitance, as it represents only the contribution from the dopants and the majority carriers, but not the minority carriers. This can cause serious errors, when the high and the low frequency capacitances are close to each other. It can be easily shown that, for p-type semiconductor, in depletion and weak inversion: $C_{sclf} = C_{schf}[1 + (n_s/N_A)]$, where C_{sclf} is the low frequency space charge capacitance, C_{schf} the high frequency space charge capacitance, n_s the electron density at the Si (cap layer) – SiO$_2$ interface, and N_A the acceptor density. This means that at the onset of strong inversion, $C_{sclf} = 2 C_{schf}$. Corrections can be made using the above relation, however, there is uncertainty in the experimental value of N_A, in the case of the substrate, containing the SiGe quantum well.

We have examined the effects of oxidation temperature [was varied between 800 and 1100 °C] and post-oxidation annealing in N_2 ambient [annealing temperature was varied between 900 and 1100 °C] on the trap density and the relaxation of the strained SiGe quantum well layer. The relaxation was monitored by x-ray diffraction. The lowest trap density was obtained for rapid thermal oxidation at 1100 °C, and there was no relaxation of the strained SiGe layer for oxidation times between 30 and 65 s. Post-metallization annealing [PMA] in the forming gas was very effective in obtaining a very low trap density, and for the samples, undergoing PMA, the post-oxidation annealing had no effect, and was not necessary. This is very useful in reducing the thermal budget, and in preventing the strained SiGe layer from relaxing. Table 1 illustrates the effects of the oxidation and the annealing conditions on the trap density. All the samples of Table 1 were prepared on the same substrate material with MBE epitaxial layers. In Table 1, the only sample, which was found to be completely unrelaxed, was the one oxidized at 1100 °C.

We have also examined the effects of the deposition conditions and the composition of the SiGe quantum layer on the interface trap density. Four different types of MBE-grown and two different types of CVD-grown SiGe layers were used: [1] MBE $Si_{1.0}Ge_{0.0}$; [2] MBE $Si_{.79}Ge_{.21}$; [3] MBE $Si_{.79}Ge_{.21}C_{.0015}$; [4] MBE $Si_{.79}Ge_{.21}C_{.0055}$;

[5] CVD $Si_{1.0}Ge_{0.0}$; [6] $Si_{.80}Ge_{.20}$. The mid-gap trap density was found to vary, in these samples, in the range of 1×10^{10} to 4×10^{10} $cm^{-2}V^{-1}$[cf. Table 2]. The incorporation of carbon into the SiGe layer did not appear to make any appreciable difference. There was no appreciable difference in the trap density, when the Ge content was nil. This indicates that the strain in the SiGe layer need not lead to generation of electron traps. The best admittance characteristics were obtained for the CVD-grown SiGe quantum well substrates, with a mid-gap trap density of less than 1×10^{10} $cm^{-2}V^{-1}$. The oxide leakage current density, in all the samples, for the voltage range of –2.5 V to 1.0 V was below $10^{-11}/cm^2$.

Table 1
Effects of oxidation/annealing conditions on trap density

Oxidation temperature, time	POA temperature, time	PMA temperature, time	Mid-gap trap density $cm^{-2}V^{-1}$
800 ºC, 149 min	---	390 ºC, 30 min	2.5×10^{11}
850 ºC, 69 min	---	---	8.0×10^{11}
850 ºC, 69 min	---	390 ºC, 30 min	1.1×10^{11}
900 ºC, 38 min	---	390 ºC, 30 min	1.4×10^{11}
950 ºC, 19 min	---	390 ºC, 30 min	1.2×10^{11}
1000 ºC, 325 s	---	---	3.0×10^{11}
1000 ºC, 7 min	1000 ºC, 16 min	---	4.0×10^{11}
1100 ºC, 50 s	---	400 ºC, 15 min	3.0×10^{10}

Table 2
Effects of the epitaxy conditions on the trap density

Epitaxy	Quantum well Composition	Mid-gap trap density $cm^{-2}V^{-1}$
MBE	Si	3.0×10^{10}
MBE	$Si_{.79}Ge_{.21}C_{.0015}$	4.0×10^{10}
MBE	$Si_{.79}Ge_{.21}C_{.0055}$	3.2×10^{10}
CVD	Si	2.6×10^{10}
CVD	$Si_{.80}Ge_{.20}$	8.0×10^{9}

4. CONCLUSIONS

We have demonstrated the growth of ultrathin [4-7 nm] gate oxides, with very low trap densities, on SiGe quantum well substrates, using a very simple substrate surface processing and a low thermal budget oxidation, that does not relax the strained SiGe layer. The simple surface preparation consisted of ultrasonic cleaning in organic solvents, followed by etching in HF, and finally rinsing in ultrapure water. Best oxidation conditions were obtained for rapid thermal oxidation at 1100 ºC for 30-65 s, under which lowest trap density and no relaxation of the strained SiGe layer were observed. Post-metallization annealing in forming gas at 400 ºC was very effective in reducing the trap density, obviating the need for the post-oxidation annealing in nitrogen at high temperatures with the attendant high thermal budget. The Ge content or the C content did not appear to have a significant effect on the trap density. The experimental results indicated that the strain in the SiGe quantum well layer need not lead to generation of electron traps. The lowest mid-gap trap density of 8×10^9 $cm^{-2}V^{-1}$ and the most ideal device characteristics were obtained for the CVD-grown SiGe layers.

REFERENCES

[1] D. K. Nayak, J. C. S. Woo, J. S. Park, K. -L. Wang, and K. P. MacWilliams, IEEE Electron Device Lett. EDL-12, 154(1991).

[2] S. S. Iyer, P. M. Solomon, V. P. Kesan, A. A. Bright, J. L. Freeouf, T. N. Nguyen, A. C. Warren, IEEE Electron Device Lett. EDL-12, 246(1991).

[3] S. P. Voinigescu, K. Iniewski, R. Lisak, C. A. T. Salama, J. -P. Noel, and D. C. Houghton, Solid-St. Electron. 37, 1491(1994).

[4] P. M. Garone, V. Venkataraman, and J. C. Sturm, IEEE Electron Device Lett. EDL-12, 230(1991).

[5] S. Kar and S. Ashok, J. Appl. Phys. 73, 2187(1993).

[6] E. H. Nicollian and J. R. Brews, MOS (Metal-Oxide-Semiconductor) Physics and Technology (Wiley, New York, 1982).

THEORY AND MODELING

Network transformation processes during oxidation of silicon

Alfredo Pasquarello

Institut Romand de Recherche Numérique en Physique des Matériaux (IRRMA),
Ecole Polytechnique Fédérale de Lausanne (EPFL),
PPH-Ecublens, CH-1015 Lausanne, Switzerland

Using density functional calculations in the generalized gradient approximation, the energetics of competing atomic processes occuring during the oxidation of silicon are evaluated. Simple molecular systems are used to model the breaking of Si-Si and Si-O bonds in the oxide. These calculations suggest that the breaking of Si-Si bonds and the formation of threefold coordinated O atoms are the most favourable transformation pathways of the bonding network of the oxide, in accord with the atomic processes observed during a recent first-principle molecular dynamics simulation.

1. Introduction

Despite several decades of intensive experimental investigations, our understanding of the silicon oxidation process has remained limited. The currently accepted picture of the oxidation process is to a large extent based on the seminal work by Deal and Grove [1]. Their model envisions a two step process in which an oxygen species diffuses through the oxide network and then attacks the silicon substrate leading to the formation of new oxide at the buried interface [1]. Although the Deal-Grove model successfully accounts for specific aspects of the oxidation process such as its kinetics, a characterization of the oxidation mechanism at the atomic scale falls well beyond its scope.

The process leading to a disordered silica network is necessarily more involved than just the incorporation of oxygen atoms in Si-Si bonds of the silicon substrate. Such incorporations produce locally high pressures and expansive motions. The structural rearrangements which follow cause bond breakings, giving ultimately rise to the amorphous nature of the oxide.

Such detailed atomic-scale behaviour has so far remained inaccessible to experimental techniques. In a nuclear reaction analysis experiment, Rochet et al. [2] provided microscopic evidence of the Deal-Grove mechanism in the thick oxide regime by oxidizing sequentially with different oxygen isotopes. In the thin film regime, deviations from Deal-Grove behavior were observed. Recent experiments carried out using medium energy ion scattering confirmed the presence of Deal-Grove oxidation, but also demonstrated the occurence of other processes which are missing from the Deal-Grove model [3,4]. In particular, these measurements show that oxygen atoms incorporated into the silica network move during subsequent oxidation and incorporate into the newly formed oxide [3,4].

Recently, in an attempt to gain insight into the fundamental structural transformation processes occuring during oxidation at the Si-SiO$_2$ interface, a first-principle molecular dynamics study was performed [5]. This investigation highlighted the role played by threefold coordinated oxygen atoms during structural rearrangements. It was found that changes in the bonding network near the interface occurred via an associative exchange mechanism wherein an oxygen atom is momentarily bonded to three silicon atoms. Such events survived several Si-O bond oscillations in the simulation, indicating that this structure is most likely metastable. However, because of the simultaneous occurrence of several atomic processes, it was impossible to extract excitation energies directly from the simulation.

It is the aim of the present contribution, to es-

timate excitation energies of competing atomic
processes occuring during oxidation at Si-SiO$_2$ in-
terfaces. Deferring the study of oxygen transport
through the film and oxygen incorporation in the
network, we here focus on those processes which
lead to a structural rearrangement of the bond-
ing network of the oxide. We consider three type
of processes: the breaking of a Si-Si bond, the
breaking of a Si-O bond, and the formation of a
threefold coordinated oxygen atom.

2. Method

Throughout this work we adopted the gener-
alized gradient approximation (GGA) to density
functional theory [6], which is here preferred over
the local density approximation (LDA). The LDA
gives accurate results for many materials includ-
ing low-pressure phases of SiO$_2$, but it has re-
cently been shown that it fails to reproduce the
correct energy difference between α-quartz and
stishovite, a high-pressure phase of SiO$_2$ [7]. This
shortcoming of the LDA is overcome by the GGA
[7]. The use of the GGA is expected to be par-
ticularly relevant when the atoms undergo coor-
dination changes.

The electronic structure was accounted for by
using a plane-wave basis set for the valence wave
functions and pseudopotentials to describe core-
valence interactions. We used a normconserving
pseudopotential for Si [8] and an ultrasoft pseu-
dopotential for O [9]. The latter pseudopoten-
tial was generated with cutoff radii of 1.3 bohr
and with two energy references in both s and p
angular momentum waves. The GGA exchange-
correlation energy is implemented as in Ref. [10].
Structural relaxations were performed using the
Car-Parrinello method [11,12]. By adopting a
damped molecular dynamics evolution for the
electronic degrees of freedom, the energy was
gradually lowered until the structure reached a
stable configuration. A detailed description of the
method used to treat the electronic structure in
the presence of ultrasoft pseudopotentials is given
in Ref. [13]. Energy cutoffs of 24 Ry and 150 Ry
were used for the wave functions and for the aug-
mented electron density, respectively. The calcu-
lations for the molecular systems were performed
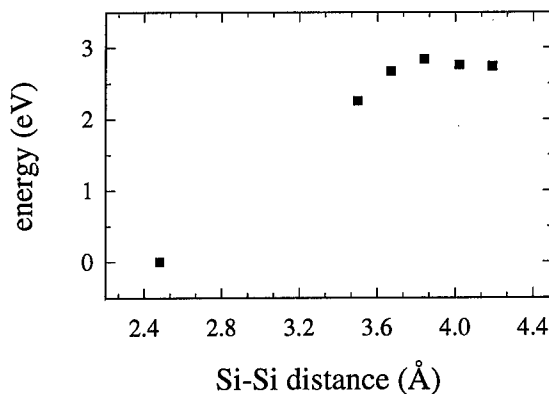
Figure 1. Energy as a function of Si-Si distance
for the neutral vacancy model in α-quartz, calcu-
lated in the generalized gradient approximation.
The energy is given with respect to that of the Si-
Si dimer configuration, which corresponds to the
lowest Si-Si distance. The puckered configuration
is obtained in correspondence of the largest Si-Si
distance. From Ref. [17]

using a periodic cubic cell with a side of 20 bohr.

3. Threefold coordinated oxygen atom

We take the E_1' defect in α-quartz as a model
system to study the threefold coordinated oxy-
gen atom. The E_1' defect has successfully been
modelled as a positively charged oxygen vacancy
[14,15]. This defect shows a bistable behavior be-
tween two competing configurations. In the first
metastable configuration, the elimination of an
oxygen atom from the regular α-quartz structure
gives rise to the formation of a Si-Si dimer bond.
In the second configuration, known as the puck-
ered configuration, the atomic network undergoes
a large local relaxation which leads to the for-
mation of a threefold coordinated oxygen atom
and of a Si dangling bond. In the positive charge
state, the puckered configuration is more stable
than the dimer configuration, explaining the ob-
servation of an electron spin resonance signal [16].
However, in the neutral charge state the dimer
configuration becomes the ground state [15].
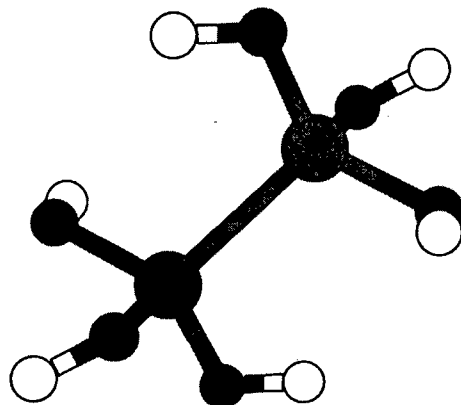
Figure 2. Ball and stick model of the molecule $(HO)_3Si$-$Si(OH)_3$.

Figure 3. Ball and stick model of fragment $(HO)_3Si\cdot$.

The energy difference between puckered and dimer configuration in the neutral charge state provides an estimate for the formation energy of a threefold coordinated oxygen atom and a dangling bond at the cost of breaking a Si-Si bond. This energy difference was calculated to be 2.3 eV in the LDA [15]. When this calculation was repeated in the GGA, a value of 2.7 eV was found [17]. The barrier which needs to be overcome to move from the dimer to the puckered configuration is only slightly higher (2.8 eV [17]). Fig. 1 shows the energy as a function of the Si-Si distance. These energy values are directly comparable with the other energies calculated in this work.

Note that the formation of a threefold coordinated oxygen atom is associated with the generation of local charges. The threefold coordinated oxygen atom gives up the sixth oxygen electron to the conduction band [18]. In the puckered configuration this electron is captured by the dangling bond, which becomes doubly occupied giving rise to a negative center. Because of the spin degeneracy of all the orbitals in this system, a local spin density calculation does not modify the total energies calculated above.

4. Bond breakings

4.1. Breaking of a Si-Si bond

In order to obtain an estimate of the energetic cost of a Si-Si bond-breaking process in the oxide, we consider the following molecular process:

$$(HO)_3Si\text{-}Si(OH)_3 \rightarrow 2\ (HO)_3Si\cdot, \qquad (1)$$
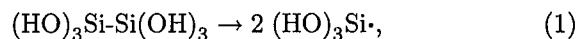
where the dot, $\cdot$, stands for a singly occupied dangling bond. The energy required for this process is obtained considering the total energies of the molecule $(HO)_3Si$-$Si(OH)_3$ and of its fragment $(HO)_3Si\cdot$ separately. In the two calculations the atomic positions are allowed to relax fully. The resulting atomic structures are illustrated in Figs. 2 and 3. The presence of singly occupied dangling bonds on the righthand side of process (1) requires the use of a spin density functional for $(HO)_3Si\cdot$. We estimate the energetic cost of process (1) to be 3.8 eV. The spin polarization energy gain associated to the Si dangling bond is found to be 0.63 eV. Hence, the consideration of spin effects is important, and their neglect would increase the energetic cost of process (1) to 5.4 eV.

These values can be compared to the bond formation energy of bulk silicon. Because two bonds can be associated with every Si atom, the bond formation energy can be obtained by dividing by two the binding energy per atom. The calculated
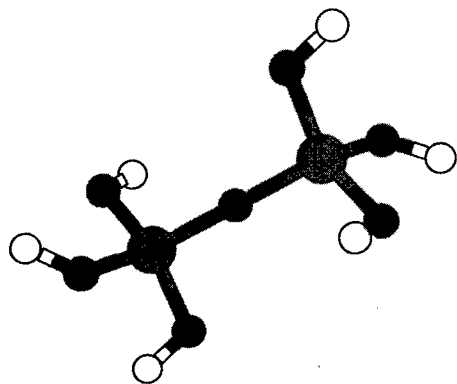
Figure 4. Ball and stick model of the molecule $(HO)_3Si-O-Si(OH)_3$.

Figure 5. Ball and stick model of fragment $(HO)_3Si-O\cdot$.

binding energy for bulk silicon is 4.65 eV [8], in excellent agreement with the experimental value of 4.63 eV [19]. We thus obtain 2.3 eV for the bond formation in bulk silicon. It is not straightforward to deduce from this the bond-breaking cost of an isolated Si-Si bond in a silicon-rich environment. Nevertheless, it appears reasonable to assume that the breaking of a Si-Si bond occurs at a lower energy cost the closer the bond is to the Si substrate.

4.2. Breaking of a Si-O bond

To obtain the bond-breaking cost of a Si-O bond we consider the process:

$$(HO)_3Si-O-Si(OH)_3 \rightarrow$$
$$(HO)_3Si-O\cdot + \cdot Si(OH)_3. \quad (2)$$

In analogy to the case of a Si-Si bond breaking, we compare the total energy of the molecule $(HO)_3Si-O-Si(OH)_3$ with those of its fragments $(HO)_3Si-O\cdot$ and $\cdot Si(OH)_3$. For simplicity, the Si-O-Si bond was kept fixed at 180° in the $(HO)_3Si-O-Si(OH)_3$ molecule, but the atomic coordinated were otherwise completely relaxed. The relaxed atomic configurations of the molecule $(HO)_3Si-O-Si(OH)_3$ and of the fragment $(HO)_3Si-O\cdot$ are shown in Figs. 4 and 5, respectively. We obtain an energy cost of 5.9 eV, in which the spin polarization energy is accounted

for. We found that the spin polarization energy amounted to 0.56 eV for the dangling bond on the oxygen atom, close to the value of 0.63 eV, found above for a dangling bond on a silicon atom.

5. Discussion

We consider excitation energies with respect to a regularly bonded interface configuration, i.e. a structure in which all the Si and O atoms are four-fold and twofold coordinated, respectively. We assume that such a configuration is metastable, and that the result of a network transformation process is another regularly bonded structure. We consider relaxation processes which lead to a topological change of the bonding network, or, in other terms, transformations which modify the list of neighbours. We also assume that the energy required to transform the bonding arrangement is much larger than the energy gain resulting from the relaxation process. We base these assumptions on observations made in previous studies, in which structural relaxations of several regularly bonded structures never led to topological rearrangements [20–24]. Furthermore, in the molecular dynamics simulation of Ref. [5], high initial temperatures were needed to break up a regularly bonded structure.

We considered here three processes which break

up the oxide network: the breaking of a Si-Si bond, the breaking of a Si-O bond, and the formation of a threefold coordinated oxygen atom. These processes and their corresponding reverse processes can be combined to obtain topological modifications of the network.

From our calculations, it appears that the breaking of a Si-Si bond can compete with the formation of a threefold coordinated oxygen atom, particularly in the neighborhood of the substrate. In fact, the process that leads to the formation of a threefold coordinated oxygen atom can also be viewed as the breaking of a Si-Si bond. The ensuing formation of a threefold coordinated oxygen atom accompanied by a doubly occupied Si dangling bond can be considered as a stabilizing process, which lowers the cost of a Si-Si bond-breaking event in the oxide from 3.8 eV to 2.7 eV. However, it is important to note that Si-Si bond breakings and bond formations alone cannot give rise to modifications of the oxide network. In fact, every oxygen atom would maintain the same neighbours during such processes. Hence, a modification of the network necessarily implies the breaking of Si-O bonds.

The appearance of transient states in which oxygen atoms are threefold coordinated, allows structural transformations to different regularly bonded structures. In fact, because any of the three O-Si bonds can break during a transient state, such a process could lead to an exchange event in the oxygen first-nearest-neighbour shell [5]. In this way, a topological modification of the oxide network can take place at excitation energies (2.7 eV) which are much lower than those required for breaking a Si-O bond in the case of regularly twofold coordinated oxygen atoms (5.9 eV). In other terms, our calculations suggest that the required Si-O bond-breakings would occur when the involved oxygen atoms are threefold rather than twofold coordinated.

From the above considerations, a picture of transient configurations emerges in which silicon dangling bonds can either be singly or doubly occupied, depending on whether they result from a Si-Si bond breaking or from the formation of a threefold coordinated oxygen atom, respectively. In such a situation, the flow of electrons between dangling bonds could play a role in network transformation processes. As the incorporation of oxygen atoms proceeds, the local structural environment undergoes changes. As a consequence, the flow of an electron from one dangling bond to another could become energetically favourable. Such processes would change the location of doubly occupied dangling bonds. In the reverse of the process leading to the formation of a threefold oxygen atom, a Si-Si bond is formed between a Si atom carrying a doubly occupied dangling bond and one of the silicon atoms in the first-nearest-neighbour shell of the threefold coordinated oxygen atom. When such a reverse process takes place after the relocation of a doubly occupied dangling bond, the topology of the network is transformed.

This picture of network transformation processes occuring during oxidation is consistent with the processes observed during a recent first-principle molecular dynamics simulation [5]. In this simulation, Si-O bonds were found to break only in correspondence of threefold coordinated oxygen atoms, in agreement with the energetic picture discussed here. Furthermore, frequent Si-Si bond breakings were observed in the neighborhood of the silicon substrate.

Further support to this picture is provided by comparison with experiment. The formation of threefold coordinated oxygen atoms is accompanied by large lattice distortions [15,17], and can lead to motion of oxygen atoms which are incorporated in the network, as clearly illustrated in the molecular dynamics simulation of Ref. [5]. The presence of network motion is consistent with recent medium energy ion scattering results which show evidence for motion of oxygen atoms after incorporation into the oxide [3,4].

In conclusion, we estimated the cost of simple atomic processes leading to modifications of the network structure. The events considered here are relatively local and do not represent an exaustive description of all the atomic processes occuring during oxidation. Nevertheless, the picture emerging from the present analysis is consistent with the results of a recent first-principle molecular dynamics simulation [5].

Acknowledgments

REFERENCES

1. B.E. Deal and A.S. Grove, *Journal of Applied Physics*, **36**, 3770–3778 (1965).
2. F. Rochet, S. Rigo, M. Froment, C. d'Anterroches, C. Maillot, H. Roulet, and G. Dufour, *Advances in Physics*, **35**, 237–274 (1986).
3. E.P. Gusev, H.C. Lu, T. Gustafsson, and E. Garfunkel, *Phys. Rev. B*, **52**, 1759–1775 (1995).
4. E.P. Gusev, H.C. Lu, T. Gustafsson, and E. Garfunkel, *"The Physics and Chemistry of SiO_2 and the Si-SiO_2 Interface - 3"*, edited by H.Z. Massoud, E.H. Poindexter and C.R. Helms (The Electrochemical Society, Pennington, 1996), Volume 96-1, p. 49–58.
5. A. Pasquarello, M.S. Hybertsen, and R. Car, *Nature*, **396**, 58 (1998).
6. J.P. Perdew and Y. Wang (unpublished); J.P. Perdew, in *Electronic Structure of Solids '91*, edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991); J.P. Perdew *et al.*, *Phys. Rev. B.* **46**, 6671 (1992).
7. D.R. Hamann, *Physical Review Letters* **76**, 660 (1996).
8. A. Dal Corso, A. Pasquarello, A. Baldereschi, and R. Car, *Phys. Rev. B* **53**, 1180 (1996).
9. D. Vanderbilt, *Phys. Rev. B* **41**, 7892 (1990).
10. J.A. White and D.M. Bird, *Phys. Rev. B* **50**, 4954 (1994).
11. R. Car and M. Parrinello, *Physcal Review Letters* **55**, 2471 (1985).
12. For a review see G. Galli and A. Pasquarello, in *"Computer Simulation in Chemical Physics"*, edited by M.P. Allen and D.J. Tildesley (Kluwer, 1993) p. 261.
13. A. Pasquarello, K. Laasonen, R. Car, C. Lee, and D. Vanderbilt, *Physcal Review Letters*

69, 1982 (1992); K. Laasonen, A. Pasquarello, R. Car, C. Lee, and D. Vanderbilt, *Physcal Review B* **47**, 10142 (1993).
14. J.K. Rudra and W.B. Fowler, *Phys. Rev. B* **45**, 8223 (1987).
15. M. Boero, A. Pasquarello, J. Sarnthein and R. Car, *Physical Review Letters* **78**, 887 (1997).
16. R.H. Silsbee, *Journal of Applied Physics* **32**, 1459 (1961).
17. A. Pasquarello, submitted to *Applied Surface Science*.
18. P. Deák, in *Early Stages of Oxygen Precipitation in Silicon*, edited by R. Jones (Kluwer, Dordrecht, 1996), p. 163–177.
19. *Physics of Group IV Elements and III-V Compounds*, edited by O. Madelung, Landolt-Börnstein, New Series, Group 3, Vol. 17, Pt. a (springer, Berlin, 1982).
20. A. Pasquarello, M. S. Hybertsen and R. Car, *Physical Review Letters* **74**, 1024 (1995).
21. A. Pasquarello, M.S. Hybertsen and R. Car, *Applied Surface Science* **104/105**, 317 (1996).
22. A. Pasquarello, M.S. Hybertsen and R. Car, *Applied Physics Letters* **68**, 625 (1996).
23. A. Pasquarello, M.S. Hybertsen and R. Car, *Physical Review B* **53**, 10942 (1996).
24. K.-O. Ng and D. Vanderbilt, *Physical Review B* **59**, 10132 (1999).

Polar Phonon Scattering at the Si-SiO₂ Interface

Jian-Qiang Lü[a] and Frederick Koch[b]

[a] Department of Electrical, Computer and System Engineering, Rensselaer Polytechnic Institute,
Troy, NY 12180, USA

[b] Physik Department E 16, Technische Universität München, 85747 Garching, Germany

In studying fine structures in the tunneling current characteristics of degraded bipolar transistors we have observed under small forward bias a unique spectral pattern of current variations at low temperature. The structures are step-like increases of the base-emitter current that repeat in well-defined voltage sequences. Among these, replicas with a spacing of 135 mV, which is equivalent to the energy of a polar optical phonon in SiO₂, are identified quite often. We believe this is the first evidence of the polar phonon scattering for transport at the Si-SiO₂ interface.

1. INTRODUCTION

Carriers moving in high electric fields of device structures along the Si-SiO₂ interface dissipate energy by phonon emission processes. Such phonon scattering limits the mobility. Velocity saturation occurs when the emission rate is just matched by the rate of energy gain in the electric field. Although this process is usually discussed in terms of interaction with the Si-phonons, it remains an open question to what extent the polar optical modes of the adjacent SiO₂ layer contribute by their long-range fields to the scattering and energy dissipation. Hess and Vogl considered this problem of electron scattering by what they called remote phonons [1]. To date there has been no convincing evidence that the SiO₂ phonons contribute to the energy loss of surface channel electrons in MOS devices, possibly because of screening and the short range of the interaction.

We examine here tunneling transport in which electrons move from the n⁺-side of a junction to the corresponding p-side, presumable via an interfacial defect that provides an energy state in the gap of Si. The p-n junction is that of the base-emitter structure of the planar bipolar transistor. It contacts the (100) Si-surface at the spacer oxide (see Fig. 1). As we shall see below, an important contribution to the forward base current involves carriers that move, assisted by traps, along the interface and just inside the Si-crystal. For this current there is found a complicated fingerprint pattern of phonon structures. Each transistor, and each of the interface configuration of traps created by hot carrier damage, has its own, unique spectral pattern in the base-emitter current. Among these one frequently finds phonon structures with energy of ~20 and ~60 meV, which probably relate to Si phonons. In addition, we identify at 135 meV a phonon replica structure that just cannot come
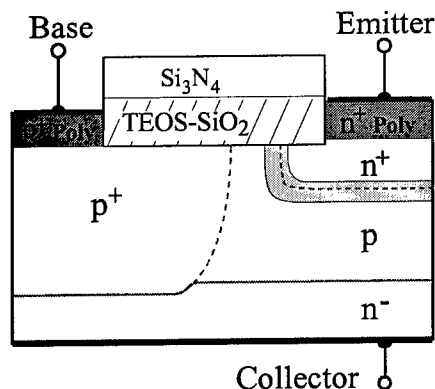
Fig. 1. Cross sectional view of the planar transistor structure in which tunneling occurs along the Si-TEOS SiO$_2$ interface.

Fig. 2. Gummel plots showing the degradation of the base current with increasing stress time, stressed at reverse bias of 3.5 V and room temperature.

from Si. We argue that tunneling which proceeds via an interfacial trap level, a deep-lying defect state, will have replica structures, which dissipate a polar phonon energy in SiO$_2$. The tunneling path, although near the interface, must lie on the Si side of the interface. In this sense, the phonon is generated as a remote excitation of the SiO$_2$ layer.

Below we briefly discuss the planar transistor and the microscopic origin of the degradation effect. This is followed by a more detailed consideration of the phonon-assisted tunneling though the trap states.

2. DEGRADATION OF THE BASE-EMITTER JUNCTION AT THE Si-SiO$_2$ INTERFACE

Fig. 1 shows the physical layout of the relevant part of the transistor in cross-section. The Siemens-made device is a self-aligned, double-poly diffused bipolar transistor [2]. The spacer oxide that contacts the base-emitter junction and separates the two electrodes is a TEOS-CVD layer followed by a nitride cap. In the publica-

tions [2, 3, 4], we have studied the degradation of the device under stress conditions where 3.5 eV is applied for a length of time in a reverse bias to the junction. Numerical simulations showed that under such stress condition there is a region of high electric field between the n$^+$-emitter and the p-base under the TEOS-SiO$_2$ [2]. After the bias stress treatment, the base current in the 0 - 1 V range is several orders of magnitude higher in the degraded state, as can be seen in Fig. 2. The microscopic mechanism, by which the base current is caused to increase, involves the creation of defect states at the interface to the spacer oxide. The defects are damage centers such as dangling bonds or others [5] at the interface and along the entire perimeter line where the p-n junction contacts the TEOS oxide. In both our own and

other related research [6, 7, 8, 9] the point defects are considered to be the source of the excess current under low voltages in forward bias. Whereas [8, 9] treat this extra contribution as a generation recombination current, our study of the temperature dependence as well as others' works [6, 7] identify it as a tunneling current. As a confirmation, see Fig. 3, the slope of the base current at low voltages does not change with the temperature.

3. PHONON REPLICA IN THE BASE-EMITTER CURRENT

With voltages in the 0 - 1 V range in the forward direction the base current rises exponentially. As in Figs. 2 and 3, at room temperatures the curves are almost featureless. The change in slope from low voltages to high voltages indicates that the dominant mechanism for transport switches from tunneling to a more strongly temperature-dependent mechanism. As the temperature is lowered in steps from 300 to 10 K the tunneling regime extends over the entire voltage range of the figure.

It is equally evident that the base current with lowered temperatures shows distinct signatures. For the sub-μm-sized transistor structure with which we are dealing, the base current shows a fingerprint pattern of voltage-tuned steps and variations. Fig. 3 shows that, as the temperature is reduced to 10 K, the I-V curve reveals definitely some fine structure and has three bending points roughly at 0.65 V, 0.78 V, and 0.92 V. The voltage difference between the bending points is $\Delta V \approx$ 130-140 mV. These are totally reproducible in an up-down sweep for the given transistor structure and given state of degradation. They are also identifiable even up to 150 K.

As discussed in Section 2, the increase
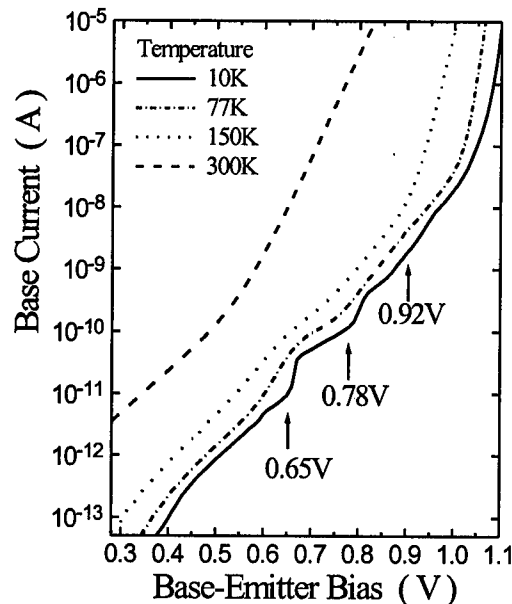
Fig. 3. Diode Characteristic of the base-emitter junction of a degradation damaged planar transistor showing SiO_2-phonon replicas at low temperatures.

current is the result of trap-assisted tunneling in the Si layer adjacent to the Si-SiO_2 interface. Because of the small size of the device, individual prominent trap sites that are particularly efficient in the contribution to the base current, can be identified [4]. The effect is totally analogous to the random patterns in the low temperature current of MOS transistors [10, 11]. Just as the case in MOSFETs, there is at finite temperature a switching type of noise, the random-telegraph signal reported in refs [3, 4].

The tunneling that leads to distinct step-like increases in the base current is a resonant process, whereby the carrier at the band-edge of the p-n junction finds a trap positioned at a suitable location in the space-charge layer and at the interface. If the energy matches up properly, resonant

tunneling occurs. The line shape of the structure reflects the availability of traps.

The step-like increases of the current repeat in well-defined voltage sequences. Among these replicas with a spacing of 20 and 60 mV are identified quite often. Others have a voltage interval of 100 and 135 mV. We believe that current steps like those in Fig. 3 are the result of phonon assisted tunneling involving intermediate mid-gap states near the Si-SiO$_2$ interface.

While the 20 and 60 mV replica structures are likely to involve Si-phonons that allow a tunneling electron to lose the transverse momentum and energy [12], the 100 and 135 mV replica only can be expected from the interfacial SiO$_2$ phonons. In particular, 135 meV is the energy of the polar stretching mode vibrations of SiO$_2$, and 100 meV.

The electrons contributing to the tunneling current move on the Si side of the interface. Nevertheless, because it is a resonant tunneling process a defect state with energy in the gap of Si is involved. This trap is a highly localized state and may well be located at the interface in the immediate vicinity of SiO$_2$. An energy transfer to the polar oxide material can occur while the carrier is in the trap. This is indeed a very different scenario than that involving ordinary channel transport as in ref. [1]. Our experiments do not provide evidence for the remote polar phonon scattering described there. Even if the energy loss were to occur along the tunneling path and inside the Si, the observation is not directly applicable to MOS channel transport, because it occurs in the depleted region and in the absence of screening effects. What the experiments show is that for transport at the Si/SiO$_2$ interface the polar SiO$_2$ phonons can be involved in special circumstances. The challenge to experimentally find and measure the SiO$_2$-phonon contributions for

hot carriers in an inversion larger is still to be met.

4. CONCLUSION

We study tunneling transport at the degradation damaged Si-SiO$_2$ interface of a planar bipolar transistor. In this transistor the base-emitter planar p-n junction contacts the Si surface where a CVD-spacer oxide is located. Under conditions where tunneling through trap states at the interface dominates the forward-bias current, we identify numerous phonon replica structures that are evidence for the interaction with the polar phonons of the adjacent oxide.

REFERENCES:

1. K. Hess and P. Vogl, Solid State Commun., 30 (1979) 807.
2. J.-Q. Lü, S. Schottl, E. Stefanov, F. Koch, R. Mahnkopf, and H. Klose, in Defect- and Impurity- Engineered Semiconductors and Devices, edited by Ashok, S.; Chevallier, J.; Akasaki, I.; Johnson, N.M.; Sopori, B.L., (1995) 743.
3. J.-Q. Lü, S. Schottl, F. Koch, R. Mahnkopf, and H. Klose, in Extended Abstracts of the 1995 International Conference on Solid State Devices and Materials, Osaka, (1995) 237.
4. J.-Q. Lü and F. Koch, Jpn. J. Appl. Phys., 35, No 2B (1996) 826.
5. Jian-Qiang Lü, Ph.D. Thesis, Technical University Munich, 1995.
6. G. P. Li, E. Hackbarth and. T. C. Chen, IEEE Trans. Electron Devices 35 (1988) 89.
7. G.A.M. Hurkx, D.B.M. Klaassen and M.P.G. Knuvers, IEEE Trans. Electron Devices, 39 (1992) 331.
8. J. D. Burnett and C. Hu, IEEE Trans. Electron Devices, 35 (1988) 2238.

9. A. Chantre, G. Festes, G. Girout-Matlakowski and A. Nouailhat, IEEE Trans. Electron Devices, 38 (1991) 1354

10. M. Bollu, Ph.D. Thesis, Technical University Munich, 1989.

11. M.J. Kirton and M.J. Uren, Adv. Phys. 38 (1989) 367.

12. A.G. Chynoweth, W.L. Feldmann and R.A. Logan, Phys. Rev. 121 (1961) 684.

Simulation of the Drain-Current Characteristics of MOSFETs with Ultrathin Oxides in the Presence of Direct Tunneling*

J. P. Shiely and H. Z. Massoud

Semiconductor Research Laboratory, Department of Electrical and Computer Engineering,
Duke University, Durham, North Carolina 27708–0291, U.S.A.

Carrier tunneling in the gate dielectric, especially in the direct-tunneling regime where large current densities flow through the gate oxide, are known to result in substantial changes in the drain-current characteristics of MOSFETs. In this paper, we present simulation results of the drain-current characteristics of MOSFETs with ultrathin oxide using Tunnel-PISCES, a MOSFET tunneling simulator that models electron tunneling through the gate dielectric in a self-consistent manner with carrier transport by drift and diffusion in the substrate. We are able to predict the experimental trends reported for the dependence of the drain current of ultrathin-oxide MOSFETs on gate-oxide thickness. This tunneling simulation capability provides a means for generating MOSFET sizing guidelines to avoid tunneling-induced drain-current degradation.

1. INTRODUCTION

By scaling a MOSFET gate oxide to thicknesses where direct tunneling is present and selecting an appropriate gate length, MOSFETs with large transconductances have been realized [1]. In direct-tunneling oxide devices, the probability that an electron traveling from the source to the drain will tunnel through the gate dielectric may be greater, under certain conditions, than its probability of reaching the drain. Therefore, a MOSFET with an ultrathin gate oxide must be designed to include the effects of the direct-tunneling gate current upon the flow of current from source to drain. This paper reports for the first time simulation results of MOSFET drain-current characteristics in the presence of direct tunneling. These simulation results are shown to correlate well with experimentally observed measurements [1].

2. TUNNELING MODEL

Simulation of direct tunneling in ultrathin oxide MOSFETs was carried out using Tunnel–PISCES [2], a modified version of the classical device simulator PISCES [3] developed to model self-consistently MOS devices in the presence of carrier tunneling in the gate dielectric. The gate tunneling current was calculated using the independent-electron tunneling model [4] and the WKB approximation was used to determine the tunneling probability. The tunneling barrier was assumed to be trapezoidal. Because Tunnel–PISCES is a modified drift-diffusion device simulator, and not a Schrödinger-Poisson solver, the effects of carrier confinement and quantization near the oxide/substrate interface upon the inversion layer [5] were not taken into account.

The tunneling current was calculated by taking into account tunneling from the substrate into the gate and from the gate into the substrate. The net tunneling current was then self-consistently matched to the net drift-diffusion current in the substrate at the oxide/substrate interface. The tunneling mechanisms present in NMOSFETs are illustrated in Fig. 1. Electrons can tunnel from

the conduction (CBET) and valence (VBET) bands and holes can tunnel from/to the valence band (VBHT).

Tunneling Mechanisms in NMOSFETs

n^+ Poly Gate SiO_2 p-type Substrate
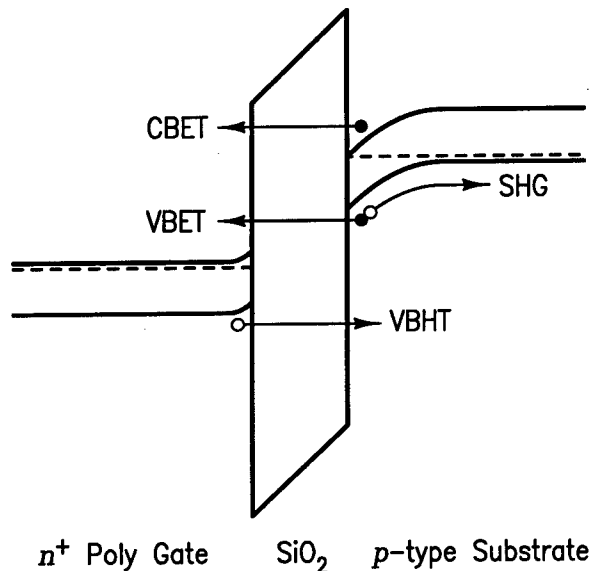
Fig. 1. Electron (CBET and VBET) and hole (VBHT) tunneling mechanisms in NMOSFETs.

3. SIMULATION RESULTS

We have simulated three MOSFET structures with different gate lengths reported by Momose et al. [1]. The gate oxide thickness for all devices was 15Å. The junction depth of the lightly doped regions of the drain and source was 30 nm. The source and drain were n-type Gaussian diffusions with peak concentrations of 4×10^{19} cm^{-3}. The substrate was doped p-type, 10^{18} cm^{-3}, and the gate was Aluminum. The gate currents were modeled using a tunneling barrier height of 3.15 eV, and a constant effective mass of $0.35m_0$ for the electron as it tunnels through the oxide. A schematic cross-section of the device is shown in Fig. 2, and the device simulation cross-section generated by Tunnel-PISCES is shown in Fig. 3.

Figures 4–6 show Tunnel-PISCES simulation results of the drain-current characteristics and Figs. 7–9 the gate- and source-current characteristics of MOSFETs with gate lengths of 1, 5,

and 10 μm, respectively. As observed by Momose et al. [1], Figs. 5 and 6 show that the direct-tunneling current through the gate oxide distorts the drain-current characteristics. Figures 8 and 9 show that, when V_{GB}=2.0 V and V_{DS} is low, the gate current is larger in magnitude than either the drain or source currents. Both the drain and the source are injecting electrons into the channel to replace electrons that have tunneled through the oxide. Figure 5 shows how the drain-current characteristics can be improved by reducing the gate length to 5 μm [1], although for $V_{GB} = 2.0$ V, the drain current remains negative until $V_{DS} > 0.17$ V. As Fig. 9 shows, the large gate currents at small values of the drain bias V_{DS} also increase the magnitude of the source current.
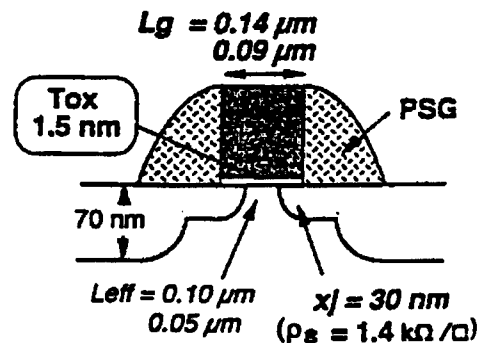
Fig. 2. Schematic cross-section of the MOSFET reported by Momose et al. [1] and simulated in this work.
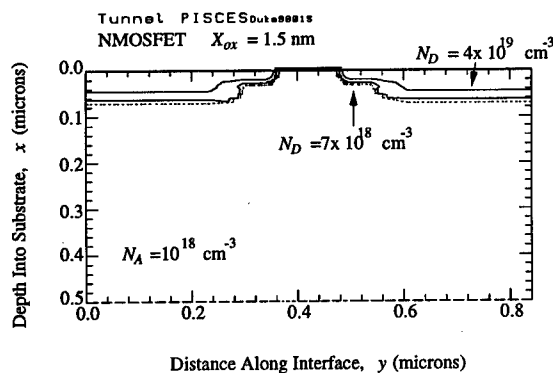
Fig. 3. Tunnel-PISCES device simulation cross-section for the MOSFETs simulated in this work.
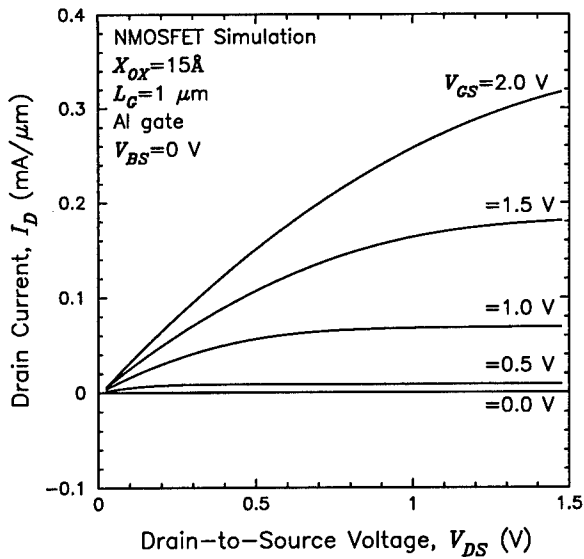
Fig. 4. Tunnel-PISCES simulation results of the drain-current characteristics of a MOSFET with a 1.5 nm-thick oxide and a 1 μm-long channel.
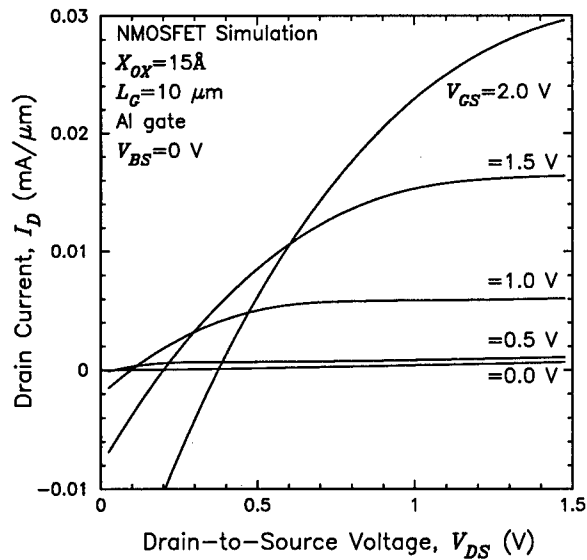
Fig. 6. Tunnel-PISCES simulation results of the drain-current characteristics of a MOSFET with a 1.5 nm-thick oxide and a 10 μm-long channel.
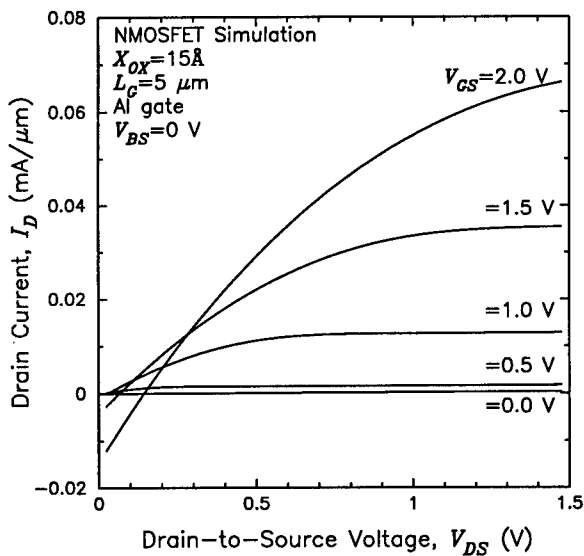
Fig. 5. Tunnel-PISCES simulation results of the drain-current characteristics of a MOSFET with a 1.5 nm-thick oxide and a 5 μm-long channel.
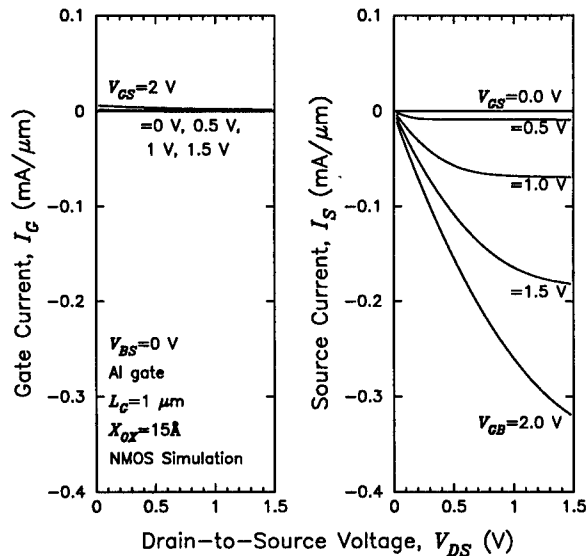
Fig. 7. Tunnel-PISCES simulation results of the gate- and source-current characteristics of a MOSFET with a 1.5 nm-thick oxide and a 1 μm-long channel.
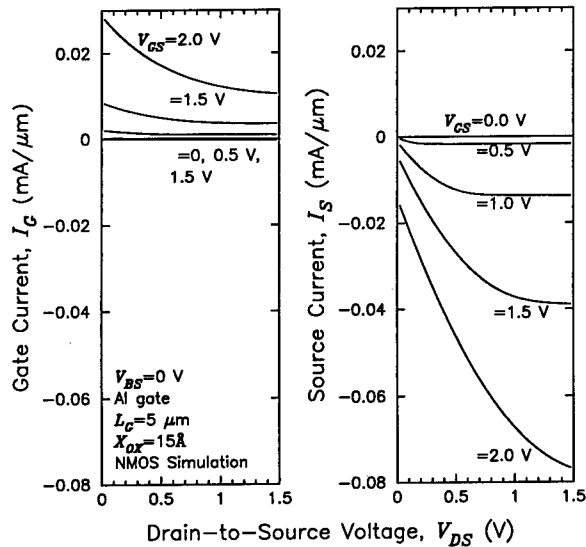
Fig. 8. Tunnel-PISCES simulation results of the gate- and source-current characteristics of a MOSFET with a 1.5 nm-thick oxide and a 5 μm-long channel.
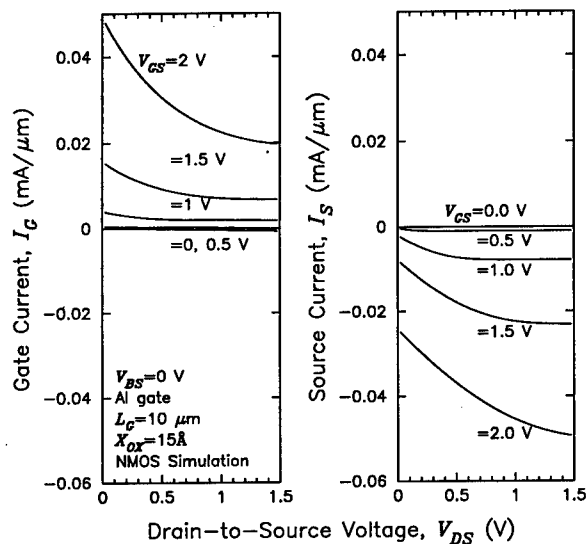
Fig. 9. Tunnel-PISCES simulation results of the gate-and source-current characteristics of a MOS-FET with a 1.5 nm-thick oxide and a 10 μm-long channel.

Finally, similar to the observations made in [1], Fig. 4 shows how a further reduction of the gate length to 1 μm produces a drain current characteristics that does not include the low-V_{DS} anomalies observed in Figs. 5 and 6. Figure 7 shows that even with small values of V_{DS} and with $V_{GS} = 2$ V, the direct-tunneling current through the gate oxide is not large enough in magnitude to dramatically perturb the source current.

4. CONCLUSIONS

We have reported simulation results of the effects of direct tunneling through the gate oxide upon the drain and source currents of deep-submicron MOSFETs. These simulations are needed to facilitate the design and optimization of such high-performance deep-submicron MOS-FETs with gate oxides scaled below the direct-tunneling limit.

REFERENCES

1. H. S. Momose, *et al.*, "1.5 nm Direct-Tunneling Gate Oxide Si MOSFET's," IEEE Trans. Electron Devices, **36**, 11 (1996).
2. J. P. Shiely and H. Z. Massoud, (to be published).
3. M. R. Pinto, C. S. Rafferty, and R. W. Dutton, "PISCES-II - Poisson and Continuity Equation Solver," Stanford Electronics Laboratory Technical Report, Stanford University, Sept. 1984.
4. J. G. Simmons, "Generalized Formula for the Electric Tunnel Effect Between Similar Electrodes Separated by a Thin Insulating Film," J. Appl. Phys., **34**, 6 (1963).
5. B. K. Ip and J. R. Brews, "Quantum Effects upon Drain Current in a Biased MOSFET," IEEE Trans. Electron Devices, **45**, 10 (1998).

ELSEVIER

The effect of polysilicon grain boundaries on MOS based devices

W. Eccleston

Department of Electrical Engineering and Electronics,
University of Liverpool, Liverpool, L693BX, UK.

Analysis is provided of the flow of current through undoped polysilicon. Charge is assumed to be trapped in grain boundaries giving potential barriers across which the free electrons and holes pass. Barriers height depends on gate and drain voltage, as does charge in the grain. At low current the electrostatic effect of the free charge is much less than trapped. There is, therefore, no gradual channel and flux of charge is controlled by the differences in concentration in adjacent grains(quasi-diffusion) as in the subtreshold region of MOSFETs. When the number of free carriers becomes comparable with that on the gate, there is an ohmic change in potential, and the channel current is drift limited (quasi-drift) by analogy with the gradual channel condition in MOSFETs. Variation field effect current with gate and drain voltages is compared with measurements at various temperatures

1. INTRODUCTION

The problem of current flow in polycrystalline semiconductors has not attracted the same interest as amorphous or single crystal materials. When it has, the focus has been heavily doped films where electron trapping is relatively unimportant (1). The uses of polycrystalline silicon have expanded rapidly in the last decade. There is, of course, its use as a gate material, for on-chip rails, and bipolar emmitters. It is important for thin film transistors in display applications and for a very important class of coulomb blockade devices. The first use of polycrystalline thin films of semiconductors was in thin film transistors of CdSe or CdS (2). The well used gradual channel approximation emerged from that work. Conveniently, that analysis avoided mention of the effects of grain boundaries on channel current. This continued to be the case until good computer device models emerged. An analytical model has until recently not been available.

The need to reduce gate oxide thickness has led to the problem of the implanted ions penetrating through poly and insulator into the channel region of MOSFETs. Reduced depths of implant,while solving the problem, leave the lower surface of the polysilicon intrinsic. There is then the possibility of large variations in the potential of the lower face of the poly and a consequent fluctuation in the field within the gate oxide. Subthreshold slope is then degraded. The same effect will also influence the operation of coulomb blockade devices, although almost nothing is known about the location of the stored charge in such devices.

2. MODEL

Whilst CV measurements may have a place in the study of polysilicon, they measure average trapped charge concentrations when the grain size is much less than the dot diameter. For this reason we have chosen first to attempt an analytical study of surface conductance, much the same as the early field effect workers (3).

We assume intrinsic grains, defect free, with grain boundaries of finite thickness, containing densities of trapped charge which increase exponentially towards the band edges. It is consistent with the presence of amorphous tissue in the boundaries, but this is not, of itself, of great significance since the theory copes with a series of different exponential functions, the sum of which can provide an adequate description of a wider range of distributions. Prior to the application of a gate voltage the fermi level is assumed to be near to the centre of the gap, both in the gap and also, therefore, within the boundary. For electrical neutrality this probably means that those states in the top half of the energy gap

are acceptor -like (negative with a trapped electron), and donor-like in the lower half of the gap (neutral when filled with an electron). Suppose the gate oxide on the surface has a metal electrode with the flat band voltage applied relative to the back of the poly. The conduction band will be flat across the diameter of the dot, for both grain and boundary regions. With a small positive gate voltage the surface potential remains constant, in spite of the pinning of the fermi level at traps in the boundary. This is because the Intrinsic Debye Length is much longer that the grain. The Debye Length falls rapidly with increasing electron concentration. There will come a point, therefore, with increase in gate voltage, where the conduction band droops in the centre of the grain giving potential barriers which impede the flow of carriers along the surface.

For low currents, where there is insufficient carrier density to affect the surface potential, there is a carrier concentration difference between the centre of adjacent grains, which maintains the current flow down the surface.

At higher currents where there is sufficient charge density to change the surface potential, the carrier density is constant between adjacent
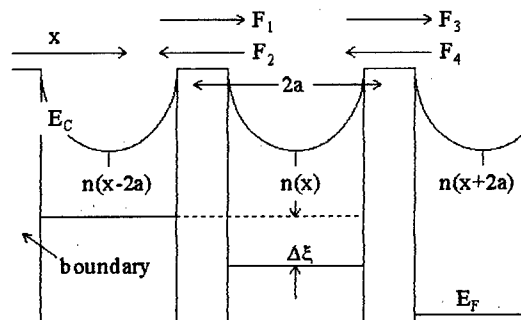
Fig.1 The variation of conduction band with distance inside a grain for low surface current.

grains, but there is a change in the level of the conduction band down the channel. Band diagrams for the low and high current conditions are shown in Fig. 1 and 2 respectively.

In Fig.1 the fermi level is drawn flat inside the grain because here the carrier concentration is high compared with that in the boundary.

There is, therefore, a rapid change in the fermi level in the grain boundaries. The current is determined by these changes.

In contrast the fermi level for high surface current remains at a constant energy difference from the mimimum in the conduction band edge. The current is determined by the difference in the barrier heights for electrons in the forward and reverse directions. This is shown in Fig.2.

With low currents (Fig.1) the net current is found by summing the fluxes F_1 to F_4 into the grain at x, where the carrier density is n(x), from the grains at either side. Note that 2a is the length of the grain.

$$\Delta F = -va^2 \exp-(\frac{q\Delta\phi_B}{kT}).[\frac{dn(x)}{dx}] -- (1)$$

where ΔF is the sum of all the fluxes into the grain at x, v is a constant, representing the frequency with which electrons attempt to move across the barrier. It has units of t^{-1}. $\Delta\phi_B$ is the height of the barrier. ΔF is proportional to the current density at some distant y from the surface. To find the total current it is necessary to integrate ΔF through the thickness of the film.

For the high current case, (Fig.2), after summing F_{e1} and F_{e2}, providing that the mean field along the surface E_X obeys the relationship:

$$E_x \leq kT / qa$$

the net flux is given by

$$\Delta F = n\mu_j E_x$$

where the mobility μ_j is an effective mobility that we have called the jump mobility and $\Delta F = F_1 + F_2$ of Fig.2. As with the low current case the flux must be integrated from the surface to the back of the film to find the current I. It is generally simpler to assume that the potential at the back of the film is zero, and a value at the front of the film, $\phi_y(0)$, that is
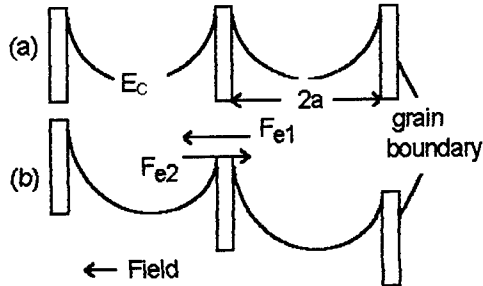
Fig. 2 The band diagram for high surface current. The barriers for high electrons moving in opposited directions are different when there is a lateral field (b).

determined by the gate voltage and also to change the integral so that it is in terms of this potential (ϕ_y)

$$ I = qW \int_o^{\phi_r(0)} \frac{\Delta F_x}{E_y} d\phi_y \quad (2) $$

where q is the electronic charge, W is the distance cross which the current flows and E_y is the value of the field at the surface.

Assumptions must be made in determining the value of E_y. We assume that the grain boundaries dominate the electrostatics in the y direction. By integrating Poisson's equation with the charge density be equal to that of the charge in the traps. The density of such traps N_T is given by

$$ N_T = N_0 \exp\left[\frac{q\phi_y}{kT_c}\right] $$

where T_C is simply a constant that defines the distribution and N_0 is the concentration that coincident with centre of the gap in the grain. The field E_y is given by

$$ E_y = \left[2\frac{N_o kT_c}{\varepsilon}\right]^{\frac{1}{2}} \left[\exp\left(\frac{q\phi_y}{kT_c}\right) - 1\right]^{\frac{1}{2}} $$

We have, for the low currents, the quasi diffusion equation (3) and for the high current condition (4).

Quasi-diffusion:

$$ I_D = \left(\frac{W}{L}\right) q n_i v a^2 \left(\frac{qkT_C N_0}{\varepsilon}\right)^{-1/2} $$
$$ \cdot \left(\frac{2kT_C}{qC}\right)\left(C_0 \varepsilon^{-2}(\frac{qkT_C N_0}{\varepsilon})^{-1/2}\right)^C V_{GT}^C \text{-----(3)} $$

Quasi-drift:

$$ I = \left(\frac{W}{L}\right)\left(\frac{2A}{C(C+1)D^C}\right)\left[(V_{GT}-V_D)^{C+1} - V_{GT}^{C+1}\right] \text{-(4)} $$

$$ C = \frac{2T_C}{T} - 1 \quad ; \qquad A = \frac{2\mu_j n_i kT_C}{K} $$

$$ D = \left(\frac{\varepsilon x_t K}{\varepsilon_o}\right)^C \qquad K = (2kT_C N_0 / \varepsilon)^{1/2} $$

L is the channel length, W is the channel width. C_0 is the capacitance per unit area of the oxide, ε is dielectric constant of silicon and n_i is the intrinsic carrier concentration.

3. THIN FILM TRANSISTOR CHARACTERISTICS
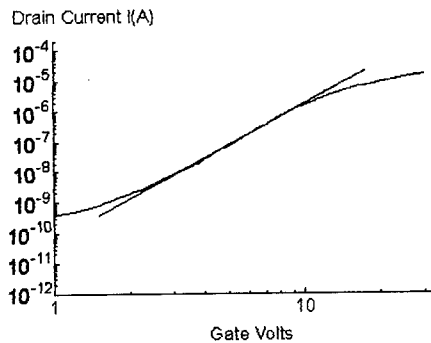
Fig.3 Logarithmic plot of current against gate voltage in the subthreshold (quasidiffusion) region indicated by the straight line.

Early studies of the field effect current involved measurement of the fractional change of resistance as a result of the field induced current in the surface of the semiconductor.
Thin film polysilicon transistors are used in advanced diplay devices. They are thin layers of polysilicon that have implanted sources and

drains. This enables the field effect current to be measured unaffected by the bulk conductivity. They have characteristics which resemble those of a conventional MOSFET and the sub- threshold region is shown in Fig. 3. The characteristic temperature T_C, in this region, is plotted against measurement and temperature in Fig.4.
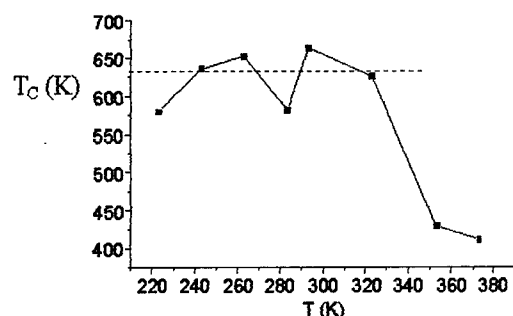
Fig. 4 Plot of characteristic temperature of the energy distribution of the traps against the measurement temperature. T_C is found to be constant within experimental error up to T=330K

In the quasi-drift region there is a linear variation of current with gate voltage as in Fig.5
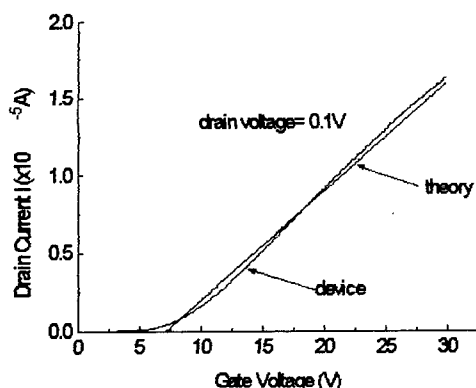
Fig. 5 Variation of drain current with gate voltage. The curve labelled 'theory' is fitted with equation (4).

The characteristic temperature in the quasi-drift region is found to be equal to the measurement temperature. There are two possible explanations. One is that there is no trapping. The second is more involved. The traps are in the grain boundary and the conduction band edge of crystalline silicon and the mobility edge of the amorphous silicon are almost coincident. Just beneath this energy

the amorphous silicon has a very high density of states. These states would be expected to fill with electrons as those at the mobility edge (4). There is under these cicumstances a proportionality between the free and trapped charge. In this treatment the trapping levels would be expected to have a characteristic equal to the temperature of the measurement.

There is a further consequence of the high current theory. The measured field effect mobility μ_{FE} is a combination of the jump mobility μ_J which is temperature dependent and the mobility at the surface of the grain μ_g. As the grains become larger the grain mobility dominates whereas with small grains it is the jump mobility. The relationship governing this behaviour is

$$\frac{1}{\mu_{FE}} = \frac{1}{\mu_j} + \frac{1}{\mu_g}$$
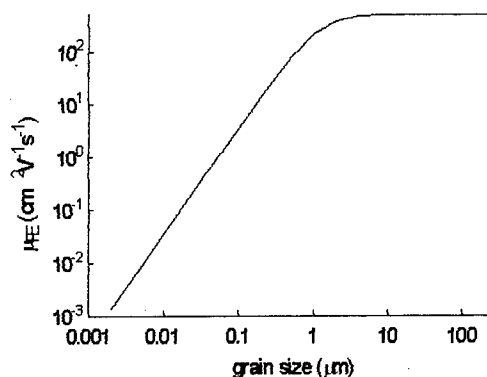
Fig. 6 Variation of field effect mobility with grain size

By parameter fitting to the characterisitics with an average grain length of 0.1μm we obtain the curve of Fig. 6 for the variation of field effect mobility with grain size.

5. REFERENCES

1. J. Y. W. Seto, JAppl. Phys, 46, No. 12, (1975), 5247
2. H. Borkan and P. Weimer, R. C. A. Review, 24, (1964), 153
3. A. Many, Y. Goldstein and N. Grover Semiconductor Surfaces," North-Holland Publishing Co., Amsterdam.
4. R. A. Street, "Hydrogenated Amorphous Silicon," Publ. Cambridge University Press.

Quantum Mechanical Characterization of E'-Type Centers in a-SiO$_x$ Films.

Shashi P. Karna[1] and Henry A. Kurtz[2]

[1]*Air Force Research Laboratory, 3550 Aberdeen Ave, SE, Kirtland AFB, NM 87117-5776*
[2]*Department of Chemistry, University of Memphis, Memphis, TN 38152*

Atomic structure and spin properties of model clusters representing neutral E'-type centers have been determined by *ab initio* Hartree-Fock calculations using extended basis sets. From a comparison of the calculated spin properties of the model E'-type centers with the experimental data, we identify the recently observed X center with O$_2$Si$\equiv$Si$^\uparrow$ atomic structure. The other E'-type species with OSi$_2$$\equiv$Si$^\uparrow$ atomic structure and identified here as the Y center has not yet been reported.

I. INTRODUCTION

Paramagnetic point defects are the main source of performance degradation of Si/SiO$_2$ based microelectronics devices [1]. Among the various point defects observed in the oxide layer, the majority is related to the oxygen-vacancy (V$_O$). The most extensively studied V$_O$ related point defects in metal-oxide-semiconductor (MOS) devices are the E' centers [2,3]. A common variant, called E'$_\gamma$, [4], is characterized by a (O$_3$$\equiv$Si$^\uparrow$ $^+$Si$\equiv$O$_3$) local atomic structure [5,6]. It is generally believed that the unpaired electron in this defect is localized in a sp^3-hybridized bond of a tetrahedral Si and the positive charge is localized on a distant, non-tetrahedral Si center, perhaps bonded to a triply coordinated O atom [5,6]. The atomic structure of another variant of the E' center, called E'$_\delta$ [7-10], is somewhat uncertain, although quantum mechanical calculations [11,12] suggest it to be a simple, symmetric variant of the E'$_\gamma$ center. Similar to E'$_\gamma$, the E'$_\delta$ center is believed to be a positively charged species when electron spin resonance (ESR) active [7-10]. There is also evidence in the literature of point defects similar to the E'$_\gamma$ center without the positive charge. These centers have been observed in Si-rich thin SiO$_2$ films near Si/SiO$_2$ interface and are called 'hemi', or surface E'$_\gamma$ (E'$_s$) centers [13,14]. The local atomic structure of this center is believed to be the electron-spin containing half, O$_3$$\equiv$Si$^\uparrow$, of the E'$_\gamma$ center, i.e., a tetrahedral Si center bonded to three O atoms and an unpaired electron (spin, s=1/2) localized in a dangling sp^3 bond. The E'$_s$ center is ESR active when neutral.

An all-Si counterpart of the E'$_s$ center, observed at the Si/SiO$_2$ interface, is the P$_b$ center of Nishi [15], characterized by a Si$_3$$\equiv$Si$^\uparrow$ local atomic configuration [16,17], in which the unpaired electron containing sp^3 orbital points in the direction of the oxide. Between the E' (O$_3$$\equiv$Si$^\uparrow$) and the P$_b$ (Si$_3$$\equiv$Si$^\uparrow$) centers, two more E'-type centers with configurations O$_2$Si$\equiv$Si$^\uparrow$ and OSi$_2$$\equiv$Si$^\uparrow$ are possible as defects in partially oxidized or Si-excess SiO$_2$ samples. These "E'-type" defects have been suggested to exist at the Si/SiO$_2$ interface [17], in a-SiO$_x$ (0<x<2) [18,19], and in Cr ion-implanted a-SiO$_2$ samples [20]. Recently, Hosono et al [21] have attributed a new doublet structure with 230 G separation in the ESR spectrum of Si-implanted a-SiO$_2$ to the E'-type centers. Experimental evidence also suggest E'-type defects with atomic structure O$_{3-n}$Si$_n$$\equiv$Si$^\uparrow$ (n=1 and/or2) to exist in x-ray irradiated phospho-silicate glass during high-temperature annealing [22] and in standard SiO$_2$ grown on Si during postoxidation high-temperature annealing in oxygen deficient ambient [23]. However, the exact nature of the atomic structure and composition of these centers remain unknown.

The goal of the present study is to understand the structural changes around a paramagnetic Si center that constitutes the E'$_s$ centers (O$_3$$\equiv$Si$^\uparrow$) in the oxide upon successive replacement of the O atoms by Si atoms by *ab initio* quantum mechanical calculations of the electronic structure and spin properties. Another goal of the present study is to develop an understanding of the electronic structure features of partially oxidized

Si centers in Si-excess a-SiO$_2$ and use it to explain the experimental observations [20,21]. As the first *ab initio* quantum mechanical study of the E'-type centers, the present set of calculations provide useful information on the distribution of unpaired electron spin in successively reduced (replacement of O atoms) E'$_s$ centers. Additionally, the results of the present calculations suggest that the Si-center responsible for the 230 G doublet, called the X center by Hosono et al. [21] in the Si-ion implanted a-SiO$_2$ samples, is bonded to two oxygen atoms and a Si atom (O$_2$Si≡Si$^{\uparrow}$). The second E'-type species in which a Si atom is bonded to two Si atoms and one O atom (OSi$_2$≡Si$^{\uparrow}$) has not yet been reported in a-SiO$_2$. This species, called the "Y center" here, is predicted to have a fairly large ESR hyperfine coupling constant to be detected and identified in Si-excess oxide samples.

II. CALCULATIONS

Calculations were performed on the O$_3$≡Si$_p$, O$_2$Si≡Si$_p$, OSi$_2$≡Si$_p$, and Si$_3$≡Si$_p$ model atomic clusters. The valency of the outer O/Si atoms was terminated by H atoms. The paramagnetic Si center is denoted in the chemical formula and in the rest of the paper as Si$_p$ to distinguish it from other Si atoms in the cluster. The O$_3$Si$_p$ model cluster represents the E' (E'$_s$) center while Si$_3$Si$_p$ represents the P$_b$ (or the D center in amorphous Si) center. The intermediate two structures represent the E'-type centers [21] generated by successively replacing one or two O atoms in the E'$_s$ structure by Si atoms. All four centers are neutral species. The stable geometries were obtained from total energy minimization of the clusters by use of an *ab initio* unrestricted Hartree-Fock (UHF) calculations with use of a 6-31G** Cartesian Gaussian basis set. The spin properties (spin density ρ^s, isotropic hyperfine coupling constant, a_{iso}, and components of the

anisotropic hyperfine coupling tensor T) of the model clusters at their optimized geometry were calculated by *ab initio* UHF method using a double-zeta plus polarization (DZP) Gaussian basis set. In order to examine the effect of the basis function on calculated results, the spin properties were also calculated using a triple-zeta plus polarization (TZP) basis set. The theory of hyperfine interaction and the details of our computational approach have been described elsewhere [11,12].

III. RESULTS AND DISCUSSION

The optimized equilibrium geometry of the model clusters used to represent the E', E'-type, and P$_b$ (D) centers are listed in Table I. The optimized Si-O bond distances are slightly larger than the average value of 1.61-1.62 Å found in a-SiO$_2$. The average Si$_p$-Si distance is about the same as that calculated for Si-Si dimer [11,12]. It is interesting to note that the Si$_p$-O bond distance increases as the number of Si$_p$-Si bonds in the cluster increases. The changes in the bond distances are accompanied by appropriate changes in the angles formed by O and Si atoms at Si$_p$. We note that the geometry around Si$_p$ center changes from a nearly-pyramidal one in O$_3$≡Si$_p$ to a nearly planar one in Si$_3$≡Si$_p$. This is consistent with the electronegativity of the atoms in an AB$_3$ cluster [24]. In accordance with the Walsh rules [24] for the AB$_3$ type structures, atom A (here, Si$_p$) moves toward the plane made by the three B atoms (here O and/or Si) as the electronegativity of the latter decreases.

Next we examine the spin properties of ^{29}Si center in our model clusters. The spin properties for the Si$_p$ center calculated with the use of a DZP (denoted A) and a TZP (denoted B) basis set are listed in Table II. Also listed under a_{iso} in column 3 of the Table are the

Table I. Calculated bond distances (Å) and bond angles (°). The paramagnetic Si center is denoted as Si$_p$.

Model Cluster	R(Si-O)	R(S-Si$_p$)	<O-Si$_p$-O	<Si-Si$_p$-O	<Si-Si$_p$-Si
O$_3$≡Si$_p^{\uparrow}$	1.643, 1.635,1.637	-	113.8, 107.2,105.5	-	-
O$_2$Si≡Si$_p^{\uparrow}$	1.653,1.652	2.355	110.3	106.2,112.7	-
OSi$_2$≡Si$_p^{\uparrow}$	1.664	2.356, 2.356	-	112.35	115.23
Si$_3$≡Si$_p^{\uparrow}$	-	2.351,2.351,2.351	-	-	115.8

experimental values [21,25] closest to the calculated results for a particular model structure. First, it is gratifying to note that the calculated spin properties change very little upon change in the basis set. Specifically, the isotropic hfc (a_{iso}) remains essentially unchanged or changes by a negligible amount between the two basis sets. The principal components of the anisotropic hfc tensor, T, exhibit a slight increase, by up to 10%, in the magnitude in going from the smaller (DZP) to the larger (TZP) of the two basis sets employed. The increased magnitude of T_{ii} in TZP basis set results from an improved description of the valence region by this basis set.

The calculated results for the spin properties reveal a number of interesting feature of the E'-type structures. First, we note that upon successive replacement of O atoms by Si atoms, the spin density, ρ^s, at the paramagnetic ^{29}Si nucleus shows a gradual decrease. Accordingly, the magnitude of the contact term, $a_{iso}(^{29}$Si), of the hyperfine coupling decreases gradually with the replacement of O atoms by Si atoms. On the other hand, the anisotropic part, T, of the hyperfine coupling that accounts for the spin density in the bonding region of the paramagnetic nucleus, shows a gradual increase upon successive replacement of the O atoms. Such changes are caused by the decrease of the $3s$ character in the spin containing orbital of the paramagnetic Si center due to a greater overlap with similar orbitals of the other Si atoms as the latter replace O atoms. At the same time, the

directional character of the dangling bond at the paramagnetic center also increases with increased overlap between the Si $3p$ orbitals. This results in an increase in the magnitude of the dipolar or the anisotropic part of the hyperfine coupling.

Upon comparison with the experimental hfc, we note that the calculated values of a_{iso} for the paramagnetic ^{29}Si$_p$ center in $O_3 \equiv Si^\dagger$ model cluster is too low by about 18 % compared to that for the E'$_s$ center [21]. In the case of the Si$_3 \equiv Si^\dagger$ model cluster, which represents the P$_b$ center, the calculated a_{iso} is about 10% too large compared to the corresponding experimental value [25]. A discrepancy of this magnitude between the theoretical values calculated on isolated molecular structures and experimental values of ESR hfc constants measured in the condensed phase is considered quite acceptable. This gives us confidence in the model clusters and the calculated results. The calculated a_{iso} for the paramagnetic ^{29}Si (Si$_p$) center in $O_2Si \equiv Si^\dagger$ cluster is in excellent agreement with the experimental value of 230 G measured for the X centers in Si-implanted a-SiO$_2$ samples [21]. The calculated value of the a_{iso} for the paramagnetic Si center in other E'-type cluster, $OSi_2 \equiv Si^\dagger$, is far too small to compare with the experimental value for the X center [21]. This comparison between the calculated results and the available experimental data on the ESR hfc constants of neutral V$_O$ related species allows us to

Table II. Calculated spin density, ρ^s, and hyperfine couplings (Gauss) for the paramagnetic (p) ^{29}Si center in E'$_s$, E'-type, and P$_b$(D) model clusters.

Model Cluster	Basis Set[a]	ρ^s	a_{iso}[b]	T_{11}	T_{22}	T_{33}
$O_3 \equiv Si^\dagger$	A	1.173380	-372.26	21.17	20.18	-41.35
	B	1.181968	-374.98 (440)	23.63	22.53	-46.16
$O_2Si \equiv Si^\dagger$	A	0.778456	-246.97	25.02	22.20	-47.22
	B	0.778458	-246.97 (230)	27.69	24.65	-52.34
$OSi_2 \equiv Si^\dagger$	A	0.612943	-194.46	26.11	24.70	-50.81
	B	0.614667	-195.01	28.80	27.32	-56.12
$Si_3 \equiv Si^\dagger$	A	0.393549	-124.85	25.95	25.95	-51.90
	B	0.395341	-125.42 (111)	28.61	28.61	-57.22

[a]The basis sets are A: DZP; B: TZP.
[b]The numbers inside parentheses are the experimental values taken from refs [21] and [25].

assign the atomic structures for the E'-type centers. We assign the X center, observed by Hosono et al. [21] with $O_2Si{\equiv}Si^{\uparrow}$ atomic structure and the Y center, which seems to have escaped the detection so far, with $OSi_2{\equiv}Si^{\uparrow}$ atomic structure.

IV. SUMMARY

We have calculated the structure and spin properties of model clusters representing $E_s'(O_3{\equiv}Si^{\uparrow})$, $P_b(Si_3{\equiv}Si^{\uparrow})$, and two intermediate species ($O_2Si{\equiv}Si^{\uparrow}$ and $OSi_2{\equiv}Si^{\uparrow}$) using *ab initio* Hartree-Fock method and extended atomic basis sets. Our calculations reveal that successive replacement of O atoms from the E'_s structure by Si atoms results in the relaxation of the central Si atoms from a nearly pyramidal position to a near-planar position. This change in the geometry accompanies a decrease in the isotropic hfc and an increase in the magnitude of the anisotropic hfc at the paramagnetic (central) Si atom. The calculated $a_{iso}(^{29}Si)$ for the model E'_s and P_b clusters are in good agreement with their experimental counterparts, lending confidence in the accuracy of atomic models and theoretical approach. A comparison of the calculated hfc for the E'-type model clusters with the recently observed X center lets us assign the latter with $O_2Si{\equiv}Si^{\uparrow}$ atomic structure. There seems to be no experimental data in the literature on the other E'-type species with $OSi_2{\equiv}Si^{\uparrow}$ atomic structure and identified as Y. The calculated hfc constants for the Y center are quite large and should be detected along with the X centers in Si-rich oxides.

References

1. D. M. Fleetwood, P. S. Winokur, L. C. Riewe, and R. A. Reber, Jr., J. Appl. Phys. **84**, 6141 (1998).
2. R. A. Weeks, J. Non-Cryst. Solids **179**, 1 (1994).
3. D. L. Griscom, Nucl. Instr. Methods B **1**, 481 (1984).
4. D. L. Griscom, E. J. Frieble and G. H. Sigel Jr., Solid State Commun. **15**, 479 (1974).
5. F. J. Feigl, W. B. Fowler, and K. L. Yip, Solid State Commun. **14**, 225 (1974).
6. K. C. Snyder and W. B. Fowler, Phys. Rev. B **48**, 13238 (1993).
7. D. L. Griscom and E. J. Frieble, Phys. Rev. B **34**, 7524 (1986).
8. R. Tohmon , Y. Shimogaichi, Y. Tsuta, S. Munekuni, Y. Ohki, Y. Hama, and K. Nagasawa, Phys. Rev. B **41**, 7258 (1990).
9. K. Vanheusden and A. Stesmans, J. Appl. Phys. **74**, 275 (1993).
10. L. Zhang and R. G. Leisure, J. Appl. Phys. **80**, 3744 (1996).
11. J. R. Chavez, S. P. Karna, K. Vanheusden, C. P. Brothers, R. D. Pugh, B. K. Singaraju, W. Warren, and R. A. B. Devine, IEEE Trans Nucl. Sci. **44**, 1799 (1997).
12. S. P. Karna, A. C. Pineda, W. M. Shedd, and B. K. Singaraju, Proc. Electrochem. Soc. (1999).
13. M. E. Zvanut, F. J. Feigl, W. B. Fowler, J. K. Rudra, P. J.Caplan, E. H. Poindexter, and J. D. Zook, Appl. Phys. Lett. **54**, 2118 (1989).
14. G. Hochstraser and J. F. Antonini, Surf. Sci. **32**, 644 (1972).
15. Y. Nishi, Japan. J. Appl. Phys. **5**, 333 (1966).
16. E. H. Poindexter, P. J. Caplan, B. E. Deal, and R. R. Razouk, J. Appl. Phys. **52**, 879 (1981).
17. P. J. Caplan, E. H. Poindexter, B. E. Deal, and R. R. Razouk, J. Appl. Phys. **50**, 5847 (1979).
18. E. Holzenkämpfer, F.-W. Richter, J. Stuke, and U. Voget-Grote, J. Non-Cryst. Solids **32**, 327 (1979).
19. I. P. Lisovskii, V. G. Litovchenko, V. B. Lozinskii, S. I. Frolov, H. Flietner, W. Fussel, E. G. Schmidt, J. Non-Cryst. Solids **187**, 91 (1995).
20. H. Hosono and R. A. Weeks, Phys. Rev. B **40**, 10543 (1989).
21. H. Hosono, H. Kawazoe, K. Oyoshi, and S. Tanaka, J. Non-Cryst. Solids **179**, 39 (1994).
22. D. L. Griscom, E. J. Frieble, K. J. Long, and J. W. Fleming, Appl. Phys. Lett. **54**, 3743 (1983).
23. A. Stesmans and V. V. Afanas'ev, Appl. Phys. Lett. **69**, 2056 (1996).
24. A. D. Walsh, J. Chem. Soc. 2266 (1953).
25. K. L. Brower, Appl. Phys. Lett. **43**, 1111 (1983).

Nature of the P_{b1} interface defect in (100)Si/SiO$_2$ as revealed by electron spin resonance ^{29}Si hyperfine structure

A. Stesmans[*] and V. V. Afanas'ev
Department of Physics, University of Leuven, 3001 Leuven, Belgium

Observation by electron spin resonance of the full angular dependence of the hyperfine (hf) interaction spectrum associated with the interfacial P_{b1} defect in thermal (100)Si/SiO$_2$ shows that the dominant interaction arises from a single ^{29}Si atom. The hf tensor displays weakly monoclinic I (nearly axial) symmetry, with the principal axes of the g and hf tensors coinciding. Observation of an identical hf structure consolidates the presence of the P_{b1} defect in porous Si. A molecular orbital analysis indicates that the unpaired electron resides for ~58% in a single unpaired Si hybrid orbital, found to be 14% s-like and 86% p-like, with the p-orbital pointing closely along a <211> direction at 35.26° with the [100] interface normal. Excluding oxygen as an immediate part of the defect, the results establish the P_{b1} core as a tilted (~20° about <011>) Si$_3$≡Si• unit. The moiety is pictured as part of a strained Si-Si dimer near the interface.

1. INTRODUCTION

The P_b-type centers are electron spin resonance (ESR) active coordination point defects inherently generated at the thermal Si/SiO$_2$ interface as a result of mismatch[1]. At least part of these were shown to be electrically detrimental charge traps, wherefrom originates the technological interest in atomic identification. Their appearance depends on the crystallographic interface orientation[1]. At the (111)Si/SiO$_2$ interface, ESR has so far isolated only one type of defect, specifically termed P_b, exhibiting C$_{3v}$ symmetry. It has been identified[2,3] as a trivalent interfacial Si, denoted as Si$_3$≡Si•. The (100)Si/SiO$_2$ interface, by contrast, exhibits two prominent types of defects, termed P_{b0} and P_{b1}, each with an inherent density[4] of ~1×10^{12} cm^{-2} for standard oxidation temperatures (800-950 °C). The initial observations[5] indicated lower than C$_{3v}$ symmetry (C$_{2v}$) –monoclinic I– for both defects, the P_{b0} symmetry, however, being nearly axial about <111>. Based on the close ESR features, there is now a consensus that P_{b0} is similar to P_b , but residing at (imperfections of) a macroscopic (100)Si/SiO$_2$ interface.[5]

The initially proposed model[3] for P_{b1} was an interfacial •Si≡Si$_2$O entity (cf. P_{b1}^P model in Fig. 1), suggesting the center to differ chemically from P_b and P_{b0}. That model, however, had appeared untenable based on calculations[6] on five model clusters, including the initial •Si≡Si$_2$O. None of the models appeared acceptable. So, up to now, the atomic identity of P_{b1} is still unknown, basically because of lacking hyperfine (hf) data. Recent work[5] on (100)Si/SiO$_2$ structures exhibiting predominantly the P_{b1} species resulted in accurate values of P_{b1} ESR parameters. The monoclinic I symmetry of the g tensor was confirmed with g_1=2.0058, g_2=2.00735 and g_3=2.0022, where, importantly, the g_3 direction is at 3 ± 1° (towards the [100] interface normal $\mathbf{n}$) with a <211> direction.

So far, only one successful P_{b1} hf observation in thermal (100)Si/SiO$_2$ has been reported[7], limited to a single orientation ($\parallel \mathbf{n}$) of the applied magnetic field $\mathbf{B}$. A hf splitting of $A_{[100]}$~157 G was inferred for P_{b1}, about 50 G larger than for P_{b0} (~105 G for $\mathbf{B} \parallel \mathbf{n}$) which in turn is comparable to the one of P_b (~117 G for $\mathbf{B} \parallel \mathbf{n}$). From this, it was speculated the P_{b1} hf structure also to arise from interaction with a single ^{29}Si, as is the case for P_{b0} and P_{b1}. Though useful, the singularity of this observation left any progress in modeling desperately speculative. Moreover, that hf observation was recently contested[8] by ESR results on porous Si (PSi), failing to resolve any hf structure attributable to P_{b1}. In the current work, we report on the successful determination of the hf tensor (A) symmetry and interaction strengths of the dominant ^{29}Si hf

structure, providing a fundamental clue to the defect's atomic identity.

2. EXPERIMENTAL

Slices of 2×9 mm^2 main face were cut from a 4 in.-diam. two side polished p-(100)Si wafer (float zone;~0.1Ωcm; ~ 29 μm thick), with the 9 mm edge along a <011> direction. After cleaning, the samples were submitted to three thermal steps: (1) Thermal oxidation at 970 °C (1.1 atm O_2; 99.9995 %; dry; d_{ox}~42 nm); (2) hydrogenation (H_2; 1 atm) at 795 °C for 1 h to maximize P_{b1} density; (3) vacuum anneal at ~620 °C for ~1 h –a treatment known to exhaustively depassivate (ESR-activate) the P_b-type defects[4]. Typically, an intensity ratio $[P_{b1}]/[P_{b0}] \approx 1.22$ is obtained, with $[P_{b1}]$ = $(7.2 \pm 0.5) \times 10^{12}$ cm^{-2}. An ESR sample bundle typically comprised about 70 slices. Porous Si was produced by stain etching[9] of p-(100)Si (~12 Ωcm) in a solution of 3g $NaNO_2$ in 100 ml HF(49%). Prior to etching, the surface was mechanically abraded to enhance surface reactivity. After having stayed in room ambient for ~24 h, the samples were annealed for ~1 h at 650 °C in vacuum (~1×10^{-6} Torr).

Conventional CW ESR (~20.1 GHz) measurements were carried out in the 1.6-4.3 K range. Routinely, it is operated in the adiabatic slow passage (incident microwave power P_μ ~20 pW) absorption mode, where first derivative absorption signals were recorded by modulation (~100 kHz; amplitude ~0.6 G) of **B**. Optimum hf structure detection, however, was obtained at higher P_μ (~0.8

nW). The rapid passage effects at 1.6 K under these partial saturation conditions resulted in recording undifferentiated absorptionlike peaks. **B** was rotated in the (011) plane with the angle φ_B of **B** with **n**, varying from 0 → 90°.

3. EXPERIMENTAL RESULTS

Typical ESR spectra observed in the low P_μ (undistorted) mode at 1.6 K are shown in Fig. 2 for three orientations of **B**. Though this detection mode is not the most sensitive one, hf structure is clearly resolved next to strong P_{b0} and P_{b1} (central) Zeeman signals. The simplest spectrum occurs for **B** $\parallel$ **n**, displaying pairs of hf doublets of splitting $\Delta B_{hf}[100]$= 105$\pm$2 and 156$\pm$2 G centred at the P_{b0} and P_{b1} Zeeman signals, respectively. The first one is the expected P_{b0} ^{29}Si hf structure, of splitting well in agreement with previous results[7,10]. The second doublet is assigned to P_{b1}. It was observed once before, with identical splitting[7]. In the latter work,
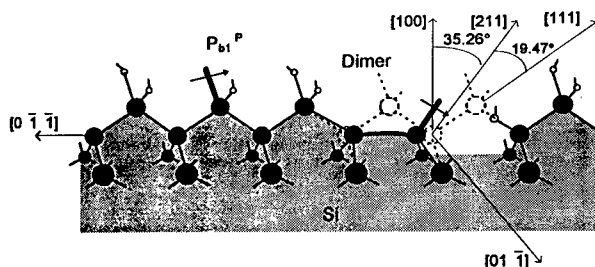
Fig. 1: Models for P_{b1} at the (100)Si/SiO$_2$ interface: (a) intial model (P_{b1}^P) after Ref. 3; (b) strained defected Si-Si dimer (similar to SB1 model in Ref. 6) with unpaired sp^3 hybrid adjusted along [211]. Dashed drawing pictures the Si lattice before the dimer formation.
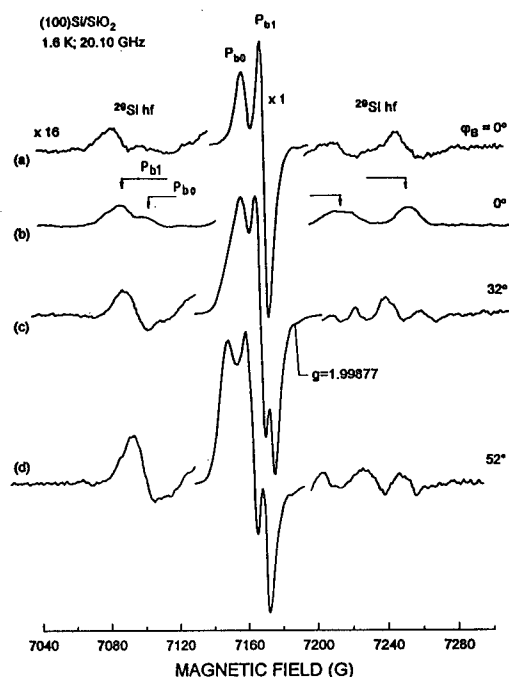
Fig. 2: Absorption derivative ESR spectra observed in the adiabatic slow (a,c,d: P_μ~20 pW) and rapid (b; P_μ~0.8 nW) passage mode in thermal (100)Si/SiO$_2$ for three directions of **B** in the (011) plane.

to maximize the signal-to-noise ratio, ESR was measured at temperatures <30 K in the dispersion mode under fast passage conditions giving absorptionlike signals. That observation (Fig. 4 of Ref. 7) is confirmed here when measuring at higher P_μ under similar circumstances [**B**∥**n**; cf. Fig. 2(b)].

Unlike previous work, the achieved signal enhancement has enabled us to perform the full angular variation of the hf structure. This is exemplified in Fig. 2 also, where the P_{b1} hf structure is seen to split into various, generally three, components. This is as expected for P_{b1} as for **B** rotating in the (011) plane, the g map exhibits three branches[3,5]. The three P_{b1} hf components exhibit different relative intensity, one being of approximately double intensity.

As determined on undistorted low P_μ spectra, the ratio in spectral intensity (area under absorption curve) of the hf doublet to the Zeeman signal is found to be 0.044±0.006 for P_{b1}. This agrees with the value of 0.049 expected for interaction with a single ^{29}Si (4.70% natural abundance) nucleus.

The entire P_{b1} spectrum can be described[11] by the simplified spin Hamiltonian $H = \mu_B \mathbf{B}.\mathbf{g}.\mathbf{S} + \mathbf{I}_j.\mathbf{A}_j.\mathbf{S}$, with effective electron spin S=½. The first term is the electronic Zeeman interaction, the second term the spin-nucleus hf interaction. Here g is the electronic g dyadic, I the nuclear spin (=½ for ^{29}Si), and $\mathbf{A}_j$ the hf tensor for interaction of the

electron spin with the jth nearby lattice site; for the present P_{b1} case, j=1. Similar to the Zeeman g map, the hf structure pattern is readily fitted with monoclinic I symmetry, with the relative branch intensities in agreement with experimental data. The optimised fitting gives the principal hf tensor values A_1(∥[011])=102±3 G, A_2(~∥[111])=112±3 G, and A_3(~∥[211])=167±3 G. The departure from trigonal (axial) symmetry thus appears small. In fact, within experimental accuracy, the data are equally well fitted by axial symmetry, giving the values A∥=A_3=167 G and $A_\perp$=107 G [=(A_1+A_2)/2]. We shall henceforth assume axial symmetry. The principal hf tensor axes are found coinciding with those of the g tensor within the experimental accuracy (~3°).

Figure 3 shows an expanded ESR spectrum observed on PSi for **B**∥**n**. When comparing with spectrum (b) in Fig. 2 for 'bulk' (100)Si/SiO₂, measured under similar conditions, the similarity is obvious, demonstrating the presence of both P_{bo} and P_{b1}. Two pairs of hf doublets my be discerned of splitting ΔB_{hf}=154 ± 3 and 110 ± 3 G, quite in agreement with those observed for P_{b1} and P_{bo}, respectively in standard (100)Si/SiO₂. However, the P_{b1} hf signals appear much broadened, i.e., ΔB_{pp}^{hf}~15 G versus 11±1 G in standard (100)Si/SiO₂. A similar broadening of the hf signals was observed[12] for P_b in (111)Si/native SiO₂. As suggested, this broadening may, at least partially, account for the previous failure to resolve ^{29}Si P_{b1} hf in PSi. The broadening is seen as a result of the greater sensitivity of hf interactions than the central Zeeman line to strain and disorder.

Interpretation within the classical linear combination of atomic orbitals (LCAO) analysis[11] indicates that the unpaired electron resides for ~58% in a single unpaired Si sp³ hybrid of 14% s and 86% p character. Remarkably, this unpaired Si hybrid points closely (within ~3°) along a <211> direction at 35.26° with n. The gyromagnetic and hf tensor symmetries are found identical within experimental accuracy. Within the LCAO framework, the results must imply that the key part of P_{b1} consists of a tilted ≡Si• entity that, under interfacial physicochemical influence, has rotated about a [011] axis over ~20° so as to bring the Si
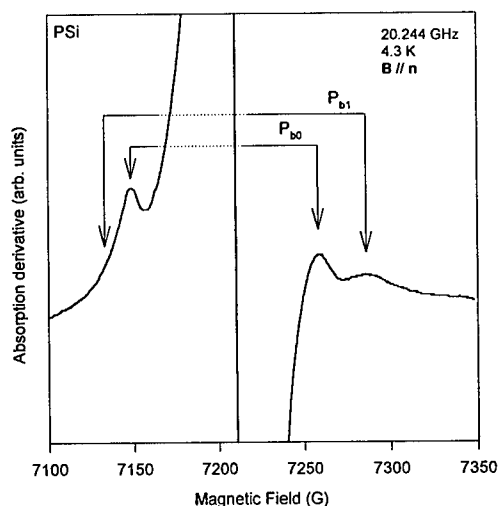
Fig. 3: ESR spectrum observed in PSi at 4.3 K in the rapid passage mode (P_μ ~2.5 nW) for **B**∥**n**.

dangling bond from its normal [111] direction towards to the nearest [211] direction.

As to modeling, the newly inferred hf data are strict. For instance, it eliminates at once all suggestions based an a $\equiv$Si• unit regularly incorporated in c-Si. Perhaps in a simplest scheme, it may be pictured as incorporated as one half of a $Si_2=Si•-Si\equiv Si_3$ defected dimer configuration (similar to SB1 in Ref. 6). As a result of the pulling of the two interfacial next nearest neighbor Si atoms together under influence of surrounding strain during the Si-Si bond reformation, the $Si_2=Si•-$ moiety with the left broken bond may be envisaged as having tilted over ~20° about the [011] axis away from [111] towards the [100] interface normal, the unbonded hybrid now pointing approximately along [211] (see Fig. 1). In this picture then, the fact that the g_2 axis direction ends up nearly along [111] is rather coincidental. Placed slightly subinterfacial, the defect structure thus being rigourously fixed by the Si lattice without much disturbance from the top SiO_2 network, this picture could, at least in principle, incorporate the various salient experimental facts thus far accumulated: The symmetry axes of the unpaired orbital at such tilted $Si_2=Si•-Si$ entity, i.e., [011], ~[211], and ~[111], agree with the measured principal g axes[5]. Also according to this symmetry, with one backbond strained, three different principal g value magnitudes are expected, that is, lower than axial symmetry. As the unpaired spin resides in a single dangling sp^3-like hybrid, the g and hf tensor symmetries are expected to coincide to first order, also as observed. While initial calculations[6] concluded the $Si_2=Si•-Si\equiv Si_2O$ dimer model to be untenable, more recent *ab initio* unrestricted Hartree-Fock calculations[13] support the dimer assignment.

4. CONCLUSIONS

Optimized ESR experiments have succeeded in the full angular mapping of the strong $^{29}Si\ P_{b1}$ hf interaction in thermal (100)Si/SiO$_2$. The P_{b1} defect is identified, like P_b, as a prototype Si dangling bond defect. If excluding oxygen as a basic part the defect, its basic entity is revealed as a <211> oriented (~20° tilted) strained $Si_3\equiv Si•$ unit. In broader sense, it appears part of a strained defected interfacial dimer. The presence of the identical

defect in porous Si is consolidated, herewith removing previous dispute.

REFERENCES

1. See, e.g., E. Poindexter and P. Caplan, Prog. Surf. Sci. 14, 211 (1983); R. Helms and E. Poindexter, Rep. Prog. Phys. 57, 791 (1994)
2. K. Brower, Appl. Phys. Lett. 43, 1111 (1983)
3. E. Poindexter, P. Caplan, B. Deal, and R. Razouk, J. Appl. Phys. 52, 879 (1981)
4. A. Stesmans and V. V. Afanas'ev, J. Phys.: Condens. Matter 10, L19 (1998)
5. A. Stesmans and V. V. Afanas'ev, J. Appl. Phys. 83, 2449 (1998)
6. A. H. Edwards, in *The Physics and Chemistry of SiO$_2$ and the SiO$_2$ interface*, edited by C. R. Helms and B. E. Deal (Plenum, New York, 1988), p. 271
7. K. L. Brower, Z. Phys. Chem. Neue Folge 151, 177 (1987)
8. J. L. Cantin, M. Schoisswohl, H. J. von Bardeleben, N. H. Zoubir, and M. Vergnat, Phys. Rev. B52, R11599 (1995)
9. N. Hadj Zoubir, M. Vergnat, T. Delatour, A. Burneau, Ph. de Donato, and O. Barrès, Thin Solid Films 255, 288 (1995)
10. W. E. Carlos, Appl. Phys. Lett. 50, 1450 (1987)
11. G. D. Watkins and J. W. Corbett, Phys. Rev. 121, 1001 (1961); 134A, 1359 (1964)
12. A. Stesmans, Appl. Surf. Sci. 30, 134 (1987)
13. A. H. Edwards and W. B. Fowler, Microelectr. Reliab. 39, 3 (1999)

Structure and electronic property of Si(100)/SiO$_2$ interface

Chioko Kaneta[a], Takahiro Yamasaki[a],
Toshihiro Uchiyama[b], Tsuyoshi Uda[b] and Kiyoyuki Terakura[c]

[a]Fujitsu Laboratories Limited,
10-1 Morinosato-Wakamiya, Atsugi 243-0197, Japan
Phone: +81-462-50-8241 Fax: +81-462-48-3473 e-mail: kaneta@flab.fujitsu.co.jp

[b]JRCAT-ATP,
1-1-4 Higashi, Tsukuba 305-8562, Japan

[c]JRCAT-NAIR,
1-1-4 Higashi, Tsukuba 305-8562, Japan

Stable structures and electronic states of Si(100)/SiO$_2$ interface have been investigated with the first-principles molecular dynamics method. Quartz, tridymite, and pseudo β-cristobalite are employed as the initial structures of SiO$_2$ at the interface to find the stable ones by structural optimization. It is found that the optimized tridymite-type SiO$_2$ on Si is most stable for a thin (about 7 Å) SiO$_2$ layer. For a thicker (about 15 Å) layer, however, this structure becomes less stable than the others, and the optimized quartz-type SiO$_2$ structure becomes most stable. Variation of the band gap perpendicular to the interface has also been investigated. In the SiO$_2$ region from the structural interface to a point about 1 Å away from it, the band gap remains as narrow as that of silicon. The dramatic change of the band gap takes place in the SiO$_2$ region from about 1 to 4 Å away from the interface.

1. INTRODUCTION

According to the requirement for thinner insulating layers in silicon-based electronic devices, the importance to characterize the Si(100)/SiO$_2$ interface region at the atomic scale is increasing.

On the Si/SiO$_2$ interface structure, a lot of experiments have been carried out, and several models which suggest the existence of crystal phases of SiO$_2$ have been proposed, for example, a tridymite layer about 5 Å thick at the interface [1], microcrystals of pseudo β-cristobalite (deformed β-cristobalite so as to match the lattice spacing to that of the Si substrate, as illustrated in Fig.1.) scattering in SiO$_2$ layers [2], and coesite [3]. The first-principles calculations employing the crystal phases as model structures of the SiO$_2$ [4,5], and the measurements of the density of the SiO$_2$ region near the interface [6], have also

been carried out. In spite of a great number of investigations, however, the basic properties of the interface itself are not well understood.

We investigated the stabilities of several model structures of the Si(100)/SiO$_2$ interface by employing the first-principles molecular
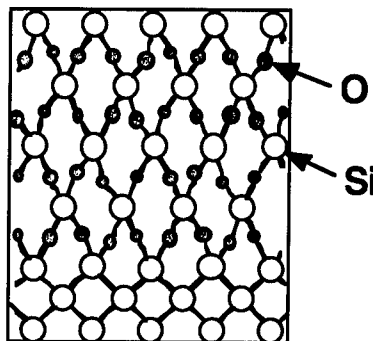
Fig. 1 Pseudo β-cristbalite on Si(100).

dynamics method [7]. The density of SiO_2 near the interface and the X-ray diffraction peaks were calculated and compared with the experiments. Energy band gap variations perpendicular to the interface were clarified for the ideally flat $Si(100)/SiO_2$ interface.

2. MODELS AND METHOD

Based on the experimental data which suggest the existence of crystal phases of SiO_2 at the interface, quartz, tridymite, and pseudo β-cristobalite were employed as the initial models of the SiO_2 structure to be optimized to find the stable ones.

The stable atomic configurations of the $Si(100)/SiO_2$ interface and their electronic structures were investigated using the first-principles method [7] based on the density functional theory [8]. We used a norm-conserving pseudopotential for silicon atoms, and ultrasoft pseudopotentials [9] for oxygen and hydrogen atoms, respectively. The generalized gradient correction [10] was taken into account for the exchange-correlation potential.

The unit cells containing 35 - 106 atoms are employed for our calculation. The Si substrate consists of 7 Si layers. We prepared two kinds of SiO_2 layers about 15 and 7 Å thick for each SiO_2 structure. The thicker (about 15 Å) layer consists of 4.5, 4.5, and 5 SiO_2 molecular layers in quartz, tridymite, and pseudo β-cristobalite, respectively. The thinner (about 7 Å) layer consists of 2 SiO_2 molecular layers in all of the types. The dangling bonds in the Si and SiO_2 surfaces are terminated with H atoms.

3. STABLE STRUCTURES OF THE Si(100)/SiO₂ INTERFACE

We optimized the interface structures with the quartz, tridymite, and pseudo β-cristobalite types of SiO_2 layers. Here, we call the optimized structures type-Q, T, and C, respectively. The structures with the thicker SiO_2 layers are illustrated in Fig. 2. In this section we discuss only the thicker layers.

We found that the pseudo β-cristobalite structure illustrated in Fig.1 is very unstable because the bond angles around Si atoms are largely different from those in sp^3 hybridization, 109.5°. Thus, a drastic structural change involving bond breaking occurs in the optimization process. The final structure (type-C) shown in Fig. 2(a) is completely different from the initial one. In spite of the bond breaking in the optimization process, there remains no dangling-bonds in the type-C structure.

For the quartz and tridymite types of SiO_2, no bond breaking occurs in the optimization process. Thus, the SiO_2 region

O Si   • O   • H
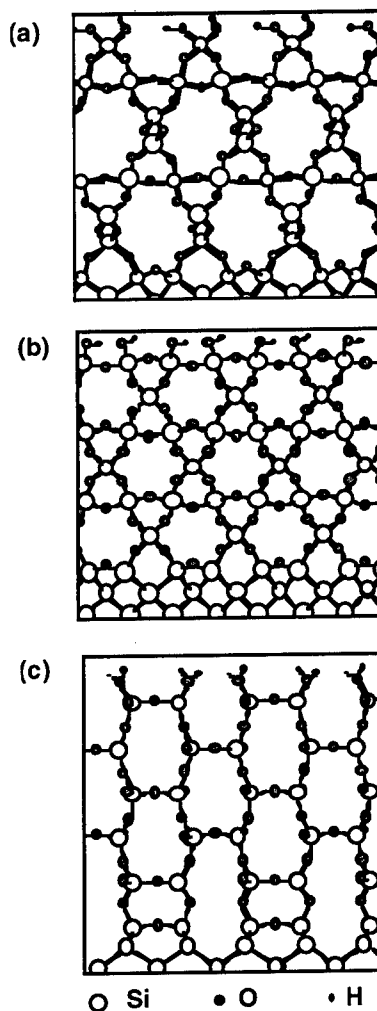
Fig. 2 Optimized interface structures,
(a) type-C, (b) type-Q, and (c) type-T.

in the type-Q and T structures basically have the characteristics of the quartz and tridymite structures, respectively. The whole structure is, however, largely distorted in the type-T (Fig. 2(c)) because of the large lattice mismatch between tridymite and silicon. The Si-O bond lengths and the X-Si-Y (X, Y = O or Si) bond angles in the type-T distribute in a wider ranges (1.58-1.68 Å and 95-125°, respectively) than in the type-Q and C. There are no dangling bonds at the interface also in the type-Q and T structures. The atomic configurations in the interface are very similar in the type-C and Q.

4. DENSITY OF SiO_2 IN INTERFACE REGION

On the basis of their X-ray reflection measurement, Awaji et al. [6] showed that the density of the SiO_2 near the interface is 2.42 g/cm^3, which is 91 % of that of α-quartz.

The densities of the SiO_2 region about 15 Å thick in type-C and Q are 93 % and 90-91 % of that calculated for α-quartz. They agree well with the experiment. For the type-T, we found the value is 99 %, which is rather larger than the experimental one.

5. X-RAY DIFFRACTION

Shimura et al. [2] reported that the extra peak at (1 1 0.45) observed in their X-ray diffraction experiment can be explained well by assuming the pseudo β-cristobalite for the SiO_2 structure. Our calculations, however have shown the structure is unstable. We investigated whether the type-C, Q, and T explain the extra peak well. We found that a strong reflection peak appears at (1 1 0.62) in type-Q. In this structure, however, another strong reflection also appears at (1 0 1.24), which has not been observed. We couldn't find consistent results with the experiment also for type-C and T. The discrepancy may be caused by the periodicity of our model structure. Further discussion is needed not only on the model for the calculation but also on the interpretation of the experiment.

6. COMPARISON OF STABILITY OF SiO_2 STRUCTURES

To clarify the origin of the difference among the total energies, we also compared the total energies between systems with the thicker and thinner SiO_2 layers for the type-C, Q, and T. In the following discussion, we neglect the difference of the configurations of the H atoms, because the contribution to the difference among the total energies was estimated to be small.

Comparing the calculated total energies, it is found that type-T structure is the most stable one when the layer is thinner (2 SiO_2 molecular layers). The difference of the total energy between that of the type-Q is -0.24 eV per one SiO_2 molecule for the type-T, and 0.17 eV for the type-C. For the thicker layer (5 SiO_2 molecular layers), however, the type-T becomes less stable because of the compressive stress in the SiO_2, and the optimized quartz-type SiO_2 (type-Q) becomes the most stable one. The difference of the total energy between that of the type-Q is 0.12 eV per one SiO_2 molecule for the type-T, and 0.28 eV for the type-C. The stabilities of the three structures for the thickness of the SiO_2 layer are compared in Fig. 3. The change of the stability for the thickness of the SiO_2 layer will cause a structural change in the whole layer or in the region away from the interface during the oxide growth.
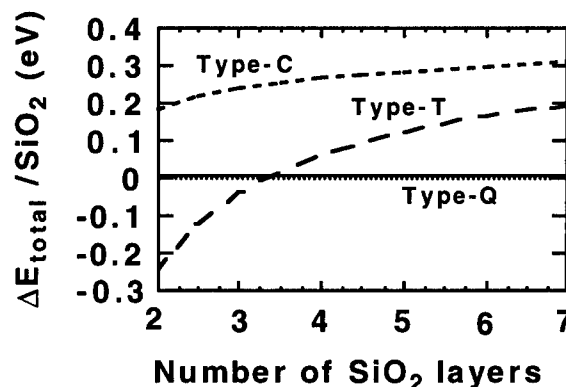
Fig. 3 The difference of the total energy between Type-C, T, and Q.

7. VARIATION OF BAND GAP

Variation of the band gap perpendicular to the interface was calculated for the thicker layers with the optimized structures. The result for the type-Q is shown in Fig. 4. In the SiO_2 region from the structural interface to a point about 1 Å away from it, which contains the first oxygen layer, the band gap remains as narrow as that of silicon. The drastic change of the band gap takes place in the SiO_2 region between about 1 and 4 Å. In type-C and T, the band gap changes similarly to that of the type-Q. But the barrier heights slightly depend on the SiO_2 structures. The difference between the position of the structural interface and that defined by the change of the band gap will give different pictures on the interface depending on the method of the observation.

8. CONCLUSIONS

We investigated the stabilities of three model structures of Si(100)/SiO_2 interface and the electronic properties using the first-principles molecular dynamics method. It was found that the optimized tridymite-type SiO_2 on Si is the most stable when the layer is thin (about 7 Å). For the thicker (about 15

Å) layer, however, this structure becomes less stable, and the optimized quartz-type SiO_2 becomes the most stable one. It was also found that the pseudo β-cristobalite is unstable. The densities of the SiO_2 region obtained for type-C and Q agree well with the experiment. We couldn't find consistent results with the extra peak at (1 1 0.45) observed in the X-ray diffraction measurement. Variation of the band gap perpendicular to the interface was calculated. In the SiO_2 region within about 1 Å from the structural interface, the band gap remains as narrow as that of silicon. The drastic change of the band gap takes place in the SiO_2 region between about 1 and 4 Å.

REFERENCES

1. A. Ourmazd, D. W. Taylor and A. Rentschler: Phys. Rev. Lett., **59**, 213 (1987).
2. T. Shimura, H. Misaki, M. Umeno, I. Takahashi and J. Harada: J. Cryst. Growth, **166**, 786 (1996).
3. V. V. Afanas'ev and A. Stesmans: Phys. Rev. Lett., **77**, 4206 (1996).
4. A. Pasquarello, M. Hybertsen and R. Car: Appl. Phys. Lett. **68**, 625 (1997).
5. H. Kageshima and K. Shiraishi : Appl. Surf. Sci. **407**, 133 (1998).
6. N. Awaji, Y. Sugita, T. Nakanishi, S. Okubo, K. Takasaki, and S. Komiya: J. Vac. Sci. Technol. **A14**, 971 (1996).
7. R. Car and M. Parrinello: Phys. Rev. Lett., **55**, 2471 (1985).
8. P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
9. D. Vanderbilt, Phys. Rev. **41**, 7892(1990).
10. J. P. Perdew: "Electronic Structure of Solids '91", ed. by P. Ziesche and H. Eschrig (Academie Verlag, Berlin, 1991).
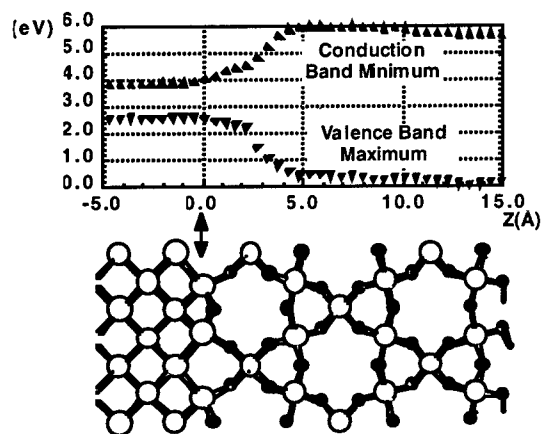
Fig. 4　Energy band gap variation in the type-Q structure perpendicular to the interface. The origin of the coordinate Z is taken at the interface.

A New Model for TDDB Lifetime Distribution of SiO₂

N. Matsukawa[a] and K. Kanebako[b]

[a] Memory Q & R Engineering Dept., Memory Division, TOSHIBA Co., Semiconductor Company
72, Horikawa-cho, Saiwai-ku, Kawasaki, 210-8572 Japan

[b] Memory Engineering Dep., Memory Division, TOSHIBA Co., Semiconductor Company
SET building, 1000-1, Kasama-cho, Sakae-ku, Yokohama, 247-8585 Japan

This paper provide a new model for TDDB lifetime distribution based on statistical considerations. The model is based on two assumptions: an assumption of hole capture cross section and an assumption of critical defect number. In this model, a small critical defect number yields the Weidull distribution and a large one changes it to the lognormal distribution. Degradation of oxide reliability corresponds to decrease of a critical defect number and so called extrinsic breakdown is induced by a small critical defect number.

1.INTRODUCTION

There have been two candidates for TDDB lifetime distribution of SiO₂ : a Weibull distribution and a lognormal distribution. In most cases, both distributions are derived empirically without a concrete physical model. In this paper, a new model is provided that can explain experimental data of both distributions.

There is another question about a power law time dependence of degradation (trap density). The cause of exponent factor (1/2~ 3/4) have been not known for long time. Our model will also statistically explain this power law time dependence.

2. ASSUMPTIONS

Our model is based on the following two assumptions, which we call a hole capture cross section model and a critical defect number model.

(1) A hole capture cross section model: In oxide, there are k number of potential defect d per certain unit area u, which is related to

breakdown. A potential defect d changes to an actual defect D, which induces breakdown, after capture of a hole with capture cross section of σ.

(2) Critical defect number model: Breakdown occurs when the number of actual defects in a unit area u becomes a critical number m which is characteristic of SiO₂ film, the material under consideration.

From these assumptions, lifetime distribution of TDDB is derived as follows.

1) Probability f of the change of a potential defect d to an actual defect D after n holes pass through a unit area u is described by the following equation.

$$f(n)=(\sigma/u)(1-\sigma/u)^{(n-1)} \qquad (1)$$

This equation also assumes no interaction/ interference between potential defects, that is $\sigma/u \ll 1$.

2) Probability P of the generation of m actual defects after n holes pass through a unit area u is described by the following equation.

$$P(n)=C(k,m)C(m,1)F(n-1)^{(m-1)}f(n) \times$$
$$(1-F(n))^{(k-m)} \qquad (2)$$

,where $C(n,q)$ is a combination number of select q number from total n number, $F(n)$ is a cumulative hole capture rate of a potential defect, that is

$$F(n)= \sum_{s=0}^{n} f(s).$$

3. CHARACTERISTICS OF THE MODEL
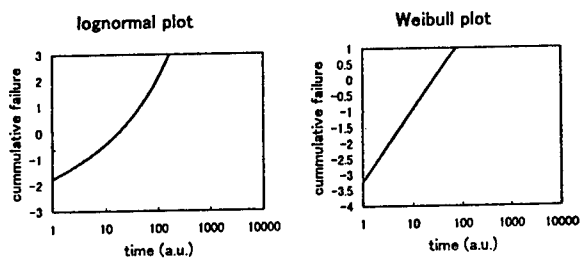
Fig. 1. TDDB lifetime distribution calculated from the model with a parameter set of (m,k, σ/u)= (1,20,0.002). In this case, lifetime has a Weibull distribution.
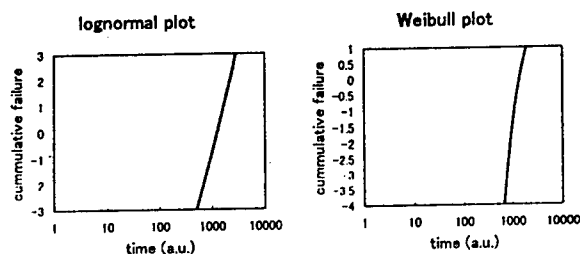
Fig. 2. TDDB lifetime distribution calculated from the model with a parameter set of (m,k, σ/u)= (19,20,0.002). In this case, lifetime has a lognormal distribution.

Figure 1 shows calculated results of equations with m=1, k=20, σ/u=0.002. In this figure a Weibull plot shows good straight line and lognormal plot shows arc-shaped curve, which means lifetime has a Weibull distribution. Application of other parameter sets reveals that m mainly determines

slope/shape of curves and k and σ/u mainly induce parallel shift of curves in time direction.

Figure 2 shows results of another parameter set of m=19, k=20, σ/u=0.002. In this case a lognormal plot shows a straighter line than a Weibull plot, which means lifetime has a lognormal distribution. In an experiment with a small number capacitors, however, it may be difficult to distinguish which distribution fit the data better.

4. COMPARISON WITH EXPERIMENTAL DATA

Figure 3 shows Weibull plots of TDDB data of 8.5nm oxide with various anneal conditions from a paper by K. Yoneda [1]. We could fit these 4 curves with 4 parameter sets of (m,k, σ/u)= (1,30,0.002), (5,30,0.002), (20,30,0.002), (28,30,0.002) for annealing time of 90min, 60min, 30min and 0min, Which means longer annealing time reduces critical number for oxide breakdown from 28 to 1.
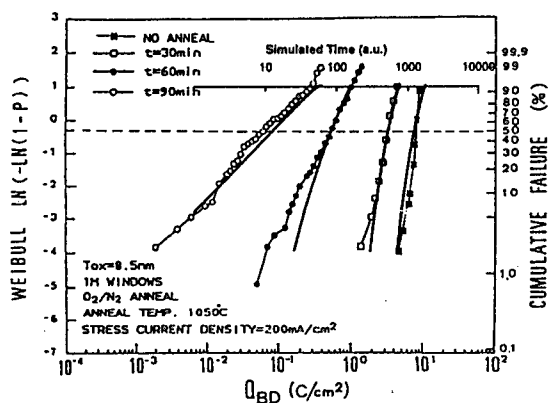
Fig. 3. Comparison of our model (four thick solid lines) with TDDB data from K. Yoneda [1]. Weibull plots of TDDB data are well fitted with parameter sets of (m.k. σ /u)= (1.30,0.002), (5,30,0.002), (20.30,0.002), (28.30,0.002) for annealing time of 90min, 60min, 30min and 0min, respectively.

In this figure, we assumed a fixed ratio

between electron flax and hole flax.

Figure 4 shows lognormal plots of other experimental TDDB data of 22.5nm oxide from a paper by J. Prendergast [2]. Here, the life time distribution consists of two parts: an extrinsic part and an intrinsic part. We could fit our equations to each part of 9MV/cm curve with parameter sets of (m,k, σ /u)=(1,80,0.002) and (7,80,0.002) (figure 5). These figures reveal that the extrinsic part has a Weibull distribution with a small critical number of defects and the intrinsic part has a rather lognormal distribution with a large critical number of defects. From this analysis, in this case, it will be concluded that extrinsic breakdown is not caused by extrinsic defects. So called extrinsic breakdown is also caused by intrinsic mechanism with small critical defect number.
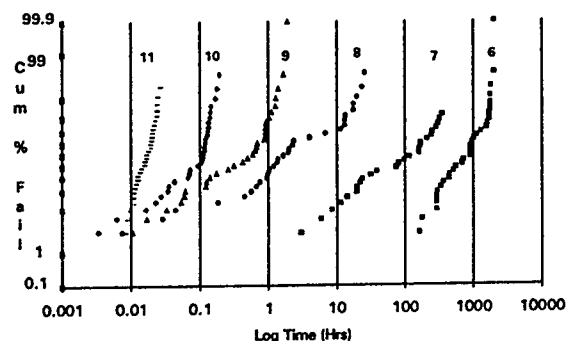
Fig. 4. TDDB data of 22.5nm oxide from J. Prendergast [2]. Numbers in the figure is applied electric field in MV/cm unit.

In this fitting, we assumed constant hole flux, which may not be so wrong assumption in the simulated time region, especially in intrinsic breakdown region. The scale of one order of time is exactly matched to that of the number of holes passing by.

The above parameter sets suggest that there are universal hole capture cross section σ and unit area u and the oxide has a rather constant number of potential defects per unit thickness, that is about 3.5/nm which is near the number of SiO_2 unit (~4/nm). The main factor determining oxide quality and TDDB

lifetime distribution is a critical number of defects, m. Weibull and lognormal distribution are reduced to small and large m, respectively.

Candidates for potential and actual defects may be Si-Si bonds and E' centers.

5. DEGRADATION POINT OF VIEW

Until now, we considered the oxide reliability problem from a breakdown point of view. In this section, it is discussed from a degradation point of view, which is evaluated by increase of trap density.
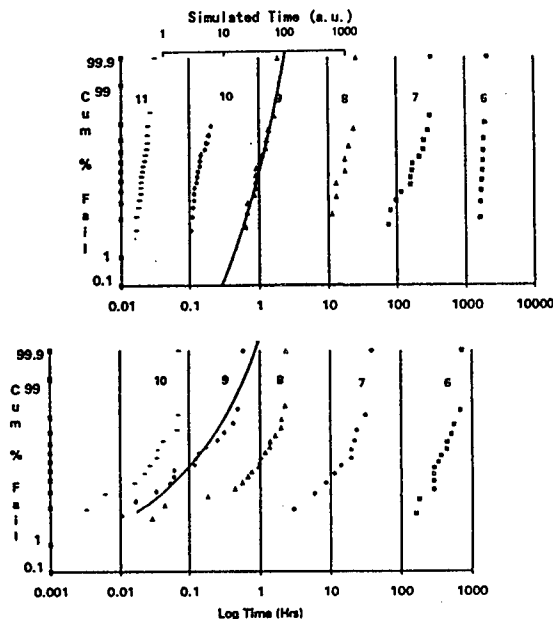
Fig. 5. Comparison of our model with separated intrinsic and extrinsic parts of 9MV/cm curve. We could fit our equations to each part with parameter sets of (m,k, σ /u)=(1,80,0.002) and (7,80,0.002). Numbers in the figure are applied electric field in MV/cm unit.

It is well known that trap density is not proportional to stress time and/or charge fluence. Many papers report power law with exponent between 1/2 and 3/4. For example, R.

Degraeve derived a exponent 0.56 against hole fluence [3].

The hole capture cross section model can derive this power law dependence, too. Average number r of real defects after n holes pass through a unit area u is calculated from n-1 step by the following equation.

$$r(n)=r(n-1)+(k-r(n-1))(\sigma/u) \qquad (3)$$

Calculated results, using a parameter set of figure 3 ((k, σ/u)= (30,0.002)), are shown in figure 6. Three curves correspond to three initial conditions of mo=0,1 and 2, where mo is numbers of initial actual defects. These initial conditions of mo=0,1 and 2 will provide experimental data with exponent about 1,0.75 and 0.5 respectively, because detailed curvature changes are usually not distinguished from experimental error. Except the case of mo=0, our model can explain reported power law dependence.

6.DISCUSSIONS

Here we discuss some perspective of other characteristics of TDDB phenomena, from the view point of our model.

Firstly, it is known that TDDB lifetime becomes short with increase of capacitor area. This may be due to increase of the initial. actual defect number of the worst unit area. If initial actual defects are scattered at random on all over the capacitor area with a certain area density, the actual defect number of the worst unit area increases with increase of capacitor area. Which correspond to decrease of critical defect number, that is shorter lifetime.

Secondly, TDDB lifetime becomes short with increase of temperature. This may be due to increase of hole capture cross section σ, because of thermal vibration of potential defects.

7.CONCLUSIONS

1) In this paper a new model for TDDB

lifetime distribution, which is based on statistical considerations, is proposed.

2) The model is based on two assumptions: a hole capture cross section assumption and a critical defect number assumption.

3) In this model, a small critical defect number yields the Weidull distribution and a large one changes it to the lognormal distribution.

4) Degradation of oxide reliability by high temperature annealing corresponds to decrease of a critical defect number.

5) So called extrinsic breakdown is also induced by a intrinsic mechanism with a small critical defect number.

6) Reported power low time dependence of trap density, that is actual defect number, is also derived from the hole capture cross section model.
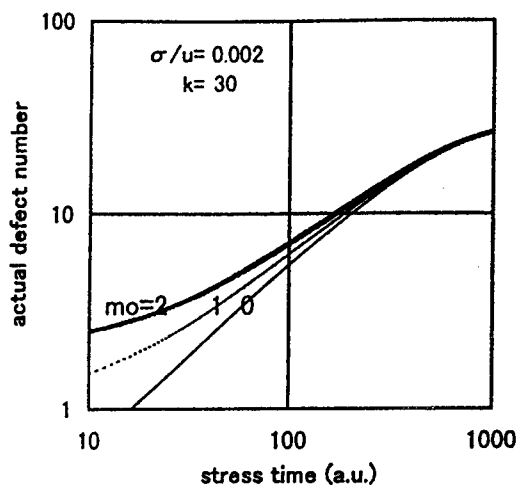
Fig.6. Time dependence of actual defect number, with parameter set of figure 3 ((k, σ /u)= (30,0.002)). Three curves are correspond to initial actual defects number of 0,1 and2.

REFERENCES

1.K. Yoneda et al., Proc. Symp. VLSI Tech.. p121, 1990

2. J. Prendergast et al.,IRPS Proc. p124, 1995

3. R. Degraeve et al.,IEEE ED Vol.45, p904, 1998

TRAPS AND DEGRADATION

Dependence of Gate Oxide Integrity on Grown-In Defect Density in Czochralski Grown Silicon

U. Lambert[a], A. Huber[a], J. Grabmeier[a], G. Obermeier[a], J. Vanhellemont[a], R. Wahlich[a] and G. Kissinger[b]

[a]Wacker Siltronic AG, P.O. Box 1140, D-84479 Burghausen, Germany

[b]Institute for Semiconductor Physics, Walter-Korsing-Str. 2, D-15230 Frankfurt (Oder), Germany

The dependence of the gate oxide integrity on the grown-in defect density is demonstrated using Czochralski grown crystals with different as grown defect densities and by varying the gate oxide thickness in the range of 5-25nm. The Poisson distribution $-\ln(1-F)=\rho_A A_{gate}$ (F: ratio of failures, ρ_A: area defect density, A_{gate}: gate area) is modified to $-\ln(1-F)=0.45\rho_B t_{ox} A_{gate}$ to be able to involve the bulk defect density ρ_B and the thickness of the gate oxide t_{ox}. The good agreement of that model with experimental data is proved comparing the density of infrared light scattering tomography defects and flow pattern defects with the gate oxide integrity yield.

1. INTRODUCTION

Defect engineering in Czochralski grown silicon gains more and more in importance since it is well known that grown-in defects cause severe problems for device yield such as early time zero dielectric breakdown, decreased reliability for time dependent dielectric breakdown, increased leakage current and decreased refresh times.

Silicon crystals exhibit in general a zoned distribution of microdefects [1] with three pronounced regions across the crystal radius. An interstitial rich region at the rim of the crystal is separated by a latent stacking fault ring from the central region characterized by an excess vacancy concentration [2]. Generally, for industrial fabrication of CZ silicon crystals the radial defect distribution can be modified by the v/G ratio (v: pulling speed, G: thermal gradient at solid/liquid interface [3]) in such a way that the entire wafer area is vacancy-rich. During cooling down of the crystal the vacancies agglomerate in grown-in microdefects which can be observed directly as light scattering tomography defects (LSTD), crystal originated particles (COP) after SC1 and flow pattern defects (FPD) after Secco etching. These microdefects in as grown silicon degrade the gate oxide integrity (GOI) on a large scale causing early breakdown of capacitor structures [4].

Recently, it has been observed that the gate oxide integrity yield increases with decreasing gate oxide thickness [5]. In particular, below 10nm the gate oxide yield grows very fast up to approximately 100% for a charge density of 5×10^{-4} C/cm² (Fig.1) as breakdown criterion.
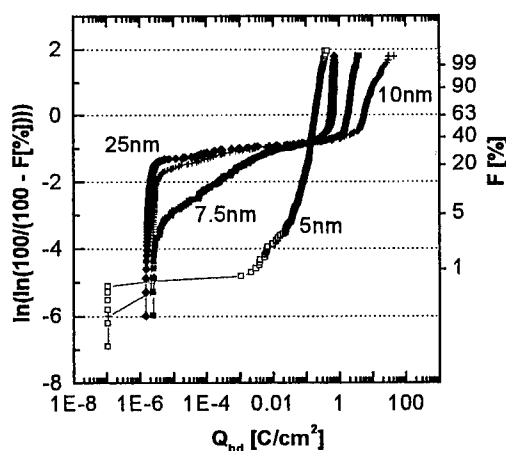
Figure 1. Cumulative breakdown curves as function of the charge density to breakdown Q_{bd} for oxide thickness' of 5-25nm on (001) p-doped 200mm wafers (D crystal, see experimental; gate area: 8mm²).

The combination of the grown-in defect density and the GOI yield dependence on the gate oxide thickness allows a description of the influence of the defect density on the gate oxide integrity yield.

2. EXPERIMENTAL

5-25 nm thin oxides were grown in dry oxygen ambient. The $8mm^2$ gate areas consisted of phosphorus doped poly-silicon. The gate oxide integrity was evaluated on 100 capacitors per wafer by applying a gradually stepped current density. The total charge density to breakdown Q_{bd} was determined by the time integral of the injected current density. The yield was calculated at a charge density of $5 \times 10^{-4} C/cm^2$ (E-field=10 MV/cm) which is well related to oxide breakdowns caused by grown-in crystal defects. Reference wafers processed simultaneously with the samples exhibited no deviation from the baseline with an constant gate oxide integrity yield in the range of 99 to 100%.

The LSTD-density (infrared light scattering tomography defects) was measured by a MILSA IRHQ-2 laser tomograph of RATOC using a NdYAG laser with a wavelength of 1.06 μm.

Flow pattern defects (FPD) were generated by vertical etching of wafers in a non-agitated Secco etchant for 30min. For density determination only the V-shaped flow patterns with an etch pit at the apex were counted as FPD.

All material used were CZ grown 150mm and 200mm boron doped (1-10 Ωcm), vacancy rich polished wafers with (001) orientation. They were produced from crystals grown in different hot zones (A,B,C,D) resulting in cooling rates of A>B>C>D.

3. RESULTS AND DISCUSSION

The relation between the yield (1-F) of the gate oxide measurement by current ramp (Q_{bd}) or voltage ramp (E_{bd}) and the area defect density is approximately given by the Poisson distribution

$$-\ln(1-F) = \rho_A A_{gate} \qquad (1)$$

where F, ρ_A and A_{gate} are the ratio of failures, the

area defect density and the gate area, respectively [6,7].

However, as the dependence of the GOI yield on the gate oxide thickness is obvious and assuming that the GOI yield is directly proportional to the number of bulk defects incorporated in the oxide equation (1) can be modified [8] by involving the thickness in

$$-\ln(1-F) = 0.45\rho_B t_{ox} A_{gate} \qquad (2)$$

with 0.45 as ratio of the contribution of the silicon substrate to the final thermal oxide and ρ_B and t_{ox} as the bulk defect density and the thickness of the gate oxide, respectively.

Under this assumption the yield of the gate oxide integrity can directly be compared with bulk defect densities and the thickness of the gate oxide is taken into account.

Fig.2 gives a graphical representation of function (2) for $t_{ox} = 5$, 7.5, 15 and 25nm. Nearly the total range between 0 and 90% GOI yield is covered by bulk defect densities between 10^6 and $3 \times 10^7/cm^3$.
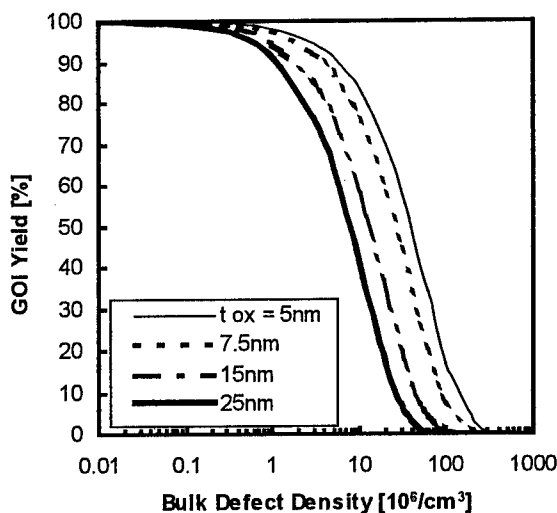
Figure 2. Calculated yield of gate oxide integrity as a function of the bulk defect density and the gate oxide thickness (gate area: $8mm^2$).

Using LST defect densities and GOI data (t_{ox}=15nm) from a series of 150mm crystals grown with different cooling conditions the same tendency

is observed which fits better after gate oxidation due to the apparently increased defect density [9] compared to as grown material (Fig.3).

From Fig.2 it can be derived that for a given bulk defect density the GOI yield increases with decreasing thickness of the gate oxide, assuming a constant bulk defect density.
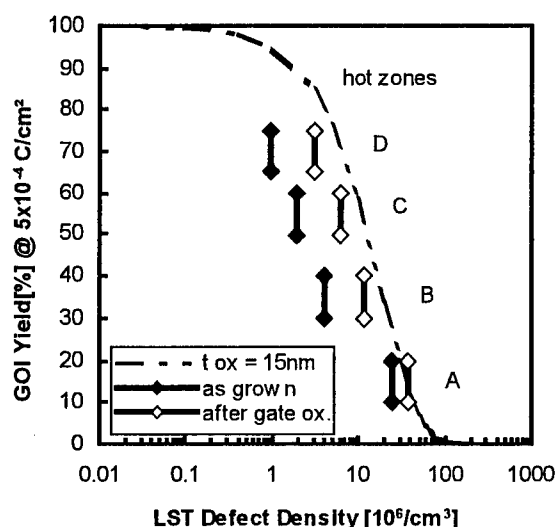
Figure 3. Measured gate oxide integrity yield range versus LST defect density for different hot zones A, B, C, D. The dashed line denotes the calculated GOI yield for t_{ox} = 15nm. (150mm CZ-crystals, as grown and after gate oxidation, t_{ox}=15nm; gate area: 8mm^2).

This is illustrated more clearly in Fig.4 showing the GOI yield as a function of the gate oxide thickness for a 200mm D-type crystal (black squares).

For a given iso-defect density line and within a gate oxide thickness range of 7.5-25nm the gate oxide integrity yield agrees fairly well with the predictions of the model taking into account a defect density variation within a range of ±1x10^6/cm^3 and a scattering of the GOI yield within ±5%. The white square represents the GOI yield for all type of crystals (A,B,C,D) for a gate oxide thickness of 5nm.

Flow pattern defects are related with clusters of vacancies (octahedral voids) because they only occur in the vacancy-rich center region of a crystal and

not in the interstitial–rich rim region [10,11]. A good linear correlation with the LST defect density can also be observed [12].
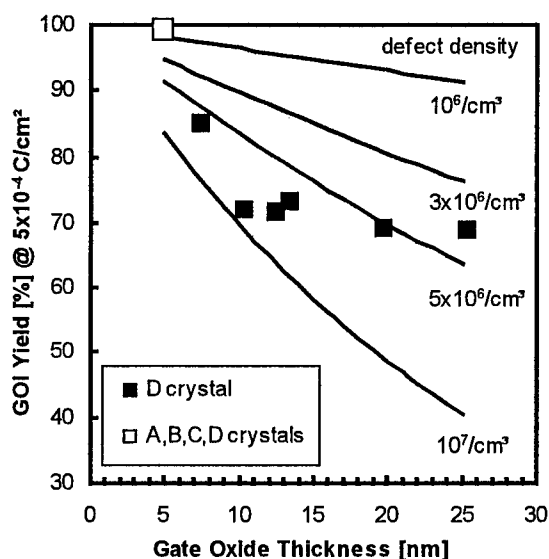
Figure 4. Gate oxide integrity yield versus gate oxide thickness in the range of 5-25nm. In the oxide thickness range of ≥7.5nm the experimental results fit fairly well with the model. For a gate oxide thickness of 5nm all types of crystals (A,B,C,D) show a high GOI yield of about 99%. (200mm crystals; gate area: 8mm^2).

In Fig.5 the measured gate oxide integrity yield for 200mm crystals with different cooling conditions (high yield=slowly cooled) is presented as function of the flow pattern defect density together with the calculated gate oxide integrity yield using equation (2). The correlation of the observed with the calculated GOI yields represents a good agreement between experimental data and calculated values.

4. SUMMARY AND CONCLUSIONS

It has been shown that for thin gate oxides between 7.5 and 25nm thickness, the relation between gate oxide integrity yield and grown-in bulk defect density can be described by modifying the Poisson distribution for area defects involving
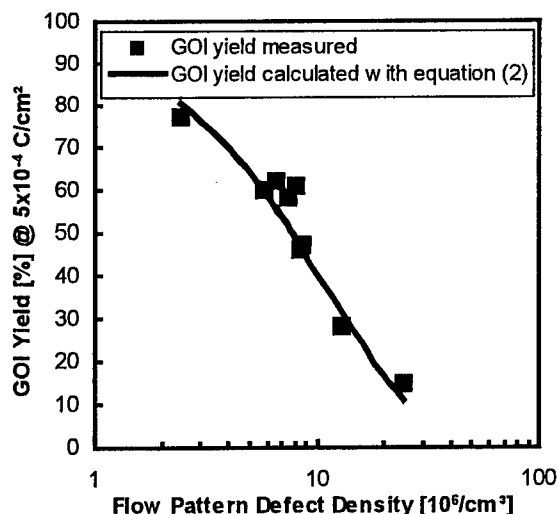
Figure 5. Observed and calculated gate oxide integri-
ty yield as function of the flow pattern de-
fect density (gate area: 8mm², t_{ox}=25nm).

the bulk defect density and the thickness of the gate
oxide.

The good correlation between the gate oxide
integrity yield and the grown-in defect density and
the thickness of the gate oxide, respectively, has
been demonstrated using light scattering tomogra-
phy and flow pattern defects from a series of CZ-
crystals grown with different cooling conditions.

The gate oxide integrity yield increases with de-
creasing grown-in defect density, however, it
increases also with decreasing gate oxide thickness.
At a thickness of approximately 5nm the influence
of the substrate defect density onto GOI becomes
negligible.

This observation may have several causes:

i) It is suggested that for thin gate oxides
(<10nm) the impact of the bulk defect density
decreases in favor of the silicon surface properties.

ii) The oxide thickness of 5nm is comparable to
the oxide thickness which has been observed on the
inner surface of the octahedral voids [13] providing
a pre-oxidized state. No significant oxide thinning
in void corners thus occurs during gate oxidation
[14].

iii) It is assumed that there is a thin silicon layer
at the surface with a defect density below 5×10^5/cm³
taking into account that only 2.3nm of the silicon is
consumed for an oxide thickness of 5nm.

REFERENCES

1. P.J. Roksnoer, J. Cryst. Growth **68** (1984), 596.
2. V.V. Voronkov, J. Cryst. Growth **59** (1982), 625.
3. W. v. Ammon, E. Dornberger, H. Oelkrug and H. Weidner, J. Cryst. Growth **151** (1995), 273.
4. M. Itsumi, O. Nakajima and N. Shiono, J. Appl. Phys. **72** (1992), 2185.
5. E.D. Grann, A. Huber, J. Grabmeier, R. Hölzl and R. Wahlich, Gate Dielectric Integrity: Material, Process, and Tool Qualification, ASTM STP 1382, D.C. Gupta and G.A. Brown, Eds., American Society for Testing and Materials, West Conshohocken, PA, 1999. (in press).
6. C.M. Osburn and D.W. Ormond, J. Electrochem. Soc. **119** (1972), 591.
7. C.M. Osburn and D.W. Ormond, J. Electrochem. Soc. **119** (1972), 597.
8. Y. Murakami, T. Shiota and T. Shingyouji, J. Appl. Phys. **75** (1994), 5302.
9. J. Vanhellemont, G. Kissinger, K. Kenis, M. Depas, D. Gräf, U. Lambert and P. Wagner, Proc."Oxygen '96. Early Stages of Oxygen Precipitation in Silicon", ed. R. Jones, Kluwer Academic Publishers (1996), 493.
10. H. Yamagishi, I Fusegawa, N. Fujimaki and M. Katayama, Semicond. Sci. Technol. 7 (1992), A135.
11. W. Wijaranakula, J. Electrochem. Soc. **141** (1994), 3273.
12. M. Hourai, T. Nagashima, E. Kajita, S. Miki, T. Shigematsu and M. Okui, J. Electrochem. Soc. **142** (1995), 3193.
13 H. Bender, J. Vanhellemont and R. Schmolke, Jpn. J. Appl. Phys. **36** (1997), L1217.
14. M. Itsumi, proc. symp. "Advances in Silicon Substrates", E-MRS 1999 Spring Meeting, to be published in Materials Science and Engineering B.

Suppression of thermal degradation in standard Si/SiO₂ by noble gases

A. Stesmans[*] and V. V. Afanas'ev

Department of Physics, University of Leuven, 3001 Leuven, Belgium

Annealing-induced degradation of Si/SiO₂ has been studied in noble gas ambients (He, Ne, Ar). The vacuum postoxidation annealing induced interface degradation in (111)Si/SiO₂, previously identified as intense creation of interfacial Si dangling bond defects (P_bs: $Si_3 \equiv Si\bullet$) from ~640 °C on, is found remarkably impeded, inversely proportional to the gas atomic diameter: He fully blocks the process up to ~800 °C, Ne suppresses partially, while the effect of Ar is marginal. This is attendant with a significant reduction of oxide degradation, including suppression of the degradation of the insulating properties of ultrathin SiO₂ layers and generation of hole trapping centers. The data support the degradation model based on interfacial SiO(g) release, showing that the degradation mechanism may just occur as the result of the existence of a typical interstice/channel structure in thermal SiO₂. The observed process represents blocking of a chemical reaction in a physical way.

1. INTRODUCTION

In the course of the evolution of metal-oxide-semiconductor device technology, the investigation of the influence of postoxidation annealing (POA) has attracted much attention. Among others, it was found that the basic Si/SiO₂ unit integrity is deteriorated when submitted to POA in vacuum or inert O-free ambient[1]. In the initial stage, electrical degradation manifests itself as reduction of breakdown strength of SiO₂ gate insulators, oxide charging, interface trap generation, and reduced radiation hardness[1-7]. In a detailed electrical study[5] of the POA influence on SiO₂ breakdown strength, the removal of interfacial SiO(g) was advanced as the primary degradation step, based on the *overall* equilibrium [SiO(g) production] reaction

$$SiO_2 + Si \Leftrightarrow 2SiO(g). \tag{1}$$

The presence of O in the anneal ambient was recognized a as crucial factor, driving the equilibrium of this reaction. By its nature, the process will also affect the interface integrity.

First atomic insight came from recent systematic electron spin resonance (ESR) analysis of vacuum POA, the information being twofold. First, one of the mechanisms of *interface* degradation in thermal (111)Si/SiO₂ was atomically

identified[8], i.e., profound creation of permanent interfacial P_b defects (trivalent interfacial Si, denoted as $Si_3 \equiv Si\bullet$) from ~640 °C onward in monotonically increasing densities N_c up to ~1.3×10^{13} cm⁻² at 1200 °C. These appear in addition to the density $N_0 = (4.9 \pm 0.4) \times 10^{12}$ cm⁻² naturally present[8,9] in as-grown Si/SiO₂ in the range 300-1000 °C. At the (111)Si/SiO₂ interface, P_b is the sole point defect so far isolated by ESR, known to be an adverse electrically-active interface trap[10]. The degradation mechanism, thus implying permanent Si bond rupture, was also interpreted as driven by interfacial SiO(g) removal. Second, ESR also traced degradation of the thermal *oxide* during vacuum POA, with major observation profound generation of a 'new' type of intrinsic defect ascribed to S centers (tentatively E'-like defects of the type $OSi_2 \equiv Si\bullet$ and/or $O_2Si \equiv Si\bullet$) from an anneal temperature $T_{an} \sim 950$ °C onward. Concomitantly, for $T_{an} > 1000$ °C, weaker amounts of other more familiar types of intrinsic point defects (EX, E'$_\gamma$ and E'$_\delta$ centers) are observed to be produced also[11,12]. So, POA degradation in vacuum is found to constitute abundant generation of atomic, predominantly intrinsic defects. However, rather than vacuum, it may appear more interesting to investigate the influence, if any, of other ambients on the defect creation mechanism(s). This is the

subject of the present work, where the influence of noble gases (He, Ne, Ar) on interface and oxide degradation is probed both by ESR and electrically. Though rather sporadically, POA in noble gases has been studied electrically before, with however often inconclusive results[1].

2. EXPERIMENTAL

Various thermal Si/SiO_2 structures were studied: a) structures grown through standard oxidation in a laboratory facility at ~970 °C (1.1 atm dry 99.9995% O_2; oxide thickness 45-65 nm) of commercial (100) and (111)Si wafers (float zone; >100 Ωcm; p-type). In agreement with previous work[8,9], a first control ESR test yielded $[P_b] \sim 4.9 \times 10^{12}$ cm$^{-2} \approx N_0$ –an important reference value as the study intends to probe the P_b creation during subsequent POA; b) ultrathin (4-6 nm) oxide (100)Si/SiO_2 entities (800 °C; dry N_2+O_2) grown in a technological clean room facility; c) (100)Si/SiO_2/poly-Si fabricated through oxidation of p-(100)Si (dry O_2; 900 °C) and subsequent deposition of a 310-nm thick poly-Si layer by chemical vapor deposition. Sets of samples were submitted to POA at T_{an} = 500-1160 °C in He, Ne, Ar, and N_2. All gases were boiled off from the liquid phase to minimize water contamination.

Conventional CW absorption-derivative mode ESR (~20.6 GHz) was performed at 4.3 K, as outlined elsewhere[8,9], with the applied magnetic

field normal (within 3°) to the (111)Si/SiO_2 interface to assure optimum sensitivity. Each ESR sample was started from freshly oxidized (~10) slices. Electrical characterization was carried out at room temperature using current-voltage and 1 MHz capacitance-voltage (CV) techniques. To study hole trapping, electron-hole pairs were generated in the oxide by 10-eV photons.

3. RESULTS

We first address interface degradation, studied previously[13]. Figure 1 shows the observed areal P_b density as a function of the temperature T_{an} of isochronal (~1 h) POA in He, Ne, and Ar. For comparison, coplotted are the data of the 'fresh oxide' sample set from previous work[8] exposing the POA creation mechanism in vacuum; similar results are found for N_2. A remarkable finding is that the interface degradation is suppressed by noble gases, and moreover, the effect depends on the type of gas: The impact decreases from He over Ne to Ar, with the effect of Ar only being marginal. The impact of He is most remarkable as, apparently, it succeeds in suppressing interface degradation in terms of P_b creation up to ~800 °C for POA times up to 1 h. Above this T_{an}, degradation occurs, albeit much reduced as compared to the vacuum or Ar POA case. At ~1100 °C, about $N_c \sim 3.5 \times 10^{12}$ cm^{-2} P_bs are created in He, which is to be compared with
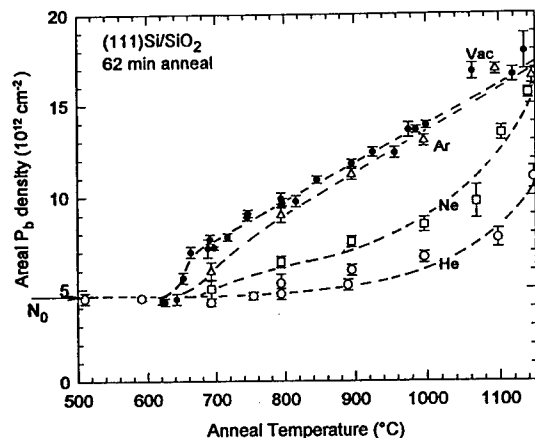
FIG. 1. Isochronal generation of ESR-active P_b defects in thermal (111)Si/SiO_2 upon POA in vacuum (•) and 1-atm He(o), Ne(□), and Ar(Δ). Each data point is obtained on a freshly oxidized Si/SiO_2 structure. The vacuum data are from Ref. 8. Curves guide the eye.
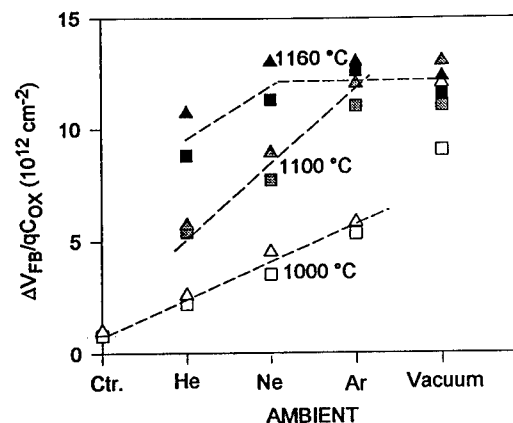
Fig. 2: Trapped hole density determined from the flatband voltage shift (ΔV_{FB}) after injection of 5×10^{14} holes/cm^2 into 60-nm thick oxides thermally grown on (111)Si (triangles) and (100)Si (squares) and annealed for 1 h in various ambients. Lines guide the eye; q is the elementary charge.

~11×10^{12} cm^{-2} created in vacuum. The blocking power of He gradually decreases for T_{an} increasing above 800 °C.

The interface degradation was studied isothermally at T_{an}~795 °C, i.e., at the upper T_{an} limit of the He blocking range. As a major result, degradation occurs fast: Common for all ambients is that P_b creation 'jumps' to saturation –when occurring– within minutes to remain unchanged for prolonged treatment up to hours. In He, the blocking still holds after 12.5 h –quite an impact. Second, and perhaps equally revealing, the saturation level $N_c + N_0$ increases with the noble gas atomic diameter σ, the data bearing out a close N_c(saturation)$\propto \sigma$ relationship.

Electrical analysis reveals that the blockage of the interface degradation is also attendant with a prominent effect on oxide degradation. This is concluded from monitoring two of the most pertinent POA-induced Si/SiO$_2$ degradation effects, i.e., generation of hole traps in the SiO$_2$ layer, responsible for enhanced radiation sensitivity, and induced low-field leakage current. Figure 2 shows the impact of the anneal ambient during 1 h POA at 1000, 1100, and 1160 °C on the positive charge generation in SiO$_2$, as sensed by the flatband voltage shift ΔV_{FB} on the 1 MHz CV curve after injection of 5×10^{14} holes cm^{-2} in the oxide. As known before[14], POA in Ar or vacuum may result in a multiple increase of hole traps as compared to the

unannealed state. We see that the degradation is weakened by Ne, while, again, He drastically suppresses it. A beneficial effect of He remains even after POA at 1160 °C, at which temperature all other ambients lead to comparable oxide damage.

The low-field leakage current (I) behavior in ultrathin oxides after POA (1000 °C; 1 h) was histogrammed for 5.8 and 4.1 nm oxides. The data were obtained using test voltages of +2 V and +1.5 V on 50 MOS capacitors of 0.4 mm^2 area, defined on n-Si(100)/SiO$_2$ by Au evaporation onto the oxide. While most of the unannealed samples exhibit negligible leaks, POA in vacuum, N$_2$, Ar, or Ne all result in severe oxide injury: No undamaged capacitor (I<10^{-10} A) is left. The effect of He, by contrast, is striking as undamaged capacitors remain after POA in He at a temperature as high as 1100 °C.

A next study addresses the impact of various ambient gases (1.1 atm) during POA on the hole trapping in (100)Si/SiO$_2$(25 nm)/poly-Si(310 nm) structures. Figure 3 shows the results for 1 h POA at 900 °C in terms of hole trap generation as reflected by the shift ΔV_{FB} after injection of 5×10^{14} holes/cm^2. Importantly, it is seen that the benign effect of the He ambient on the hole trapping properties is also observed in poly-Si covered structures, of much relevance to device processing.

4. INTERPRETATION

Modeling may be guided by the extreme inertness of noble gases which would indicate their impact to be physical rather than chemical. A plausible picture may be attained based on only two pillars: First, we admit that the degradation initiates at the Si/SiO$_2$ interface through SiO formation and that its removal controls the overall degradation process. There is little doubt that this chemical interface etching reaction does actually occur at the Si/SiO$_2$ interface, even for a thick oxide coating[7,15]. The mechanism has been successfully invoked in the explanation of various physical facts, such as thermal oxide degradation[5] and vacancy supersaturation[16] in the Si underneath SiO$_2$ during POA in Ar. The essence of the impact of intruded noble gas atoms is to be situated here, where, within a configurational picture, it might effectuate through a concerted action of various mechanisms: The latter may include suppression of interfacial bond breaking through modification of the thermal
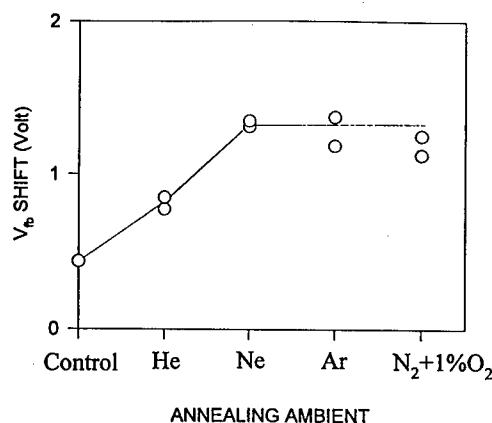
Fig. 3. Flatband voltage shift after injection of 5×10^{14} holes/cm^2 into (100)Si/SiO$_2$ (25nm)/poly-Si (310nm) structures annealed for 1 h at 900 °C in various ambients (pressure 1.1 atm). The lines guide the eye.

excitation spectrum, bond strengthening in interfacial layers through their steric presence, affecting the attempt frequency for diffusive site occupation by SiO, and impact on transport channel dynamics.

Second, there is the fact that vitreous silica, the phase closest to thermal oxide, exhibits a distribution of gas accessible sites[17] modeled by a log-normal distribution of mean (average interstice size) of 1.96 Å. The estimated density is ~2.2x10^{22} cm^{-3} based on calculations for cristobalite, a crystalline polymorph of SiO_2. The accessible fraction, however, depends on the size of the penetrating atom or molecule, respectively estimated[17,18] for He, Ne, and Ar as 10.5, 5.9, and 0.5% of the total site density.

A tentative scheme then is that upon admittance, gas atoms rapidly flood the SiO_2 film (at 600 °C, He permeates a 500-Å thick SiO_2 film in ~1 μs). From then on, the SiO removal and its adverse action is countered by the dissolved gas atoms with an efficiency inversely proportional to σ. Previously, it was concluded[19] that the degradation of the oxide at elevated T_{an}'s results from aggressive interaction with penetrating SiO. So, the presently observed blockage effect would indicate He (and, less efficiently, Ne) to counter the very source of degradation, i.e., interfacial SiO removal. As T_{an} increases above 800 °C, the He blocking power gradually weakens, possibly as a result of the formation of percolation paths for SiO. It is conceivable that the blocking efficiency will decrease the smaller $σ_{gas}-σ_{SiO}$; for inert gases, no effect is expected for $σ_{gas} \geq σ_{SiO}$. This is confirmed by the Ar case, where, with $σ_{Ar}$=3.44 Å, that is, comparable to $σ_{SiO}$~3.7-4.3 Å, the impact on degradation is only marginal, indeed (cf. Fig. 1).

5. CONCLUSIVE REMARKS

The POA-induced degradation of the thermal Si/SiO_2 entity is found to be beneficially affected by noble gases in efficiencies decreasing with increasing gas atomic number. As prime POA ambient, He is found to suppress interface degradation in terms of P_b creation up to ~800 °C for POA times up to 12 h. As to the oxide, He significantly reduces creation of intrinsic defects responsible for hole trapping and impairment of insulating properties. He also affects beneficially (100)Si/SiO_2/poly-Si structures, promising benefits

for the preservation of the oxide integrity throughout device processing. The combined beneficial effect of He on both interface and oxide degradation support the model where He efficiently suppresses *interfacial* SiO removal. Though the inertness of He would refer to a physical mechanism, detailed insight is still lacking. Likely, it concerns a configurational picture in which noble gases, in proportion to their occupation of accessible sites, affect the physical properties of the SiO_2 network.

REFERENCES

1. See, e.g., Y. C. Cheng, Progr. Surf. Sci. 8, 181 (1977) and references therein
2. D. M. Brown and P. V. Gray, J. Electrochem. Soc. 115, 760 (1968)
3. F. Montillo and P. Balk, J. Electrochem. Soc. 118, 1463 (1971)
4. B. H. Vromen, Appl. Phys. Lett. 27, 152 (1975)
5. K. Hoffmann, G. W. Rubloff, and D. R. Young, J. Appl. Phys. 61, 4584 (1987)
6. R. E. Walkup and S. I. Raider, Appl. Phys. Lett. 53, 888 (1988)
7. G. F. Derbenwick and B. L. Gregory, IEEE Trans. Nucl. Sci. NS-22, 2151 (1975)
8. A. Stesmans and V. V. Afanas'ev, Phys. Rev. B 54, R11 129 (1996)
9. A. Stesmans, Phys. Rev. B 48, 2418 (1993)
10. See, e.g., R. Helms and E. H. Poindexter, Rep. Prog. Phys. 57, 791 (1994)
11. A. Stesmans and V. V. Afanas'ev, Appl. Phys. Lett. 69, 2056 (1996)
12. M. E. Zvanut and T. L. Chen, Appl. Phys. Lett. 69, 29 (1996)
13. J. R. Schwank and D. M. Fleetwood, Appl. Phys. Lett. 53, 770 (1988)
14. A. Stesmans, V. V. Afanas'ev, and A. G. Revesz, Appl. Phys. Lett. 74, 1466 (1999)
15. Y. Takakuwa, N. Mizuhisa, and N. Miymoto, Jpn. J. Appl. Phys. 32, L480 (1993)
16. S. T. Ahn, H. W. Kennel, W. A. Tiller, and J. D. Plummer, J. Appl. Phys. 65, 2957 (1989)
17. J. F. Shackelford and J. S. Masaryk, J. Non-Cryst. Solids 30, 127 (1978)
18. J. E. Shelby, J. Appl. Phys. 47, 135 (1976); G. S. Nakayama and J. Shackelford, J. Non-Cryst. Solids 126, 249 (1990)
19. A. Stesmans and V. V. Afanas'ev, Microelectr. Eng. 36, 201 (1997)

Relation between hydrogen and the generation of interface state precursors

H.K.Sii[a], J.F.Zhang[a], R.Degraeve[b], and G.Groeseneken[b]

[a]School of Engineering, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, UK[*]

[b]IMEC, Kapeldreef 75, B3001 Leuven, Belgium

Plasma charging and electrical stresses create new interface state precursors, which accelerate the device degradation during the subsequent stress. This paper investigates the mechanism responsible for the precursor generation. The attention is focused on the roles played by the hydrogen species and the holes trapped in the oxide. The properties of the generated precursor are studied and compared with those of precursors originally existed in the device.

1. INTRODUCTION

It is well known that hydrogen plays an important role in the instability of MOS devices [1-6]. The interface state precursor is generally believed to be a hydrogen-related bond (e.g. ≡Si-H). At relatively low temperature (e.g. <200°C), this Si-H bond can be broken by hydrogen atoms, leading to the generation of interface states [1-3]. It has been proposed that this is the mechanism responsible for interface state creation under electrical stress [5] and irradiation [6]. As the temperature increases further (e.g. >350°C), however, hydrogen will anneal the interface state by restoring the bond [1]. The number of interface states at the thermal equilibrium is determined by the balance between the reactions for breaking and restoring the hydrogen bond. An implicit assumption in this model is that the total number of the defect, which is the sum of interface states and their precursors, remains a constant.

While there is no clear evidence against the above assumption when tests were carried out at room temperature, the generation of new interface state precursors has been reported when a stressed device was exposed to higher temperature (e.g., 400°C) [7-9]. Although a device rarely experiences the temperature at this level after fabrication, it is commonly used during the fabrication. For a modern CMOS process, MOSFETs are stressed by plasma charging and then annealed at high temperature to remove the damage. When the MOSFETs were electrically stressed during the operation after the device fabrication, these generated precursors can enhance the degradation significantly. It is therefore important to understand this generation process, in order to control the plasma processing induced damage. The objective of this paper is to investigate the roles played by hydrogen in the precursor generation.

2. DEVICES AND EXPERIMENTAL

The substrate hot hole injection (SHI) technique was used to stress the pMOSFETs fabricated by a 0.35μm CMOS process. The oxide was grown at 900°C to a thickness of 7.1nm. After a pre-set time, the injection was interrupted and the device was exposed to an elevated temperature with all terminals floating, either in N_2 or forming gas (FG, 10%H_2). The SHI was then resumed. The interface state density (Nit) and oxide charge density (ΔNot) were monitored during the experiment by the charge pumping and subthreshold I-V measurements, respectively.

3. RESULTS

3.1 Roles played by hydrogen

Fig.1 shows a typical result. Two fresh pMOSFETs were used here. After a hole injection of 5,000sec (the 1st SHI), the two devices were exposed to 400°C in N_2 and FG, respectively. The hole injection was then resumed (the 2nd SHI). It can be seen that the number of interface states generated during the 2nd SHI is much larger than that during the 1st SHI. This indicates that new precursors were created. To facilitate the description, the precursor responsible for the interface state generated during the 1st SHI will be referred to as the "original

precursor" and the defect responsible for the enhanced generation during the 2nd SHI will be called as "generated precursor", hereafter.
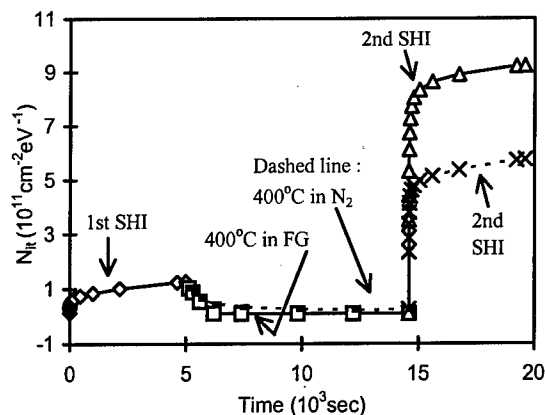
Fig.1 Effects of ambient gases on the precursor generation. Both the 1st and the 2nd SHI were carried out at Eox=-5MV/cm, Vwell=8.8V and Vsub=10V.

Fig.1 also shows that the external supply of H_2 at 400°C enhances the precursor generation. It is not clear, however, if the generation in N_2 is also related to hydrogen. Although hydrogen was not externally supplied during the N_2 exposure, it can be pre-stored in the device and released at 400°C, leading to the precursor generation.
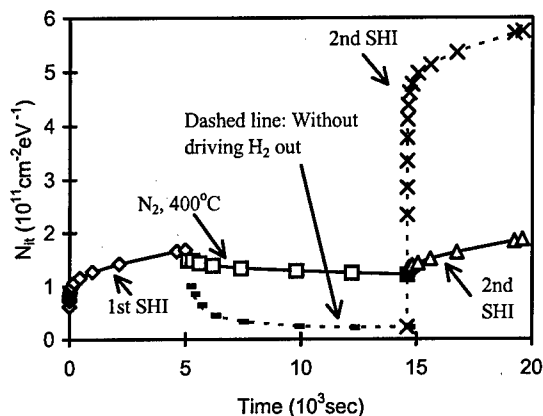
Fig.2 Solid line: the device was annealed in N_2 at 450°C for 17hours before the 1st SHI to drive out the hydrogen. Dashed line: without this anneal.

In Fig.2, it is attempted to drive the pre-stored hydrogen out of the device by a long anneal in N_2 (450°C, 17hours) before the 1st SHI. The result unambiguously shows that this long pre-stress N_2 annealing effectively suppressed the precursor generation. We conclude that the presence of

hydrogen is a necessary condition for the precursor generation.

3.2 Roles played by the trapped holes

Although hydrogen is needed for precursor generation, it is not clear whether the presence of H_2 is a sufficient condition. To answer this question, one nMOSFET was stressed by the substrate hot electrons injection (SHE) and then exposed to FG at 400°C. Fig.3 shows that the interface state generation during the 2nd SHE is essentially the same as that during the 1st SHE. There is little precursor generation after electron injection, even if H_2 was made available at 400°C. Thus, exposing a stressed device to H_2 does not guarantee the creation of precursors. The stress condition also plays a key role. The results shown in Figs.1&3 indicate that the cause for precursor generation is the hole injection.
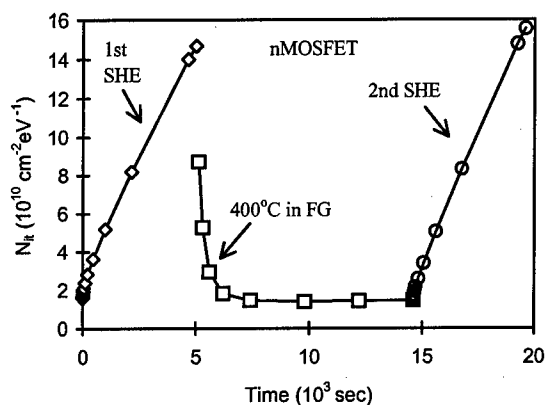
Fig.3 Effects of the substrate hot electron injection (SHE). Eox=3.5MV/cm during the SHE.

To find why precursors can be created after hot hole injection, the detrapping of trapped holes is studied. In Fig.4a, the hydrogen was first driven out of the device before the 1st SHI. After the 1st SHI, the device was exposed to N_2 at 400°C and the detrapping was considerable through tunnelling. The rapid decline of detrapping rate is caused by the increase of tunnelling distance between traps and the substrate. If the 2nd SHI were started at the end of this N_2 exposure, the solid line in Fig.2 showed that the generation of precursors were negligible. Detrapping of trapped holes does not necessarily lead to precursor generation, therefore.

If we switch the gas from N_2 to FG, Fig.4a shows that the detrapping rate is significantly enhanced and the detrapping can approach 100% now. The presence of H_2 introduces an additional detrapping mechanism, therefore. It has been proposed that the trapped holes can be detrapped by cracking the H_2 and one product of this reaction is a hydrogen ion

(H$^+$) or a hydrogen atom (H°) [2]. Fig.4b shows that the precursor generation is significant, after the gas switching. It can be concluded that the H$_2$ induced detrapping leads to the creation of precursors, while the tunnelling induced detrapping does not. We believe that H$_2$ cracking by trapped holes also occurs here and the resultant H$^+$/H° causes the precursor generation. Thus, the role played by the trapped holes is supplying the H$^+$/H° by cracking H$_2$.
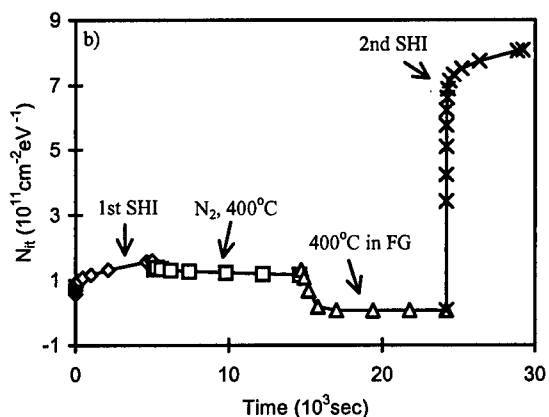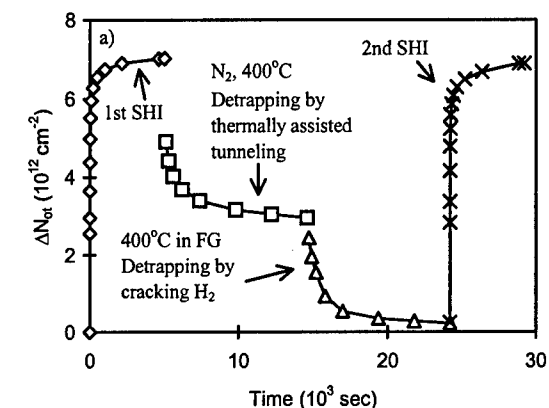
Fig.4 Behaviour of trapped holes (a) and interface states (b), when the gas was switched from N$_2$ to FG. Before the 1st SHI, the hydrogen pre-stored in this device was driven out.

3.3 Properties of generated precursors

As mentioned earlier, the hydrogen required for precursor generation can be removed from the device by annealing in N$_2$ at 450°C for 17 hours. These hydrogen species are unstable at 450°C, therefore. However, the hydrogen bond responsible for the original precursor changes little during this anneal. This indicates that hydrogen can be stored within the device in different ways. Since the

presence of hydrogen is essential for creating precursor, it is reasonable to assume that the generated precursor also contains a hydrogen-related bond. It is interesting to investigate whether this bond can survive the above anneal.
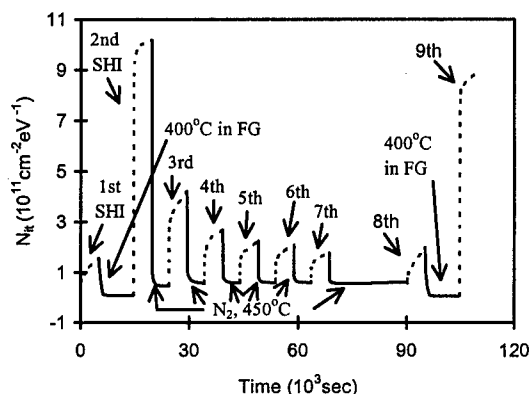
Fig.5 Precursors were created by exposing the device to FG after the 1st SHI. The device was then annealed at 450°C for 17 hours in N$_2$ before the 2nd SHI, which was not shown in this figure.

In Fig.5, precursors were generated by exposing a device to the FG at 400°C after the 1st SHI. The device was then subjected to the 17 hour N$_2$ anneal at 450°C. The interface states generated during the subsequent 2nd SHI is not reduced, when compared with the 2nd SHI in Fig.1 which did not experience the 17 hour anneal. Once a precursor is generated, it is thermally stable at least up to 450°C, therefore.

In a previous investigation [4], precursor generation was observed when a sample was exposed to H$_2$ at a temperature over 550°C without electrical stress. It was reported that the original and the generated precursor have the same properties. Here we observed that both the original and the generated precursors involve hydrogen bonds and both are thermally stable at 450°C. However, whether these two are the same should be investigated further.

In Fig.2, some of the original precursors were converted into interface states during the 1st SHI. In the subsequent thermal exposure, their annealing rate is drastically reduced, if the hydrogen was driven out of the device. This indicates that the restoration of original precursors can only take place when hydrogen is available. The defect responsible for the original precursor can only exist in two forms: either as a precursor with hydrogen or as an interface state without hydrogen.

To find if the above applies to the generated precursor, the SHI/N$_2$ exposure (400°C) sequence was cycled several times after the 2nd SHI. Fig.5

shows that, as the cycle number increases, the interface state generated during the SHI decreases and approaches the level reached during the 1st SHI. This means that the number of generated precursors is reducing. The hydrogen pre-stored in the device has been driven out here, during the 17hour N_2 exposure at 450°C immediately before the 2nd SHI. Thus, the removal of generated precursors is achieved without supplying hydrogen. The defect responsible for the generated precursor can be transformed from an interface state into a "non-precursor", which can not be converted back into an interface state during the following SHI. This, however, does not mean that the defect has been eliminated from the device. Once hydrogen is supplied, the non-precursor can be converted into a generated precursor again (see the 9th SHI in Fig.5). Unlike the defect responsible for the original precursor, the defect responsible for the generated precursor can exist at least in three forms: a precursor with hydrogen, an interface state and a non-precursor without hydrogen.

4. DISCUSSIONS

On the structure of generated precursors, the following model has been proposed [7]:
a. The stress breaks the strained ≡Si-O-Si≡ bond into ≡Si-O• and •Si≡;
b. The ≡Si-O• migrates to the interface at the elevated temperature and reacts with hydrogen to form ≡Si-OH, which is the generated precursor;
c. During the subsequent stress, the hydrogen bond is broken and the precursor ≡Si-OH is converted into an interface state, ≡Si-O•.

Although this model correctly predicts that the generated precursor contains a hydrogen bond, it does not agree with the present result on the following aspects:
a. According to this model, the ≡Si-O• has to react with hydrogen at the interface to form a precursor. In the case that there is a lack of hydrogen species, one would expect a significant increase of interface states during the 400°C exposure, because of the pile-up of ≡Si-O• at the interface. This is against the solid lines in Fig.2.
b. This model does not predict that the defect can exist in a non-precursor form.
c. The H_2 cracking by trapped holes is not included in this model.

At this stage, we do not have enough information to predict the microstructure of the defect. However, the reported data have improved our understanding of the generation process, as is shown here:

Stress Defects ——► Non-precursors ——High Temp. in hydrogen——► Precursors

Stress ——► Interface States —— High Temp. in hydrogen ——► Precursors / High Temp. without hydrogen ——► Non-precursors
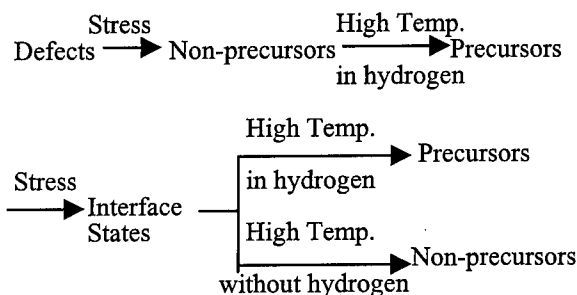
Since a non-precursor is electrically inactive, it is unlikely that it is a broken bond at the interface. Since it can be formed by annealing an interface state without hydrogen supply, it is unlikely that the non-precursor involves a hydrogen bond. These suggest that stress changes the properties of an interfacial defect without creating a broken bond. This defect does not have a hydrogen bond initially, but it can react with atomic hydrogen to form a precursor. It is likely that the defect involves some form of strained Si-O bonds.

5. CONCLUSIONS

Hydrogen is an essential reactant for the interface state precursor generation. The defect responsible for the precursor generation does not react with H_2 directly and the H_2 has to be cracked by the trapped hole into H^+/H° first. Unlike the original precursors, the generated precursor can be converted into a non-precursor form. This indicates that the generated precursor has a microstructure different from that of original precursors.

REFERENCES

1. L.Do Thanh and P.Balk, J.Electrochem.Soc., **135**, p.1797, 1988.
2. R.E.Stahlbush, A.H.Edwards, D.L.Griscom and B.J.Mrstik, J.Appl.Phys., **73**, p.658, 1993.
3. E.Cartier and J.H.Stathis, Microelectronic Eng., **28**, p.3, 1995.
4. A.Stesmans and V.V.Afanas'ev, Appl. Phys. Lett., **72**, p.2271, 1998.
5. D.J.DiMaria and J.W.Stasiak, J.Appl.Phys., **65**, p.2342, 1989.
6. N.S.Saks, C.M.Dozier, and D.B.Brown, IEEE Trans. Nucl.Sci., **35**, p.1168, 1988.
7. M.Berger, O.Schwartsglass, and J.Shappir, J.Electrochem.Soc., **141**, p.2140, 1994.
8. J.F.Zhang, I.S.Al-kofahi, and G.Groeseneken, J.Appl.Phys., **83**, p.843, 1998.
9. M.Creusen, H.C.Lee, S.Vanhaelemeersch, and G.Groeseneken, Proc. Plasma Induced Damage Conf., p.217, 1998.

Hydrogen migration in wet-thermally grown silicon dioxide layers due to high dose ^{15}N ion beam irradiation

K. Maser[a], U. Mohr[a], R. Leihkauf[a], K. Ecker[b], U. Beck[b], D. Grambole[c], R. Grötzschel[c], F. Herrmann[c], J. Krauser[d], and A. Weidinger[d]

[a]Technische Universität Berlin, Institut für Mikroelektronik und Festkörperelektronik, Jebensstr. 1, D-10623 Berlin, Germany, Fax: ++49 (0)30 314 24597, Email: leihkauf@imf.ee.TU-Berlin.de

[b]Bundesanstalt für Materialforschung und -prüfung (BAM), Unter den Eichen 87, 12205 Berlin, Germany

[c]Forschungszentrum Rossendorf, Institut für Ionenstrahlphysik und Materialforschung, Dresden, Germany

[d]Hahn-Meitner-Institut Berlin, Glienicker Str. 100, D-14109 Berlin, Germany

Three effects have been observed in wet-thermally grown SiO_2/Si structures under bombardement with energetic ^{15}N ions:
(1) Out-diffusion of H from the SiO_2/Si system through the surface,
(2) Accumulation of H at the SiO_2/Si interface,
(3) Changes in the optical fingerprint spectra of Ψ and Δ in dependence on λ (wavelength) obtained by ellipsometric measurements.

1. INTRODUCTION

Hydrogen dynamics in wet-thermally grown SiO_2/Si layers are of interest for modern electronic devices. Nuclear Reaction Analysis (NRA) using the ^{1}H(^{15}N,$\alpha\gamma$)^{12}C nuclear reaction is a powerful method to determine hydrogen in thin layers. However due to the irradiation of the sample with energetic ^{15}N ions during the measurement the distribution of hydrogen can be changed. It was the aim of the present investigations to clarify some aspects of the system H in SiO_2 on Si under irradiation conditions which occur during H depth profiling with this method.

2. EXPERIMENTAL DETAILS

Samples are prepared as follows: CZ-Si crystal, 100 mm diameter, (100)-orientation, p-type (B), 1-10 Ωcm, IC quality, standard RCA cleaning, oxidation at 700 °C in a water steam atmosphere to a 173 nm thick oxide layer. Some of the samples were subsequently thinned by chemical etching. The distribution of hydrogen in the samples was investigated by NRA measurements at temperatures around 300 K. Repeated measurements were performed on the same sample, so that the development of H profiles can be observed as a function of the actual ^{15}N ion dose Q.

Oxide thickness and information on degradation effects were obtained from ellipsometric measurements. Amplitude ratio $\Psi(\lambda)$ and phase shift upon reflection $\Delta(\lambda)$ were used in terms of fingerprint functions, which are related to the electronic structure of the material and design features of the coating-substrate system. A rotating polarizer ellipsometer and a rotating analyzer ellipsometer took data every 5nm in the wavelength

range from 300 nm to 800 nm, and from 270 nm to 1100 nm respectively. Xenon white light sources result in effective spot areas in the mm^2 range.

3. RESULTS

Fig. 1 shows the γ-ray yield for a sample with a 173 nm thick oxide layer as a function of the energy loss ΔE, the difference of the actual ion beam energy to the resonance energy (6.385 MeV) of the $^1H(^{15}N,\alpha\gamma)^{12}C$ nuclear reaction. Three profiles are depicted for runs with increasing total irradiation doses. The symbols show measured values, solid lines are the result of convolution analysis as described in [1,2]. For clearer view, the surface peaks due to surface adsorbates at ΔE = 0 are not plotted. The first run of measurements limited to Q = 6.2×10^{14} $^{15}N^+/cm^2$ gives an almost constant H concentration of about 2.2×10^{20} cm^{-3} throughout the SiO$_2$ layer. At such a low irradiation dose we assume that this profile is close to the H distribution of the unirradiated sample. The last run of NRA referring to the dose interval Q = 5.3 - 7.6×10^{16} $^{15}N^+/cm^2$ shows an increase in the H concentration in the SiO$_2$/Si transition region as well as a marked decrease of the H concentration inside the oxide layer.
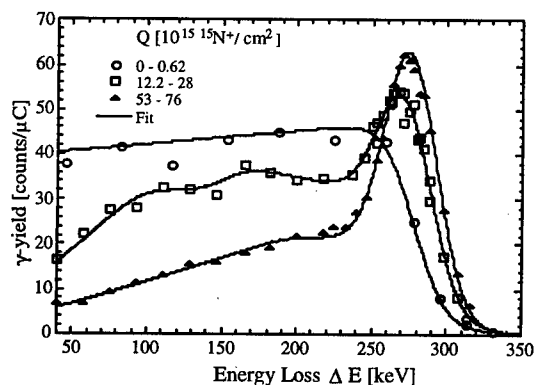
Fig. 1: Measured γ-yield versus ΔE for several $^{15}N^+$ ion dose intervals for a sample with a 173 nm thick SiO$_2$ layer. The solid lines are the result of a convolution analysis.

This loss of hyrogen from the SiO$_2$ region and the accumulation at the SiO$_2$/Si interface as a function of the ^{15}N ion irradiation dose are shown in Fig. 2.

The points are obtained from deconvoluted profiles. The lines are only intended to guide the eyes.
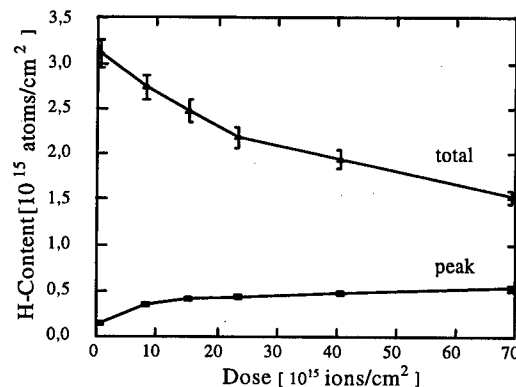
Fig. 2: Integral hydrogen content in the total SiO$_2$ layer and in the SiO$_2$/Si transition region (the latter is assumed from 172 to 177 nm) as a function of the ^{15}N ion dose.

Figure 3 refers to an oxide layer which was wet - chemically thinned from 173 to 32 nm. The γ-yield of the sample irradiated with ^{15}N ions with energy 6.438 MeV is shown as a function of the ion dose. This energy corresponds to a depth of 34 nm assuming an energy loss of 1.54 keV/nm [2] in the SiO$_2$. Although in this case no convolution of profiles could be performed the γ-yield obtained with this irradiation energy is a qualitative measure of the H-content in the transition region. Initially, an accumulation of interfacial hydrogen by the ion irradiation is observed as in the thicker sample (Fig. 2). However, at higher irradiation doses the H-content at the interface drops again slowly.
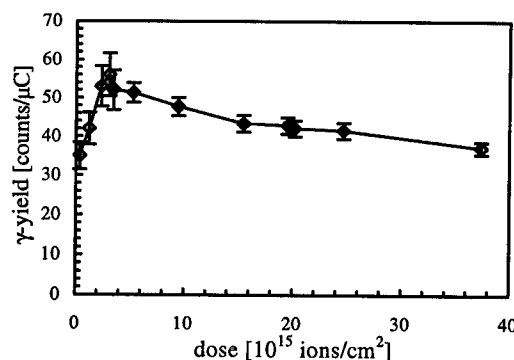
Fig.3: Interfacial γ-yield of the thinned oxide layer irradiated with 6.438 MeV ^{15}N ions as a function of ion dose.

Figures 4a and 4b show fingerprint spectra of the ellipsometric quantitities $\Psi(\lambda)$ and $\Delta(\lambda)$ for the sample with an oxide thickness of 173 nm prior to and after ^{15}N ion irradiation with a dose of 7×10^{16} ions/cm^2. The difference of these two functions on λ is significant. No significant effects were found for samples with lower oxide thickness (wet-chemically thinned to 90 nm, and 32 nm respectively). Hence, for thicker oxide films, there is a degradation effect in the optical functions as a result of ^{15}N ion irradiation. Calculating the oxide thickness based on a simple one-layer model (Jellison data) the uncertainity of the measurement is about 1 nm.
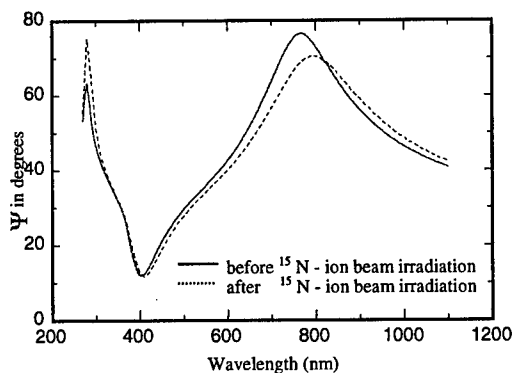
Fig. 4a: Fingerprint function $\Psi(\lambda)$ before and after ^{15}N ion irradiation. Angle of incidence: 75°.
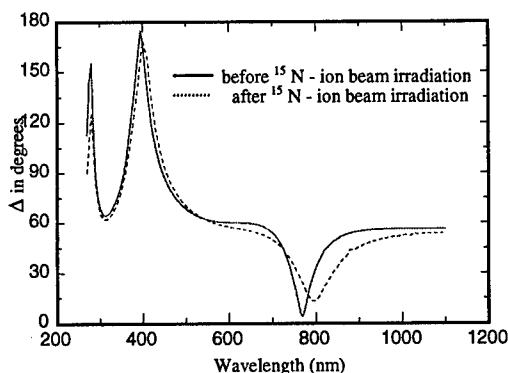
Fig. 4b: Fingerprint function $\Delta(\lambda)$ before and after ^{15}N ion irradiation. Angle of incidence: 75°.

4. DISCUSSION

In view of the present measurements the interpretation of previous investigations with high ^{15}N doses [3-5] have to be somewhat revised. The analysing beam undoubtedly changes the hydrogen profiles. The accumulation of hydrogen near the SiO_2/Si transition zone and the loss of hydrogen from the oxide layer are related to the phenomenon that the initially fixed H in the oxide becomes mobile by the energetic irradiation of the analysing beam [1,2]. The released hydrogen migrates to the surface and to the interface region. The surface acts as an unsaturable sink where hydrogen is lost. The interface region should contain saturable trap sites for hydrogen to explain the observed increase in H-content in this region. In the case of thermally grown SiO_2/Si(100) transition zones, H-trapping sites may arise from following reasons:
(a) Compositional modifications involving the suboxides Si_2O, SiO and Si_2O_3 [6-8],
(b) a mass density gradient as well as interfacial strain arising from the lattice mismatch between Si and SiO_2 [9,10],
(c) structural defects forming during the oxidation process, e. g. trivalent silicon atoms (silicon dangling bonds, P_b center), non-bridging oxygen atoms (oxygen dangling bonds), strained (or weak) Si-Si bonds (E' center), strained (or weak) Si-O bonds [11-17],
(d) irradiation-induced structural damages and defects arising from measuring methods and technological steps [18-25] .
Recent stress measurements on SiO_2/Si-structures prepared very similar to those of the present investigations indicate the highest oxide stress levels (around 1GPa) near the interface of these SiO_2/Si-sandwiches [26]. It is supposed, that hydrogen accumulates preferentially there to reduce the corresponding strain. Prolonged irradiation again reduces the interfacial H-content (see Fig. 3). This result indicates that the traps at the interface are shallow enough to enable a small release by ion irradiation. The change in the hydrogen profile might be responsible for the change in the ellipsometric data. Other phenomena (e.g. change of the oxide density, increasing surface roughess) would occur also for thinner oxides and would even more influence ellipsometric data. A detailed ellipsometric analysis would require a model

calculation including the H-distribution in the SiO_2. This, in fact, would require also a multi-sample analysis.

The dynamics of hydrogen and the energy levels of the traps are still unsolved questions. It is suggested, that the interfacial H-accumulation can be explained at least qualitatively using a non-Fickian dopant (H) transport model proposed in [27,28]. Accordingly, two non-Fickian H-flux components are responsible for the development of the H-peak. They are thought to arise from both, a maximum concentration of trap sites as well as a minimum dopant jump rate, which are assumed to be located in the SiO_2/Si transition zone. Problems, however, remain to develop the overall functional dependence of these two parameters on position so that up to now no numerical check of this theory has been done.

In order to obtain more insight into micro-structural and/or stoichiometric changes in SiO_2/Si structures during ^{15}N ion irradiation further work is needed.

5.CONCLUSION

Since the analysing beam of ^{15}N ions undoubtedly changes the hydrogen profiles in SiO_2 layers, the lowest irradiation dose should be taken in order to minimize measurement artefacts.

REFERENCES

1. M. A. Briere, F. Wulf, and D. Bräunig, Nucl. Instr. and Meth. in Phys. Res., B 45 (1990) 45
2. M. A. Briere, Berichte des Hahn-Meitner- Institutes HMI-B 509 (1993)
3. K. Maser, H. H. Berger, J. Krauser, and D. Bräunig, Exp. Tech. Phys. 41 (1995)101
4. K. Maser, J. Krauser, H. Berger, and D. Bräunig, Microelectron. Engg. 28 (1995) 129
5. J. Krauser, A. Weidinger, and D. Bräunig. In:The Physics and Chemistry of SiO_2 and the Si-SiO_2 Interface, Proc. Vol. 96-1(1996)184
6. H. Fukuda, M. Yasuda, T. Iwabuchi,
7. S. Kaneko,T. Ueno, and I. Ohdomari, J. Appl. Phys. 72 (1992) 1906
7. H. Ono, T. Ikarashi, K. Ando, and T. Kitano, J. Appl. Phys. 84 (1998) 6064
8. F. J. Himpsel, F. R. McFeely, A. Taleb-Ibrahimi, J. A. Yarmoff, and G. Hollinger, Phys. Rev. B, 38 (1988) 6084
9. B. J. Mrstik, and P. J. McMarr, Phys. Rev. B, 48 (1993) 17972
10. T. Hattori. In: The Phsics and Chemistry of SiO_2 and the Si- SiO_2 Interface-3. Proc. Vol. 96-1 (1996) 392
11. Y. Nishi, Japan. J. Appl. Phys. 10 (1971) 52
12. G. J. Gerardi, E. H. Poindexter, P. J. Caplan, and N. M. Johnson, Appl. Phys. Lett. 49 (1986) 348
13. T. Hori, Gate Dielectrics and MOS ULSIs.-Springer (1997)
14. M. L. Reed, and J. D. Plummer, J. Appl. Phys. 63 (1988) 5776
15. P. J. Caplan, E. H. Poindexter, B. E. Deal, and R. R. Razouk, J. Appl. Phys. 50 (1979) 5847
16. D. Sands, K. M. Brunson, and M. H. Tayarani-Najaran, Semicond. Sci. Technol. 7 (1992) 1091
17. E. H. Poindexter, G. J. Gerardi, and D. J. Keeble. In: The Physics and Chemistry of SiO_2 and the Si- SiO_2 Interface-3. Proc. Vol. 96-1 (1996) 172
18. O. O. Awadelkarim, S. J. Fonash, P. I. Mikulan, and Y. D. Chan, J. Appl. Phys. 79 (1996) 517
19. S. J. Fonash, J. Electrochem. Soc. 137 (1990) 3885
20. W. L. Warren, E. H. Poindexter, M. Offenberg, and W. Müller- Warmuth, J. Electrochem. Soc. 139 (1992) 872
21. M. J. Uren, K. M. Brunson, and A. M. Hodge, Appl. Phys. Lett. 60 (1992) 624
22. R. A. Sigsbee, and R. H. Wilson, Appl. Phys. Lett. 23 (1973) 541
23. C. W. Gwyn, J. Appl. Phys. 40 (1969) 4886
24. R. Gale, F. J. Feigl, C. W. Magee, and D. R. Young, J. Appl. Phys. 54 (1983) 6938
25. A. Losavio, B. Crivelli, F. Cazzaniga, M. Martini, G. Spinolo, and A.Vedda, Appl. Phys. Lett. 74 (1999) 2453
26. R. Leihkauf, U. Mohr, unpublished
27. K. Maser, Exp. Tech. Phys. 39 (1991) 169
28. K. Maser, Appl. Phys. Lett. 63 (1993) 2576

Relationship between hole trapping and oxide density in thermally grown SiO$_2$

B. J. Mrstik[a], V.V. Afanas'ev[b], A. Stesmans[b], and P.J. McMarr[c]

[a] Naval Research Laboratory, 4555 Overlook Ave., Washington, DC 20375, USA

[b] Dept. of Physics, University of Leuven, Celestijnenlaan 200D, 3001 Leuven, Belgium

[c] SFA, Inc., Landover, MD 20785, USA

SiO$_2$ films were grown in dry oxygen on Si (100) substrates at various temperatures, then annealed for various times and temperatures. The density of the oxides was found to depend on the growth temperature, and on the anneal time and temperature. The amount of positive charge trapped in these oxides during vacuum ultraviolet hole injection was found to be inversely related to the oxide density, regardless of how the oxide was prepared. Comparison of the density of trapped holes with the density of E' centers detected by electron spin resonance indicates that E' centers cannot account for most of the trapped positive charge. A model to explain these observations is proposed.

1. INTRODUCTION

SiO$_2$ films prepared by oxidizing Si substrates in dry oxygen trap positive charge when x-rayed or when injected with holes. Although it is well known that the extent of charge trapping is enhanced by high growth temperatures or by post-oxidation anneal (POA), there is no consensus as to the mechanism for this enhancement. In an early model of EerNisse and Derbenwick [1] the presence of charge traps was attributed to the breaking of bonds during the relaxation of the oxide by viscous flow. Kim and Lenahan [2] suggested that a major hole trap is the E' center, which is an oxygen vacancy in SiO$_2$. Warren et al. [3] proposed that oxygen vacancies result from the outdiffusion of oxygen during high temperature processing. Lenahan et al. [4] suggested that the increase in the number of hole traps during POA can be explained by statistical thermodynamics.

In this paper it is shown that the hole trapping efficiency can be correlated to the oxide density, ρ_{ox}, and that ρ_{ox} is affected by the growth temperature, T_g, and by the POA time and temperature. It is also shown that the increase in hole trapping efficiency following high temperature anneals occurs *without* a concomitant increase in the number of E' centers.

2. EXPERIMENTAL DETAILS

Oxides were grown on (100) oriented p-type Si substrates in dry oxygen at temperatures of 900, 975, 1050, 1125, and 1200 °C. The growth times were adjusted to produce oxides approximately 120 nm thick. The oxides were then annealed in Ar at 900, 975, 1050, or 1125 °C for 0, 10, 60, or 360 min.

The density and thickness of each oxide were determined from spectroscopic ellipsometry measurements at an angle of incidence of 61.42° using the procedure described previously [5]. Briefly, the measured values of the ellipsometric parameters Δ and Ψ are compared to values calculated assuming various oxide structures. For these oxides, the best agreement was obtained by assuming that the oxide was a homogeneous layer of densified SiO$_2$.

In preparation for hole trapping measurements, 15 nm thick Au dots 0.005 cm^2 in area were evaporated onto the oxides. Electron-hole pairs were formed in the top ~10 nm of the oxide by shining vacuum ultraviolet (VUV) irradiation through the Au dots. 10^{14} holes/cm^2 were injected into the oxide by applying a positive voltage to the Au dot to create a field of 1 MV/cm in the oxide. The resulting buildup of positive charge in the oxide was monitored using 1 MHz capacitance-voltage measurements to determine the (negative) shift in

the flatband voltage, ΔV_{fb}. ΔV_{fb} has two components, one arising from hole trapping in the oxide, and one from generation of interface traps. To separate these two components, UV light was used to photoinject electrons into the oxide to neutralize the trapped holes. The remaining shift in the flatband voltage, $\Delta V_{fb(int)}$, is due to interface traps, and the shift $\Delta V_{fb} - \Delta V_{fb(int)}$ is due to hole trapping in the oxide.

Electron spin resonance (ESR) measurements were used to determine the number of paramagnetic defects. Measurements were made at 4.3 K using conventional absorption mode ESR at 20.1 GHz, as described previously [6]. 10^{14} holes/cm^2 were injected by shining VUV light through a Au film positively biased to produce an oxide field of 2 MV/cm. The Au film was then removed, and ESR measurements were made to determine the number of defects that were paramagnetic in the charged state. The trapped charge was then neutralized using electron photoinjection, and ESR measurements were repeated to determine the number of defects that were paramagnetic in the neutral state.

3. RESULTS

All the oxides were found to be denser than fused silica. The oxide density was found to depend on T_g, and on the POA time and temperature. Figure 1 shows ρ_{ox} of as-grown oxides as a function of T_g. Oxides grown at 900 °C are seen to have a density about 3.2% greater than fused silica (%gfs). As T_g is increased, ρ_{ox} decreases, at least up to ~1100 °C. A similar decrease of the refractive index with growth temperature was observed earlier by Taft [7]. For $T_g \leq 1050$ °C, POA treatments decrease ρ_{ox}, as shown in Fig. 2 for an oxide originally grown at 900 °C. After high temperature POA, or very long POA at lower temperatures, ρ_{ox} becomes similar to that of oxides initially grown at large T_g. A similar decrease in the oxide refractive index after POA was reported earlier on the basis of single wavelength ellipsometry [8,9]. These results indicate that the oxide relaxes during the POA until it attains the density of an oxide grown at high T_g. Infrared measurements [10] indicate that low T_g oxides have smaller Si-O-Si bridging bond angles than bulk oxide, so the structural relaxation (decrease in ρ_{ox}) occurring during POA likely involves an increase in this bond angle.
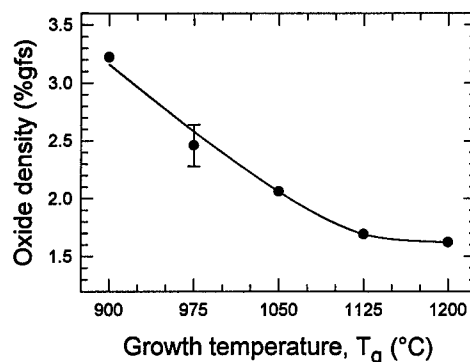
Figure 1. Density of as-grown oxides as function of growth temperature.
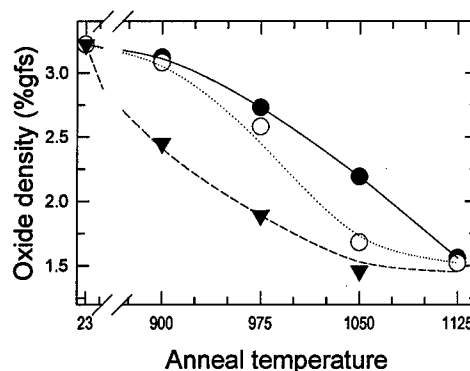
Figure 2. Density of oxide grown at 900 °C, then annealed at various temperatures for 10 (●), 60 (o), or 360 (▼) min.

Positive charge trapped in the oxide after injection of 10^{14} holes/cm^2 results in a negative shift of the flatband voltage. Figure 3a shows that as ρ_{ox} is decreased by increasing T_g, ΔV_{fb} is increased. Figure 3b shows that for oxides grown at 900 °C, as ρ_{ox} is decreased by various POA treatments, ΔV_{fb} also increases. Finally, Fig. 3c shows ΔV_{fb} as a function of ρ_{ox} for all the oxides studied. From these data, it is clear that ΔV_{fb} is correlated to ρ_{ox}, regardless of the specific fabrication conditions by which the oxide attained that density.

Contributions to ΔV_{fb} due to interface traps and to oxide traps can be separated as discussed above. From the individual voltage shifts, the density of interface traps, N_{it}, and the density of oxide traps, N_{ot}, can be determined. The results, shown in Fig. 4, indicate that N_{it} and N_{ot} *both* are correlated with ρ_{ox}, and, again, not on the specific fabrication conditions by which the oxide achieved that value of ρ_{ox}.
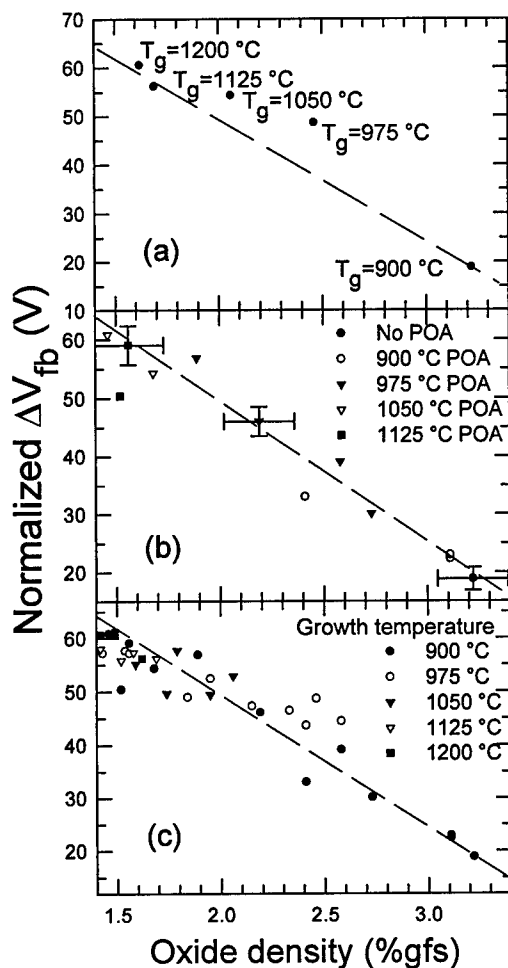
Figure 3 Dependence of ΔV_{fb} on oxide density for a) as-grown oxides, b) oxides grown at 900 °C then annealed for 10, 60, or 360 min at indicated temperature, and c) oxides grown at indicated temperature, then annealed for 0, 10, 60, or 360 min at 900, 975, 1050, or 1125 °C.

Figure 4. Dependence on oxide density of (a) number of oxide traps and (b) number of interface traps.

The only paramagnetic defect observed by ESR was the E' center. Figure 5 shows the number of paramagnetic defects observed in the oxide grown at 900 °C then either annealed at higher temperatures for 1 h, or left unannealed. The number of defects that are paramagnetic in the positive charge state and in the neutral state are shown, along with the total. Also shown is the number of positively charged defects determined electrically. Note that the number of paramagnetic defects is essentially unchanged by the anneal, whereas the number of defects observed

electrically is increased by the anneal. Also note that, except for the unannealed sample, the total number of defects observed electrically is much larger than the total number of paramagnetic defects. Similar trends were also observed for unannealed samples as T_g was increased from 900 °C to 1125 °C.

4. DISCUSSION

The above results indicate that hole trapping in thermal oxides is correlated with the macroscopic structure of the oxide, as characterized by ρ_{ox}. In this section a model is proposed to explain this correlation.

Figure 4 indicates that N_{ot} and N_{it} are *both* correlated to ρ_{ox}, suggesting that a similar mechanism might be responsible for charge trapping in the bulk and at the interface. It has been shown [11] that interface traps formed as a result of VUV irradiation are donor-type states that are stable when positively charged, but that decay when neutral. It has been suggested [12] that these states result from the trapping of hydrogen at network oxygen sites near the interface. When positively charged to form
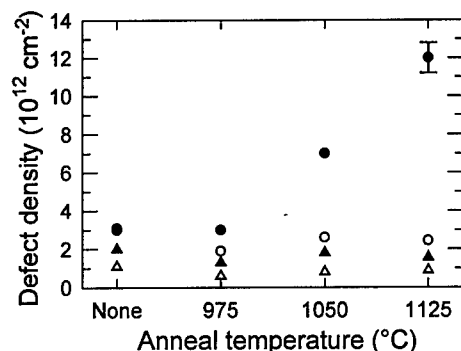
Figure 5. Dependence of defect density on anneal temperature for oxide grown at 900 °C. Density of trapped holes (●) and total paramagnetic defects (o) are shown. Also shown are density of defects that are paramagnetic when the oxide is positively charged (□), or when neutral (▲).

$\equiv$Si-OH$^+$-Si$\equiv$, these sites are stable. But when neutral, the H atom is released into the network where it may be dimerized by interaction with another H atom, and the donor-type states anneal.

Similar trapping of H (and subsequently a hole) at $\equiv$Si-O-Si$\equiv$ sites in the oxide could explain the observed dependence of N_{ot} on ρ_{ox}, since it has been suggested by Edwards and Germann [13] that H bonding at network O sites increases as the Si-O-Si bond angle increases, i.e., as ρ_{ox} decreases. Such trapped charge would be stable unless neutralized by an injected electron, which would cause the trapped H to be released.

This model is consistent with the ESR results discussed above that indicate that much of the oxide charge is trapped at sites that are *not* E' centers. It is also consistent with the ESR observation that no other paramagnetic defects are seen when the oxide is *either* positively charged *or* neutral: When charged, the $\equiv$Si-OH$^+$-Si$\equiv$ defect is diamagnetic, and therefore not seen by ESR. After neutralization the atomic H released by electron capture quickly dimerizes, so no ESR signal is observed after neutralization either.

5. SUMMARY

It has been shown that as-grown thermal oxides are denser than fused silica, and that ρ_{ox} decreases as T_g is increased. It has also been shown that ρ_{ox} can

be decreased by POA to a value similar to that of oxides grown at very high temperatures.

Both N_{ot} and N_{it} have been shown to be related to ρ_{ox}, regardless of the oxide's specific fabrication conditions (T_g, or POA time and temperature). The number of E' centers, on the other hand, has been found to be relatively insensitive to these processing parameters, and, in most cases, smaller than the number of electrically observed defects.

A model has been proposed to explain the observed dependence of hole trapping on oxide density. In this model, which is similar to that proposed earlier for donor-type interface states formed by VUV irradiation, holes are trapped at network oxygen sites that have captured a hydrogen atom.

REFERENCES

1. E.P. Eernisse and G.F. Derbenwick, IEEE Trans. Nucl. Sci. NS-23, 1534 (1976).
2. Y.Y. Kim et al., J. Appl. Phys. 64, 3551 (1988).
3. W. L. Warren et al., Appl. Phys. Lett. 64, 3452 (1994).
4. P.M. Lenahan et al., J. Appl. Phys. 81, 6822 (1997).
5. B.J. Mrstik, P.J. McMarr, J.R. Blanco, and J.M. Bennett, J. Electrochem. Soc. 138, 1770 (1991).
6. A. Stesmans, Phys. Rev. B 48, 2418 (1993).
7. E.A. Taft, J. Electrochem. Soc. 127, 993 (1980).
8. K. Taniguchi, M. Tanaka, C. Hamaguchi, and K. Imai, J. Appl. Phys. 67, 2195 (1990).
9. L.M. Landsberger and W.A. Tiller, Appl. Phys. Lett. 51, 1416 (1987).
10. G. Lucovsky, M.J. Martinet, J.K. Srivastava, and E.A. Irene, J. Vac. Sci. Technol B 5, (1987).
11. K.G. Druijf, J.M.M. de Nijs, E. van der Drift, E.H.A. Granneman, and P. Balk, Appl. Phys. Lett. 65, 2428 (1994).
12. J.M.M. de Nijs, K.G. Druijf, V.V. Afanas'ev, E. van der Drift, and P. Balk, Appl. Phys. Lett. 65, 347 (1994).
13. A.H. Edwards and G. Germann, Nucl. Instrum. Method Phys. Res. B 32, 238 (1988).

ELSEVIER

Evidence for Spatial Distribution of Traps in MOS Systems after Fowler Nordheim Stress

M. Kerber, U. Schwalke*, and G. Innertsberger

Infineon Technologies AG., Corporate Development,* Corporate Research, 81730 Munich, Germany

The static C-V characteristics allows to separate interface traps and slow traps in MOS systems after high field injection stress. A procedure is developed to deduce the spatial distribution of rechargeable traps in the gate dielectric. The method is based on charge conservation and voltage boundary conditions at the interfaces.

1.INTRODUCTION

High electric fields are known to cause device parameter degradation due to generation of interface traps and charged bulk oxide traps. The classification of rechargeable sites in the MOS system usually is based on the time constants to recover steady state after changes of the internal potential due to external biasing. Traps with relaxation times <0.1 sec are classical interface traps while extremely long time constants are due to fixed oxide charges. Charges trapped in the vicinity of the interfaces with a relaxation time of seconds up to hours are termed slow traps or border traps [1].

Recently, slow traps gain attention as they cause medium term parameter drift [2] and also play a significant role in charge loss mechanisms of non volatile memories [3]. However, the experimental investigation so far primarily is focused on indirect measures of the charge states [4], e.g. threshold voltage shifts or transistor current drifts. These indirect methods sense the electric fields at the substrate interface but miss the charge densities causing the instability.

We have developed an adaptive static C-V procedure with equilibrium controlled feedback of the voltage sweep which allows the separation of interface traps from slow traps in MOS systems and provides both independent information of potential and charge. Hence it permits to address the spatial distribution of various species of traps.
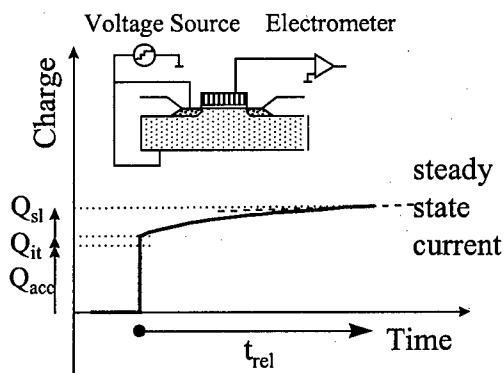
Figure 1 Schematics of the static C-V setup and measurement procedure.

2. THE STATIC C-V MEASUREMENT PROCEDURE

The equilibrium controlled static C-V setup [3] uses commercial equipment as voltage source and electrometer. During the sweep the system waits for steady state after each biasing step. Figure 1 illustrates the setup schematics and the displacement charge versus time after incrementing the gate bias. After the fast relaxation of displacement charge and interface traps the total signal slowly increases and approaches the steady state value when $\Delta Q/\Delta t$ drops below a steady state current. A current resolution of 1 fA is provided by the advanced equipment which allows to detect slow traps even in capacitors of $10^4 \ \mu m^2$ active area.
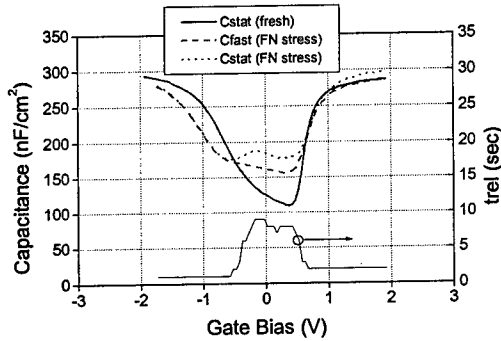
Figure 2 Static C-V curves of the fresh MOS capacitor on p-substrate and after high field injection stress of $0.01C/cm^2$ @ $0.1mA/cm^2$, 85°C and negative gate bias.

In Fowler Nordheim (FN) stressed MOS devices pronounced increase in relaxation time (t_{rel}) is observed between fast response in accumulation and inversion (Figure 2). This additional charge is attributed to charging of slow traps [4] and the amount of charge is a quantitative measure of the density of slow traps.

3.　EXTRACTION　OF　SPATIAL DISTRIBUTION OF TRAPS FROM C-V DATA

The band bending Ψ in a MOS capacitor can be calculated according to [5]:

$$\Psi = \int \left(1 - \frac{C}{C_{ox}}\right) \cdot dV \qquad (1)$$

With the oxide capacitance $C_{ox} = \varepsilon\varepsilon_0/t_{ox}$ and $dQ = C\,dV$

$$\Psi = \int dV - \frac{t_{ox}}{\varepsilon \cdot \varepsilon_0} \cdot dQ \qquad (2)$$

In an ideal MOS system sweeping from accumulation to inversion or vice versa $|\Psi|$ should always results in 1.12V (band gap of silicon). In reality the result is affected by the doping density which may cause poly depletion $\Psi = \Psi_{sub} - \Psi_{polydep}$ and also by the measurement temperature. Since the band bending Ψ (for constant temperature) is not affected by degradation due to
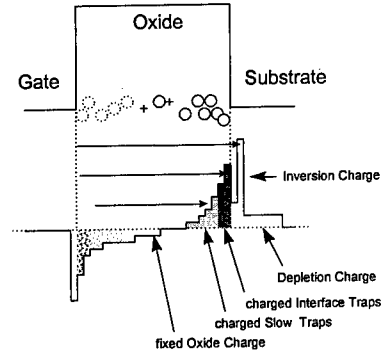
Figure 3 Schematic distribution of charges in the MOS system after high field stress.

electrical stress the values deduced from C-V sweeps taken prior to and after degradation need to be identical. However, for degraded MOS devices this evaluation according to (2) generally leads to a smaller value compared to that deduced for the fresh device. This discrepancy is attributed to the simplified charge distribution in (2) which is not generally appropriate for MOS systems after electrical degradation.

A more detailed model of MOS capacitors illustrated in Figure 3 assumes that the charge of the accumulation/inversion layer in the substrate is separated from the gate interface by the effective oxide thickness t_{ox}. This is the oxide thickness deduced from the accumulation capacitance. The centroid of interface traps is preferentially located at the physical interface $t_{it} < t_{ox}$. Slow traps, finally, are shifted inside the oxide $t_{sl} < t_{it}$ (cf. Fig. 5). With this charge distribution the integration of the band bending turns into

$$\Psi(V) =$$

$$\int_{sweep} dV - \frac{dQ_{acc} \cdot t_{ox}}{\varepsilon\varepsilon_{ox}} - \frac{dQ_{it} \cdot t_{it}}{\varepsilon\varepsilon_{ox}} - \frac{dQ_{sl} \cdot t_{sl}}{\varepsilon\varepsilon_{ox}}$$

$$(3)$$

The static C-V measurement records the fast responding displacement charge which includes dQ_{acc} and dQ_{it} and the slow trapping charge dQ_{sl}. The fast charges dQ_{acc} and dQ_{it} can be separated with the help of the initial sweep: in the fresh device the interface trap density is very small, hence dQ_{fast} only includes dQ_{acc} but not dQ_{it}. dQ_{acc} corresponds to the conventional high-frequency C-V curve. After determining the charges in (3) there
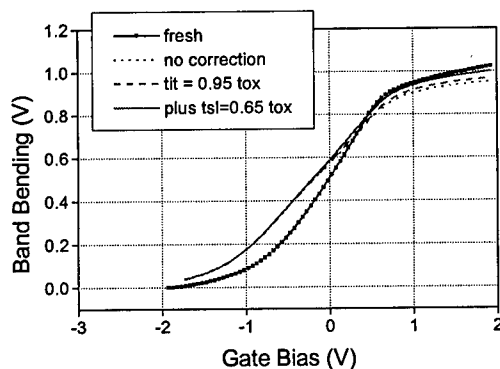
Figure 4 Interface potential of the fresh device and after high field stressing assuming different spatial trap profiles.
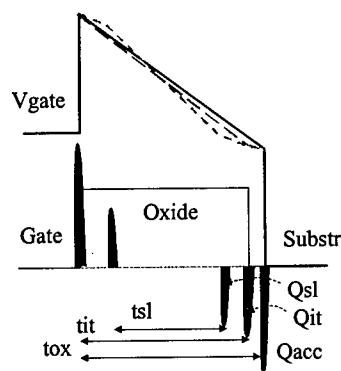
Figure 5 Schematic charge distribution of interface traps and slow traps to deduce the band bending of the C-V characteristics consistently.

are two fit parameters left, t_{it} and t_{sl} which denote the equivalent oxide thickness associated to interface traps and to slow traps respectively. This procedure resolves the above mentioned discrepancy and allows to consistently describe the band bending by adjusting one or both parameters.

4. TRAP DISTRIBUTIONS AFTER FN INJECTION STRESS

Planar MOS capacitors (11nm gate oxide, p-substrate, 0.08mm² active area) have been characterized by the static C-V procedure. After they have been exposed to high field FN injection stress of 0.01 C/cm² at 0.1mA/cm² and negative gate polarity (accumulation) again a static C-V characteristic was taken. Due to the possible
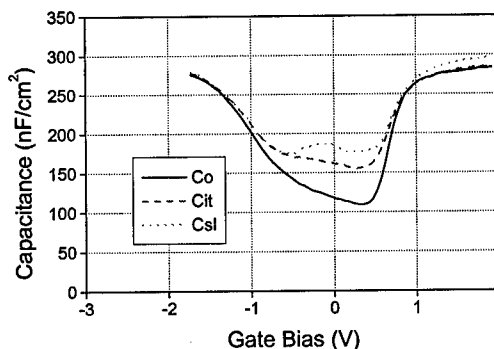
Figure 6 Measured C-V curve and transformation of the characteristics from the virgin device.

separation of fast responding capacitance and steady state capacitance, it is evident from Figure 2 that interface traps have been generated and also slow traps are present. The band bending calculations from (2) and also the using the detailed charge model (3) for the degraded device are shown in Figure 4. It becomes evident that the conventional approach results in a smaller band bending of the degraded device. In a first step the interface traps are shifted by 0.6nm (t_{it} = 0.95 t_{ox}) which increases the band bending. The amount of shift corresponds to the difference of oxide thickness deduced by electrical and optical methods [6] and thus assumes the interface traps at the physical interface. Then the effective distance of the center of slow traps is fitted to increase the band bending to that of the fresh device. An effective distance of t_{sl} = 0.65 t_{ox} is obtained. Assuming slow traps on both interfaces an effective depth of 0.15 t_{ox} = 1.6nm from either interface is deduced. This is a feasible distance for charge exchange by direct tunneling processes. The extended charge distribution model allows to evaluate C-V data consistently and provides experimental evidence for the spatial location of traps in MOS systems. A schematic summary of the charge distribution is presented in Figure 5.

With the correct band bending the C-V curve without interface traps Co is calculated in Figure 6. Thus, the static C-V procedure alone provides all necessary information to evaluate the interface trap densities. The results are plotted in Figure 7 and show a broad distribution of interface traps while the slow traps are confined closer to the conduction
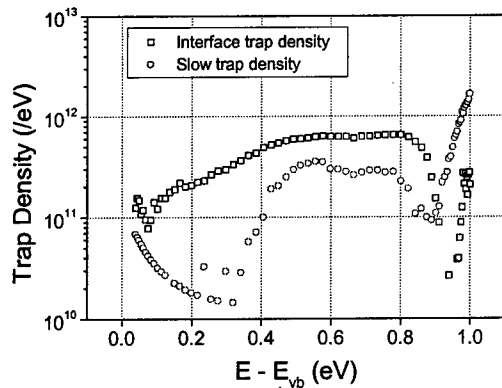
Figure 7 Densities of interface traps and slow traps deduced from Figure 6. The integrative densities of interface traps and slow traps are $N_{it} = 1.3 \ 10^{12}$ and $N_{sl} = 1.4 \ 10^{11} \ cm^{-2}$ respectively

band [7] which is also suggested from Figure 5 directly. Since the band bending is consistent using the improved charge distribution model also the trap density spectra are more realistic.

5. SUMMARY

The equilibrium controlled static C-V procedure is used to experimentally resolve slow traps in MOS capacitors after high field injection. Consistent data analysis leads to a refined charge distribution model which results in smaller effective distances of interface traps and slow traps from the gate electrode as compared to the effective oxide thickness. In Fowler-Nordheim stressed devices, the charge center of slow traps is located about 1.6 nm from the interfaces if slow traps are assumed to be generated at both interfaces. The static C-V procedure thus provides experimental evidence for the spatial distribution of traps in MOS systems.

6. REFERENCES

[1] D. M. Fleetwood, IEEE Trans. Nucl. Sci., NS-39 (1992) 269.

[2] K. Sakakibara et. al., IEEE Trans Electr. Dev., ED-44 (1997) 986.

[3] J. DeBlauwe et al., IEEE Trans Electr. Dev., ED-45. (1998) 1751.

[4] M. Kerber and U. Schwalke, Solid State Electr., 34 (1991) 1141.

[5] M. Kuhn, Solid State Electr. 13 (1970) 873

[6] H. Reisinger, H. Oppolzer and W. Hönlein, Solid State Electr., 35 (1992) 797.

[7] M. Kerber, J. Appl. Phys., 74 (1993) 2125.

Separation of Electron and Hole Traps by Transient Current Analysis

A. S. Spinelli[a] A. L. Lacaita[b], M. Rigamonti[b], D. Ielmini[c] and G. Ghidini[d]

[a]Dipartimento di Scienze Chimiche, Fisiche e Matematiche, Università degli Studi dell'Insubria
via Lucini 3, 22100 Como, Italy

[b]Dipartimento di Elettronica e Informazione, Politecnico di Milano
piazza L. da Vinci 32, 20133 Milano, Italy

[c]Dipartimento di Ingegneria Nucleare, Politecnico di Milano
piazza L. da Vinci 32, 20133 Milano, Italy

[d]STMicroelectronics, via C. Olivetti 2, 20041 Agrate Brianza, Italy

An experimental method capable of separating electron and hole transient currents in MOS devices is presented. The measurements of transient discharge resulting from voltage step above threshold voltage allow a quantitative evaluation of the trap energy distribution and of its dependence on stress. Traps aligned with the oxide band gap are found to increase strongly after stress. The separation of steady-state currents has been also performed, and discussed in terms of the stress-dependent trap distribution obtained from the transient separation analysis.

1. INTRODUCTION

The increased importance of the stress-induced leakage current (SILC) in ULSI low-power and non-volatile memory applications [1] has long been recognized. The SILC has been described by a process of trap-assisted tunneling (TAT) through stress generated oxide traps [2,3]. The experimental measurement of the trap energy distribution represents therefore a primary concern in order to quantitatively understand the SILC process. Although the transient SILC current has been shown to provide useful informations about the oxide trap distribution [4], a careful analysis of the transient processes, taking into account both electron and hole tunneling, is still lacking.

In this work, an experimental procedure for separation between electron and hole transient currents on the same device is presented for the first time. The measurements have been analyzed on the basis of the tunneling front model [5] and the trap distribution inside the oxide energy gap has been computed. The dependence of such distribution on stress has been studied, and evidence for a marked increase in the concentration of traps aligned with the Si energy gap has been

found. Finally, the impact of hole tunneling on the steady-state SILC has also been discussed.

2. SEPARATION TECHNIQUE

Transient current measurement is a valuable tool for probing the oxide trap distribution [4]. Conventionally, this technique requires the recharge of oxide traps at a fixed voltage until equilibrium is reached (injection stage); then, the oxide field is removed by switching the MOS capacitor at flat bands, and the current resulting from trap emission is measured (discharge stage) [6]. Application of the tunneling front model can lead to straightforward computation of the trap density as a function of energy inside the oxide. The technique has been improved here by allowing for carrier separation at the substrate side, thus providing information on the detrapping dynamics of electrons and holes separately. This is accomplished by applying a gate voltage above threshold [7], instead of the flat-band voltage, during the discharge stage.

The procedure is schematically shown in Fig. 1, displaying the tunneling and trapping processes occurring in an n^+-polysilicon/n-MOS structure.
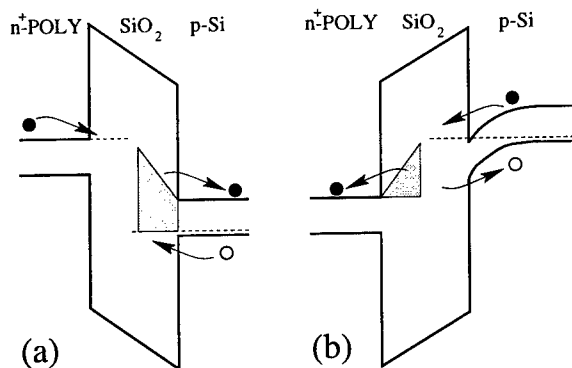
Figure 1. Schematics of the experimental procedure: (a) injection and (b) discharge.

Figure 2. Gate discharge currents for the virgin and stressed samples as a function of the discharge time, for $V_{inj} = -1$ V, -3 V and -5 V.

In (a) the injection dynamics is represented. The application of a negative gate voltage V_{inj} implies that, at equilibrium, traps close to the gate-oxide interface become filled if their energy is lower than the gate Fermi energy. These traps are initially empty due to the previous discharge stage, and are filled as a consequence of electron injection from the n^+-polysilicon. Conversely, traps near the substrate interface are initially filled (shadowed region in Fig.1(a)) and electrons escape from defect levels above the substrate Fermi energy. Electron emission occurs toward both the conduction and valence band of the substrate, and can be treated as hole trapping in the case of emission toward the valence band. Therefore, both trapping of gate electrons and substrate holes occurs in the injection stage.

In the discharge stage (b) application of a positive gate voltage allows for carrier separation at the substrate. Electrons are injected from the source-drain contacts to the initially empty oxide traps, while detrapped holes are collected at the substrate. At the gate-oxide interface electrons are emitted from oxide traps filled during the previous discharge stage (shadowed region in Fig.1(b)).

In order to evaluate the oxide trap distribution, the gate and substrate currents are measured in the discharge stage. Whereas the gate transient current includes both displacement and conduction contributions [8], the substrate current represents the net hole conduction current. In fact, inversion layer electrons effectively screen majority carriers in the p-substrate from interaction with trapped charge in the oxide layer. Therefore, the density of traps aligned with the valence band near the substrate-oxide interface is computed directly from the substrate current, whereas trap levels above the Si conduction band near the gate-oxide interface is calculated from the electron conduction current, obtained from the measured gate current. Changing V_{inj} allows for sampling the trap distribution in a large energy range within the oxide band gap.

3. RESULTS

Figs. 2-3 show transient currents measured after injection at $V_{inj} = -1$, -3 and -5 V gate voltage at the gate and substrate electrodes, respectively. Carrier discharge was measured under +2 V gate voltage. Transient currents shown in Figs. 2-3 behave very differently upon application of an electrical stress. Namely, gate currents in Fig. 2 are strongly enhanced after stress, whereas hole currents in Fig. 3 display only a weak dependence. Better evidence for the different stress sensitivity can be obtained by integrating the transient currents at the substrate and gate electrodes; the
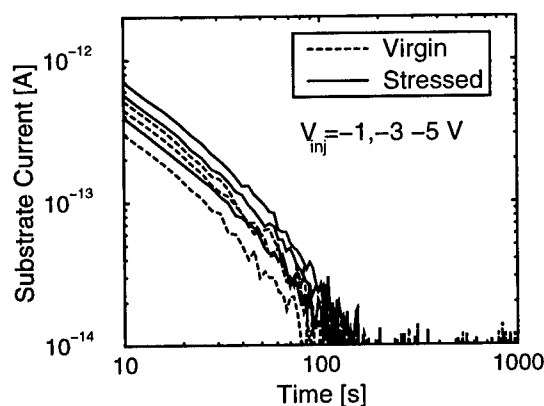
Figure 3. Same as Fig. 2 but referred to the substrate current.

Figure 4. Integrated charge at the substrate and gate, before and after stress.

resulting charge values Q_{tot} are reported in Fig. 4 as a function of V_{inj}. It can be seen that the integrated discharge currents in the virgin oxide are approximately equal at the gate and the substrate, and that application of a stress leads to the addition of an almost constant factor to the total discharge, irrespective of V_{inj}. The additional quantity for gate currents is nearly one order of magnitude larger than that for substrate currents, demonstrating that gate transient current is much more sensitive to the oxide damage conditions.

The determination of trap density has been carried out on the basis of the tunneling front model [5,6]. In this model, which generally refers to the flat-band condition, the application of an external electric field in the discharge configuration must be taken into account. Also, the presence of a displacement contribution affecting the gate current has been properly considered in order to obtain the net particle current. The trap distribution obtained in this way is reported in Fig. 5. The trap density is sampled in energy regions aligned both with the Si bands and with the band gap. Whereas the number of defect states aligned with the Si band gap clearly increase after stress, other energy locations display only small fluctuations upon application of the stress.

The transient charge separation presented so

far allows to measure the carrier discharge after the oxide traps have been populated up to equilibrium. This method has shown the bipolar nature of transient currents, since both electron and hole detrapping processes have been detected. To furtherly characterize the trapping and tunneling of both types of carriers, we have performed I-V measurements on n^+-poly/p-MOS structure. In these type of device, carrier separation is possible at negative voltages below the threshold for substrate inversion: holes can be collected at the source-drain contacts, while electron current is measured at the substrate. Fig. 6 shows the hole and electron currents in an n^+/p-MOSFET with 8.2 nm oxide thickness. In this figure, conventional I-V characteristics (lines) obtained at a ramp-rate of 1 V/400 s are compared to stationary values (symbols), in which the steady-state current at fixed voltage was measured. Two voltage regions can be identified with respect to the sign of the hole current: for low gate voltages up to 5.5 V, a significant hole current is measured flowing toward the gate electrode [9], whereas in the upper voltage range energetic gate electrons generates electron-hole pairs by impact ionization [7], leading to an excess hole current collected at the source-drain contacts. The stationary results in Fig. 6 demonstrates that hole current is
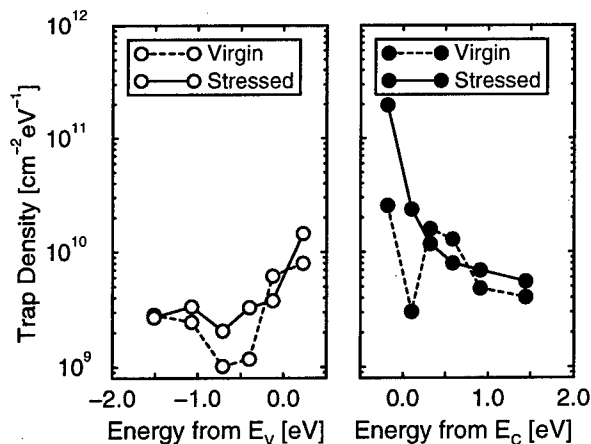
Figure 5. Energy distribution of electron and hole traps for the virgin and stressed samples.

Figure 6. Source/drain (dashes) and substrate (solid) currents in a p-MOSFET. Circles represents steady-state points.

not merely a transient feature, but consists of a steady-state flow through the gate oxide layer, analogous to the well-known electron SILC.

Finally, the trap distribution obtained from transient measurements and the dependence of the trap density on stress, allows to recognize the trap states which might play a major role in the stationary leakage current. Since trap states aligned with the Si band gap display the larger sensitivity to stress, the same states might be responsible for the trap-assisted process allowing for the bipolar leakage observed in Fig. 6. The relationship between concentration of oxide deep-levels aligned with the Si band gap and SILC has been recently predicted, based on an inelastic trap-assisted tunneling model and on calculation of I-V curves in both n- and p-MOS devices [10].

4. CONCLUSION

A novel experimental technique capable of probing electron and hole trap states in virgin and stressed oxide has been developed. It allows determination of trap distribution in a large region of the oxide gap, requiring only minimal computational effort. Results demonstrates that the stress-induced damage is mostly confined in energy locations aligned with the Si band gap, in agreement with recent SILC calculations based on

an inelastic trap-assisted model.

REFERENCES

1. J. D. Blauwe *et al.*, IEEE Trans. Electron Devices **ED-45**, 1745 (1998).
2. R. Moazzami and C. Hu, in *IEDM Tech. Dig.*, 1992, pp. 139–142.
3. D. J. DiMaria and E. Cartier, J. Appl. Phys. **78**, 3883 (1995).
4. D. J. Dumin and J. R. Maddux, IEEE Trans. Electron Devices **ED-40**, 986 (1993).
5. S. Manzini and A. Modelli, *Insulating films on semiconductors* (Elsevier Science, ADDRESS, 1983), pp. 112–115.
6. A. S. Spinelli, A. L. Lacaita, M. Rigamonti, and G. Ghidini, IEEE Electron Device Lett. **20**, 106 (1999).
7. C. Chang, C. Hu, and R. W. Brodersen, J. Appl. Phys. **57**, 302 (1985).
8. D. A. Buchanan, M. V. Fischetti, and D. J. DiMaria, Phys. Rev. B **43**, 1471 (1991).
9. J. D. Bude, B. E. Weir, and P. J. Silverman, in *IEDM Tech. Dig.*, 1998, pp. 179–182.
10. D. Ielmini, A. S. Spinelli, M. Rigamonti, and A. L. Lacaita, submitted for publication.

Validation of the Voltage Step Technique for Determination of Slow State Density in MOS Gate Oxides

M.P.Spillane, S.Taylor and M.J.Uren[*]

Department of Electrical Engineering and Electronics, Liverpool University, Brownlow Hill, Liverpool L69 3BX

[*] DERA St Andrews Road, Malvern Worcs. WR14 3PS, UK

Presented here is a technique for profiling near interfacial oxide traps based on the McWhorter tunneling model. The technique is verified by comparing slow state densities obtained using the G_p/ω conductance technique, which is an established method for slow state density determination.

1. INTRODUCTION

As MOS device dimensions become smaller, reliability becomes a key issue. Long term reliability is affected by the presence of slow states, or border traps, in the oxide close to the Si-SiO$_2$ interface. During measurement of $1/f$ noise in MOS transistors, Meisenhimer and Fleetwood[1] found that radiation induced $1/f$ noise correlated with oxide trap charge through irradiation and anneal, but not with interface trap charge. This implies that $1/f$ noise is directly related to the oxide trap charge in irradiated or otherwise electrically damaged devices. McWhorter[2] proposed that near-surface states are responsible for the $1/f$ noise seen in MOS transistors, with a distribution of trapping times at the semiconductor-oxide interface which would arise naturally from a spatially uniform distribution of tunnelling depths to traps in the oxide. Measurements of Random Telegraph Signals in MOSFETs confirmed that slow surface states are responsible for the $1/f$ noise, but with the distribution of trapping times being due to both a distribution of tunnelling depths and a distribution

of activation energies for capture[3]. Nevertheless, the simple McWhorter tunnelling model is usually used to interpret $1/f$ and slow state measurements because of its easy applicability.

In the study of the effects of oxide traps on MOS capacitance, Heiman and Warfield[4] used the McWhorter model and concluded that very little information about the energy distribution and capture cross sections of the oxide traps is obtained from the analysis of MOS capacitance curves. Kerber[5] found experimental evidence of slow traps by using an improved static C-V method which keeps the MOS system close to equilibrium throughout the C-V sweep. The decay of the displacement current after small gate bias steps was attributed to the charging / discharging of slow traps.

Paulsen[6] et al encountered low frequency increase in charge recombination per cycle in studies of MOS devices using the charge pumping technique which they attributed to the charging and discharging of traps located within tunnelling distance of the Si/SiO$_2$ interface, i.e. border traps. Uren[7] et al observed a low frequency plateau in conductance measurements on MOS capacitors

which was consistent with slow states observed by other techniques in particular those states responsible for 1/f and telegraph noise in MOSFETs. More recently the study of current transients following the application of a small voltage step has been used by Tanner[8] *et al* to characterise traps generated in the oxide of MOS capacitors by constant current stressing.

Here we present an analysis of the experimental method presented by Tanner et al using the McWhorter model, and apply it to radiation damaged oxides. The measurement of slow state density is validated using the conductance technique. A feature of the technique used in this analysis of border traps is that it allows the relationship of time with distance into the oxide enabling 3D profiling of states.

2. THEORY

For a set of slow states at the Fermi level at equilibrium, the capture time equals the emission time ($\tau_c = \tau_E$) where :

$$\tau_c = \sigma \upsilon n^{-1} \tag{1}$$

where σ is the capture cross section, υ the carrier velocity and n the carrier concentration in the conduction band at the surface, given by :

$$n = N_C \exp(-E_{CF} / kT) \tag{2}$$

where N_C is the effective density of states in the conduction band and E_{CF} is the Fermi energy relative to the conduction band.
Assuming that the population of oxide traps is distributed uniformly in distance into the oxide then according to McWhorter's model

$$\sigma = \sigma_o \exp(-2\kappa x) \tag{3}$$

where κ is the wavefunction inverse decay length in the oxide. The capture time for traps at a depth x into the oxide is:

$$\tau_c = \exp(2\kappa x + E_{CF} /kT) / \upsilon\sigma_o N_C \tag{4}$$

and

$$x = 1/2\kappa[\ln(\upsilon\sigma_o N_C\tau) - E_{CF} / kT] \tag{5}$$

We assume the traps are distributed through the oxide both in energy and space with density N_T m^{-3} joule^{-1} as is required to produce a 1/f noise spectrum. The method consists of stepping the gate voltage resulting in a change in the Fermi energy ΔE (fig1). If the device is in equilibrium before the step, then following the step traps up to a depth x will have changed charge state, and depth x will increase with time, thus:

$$N(x) = xN_T\Delta E. \tag{6}$$

where N(x) is the number of states which will have changed state per unit area. Now all states with time constants greater than the time constant for x = 0 ie $\tau > \tau_{c0}$ where

$$\tau_{c0} = \exp(E_{CF}/kT)/N_C\upsilon\sigma_0 \tag{7}$$

and less than the time constant at depth x, ie the measurement time, will have changed charge state. Substituting for x

$$N(t) = 1/2\kappa [\ln(\upsilon\sigma_0 N_C t) - E_{CF}/kT] N_T\Delta E \tag{8}$$

the total charge is

$$Q(t) = N(t)q \tag{9}$$

ie charge in the slow states changes logarithmically with time.
The current density at time t after the step is

$$J(t) = dQ(t)/dt \tag{10}$$

$$J(t) = qN_T\Delta E/2\kappa t \tag{11}$$

And so

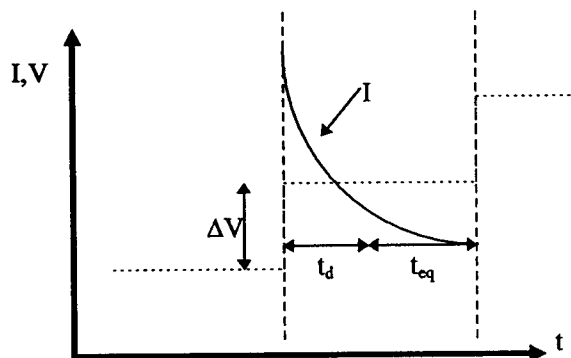$$N_T = 2\kappa J(t)t/q\Delta E \tag{12}$$

Figure 1. Following voltage step ΔV, current I is measured after a delay t_d. A further delay t_{eq} is allowed for device to return to equilibrium before next voltage step.

3. EXPERIMENT AND RESULTS

Both n-type and p-type aluminium gate MOS capacitors, with an oxide thickness of 26nm and a gate area of 6.97 x 10^{-3} cm^2 were X-ray irradiated with 400krad(Si) dose to produce both fast and slow states. These were examined using both the above technique and the conductance technique. For the equilibrium voltage step technique measurements were taken using the Keithley 595 Quasistatic C-V meter. The measurements taken provided the information to plot both the quasistatic C-V and the I-V curves used in the determination of energy levels and trap densities. The gate bias voltage was swept from +4V to -4V using a staircase waveform with a step voltage of 10mV per step, which was supplied by the Keithley 595.

Following the application of the voltage step the resulting current transients were measured after a pre-determined time delay ranging from 100ms on the first sweep up to 2secs on subsequent sweeps. Following each measurement a delay of 6secs was allowed for the device to return to equilibrium before the next voltage step was applied. This delay time necessary to ensure equilibrium was determined by biasing the device into inversion and monitoring the capacitance and current at regular time intervals until the capacitance had risen to the value of C_{ox} and the current had fallen to zero. Fig.2 shows an
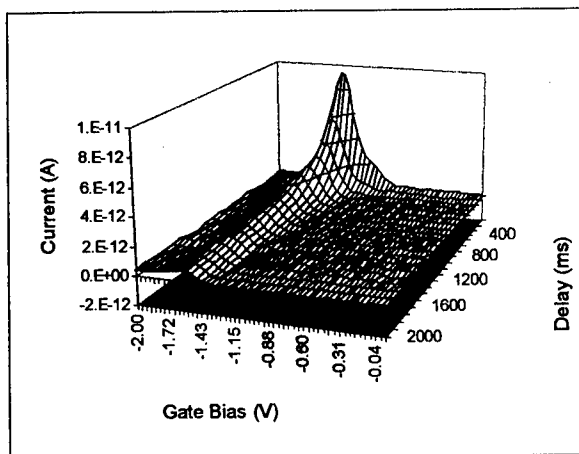
Figure 2. Measured gate current following the application of 10mV step. The current is plotted as a function of time delay which is related to the distance into the oxide.

experimental data set where the peak corresponds to the fast states and the slow states appear as a low level plateau to the right of the peak.

The relation of Fermi energy to the gate bias was determined from the equilibrium C-V information and obtained by using analysis presented by Nicollian and Brews[9]

$$\psi_s(V_g) - \psi_s(V_{fb}) = \int_{V_{fb}}^{V_g} [1 - C_{lf}/C_{ox}]dV \qquad (13)$$

where the flatband voltage V_{fb} is defined by $\psi_s(V_{fb}) = 0$

Fig.3 shows the calculated density of slow states using Eqn. (12), around the mid gap region for the n-type capacitor (right of the prominent peak). It can be seen that the level of states into the oxide remains constant. This agrees with the McWhorter model and recent suggestions made by Bauza and Maneglia[10] in their study of oxide traps using the charge pumping technique. The prominent peak in the slow state density corresponds to the limiting case of traps located at the surface (ie to the fast states) indicating a close relation between fast and slow states as concluded by Kerber[2].

The conductance technique (G_P / ω) was applied to the same device, where the density of

'plateau' region below the normally analysed fast state peak.
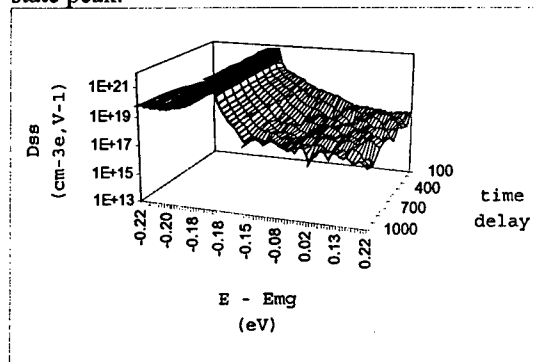
Figure 3. Density of slow states plotted as a function of time delay mid gap region of n-type capacitor.

Fig.4 shows the results of this comparison for the MOS capacitor in depletion taken from the experimental measurements at 1 sec time delay using EVS and experimental measurements taken at a frequency of 1 Hz using G_P / ω. It can be seen that the two sets of results are broadly in agreement with each other.
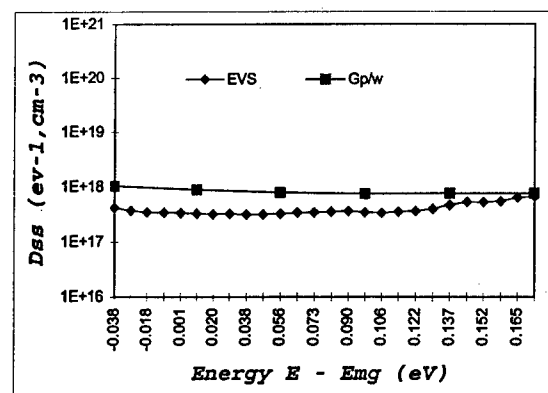
Fig. 4 Density of slow states for Equilibrium Voltage Step (EVS) technique with a time delay (t_d) of 1 second and Conductance technique (G_P /ω) at a frequency of 1Hz

4. CONCLUSIONS

A technique for profiling near interfacial oxide traps is demonstrated based on application of a voltage step to an MOS gate, assuming the McWhorter model of traps distributed into the oxide. The technique has been verified by comparing the slow state densities obtained with the results of the G_p/ω conductance technique which is an established method for slow state determination. The equilibrium voltage step technique is simpler (although of lower resolution) than conventional methods such as G_p/ω and charge pumping. A feature of the technique is that it allows the relationship of time with distance into the oxide enabling 3D profiling of states.

REFERENCES

1. T.L. Meisenhier and D.M. Fleetwood, IEEE Trans Nucl. Sci. NS-37, 1696 (1990)
2. A.L. McWhorter, Semiconductor Surface Physics (1957)
3. M. J. Kirton and M. J. Uren, Adv. in Phys. 38, 367 (1989).
4. F.P. Heiman and G. Warfield, IEEE Trans Electron Devices ED-12, 167 (1965)
5. M. Kerber , J. Appl. Phys. 74, 2125 (1993).

6. R.E. Paulsen, R.R. Siergiej, M.L. French and M.H. White, IEEE Electron Device Lett. EDL-13, 627 (1992)
7. M.J. Uren, S. Collins and M. J. Kirton, Appl. Phys. Lett. 54, 1448 (1989)
8. P.Tanner, S.Dimitrijev and H.B.Harrison, Electronics Lett 31,1880 (1995)
9. E.H Nicollian and J.R. Brews, MOS Physics and Technology (1982)
10. Y. Maneglia and D. Bauza, J. Appl. Phys. 79, 4187 (1996)

Two types of traps at the Si/SiO₂ interface characterized by their cross sections

J. Albohn, W. Füssel, Ngo Duong Sinh*, K. Kliefoth, H. Flietner, and W. Fuhs

Hahn-Meitner-Institut Berlin, IT Adlershof, Kekulestraße 5, D-12489 Berlin, Germany

The dispersion behavior of traps at the Si/SiO₂ interface has been investigated applying the modulation capacitance voltage method. N-type MOS samples of different material parameters and preparation conditions were used to vary the state density distribution. Two dispersion regions have been resolved in the upper half of the gap which can be described by two Gaussian broadened time constants. The evaluated capture cross sections can be classified in two groups independent of the specific properties of the samples. The cross sections were assigned to two types of silicon dangling bond defects with different back bond configurations.

1. INTRODUCTION

The static properties of traps at the Si/SiO₂ interface such as density and energetic distribution have frequently been studied as a function of preparation conditions and this knowledge is widely used in device technology. There is far less information about the dynamic properties e.g. emission and capture times or cross sections. A critical review shows that in the upper half of the gap the published data scatter over more than 4 orders of magnitude without definite relationship to preparation conditions [1]. This paper presents investigations of the electron capture process of Si/SiO₂ interface traps in the upper half of the gap applying the modulation capacitance voltage (MCV) method [2].

2. EXPERIMENTAL

In the used set up the MOS capacitance is the tuning part of an oscillator circuit (f_{HF} = 5 MHz) which transforms the capacitance change to a frequency shift. Adding a small sinusoidal LF signal (v_{Mod} = 1 Hz...100 kHz) to the bias voltage the complex response of the demodulator is measured using the lock-in technique. The response $\Delta C(v_{Mod})$ represents the derivative of the MOS HF-capacitance to the modulation voltage dC_{MOS}/dV_{Mod} for a constant

amplitude of the LF signal and is recorded in the real and the imaginary component. The MCV spectrum shows (similar as in admittance spectroscopy) the dispersion behavior of interface traps at the energy level determined by the bias voltage. The spectra were recorded as a function of the gate voltage V_g and in the temperature range T= 210 ... 297 K to shift the total dispersion region of interface traps into the available frequency window.

To get a representative overview, n-type MOS samples of different silicon material (Czochralski, Floatzone), surface orientation (<111>, <100>), oxidation conditions (wet, dry), and annealing treatments (high temperature annealing, low temperature annealing in hydrogen atmosphere, without annealing) prepared by three different laboratories were investigated (Tab. 1).

3. EXPERIMENTAL RESULTS

As to be expected, the samples show a considerable variety in the density and shape of the energetic distributions $D_{it}(E)$ of the interface traps determined from the HF-CV data (Fig. 1). Samples 1 and 2 represent the as-oxidized state for wet oxidation and samples 3 and 5 that for dry oxidation with 3 vol% TCA supply. Annealing at 800 °C in H₂ does not significantly change the interface state density (sample 2).

---

*Present address: Institut für Physikalische Hochtechnologie, PF 100 239, D-07702 Jena, Germany

Table 1
Sample preparation

Sample	Wafer material	Doping density [cm^{-3}]	Oxidation	Annealing	Oxide thickness [nm]
1	CZ <111>	1.3x10^{15}	1180°C (wet)	no	400
2	CZ <111>	9x10^{14}	1180°C (wet)	800°C in H$_2$ quenched in N$_2$	400
3	FZ <111>	1.3x10^{14}	1100°C (dry)	no	100
4	CZ <111>	6x10^{14}	1000 °C (dry/wet/dry)	400°C in H$_2$ after metallization	850
5	FZ <100>	1.4x10^{15}	1100°C (dry)	no	95
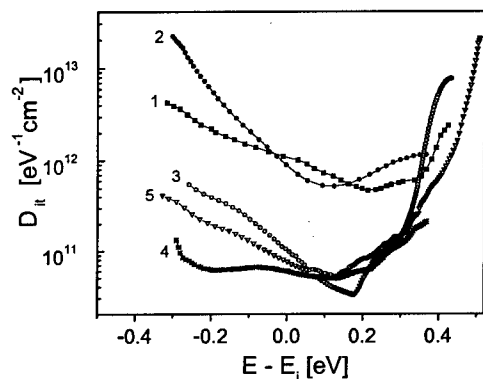
Fig. 1. Energetic distribution of the interface state density for samples 1 to 5.

These distributions are characterized by a pronounced fraction of states unsymmetrically shaped around midgap with the major contribution in the lower half of the gap. The more symmetric distribution of sample 4 show the well known effect of hydrogen annealing at low temperatures decreasing the state density by hydrogen saturation of dangling bond defects.

The measured dispersion spectra show typically broadened dispersion regions instead of Debye curves for a single time constant. For samples 1 and 2 the dispersion can be described by a simple Gaussian broadening of time constants which is usually attributed to potential fluctuations. For samples 3 and 5 at all investigated trap energy levels two distinct dispersion regions can be separated in the spectra each describable by a Gaussian broadened Debye formula. Sample 4 exhibits also two disper-

sion regions but the one at lower frequencies can be resolved only in the energy range of E – E$_i$ = 0...0.2 eV. This results point to the superposition of two different types of states in the related energy range. The characteristic time constants τ_1 and τ_2 of the distinct dispersion regions for a certain energy level differ by about 1 to 2 orders of magnitude.
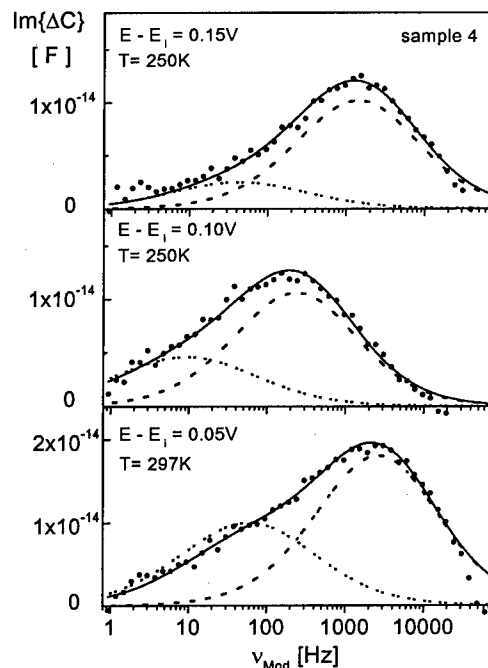
Fig. 2. The imaginary part of the spectral response of sample 4 for three selected trap energies including the fractions for both time constants and the total fit. The lines show the result of the fitting procedure described in section 4: • experiment, fits: --- τ_1, ··· τ_2, —— total.

4. DATA ANALYSIS

The fitting procedure for analyzing the spectra were based on the calculation of the equivalent network for MOS samples in accumulation and depletion modes. Interface traps were represented in the usual way by series capacitor/resistor links [3]. The observed broadening is reflected by Gaussian weighted sets of links with time constants equidistantly distributed on a logarithmic scale.

To calculate the characteristic dispersion frequencies

$$\nu_{char} = 1/2\pi\tau \qquad (1)$$

the response of the complete network must be taken into account:

$$\nu_{char} = \nu_{peak} (C_{sc} + C_{ox})/(C_{it} + C_{sc} + C_{ox}) \qquad (2)$$

where C_{it}, C_{sc}, and C_{ox} denote the capacitances of interface states, space charge layer, and oxide layer, respectively. According to our results the definition of the characteristic frequencies ν_{char} by the peak positions ν_{peak} of the spectrum would lead to errors between the factor 3 up to two orders of magnitude.

The fitting procedure consists of three steps, using i) the normalized real component and the phase angle for a first approximation, ii) the Cole-Cole plot for improving the fit, and iii) the measured real and imaginary response for a final adjustment.

Electron capture cross sections have been calculated equating emission and capture rates:

$$\sigma_n = \nu_{char} / (v_{th} N_c \exp((E_F-E_c)/kT) \qquad (3)$$

where v_{th} denotes the thermal velocity, N_C the effective density of states in the conduction band, and E_F the Fermi energy at the interface.

5. RESULTS AND DISCUSSION

As surprising result, we found that the dependence of the evaluated capture cross sections on the energy position can be classified in two curves (Fig. 3). This classification is unaffected by the variety of material, the preparation parameters, and the resulting trap densities. Therefore, we conclude that the two sets of cross sections represent a basic feature of the Si/SiO$_2$ interface. This result suggests two different defect centers the distribution of which are superimposed in a broad energy range in the gap. Fig. 5 shows a model for the fractional state density distributions of defects at the Si/SiO$_2$ interface which has been developed predominantly on CV and surface photovoltage (SPV) measurements [4].

There are exponential tails (U_T) and various types of dangling bonds (U_M, P_L, P_H) which differ chemically by the number of backbonded Si and O atoms.

A first argument for an assignment of the σ_n to specific defect types comes from the charge state of the centers involved. The donor-like state due to the (+/0) transition of the P_L defect is coulomb attractive for electrons and should have a larger capture cross section than the acceptor-like U_M states (0/-) of the singly occupied neutral dangling bond defect. However a defect center in a higher charge state does not in any case lead to a higher cross section [5].
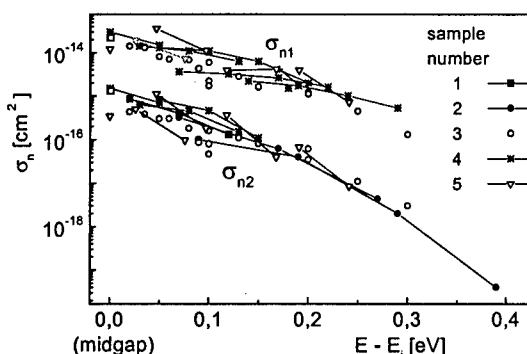
Fig. 3 Capture cross sections of five samples as a function of energy.
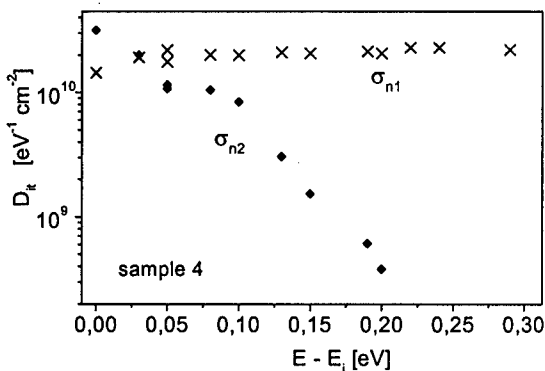
Fig. 4. Fractional state distributions $D_{it}(E)$ related to the two cross sections for sample 4.

Considering the partial $D_{it}(E)$ distributions (Fig. 4) connected with the two groups we arrive at the opposite conclusion. In Fig. 5 we have highlighted the energy window accessible in the present study. It is obvious that the distributions of Fig. 4 resemble the distribution of the U_M and the P_L group in Fig. 5.

We suggest that the larger cross section σ_{n1} is related to the -Si≡Si$_3$ defect (U$_M$ group of traps identical with states of the P$_b$ center) which is known from EPR and CV measurements to show a symmetric state distribution around midgap [6, 7] and the smaller σ_{n2} to the -Si≡Si$_2$O defect (P$_L$ group of traps).
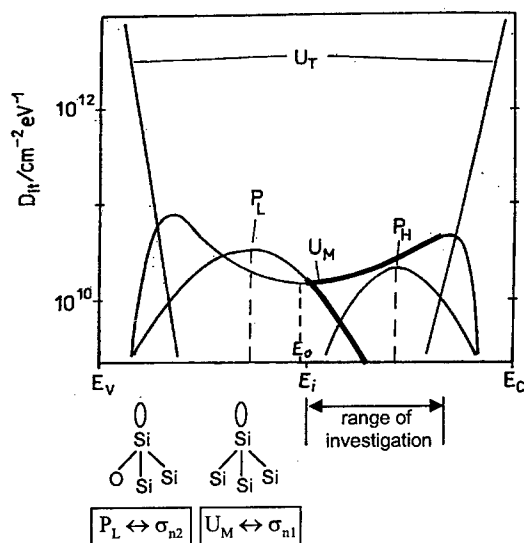
Fig. 5. Model for fractional state contributions of different interface defects [8]. The parts of the fractional contributions which correspond to the experimental results of figure 4 are marked.

This assignment is further supported by the detailed influence of the preparation conditions on the D$_{it}$ distributions. For the samples 1 and 2 the contribution of the P$_L$ states is such high that it dominates the spectral response also in the investigated upper part of the gap. Only the dispersion at lower frequencies related to σ_{n2} can be resolved for these samples. For the samples 3 and 5 both dispersion regions have been observed in the whole investigated energy range. Sample 4 show only a small contribution of the asymmetric P$_L$ fraction. In fact, we observed the dispersion at lower frequencies only in a limited energy range of 0.2 eV at midgap. Summarizing these argumentations the assignment of the larger cross section σ_{n1} to be related to the U$_M$ group (P$_b$ centers) and the smaller σ_{n2} to the P$_L$ group of defect states is the most convincing.

The remarkable result of decreasing cross sections towards the conduction band cannot be understood in the kinetic model of capture process. Dis-

cussing non-radiative transition probabilities in a configuration coordinate diagram under the assumption of strong electron phonon coupling (Huang-Rhys factor S »1) the decreasing depth of the defect state leads to an increasing energy barrier for the capture process. Thus the $\sigma_n(E)$ values decrease towards the conduction band. This topic will be discussed in more detail in a forthcoming paper.

6. CONCLUSIONS

Two distinct types of capture cross sections for traps in the upper half of the gap have been determined using modulation capacitance voltage spectroscopy. Investigations of different samples show that the magnitude and the energy dependence of the cross sections are not affected by material parameters and preparation conditions. We conclude that a basic feature of defects at the Si/SiO$_2$ interface has been observed. The larger cross section can be assigned to the dangling bond defect -Si≡Si$_3$ and the lower cross section to -Si≡Si$_2$O.

REFERENCES

1. A. Aberle, S.Glunz, and W. Warta J. Appl. Phys. 71 (1992) 4422.
2. N.D. Sinh and H. Flietner, in Proc. 4th ICFSI, eds. B. Lengeler, H. Lüth, W. Mönch, J. Pollmann, World Scintific, Singapore, 1993.
3. E.H. Nicollian and J.R. Brews, MOS Physics and Technology, J. Wiley & Sons, NewYork, 1982.
4. H. Flietner, Mat. Science Forum 185-188 (1995) 73.
5. N.N. Abakumov, V.I.Perel, and I.N. Yassievich, Sov. Phys. Semiconductors 12 (1978) 1.
6. P. M. Lenahan and P. V. Dressendörfer, J. Appl. Phys. 55 (1984) 3495.
7. J. H. Stathis and E. Cartier, Phys. Rev. Lett. 72 (1994) 2745.
8. W. Füssel, M. Schmidt, and H. Flietner, Nucl. Instr. and Methods in Phys. Research B 65 (1992) 238.

TURN-AROUND EFFECTS DURING DYNAMIC OPERATION IN 0.25μm CMOS TECHNOLOGY FROM LOW TO HIGH TEMPERATURE

A. Bravaix [a], D. Goguenheim [a], N. Revil [b], E. Vincent [b]

[a] ISEM, Maison des Technologies, place G.Pompidou, 83000 Toulon, France.
[b] STMicroelectronics, 850 rue J.Monnet BP16, 38926 Crolles cédex, France.

The hot-carrier reliability of a 0.25μm CMOS technology with a 5nm-thick gate-oxide is studied from low (-40°C) to high temperature (125°C). Channel hot-carriers are investigated using DC and AC experiments on single transistors and on ring oscillators. Turn-around effects are observed in the degradation of the ring oscillator frequency at short times when P-MOSFET's degradation totally compensate the N-MOSFET's current reduction. As the electron trapping becomes much less efficient in thin gate-oxide P-MOSFET's with respect to field enhanced detrapping, the influence of the generated interface traps dominates leading to a cumulative current reduction in both device types at long stress time which results in a reduction of the oscillator frequency. The degradation behavior of the oscillator frequency exhibits a clear dependence on the P-MOSFET channel-width and on temperature as negative bias temperature instability occurs at 125°C during the high state in P-MOSFET's and the field-assisted discharge of electron traps occurs at -40°C.

1. INTRODUCTION

The different damage mechanisms involved in the hot-carrier reliability of MOSFET's have been largely studied in the literature [1-3], but much effort is still required to assess the interaction between the different mechanisms during AC operations in recent technologies with thin gate-oxide ($\leq$ 5nm). The degradation of circuit performances correlated to device parameter variations is required in order to determine the impact of the device degradation in a realistic way. As devices operate over a wide range of temperature, it is crucial to determine the effect of temperature on the degradation because other mechanisms like Negative Bias Temperature Instability (NBTI), thermal emission, field-enhanced charge detrapping may contribute to the resultant effect on the transistor parameters [4,5].

In this work, we investigate AC/DC degradation in single N- and P-MOSFET's and in 56-90 stage ring oscillators from a 0.25μm CMOS technology with a 5nm gate-oxide thickness. The process presents Medium Doped Drain N-MOSFET's and Lightly Doped Drain P-MOSFET's with $n^+(p^+)$ polysilicon gates, twin retrograde wells and a Shallow Trench

Isolation. The oscillator frequency (F_{osc}) is measured after a 256 divider down to a 0.01% resolution with a HP5316B. DC and AC results are correlated to analyze the degradation of circuit performance.

2. EXPERIMENTAL RESULTS

2.1 Alternating degradation mechanisms

The trapping-detrapping balance is studied (Fig.1) during AC stresses performed with a constant hot
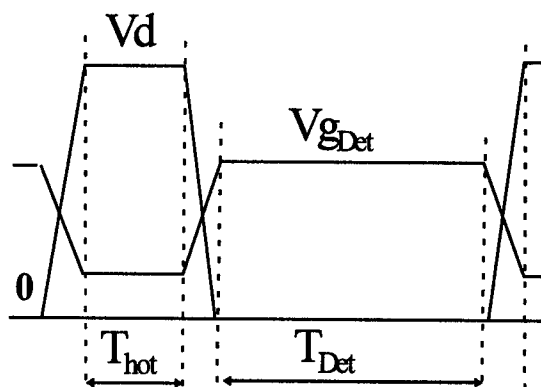

Fig.1: Illustration of AC stresses with constant T_{hot} phases varying the detrapping period T_{Det} for various Vg_{Det} and a low duty-cycle fixed at 25%.
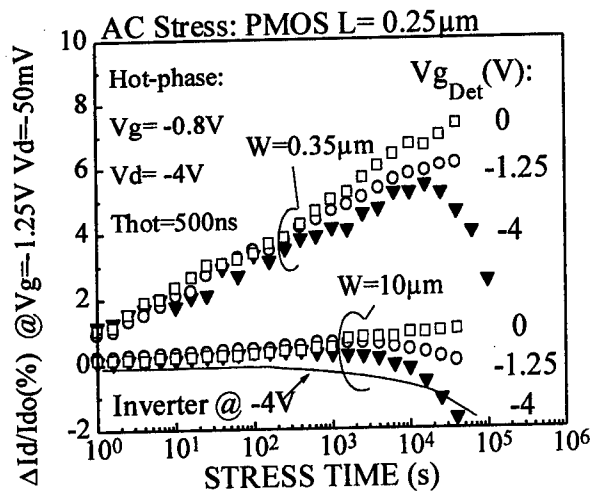
Fig.2: Drain current degradation during AC stresses in wide and narrow PMOSFET's (Leff=0.2μm) with Vg$_{Det}$ as the variable during the detrapping phase (T$_{Det}$).

Fig.3: Drain current reduction in NMOSFET's for the same time sequence than in Fig.2 and for two Vg$_{Det}$ values during the detrapping period. The case inverter (filled squares) is for a 50% duty cycle

phase T$_{hot}$ and detrapping phase T$_{Det}$, for various gate-voltages Vg$_{Det}$. During T$_{hot}$ periods in P-MOSFET's, Vg is chosen to yield the maximum gate-current (Fig.1) where electron trapping (ΔNot-) occurs under hot electron (HE) injection (Fig.2). Positive increases of transistor parameters as the linear drain current, the transconductance or the threshold voltage are related to ΔNot- which induces the typical channel shortening effect. ΔId/Id$_o$ is chosen here although *same behaviors* are observed with other parameters. The increase in the oxide field (F$_{ox}$) with Vg$_{Det}$ enhances (Fig.2) the electron detrapping which is observed by a saturating logarithmic time-dependence and a turn-around at Vg$_{Det}$= -4V (Vd=0). This turn-around also arises from the contribution of the interface trap generation (ΔNit) when Vg$_{Det}$ is high (T$_{Det}$ period) where negative charge can partly detrap [3] leading to a |Id| reduction. Inverter-like stressing on single p-devices shows (solid line in Fig.2 with 50% duty cycle) only a current decrease (ΔId/Id$_o$ <0) which confirms the effect of charge detrapping *and* ΔNit. Similar degradation results have been evidenced during pass transistor operation using the charge-pumping technique [6] due to these counteracting damage mechanisms. This turn-around is clearly enhanced in narrow P-MOSFET's where ΔNot- is increased as much as the charge detrapping and ΔNit in the gate-drain overlap region. This can be ascribed to the large proportion of the degraded region (ΔNot- and ΔNit) with respect to the active region in narrow

devices. During T$_{hot}$ in N-MOSFET's (Fig.3), Vg is chosen to yield the maximum substrate current (IB) where ΔNit occurs following a typical power-law with a 0.55 time exponent. Contrary to P-MOSFET's (Fig.2), no significant dependence is observed with Vg$_{Det}$ because when Vd is high and Vg pulsed with long T$_{Det}$ period, no positive charge is present and we only observe the influence of ΔNit [3,5]. A small dependence on channel-width is seen (Fig.3) which comes from ΔNit until the overlap region as the electron trapping (and detrapping) does not occur here due to the non-favoring electric-field.

2.2 Temperature dependence of DC stresses

DC stresses are investigated (Figs.4,5) at low temperature (LT), room temperature (RT) and high temperature (HT). In P-MOSFET's, the turn-around occurs at the HE condition and *is enhanced at* -40°C *and* 125°C. This shows competing interactions between ΔNit and the trapping-detrapping balance assisted by F$_{ox}$ at LT where the injected current is twice the one at RT. At HT thermal emission from shallow oxide traps dominates the resulting damage. The turn-around is clearly enhanced at -40°C which indicates that the fraction of the detrapped charge increases suggesting a field-assisted thermal discharge of electron traps [5]. This is explained by the trapping-detrapping balance at relatively low oxide field F$_{ox}$ and the resulting increase of the influence of ΔNit with electron detrapping [6].

Fig.4: Drain current degradation in PMOSFET's as a function of temperature for stressing time 1000s (open labels) and 4000s (filled) at the HE condition (Vg ≈ Vd/5) and the HH condition (Vg=Vd). At each temperature, stressing and measurements are performed at the same temperature.
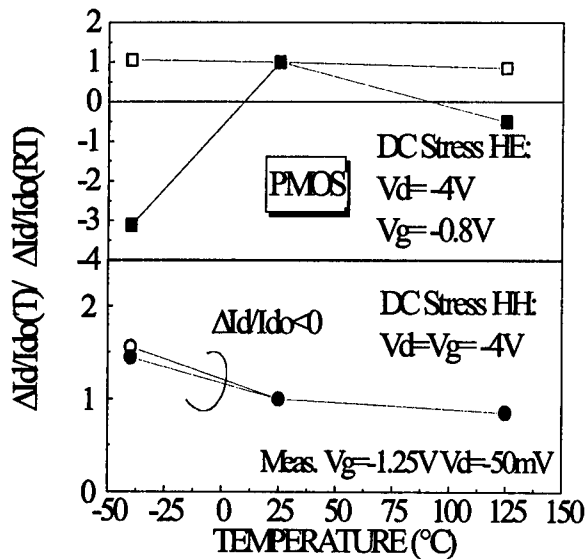
Fig.5 Same plot as in Fig.4 for NMOSFET's (W/L=10/0.25μm) at the voltage condition of each damage mechanism *i.e.* at the HE condition (Vg = Vd), IB (Vg ≈ Vd/2) and HH (Vg ≈ Vd/5) and for stressing time 1000s (open labels) and 4000s (filled).

Under hot hole (HH) injection at Vg = Vd (Fig.4) the P-MOSFET degradation originates from the increase in donor-type ΔNit leading to the |Id| reduction. No large difference is found between RT and HT results at the HH condition in both devices (Figs.4-5) indicating that ΔNit is not affected at HT. NBTI damage is known to play a significant role in P-MOSFET degradation at HT [4]. In contrast to hot-carrier damage, NBTI damage typically arises above 100°C when Vgs is high causing significant degradation of the transistor characteristics at low oxide field (F_{ox} = 1-2MV/cm). Consequently NBTI damage in the oxide can occur at the source side during a HH stress and above the whole active region during the Vg_{Det} period during inverter operation (Figs.1,2) [4]. It is clear in Fig. 4 that NBTI has no effect here on the resultant HH degradation. NBTI damage shows a larger impact at the high state in P-MOSFET's during inverter operation with an activation energy between 0.35-0.45eV. In this latter case, the NBTI effect on transistor parameters is analogous to oxide hole trapping as checked by the parallel shift in the degraded subthreshold drain current and by charge-pumping measurements (not shown here).

2.3 Degradation of circuit performance

Ring oscillators with different number of stages and load capacitances (Table 1) are tested with the shift of the oscillator frequency ($\Delta F_{osc}/F_{osc}$). Fig.6 shows the F_{osc} reduction ($\Delta F_{osc}/F_{osc}$ < 0) for ring B with large P-MOSFET's where ΔNot- remains ineffective and where the F_{osc} change originates from

Table I : Description of the ring oscillators

L=0.25μm	Wn(μm)	Wp(μm)	Fan Out	Stage Nb.
RING A	0.5	1	1	56
RING B	5	10	5	90
RING C	0.5	1	5	90

the Id reduction in N-MOSFET's (Fig.3). A frequency speed-up and down of the ring oscillator has been previously explained as caused by P-MOSFET drain current enhancement [7]. The turn-around is observed when $\Delta F_{osc}/F_{osc}$ >0 at the beginning of the stress in ring oscillators *only with narrow* P-MOSFET's (Wp=1μm) with different capacitive loads. This is explained by the ΔNot-predominance (Fig.2) on the charging current and by the consequent decrease in delay time observed in the subsequent stages with narrow P-MOSFET's. As

Fig.6 Comparison of the ring oscillator frequency (F_{osc}) shifts at 25°C for the three rings A, B and C of geometry ratios given in table 1

Fig.7 Turn-around in the oscillator frequency as a function of temperature for ring oscillators B and C (see Table 1).

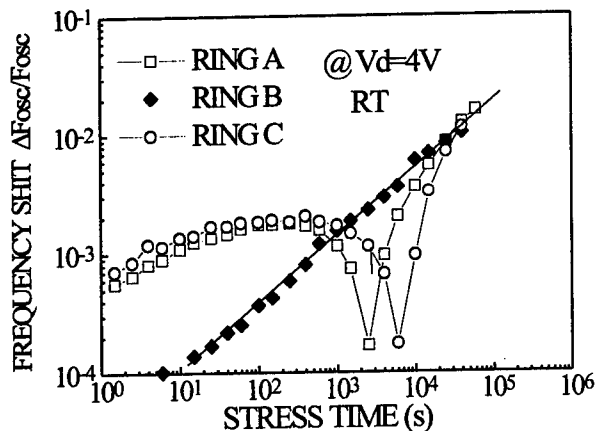the time proceeds $\Delta F_{osc}/F_{osc}$ becomes negative which corresponds to the detrapping effects leaving the effect of donor-type ΔNit in P-MOSFET's. In addition, due to the distinct time dependencies (power-law *vs.* logarithmic law for N and P- devices respectively), the three curves tend to merge at long stress time with approximately the same degradation when both transistor types degrade significantly (ΔNit dominant), and when the N-MOSFET degradation finally dominates the degradation. At 125°C the N-MOSFET degradation is dominant in ring B (Fig.7) with $\Delta F_{osc}/F_{osc}$ <0 confirming that P-MOSFET's (Wp=10μm) induce neither damage compensation nor significant temperature influence in the inverter chain. This is due to the small electron trapping during switching with respect to ΔNit and charge detrapping (Vg$_{Det}$ phase) and to the smaller I-V sensitivity to a given damage at HT. In contrast in ring C (Wp=1μm), the turn-around at 125°C occurs at a shorter time due to the smaller effect of the electron trapping at elevated temperature with respect to RT results (Fig.4). This reduces the time window at which the compensation occurs. This behavior is also enhanced by the cumulative contribution of NBTI damage on the current reduction in P-MOSFET during the high state of the input pulse. At –25°C the F_{osc} change exhibits a smaller turn-around effect compared to RT results as the trapping is increased and interacts with the field-assisted

discharge of the electron traps at low temperature.

3. CONCLUSION

New interacting hot-carrier damage mechanisms have been identified as a function of temperature in actual CMOS circuits. These effects arise from the compensation behaviour between P- and N-MOSFET's degradation found in ring oscillator with narrow p-devices. This is explained by the charge trapping-detrapping balance and interface trap generation leading to turn-around effects in the degradation of the oscillator frequency. The turn-around effect is enhanced at low and high temperature with the detrapping effects and the influence of NBTI damage during AC operation.

REFERENCES

[1] M.Brox, W. Weber, Trans. Electron Dev., Vol. 38, p. 1852, 1991.

[2] K. R. Mistry, B.S. Doyle, Trans. Electron Dev., Vol. 40, p. 96, 1993.

[3] R.Bellens, G. Groeseneken, P. Heremans, H.E. Maes, Trans. Electron Dev., Vol. 41, p. 1421, 1994.

[4] B.S.Doyle, B.J. Fisbein, K.R. Mistry, in IEDM Proc., p. 529, 1991.

[5] P.Heremans, R. Bellens, G. Groeseneken, A.V. Schwerin, H.E. Maes, M. Brox, W. Weber, Van Nostrand Reinhold NY , 1992.

[6] A. Bravaix, D Vuillaume, D. Goguenheim, V. Lasserre, M. Haond, in IEDM Proc., p. 873, 1996.

[7] K.N.Quader, P. Feng, J.T. Yue, P.K. Ko, C. Hu, Trans. Electron Dev., Vol. 41, p. 681, 1994.

A study of proton generation in Si/SiO$_2$/Si structures

V. Girault[a], R. A. B. Devine[a], W. L. Warren[b] and K. Vanheusden[c].

[a]France Telecom-CNET,
BP 98, 38243 Meylan, France.

[b]Sandia Laboratories,
Albuquerque, NM USA.

[c]US Air Force Research Laboratories,
Kirtland AFB, NM USA.

The results of a study of the proton generation reaction in Si/SiO$_2$/Si structures during forming gas (N$_2$, H$_2$) annealing are presented. The main objective of the experiment was to determine the type of the reaction, i.e. to ascertain whether the oxide surface or oxide volume is predominant. We conclude that the oxide surface related generation mechanism appears to be the most important. At the same time as protons are generated, we observed a change in the top Si layer resistance. This may be associated with the reaction with hydrogen during the forming gas anneal.

1. INTRODUCTION

The present work endevours to bring some new elements to the understanding of a new type of non-volatile memory called the proton memory [1], whose basis is the use of mobile protons generated in the gate oxide of a classical Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET). This memory is potentially a multi-level system which has good stability, very low energy consumption and very simple technological fabrication. For a two level system the states are obtained by placing all the protons either at the gate/oxide interface of a MOSFET (state OFF for an n-channel device, state ON for a p-channel one) or at the oxide/substrate interface (state ON for an n-channel device, state OFF for a p-channel one). After their positioning the protons remain immobile without the requirement of a "maintaining" polarisation, thus resulting in very low energy consumption. Starting from a classical MOSFET device process, in principle the only necessary technological step is a forming gas anneal (N$_2$, H$_2$: 92, 8) around 700°C [2] to generate the protons.

The mechanism of proton creation due to annealing is presently unknown. In this work we have carried out a series of experiments designed to increase our understanding of relevant parameters both to help in identifying physically relevant mechanisms and to enable us to lay down a set of technologically relevant conditions.

2. EXPERIMENT

For the present experiments we used two types of commercially available silicon on insulator (SOI) substrates manufactured using the Separation by IMplantation of OXygen (SIMOX) method. The buried oxide thicknesses were 400nm and 80nm. In both cases the thickness of the top Si layer was around 20 nm. Isolated Si islands ~ 0.8mm×2.5mm were defined in the Si top layer using simple photolithography, then excess Si was removed by etching in XeF$_2$ gas. Pseudo-MOSFET [3] transistors were constructed by placing two tungsten needles (spacing ~ 1 mm) as source and drain electrodes on the Si island and using the Si substrate as the gate electrode.

2.1. Oxide modification

The samples were annealed at 700°C and quenched to room temperature by rapid pulling to the extremity of the furnace tube. Anneals were accumulated in steps of 4 minutes. After each anneal

the quantities of mobile and fixed H^+ ions were estimated using electrical measurements with the pseudo-MOSFET. Two measurements were required to determine the mobile H^+ density. First the "gate (substrate)" polarisation was swept from high positive voltage to high negative voltage and the source-drain current (I_{DS}) between the tungsten probes in the top Si island layer measured, for a source-drain voltage of 0.2V. Secondly, the voltage bias was swept in the opposite sense. To be sure that the proton population was either all at the gate electrode or at the oxide/island interface, extensive periods of pre-polarisation were used before sweeping the gate voltage. Typically, for the down ramp (sweep from positive voltage to negative) we polarised the 400nm oxide pseudo-MOSFET for 300s at 70V. The 80nm oxide pseudo-MOSFET was polarised at 10 V. The source-drain current, I_{DS}, was then measured at each gate voltage. Note that the pre-polarisation stage was undergone for each subsequent gate voltage data point down to –100V (for the case of the 400nm oxide) or –20V (for the 80nm oxide).

This first measurement allowed us to determine a threshold voltage taking into account the fixed and mobile charges together.

$$V_T^{DOWN} = V_{T,mobile} + V_{T,fixed} \qquad (1)$$

For the second ramp (up) the gate voltage was held at –100V (-10V) for 300s for the 400nm (80nm) samples, this voltage blocked the protons at the gate/oxide interface. The current I_{DS} was measured at the gate voltage change, as before, by applying a superior voltage compared to –70V and -10V depending on the samples. We repeated the measurement mode for many different voltages to reach 0V. From this sweep, a threshold voltage only depending upon the contribution of the fixed charges was determined.

$$V_T^{UP} = V_{T,fixed} \qquad (2)$$

The quantity of mobile H^+ charges was then calculated from the difference between V_T^{DOWN} and V_T^{UP}.

$$V_{T,mobile} = V_T^{DOWN} - V_T^{UP} \qquad (3)$$

The current I_{DS} and the voltage applied on the gate were measured using a Keithley 617 electrometer, the voltage source was a HP 6628A DC power supply and a HP 900, series 300 computer was used to pilot the measuring instruments.

2.2. Silicon modification

Since we were interested in the oxide modifications through the forming gas anneal, this study lead us to investigate as well changes in the Si film. We measured its resistance with the same two-probe system by applying a V_{DS} voltage between the source and the drain and by measuring the current I_{DS} before and after a forming gas anneal at 700°C.

3. RESULTS

Figures 1 and 2 show the results of I_{DS} versus gate voltage (V_G) for down and up sweep modes in 400nm thick oxide samples annealed for different times up to 56 minutes. Similar curves were obtained for the 80 nm thick oxide samples (though the voltage sweep range was smaller).

Figure 1. Down sweep mode measurement of I_{DS} (V_G) for different anneal times (curves separated by 4 minute increments starting at 4min up to 36min, one at 46min and one at 56min).

From figure 2 we see that for long time anneals the current I_{DS} never goes to zero for any value of V_G. This is shown more clearly in Fig. 3 where down sweep and up sweep curves of I_{DS} are shown for an 80nm oxide sample annealed for a total of 36 minutes. Included is a curve for the sample prior to annealing where clear current cancellation is observed.
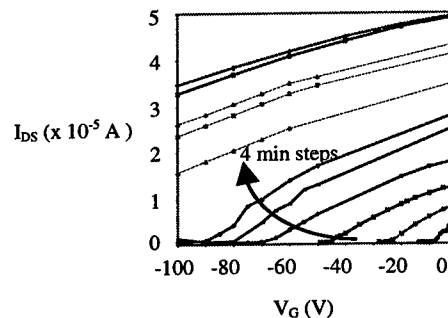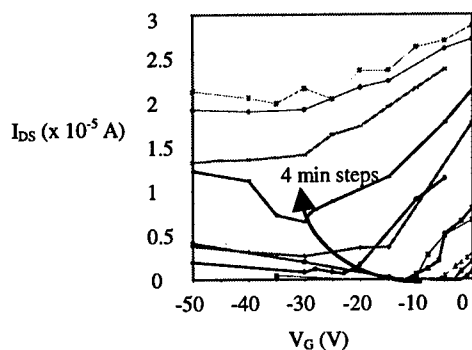
Figure 2. Up sweep mode measurement of I_{DS} (V_G) for different anneal times (curves separated by 4 minute increments starting at 4min up to 36min, one at 46min and one at 56min).
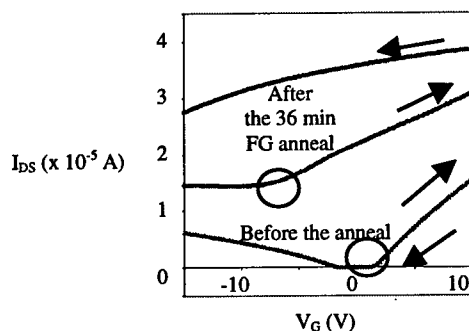

Figure 3. I_{DS} (V_G) curves for down and up sweep modes, before and after a 36min forming gas anneal at 700°C.
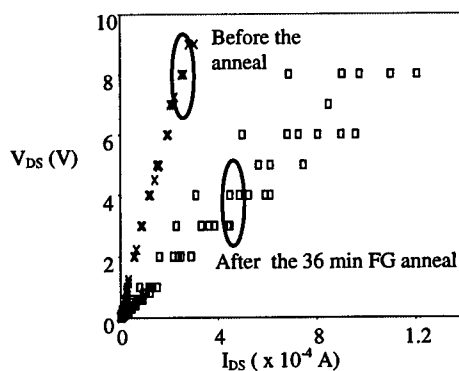

Figure 4. Si sheet resistance evolution shown in V_{DS} (I_{DS}) measurement before and after a 36min forming gas anneal at 700°C.

The resistivity of the Si island film was measured for this same sample, the voltage (V_{DS})/current (I_{DS}) curves are shown in Fig. 4. We see clearly that

extended annealing in an H_2 ambient results in a significant decrease in the inherent resistivity of the film.

4. DISCUSSION

From figures 1, 2 and the data for the 80nm samples we calculated the quantities of mobile and fixed charges at each step of experiment. We then distinguish the two different kinetics associated to the two types of charges. Figure 5 shows the kinetics of creation of fixed and mobile charges for the two types of samples.

There are several observations. First, it would appear that the density of fixed oxide charges saturates after ~ 25 minutes annealing at a value ~ $2 \times 10^{12} cm^{-2}$. This value is independent of the oxide thickness and suggests that it is indicative of the maximum number of interfacial sites available which can trap protons to form a pseudo "fixed charge" state (we do not yet know the temperature stability of these charges). If the rate of trapping of protons at the fixed sites were proportional to the total number created (i.e. the volume of the oxide) then we would expect the 400nm oxide behaviour to be more rapid than for the 80nm, this is not the case. In consequence, it appears that fixed charge in the form of protons trapped near the oxide/Si interface is a purely interfacial effect.

For the mobile charges we observe no clear saturation effect for anneal times up to ~ 30 minutes. Measurement beyond this value is not possible for the 400nm thick oxide because our voltage source could not go beyond –100V which was necessary for such a charge density ($>3 \times 10^{12} cm^{-2}$). For the 80nm samples there was a risk of oxide break down. The data shown in Fig.5 for 400nm samples annealed 8 days after fabrication and 80nm samples annealed 26 days after fabrication do show different mobile charge creation kinetics. If we examine the mobile charge creation after 10 minutes of annealing we observe $\rho(400nm)/\rho(80nm) \cong 10$, this is a factor of two larger than the volume ratios. For 20min anneals, $\rho(400nm)/\rho(80nm) \cong 3$, now smaller than the oxide volume ratios. Neither of these results enable us to distinguish clearly if the mobile proton generation is a volume or a surface effect. However, since proton generation involves the release of an electron and the capture of a proton, probably at an Si-O-Si bonded oxygen[2], it is unlikely that the oxide volume is involved. There would be

no escape mechanism for the electron and proton-electron recombination would occur. Therefore, we argue that proton generation results from surface/interface mediated reactions.

Finally, the data presented in Fig.5 do give evidence for a sample age effect in proton generation efficiency. We observe that proton generation in samples annealed 26 days after fabrication is significantly less efficient than samples annealed 8 days after fabrication. This can be criticised since the oxide thickness differed in the two samples. However, we have performed less extensive measurements (fixed anneal time) on 400nm SIMOX samples and these clearly show reduced mobile proton generation in samples left extended times after preparation, before annealing.
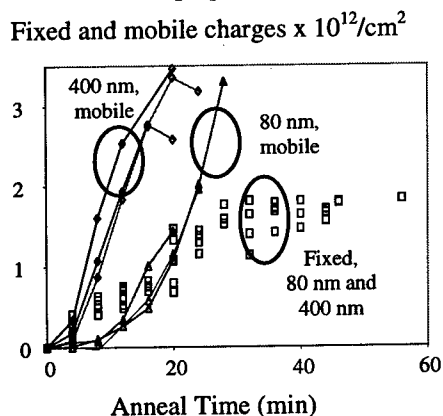
Fixed and mobile charges x 10^{12}/cm^2

Figure 5. Density of fixed and mobile charges versus anneal time for 400nm and 80nm thick oxide samples in the 4min step anneal experiments.

Figures 1 and 2 also show an unexpected phenomenon, the non-cancellation of the current I_{DS} at any applied gate voltage at the same time as an increase of its minimum residual level. Furthermore, whereas the pseudo-MOSFET should be always controlled by the applied gate voltage [3], we see that after about 30 min of anneal the current I_{DS} is no longer modulated by the gate voltage. It seems that a leakage current superposes on the normal behaviour of a pseudo-MOSFET and becomes preponderant.

No leakage current was noticed through the oxide and anneals carried out for various times under vacuum at 700°C eliminated the idea that the temperature alone could be at the origin of a modification of the Si film properties. Consequently, we conclude that the residual current level is due to H_2 annealing induced increased conductivity in the

Si island film, this is confirmed by the resistance data presented in Fig. 4. Such an effect has been observed previously [4]. These observations are still under investigation.

5. CONCLUSION

The question of volume or surface origin in proton creation has been investigated in this work. By using two different oxide thicknesses in the pseudo-MOSFETs in the experiment, we have obtained some evidence that proton creation is not a volume phenomenon but rather a reaction taking place at the oxide/silicon interfaces.

Two reaction kinetics have been observed, the first associated with fixed charge creation, which saturates after 30min anneal at 700°C in forming gas. The second, relevant to mobile proton creation, has no observed saturation in our experiments.

Comparison of the proton generation kinetics has enabled us to demonstrate the importance of the oxide surface. We have shown that the "age" of the sample, prepared prior to N_2H_2 annealing is extremely important in mobile charge generation. The larger the time lapse, the less efficient the proton creation. Note that this time lapse does not seem to affect fixed charge creation. The former result is symptomatic of the surface becomes less active with time, presumably due to pollution. Further experiments are necessary to elucidate this.

Finally, a secondary effect has been detected, annealing induced modification of the Si film resistivity. Forming gas annealing appears to result in a decrease of the sheet resistance, presumably due to solution of H/H_2 in the Si. This result also requires further investigation.

REFERENCES

1. K. Vanheusden, W. L. Warren, R. A. B. Devine, D. M. Fleetwood, J. R. Schwank, M. R. Shaneyfelt, P. S Winokur, and Z. J. Lemnios, Nature, 386 (1997) 587.
2. V. Girault and R. A. B Devine, J. Non Cryst. Sol., (in publication, 1999).
3. S. Cristoloveanu and S. S. Li, "Electrical Characterization of Silicon-On-Insulator Materials and Devices", (Kluwer, Boston, 1995) 104.
4. T. I. Kamins and K. Nauka, Electrochem. Sol. State. Lett., 1 (2) (1998) 100.

A common framework for soft and hard breakdown in ultrathin oxides based on the theory of point contact conduction

E. Miranda[*], J. Suñé, R. Rodríguez, M. Nafría and X. Aymerich

Departament d'Enginyeria Electrònica. Universitat Autònoma de Barcelona,
08193-Bellaterra, Spain.

We present an alternative explanation for the soft (SBD) and hard (HBD) breakdown conduction modes in ultrathin gate oxides based on the physics of mesoscopic conductors. It is shown that the Landauer approach applied to this system not only gives the observed HBD conductance values, which are those expected for a ballistic constriction in the linear and nonlinear regimes, but also a plausible mechanism for the SBD conduction.

1. INTRODUCTION

It has been long recognized that the breakdown of a gate oxide in a MOS structure consists in the formation of a low-resistance conduction path running between the electrodes. But, what else is beyond this widely invoked assertion? Has the highly local nature of this event any implication in the charge transport characteristics?. In this regard, we have recently shown that the I-V characteristic of a broken down (HBD) oxide can be understood in terms of the point contact conduction theory [1]. In addition, other failure mechanisms such as Stress Induced Leakage Current (SILC) and soft breakdown (SBD) have been experimentally found in these oxides. In this paper, we propose a common framework for SBD and HBD in which the electron confinement within the breakdown conduction path plays a fundamental role. Theoretical considerations are based on the linear and nonlinear Landauer approaches to electrical transport in narrow constrictions [2,3].

2. EXPERIMENTAL AND DISCUSSION

The devices used in this study are n$^+$ polysilicon-gate MOS capacitors fabricated on N- and P-type (100) Si substrates with oxide thickness in the range 3-5 nm and areas of about 10^{-5}-10^{-4} cm^2. Fig.1 shows the typical failure modes of an ultrathin gate oxide.
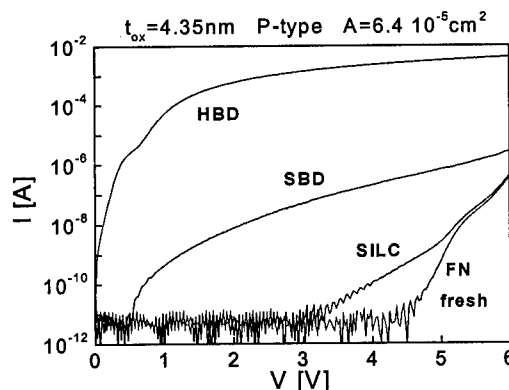
Figure 1. Typical stages in the I-V characteristic as the degradation proceeds.

Several points must be highlighted: i) SILC increases continuously during stress and it has been usually ascribed to trap assisted tunneling uniformly distributed over the capacitor area; ii) SBD corresponds to an abrupt change in the conduction mechanism which exhibits many striking features: as shown in Fig. 2, it is essentially independent of the device area (local area phenomenon) and oxide thickness (localized blocking

mechanism). Besides, it does not depend on the sign of the applied voltage (symmetrical conduction structure) -see Fig.3-; iii) HBD is also associated with a sudden and dramatic local change in the conduction mechanism.
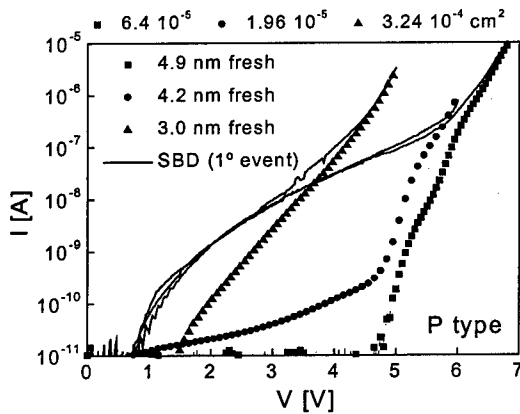

Figure 2. Fresh and SBD I-V characteristics measured on oxides with different thickness and gate area.
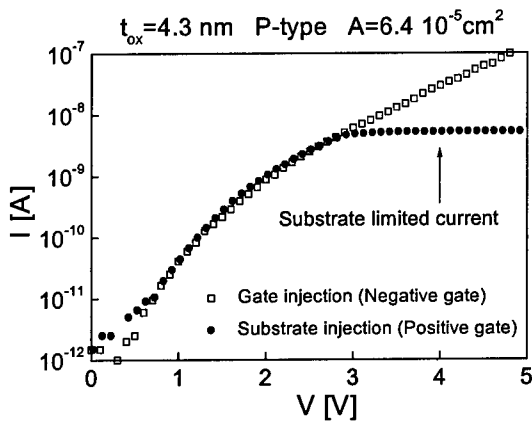

Figure 3. Effect of the bias sign on the SBD I-V characteristic.

Electrical and optical experiments not only reveal that several SBD but also HBD spots can be induced on the same sample by a high-field stressing -see Fig.4-. For the latter mode, the differential conductances as a function of the applied bias (measured after the detection of each resistance jump) -see Fig.5- exhibit ohmic-type plateaus of the order of the quantum conductance unit $G_0=2e^2/h$ (≈ 13 kΩ⁻

¹), e and h being the electron charge and the Planck's constant respectively. This can be recognized as the signature of point contact conduction [4].
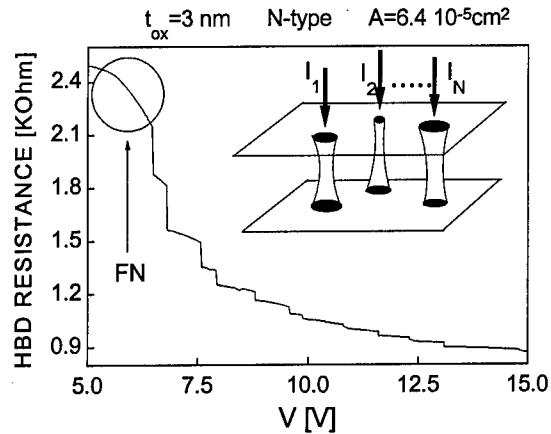
Figure 4. Steplike behavior of the HBD resistance (V/I) as a function of the applied voltage.
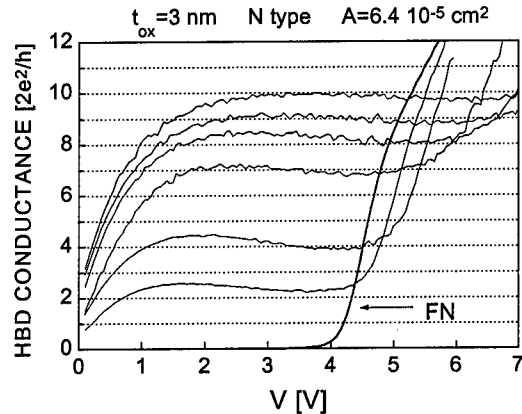
Figura 5. Ohmic-type plateaus in the HBD conductance (dI/dV) characteristics.

At the lowest voltages the curves are affected by the silicon band bending while the steep current increase for $V_{app}>4V$ is due to Fowler-Nordheim (FN) tunneling through the non-damaged capacitor area. It is worth mentioning that in a previous work [1], we ascribed this steep current increase to the appearance of a valence band to conduction band current component. Each one of the curves shown in Fig. 5 correspond to the overlap of the currents associated with the

open spots up to the moment of the measurement and not to individual characteristics. This is a distinctive feature of the system under study which is different from other point contact systems such as those based on split-gate devices, STM or mechanically controllable break-junctions [4-6]. In these cases, the current through a single constriction is always analyzed and, generally, conclusions are drawn from statistical data ensembles. When the difference between successive curves in Fig.5 is considered, conductance plateaus of approximately $1G_0$, $2G_0$ and $3G_0$ are obtained. As will be discussed below, these are the values expected for a ballistic constriction in the quantum approach.
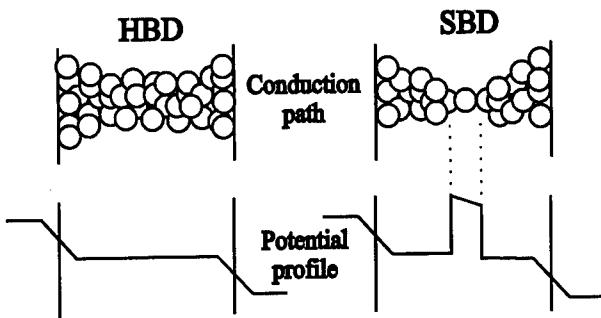

Figure 6. Conduction path and idealized electrical potential profile between the electrodes.

Figure 6 represents an oversimplified version of the HBD and SBD conduction paths. If ballistic transport is assumed within the constriction, the electrical potential largely drops at the interfaces due to mode mismatching effects [2,7,8]. Now consider, also for simplicity, that the geometry of the lateral confining potential is cylindrical. As in an infinite quantum well, the splitting between transverse electron levels depends, crudely speaking, on the inverse of the constriction area. Those discrete states below the energy of the injected electrons behave as transmitting channels, but those above the electron energy act as backscattering channels. If the cross-section of the constriction is very small we can have no high-transmission channels and conduction can only be due to tunneling. Theoretical analysis reveals that the

narrowest part of the constriction can behave as an effective one-dimensional tunneling barrier as schematically depicted in Fig.6. This is totally consistent with the SBD experimental features described earlier. Furthermore, this mechanism has been thoroughly studied in nanowires by means of molecular dynamics simulations [6] and in quantum point contacts by solving Schrödinger's equation in nonuniform constrictions [9]. It is worth emphasizing that this tunneling barrier has nothing to do with previous approaches based on direct tunneling calculations through thinned down SiO_2 barriers. In this regard, some of the fitting values considered for the cathode barrier height (6.2 eV [10], 4.2 eV [11]) were found to be unrealistic to be related to a conduction band offset between Si and SiO_2. Perhaps they might be better understood in terms of a cross-section related energy barrier.

We want to remark that our approach is radically different from those intended before. We are considering the transport properties of an atomic-sized conducting chain across the insulator. Similar configurations were demonstrated to have its conductance quantized, the number of atoms (defects) contacting the electrodes and its structural arrangement being the main factors governing the conductance. The phenomenon has also been detected both at room temperature and high applied bias [12]. In this framework, the SBD mode appears as a particular case of the theory linked to the lateral size of confinement. As mentioned in the Introduction, this approach is fairly supported by the Landauer's conductance formula:

$$G = \left(2e^2/h\right)Tr(t^+t) \qquad (1)$$

where Tr is the trace and t the transmission probability matrix. For ballistic conduction $((t^+t)_{ij}=\delta_{ij})$ the conductance of the system reads

$$G = \left(2e^2/h\right)N \qquad (2)$$

with N the number of available conduction channels. If this number is large enough, the

latter expression reduces to the Sharvin's conductance formula corresponding to a classical point contact. Eqn. (2) can be further developed to account for the nonlinear transport regime [13]:

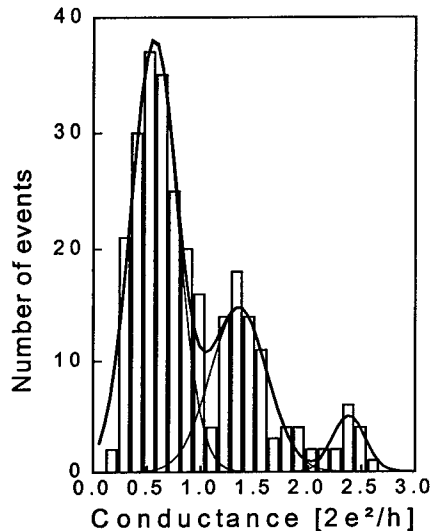$$G = \left(2e^2/h\right)\{\beta N[\rightarrow] + (1-\beta)N[\leftarrow]\} \qquad (3)$$

Figure 7. Histogram of conductance values (normalized to $G_o = 2e^2/h$) corresponding to single breakdown spots.

where β is a parameter between 0 and 1 describing the fraction of the applied voltage that drops on the connection between the cathode and the constriction. If we consider a symmetrical constriction then $\beta = \frac{1}{2}$. $N[\rightarrow]$ and $N[\leftarrow]$ are the number of available right- and left-going transverse conduction modes respectively. If the difference between these numbers is an odd number, Eqn. (3) predicts that half plateaus will appear at the values of the half-integer multiples of G_0. This can explain the distribution of conductance data shown in Fig.7. These values were obtained from measurements similar to those shown in Fig. 5 after subtracting successive curves. We have detected up to fifteen HBD events in the same sample (more than fifty samples) but we have only considered for the histogram in Fig. 7 those curves exhibiting clear conductance plateaus. Nevertheless, it is worth pointing out that the large contribution of half integer

conductance plateaus might be also associated with preferred atomic configurations within the constrictions (some authors have claimed that the conductance per atom is of the order of $G_0/2$ [14]).

3. CONCLUSION

The electrical transport through quantum constrictions has been extensively investigated in the last decade. However, the concepts involved have never been applied to a broken down oxide in a MOS structure. In this work, we presented experimental results compatible with the predictions of the point contact conduction theory. We have also suggested, for the first time, a new mechanism for the SBD conduction consistent with this theoretical framework.

REFERENCES

1. Suñé et al., IEDM'98, p. 191.
2. Landauer, Phys. B - Condensed Matter 68, 217 (1987).
3. Imry, in *Directions in Condensed Matter Physics*, Ed. G. Grinstein and G. Mazenko (World Scientific, Singapore, 1986), p. 101.
4. Krans et al., Nature 375, 767 (1995).
5. Thomas et al., Phys. Rev. B 58, 4846 (1998).
6. Brandbyge et al., Phys. Rev. B 52, 8499 (1995).
7. Payne, J. Phys. Con. Mat. 1, 4931 (1989).
8. Pernas et al.,Phys. Rev. B 41, 8553 (1990).
9. Tekman et al.,Phys. Rev. B 43, 7145 (1991).
10. Haimaoui et al., Micro. Engin. 36, 157 (1997).
11. Houssa et al., J. Appl. Phys. 84, 4351 (1998).
12. Yasuda et al., Phys. Rev. B 56, 1069 (1997).
13. Xu, Phys. Rev. B 47, 15630 (1993).
14. Olesen et al., Phys. Rev. Lett. 72, 2251 (1994).

Two-fold coordinated nitrogen atom: an electron trap in MOS devices with silicon oxynitride as the gate dielectric

Yu.N. Morokov[a], Yu.N. Novikov[b], V.A. Gritsenko[b], and H. Wong[c]

[a]Institute of Computational Technologies, Novosibirsk, 630090, Russia
[b]Institute of Semiconductor Physics, Novosibirsk, 630090, Russia
[c]Department of Electronic Engineering, Cityu, Tat Chee Avenue, Hong Kong

Having conducted semiempirical quantum-chemical simulation (MINDO/3) of several clusters at different charge states, we identify that the two-fold coordinated nitrogen atom with an un-paired electron ($\equiv Si_2N\bullet$) is the most responsible trap center for the observation of large electronic capturing in SiO_xN_y. Our calculations also show that electron localized in this defect will result in spin dissipation. Trap formation and removal mechanisms during nitridation and re-oxidation are also discussed in this work.

1. INTRODUCTION

The properties of gate dielectric determine the reability of metal-oxide-semiconductor (MOS) devices. It was suggested that the conventional silicon dioxide gate dielectric should be replaced by silicon oxynitride (SiO_xN_y) in the future nanoscale MOSFET to minimize the hole and electron trapping on Si-Si defects at the Si/SiO_2 interface. It was demonstrated that the Si-Si defects could be removed effectively by high-temperature annealing of thermal SiO_2 in ammonia. This process can be described as follow [1]:

$$3 \equiv Si\text{-}Si \equiv + 2 N \rightarrow 2 \equiv Si_3N \qquad (1)$$
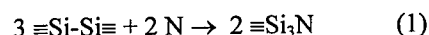
The symbols –, = and $\equiv$ in (1) represent one single bond, two single bonds, and three single bonds, respectively. With oxynitride as the gate dielectric, hole trapping was reduced considerably. However, large amount of electron traps are still found in oxynitride and degradation of MOSFET due to the capturing of hot channel electrons is often reported. In present paper we show, using semiempirical quantum-chemical method MINDO/3, that the two-fold coordinated nitrogen atom with unpaired electron $\equiv Si_2N\bullet$ is the trap responsible for the capturing of electrons in ammonia-nitrided SiO_xN_y. Here $\bullet$ denotes the one unpaired electron.

MINDO/3 had been shown to be a powerful and informative method for simulating electronic structure and bulk defect in solid [2-3].

2. SIMULATION RESULTS

In this work, we used the cluster approximation to study the electronic structure of several different clusters in silicon oxynitride. Atomic relaxation in different charge states of defect was considered in the simulation. To simulate the effect of chemical composition on the capturing properties of the $\equiv Si_2N\bullet$ defect in silicon oxynitride, clusters with different numbers of oxygen and nitrogen atoms, i.e. $\bullet NSi_2(N_6)SiH_{12}$, $\bullet NSi_2(N_4O_2)SiH_{10}$, and $\bullet NSi_2(O_6)H_6$, in the second coordination sphere were considered. For simulation of the Si_3N_4 bulk electronic structure we used the $Si_{20}N_{28}H_{36}$ cluster. The atomic structures of these clusters are shown in Fig.1. MINDO/3 parameters used in this work are same as those used in Ref. [3-4] which are $\alpha_{SiN} = 1.053011$ and $\beta_{SiN} = 0.434749$.

Simulation of the neutral $\equiv Si_2N\bullet$ defect shows that unpaired electron is localized for all considered clusters in the N $2p_\pi$ nitrogen non-bonding orbital which is oriented normally to the Si_2N plane. This result agrees with the previous theoretical simulation of this defect in silicon nitride [5]. As

obtained in ref. [6] for the $\equiv$Si$_2$N$\bullet$ defect in Si$_3$N$_4$, the wave function of the unpaired electron consists from 90% of p-type and 10% of s-type atomic functions. Similar result was also obtained from the analysis of ESR signal hyperfine splitting in SiO$_x$N$_y$ [7]. Our calculations show that the distribution of the unpaired electron in the defect nitrogen atom are 96% and 4% for p-type and s-type wave function, respectively. The difference between experimental and simulation result may be due to the small cluster being used in the simulation.
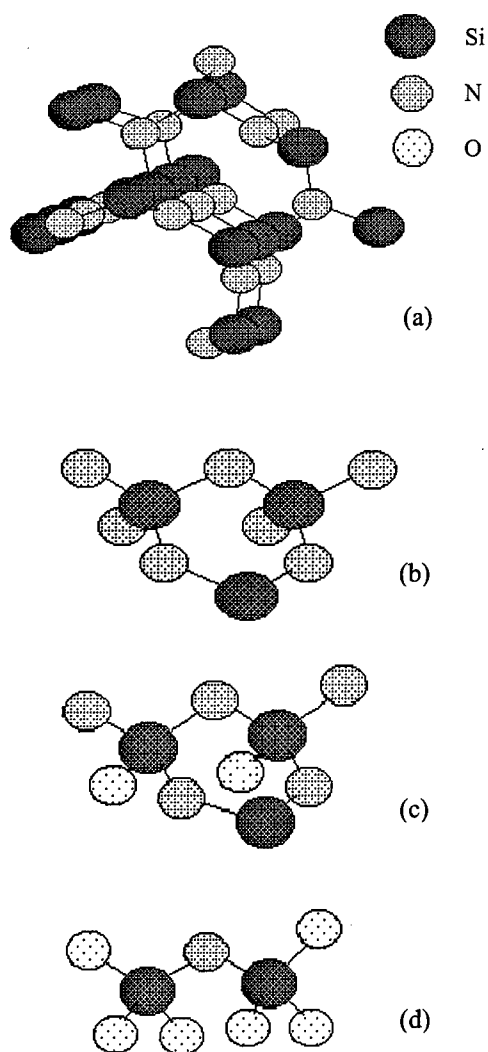
It was found experimentally that cross-section for the capture of electron (σ) in SiO$_x$N$_y$ is about 10^{-17} cm^2. This value corresponds to an effective capture radius of the trap $R = \sqrt{\sigma/\pi} \approx 0.2$ Å. This value reflects the highly localized character of the non-bonding N 2p$_\pi$ wave function and supports our simulation results.

We also estimated the energy gain for the electron or hole capturing in the $\equiv$Si$_2$N$\bullet$ defect by calculation of the differences between the total energies of clusters in different charge states. Figure 2 shows the calculated energy diagram for $\equiv$Si$_2$N$\bullet$ and $\equiv$SiO$\bullet$ defects in oxynitride. Similar defects in oxide and nitride are also shown for comparison. The obtained values of the energy gain for the capture of electron from the bottom of the conduction band (E$_c$) is about 1.0 eV. That is the thermal delocalization electron trap energy is about 1.0 eV. However, the capture of hole from the top of the valence band (E$_v$) is energetically unfavorable. In addition, simulation results also show that the electron trap energy is almost independent on the cluster chemical composition. The energy gain results indicated that the $\equiv$Si$_2$N$\bullet$ defect in silicon nitride and oxynitride with high concentration of nitrogen cannot capture a hole but an electron. The captured electron is localized in the N 2p$_\pi$ non-bonding orbital of the $\equiv$Si$_2$N$\bullet$ defect (see Fig.3).

Fig. 1. Clusters used in simulating the defect capturing properties of (a) Si$_{20}$N$_{28}$H$_{36}$, (b) $\bullet$NSi$_2$(N$_6$)SiH$_{12}$, (c) $\bullet$NSi$_2$(N$_4$O$_2$)SiH$_{10}$, and (d) $\bullet$NSi$_2$(O$_6$)H$_6$.
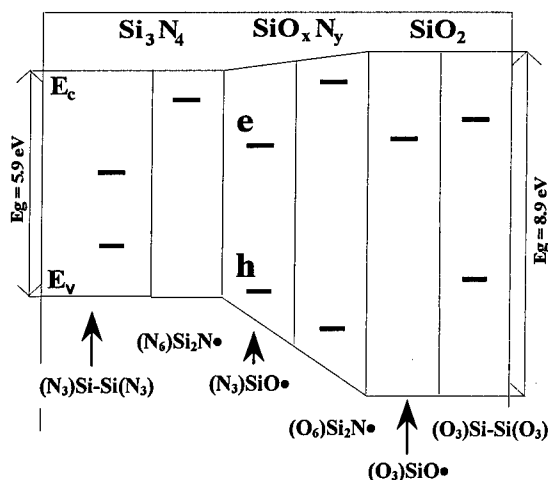

Fig. 2. Calculated energy diagram for the major defects in oxynitride. Similar defects in oxide and nitride are also shown for comparison.
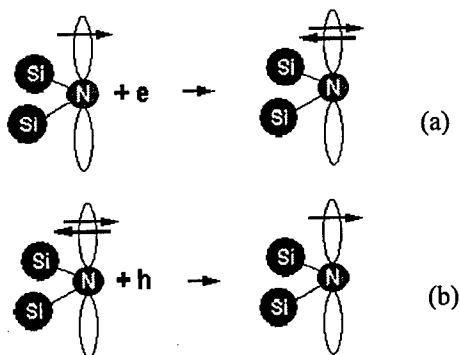
Fig. 3. Model of three-fold coordinated nitrogen atom as an electron trap in silicon nitride and oxynitride (a) capture of electron, and (b) capture of hole.
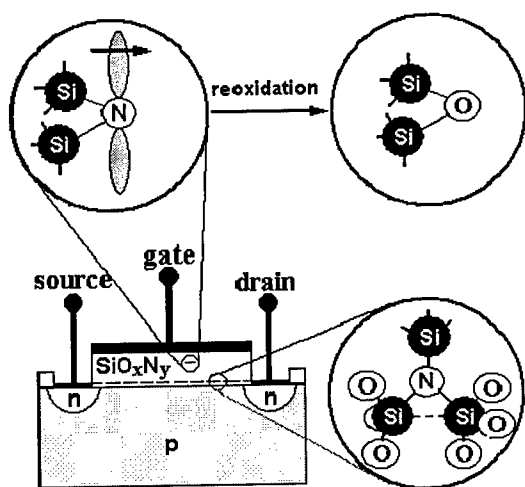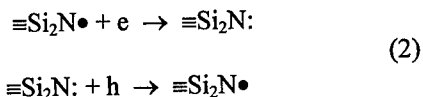

Fig. 4. Illustration of $\equiv Si_2N\bullet$ centers in oxynitride and the removal of electron traps during re-oxidation of oxynitride in MOSFET.

3. DISCUSSION

According to the obtained results, the capture (localization) and following decapture (delocalization) of electron on the $\equiv Si_2N\bullet$ defect in

Si_3N_4 and SiO_xN_y are described by the following reactions:

$$\equiv Si_2N\bullet + e \rightarrow \equiv Si_2N:$$

$$\equiv Si_2N: + h \rightarrow \equiv Si_2N\bullet \qquad (2)$$

Electron localization results in the transfer of paramagnetic defect $\equiv Si_2N\bullet$ to diamagnetic defect $\equiv Si_2N:$. The corresponding model is picturized in Fig.3. Our simulation predicts the electron paramagnetic signal dissipation at the electron capture by three-coordinated nitrogen atom, which was observed experimentally in nitrided oxide [7].

It has been confirmed that the $\equiv Si_2N\bullet$ is the major defect in SiO_xN_y. The $\equiv Si_2N\bullet$ defect creation in SiO_xN_y could be resulted from the breaking of the Si_2N-H bond according to the following reaction

$$\equiv Si_2NH \rightarrow \equiv Si_2N\bullet + H \qquad (3)$$
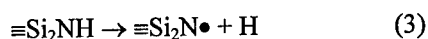
Another mechanism for the creation of $\equiv Si_2N\bullet$ defect is the breaking of N-N bond with the following reaction

$$\equiv Si_2N\text{-}NSi_2\equiv \rightarrow \equiv Si_2N\bullet + \bullet NSi_2\equiv \qquad (4)$$
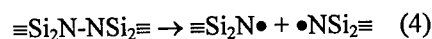
The $\equiv Si_2N\bullet$ defect can be removed by re-oxidation of gate oxynitride, i.e.

$$\equiv Si_2N\bullet + O \rightarrow \equiv Si\text{-}O\text{-}Si\equiv + N \qquad (5)$$

Since the nitrogen atom and the nitrogen defect $N\bullet$ have the same coordination number, this replacement can occur without other atom re-arrangement (see Fig. 4).

4. CONCLUSION

The electronic structure of two-fold coordinated N atom ($\equiv Si_2N\bullet$) in different charge states in SiO_xN_y was simulated and the results indicate that this defect is an electron trap. This observation agrees with the experimental results [7]. In addition, our results also agree with Powell and Robertson [8] that the negatively charged nitrogen defect $\equiv Si_2N:$ can be a hole trap. We also rule out the possibility that the neutral $\equiv Si_2N\bullet$ defect may also

act as a hole trap in Si_3N_4 which was proposed by Kirk [9]. On the other hand, simulation results also suggest that the electron localization by the $\equiv Si_2N\bullet$ defect will result in spin dissipation. This effect was experimentally observed earlier [10]. Based on the simulation results, the nature of electron traps removal during gate oxynitride reoxidation is also discussed in this work.

REFERENCES

1. V.A. Gritsenko, J. B. Xu, I. H. Wilson, R. M. Kwok, Y.H. Ng, *Phys. Rev. Lett.* 81 (1998) 1054-1057.
2. A. H. Edwards and W. B. Fowler, J. Phys. Chem. Solids, 46 (1985) 841-857.
3. V. A. Gritsenko, Yu. N. Novikov, Yu. N. Morokov and H. Wong, Microelectron. Reliab. 38 (1998) 1457-1464.
4. V. A. Gritsenko, Yu. N. Morokov and Yu. N. Novikov, Solid State Phys. 39 (1997) 1191-1196.
5. W. L. Warren, J. Robertson, and J. Kanicki, J. *Appl. Phys. Lett.* 63 (1993) 2685-2687.
6. W. L. Warren, P. M. Lenahan, and S. E. Curry, *Phys. Rev. Lett.* 65 (1990) 207-210.
7. J. T. Yount and P. M. Lenahan, *J. Non-Cryst. Solids*, 164-166 (1993) 1069-1072.
8. J. Robertson and M. J. Powell, *Appl. Phys. Lett.* 44 (1984) 415-417.
9. C. T. Kirk, *J. Appl. Phys.* 50 (1979) 4190-4195.
10. I. A. Chaiyasena, P. M. Lenahan, and G. J. Dunn, *Appl. Phys. Lett.* 58 (1984) 2141-2143.

MEASUREMENTS

Characterization of Conductive Oxides on Silicon Using Non-Contact Surface Charge Profiling

P. Roman, D-O. Lee, M. Brubaker, E. Kamieniecki* and J. Ruzyllo

Department of Electrical Engineering, Penn State University,
University Park, PA 16802, USA

* QC Solutions, Inc., Woburn, MA 01801, USA.

The density of charge associated with the silicon-oxide structure remains a figure of merit for the process as well as an indicator of the eventual MOS device performance. As the effective oxide thickness drops below 2 nm significant currents limit the effectiveness of oxide charge determination from C-V measurements. One potential solution is the SPV-based SCP (Surface Charge Profiler) method which allows measurement of total oxide/interfacial charge without making a contact to the oxide. In this paper the effectiveness of the SCP in measuring charges in ultra-thin thermal SiO_2 and conductive Ta_2O_5 films is evaluated.

1. INTRODUCTION

The density of charge associated with the silicon-oxide structure is an indicator of the quality of the fabrication process as well as a predictor of the characteristics of a MOSFET. Hence, the measurement of oxide and interface charges in silicon-oxide structures is a very common procedure in process monitoring and diagnostics. Typically, the density of charge is deconvoluted from C-V measurements which are implemented using either a permanent contact to the oxide surface (e.g. poly Si or Al), a temporary solid contact (e.g. Hg probe), or a "charge contact" enforced by corona discharge. In each case due to the potential difference across the oxide an electric field is established. In the case of "conductive" oxides, such as ultra-thin SiO_2 allowing tunneling current, or new generation high-k oxides which may feature substantial leakage current, the determination of the oxide charge from C-V measurements under those circumstances is not effective.

Free from these limitations is the method of Surface Charge Profiling (SCP) which allows measurement of the charge density without any bias, and hence, without any current flow across the oxide. Therefore, measurement of charge density in the oxide is possible regardless of its conductivity which offers new possibilities in process development and diagnostics. The performance of the SCP method in monitoring charge on Si surfaces subjected to various treatments was evaluated earlier [1]. The goal of this experiment was to study the usefulness of the SCP method in monitoring charge density in oxides which by virtue of being too conductive do not allow effective determination of the oxide charge using the conventional C-V method.

Two different aspects of this experiment include: (i) non-contact monitoring of charge associated with thermally grown ultra-thin SiO_2 on Si substrates and (ii) measurements of charge density in Ta_2O_5 on Si as a function of deposition process conditions and interface characteristics. In the former case significant tunnel currents preventing C-V measurements may flow at very low voltages, e.g. for a 2.7 nm thick oxide a current density of 1 A/cm² is reached at 1.5 volts [2]. In the case of 14 nm PECVD Ta_2O_5 the same level of leakage current may be reached at 3 volts [3]. Therefore, in both cases a thorough C-V analysis needed to derive values of oxide charge may not be practical. A premise upon which this experiment was initiated was that with SCP method measurement of surface/oxide charge density is possible down to bare Si surfaces, and hence, the SCP method should allow monitoring of surface charge evolution during the early stages of thermal oxidation of silicon.

2. EXPERIMENTAL PROCEDURES

The SCP is a non-contact method in which the silicon surface is illuminated by a collimated beam of chopped blue light. The acquisition of the resulting ac-SPV signal is accomplished by means of capacitive coupling through an air gap (Fig. 1). The method measures width of the surface depletion region (W_d) and surface recombination lifetime. Since under depletion conditions the surface charge, Q_s, is balanced by the depletion layer charge the surface charge can be determined as $Q_s=N_{sc}W_d$ where N_{sc} represents the known dopant concentration in the space charge region. This charge represents the total charge in equilibrium with the depletion layer, and hence, includes the total oxide and surface state charges. The surface carrier lifetime is measured from the phase delay of the SPV signal and is thought to represent surface recombination velocity. The principles of SCP measurement are presented in detail elsewhere [4,5].

The measurements in this study were carried out on thermally grown SiO_2 and Ta_2O_5 reactively sputtered on silicon substrate. Additional measurements were carried out on CVD deposited Ta_2O_5 films. In the case of SiO_2 oxidations in dry oxygen at various temperatures were performed to study oxide charge evolution during the early stage of thermal oxidation of silicon as well as to fabricate Al-gate MOS capacitors for C-V characterization. Similarly, measurements of Ta_2O_5 involved non-contact SCP measurements as well as C-V measurements of Al-gate MOS capacitors.
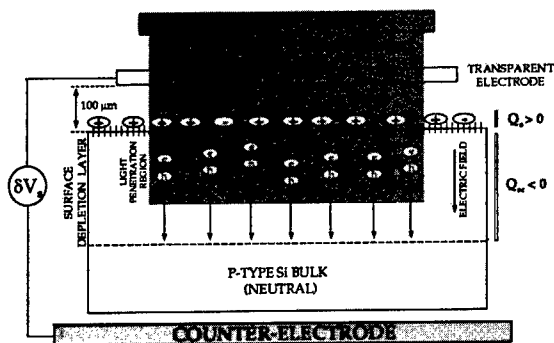
3. RESULTS AND DISCUSSION

Figure 1. SCP schematic.

3.1. SiO_2 Monitoring

During the initial stage of this experiment the correlation between charge densities determined from C-V measurements and SCP measurements was investigated. In the former case density of effective charge was determined from the measured V_{FB} value for an assumed Al-Si work function difference and the known oxide thickness. Changes in charge density were enforced by post-oxidation anneals in nitrogen of as-grown oxides which are known to affect the density of oxide fixed charge [6]. This experiment did not yield conclusive results. Both C-V and SCP measurements resulted in oxide charge densities which were decreasing slightly with the time of the anneal, but the decrease was not enough to determine a consistent correlation between these parameters. The reason for this weak response may be that the temperature (900 °C) of the anneal was insufficient to enforce the desired changes.

In subsequent experiments in this study the attention was focused on measurement of oxide charges using the SCP method. Figure 2 shows evolution of the surface charge during the early stage of thermal oxidation of p-type silicon surfaces. As seen, the density of surface charge immediately drops to half its value after cleaning with only a push into the furnace and remains at this new value with increasing time of oxidation. These data indicate a reduction of the positive surface charge taking place during initiation of the oxidation. It should be pointed out that these changes could be
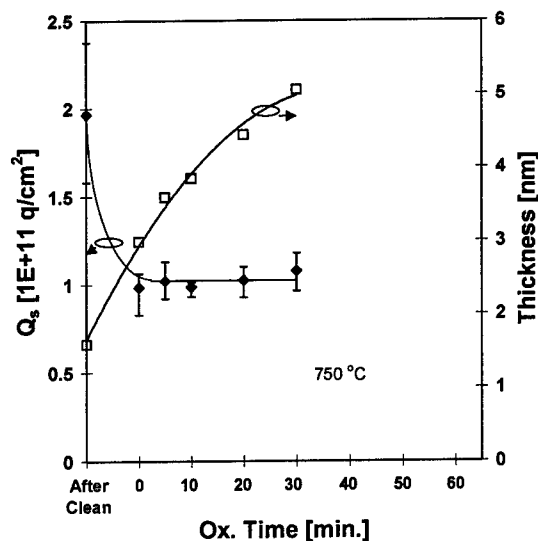

Figure 2. Surface charge as a function of oxidation time.

detected only for Si wafers with a doping level high enough to prevent inversion of the surface by the surface charge.

In an attempt to determine the location of the charge centroid in the oxide an etch back experiment was carried out. In this experiment a 4 nm oxide was etched in steps in HF(1) : H$_2$O(100) solution and oxide charge density was measured using SCP following each etching step. As shown in Fig. 3a the surface charge initially changes only slightly with decreasing thickness of the oxide but at 2 nm begins to increase steadily and becomes equal to ~2E+12 q/cm^2 for a hydrophobic, oxide free surface. This indicates that the charge added during oxidation is contained within approximately 2 nm from the interface. The nature of this charge will be borne out with further investigation. The lifetime, on

the other hand, remains constant until the thickness drops below 1 nm (Fig. 3b). Here, the lifetime begins to increase sharply to 8 μs for an oxide-free surface. This indicates the removal of recombination centers located within this region of the oxide. As seen in this figure prolonged exposure to HF causes the lifetime to decrease, which may be due to surface microroughening. To further investigate the location of the charge in the oxide a 12 nm oxide was subjected to the same etch back experiment. The results shown in Fig. 4 indicate again an increased density of charge (Fig. 4a) and an independent increase in lifetime (Fig. 4b) in the region adjacent to the SiO$_2$-Si interface.

3.2. Ta$_2$O$_5$ Monitoring

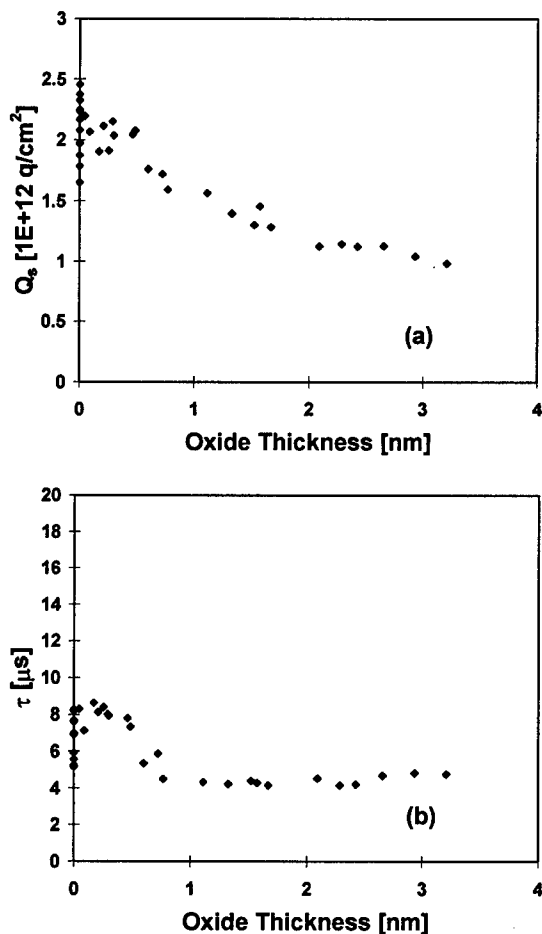In light of growing interest in high-k

Figure 3. Surface (a) charge density and (b) recombination lifetime as a function of thickness during an etchback of a 3.5 nm oxide.
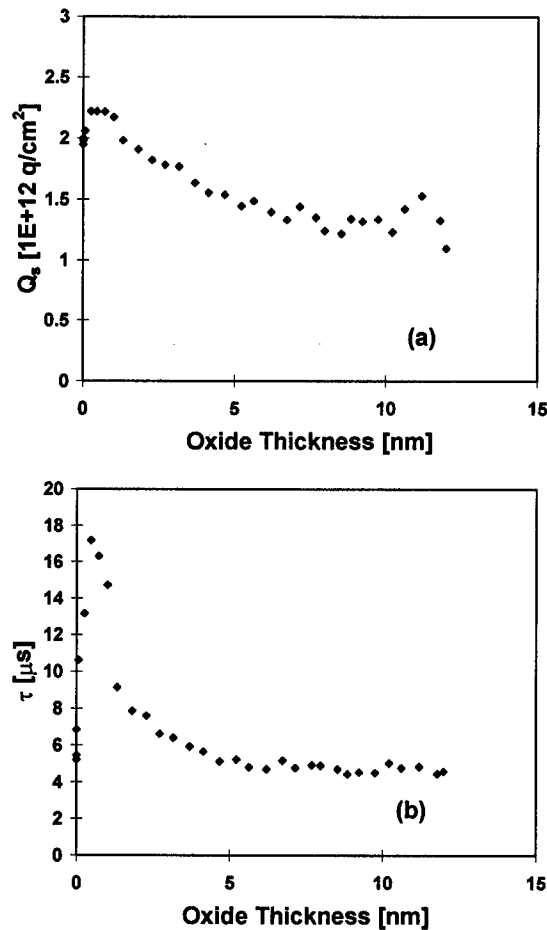

Figure 4. Surface (a) charge density and (b) recombination lifetime as a function of thickness during an etchback of a 12 nm oxide.

dielectrics for both MOS gate as well as storage capacitor applications adequate methodologies for material characterization for process development and monitoring should be available. Due to the properties of these materials, such as leakage currents typically higher than thermal SiO_2 as well as difficulties with selection of the gate contact material a non-contact method that can measure charge density across a range of thicknesses of high-k dielectrics without an electric field stress may play a very useful role. In this section thin layers of Ta_2O_5 on Si are studied using the SCP method.

In this study 40 nm thick films of Ta_2O_5 deposited on silicon by reactive sputtering were studied. It is known that that the reactive sputter deposition method can not produce Ta_2O_5 films featuring adequate electrical characteristics, but this deficiency was what was needed in this experiment. Using the SCP method the oxide charge density was effectively measured as shown in Fig. 5. In this figure variation of charge density in Ta_2O_5 with time of post-deposition anneal at 550 °C in both N_2 and O_2 is shown.

Additional measurements were performed on 4 nm thick films of Ta_2O_5 deposited by CVD on Si surfaces which were either thermally nitrided before deposition or not [7]. Different values of charge density, 0.8 and 3.3×10^{10} cm^{-2}, respectively, (Fig. 6) were detected in spite of the observed substantial conductance of these films.

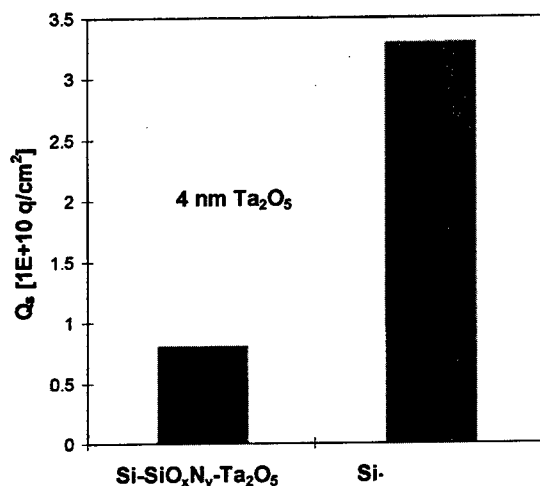## 4. CONCLUSIONS

The results obtained indicate a very good potential of the non-contact SPV based method of Surface Charge Profiling in the electrical characterization of oxides which either due to the negligible thickness and resulting tunneling, or structural imperfections and resulting leakage are too conductive to allow C-V characterization.

REFERENCES

[1] P. Roman, M. Brubaker, J. Staffa, E. Kamieniecki and J. Ruzyllo in "Characterization and Metrology for ULSI Technology", D.G. Seiler, A.C. Diebold, W.M. Bullis, T.J. Shaffner, R. McDonald and E.J. Walters, Editors, AIP Conference Proceedings, 449, Woodbury, NY, 1998, p. 250.
[2] A. Chin, B.C. Lin, W.J. Chen, Y.B. Lin and C. Tsai, IEEE El. Dev.Lett., 19,426 (1998).
[3] S.A. Jang, Y.B. Kim, I.S. Yeo and S.K. Lee, J. Electrochem. Soc., 146, 270 (1999).
[4] SCP, QC Solutions, Inc., Woburn, MA, USA.
[5] E. Kamieniecki, J. Appl. Phys. 54, 6481 (1983).
[6] B. Deal, B.E. Deal, M. Sklar, A.S. Grove and E.H. Snow, J. Electrochem Soc., 114, 266 (1967).
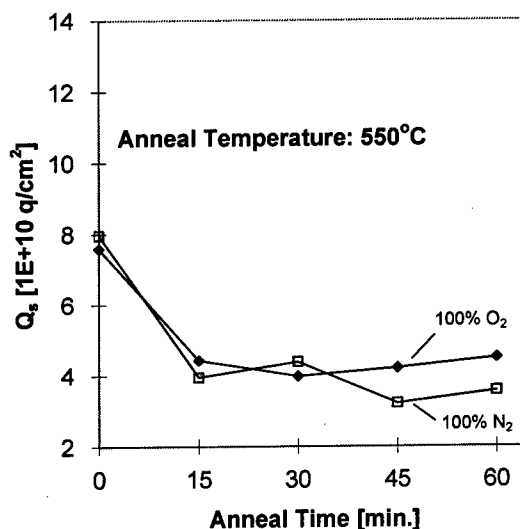[7] Wafers were obtained from Lucent Technologies Bell Labs.

Figure 6. Surface charge measured in 4nm CVD Ta_2O_5 films.

Figure 5. Surface charge as a function of anneal time in 40 nm sputter-deposited Ta_2O_5 films.

RTS capture kinetics and Coulomb blockade energy in submicron nMOSFETs under surface quantization conditions

N.B.Lukyanchikova[a], M.V.Petrichuk[a], N.P.Garbar[a], E.Simoen[b] and C.Claeys[b,c]

[a]Institute of Semiconductor Physics,
45 Prospect Nauki, 252650, Kiev, Ukraine

[b]IMEC, Kapeldreef 75, B-3001 Leuven, Belgium

[c]KU Leuven, ESAT-INSYS, B-3001 Leuven, Belgium

By measuring the RTS capture kinetics in submicron nMOSFETs under conditions of a constant drain current but changed gate and substrate bias in linear operation, it has been found that a high surface electric field E_s is the important factor that controls the RTS capture kinetics and the Coulomb blockade energy ΔE. In fact, it is shown that the capture time constant and the value of ΔE may be presented as unique functions of the value of $N_s E_s$ and $N_s E_s^2$, respectively, where N_s is the inversion surface charge density in the channel. This effect is attributed to the fact that some distance exists between the centroid of the inversion layer and the Si/SiO$_2$ interface under conditions of surface quantization occuring at sufficiently high values of E_s and this distance decreases with increasing E_s.

1. INTRODUCTION

The random telegraph signal (RTS) noise observed in submicron MOSFETs offers a unique way of studying the single-electron trapping process in an individual interface trap [1]. Investigations of this noise in state-of-the-art nMOSFETs characterised by ultrasmall thicknesses of the oxide layer t_{ox} ($\leq$ 10 nm) and high doping concentrations in the channel region N_A ($\geq 10^{17}$ cm^{-3}) seem to be of a special interest. First of all, due to such small values of t_{ox} the high Coulomb blockade energy ΔE influences significantly the RTS kinetics [2-4]. This effect implies that the image charges of the RTS trap charge have to establish on the gate electrode, in the channel and in the substrate. To do this, the additional energy ΔE called Coulomb energy and related to the movement of charge against the potential distribution in the structure is required. Moreover, due to high values of N_A the surface electric fields at the Si/SiO$_2$ interface appear to be so high that the RTS process occurs under conditions of a two-dimensional quantization of the c-band near the interface where the following effects can affect the RTS

kinetics: (i) centroid of the inversion charge layer is a finite distance away from the Si/SiO$_2$ interface and this distance decreases with increasing surface electric field E_s; (ii) the electrons emitted from the interface trap must have an additional energy to be captured in the conduction band and this energy increases with increasing E_s [5].

In this paper the results of investigations of the RTS noise in the nMOSFETs with t_{ox} = 7 nm and $N_A = 2.1 \cdot 10^{17}$ cm^{-3} are presented. In weak inversion this noise is characterised by the superlinear reduction of the capture time constant τ_c with increasing drain current I in linear operation, namely: $\tau_c \sim I^{-3.7}$, while the emission time constant τ_e is practically constant in this case (Fig.1).

Within the framework of the Coulomb blockade theory, such a behaviour of τ_c and τ_e is typical for a repulsive center (which is neutral when empty) for which [2-4]

$$\tau_c = (1/C_n n_s)\exp(\Delta E/kT), \qquad (1)$$

$$\tau_e = (1/C_n N_c)\exp(E_T/kT) \qquad (2)$$

where C_n is the capture coefficient for electrons, n_s is the effective free electron concentration per unit volume at the Si/SiO$_2$ interface, k is the Boltzmann constant, T is the temperature, N_c is the effective density of states in the conduction band of a semiconductor and E_T is the energy distance between the RTS level and the c-band level where the electron can be emitted.
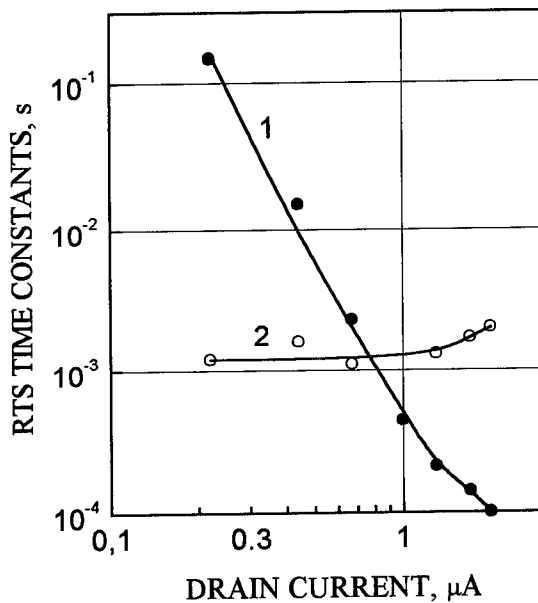

Figure 1. Capture (1) and emission (2) time constant in function of the drain current at $U_{BS} = 0$ V.

The essential property of the Coulomb energy is that ΔE reduces with increasing inversion surface charge density Ns in the channel (where Ns is the number of free carriers per unit area at the interface) due to the increased screening of the gate image charge by the inversion layer [3,4]. The superlinear decrease of τ_c with increasing I is caused by the fact that ΔE decreases with increasing N_s while I increases linearly with increasing N_s. It will be shown that under conditions investigated the behaviour of τ_c and ΔE was determined not only by the value of N_s but also by the surface electric field E_s that was changed from 200 to 366 kV·cm^{-1} giving rise to the surface quantization where the ground state resided (60 to 88) meV above the surface c-band energy level and the centroid of the inversion layer was a distance (2.0 to 1.7) nm away from the interface.

2. EXPERIMENTAL

The RTS noise has been studied in submicron nMOSFET fabricated in a 0.35 μm CMOS technology. The effective channel length and width are 0.48 μm and 0.75 μm, respectively. The noise spectra in the frequency interval f = 1 Hz to 100 kHz as well as the RTS óscilloscope pictures have been measured and analysed in detail in order to determine the values of τ_c and τ_e under different operating conditions. Detailed measurements of τ_c and τ_e in function of the drain current and of the gate (U_{GS}) and substrate (U_{BS}) bias were performed in linear operation at room temperature. Particular attention was given to the experiments where E_s and n_s were changed under conditions of N_s = const. These measurements were carried out by changing U_{BS} and U_{GS} while keeping the value of I unchanged. The gate bias U_{GS} was changed from 0.45 V to 0.85 V; the substrate bias U_{BS} was changed from -1.39 V to +0.30 V and the current interval investigated varied between 0,15 μA to 2.0 μA.

3. RESULTS AND DISCUSSION

The behaviour of τ_c under conditions where N_s = const while U_{BS} and U_{GS} change is shown in Fig.2 where the dependences of τ_c on U_{BS} measured at different drain currents are plotted.
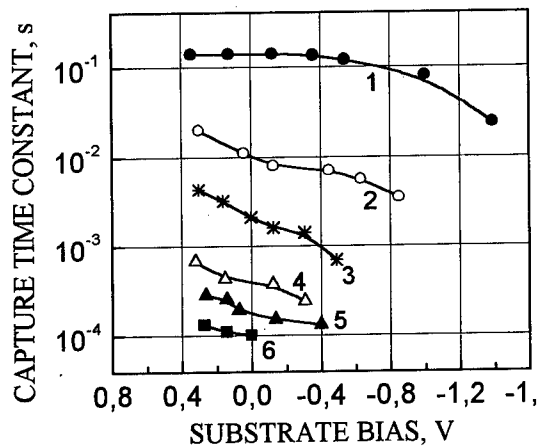

Figure 2. Capture time constant as a function of the substrate bias measured at I = 0.22 μA (1), 0.44 μA (2), 0.67 μA (3), 1.0 μA (4), 1.3 μA (5) and 2.0 μA (6), corresponding to six different constant values of N_s where the conversion factor is $N_s/I = 3.2 \cdot 10^{17}$ cm^{-2}A^{-1}.

As is seen, the value of τ_c is not constant at N_s = const but is a strong function of the substrate bias where τ_c decreases with increasing value of $|-U_{BS}|$. From (1) it follows that one of the reason for such a decrease of τ_c can be the increase of the value of n_s due to the rise of the surface electric field with increasing negative substrate voltage $|-U_{BS}|$ at N_s = const. To check whether or not this increase of n_s is responsible for the decrease of τ_c considered, the values of E_s and n_s were calculated and then the dependences of $n_s\tau_c$ on U_{BS} were analysed. The values of E_s and n_s were determined by solving the set of equations describing the situation in the inversion channel under surface quantization [5]. The dependences of $\tau_c n_s$ on U_{BS} for different values of I are shown in Fig.3. As is seen, the value of $n_s\tau_c$ appears to be also a function of U_{BS}.
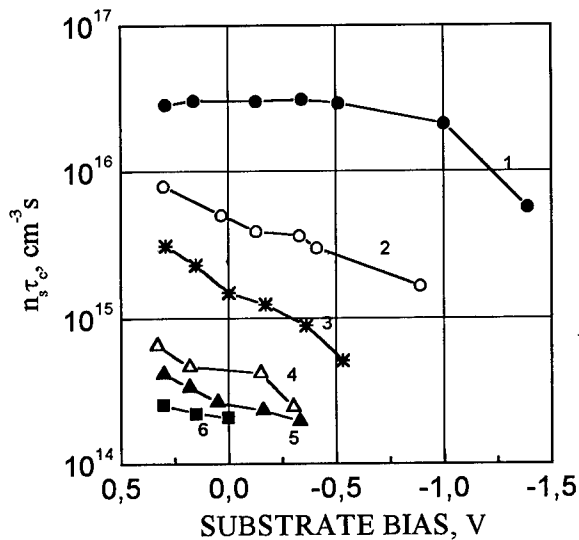

Figure 3. Dependences of $\tau_c n_s$ on U_{BS} determined for data of Fig.2.

This points to the fact that the decrease of τ_c with increasing $|-U_{BS}|$ shown in Fig.2 can not be attributed to the small increase of n_s under conditions investigated. Therefore, along with n_s and N_s, some additional factor controls the RTS capture kinetics investigated. It has been found that such a factor is the value of E_s. This is demonstrated by plotting τ_c versus N_sE_s (Fig.4).

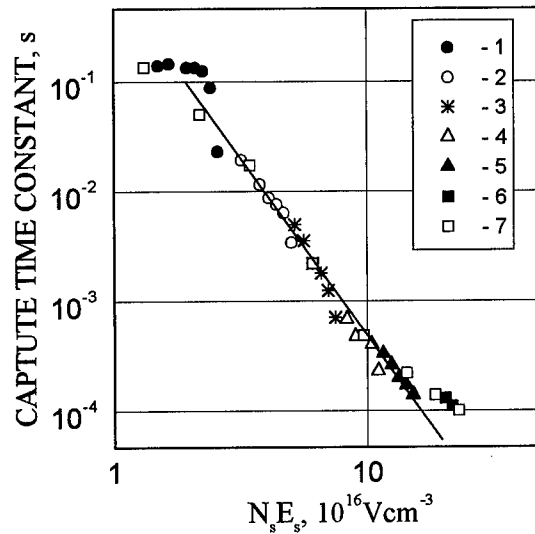As is seen from Fig.4, all the results shown in Figs.1 and 2 fall on a single line.

Figure 4. Capture time constant as a function of N_sE_s; labelling of the curves 1 to 6 is identical to that in Fig.2; curve 7 corresponds to curve 1 in Fig.1.

By attributing the behaviour of τ_c shown in Fig.1 and Fig.2 to some specific behaviour of the Coulomb blockade energy ΔE under conditions considered, the results shown in Figs. 5 to 7 have been obtained where curves 1 to 6 correspond to curves 1 to 6 in Fig.2 measured at I = const while curve 7 corresponds to curve 1 in Fig.1.
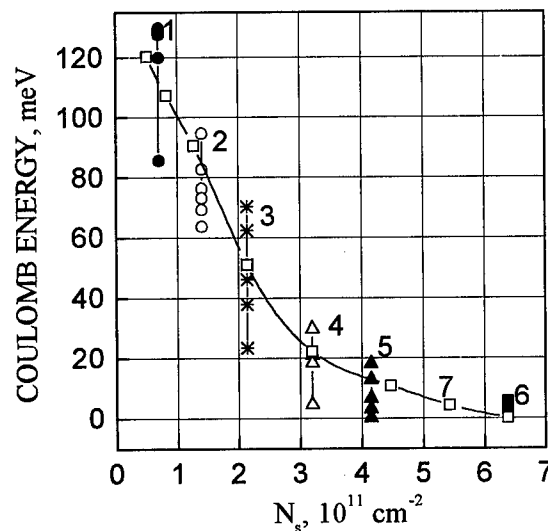

Figure 5. Coulomb blockade energy as a function of N_s; labelling of the curves is identical to that in Fig.4

The value of ΔE was found by means of the formula $\Delta E = kT\ln(\tau_c C_n n_s)$ derived from (1) where the capture coefficient C_n was considered to be independent of I.
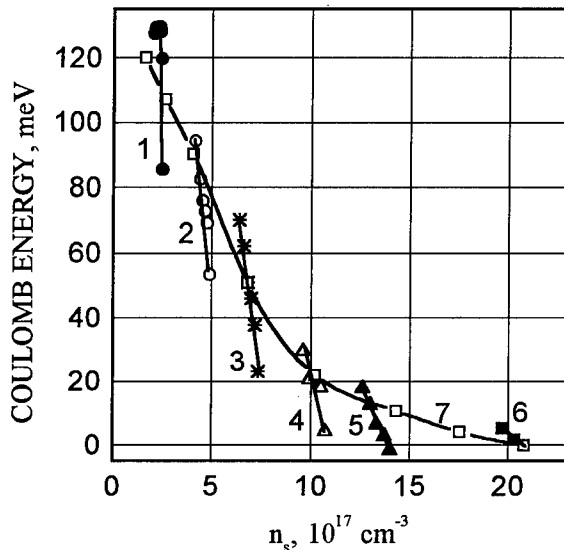

Figure 6. Coulomb blockade energy as a function of n_s; labelling of the curves is identical to that in Fig.4
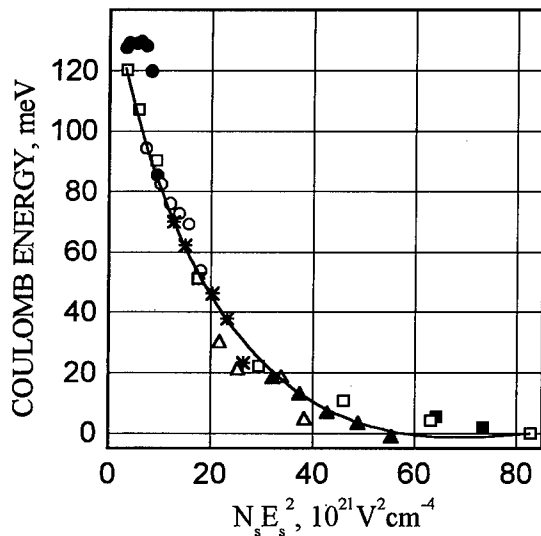

Figure 7. Coulomb blockade energy as a function of $N_s E_s^2$; labelling of the curves is identical to that in Fig.4

The reason is the constant value of τ_e in the region of the rapid decrease of τ_c with increasing I (Fig.1, curve 2) that can be described by (2) with unchanged value of C_n. The value of C_n was found

from $C_n = (\tau_c n_s)^{-1}$ where τ_c and n_s correspond to the highest current investigated where $\tau_c \sim I^{-1}$ is observed (Fig.1, curve 1) and, hence, the simple relation $\tau_c = (C_n n_s)^{-1}$ is valid. The result is $C_n = 4.8 \cdot 10^{-15}$ cm^{3}s^{-1} and $\sigma_n = C_n v_{th}^{-1} = 4.8 \cdot 10^{-22}$ cm^2 where σ_n is the capture cross-section for electrons and v_{th} is the thermal velocity. Such a cross-section is typical for the electron repulsive traps [1].

As is seen from Figs.5 to 7, the value of ΔE is rather high (up to 130 meV) at a relatively small value of N_s in the devices considered. It is also seen that ΔE is not a unique function of either N_s (Fig.5) or n_s (Fig.6). At the same time, Fig.7 shows that ΔE may be presented as a unique function of the value of $N_s E_s^2$.

4. CONCLUSIONS

Not only the inversion surface charge density N_s but also the high surface electric field E_s are factors that control the RTS capture kinetics and, perhaps, the value of the Coulomb blockade energy in submicron nMOSFETs under surface quantization conditions. This fact can be explained intuitively considering the distance existing between the centroid of the inversion layer and the Si/SiO$_2$ interface under conditions of surface quantization. This distance decreases with increasing E_s, explaining the observed trends.

REFERENCES

1. M.J.Kirton and M.J.Uren, Adv. Phys., 38 (1989) 367.

2. M.Schulz, J.Appl.Phys., 74 (1993) 2649.

3. H.H.Mueller, D.Wörle and M.Schulz, J.Appl. Phys., 75 (1994) 2970.

4. M.Schulz and H.H.Mueller, in: Festkörperprobleme/Advances in Solid-State Physics, 35, ed. by R.Helbig (Vieweg, Braunschweig/Wiesbaden), pp.229-241.

5. R.R.Siergiej, M.H.White and N.S.Saks, Solid-St.Electron., 35 (1992) 843.

On the origin of flicker noise in MOSFET's

Y. Maneglia and D. Bauza,

Laboratoire de Physique des Composants à Semiconducteurs,
ENSERG/INPG, UMR CNRS 5531,BP 257, 38016 Grenoble France.

Results extracted using noise spectroscopy, i.e. trap density and slope of noise spectra, are compared, as a function of Fowler-Nordheim stress, with those obtained using a technique called tunnel charge pumping (TCP) which allows the extraction of the Si-SiO$_2$ interface trap depth concentration profiles. The trap densities extracted using the two techniques agree for the virgin device, but increase much faster with stress for noise spectroscopy and seem excessively high at high injection doses for this technique. However, a clear correlation between the evolution of the slope of noise spectra and that expected from the trap parameters measured using TCP is evidenced. This emphasizes that flicker noise originates from carrier number fluctuations du to trapping into the oxide traps in the p-channel device studied.

1. INTRODUCTION

Since the late sixties, the origin of flicker noise in MOSFET's has been extensively discussed. Nevertheless, in recent papers, for some authors and from the measurement of large area MOS transistors, flicker noise originates from carrier mobility fluctuations in p-channel MOS transistors and from carrier number fluctuations du to trapping into the oxide traps in n-channel devices [1]. For others, it results in both device types from carrier number fluctuations [2]. Criteria for determining the origin of flicker noise have been defined [3,4] and for MOS devices with conventional architectures, it is found to result mainly from carrier number fluctuation in both device types [3-5]. From the measurement of devices with various areas, noise spectra have been very convincingly decomposed into their individual trapping events [6]. However, assuming that noise results from carrier trapping, deviations from the expected Shockley-Read-Hall behaviour have been observed in micron sized devices [5,7], leading some author to doubt this interpretation [7].

Concerning large area MOS transistors and the carrier number fluctuations approach [8,9], theoretical calculations derived from the McWhorter model [10,11] after introducing various energy and depth dependencies of the trap properties have been proposed. The bias dependence of the noise power spectral density and of the slope γ of noise spectra were studied [10,11]. The main problem concerning this approach was the lack of information about the energy and depth evolution of the trap properties in the near Si-SiO$_2$ interface region (for instance [10-12]).

A technique based on charge pumping measurements and on the tunnelling of carriers into the oxide traps has been proposed recently [13,14]. It allows the extraction of the trap depth concentration profiles, from the Si-SiO$_2$ interface up to around ten angstroms in the oxide. Trap profiles recorded from numerous devices are found of the form $N_t(x) = N_{ts} \exp(-x/d) + N_{to}$ [14]. The trap capture cross sections at the interface can also be extracted [13].

In this paper, the evolution, with Fowler-Nordheim (FN) injection, of the results obtained from noise measurements (trap concentration obtained using the McWhorter model and slope γ of the 1/f$^{\gamma}$ noise spectra), are compared with those extracted using TCP.

2. EVOLUTION OF THE TRAP PROFILES WITH FOWLER-NORDHEIM STRESS

The device studied is a p-channel transistor. The injection has been carried out from the gate.

For the virgin device, the trap depth profile is of the form (dashed line in Fig. 1):

$$N_t(x) = N_{ts} \exp(-x/d) + N_{to} \qquad (1).$$

In state-of-the-art MOS transitors, N_{ts}/N_{to} is of the order of 100 while d lies around 1 Å. The trap capture cross sections at the interface, $\sigma(0)$, of the order of 10^{-15} to 10^{-16} cm^2, agree with the values reported in the literature.
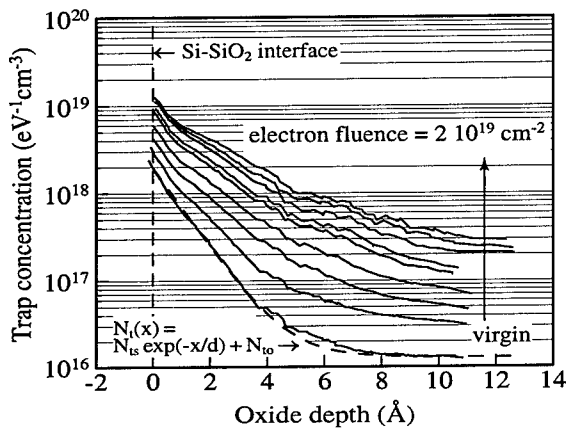

Fig. 1: Evolution of the Si-SiO₂ interface trap depth profile with FN injection.

As the stress proceeds and apart from the overall increase of the trap density, the trap profile always follows Eqn.1 but N_{ts}/N_{to}, the ratio of the trap density at the interface with respect to that in the oxide, decreases, d, the decay characteristic length of the exponential part of the profile increases (Fig. 1) and the trap capture cross section strongly increases. More details about these points can be found in [15].

3. TRAP DENSITIES OBTAINED USING THE TWO METHODS

In Fig. 2 the trap density, N_{to}, measured in the plateau region of the trap profiles in Fig. 1 is compared with that, $N_t(E_f)$, extracted from noise measurements using the McWhorter model [8,9]:

$$N_t(E_f) = S_{vg} \frac{A C_{ox}^2 f}{q^2 kT\lambda} \qquad (2),$$

owning to the fact that in weak inversion, the measured gate voltage spectral density, S_{vg}, is equal

to the flat band voltage spectral density, S_{vfb}, which can be expressed as a function of the parameters of the device [16].

In Eqn. 1, E_f is the Fermi level, A is the device area, C_{ox} is the gate oxide capacitance, f is the frequency, q is the absolute electron charge, k is the Boltzman constant, λ is the tunnel distance ($\lambda \approx 1$ Å).
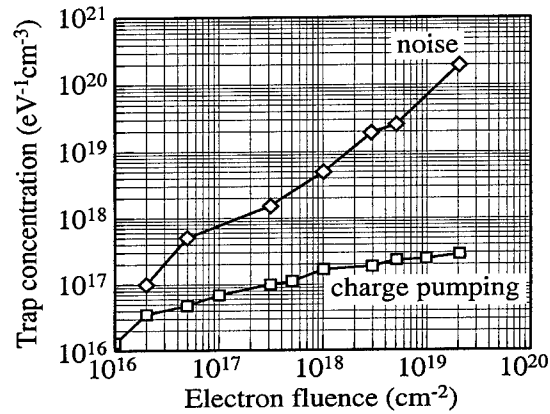

Fig. 2: Comparison between the trap densities extracted from TCP and from noise spectroscopy.

For the virgin device the agreement between the two techniques is good. However, the trap concentration measured using noise spectroscopy increases much faster than that obtained using TCP when increasing the electron fluence.

At high electron fluence, the N_t values extracted from noise measurements seem unrealistic [18]: starting from $N_t = 2 \ 10^{20}$ eV^{-1}.cm^{-3}, as measured at 2.10^{19} q.cm^{-2} in Fig. 2, and assuming that the traps in the first 15 Å in the oxide mainly contribute to noise spectra [9] or TCP, an overall interface trap density $N_{ss} = 3 \ 10^{13}$ eV^{-1}.cm^{-2}, much larger than $N_{ss} = 5.6 \ 10^{11}$ eV^{-1}.cm^{-2} measured on the device, should be obtained; or, to get $N_{ss} = 5.6 \ 10^{11}$ eV^{-1}.cm^{-2} with $N_t = 2 \ 10^{20}$ eV^{-1}.cm^{-3}, it should be assumed that the observed noise originates from a layer 0.3 Å thick.

3. SLOPE γ OF THE 1/f$^\gamma$ NOISE SPECTRA

Fig. 3 shows noise spectra calculated in the case of trap concentrations given by Eqn. 1 using the trap parameters, N_{ts}, N_{to}, d and $\sigma(0)$ measured at low (Fig. 3a) and high (Fig. 3b) fluence, for three surface potential, ϕ_s, values, between weak and the onset of

strong inversion. The exponential part of the profiles induces a hump on the spectra before the occurrence of the $1/f^2$ dependence when the traps at the interface have not the time to respond [9]. The frequency range, 1-200 Hz, where measurements are carried out is noted in the figures.

Fig. 3: Simulation of noise spectra using the trap parameters extracted from TCP after a (a) low and (b) high electron fluence.

At low fluence (Fig. 3a), $\gamma = 1$ in the 1-200 Hz region for a large range of ϕ_s values (ϕ_f is the Fermi potential). It diminishes only at very weak inversion. At high fluence (Fig. 3b), $\gamma < 1$ in a large part of the weak inversion region except near $\phi_s = \phi_f$ where the nearness of the $1/f^2$ region increases the slope.

Figs 4a and 4b show $N_t(f)$ curves directly extracted from the measured noise spectra using the McWhorter model (Eqn. 2). They correspond to

spectra recorded from weak to the onset of strong inversion. In Fig. 4a and 4b, the electron fluencies are those used in Fig. 3a and Fig. 3b, respectively. An N_t value independent of frequency reveals an 1/f noise spectrum; an N_t value increasing with f corresponds to an $1/f^\gamma$ ($\gamma < 1$) noise spectrum.

At low fluence (Fig. 4a) and in agreement with Fig. 3a, the slope of noise spectra is near unity except in very weak inversion while at high fluence (Fig. 4b) it is strongly bias dependent and increases, as expected, toward weak inversion.

This is summarized in Fig. 5, where the values of γ obtained from curves such as those in Fig. 3, after simulating noise spectra using the parameters of the trap profiles extracted from TCP (Fig. 5a) are compared, as a function of the injection dose, with those recorded from noise spectra (Fig. 5b).
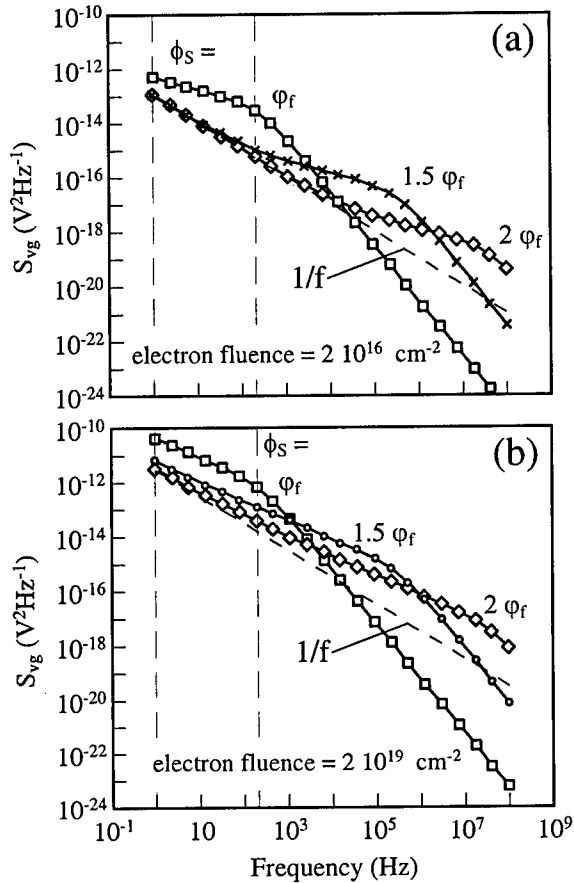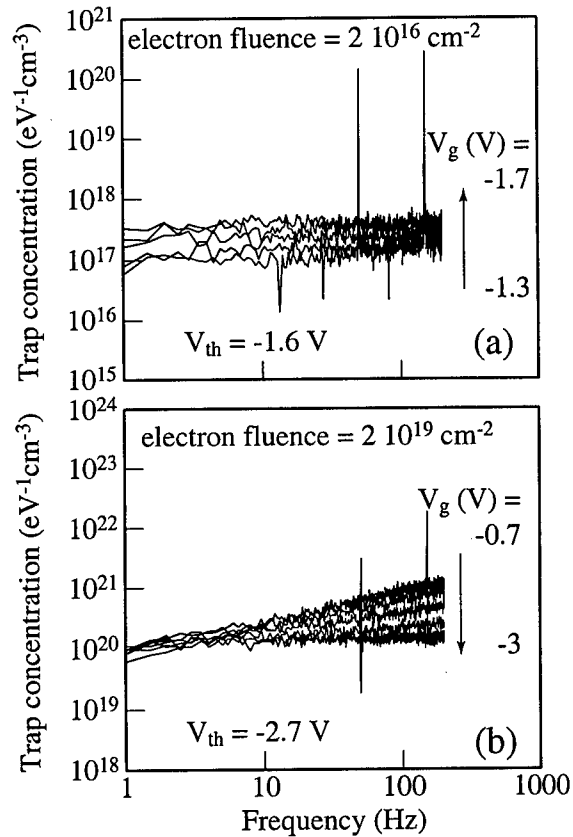

Fig. 4: Trap concentration directly calculated from noise spectra using the McWhorter model (Eqn. 2) at (a) low and (b) high electron fluence.
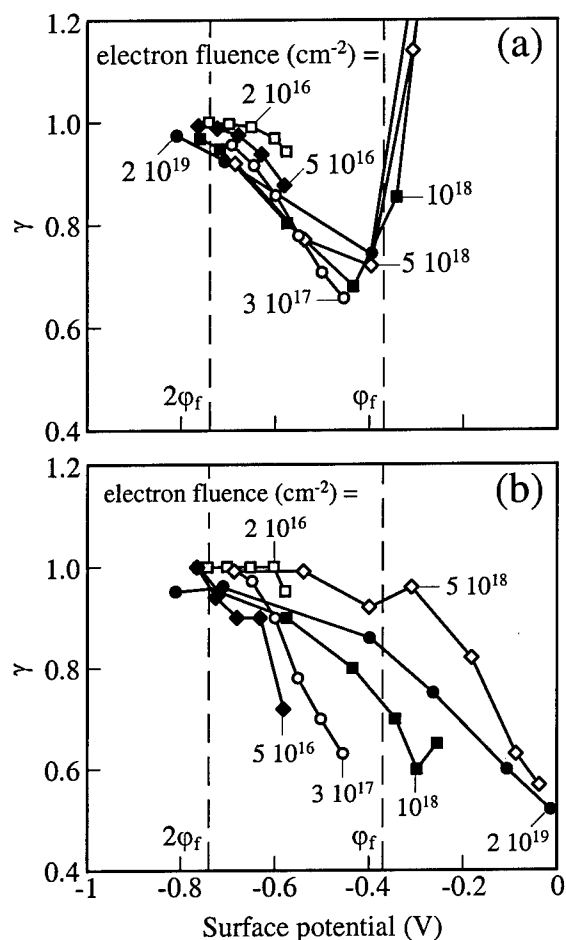
Fig. 5: Slope of noise spectra (a) obtained from simulation after introducing the trap parameters measured using TCP and (b) measured.

A clear correlation can be seen between Figs. 5a and 5b, especially at low fluencies (up to 10^{18} q.cm^{-2}). At high fluencies, the experimental curves are stretched in regard to those simulated. This mainly results from the simulation of $\phi_s(V_g)$, at high trap density values, which is probably far from that of the device. In particular, the abrupt increase of γ near $\phi_s = \varphi_f$ in Fig. 5a, which corresponds to the occurrence of the $1/f^2$ noise regions in Fig. 3 is not observed in Fig. 5b.

Figs. 4 and 5 show that γ follows the evolution of the trap parameters measured using TCP. As a result, flicker noise in the p-channel device studied results from carrier number fluctuations due to trapping into the near Si-SiO$_2$ interface traps.

4. CONCLUSION

It has been shown that the oxide trap densities measured using TCP and noise spectroscopy, which agree for virgin devices strongly differ at high FN stress, those extracted from noise measurements being extremely high. However, a clear correlation has been evidenced between the trap parameters extracted from TCP and the slope of noise spectra, showing that flicker noise originates from carrier number fluctuations in the p-channel device studied.

REFERENCES

1. J. Chang and A.A. Abidi and C. R. Viswanathan, IEEE Trans. Electron Devices, ED-41 (1994) 1965.
2. H. E. Maes, S. H. Usmani and G. Groeseneken, J. Appl. Phys., 57 (1985) 4811.
3. G. Reimbold, IEEE Trans. Electron Devices ED-31 (1984) 1190.
4. T. Boutchacha and G. Ghibaudo, J. Non-Cristal. Solids, 216 (1997) 192.
5. M. J. Kirton and M. J. Uren, Adv. Phys., 38 (1989) 367.
6. M. J. Uren, D. J. Day and M. J. Kirton, Appl. Phys. Lett. 47 (1985) 1195.
7. K.R. Farmer and R. A. Buhrman, Semicon. Sci. Technol., 4 (1989) 1084.
8. A. L. McWhorter, Semiconductor Surface Physics, Pennsylvania University Press, Philadelphia, 1957, p. 207.
9. S. Christensson, I. Lundström and C. Svensson, Solid-St. Electron., 11 (1968) 797.
10. H. E. Maes and S. H. Usmani, J. Appl. Phys., 54 (1983) 1937.
11. Z. Celik-Butler and T. Y. Hsiang, Solid-St. Electron, 30 (1987) 419.
12. A. Nemirovsky and A. Ron, Solid-St. Electron, 41 (1997) 1811.
13. D. Bauza and Y. Maneglia, IEEE Trans. Electron Devices, 44 (1997) 2262.
14. Y. Maneglia and D. Bauza, International Conference on Microelectronic Test Structures, Kanazawa, Japan, 1998. Proc. IEEE 98CH36157.
15. D. Bauza and Y. Maneglia, Electrochemical Society Spring Meeting, Seattle, USA, 1999. Proc. to be published.
16. G. Ghibaudo, Solid-St. Electron, 32 (1989) 563.
18. D. Vuillaume, IEEE Trans. Electron Devices, 37 (1990) 824.

A study of 100 nm channel length asymmetric channel MOSFET by using charge pumping

S. Mahapatra[a]*, V. Ramgopal Rao[a], C. D. Parikh[a], J. Vasi[a], B. Cheng[b], and J. C. S. Woo[c]

[a]Department of Electrical Engineering, Indian Institute of Technology, Bombay-400076, India.

[b]Advanced Products Research & Development Lab, Motorola, 3501 Ed Bluestein, Austin, TX 78721, USA

[c]Department of Electrical Engineering, University of California, Los Angeles, CA 90095-1594, USA.

Lateral Asymmetric Channel (LAC) MOSFETs with channel lengths down to 0.1 μm have been fabricated and characterized for their electrical performance. Using charge pumping, we show, for the first time, channel V_T profiles obtained experimentally, demonstrating realization of asymmetric channel MOSFETs down to 0.1 μm channel lengths. Our detailed experimental characterizations show improved performance for LAC MOSFETs over conventional MOSFETs, in addition to excellent hot-carrier reliability. Based on 2-D device simulation results, we attribute the improved hot-carrier reliability in LAC MOSFETs to the reduced peak lateral electric field in the channel.

1. INTRODUCTION

Hot-carrier degradation (HCD) involving carrier heating, injection and subsequent gate-oxide damage can become a severe reliability concern when realizing deep sub-micrometer MOSFETs [1]. With the advent of high-k dielectrics for use in sub 0.1 μm channel MOSFETs, and the invariably lower barrier heights that would result by the use of these dielectrics [2], HCD is going to become a major concern. One attractive way of alleviating hot-carrier problems is by adopting suitable channel engineering techniques in the sub 0.1 μm channel regime [3]. In this paper, we present experimental results on Lateral Asymmetric Channel (LAC) MOSFETs having channel lengths down to 0.1 μm. Using charge pumping in LAC and conventional MOSFETs, fabricated on the same wafer, we show that asymmetric channel doping improves hot-carrier reliability by more than a factor of two in comparison to conventional MOSFETs. 2-D device simulations show that reduced peak lateral electric field near the drain

junction in LAC MOSFETs is responsible for the observed improvement in reliability.

2. DEVICE FABRICATION

Both the conventional (CON) and LAC n-MOSFETs were fabricated on the same wafer for realistic comparisons [3]. E-beam lithography was used to define gate lengths down to 0.1 μm (effective channel lengths down to 0.08 μm). Conventional SiO_2 having an electrical oxide thickness (EOT) of 3.9 nm is used as the gate dielectric. A large angle tilt implant was employed for V_T adjustment in LAC MOSFETs after the poly gate formation. The V_T adjustment in CON devices was done before the gate oxidation. A two-step low-temperature Ti silicidation process with Ge preamorphization was implemented to control the silicide depth and to reduce series resistance. The schematic of the LAC MOSFET structure its and doping profile is shown in Fig. 1.

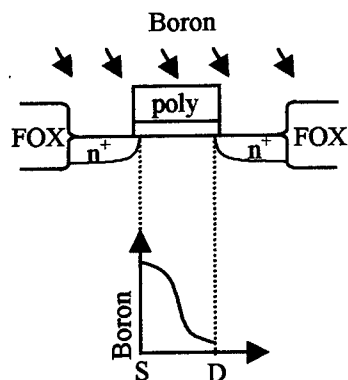
* Recipient of fellowship from Siemens AG, Germany.

Figure 1. Schematic of a LAC n-MOSFET along with the channel doping profile.

3. RESULTS AND DISCUSSION

Fig 2 shows the output characteristics of a 0.1 μm LAC and conventional MOSFET. The higher drain current and improved transconductance saturation in the LAC MOSFET can be attributed to the early electron velocity overshoot at the source end of the channel [3]. Fig 3 shows the DIBL for conventional and LAC MOSFET as a function of channel length. DIBL in the forward mode of operation is less for LAC MOSFETs compared to conventional MOSFETs of identical channel length.
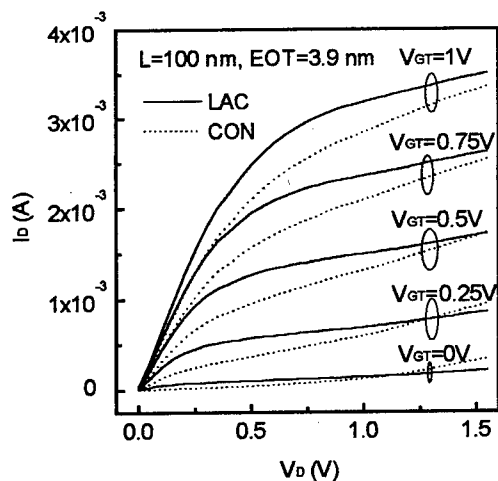
Figure 2.Output characteristics for a 0.1 μm LAC and conventional MOSFET.

Charge pumping (CP) measurements were performed in LAC and conventional MOSFET using the setup shown in Fig. 4. The gate of the MOSFET was pulsed using a 1 MHz trapezoidal waveform. The substrate was shorted to ground, while the dc recombination current (I_{cp}) was measured at the source and drain [4].
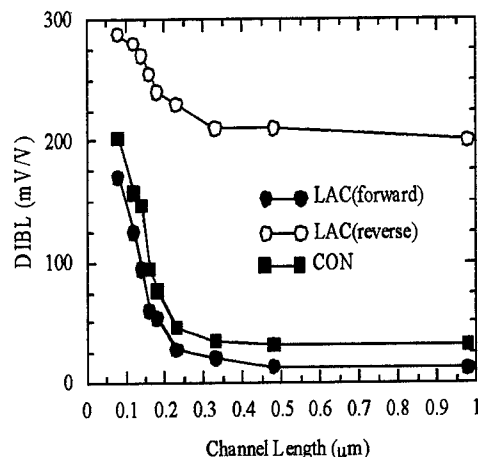

Figure 3. DIBL as a function of channel length for conventional and LAC MOSFET.
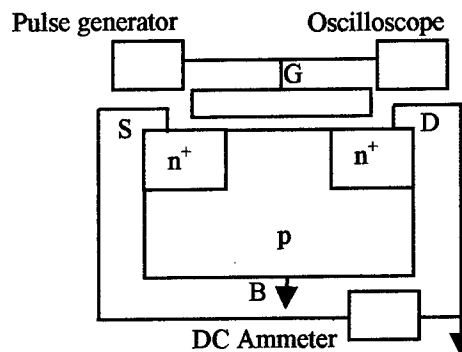

Figure 4. Charge pumping setup.

Fig 5 shows the pre- and post-stress CP current characteristics as a function of pulse top level (V_{top}) for a 0.1 μm LAC and conventional MOSFET. The identical pre-stress I_{cp} in LAC and conventional devices indicates no additional interface-state density (N_{it}) generation due to V_T implant through gate oxide in the LAC MOSFETs. The relatively smaller increase in post-stress I_{cp} in

LAC MOSFETs indicates lesser post-stress N_{it} generation in such devices, as shown later.
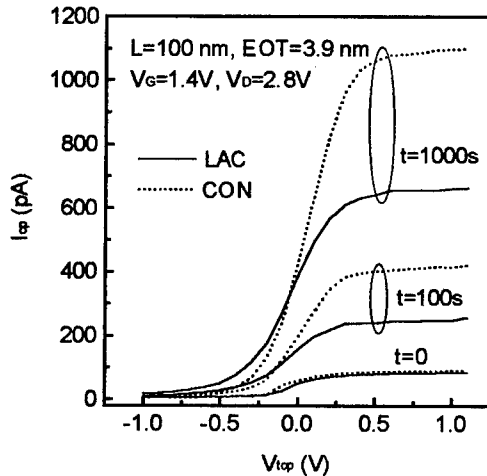

Figure 5. Pre- and post-stress charge pumping currents as a function of pulse top level for conventional and LAC MOSFETs.

The slope of the fitted straight line through the pre-stress I_{cp} versus drawn gate length (L) plot furnishes average pre-stress N_{it}. This along with I_{cp} from open drain and open source measurements provides the threshold voltage (V_T) profile [4], as shown in Fig 6. The asymmetric nature of the LAC devices can be clearly seen even down to 0.1 µm channel length.
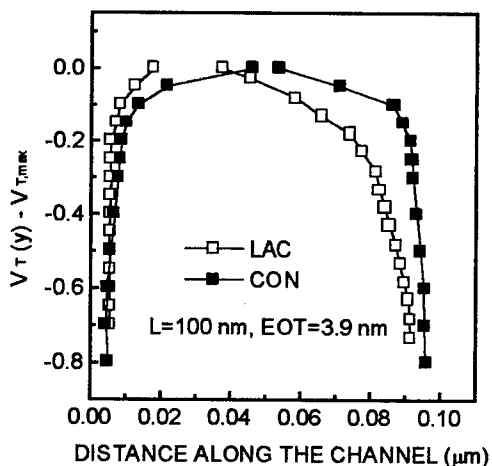

Figure 6. Pre-stress threshold voltage profiles along the channel for conventional and LAC MOSFETs obtained using charge pumping.

Fig 7 shows the stress-induced N_{it} distribution along the channel for a 0.1 µm conventional and LAC MOSFETs. Stressing was done at $V_G = V_D/2 = 1.4V$ for 100s. The N_{it} profiles were obtained by
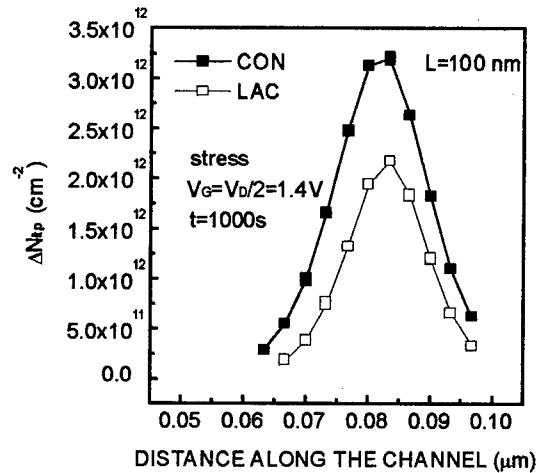

Figure 7. Stress-induced interface-state density profile for conventional and LAC MOSFETs.

differentiating the measured ΔI_{cp}-V_{top} data (ΔI_{cp} is the incremental CP current) and using the pre-stress V_T distribution [4]. The reduced peak and spread of the stress induced N_{it} profiles in the LAC MOSFET compared to the conventional device can be clearly seen.

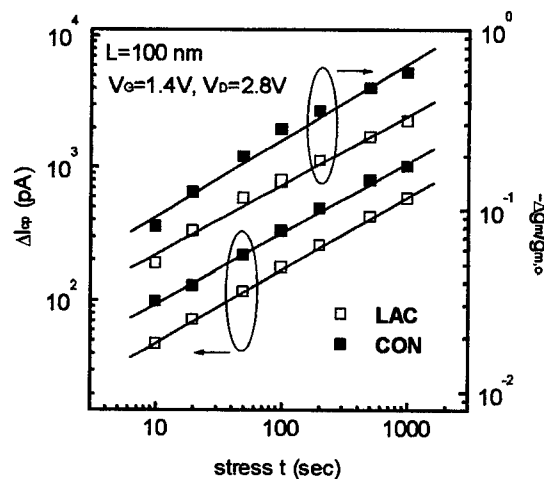Figs. 8 through 10 show the incremental CP

Figure 8. Incremental CP current and transconductance degradation as a function of stress time.

current (ΔI_{cp}) and transconductance degradation ($\Delta g_m/g_{m,o}$) for LAC and conventional devices as a function of stress time, supply voltage and channel length respectively. It is clear that ΔI_{cp} and Δg_m are
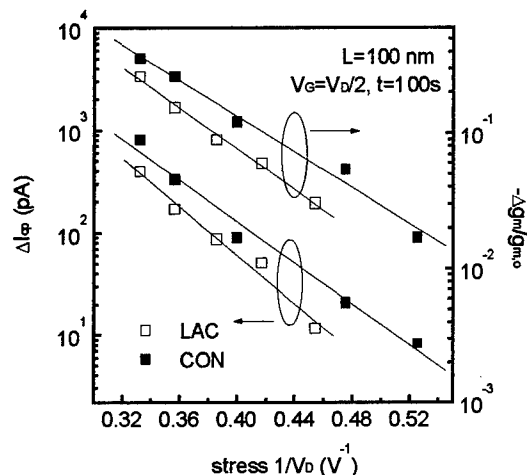

Figure 9. Incremental CP current and transconductance degradation as a function of stress drain bias.

higher in conventional devices compared to the LAC ones for all stress times, supply voltages and channel lengths down to 0.1 μm. Note that both ΔI_{cp} and Δg_m vary as t^n with stress time (Fig 8) and exp $(-a/V_D)$ with stress voltage (Fig 9), with similar values of n (0.54 and 0.43 for ΔI_{cp} and Δg_m) and a (25V and 16V for ΔI_{cp} and Δg_m) for the conventional and LAC MOSFETs.

Fig 11 shows the simulated lateral electric field along the channel for conventional and LAC MOSFETs. The improvement in hot carrier performance in LAC MOSFETs can be directly attributed to the reduced peak lateral field near the drain junction in such devices.

4. CONCLUSION

The asymmetric nature of the channel and its hot-carrier performance is investigated in LAC MOSFETs for channel lengths down to 100 nm. The charge pumping results show excellent hot-carrier immunity of LAC MOSFETs compared to the conventional devices fabricated on the same wafer. The improved hot-carrier reliability in LAC

MOSFETs is attributed to lower peak lateral electric fields near the drain end as seen from 2-D device simulations.
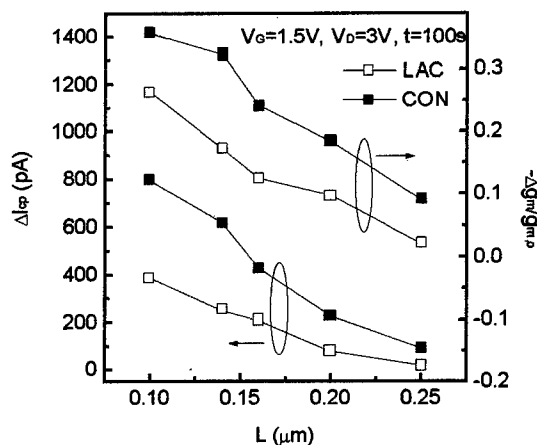

Figure 10. Incremental CP current and transconductance degradation as a function of channel length.
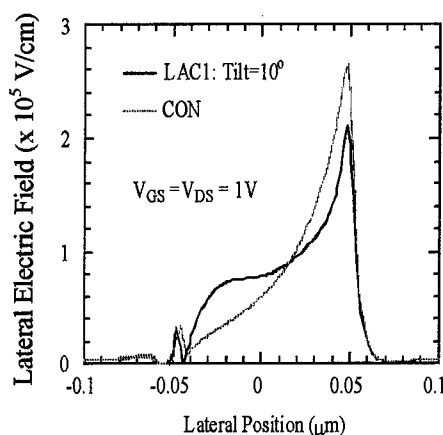

Figure 11. Simulated lateral electric field profiles for LAC and conventional MOSFET.

REFERENCES

1. G. Groeseneken et. al., Semicon. Sci. Technol., 10 (1995) 1208.
2. T. P. Ma, IEEE Trans. on Elect Devices, 45 (1998) 680.
3. B. Cheng et al., IEEE SOI Conference, 1998.
4. C. Chen et al., VLSI Tech. Symp, (1996) 230.

TECHNOLOGY

Advanced cleaning for the growth of ultrathin gate oxide

P.W. Mertens, T. Bearda, M. Houssa, L.M. Loewenstein, I. Cornelissen, S. De Gendt, K. Kenis, I. Teerlinck, R. Vos, M. Meuris, and M.M. Heyns [a] *

[a]Imec, Kapeldreef 75, B-3001 Leuven, Belgium

1. INTRODUCTION

For decades cleaning processes in IC manufacturing were very poorly characterized and specified. Cleaning recipes, such the RCA-recipe [1], have been taken for granted. In order to meet the stringent future gate-oxide defect density requirements [2], cleaning strategies with a higher performance have to be developed. Today wet cleaning takes up a significant fraction of the production cost and generates large amounts of waste. Improving the pre-gate cleaning intrinsic quality and at the same time reduce the consumption of chemicals and water urges us to drastically review the cleaning processes. The development of a better understanding of the basic mechanisms involved is believed to be mandatory.

In this overview, at first the effect of the chemical state of the surface prior to gate oxidation is studied. Then the detrimental effects of metallic contamination will be highlighted. This provides a basis for the specifications for new cleaning recipes. Rinsing is re-examined. It will be shown that the final rinsing has a definite impact on the residual metallic contamination. Finally a new process for the in-situ use of organic precursors for Cl in dry cleaning during gate oxidation is proposed.

2. EFFECT OF SURFACE PASSIVATION

It is commonly assumed that in order to obtain good thickness control for ultra thin oxides, the initial silicon surface should not have any chemical oxide but should be hydrogen-passivated as obtained after a proper HF-immersion and rinse. Experiments [3] have shown however, that the thermal oxide grown on chemical oxide passivated silicon (100) surfaces is less than 0.1 nm thicker than on HF-last treated surfaces, as shown in Fig. 1. In fact the oxide passivated surface is preferred as it results in slightly lower oxide thickness non-uniformity. In addition the oxide passivated surface is thermodynamically more stable and less vulnerable to particle contamination.
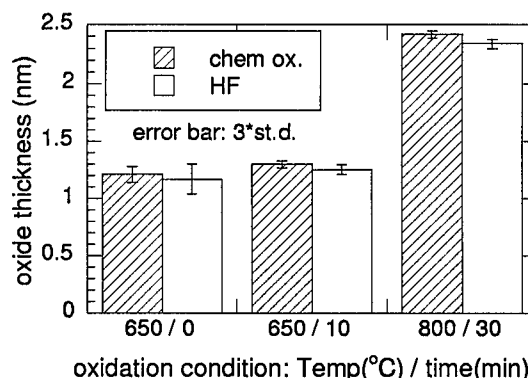

Figure 1. Ellipsometric oxide thickness and 3-σ non-uniformity (= error bars) for three different oxidation conditions (10% O_2 and 90% N_2) in the case of chemical-oxide passivation and of hydrogen passivation of the initial silicon surface.

3. EFFECT OF CONTAMINATION

3.1. Redistribution

During thermal oxidation (or treatment) the initial surface contamination will be redistributed
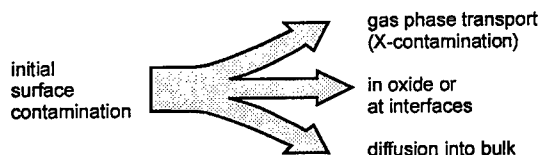
Figure 2. Schematic representation of the redistribution of metallic contamination that is initially present on the wafer surface upon thermal oxidation [4,5].

in three different parts, as indicated in Fig. 2 [4,5]. The surface contamination can either diffuse into the bulk silicon, be incorporated into the oxide or be transported through the gas phase and eventually cause cross-contamination. Each of these three components of material transport have been studied with an appropriate method. The concentration in the oxide and at the interfaces can be measured with vapour phase decomposition - droplet collection (VPD-DC) [6] in combination with either total-reflectance X-ray fluorescence (TXRF) or AAS. The bulk concentration, in the case of Fe has been measured with the SPV method. The "longitudinal out-diffusion" of a box-shaped longitudinal contamination profile in the furnace (i.e. the cross-contamination) gives an indication of the gas phase transport [7].

In this section the redistribution of metallic contamination in the absence of Cl is treated. Fe is not volatilized during typical oxidation process. The redistribution of Fe upon oxidation is relatively complex as it changes significantly with changes in the oxidation process [4,8]. The more oxygen is present in the furnace ambient in an early stage of the heating cycle the smaller the fraction of Fe diffuses into the silicon. The Fe that diffuses in the wafer results in enhanced carrier recombination. More recent tests have been done with (10^{12} at./cm^2) metallic contamination applied prior to the growth of 4.5 nm gate oxide at 800°C in a dry ambient (Fig. 3) [9]. Only Pb, Na and Zn showed significant volatilization. In general metals with a strong tendency for oxide or silicate formation tend to be incorporated in the oxide. Metals with sufficiently high diffusion constant and solubility tend to diffuse in to the

silicon bulk. In general the metals that diffuse more into the bulk yielded the largest carrier lifetime degradation.
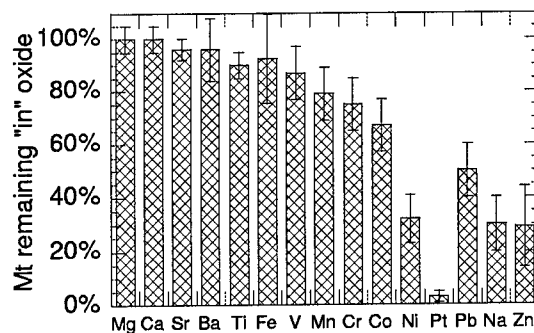

Figure 3. Fraction of the total metal contamination (approximately 10^{12} at./cm^2) that is incorporated into the oxide 4.5 nm oxide grown at 800°C in dry ambient or at one of its two interfaces [9].

3.2. Effect on poly-silicon

Metallic contaminants have also been reported to affect the growth of poly silicon [10]. In case of Fe contamination prior to gate oxidation, an increase of thickness and surface roughness of the poly silicon gate layer has been observed [4,8]. In more recent tests [9] with (10^{12} at./cm^2) metallic contamination applied prior to the growth of gate oxide, mentioned earlier, the light scattering haze on 450 nm polysilicon layers deposited on top of the oxides was measured using a Censor ANS100, as shown in Fig. 4. From all the elements under study only Cr contamination had a major effect on the poly-Si haze.

3.3. Effect on oxide integrity

In general, metallic contamination is expected to degrade device performance. Over the past years, it has been demonstrated that metallic contamination on the wafer surface prior to gate oxidation has a distinct negative effect on the dielectric integrity of thin gate oxides [11,12]. Particularly Ca has been identified as one of the most detrimental metals in that respect [11,12]. Also Fe contamination present on the wafer sur-
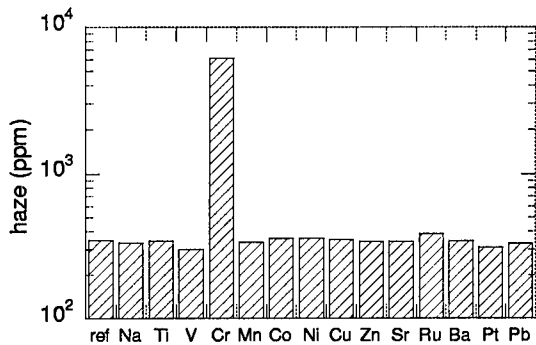
Figure 4. Light scattering haze (Censor ANS100) measured on 450 nm LPCVD poly-silicon layers deposited on 4.5 nm oxide grown at 800°C in dry ambient on wafers that were metal contaminated with approximately 10^{12} at./cm^2 [9].

face prior to the gate oxidation has been correlated with gate oxide integrity (GOI) degradation [8,12,13]. GOI degradation obtained in case of Fe contamination has been found to depend on the silicon crystal growth process as well [13]. A recent evaluation, shown in Fig. 5, shows that the column II metals are most detrimental for gate oxide dielectric strength, followed by V and Co [9]. Metal contamination can also introduce fur-

Figure 5. E_{bd} defect density of 4.5 nm gate oxides at 11 MV/cm obtained from ramped voltage stress for different conditions of intentional metal contamination [9].

ther degradation during post oxidation anneals

[12,14]. Metals are reported to act as catalyst for the SiO$_2$ decomposition reaction during heat treatment in an ambient with a very low oxygen concentration [15]. Another mechanism for degradation could be the formation of stable silicates [14].

4. IMEC CLEAN AND FINAL RINSING

4.1. Imec clean

The RCA-recipe [1] has been commonly used to clean wafers. Today wet cleaning takes up a significant fraction of the production cost and generates large amounts of waste. A strong reduction in chemical and UP-water consumption can be obtained by using simplified cleaning schemes, such as the imec clean^{TM} [16,17]. A typical implementation is presented in table 1.

Table 1
Overview of different steps of a typical implementation of the imec clean^{TM} in an immersion tank system [16,17].

description	temp. (°C)	time (min)
H$_2$SO$_4$/O$_3$	90	5
3-fold quick-dump rinse	60, 20	8
dHF (0.5%) + dHCl (0.5M)	22	2
rinse, reduced pH (+ O$_3$)	20	10
Marangoni dry, reduced pH	20	8

In the first step the organic contamination is oxidized. In the same step a well controlled thin chemical oxides is obtained on the silicon surface by a self-limiting growth. From an environmental perspective it is desirable to perform this oxidation step using ozone - water mixtures [18], without sulphuric acid. In the latter case further savings in rinsing can be realized.

The HF/HCl-mixture etches off the chemical oxide very selectively, thereby under cutting particles. This optimized mixture provides protection against noble metal outplating from the solution [19].
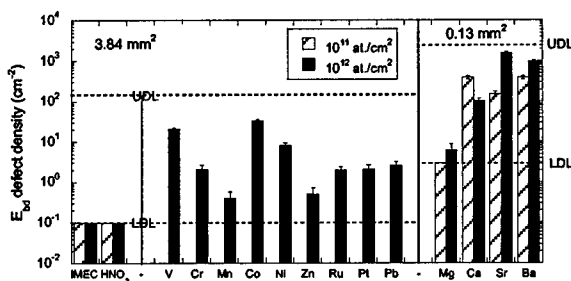
Then rinsing is performed at reduced *pH* in order to avoid redeposition of metallic contamination, as will be explained below (section 4.2). If allowed for the application, a reoxidation is performed during this rinsing step by injecting O_3 in the rinsing solution. The hydrophilic surface thus obtained is less sensitive to particle contamination and allows for easier drying without the generation of water marks. As mentioned in section 2 the reoxidized surface is preferred for gate oxidation. During this reoxidation step, care must be taken to avoid gas bubble formation in this mixture, since bubbles introduce particle deposition. Drying is preferably performed using the Marangoni technique [20]. The *pH* of the water used during Marangoni drying is also reduced by spiking small amounts of acid.

4.2. final rinsing

It has been recently reported that rinsing clean wafers in ultra pure water results in a built up of metallic surface contamination [21,22] on hydrophilic wafers, particularly of Ca [23–25]. Even after state of the art wet cleaning the surface concentration of Ca ranges typically from 1×10^{10} to 5×10^{10} at./cm² [23].

Fig. 6 shows the results of a first order model of a designed experiment in which the rinse water in the Marangoni drier was intentionally spiked with 2 w-ppb of Ca and varying amounts of HNO_3 [23,26]. These final rinses were performed directly after immersion of clean wafers in either an SC1 or an SC2 mixture. It was found that the metal surface concentration drastically increases with decreasing HNO_3 spiking. The higher concentration found on SC1-last cleaned wafers with respect to SC2-last cleaned wafers was attributed to the carry over of basic SC1 chemical into the final rinse and dry tank.

Another set of designed experiments was performed. Also in these experiments higher [H⁺] results in lower metal surface concentration, σ_M The time dependence of deposition was negligible for most elements. Only Cr, Fe, and Al surface concentrations showed time dependence at *pH* > 5 only. This time dependence is attributed to transport (typically diffusion) limited deposition [24].
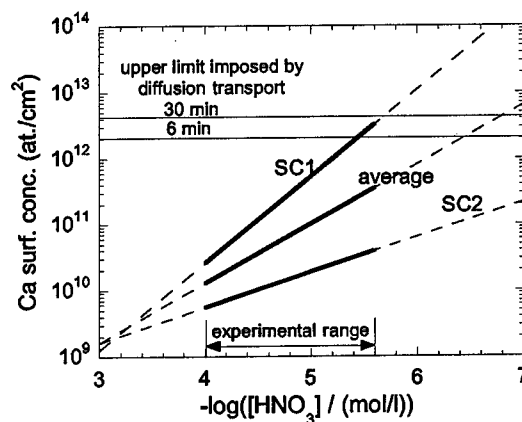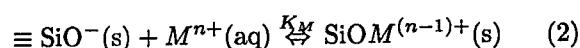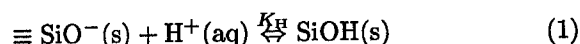

Figure 6.   Dependence of the final Ca surface concentration (equilibrium values) as a function of pH of the final rinse solution and of the last cleaning step prior to the rinsing [23,26]. The rinse water was spiked with approximately 2 w-ppb of Ca.

A model for the adsorption of metallic ions onto hydrophilic surfaces resulting from wafer cleaning has been developed. The acid-base behavior of Si-OH groups act as weakly acidic ion exchangers [27–30]. The chemical oxide undergoes reactions with H⁺ and other cations (e.g., metal ions), M^+, in solution [25]:

$$\equiv SiO^-(s) + H^+(aq) \overset{K_H}{\rightleftharpoons} SiOH(s) \tag{1}$$

$$\equiv SiO^-(s) + M^{n+}(aq) \overset{K_M}{\rightleftharpoons} SiOM^{(n-1)+}(s) \tag{2}$$

K_H and K_M are the constants describing the attraction of H⁺ and M^{n+} to the surface, respectively. It is assumed that multiple cationic species can adsorb to a limited number of surface sites with concentration, σ_0 [32]. Cations, M^{n+}, attach singly to these surface sites. This results in the following surface concentration balance:

$$\sigma_0 = \sigma_{SiOM} + \sigma_{SiOH} + \sigma_{SiO^-} \tag{3}$$

where σ_{SiOM} is the surface concentration of the metal on the wafer surface. Combining the equilibrium expressions corresponding to reactions (1) and (2) with equation (3) yields the following expression for the metal surface concentration:

$$\sigma_{SiOM} = \frac{K_M[M^+]}{1 + K_H[H^+] + K_M[M^+]}\sigma_0 \tag{4}$$

The values of K_H, K_M and σ_0 can then be obtained as fit parameters. Fig. 7 shows the agreement between experimental data [31] and the model of eq. (4) in a conventional deposition plot. This plot also shows why for the range of experimental conditions a sublinear relationship was obtained from the DOE-model. Fig. 8 illustrates the model given by eq. (4) for different pH values. These curves can be used to design an optimal rinsing process in terms of the level of acidification required and purity specification for the UPW. For example at pH 2 the Ca^{2+}-weight concentration should be below 10 ppt in order to keep the Ca surface concentration below 5×10^9 at/cm^2. This model has also been expanded
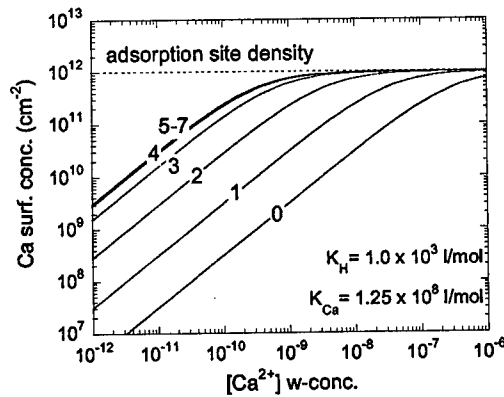

Figure 8. Ca-surface concentration as a function of $[Ca^{2+}]$ weight-concentration for $pH = 0 - 7$ according to eq. (4).
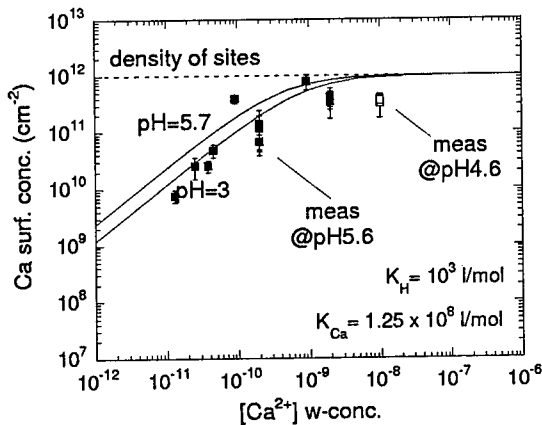

Figure 7. Experimental results [31] and the model of eq. (4) for σ_{Ca} as a function of $[Ca^{2+}]$ weight-concentration.

to describe more realistic situations with different metallic contaminants being present simultaneously [32]:

$$\sigma_{SiOM_i} = \frac{K_{M_i}[M_i^+]}{1 + K_H[H^+] + K_{M_i}[M_i^+]}\sigma_0 \qquad (5)$$

This work contributes to the design of an optimal rinsing processes and helps to specify the required metallic purity of ultra pure water systems. This optimized rinsing has been implemented in the imec clean^{TM} and has demonstrated to yield excellent results in a production line over an extended period of operation.

5. Cl FOR THIN OXIDES

Historically Cl has been introduced in the oxidation ambient mainly in order to reduce the electronic instabilities attributed to the presence of mobile ions, mainly from Na but also to reduce of the density of dielectric breakdown defects [33,34] The beneficial effect on dielectric integrity of thin gate oxides is attributed to the removal of residual metallic contamination (Fig. 9), in particular Ca originating from the wet clean (see above) or from the quartz furnace walls in particular, [35]. In order to meet the stringent future gate-oxide defect density requirements, a well optimized (wet) cleaning process should be used in combination with the *in situ* use of a small amount of Cl during oxidation. Because of corrosion problems, associated with the direct use of HCl [5,37] and Cl$_2$, volatile organic chlorocarbon precursors such as 1,1,1-trichloroethane (TCA) $C_2H_3Cl_3$, *trans*-1,2-dichloroethylene (DCE) $C_2H_2Cl_2$ and oxalyl chloride (OC) $C_2Cl_2O_2$ were introduced.

The use of organic Cl-precursors becomes more difficult for the growth of ultra thin gate oxides with a thickness in the range of 2 to 6 nm. In order to obtain good thickness control, the oxi-
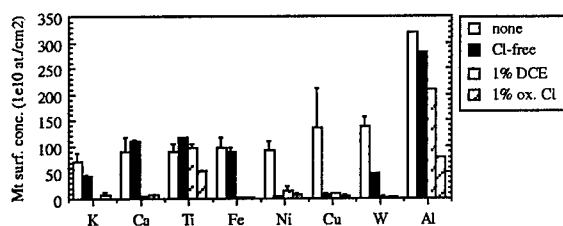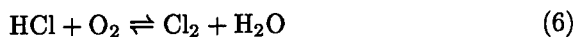
Figure 9. The concentration of metals remaining at the surface or in the grown oxide was measured with VPD - DC - TXRF/GFAAS, directly after contamination as well as after growth of 8 nm oxide at 850°C in Cl-free O_2 or O_2 with DCE or OC at a $C_{Cl-eq.}$ of 1 % [36,35].
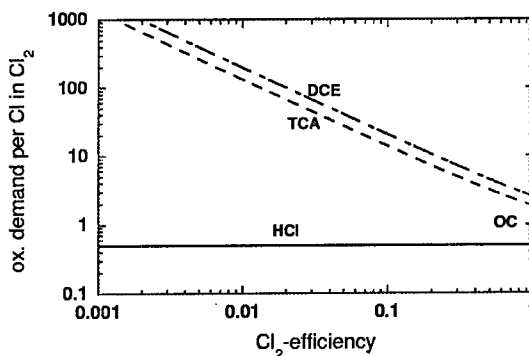

Figure 10. The stoichiometric oxygen demand of the Cl-precursor reactions under consideration as a function of the fraction of the number of Cl atoms transformed into Cl_2 over the total number of Cl atoms in the system [35].

dation time cannot be too short. Therefore, usually the oxidation is performed at relatively low temperature (e.g. 650 to 800 °C) using an ambient consisting of 2 to 10 % O_2 in an inert filler gas, such as N_2. These processes could pose difficulties for the use of chlorocarbon precursors, namely incomplete combustion. Under such cases TCA and particularly DCE are known to result in soot formation. Moreover the combustion of those two precursors should preferably take place well above 700°C.

The overall combustion chemistry equilibrium of TCA and DCE is very similar since for both molecules the number of hydrogen atoms equals the number of Cl atoms. Under oxidizing ambient, such as used in typical gate oxidation processes, TCA and DCE can be seen as direct precursors for HCl. Only a fraction of the HCl is (further) oxidized to form Cl_2 and H_2O, according to the equilibrium of the reaction [33]:

$$HCl + O_2 \rightleftharpoons Cl_2 + H_2O \qquad (6)$$

In case of OC, all Cl is transformed into Cl_2 under oxidizing ambient [38]. Cl_2 and not HCl is believed the active species [35]. Therefore OC is preferred over TCA and DCE for ultra thin oxide growth as it will require less oxygen to generate a given concentration of Cl_2 (see Fig. 10).

A process has been demonstrated for growing oxides in the range of 3nm based on OC making use of a conventional oxidation furnace. Using OC at a $C_{Cl-eq.}$ of only 0.05 to 0.3 % yields the same performance as with 3 % $C_{Cl-eq.}$ HCl (Table 2).

6. CONCLUSIONS

A chemical oxide passivated initial surface is preferred because of its higher stability, more robust particle performance and good thickness control for the growth of ultra thin oxides. Column II metals were found to be most detrimental with respect to the dielectric quality of the gate oxide. Also in case of contamination with Fe, V or Co significant degradation was observed. The imec clean results in very good cleaning performance at low chemical and water consumption. The deposition of metals such as Ca on hydrophilic wafer surfaces, during the final rinse can be suppressed by spiking minute amounts of acid into the rinse water. For growth of ultra thin gate oxides or for *in situ* Cl anneal, the process is performed with reduced oxygen concentration and at reduced temperature, severely limiting the use of common organic Cl-containing precursor such as TCA or DCE. It is demonstrated that under such circumstances the use of small amounts of oxalyl chloride results in successful removal of metallic contamination.

Table 2
Overview of experimental conditions and results of the 30 min 650°C furnace processes [35]. The experimental conditions include the Cl-source gas, the O_2 concentration the Cl-eq. concentration of the Cl-source and the O/Cl ratio. The measured responses are: the thickness of the grown thermal oxide, the average haze lightscattering-fraction, the total amount of carbon present in the thin oxide layer and the reduction of the Ca and Fe surface concentration.

source	O_2	$C_{Cl-eq.}$	t_{ox}	haze	C incorp.	Ca reduc.	Fe reduc.
	(%)	(%)	(nm)	(ppm)	(10^{13} at./cm^2)	(%)	(%)
-	2	0	1.5	0.067	0.016	< 14	< 21
-	2	0	1.86	0.066	1.3	0	< 23
HCl	2	3	1.7	0.067	0.075	85	49
HCl	2	3	2.1	0.066	0.24	> 98	61
OC	2	0.3	3.6	0.080	2	> 98	57
OC	2	0.05	2.3	0.067	0.2	> 98	55

REFERENCES

1. W. Kern, D. Puotinen, RCA Review **31**, 187-206 (1970).

2. The National Technology Roadmap for Semiconductors, (Semiconductor Industry Association, San Jose, 1994), p. 113, 116.

3. M. Depas, T. Nigam, K, Kenis, M.M. Heyns, H. Sprey and R. Wilhelm, in *Proceedings of the Third International Symposium on Ultra-Clean Processing of Silicon Surfaces* (Acco, Leuven, 1996), p. 291.

4. P.W. Mertens, A.L.P. Rotondaro, M. Meuris, H.F. Schmidt, M.M. Heyns and D. Gräf, in *Institute of Environmental Science 40th Annual Technical Meeting 1994 Proceedings* Vol. 1 (IES, Mount Prospect, Illinois, 1994), p. 325.

5. P.W. Mertens, T.Q. Hurd, D. Gräf, M. Meuris, H.F. Schmidt, M.M. Heyns, L. Kwakman, M. Hendriks and M. Kubota, in *Proceedings of the Symposium on Contamination Control and Defect Reduction in Semiconductor Manufacturing III* (The Electrochemical Society, Pennington, NJ), **94-9**, p. 241.

6. SEMI Doc. 2468: *Test Method for the Determination of Residual Surface Contamination on Silicon Wafers by Means of Total Reflection X-Ray Fluorescence Spectroscopy (TXRF)* (Semiconductor Equipment and Materials International, Mountain View, CA, 1997).

7. P.W. Mertens, M. Meuris, H.F. Schmidt, S. Verhaverbeke, M.M. Heyns, P. Carr, D. Gräf, A. Schnegg, M. Kubota, K. Dillenbeck and R. de Blank in *Crystalline Defects and Contamination: Their Impact and Control in Device Manufacturing*, ECS-Proc. Vol. **93-15** (The Electrochemical Society, Pennington, NJ, 1993), p. 87.

8. P.W. Mertens, S. De Gendt, M. Depas, K. Kenis, A. Opdebeeck, P. Snee, D. Gräf, G. Brown and M.M. Heyns, in same as [3], p. 33.

9. T. Bearda, S. de Gendt, L. Loewenstein, M. Knotter, P. Mertens and M. Heyns, in *Proceedings of the Fourth International Symposium on Ultra-Clean Processing of Silicon Surfaces* (Scitec Publications Ltd, Zuerich-Uetikon, Switzerland, 1999), p. 11.

10. W.A. Syverson and P.J. Schubring, in *Proceedings of the Second International Symposium on Cleaning Technology in Semiconductor Device Manufacturing*, 180th Electrochemical Society Meeting, **PV 92-12**, (The Electrochemical Society, Pennington, NJ, 1992), p. 204.

11. P.W. Mertens, M. Meuris, S. Verhaverbeke, M.M. Heyns, A. Schnegg, D. Gräf and A. Philipossian, in *IES 38^{t}h Annual Technical Meeting 1992 Proceedings* Vol. 1, (IES, Mount Prospect, Illinois, 1992), p. 475.

12. M. Takiyama, S. Ohtsuka, S. Hayashi and M. Tachimori, in *Proc. of 19th Workshop on ULSI Ultra Clean Technology; Advanced Silicon Wafer Qualities for ULSI*, (Ultra Clean Soc., 1992), p. 96.

13. P.W. Mertens, M. Meuris, H.F. Schmidt, S. Verhaverbeke, M.M. Heyns, D. Gräf, A. Schnegg, M. Kubota, K. Dillenbeck and A. Philipossian, in *IES 39th Annual Technical Meeting 1993 Proceedings* Vol. 1, (IES, Mount Prospect, Illinois, 1993), p. 231.

14. M. Takiyama, S. Ohtsuka, S. Hayashi and M. Tachimori, in *Proceedings of the Third International Symposium on Silicon Materials Science and Technology*, 185th Electrochemical Society Meeting, **PV 94-10**, (The Electrochemical Society, Pennington, NJ, 1994), p. 346.

15. M. Kobayashi, Y. Ogawa and K. Wada, Extended Abstracts of the Electrochemical Society Spring Meeting 1985 Vol. 85-1 (The Electrochemical Society, Pennington, NJ, 19985), p. 94; M. Liehr, H. Dallaporta and J.E. Lewis, Appl. Phys. Lett. **53**, 7 (1988).

16. M. Meuris, P.W. Mertens, A. Opdebeeck, H.F. Schmidt, M. Depas, G. Vereecke, M.M. Heyns and A. Philipossian, Solid State Technology, July 1995, p. 109.

17. M. Meuris et al. in 1999 IEEE International Symposium on Semiconductor Manufacturing Conference Proceedings (IEEE, Piscataway, NJ, 1999) to be published.

18. S. De Gendt, P. Snee, I. Cornelissen, M. Lux, R. Vos, P.W. Mertens, D.M. Knotter and M.M. Heyns, 1998 Symposium on VLSI Technology Digest of Technical Papers, 168 (1998).

19. W. Krusell et al, Electrochemical Society Meeting, **PV 90-9** (The Electrochemical Society, Pennington, NJ, 1990); I. Teerlinck, P.W. Mertens, R. Vos, M. Meuris and M.M. Heyns, in same as [3], p. 21.

20. A.F.M. Leenaars, J.A.M. Huethorst and J.J. van Oekel, Langmuir **6**, 1701 (1990).

21. I. Lampert and L. Fabry, German patent DE 42 09 865 C2 (30 June 94).

22. L. Fabry, S. Pahlke, L. Kotz, P. Blöchl, T. Ehmann and K. Bächmann, in same as [3], p. 163.

23. L.M. Loewenstein and P.W. Mertens in *Proceedings of the Fifth International Symposium on Cleaning Technology in Semiconductor Device Manufacturing*, 192nd Electrochemical Society Meeting, **PV 97-35**, (The Electrochemical Society, Pennington, NJ, 1998) p. 89.

24. P.W. Mertens, L.M. Loewenstein, R. Vos, S. De Gendt, T. Bearda and M.M. Heyns, in *Proceedings of the Eighth International Symposium on Silicon Materials Science and Technology*, 193rd Electrochemical Society Meeting, **PV 98-1 Vol. 1** (The Electrochemical Society, Pennington, NJ, 1998), p. 592.

25. L.M. Loewenstein and P.W. Mertens, J. Electrochem. Soc. **145**, 2841 (1998).

26. P.W. Mertens, T. Bearda, L.M. Loewenstein, A.R. Martin, W. Hub, B.O. Kolbesen, I. Teerlinck, R. Vos, M. Baeyens, S. De Gendt, K. Kenis and M.M. Heyns, in *Proceedings of the Third International Symposium on defects in Silicon*, 195th Electrochemical Society Meeting, **PV 99-1** (The Electrochemical Society, Pennington, NJ, 1999), p. 401.

27. M. Grundner and H. Jacob, Appl. Phys. **A39**, 73 (1986).

28. T.Q. Hurd, H.F. Schmidt, A.L.P. Rotondaro, P.W. Mertens, L.H. Hall and M.M. Heyns, in *Proceedings of the Fourth International Symposium on Cleaning Technology in Semiconductor Device Manufacturing*, 188th Electrochemical Society Meeting, **PV 95-20**, (The Electrochemical Society, Pennington, NJ, 1996), p. 277.

29. R.K. Iler, *The Chemistry of Silica: Solubility Polymerization, Colloid and Surface Properties, and Biochemistry*, (Wiley, New York, 1979), p. 180-189 and 663-676.

30. B.A. Krishna and J.R. Regalbuto, J. Colloid and Interf. Science **185**, 174 (1997).

31. A. Misra, B. Schmidt, L. Hall, J. Sees and T. Hurd, Semiconductor Fabtech, Winter 1998/1999, 9th ed. p. 173 (1999).

32. L.M. Loewenstein, F. Charpin and P.W. Mertens, J. Electrochem. Soc. **146**, 719 (1999).

33. R.E. Tressler, J. Stach and D.M. Metz, J. Electrochem. Soc. **124**, 607 (1977).

34. J. Monkowski, Solid State Technology July 1979, p. 58; and refs in Ref. [35].

35. P.W. Mertens, M.J. McGeary, M. Schaekers, H. Sprey, B. Vermeire, M. Depas, M. Meuris and M.M. Heyns; in Mater. Res. Soc. Proc. (Mater. Res. Soc., Pittsburgh, PA, 1997) **Vol. 473**, p. 149; or in **Vol. 477**, p. 89.

36. B. Vermeire, P.W. Mertens, M.J. McGeary, K. Kenis, M.M. Heyns, M. Schaekers and A. Lubbers, in *Proceedings of the Second International Symposium on Ultra-Clean Processing of Silicon Surfaces* (Acco, Leuven, 1994), p. 143.

37. G. Vereecke, N. Anderson, C. Elsmore, P. Espitalier-Noel and M.M. Heyns, in same as [3], p. 69.

38. M.J. McGeary, P.W. Mertens, B. Vermeire, M. Heyns, H. Sprey, A. Lubbers and M. Schaekers, in Mater. Res. Soc. Proc. (Mater. Res. Soc., Pittsburgh, PA, 1997) Vol. **447**, p. 115.

Thermal dry oxidation of $Si_{1-x-y}Ge_xC_y$ strained layers grown on silicon

B. Garrido[a*], A. Cuadras[a], C. Bonafos[a], J.R. Morante[a], L. Fonseca[b], M. Franz[c] and K. Pressel[c]

[a]Departament d'Electrònica, Universitat de Barcelona, Martí i Franquès, 1, 08028 Barcelona, Spain
*Corresponding author: tel: +34-93-4029068, fax: +34-93-4021148, e-mail: blas@el.ub.es

[b]Institut de Microelectrònica de Barcelona (CNM-CSIC), Campus UAB, 08193 Bellaterra, Spain

[c]Institut of Semiconductor Physics, Walter-Korsing-Str. 2, 15230 Frankfurt (Oder), Germany

The incorporation of C into strained $Si_{1-x}Ge_x$ to form partially compensated $Si_{1-x-y}Ge_xC_y$ layers improve their critical thickness and thermal stability against relaxation. Thus, these ternary alloys are attractive for the realization of MOS-gated HFETs with the gate grown by thermal oxidation. For this purpose, we present a detailed study of the growth kinetics of SiO_2 in the thin oxide regime for tensile and compressive layers. The oxides have been analyzed by FTIR. The modification of the $Si_{1-x-y}Ge_xC_y$ layers after oxidation has been studied by FTIR (substitutional carbon, β-SiC precipitation) and SIMS (Ge and C depth profiles). From these analyses, suitable process windows for dry thermal growth of oxides with good quality are defined. Preliminary results of the electrical characterization performed on test capacitors are shown.

1. INTRODUCTION

At the present time, the SiGe HBT technology has matured so far that suitable modules for integration into mainstream BiCMOS technology are generally available. A SiGeC HBT is presently under development and will probably reach the market before the end of this year. The main advantages of the SiGeC technology rely on the flexibility in growing the layers with less stress and the suppression of Boron outdiffusion from the base [1]. Generally speaking, the success of SiGe and the promising future SiGeC technologies is a consequence of design transparency with conventional CMOS. Thus, they are compatible with further improvements in CMOS.

The major limiting factor in CMOS is the inferiority of p-MOS devices in terms of current drive capability and speed performance. This is obviously a consequence of the lower mobility of holes than electrons in Si. The transistors in CMOS have thus to be scaled to balance this assymmetric behavior.

Theoretically, it is known that strained $Si_{1-x}Ge_x$ alloys grown on (100) Si could have in-plane hole mobilities significantly greater than Si (from 3 to 10 times that of Si depending on x) in spite of the alloy fluctuation scattering [2]. This increase is attributed to the reduction of hole effective mass coupled with strain-induced heavy/light hole band splitting. Moreover, the valence band offset provides hole confinement.

However, up to date experimental results have demonstrated hole mobilities only slightly superior than those of Si. Nevertheless, encouraging results from the performance point of view have been reported from SiGe p-MOS-gated MODFETs. Impressive hole mobilities of about 1665 cm^2/Vs at room temperature have been reported recently for $Si_{1-x}Ge_x$ strained channels grown by MBE on strain relieved buffers of $Si_{1-y}Ge_y$ (x>y) [3].

2. OXIDATION OF IV-IV STRAINED LAYERS

The growth of a high quality oxide is an important prerequisite before attempting the fabrication of any MOS related device. This is one of the main technological issues that are at stake for full CMOS compatibility of $Si_{1-x}Ge_x$ HFETs. Other issues are also related to the large thermal budgets imposed by CMOS: implant activation and dopant diffusion anneals. The limitation of $Si_{1-x}Ge_x$ strained layers come from the metastable nature of the alloys, the limited critical layer thickness and the thermal cycling that can induce lattice relaxation. This is why p-MOS gated HFETs have been processed mainly with low temperature plasma deposited oxides. Other alternative is to oxidize a Si cap layer, but this procedure has been shown to lead to a competing parasitic Si channel. Within this scenario, we believe that the use of $Si_{1-x-y}Ge_xC_y$ strained layers can alleviate some of these

problems: i) The critical thickness increases with the addition of C due to strain compensation; ii) C incorporation increases the stability against relaxation due to an effective dislocation pinning [4]; iii) Indeed, post-growth moderate annealings at 800 °C have been shown effective to improve crystal quality [5]; iv) Finally, as stated above, C completely freezes B outdiffussion from the doped regions even after high thermal budgets. On the contrary, the drawback is still a thermal budget one: high temperature anneals (let's say T>950 °C) degrade the integrity of the $Si_{1-x-y}Ge_xC_y$ layers by inducing β-SiC precipitation at the expense of the substitutional carbon content, C_s. But precipitates are coherent with the matrix, formation of dislocations is negligible and the in-plain lattice parameter remains constant [10]. Thus, relaxation in $Si_{1-x-y}Ge_xC_y$ is not a mechanical process like in $Si_{1-x}Ge_x$, but a chemical driven one without crystal degradation.

We have focused this work in exploring the dry oxidation behavior of $Si_{1-x-y}Ge_xC_y$ strained alloys as a first step towards defining their ability to become a suitable material for the fabrication of MOS-gated HFET devices. We report reliable dry oxidation kinetics of IV-IV compounds in the thin oxide regime. The structural modification of the $Si_{1-x-y}Ge_xC_y$ layer below the oxide after the oxidation process is studied in detail. Finally, the quality of the oxide and the interface is assessed from the structural and electrical point of view.

3. EXPERIMENT AND OXIDATION KINETICS

Epitaxial fully strained layers of Si, $Si_{1-x}Ge_x$, $Si_{1-y}C_y$ and $Si_{1-x-y}Ge_xC_y$ (x=0.05-0.10, y=0.0025-0.017) of about 250 nm thick have been grown at 400 °C in a Molecular Beam Epitaxy (MBE) system (DCA Instruments) on 100 nm thick silicon buffer layers grown on (100) silicon wafers. Details can be found elsewhere [1,6]. All the subsequent cleaning, oxidation and annealing procedures were performed in a class 100 clean room facility. Prior to oxidation all the samples were submitted to the standard piranya cleaning (H_2SO_4:H_2O_2) followed by a dip into HF (10%) to get rid of the chemical and native oxides. The dry oxidations were performed in a rapid thermal furnace at temperatures between 900 °C and 1100 °C for times long enough to get oxides with thicknesses distributed from few tens of Å to few hundreds of Å (thin oxide regime).

Several studies on wet and dry oxidation of $Si_{1-x}Ge_x$ layers have been reported in the past [7]. There was a consensus, and still there is, on the idea that Ge enhances wet oxidation rate in comparison with that of pure Si. The stoichiometry of the oxides is always SiO_2 for temperatures above 700 °C and the oxidation front rejects all the Ge, that piles up at the interface SiO_2/$Si_{1-x}Ge_x$. On the contrary, the same authors claimed that dry oxidation kinetics of Si and $Si_{1-x}Ge_x$ is identical, even though they also observed pile up of Ge at the interface. Legouges et al. [7] have tentatively explained this discrepancy by considering that in wet oxidation (in the linear regime) the compressively strained Ge rich layer relaxes reducing the stress between the oxide and the substrate. As the oxidation front advances very fast, the formation of Si interstitials is prevented and the oxidation rate increases. No clear explanation is still available for the different behavior during dry oxidation. On the contrary, Tételin et al. [8] have found recently that Ge enhances oxidation rate for low temperature atomic oxygen oxidation of $Si_{1-x}Ge_x$.

Only a few workers have studied the oxidation behavior of carbon containing $Si_{1-x-y}Ge_xC_y$ and $Si_{1-y}C_y$ strained layers. Wet oxidations of $Si_{1-y}C_y$ strained layers have been reported by us in previous papers [6]. We found that concentration of C up to 2% does not affect the wet oxidation rate. Most of the C of the initial layer is consumed due to reaction at the interface and outdiffusion as CO, although some carbon remained pinned at the interface in the form of C-O and clustered C-C. Regarding $Si_{1-x-y}Ge_xC_y$, Atzmon et al. [9] studied wet oxidation at 900 °C and reported Ge pile up. Xiang et al. [10] studied dry oxidation of $Si_{1-x-y}Ge_xC_y$ at 1000 °C and found Ge pile up but no traces of C in the oxide. Neither of them reported data on dry oxidation kinetics.

Our experiments have shown that the measurement of oxide thickness by ellipsometry is somehow tricky. If one relies on the refractive index of the substrate measured before oxidation, the resulting dry oxidation kinetics does not depend on Ge and/or C concentration (details in [11]). Then, the only way to obtain a reliable kinetics is to etch-off part of the oxide and measure the complex refractive index of the substrate after oxidation (there is a significant increase in n and k). Then, by proceeding this way, we obtain the dry oxidation kinetics shown in Fig. 1. It is noticed a strong dependence on Ge and C. The SiO_2 thicknesses agree with HREM results [11].
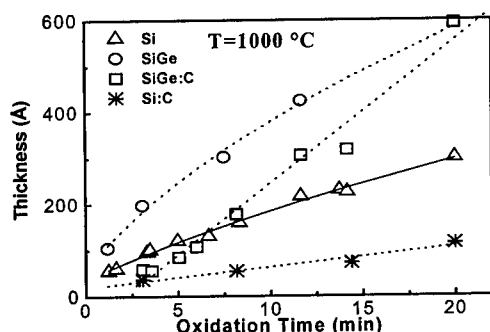
Fig. 1. Dry oxidation kinetics at 1000 °C.

The modification of the complex refractive index of the substrate after oxidation is unambiguously linked with the rejection of Ge from the oxidation front and the always reported piling-up of Ge at the interface. The substrate is enriched of Ge in a region of 200-400 Å. The sample with 10% of Ge has up to 18% at the interface (Fig. 2). Carbon does not pile-up but rather is consumed during oxidation. A small amount of C remains pinned at the interface, the same as reported previously by us in $Si_{1-y}C_y$ [6].
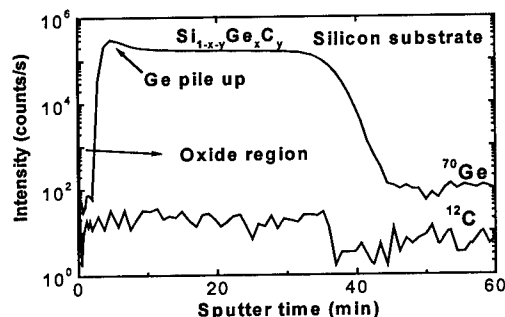
Fig. 2. SIMS profiles after oxidation at 1000 °C.

The dry oxidation kinetics of Fig. 1 demonstrate that: i) Ge acts as a "catalyzator" in both wet and dry oxidation; ii) $Si_{1-y}C_y$ layers are much more difficult to oxidize so C inhibits dry oxidation; iii) The partially compensated $Si_{1-x-y}Ge_xC_y$ layers (Ge/C=20) behave as $Si_{1-y}C_y$ for short oxidation times, overpasse afterwards the Si kinetics and behave assimptotically as $Si_{1-x}Ge_x$ for relatively long oxidation times. In the light of the old models for wet oxidation [7], dry oxidations results in i) and ii) are easily explained. The compressive Ge-rich $Si_{1-x}Ge_x$ layer at the interface partially relaxes. Then, Ge affects the point defect concentration near the oxidizing interface, changing the mechanism from excess Si interstitial production, which occurs for pure Si oxidation, to excess vacancy production via

interdiffusion between Si and Ge. This has the result of promoting easier diffusion of Si through the Ge-enriched layer. Furthermore, Ge-Ge bonds (2.86 eV) and Si-Ge bonds (3.12 eV) are weaker than Si-Si bonds (3.30 eV). Thus, even if some GeO_2 is formed during oxidation, the reaction $GeO_2+Si->SiO_2+Ge$ is thermodynamically biased to the right because of the great differences between Gibbs energy of formation of both oxides (at 1000 °C is $\Delta G_{SiO2}-\Delta G_{GeO2}=$ -299.1 $kJmol^{-1}$). In fact, GeO_2 is unstable and the great interdiffusion and the complete miscibility between Si and Ge can account for the fast kinetics of Fig. 1. The retardation that C imposes to oxidation is explained by just putting the same argument the other way round. The tensile $Si_{1-y}C_y$ layer does not help at all to relax the compressive stress of the oxide and Si interstitials in excess are generated and diffuse down. The strong Si-C bond (4.51 eV) and the very low solubility of C in Si and $Si_{1-x}Ge_x$ limit interdiffusion and the result is the slow down of kinetics.

The $Si_{1-x-y}Ge_xC_y$ oxidation kinetics can be easily understood as well. We have stated above that if a high thermal budget is applied to the $Si_{1-x-y}Ge_xC_y$ layers, β-SiC precipitates are formed. We have shown this effect to occur in oxidation of $Si_{1-y}C_y$ layers [6]. At the initial step of oxidation, the C_s content slows down the oxidation rate. But as oxidation proceeds, more C_s is lost into β-SiC and the $Si_{1-x-y}Ge_xC_y$ layers will eventually behave as $Si_{1-x}Ge_x$. The next section is devoted to quantify by FTIR this effect.

4. FTIR ANALYSIS OF THE SiO_2/ $Si_{1-x-y}Ge_xC_y$

Infrared spectroscopy has allowed us to perform a detailed analysis of the SiO_2/ $Si_{1-x-y}Ge_xC_y$ system. The oxide is characterized from its strong absorption modes (TO_1-TO_4). After the oxide is etched-off, the C_s local vibrational mode (LVM) at 607 cm^{-1} serves to quantify (by using the ASTM standard) the substitutional carbon content of $Si_{1-x-y}Ge_xC_y$. Moreover, the absorption peak of β-SiC at 820 cm^{-1} is used to quantify the amount of precipitates (for a full description of FTIR analyses see [6]). HREM has also been used to quantify independently the amount of β-SiC precipitates. They are scattered within the whole volume of $Si_{1-x-y}Ge_xC_y$ and have a mean diameter of 25 Å with a standard deviation of 8 Å. The inset of Fig. 3 shows the TO_3-TO_4 absorp-tion peaks of oxides grown at 1000 °C. The infrared spectra of Si, $Si_{1-x}Ge_x$ and $Si_{1-x-y}Ge_xC_y$ oxides are remarkably similar.
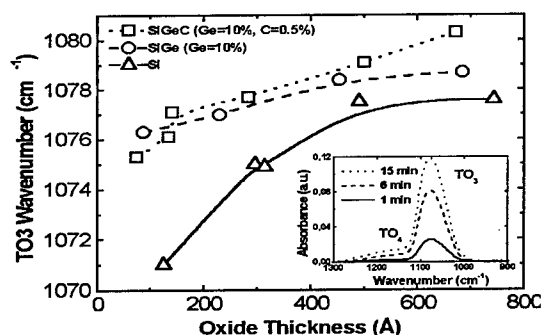
Fig. 3. TO_3 absorption mode of $Si_{1-x-y}Ge_xC_y$ oxides.

Fig. 4. Relative C_s loss and β-SiC absorption (inset).

From the optical and vibrational point of view, the quality is much better than that for deposited or low temperature oxides. The compressive stress in thermally grown oxides is well known and intimately related to the mechanism of oxidation. The position of the TO_3 band is used to quantify the degree of relaxation for oxides of similar thickness. From Fig. 3 we can notice an increase of TO_3 position with thickness for all the samples. This is expected, as for thicker layers the upper part of the oxide relaxes by viscous flow and the thickness of the compressively stressed region is smaller in percentage. But by far the most interesting result is that $Si_{1-x}Ge_x$ and $Si_{1-x-y}Ge_xC_y$ oxides are significantly more relaxed than Si oxide. The TO_3 frequency of them is 5-6 cm^{-1} higher in the low thickness range. Indeed, this is in agreement with the model reported above of a relaxed transition region. For the thinnest oxide $Si_{1-x}Ge_x$ is more relaxed due to the absence of C. But the curve for $Si_{1-x-y}Ge_xC_y$ overpasses the one for $Si_{1-x}Ge_x$ at certain thickness and increases afterwards with higher slope. With no doubt this behaviour is connected with the more effective relaxation of $Si_{1-x-y}Ge_xC_y$ by β-SiC precipitation.

Part of the oxide was etched-off to perform FTIR analyses on the layer below the oxide. This allowed us to quantify the loss of C_s (from LVM at 607 cm^{-1}) as a function of the thermal budget imposed by the oxidation. The results are in Fig. 4, where the vertical scale is referred to the 100% of the initial concentration (0.5%). These results suggest that for minimizing C_s loss, oxidation temperature must be low. From the inset of this Fig. 4 is clear that loss of C_s correlates with an increase of absorption by β-SiC precipitates. We do not know whether these precipitates affect electrical properties. Quantification of C in the form of β-SiC has been performed from FTIR and HREM. We estimate that about 50% of the C_s lost is in the form of β-SiC.

5. ELECTRICAL CHARACTERIZATION
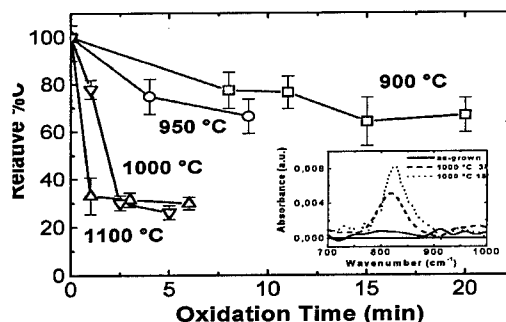
MOS capacitors have been fabricated with low thermal budget sequence. We chose a conservative approach so that composition and stability of epilayers was not compromised. PECVD oxide was deposited for isolation and Al for electrodes. But, due to the PECVD oxide, all the samples, even the Si ones, exhibited a leaky behavior that prevented measurement of quasi-static curves. But some trends were inferred from the high-low frequency C(V) curves. The epilayers were grown on 10^{15}cm^{-3} n-type Si. We find Si and $Si_{1-x}Ge_x$ have underdoping (4×10^{14}cm^3). But the samples containing C exhibited a huge doping level (10^{18}-10^{19}cm^{-3}). $Si_{1-y}C_y$ shows the highest (9×10^{18}cm^{-3}) while samples with Ge and C show less but still high. Regarding V_{FB}, $Si_{1-x}Ge_x$ tends to give negative values while C contributes with positive pointing out opposite sign of fixed charge. We found $D_{it}=5 \times 10^{11}$ cm^{-2}eV^{-1} for $Si_{1-x}Ge_x$ oxide, 10 times that of Si oxide. High R_s and doping level of samples with C made the C(V)'s meaningless. As a conclusion, the dry oxidation kinetics of IV-IV compounds has been revealed. Ge is rejected and C outdiffuses but some C remains at interface. A low thermal budget process must be optimized.

References
[1] H.J. Osten, B. Heinemann, D. Knoll, G. Lippert, H. Rücker, J. Vac Sci. Technol B, **16**, 1750 (1998).
[2] T.Manku, A.Nathan, IEEE E. Dev. Lett.,**12**,704 (1991).
[3] G. Köck, M. Glück, T. Hackbarth, H. Herzog, E. Kohn, Thin Solid Films, **336**, 141 (1998).
[4] P. Zaumseil, G.G. Fischer, K. Brunner, K. Eberl, J. Appl. Phys., **81**, 6134 (1997).
[5] K. Pressel, G. Fischer, P. Zaumseil, M. Kim, H.J. Osten, Thin Solid Films, **294**, 133 (1997).
[6] B. Garrido, J.Morante, M. Franz, K. Pressel, D. Krüger, P. Zaumseil, HJ. Osten, J. Appl. Phys., **85**, 833 (1999).
[7] F.K. Legouges, R. Rosenberg, B.S. Meyerson, Appl. Phys. Lett., **54**, 644 (1989).
[8] C. Tételin et al , J. Appl. Phys., **83**, 2842 (1998).
[9] Z. Atzmon et al., Appl. Phys. Lett., **66**, 2244 (1995).
[10] J. Xiang et al. J. Appl. Phys., **80**, 1857 (1996).
[11] A. Cuadras et al., E-MRS (Symp. P), June 1999.

Electrical characterization of n-channel MOSFETs with oxynitride gate dielectric formed by Low-Pressure Rapid Thermal Chemical Vapor Deposition

P. Masson[a-b], P. Morfouli[b], J.L. Autran[a] and J.J. Wortman[c]

[a] LPM, UMR CNRS 5511, INSA de Lyon, Bâtiment 502, F-69621 Villeurbanne, France
[b] LPCS, UMR CNRS 5531, ENSERG, BP 257, F-38016 Grenoble, France
[c] DECE, NCSU, 404 Daniels Hall, Box 7911, Raleigh, NC 27695-711, USA
  E-mail : masson@insa-lyon.fr - Tel : + 33 472 438 730 - Fax : +33 472 438 531

This study reports on n-channel MOS transistors using very thin oxynitride gate dielectrics deposited by Low-Pressure Rapid Thermal Chemical Vapor Deposition (LPRTCVD). The threshold voltage, the transconductance, the fast and slow trap densities, the low field mobility and its reduction factors have been investigated as a function of the nitrogen concentration in the film. The threshold voltage was found to decrease linearly with the nitrogen concentration, primarily because of the increase of the positive insulator charge density. At the same time, it has been shown that the quadratic mobility reduction factor is lower for a LPRTCVD oxynitride (and nearly constant with the nitrogen concentration) in comparison to a thermal oxide. From this high field mobility attenuation, a lower transconductance is then expected for oxynitride-based transistors. Finally, we have shown that the presence of nitrogen in the oxide seems to induce more donor than acceptor traps at the oxynitride-silicon interface.

1. INTRODUCTION

Increasing the level of integration in present technologies requires the use of MOS devices with thinner and thinner gate oxides. Oxynitrides are one of the most extensively studied alternative ultrathin gate dielectrics for very large scale integrated (VLSI) circuits. Oxynitrides formed by Low Pressure Rapid Thermal Chemical Vapor Deposition (LPRTCVD) present several advantages when compared with silicon dioxide, such as radiation hardness, larger barrier to impurity diffusion as well as a reduction of mobility attenuation at high electrical field [1-3]. The aim of this work is to investigate the electrical properties of thin gate dielectrics (5-7 nm) in n-channel MOS transistors as a function of the nitrogen concentration. The main electrical parameters (threshold voltage, low-field mobility) as well as the slow trap and interface trap densities, respectively extracted from low frequency noise and charge pumping measurements, have been studied. The correlation between these two techniques has been also investigated.

2. TECHNOLOGICAL DETAILS

Isolated enhancement mode n-channel MOS transistors have been fabricated on p-type (0.05-1 Ω.cm) <100> silicon substrate with phosphorus-doped polysilicon gate. The oxynitride films were deposited in a cold wall LPRTCVD system by using a mixture of SiH_4, N_2O and NH_3 at 800 °C. The nitrogen concentration, determined from Secondary Ion Mass Spectroscopy (SIMS) analysis, is varying from 0 % up to 11 %. In addition, nitrogen has been found to be homogeneously distributed in the bulk of the oxide [2]. Reference devices with thermal gate oxide grown at 900 °C under dry atmosphere have been also fabricated.

3. RESULTS AND DISCUSSION

All electrical parameters reported in the present paper have been determined on n-channel MOSFETs with W/L = 10/3 μm. The initial electrical parameters of these devices have been previously investigated [4].
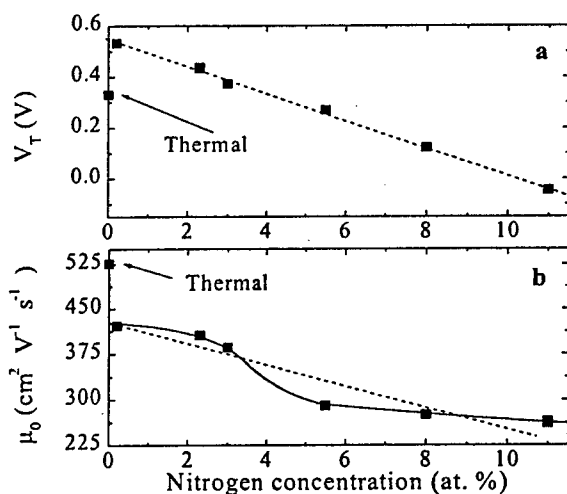
Fig. 1. Experimental evolution of the threshold voltage (a) and the low field electron mobility (b) versus nitrogen concentration in the gate dielectric.

3.1. Threshold voltage and low field mobility

As usually reported [8], the fixed charges present in the dielectric have been found positive. The linear decrease of the threshold voltage (V_T), shown in Fig. 1.a, suggests that the nitrogen incorporation in the oxide results in a large amount of positive charges (Q_{ox}) which dominate the opposite contribution of interface traps. Indeed, these traps are negatively charged in strong inversion and for a n-channel MOSFET (the acceptor and donor trap charges are negative and neutral respectively). For a p-channel MOSFET, in strong inversion regime, the charge due to interface traps is positive, accentuating the decrease of V_T [2, fig. 8].

The extraction of Q_{ox} has been performed from the MOSFET $I_{DS}(V_{GS})$ characteristics in weak inversion regime by determining the value of the gate potential when the semiconductor surface is between the depletion and weak inversion regimes (midgap voltage V_{mg}). This particular voltage corresponds to a neutral charge into the interface traps if we assume that these states are acceptors in the upper part of the bandgap and donors in the lower part. The extracted values of Q_{ox} variations (ΔQ_{ox}) present a linear dependence with V_T, as previously shown in Ref. [8].

Fig. 1.b shows the evolution of the low field mobility (μ_0) with the nitrogen concentration in the gate dielectric. This quasi-linear decrease of μ_0 is due to the increase of the charge density near the
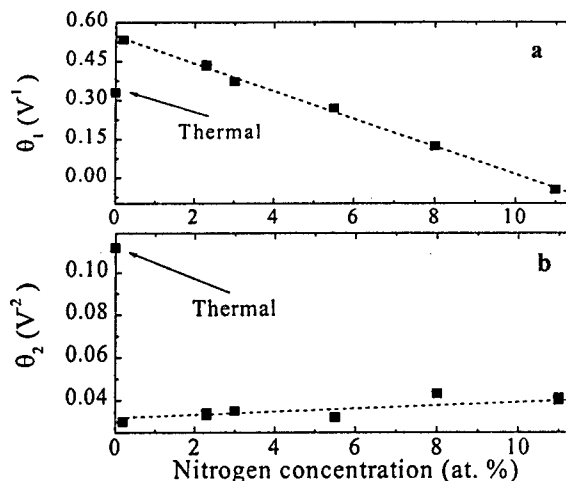
Fig. 2. Evolution of the linear(a) and quadratic (b) mobility reduction factors versus nitrogen concentration. The dot lines underline the evolution of θ_1 and θ_2.

conduction channel and also to the (linear) increase of Q_{ox} versus nitrogen concentration [10].

As presented in Fig. 1.b, the decrease of μ_0 is more pronounced for nitrogen concentration range between 3 and 4 %. This phenomenon cannot be explained on the electrical point of view, even if it was reported in numerous studies [3] that a nitrogen concentration in the range of 3-4 % represents a good compromise between dielectric quality improvement in terms of impurity diffusion or stress resistance and transistor characteristic degradation (charge density, threshold voltage, etc.)

3.2. Mobility reduction factors

The linear (θ_1) and quadratic (θ_2) mobility reduction factors are respectively attributed to the coulombic and roughness scattering [11]. The decrease of θ_1 with nitrogen concentration, shown in Fig. 2.a, can be interpreted as an increase of the charge density near the conduction channel. This result confirms the increase of both the oxide charge and the interface trap density with the nitrogen concentration. It also confirms the evolution of V_T.

Fig. 2.b shows a slight increase of θ_2 with the nitrogen concentration. In the case of LPRTCVD oxynitrides, we notice a very low value of θ_2 by comparison of thermal oxide. Such a result suggests that the presence of nitrogen at the dielectric-silicon interface may reduce the stress and consequently the interface roughness, or has a similar action on the transport of the carriers.
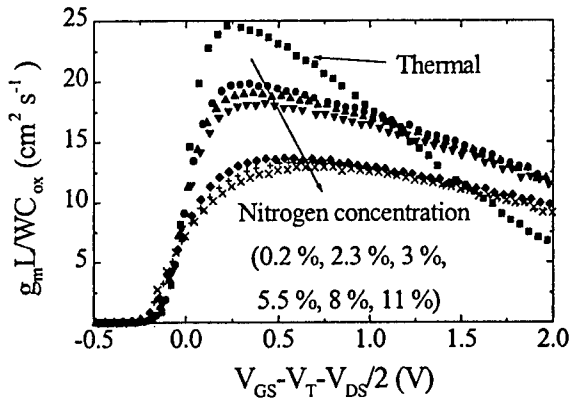
Fig. 3. Normalized channel transconductance (with respect to oxide capacitance and channel geometry) as a function of the gate voltage.

3.3. Channel transconductance

The presence of nitrogen atoms in the oxide is well known to improve the transconductance g_m at high electric field [3]. Fig. 3 shows the normalized transconductance as a function of the gate voltage for different devices. From this data, we can observe that oxynitrides present a lower maximal transconductance (g_{mmax}) as compared to thermal oxide. Moreover, as evidenced for μ_0 in the case of oxynitrides, we point out the same fast decrease of g_{mmax} between 3 and 4 % of nitrogen in the film. Finally, the decrease of g_m is found slower at high electric field for oxynitrides than for thermal oxide. This result is in good agreement with the value of θ_2 which has been found lower and sensibly constant with the nitrogen concentration for the oxynitrides.

3.4. Fast and slow trap densities

The fast and slow trap densities at or near the dielectric-silicon interface have been respectively determined by charge pumping (CP) and low frequency (1/f) noise. In its standard (2-level) implementation, the first technique gives quantitative information about the fast interface trap density (D_{it}) [12] whereas the 3-level method is used to extract the energy distribution $D_{it}(E)$ of the interface traps in the silicon bandgap [7]. The second technique allows the determination of the energy trap volume density (N_t) at ~1 nm from the interface and for traps energetically located at a few kT around the Fermi energy level in strong inversion regime [13].
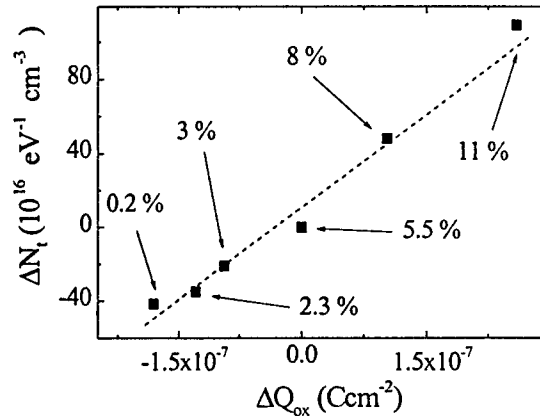
Fig. 4. Correlation between the variations of slow trap density and fixed oxide charge.

The variation of the slow trap density (ΔN_t) versus the variation of the fixed oxide charge (ΔQ_{ox}) are shown in Fig. 4. The good correlation between ΔQ_{ox} and ΔN_t highlights a possible common origin for these quantities.

Another result concerns the mean fast trap density (D_{it}) which has been found to be linear with nitrogen concentration [8]. In Fig. 5, the correlation between the slow traps density and the interface traps density shows a good linearity. That suggests an underlying link between the number of traps in the volume and at the interface of the dielectric. In an other way, an increase of trap density in the dielectric is followed by a proportional increase of interface traps and also by an increase of the fixed oxide charge, as clearly shown in Fig. 4.
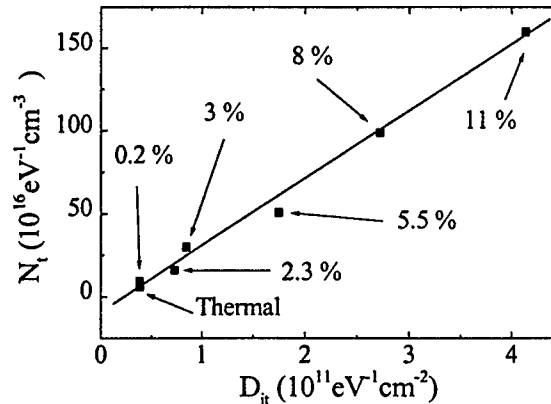

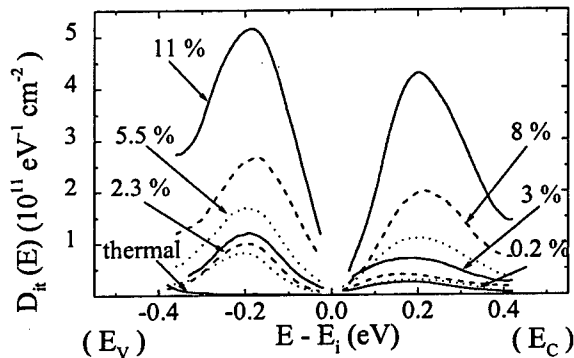Fig. 5. Correlation between interface trap and slow trap densities.

Fig. 6. Interface trap energy distribution in the silicon bandgap determined by three-level charge pumping as a function of the gate dielectric.

Finally, the variation of the interface state density with the nitrogen concentration, as deduced from 3-level charge pumping, is shown in Fig. 5. Assuming acceptor interface traps in the upper part of the bandgap and donors traps in the lower part, data show that the presence of nitrogen in the oxide seems to induce more donor than acceptor traps. This results in an asymmetry of the energy-resolved D_{it} distribution, as already observed in previous work using capacitance-voltage measurements [14].

4. CONCLUSION

In this paper, n-channel MOS transistors using oxynitride films deposited by low-pressure RTCVD as gate dielectric have been studied. We have shown that the threshold voltage of the devices decreases linearly with the nitrogen concentration. Such a behavior has been interpreted as the direct consequence of the increase of the fixed charge in the film induced by nitrogen. Moreover, we have established that the low field mobility rapidly decreases with varying nitrogen concentration from 3 to 4 %. In addition, this work shows that the quadratic mobility reduction factor is lower for a LPRTCVD oxynitride (and relatively constant versus the nitrogen concentration) than for a thermal oxide. This important difference is certainly at the origin of the lower decrease of g_m at high electric field for the oxynitrides, by comparison to the rapid degradation observed in the case of the thermal

reference oxide. Results also show that the oxide fixed

charge, the interface and the volume trap densities linearly increase with the nitrogen concentration in the oxide. Finally, we have established that nitrogen in the oxide induces more donor than acceptor traps at the oxide-silicon interface.

REFERENCES

[1] M.J. Rand and J.F. Roberts, J. Electrochem. Soc. 120, 446 (1973).

[2] W.L. Hill, E.M. Vogel, V. Misra, P.K. McLarty and J.J. Wortman, IEEE Trans. Electron Devices 43, 15 (1996).

[3] B. Balland and A. Glachant, in *Instabilities in Silicon Devices*, edited by G. Barbottin and A. Vapaille (Amsterdam: Elsevier North-Holland, 1999), chapter 1.

[4] P.K. McLarty, S. Cristoloveanu, O. Faynot, V. Mistra, J.R. Hauser and J.J. Wortman, Solid-State Electron. 38, 1175 (1995).

[5] P. Masson, J.L. Autran, C. Raynaud, O. Flament and P. Paillet, IEEE Trans. Nuclear Science 45, 1355 (1998).

[6] P. Masson, G. Ghibaudo, J.L. Autran, P. Morfouli and J. Brini, IEE Electron. Lett. 34, 1977 (1998).

[7] J.L. Autran, B. Balland, Rev. Sci. Instrum. 65, 2141 (1994).

[8] P. Masson, PhD thesis n° 99ISAL007, INSA de Lyon, France (1999).

[9] G. Lucovsky and D.J. Chadi, in *The physics of MOS insulator* (New-York : Pergamon Press, 1981) 301.

[10] S.C. Sun and J.D. Plummer, IEEE Trans. Electron Devices 27, 1497 (1980).

[11] G. Reichert and T. Ouisse, IEEE Trans. Electron Devices 43, 1934 (1996).

[12] G. Groeseneken, H.E. Maes, N. Beltran and R.F De Keersmaecker, IEEE Trans. Electron Devices 31, 42 (1984).

[13] S. Christensson, I. Lundstrom and C. Scensson, Solid-State Electron. 11, 797 (1968).

[14] M. Severi, L. Dori, M. Impronta and S. Guerri, IEEE Trans. Electron Devices 36, 2447 (1989).

The effects of radiation-induced defects on H^+ transport in SiO_2

R. E. Stahlbush

Naval Research Laboratory, Code 6816, Washington, DC 20375 USA

The effects of radiation on H^+ transport in buried oxides have been studied. Mobile H^+ was introduced into the buried oxides in SIMOX and Unibond by hydrogen annealing. The position of the H^+ centroid was monitored by rapid point-contact transistor measurements. After irradiation, H^+ transit across the oxide is significantly slower. The concentration of mobile H^+ in the oxide remains nearly the same after irradiation. The significance of these findings to the contradictions between the radiation and hydrogen-annealing models are discussed.

1. INTRODUCTION

Despite three decades of study, recent experiments have shown that our understanding of the interactions between hydrogen and ionizing radiation in MOS systems is far from complete. In particular, recent experiments studying hydrogen-annealed buried oxides by Vanheusden et al. [1] challenge the hydrogen-based model for radiation-induced interface state formation by McLean [2]. Both the hydrogen-annealing and radiation models rely on the same H^+ entity. However, the properties attributed to H^+ are different in the models [3, 4].

The radiation model is based on the results of pulse irradiation studies and explains the buildup of radiation-induced interface states. The mechanism for the buildup is a multi-step process in which H^+ is created during the pulse, drifts to the Si/SiO_2 interface and reacts to form interface states. This model successfully describes how the time dependence of the interface state buildup depends on physical parameters including electric field, oxide thickness and temperature [5]. A prominent feature of the model is that the rate-limiting step is the transport of H^+ from the bulk of the oxide to the Si/SiO_2 interface. Another feature is that most of the H^+ reacts at the interface forming interface states.

The radiation model is very consistent with experiments in which interface states are formed by introducing atomic hydrogen into undamaged oxides [6] and with experiments in which interface state formation is caused by introducing molecular hydrogen into irradiated MOSFETs [7]. The radiation model is also consistent with hot-electron studies that concluded that damage within the oxide

bulk and at the Si/SiO_2 interface is due to hydrogen released by the hot electrons [8].

While the evidence for the radiation model is very compelling, recent hydrogen annealing experiments, which are equally persuasive, have attributed different characteristics to the same H^+ entity. Annealing oxides, such as in Unibond and SIMOX, in hydrogen between 500 and 800 °C introduces mobile positive charge that has been attributed to H^+ [1]. The same effect has been observed in thermal oxides that have been covered by polysilicon and annealed at 1100 to 1200 °C [9]. The characteristics attributed to H^+ in these experiments are different from the characteristics derived from the radiation model. In the hydrogen-annealing model, H^+ is very stable and does not react at Si/SiO_2 interfaces. In most oxides, it can be repeatedly cycled between the Si/SiO_2 interfaces without reacting or being trapped [4]. Its field-induced drift through the oxide is three orders of magnitude faster than the H^+ drift derived from the radiation experiments [3, 4].

Previous studies also suggest possible resolutions to the discrepancies between the two models. Our previous work has shown that H^+ trapping and transport is strongly affected by process-induced defects [4]. The question arises about the effects of radiation-induced defects on H^+ transport. It is also known that irradiation affects the reactivity of molecular hydrogen in silicon dioxide [7]. Before irradiation H_2 is inert at room temperature whereas after irradiation H_2 introduced into the oxide reacts and forms interface states.

In this paper, the effects of radiation on H^+ transport and on its reactivity at the Si/SiO_2

interfaces are addressed for Unibond and triple-implant SIMOX. The transport time can be increased by an order of magnitude at relatively low irradiation dose (10-20 krad(SiO$_2$)). This increase is sensitive to the irradiation bias and to defects present before the irradiation. The increase provides a possible resolution to the vastly different transit times ascribed to H$^+$ by the radiation and hydrogen-annealing models. However, the discrepancy concerning the reactivity of H$^+$ at the Si/SiO$_2$ interfaces remains. H$^+$ reactivity is not noticeably changed after irradiation and the amount of mobile H$^+$ is nearly the same before and after irradiation.

2. EXPERIMENTAL DETAILS

Hydrogen transport in Unibond and triple-implant SIMOX material was examined. Unibond is made by a wafer-bonding process in which a thermally grown oxide becomes the buried oxide. After bonding, the bond is strengthened by annealing at 1100 °C. The SIMOX had two oxygen doses of 0.5x10^{18}/cm^2 and a third dose of 0.5x10^{18}/cm^2. It was annealed at 1325 °C after each dose.

Structures for point-contact transistor measurements [10] were made by forming 0.8 mm diameter islands in the top silicon layer. The islands were fabricated by masking the surface with aluminum dots, etching the silicon layer in hydrazine and then removing the aluminum.

Mobile positive charge was introduced by annealing at 700 °C in H$_2$ for 30 minutes. During the cool down, the hydrogen was left flowing until the sample temperature was below 200 °C.

After H$_2$ annealing, devices were irradiated with x-rays from a tungsten tube.

The position of the H$^+$ centroid was monitored by rapid I-V measurements of the point-contact transistors. In these measurements, the wafer substrate serves as the gate and probes contacting the islands in the top silicon layer act as the source and drain. The voltage near midgap, V$_{midgap}$, is used to measure the oxide charge. The charge is sensed from the top Si/SiO$_2$ interface and since the amount of charge is known to be constant[4], the centroid can be determined. I-V measurements were chosen because it was convenient to make rapid I-V measurements, typically 0.4 sec. A rapid I-V sweep minimizes the perturbation of the H$^+$ distribution within the oxide during the sweep.

3. RESULTS AND DISCUSSION

The effect of x-ray irradiation on H$^+$ cycling in Unibond material is shown in Fig. 1. The three frames of Fig. 1 show cycling (a) before irradiation, (b) after 15 krad(SiO$_2$), -100 V to substrate, and (c) after 15 krad(SiO$_2$), +60 V to substrate. The corresponding applied fields are 2.5 and 1.5 MV/cm. During the irradiation, the H$^+$ is near the bottom Si/SiO$_2$ interface in Fig. 1b and near the top interface in Fig. 1c. The H$^+$ was cycled by alternately applying +60 and –100 V to the substrate with the top silicon layer grounded. Before and after irradiation, H$^+$ can be repeatedly cycled without any reduction in its concentration (6.5x10^{12}/cm^2). The cycling after irradiation is offset by trapped positive charge. The offset in Fig. 1b is small, > 10 V, and larger in Fig. 1c, 50 V. This is consistent with holes collecting near the bottom Si/SiO$_2$ interface for negative polarity and collecting near the top interface for the opposite polarity.

The individual data points within each cycle are very reproducible. Furthermore, the H$^+$ concentration revealed by the cycling is nearly the same before and after irradiation. The small reduction of the cycling amplitude in Fig. 1c is probably due to the shorter cycling time that prevents the complete transit of H$^+$ across the oxide. These results indicate that H$^+$ is not more reactive at the Si/SiO$_2$ interfaces after irradiation.

The I-V curves are stretched after irradiation. However, this is not strong evidence for interface trap formation because lateral nonuniformity of trapped positive charge can produce the same stretchout. Thus, the discrepancy between the radiation and hydrogen-annealing models concerning the reactivity of H$^+$ at the Si/SiO$_2$ interfaces and formation of interface traps is not resolved.

To examine the effect of irradiation on the H$^+$ transit time, the data in Fig. 1 is replotted using a log(time) scale and a normalized H$^+$ centroid during a single cycle. In Fig. 2a, H$^+$ is moving from the top Si/SiO$_2$ interface to the bottom interface and in Fig. 2b, H$^+$ is returning to the top. In each frame, the transit is compared before irradiation and after the two irradiation conditions. For H$^+$ motion in both directions, its drift is slower after irradiation. This result suggests a resolution to the discrepancy between the radiation and hydrogen-annealing models concerning the speed of H$^+$ drift. In the
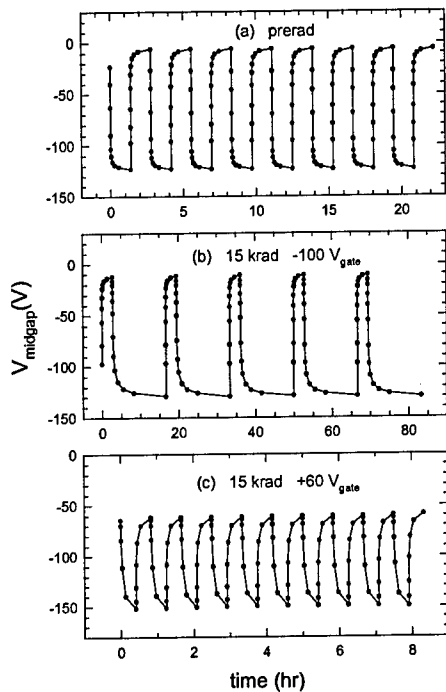
Figure 1. Cycling of H^+ between the top and bottom interfaces of the buried oxide in Unibond.

Figure 2. H^+ transport across the Unibond oxide.

pulse-irradiation experiments the dose was 50 krad(SiO_2) [11]. A three order of magnitude slowing of H^+ transport is consistent with the above results.

The slowing of H^+ motion is more pronounced as it drifts from the bottom interface to the top interface. The H^+ transit curve in Fig. 2b is shifted to the right by an order of magnitude. This suggests that the irradiation produces more defects near the bottom interface than near the top and that these defects govern the H^+ transport. The traps are shallow enough so that H^+ escapes for the cycling times shown in Figs. 1 and 2. More traps may be formed near the bottom interface because the wafer bond is at this interface.

The data also suggest that more traps are produced near an interface when H^+ is at that interface during the irradiation as shown in Figs. 2a and 2b. During the irradiation with -100 V (+60 V) applied to the substrate, H^+ is at the bottom (top) interface. In Fig. 2a, H^+ transport from the top to the bottom interface is slower after irradiation with H^+ at the top interface. In contrast, in Fig. 2b, H^+ transport from the bottom to top interface is slower after irradiation with H^+ at the bottom interface.

Figures 3 and 4 show similar results for a triple-implant SIMOX sample. Figure 3 shows cycling (a) before irradiation, (b) after 20 krad(SiO_2), -100 V to substrate, and (c) after 20 krad(SiO_2), +60 V to substrate. The effects of irradiation on H^+ in SIMOX are very similar to the effects on H^+ in Unibond. The cycling of H^+ is very reproducible from cycle to cycle both before and after irradiation. There is a 20% reduction in the H^+ concentration from the initial $6.0 \times 10^{12}/cm^2$ value after irradiation with either bias. Similar to Unibond, the is a stretchout of the I-V curves after irradiation. These effects are probably not due to H^+ reacting to form interface traps. If these effects were due to the reaction of H^+ at the Si/SiO_2 interfaces to form interface traps, the implication would be that the reaction stops after 10^{12} /cm^2 H^+ have reacted despite the presence of much more H^+. This is not consistent with radiation-induced interface traps which can well exceed this level.

The voltage offset to the cycling of V_{midgap} after irradiation is about 40 V regardless of the polarity of the applied voltage during irradiation. This is consistent with the observation that the buried oxide in SIMOX has a much higher concentration of hole traps than the buried oxide in Unibond. [12]

The transit of H^+ across the SIMOX buried oxide shown in Fig. 4 is also similar to the transit in Unibond shown in Fig. 2. In particular, the slowing of H^+ after irradiation is more pronounced for
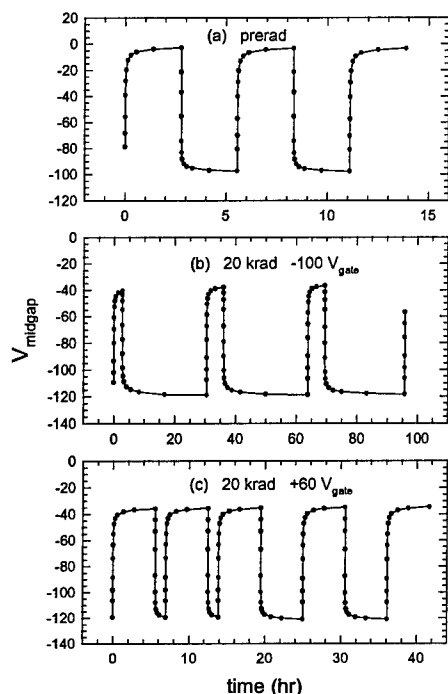
Figure 3. Cycling of H^+ between the top and bottom interfaces of the buried oxide in SIMOX.

Figure 4. H^+ transport across the SIMOX oxide.

motion from the bottom to top Si/SiO_2 interface. Also, H^+ motion in this direction is slowed more following irradiation with H^+ present at the bottom interface. The slowing of H^+ transport from the top to bottom interfaces is barely noticeable. These results suggest that there is a higher concentration of defects near the bottom interface that interact with the hydrogen during irradiation to form shallow H^+ traps. This is consistent with previous observations of H^+ trapping near the bottom interface. [3]

4. SUMMARY

Radiation-induced defects have been shown to have strong effects on H^+ transport and trapping in Unibond and SIMOX buried oxides. The details of these effects depend on fabrication-induced defects already present in the oxide and on the dose and bias conditions of the irradiation. The slowing of H^+ transport after irradiation helps to resolve the apparent contradictions between the radiation and hydrogen-annealing models. However, the discrepancy between the assumed reactivity of H^+ at Si/SiO_2 interfaces is not resolved.

REFERENCES

1. K. Vanheusden, W. L. Warren, R. A. B. Devine, D. M. Fleetwood, J. R. Schwank, M. R. Shaneyfelt, P. S. Winokur, and Z. L. Lemnios, Nature **386**, 587, (1997).
2. F. B. McLean, IEEE Trans. Nucl. Sci. **NS-27**, 1651, (1980).
3. K. Vanheusden, R. A. Devine, J. R. Schwank, D. M. Fleetwood, R. G. Polcawich, W. L. Warren, S. P. Karna, and R. D. Pugh, IEEE Trans. Nucl. Sci. **NS-44**, (1997).
4. R. E. Stahlbush, R. K. Lawrence, and H. L. Hughes, IEEE Trans. Nucl. Sci. **NS-45**, 2398, (1998).
5. D. B. Brown and N. S. Saks, J. Appl. Phys. **70**, 3734, (1991).
6. E. Cartier, J. H. Stathis, and D. A. Buchanan, Appl. Phys. Lett. **63**, 1510, (1993).
7. R. E. Stahlbush, B. J. Mrstik, and R. K. Lawrence, IEEE Trans. Nucl. Sci. **NS-37**, 1641, (1990).
8. D. J. DiMaria and J. W. Stasiak, J. Appl. Phys **65**, 2342, (1989).
9. W. L. Warren, D. M. Fleetwood, J. R. Schwank, K. Vanheusden, R. A. B. Devine, L. B. Archer, and R. M. Wallace, IEEE Trans. Nucl. Sci. **NS-44**, 1789, (1997).
10. S. Cristoloveanu and S. S. Li, *Electrical Characterization of Silicon-On-Insulator Material and Devices*. Boston: Kluwer Academic, 1995.
11. N. S. Saks and D. B. Brown, IEEE Trans. Nucl. Sci. **NS-36**, 1848, (1989).
12. R. E. Stahlbush, IEEE Trans. Nucl. Sci. **NS-44**, 2106, (1997).

Surface Potential Influence on Defect Passivation Kinetics Probed by Chromium Gated Metal - Oxide - Silicon Devices

L.-Å. Ragnarsson and P. Lundgren
Microtechnology Centre at Chalmers and Solid State Electronics Laboratory,
Department of Microelectronics, Chalmers University of Technology,
SE-412 96 Göteborg, Sweden, phone: +46 31 772 18 67, fax: +46 31 772 36 22, E-mail: loke@ic.chalmers.se

In this study we present a new scheme for obtaining information on the passivation mechanisms of as-grown Si—SiO$_2$ interface defects by resolving the effects of bias on the defect passivation reaction. This is possible when using chromium gated metal-oxide-silicon devices for which the effect of non-biased passivation is very small compared to devices with for example aluminum gates. It is shown that a negative surface potential promotes the passivation of interface states while a positive surface potential reduces it.

The success of CMOS technology relies on efficient passivation of dangling bonds at the Si-SiO$_2$ interface, which motivates research on this subject [1,2]. Previous studies have shown that the rate of passivation of these defects is influenced by the applied voltage [3,4]. In these studies of Al-gated metal-oxide-silicon (MOS) devices on both <111> and <100> oriented p-type substrates it was argued that the charge-state of the P_b-center influences its passivation rate as predicted in theoretical calculations [5].

However, the Al-gated devices display a high passivation rate without any applied bias which makes it difficult to draw accurate quantitative conclusions regarding the effects of biased annealing. For Cr-gated devices the passivation rate without bias is considerably slower making it easier to resolve the effects of the bias. The slower passivation rates could be explained by a smaller amount of passivating species (hydrogen) in the Cr devices. It could also be due to the fact that an oxide reduction reaction which has been shown to *reduce* interface defects [6,7] occurs in Al-gated MOS devices but not in Cr-gated ones. This oxide reduction is believed to be caused by a partial reduction of the SiO$_2$ by Al into Al$_2$O$_3$ [7-9].

In this study, with the use of Cr instead of Al as the gate material, we are not only able to resolve the bias dependence of the passivation kinetics but also to show that the passivation rate is dependent on the charge-state of the defects.

MOS devices were fabricated on <100> oriented (1-10 Ωcm) n- and p-type silicon wafers. The wafers were oxidized to form a thick field oxide followed by a lithography step where devices with areas ranging from 2 x 2 μm^2 to 100 x 100 μm^2 were defined. After a standard RCA clean step at 60 °C the gate dielectric was thermally grown in a 2 % O$_2$ in N$_2$ mixture at 900 °C for 20 minutes. The gate metal, Al or Cr, was evaporated to form the top contacts of the MOS structures. The oxide thicknesses, approximately 29 Å for the Cr devices and 26 Å initially for the Al-gated devices, were extracted from C-V measurements following the method of Maserjian [10,11]. Post metallization annealing (PMA) was carried out in air at a temperature of 170 or 260 °C using a hot plate. The device was put on the preheated surface and the bias was applied for the desired time. Afterwards the device was placed on a cold (room temperature) metal plate. This strategy enables relatively short annealing times.

High-frequency (HF) C-V measurements were obtained at room temperature before heating and after cool-down at a frequency of 10 kHz using an HP4284A LCR meter. In Figure 1, HF C-V measurements on n-type Cr-gated and p-type Al-gated devices are shown. The Cr devices show a high initial interface state density (D_{it}) estimated to 2.5x10^{12} cm^{-2}eV^{-1} both at 0.25 eV and 0.85 eV above the valence band edge using data from measurements on p- and n-type devices respectively. For the Al-gated devices the initial peak D_{it} is approximately 1.2x10^{12}
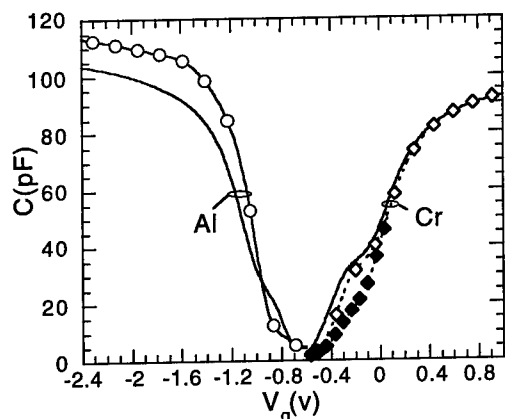
Figure 1) HF (10kHz) *C-V* measurements on *n*-type Cr and *p*-type Al MOS devices prior to PMA(—). Also shown are measurements after a 1000 s PMA at 260 °C without bias (◇) and with an applied bias of -1.0V (◆,O).
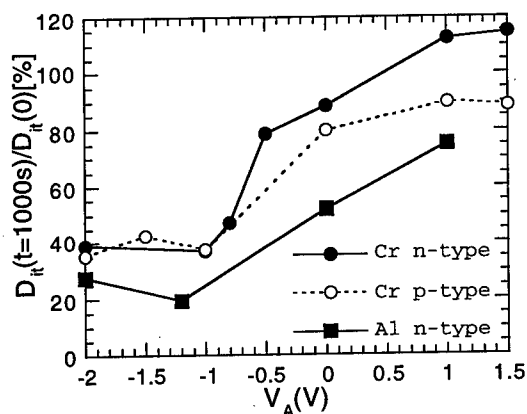

Figure 2) The relative peak D_{it} at E-$E_v \approx 0.25$ eV (*p*-type) or E-$E_v \approx 0.85$ eV (*n*-type) after a 1000 s PMA at 260 °C.

$cm^{-2}eV^{-1}$ and 2.3×10^{12} $cm^{-2}eV^{-1}$ at 0.25 eV and 0.85 eV above the valence band edge (E-E_v) respectively. The D_{it} values were estimated from the HF *C-V* curve by comparing it to a theoretical one. This method relies on the fact that a large fraction of the fast interface states follows the 10 kHz signal [4]. Values in the middle of the bandgap are unavailable since the charges are in non thermal equilibrium in the inversion region due to direct tunneling through the oxide. Evident from Figure 1 is an increase of the oxide capacitance during PMA(O) due to the aforementioned reduction of the oxide, whereas for the Cr devices the capacitance in accumulation is unaltered after PMA (◇).

From Figure 1 it is clear that the passivation rate of the Cr-gated MOS devices is significantly influenced by the bias during anneal. The peak D_{it} is considerably lower after PMA *with* bias (◆) than without (O). In Figure 2 where the relative peak D_{it} at E-$E_v \approx 0.25$ eV (*p*-type) or E-$E_v \approx 0.85$ eV (*n*-type) after a 1000 s PMA at 260 °C is shown, a clear bias dependence between approximately -1.0 and 1.0 V is observed both for Al- and Cr-gated MOS devices.

To show that this is due to the change in charge-state of the defects, the relation between applied bias and the approximate Fermi-level position (E_{fs}-E_v) at the surface was calculated from a *C-V* measurement using the Berglund-integral [12]. The Fermi-level position was found to be in the middle of the band gap at V_A=-0.5 V and close to the valence band edge at V_A=-1.0 V. *If* the charge-state of the defect has an

impact on the passivation rate, a further increase of the negative bias below -1.0 V would not increase the passivation rate: most of the additional bias is lost across the oxide. This is indeed observed in Figure 2 for $|V_A| > 1$. Since the oxide electric field and the tunneling current density vary considerably with bias in the regions where the passivation rate is largely unaffected, neither oxide field nor electron transport through the oxide is capable of explaining the bias dependency. In these devices, the energy levels where the defects change their charge from positive to neutral (+/0) and from neutral to negative (0/-), correspond to approximately -1.0 V and 0.0 V respectively. Since a rapid change of the passivation rate is observed between these two biases the charge-state of the defects is the most likely explanation to the observed difference in passivation kinetics.

Also to be noted in Figure 2 is that at positive biases the passivation rate is *reduced* both for Cr- and Al-gated devices. In fact for *n*-type Cr devices the D_{it} *increases* when annealed at positive biases. This might be used to get information regarding the defect density of the non-passivated Si—SiO₂ interface.

These results are in direct contrast to what Reed and Plummer [13] found from their extensive study on the passivation of (mainly) Al-gated MOS devices with various thicknesses. On the matter of bias dependence they conducted an experiment where the passivation dynamics of MOS devices with either *n*- or *p*-type substrates were compared. No significant differences between the two different types were
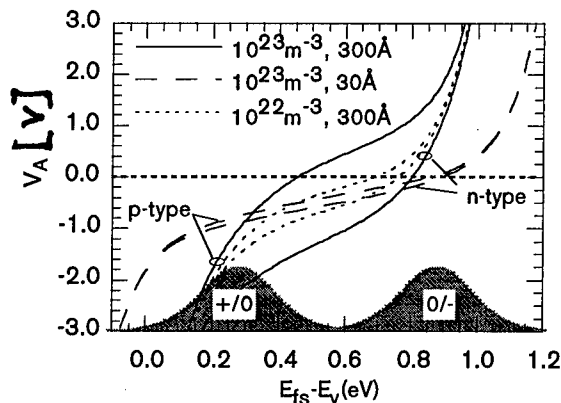
Figure 3)   Applied voltage vs the Fermi-level position at the surface of simulated Al-gated MOS devices with different doping and oxide thicknesses. Also shown is the (+/0) probability to change charge-state at 0.28 eV and (0/-) at 0.88 eV. The temperature was T=220 °C.
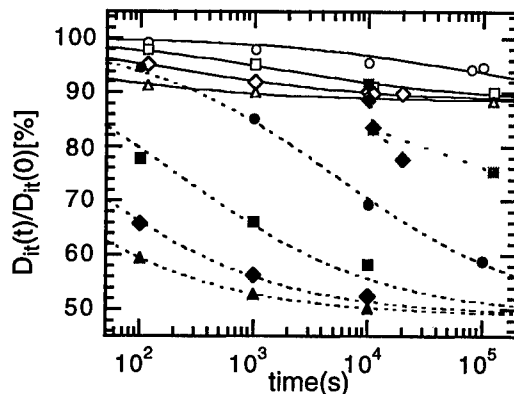

Figure 4)   The decrease of the relative peak D_{it} of normal (open) and biased (-1.0V) (closed) anneal of Cr-gated n-type MOS devices. The temperature was 170 °C (O), 200 °C (□), 230 °C (◇), and 260 °C (△). Also shown is the passivation of previously non-biased devices (◆,■). Symbols are measurements and lines are guides to the eye.

found and they concluded that since the difference in surface potential was sufficient to change the net charge from negative to positive, the passivation is *independent* of the charge-state of the defects.

To explain the apparent discrepancy of these results, Al-gated MOS devices with different doping type and concentrations were simulated. In Figure 3, the relation between applied bias and the Fermi-level position at the surface $V_A(E_{fs}-E_v)$ is shown. In the simulations, a peak P_b concentration of 2×10^{12} cm^{-2}eV^{-1} was used. The energy levels where the P_b centers change their charge-state was 0.28 eV (+/0) and 0.88 eV (0/-) with a Gaussian spread of 0.1 eV. The doping concentration was 10^{22} or 10^{23} m^{-3} and the oxide thickness was 30 or 300 Å. No quantization effects at the surface were taken into account and it was assumed that the charges in the substrate are in thermal equilibrium (i.e. direct tunneling is prohibited), a simplification which is acceptable for medium biases. Also shown in Figure 3 (gray) is the probability for a P_b center to change its charge-state at a given Fermi-level position. The temperature in these simulations was T=220 °C.

A substrate doping type influence on $V_A(E_{fs}-E_v)$ is evident from Figure 3 for *high doping* concentrations and relatively *thick* oxides (10^{23} m^{-3}, 300 Å). At V_A=0.0 V the difference in $E_{fs}-E_v$ is approximately 0.35 eV, and a large fraction of the P_b centers would thus be of different charge-states in the two cases. On the other hand, for lower doping concentrations the

difference between n- and p-type substrates is smaller and at even lower doping concentrations (not shown) it is absent. To be able to measure the influence of the P_b charge-state by means of a change of substrate type, sufficiently high doping concentrations must thus be used. For these simulated devices, a doping concentration of 10^{23} m^{-3} is sufficient for 300 Å thick oxides. However for thinner oxides, a considerably higher doping level is needed. For thicknesses between 200 and 300 Å and doping levels around 1×10^{21} m^{-3} the difference in $E_{fs}-E_v$ is approximately 0.05 eV at 0.0 V. A change of the net charge is thus *not* expected and the passivation kinetics are anticipated to be equal, even if the charge-state *has* an impact. This is also true for "normal" PMA temperatures above about 400 °C, even for thicker oxides and higher doping levels.

The similarity we observe for p- and n-type substrates in Figure 2 is thus consistent with the charge-state dependence for the defect passivation dynamics in these ultrathin oxide MOS devices.

To study the charge-state influence on the passivation kinetics in more detail, in Figure 4 the decrease of the relative peak D_{it} of n-type Cr devices at approximately 0.85 eV above the valence band edge is shown during PMA. The temperature was in the range 170 °C to 260 °C. Some devices were annealed without bias (open symbols) while others were annealed with V_A=-1.0V (closed symbols).

Without bias, the rate of passivation is low and

after 10 000 s more than 90 % of the interface states remain unpassivated even at T=260°C. With V_A=-1.0 V on the other hand only about 50 % of the defects remain unpassivated at 10 000 s.

An additional feature of the passivation kinetics in Figure 2 is that the passivation saturates. This is in contrast to other studies on the passivation of P_b centers in non-metal [14] and Al-gated [3,15,16] structures, in which no saturation of the passivation was noted. To investigate the influence of temperature and bias on the saturation value, a few devices initially annealed without bias at 200 and 230 °C were given an extra anneal *with* bias at -1.0 V. The relative peak D_{it} of these devices is shown in Figure 4 (◆,■), and although the value decreases with time — down to approximately 75 % after 100 000 s — the amount of decrease is substantially smaller compared to that of devices annealed with bias from the start. From this it appears that the concentration of the passivating species, presumably hydrogen, at the reacting surface might be reduced not only by the passivation reaction but also by some other process. One possible H_2 consuming process is out-diffusion of H_2 into the field-oxide or the metal. Another competing hydrogen consuming process would be the passivation of P_{b1} centers [17]. This defect, which is present on (100) surfaces, was recently shown to be electrically inactive [2,18] and would consequently not be detected by the C-V technique used in this study. Further, since the charge-state of the P_{b1} center will be indifferent to the Fermi-level position at the surface (at least between the Si band-edges) the passivation rate of P_{b1} centers is not expected to be sensitive to the bias. The large difference between the saturation levels for biased and non-biased devices could thus be explained by the presence of P_{b1} centers that are equally "slowly" passivated at V_A=-1.0V which results in a larger amount of hydrogen available for the passivation of P_{b0} centers.

In this study we have observed the effects of bias during post metallization anneal of Cr-gated MOS devices. The rate of passivation of P_b centers is found to be faster for positively charged defects and slower for negatively charged defects compared to neutral ones. The bias dependency for *p*- and *n*-type substrates is found to be similar, due to the fact that the relation between applied voltage and Fermi-level position at the surface and thus the charge-state of the defects is close to equal for both substrate types.

1. E. H. Poindexter, Semicond. Sci. Technol., **4**, 961 (1989).
2. A. Stesmans and V. V. Afanas'ev, Phys. Rev. B, **57**(16), 10030 (1998).
3. L.-Å. Ragnarsson, P. Lundgren, Z. Ovuka, and M. O. Andersson, J. Electrochem. Soc., **144**(5), 1866 (1997).
4. M. O. Andersson, A. Lundgren, and P. Lundgren, Phys. Rev. B, **50**(16), 11666 (1994).
5. A. Edwards, Phys. Rev. B, **44**, 1832 (1991).
6. R. S. Bauer, R. Z. Bachrach, and L. J. Brillson, Appl. Phys. Lett., **37**(11), 1006 (1980).
7. R. McGrath, J. F. McGilp, I. T. McGovern, S. J. Morgan, W. G. Herrenden-Harker, and R. H. Williams, Semicond. Sci. Technol., **3**, 937 (1988).
8. L.-Å. Ragnarsson, E. Aderstedt, and P. Lundgren, J. Electrochem. Soc. (to be published).
9. L.-Å. Ragnarsson, E. Aderstedt, and P. Lundgren, presented at Ultrathin SiO_2 and High-K Materials for ULSI Gate Dielectrics, San Francisco, 1999.
10. J. Maserjian, G. Petersson, and C. Svensson, Solid-State Electronics, **17**, 335 (1974).
11. M. H. Hecht, R. P. Vasquez, F. J. Grunthaner, N. Zamani, and J. Maserjian, J. Appl. Phys, **57**(12), 5256 (1985).
12. C. N. Berglund, IEEE Trans. Electron. Devices., **ED-13**, 701 (1966).
13. M. L. Reed and J. D. Plummer, J. Appl. Phys, **63**(12), 5776 (1988).
14. A. Stesmans, Appl. Phys. Lett., **68**(19), 2723 (1996).
15. M. O. Andersson, P. Lundgren, O. Engström, and K. R. Farmer, Microelectronic Engineering, **22**(1-4), 235 (1993).
16. P. Lundgren and M. O. Andersson, J. Appl. Phys, **74**(7), 4780 (1993).
17. A. Stesmans, Appl. Phys. Lett., **68**(15), 2076 (1996).
18. A. Stesmans and V. V. Afanas'ev, J. Phys.: Condens. Matter, **10**, L19 (1998).

Low Temperature-High Pressure Grown Thin Gate Dielectrics for MOS Applications

V. Ramgopal Rao[a], W. Hansch[b], S. Mahapatra[a], D.K.Sharma[a], J.Vasi[a], T.Grabolla[c], and I.Eisele[b]

[a] Indian Institute of Technology, Bombay, Powai, Mumbai- 400076, India

[b] Fakultät für Electrotechnik, Universität der Bundeswehr Munich, D-85577, Neubiberg, Germany

[c] Institute for Semiconductor Physics, Walter-Korsing Strasse 2, D-15230, Frankfurt (Oder), Germany

In this study we propose high pressure grown oxide (HIPOX) as an alternative low-temperature MOS gate insulator and show that it performs excellently in comparison to other widely reported low-temperature deposited oxides. Our optimized process conditions for HIPOX result in high quality gate dielectric comparable in quality to the standard thermal dry oxide in terms of initial properties as well as under various stress conditions. Sub 100 nm channel length vertical MOSFETs with HIPOX as a gate dielectric are fabricated and characterized to demonstrate the suitability of HIPOX as a low-temperature MOS gate dielectric.

1. INTRODUCTION

Reduced thermal budget processing of deep sub-micron MOSFETs and thin film technologies have brought about increased interest in low-temperature oxides. Much of the low-temperature oxide study has however been concentrated on the deposited oxides with a wide variety of annealing conditions, generally a high temperature anneal to improve the quality of the oxide [1]-[5]. In this work we report for the first time high-pressure grown oxide (HIPOX) as an alternative MOS gate dielectric and compare its performance with the deposited oxides. However, one of the main factors limiting VLSI device reliability is the degradation caused by hot carriers. These hot carriers are known to cause interface state generation, electron trapping at the Si/SiO$_2$ interface and also charge trapping in the oxide bulk [6]-[10]. For these reasons, not only must any SiO$_2$ dielectric possess good initial quality, but should also perform well under accelerated stress conditions. For this reason, we have looked at the high-field performance of all these different oxides, namely: low-pressure chemical vapor deposited (LPCVD) oxide, remote plasma enhanced chemical vapor deposited (RPECVD) oxide and finally, the high pressure pyrogenic steam grown oxide (HIPOX) and compared their degradation characteristics with the standard thermally grown dry oxide under different stress conditions. Finally, vertical MOSFETs with channel lengths below 100 nm are fabricated using HIPOX as the gate dielectric to demonstrate the suitability of HIPOX for low-temperature MOS applications.

2. EXPERIMENTAL CONDITIONS

For the initial optimization of all the gate oxides, extensive work has been done on poly-silicon gate MOS capacitors with all the above gate oxides, namely: LPCVD oxide, RPECVD oxide, HIPOX and the thermal dry oxide. The thermal dry oxide is used as the control oxide. The MOS capacitors (of area 0.04x10^{-2} cm^2) used in this study were fabricated on 4" p-type (100) Si substrates using standard polysilicon gate technology. The control SiO$_2$ was thermally grown in O$_2$ at 800° C. LPCVD oxide deposition was done at 420° C using silane and oxygen reaction while for the remote plasma enhanced SiO$_2$, the substrate temperature was 300^0 C and the deposition was done using tetraethylorthosilane (TEOS) and oxygen. High-pressure oxidation was carried out at 650°C in a commercial high-pressure

oxidation reactor in pyrogenic steam at a pressure of 10^5 Pa using O_2. The oxide thickness for all the types of oxides was 10 nm. An anneal at 750° C was given for all the oxides to a total time of 40 minutes in N_2/O_2 as a densification step. This anneal temperature was chosen based on our simulation and experimental results on MOSFETs with channel lengths of 150 nm, which show that oxidation temperatures above 750° C would cause thermal out-diffusion of dopants making the devices dysfunctional [11]. Except for the gate insulator growth process, all other processing conditions were kept the same for all the wafers. The above optimized process conditions for all the oxides are summarized in Table 1 and the measured initial characteristics on poly-silicon gate MOS capacitors are shown in Table 2. The excellent initial properties of all these oxides shows the effectiveness of the process optimization done for these oxides.

For all the high-field stress experiments, the capacitors were stressed by a constant current. The configuration used for constant current stressing consisted of a Keithley 220 current source and a Keithley 617 electrometer, which was used to monitor the voltage across the capacitor. The stress current was interrupted at regular intervals to monitor the degradation. HF-LF CV technique was used for the measurement of midgap interface state density (ΔD_{itm}) and the midgap voltage shift (ΔV_{mg}). Experiments were performed by injecting charge from the metal as well as from the substrate. For the experiments where silicon was the injecting electrode, the devices were exposed to light of sufficient quantity such that the injection was not limited by the rate of generation of minority carriers.

Type of Oxide	Oxidation Parameters
Thermal dry	T=800°C, Dry oxidation in O_2
HIPOX	T=650°C, P= 10 bar
LPCVD Oxide	T=420°C, reactive gases SiH_4 and O_2. Ratio between SiH_4 to O_2 is 1:5. Pressure: 35 Pa
Remote PECVD Oxide	T=300°C, TEOS process

Table 1: Optimized process conditions for various gate dielectrics.

Type of Oxide	V_{fb} (V)	D_{itm} (eV^{-1}. cm^{-2})	Q_{BD} (coul. cm^{-2})	T_{ox} variation across Wafer (%)
Thermal Dry	-0.87	1.3×10^{10}	> 50	2
Remote PECVD oxide	-0.86	1.5×10^{10}	12.5	7
LPCVD oxide	-0.86	1.9×10^{10}	10	10
HIPOX	-0.87	9×10^9	> 50	9

Table 2: Measured characteristics on the virgin poly Si gate MOS capacitors with various dielectrics.

3. ELECTRICAL CHARACTERIZATION

The initial I-V characteristics are shown in Fig. 1 and the breakdown field distribution is shown in Fig. 2 for all the optimized 10nm gate oxide capacitors with different gate dielectrics. Constant current injection experiments were also carried out on all these different gate oxides to assess their hot-carrier reliability. Figure 3 shows the ΔV_{mg} in all these oxides as a function of stress time for two different current densities. As can be seen, deposited oxides show electron trapping while the HIPOX and thermal oxides show conventional hole trapping. The excess Si-H and Si-OH centers in deposited oxides are known to act as electron trap centers [9]. HIPOX, as can be seen, shows shifts comparable to that of a thermal dry oxide.
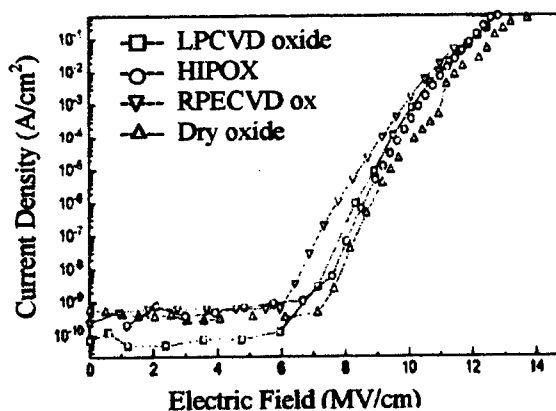
Figure 1: Current density (J) versus Electric field (E) for different oxides in accumulation.
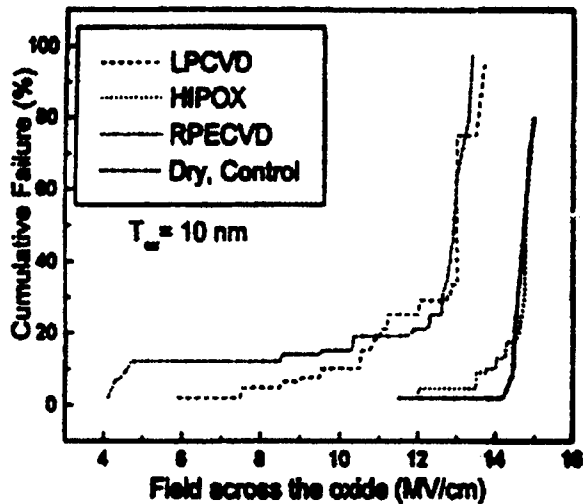
Figure 2: Breakdown field distribution in various gate oxide capacitors.
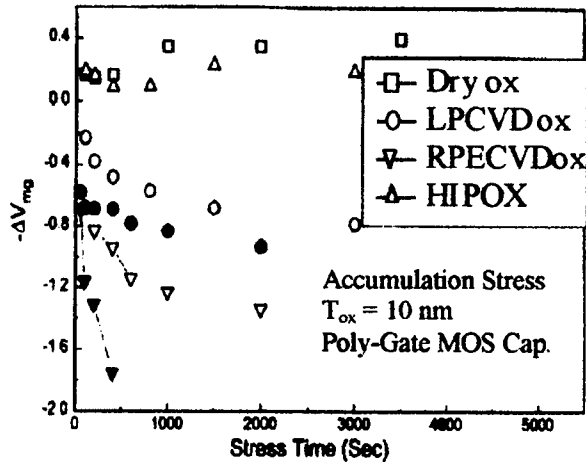
Figure 3: Midgap voltage shifts (ΔV_{mg}) as a function of stress time for various oxide capacitors. Open symbols are for a stress current of 0.25 mA/cm^2 and filled symbols are for corresponding oxides (LPCVD and RPECVD ox) at a stress current of 2.5 mA/cm^2.

The change in gate voltage during constant current stress in all the four types of oxides is shown in Fig 4. The higher trapping rate in RPECVD oxides as observed from the larger change in V_G during the constant current stress is also responsible for the low Q_{BD} values observed in these oxides. We also observe a good correlation between Q_{BD} and ΔV_G in this work. In general,

oxides with lower Q_{BD} show higher ΔV_G during constant current stress. The stress induced mid-gap interface state generation (ΔD_{itm}) is shown in Fig 5. In LPCVD oxides, it can be seen that there is an order of magnitude higher ΔD_{itm} compared to HIPOX for identical stress conditions. RPECVD oxides are much worse compared to even the LPCVD oxides and are therefore not shown.
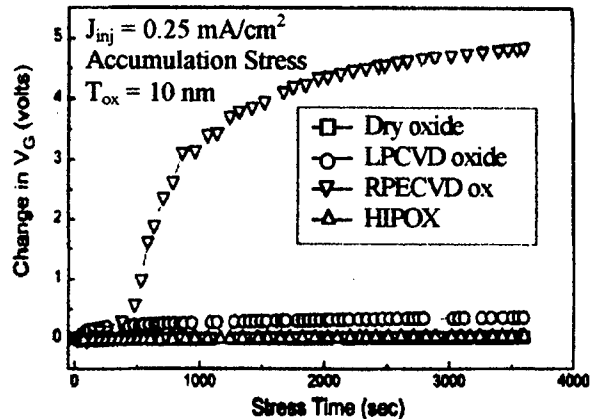
Figure 4: Change in gate voltage (ΔV_G) during constant current injection.
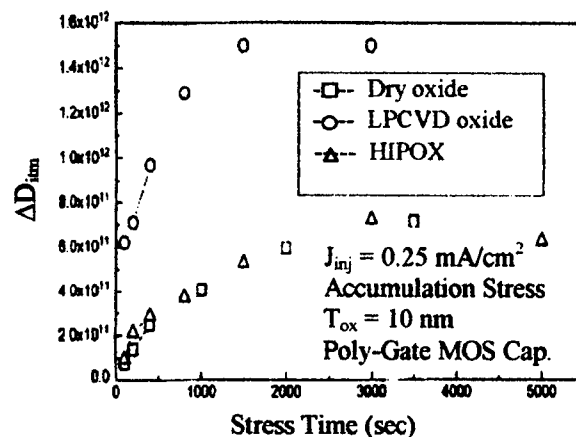
Figure 5: Change in mid-gap interface state density as a function of stress time.

In order to demonstrate the applicability of HIPOX as a MOS gate dielectric, vertical sub 100nm Si MOSFETs are fabricated using HIPOX as the gate dielectric. The schematic of the vertical MOSFET structure with gate oxide surrounding the mesa is shown in Fig. 6 (a) while the fabricated MOSFET structure is shown in Fig. 6 (b). HIPOX is used as a gate dielectric with a thickness of 14nm. HIPOX was grown at 600°C at P=10 bar in

pyrogenic steam without any additional high-temperature anneal. The gate oxide thickness of 14nm was chosen mainly to avoid corner breakdown at the mesa edges. The detailed process sequence for these vertical MOSFETs is reported elsewhere [12]. Measured input characteristics are shown in Fig. 7 [12]. The measured value of low-field mobility in these MOSFETs is $420 cm^2/V$-s indicating a good HIPOX/Si interface. This excellent low-field mobility has also been recently shown to give rise to high velocity overshoot effects in these vertical MOSFETs [13].

4. CONCLUSIONS

High-pressure grown oxide is optimized to exhibit excellent characteristics as a MOS gate dielectric. Vertical sub 100nm channel length MOSFETs employing the optimized low-temperature HIPOX are fabricated and shown to exhibit excellent mobility and I-V characteristics.

ACKNOWLEDGEMENTS
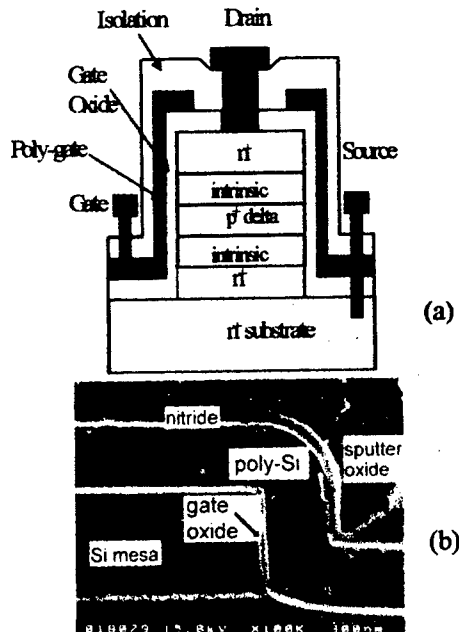
Figure 6: (a) Schematic of a vertical MOSFET with gate oxide surrounding the mesa forming a MOSFET structure. (b) SEM cross-sectional view of the fabricated vertical MOSFET structure.
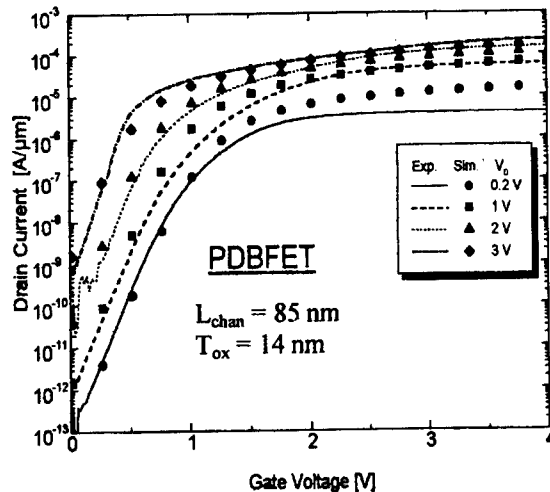
Figure 7: I_d-V_g characteristics for the vertical MOSFET with a channel length of 85 nm by using HIPOX as a gate dielectric.

REFERENCES

1. J.Ahn, W. Ting, and D. L. Kwong, IEEE Electron Device Lett., vol. EDL-13, p. 186, 1992
2. S. Ang and S. Wilson, J. Electrochemical Soc., vol. 134, p. 1254, 1987
3. J. Lee, C. Hegarty, and C. Hu, IEEE Electron Device Lett., vol. EDL-9, p. 324, 1988
4. J. Batey, E. Tierney, and Thao N. Nguyen, IEEE Elec. Dev. Lett, v. 8, p. 148, 1987
5. K.H.Lee, W. H. Liu, S. A. Campbell, and I. Banerjee, JECS, vol. 140, p. 501, 1993
6. R. Rofan and C. Hu, IEEE Electron Device Lett., vol.12, p. 632, 1991
7. D. J. DiMaria, and J. W. Stasiak, J. Appl. Phys., vol. 65, p. 2342, 1989
8. D. A. Buchanan and D. J. DiMaria, J. Appl. Phys., vol. 67, p. 7439, 1990
9. T. Hori, H. Iwasaki, Y. Yoshioka, and S. M. Sato, Appl. Physics Lett., vol. 52, p. 736, 1988
9. V. Ramgopal Rao and J. Vasi, v. 71, p.1029, J. of Appl. Physics, 1992
10. V. Ramgopal Rao, D. K. Sharma, and J. Vasi, IEEE Trans. on Elect. Dev. , September, 1996
11. H.Gossner, Ph.D thesis, 1997,UniBw Muenchen, Germany
12. H.Gossner et al., Jpn.J.Appl.Phys.,v.33, p.2423, 1994
13. V.Ramgopal Rao et al., IEDM Tech Dig., p. 811, 1997

SiGe nMOSFETs with gate oxide grown by low temperature plasma anodisation

L.S.Riley[a], S.Hall[a], J. Harris[b], J. Fernandez[b],B. Gallas[b], A.G.R.Evans[c], J.F. Clarke[c] J. Humphrey[c], R.T.Murray[d] and C. Jeynes[e]

[a]Dept. of Electrical Eng. &Electronics, University of Liverpool, Liverpool, L69 3GJ,e-mail: L.S.Riley@liverpool.ac.uk ,S.Hall@liverpool.ac.uk
[b]IRC for Semiconductor Materials, Imperial College, London.
[c]Dept. of Electronics & Comp. Sci., University of Southampton, [d]Dept. Mat. Sci. Liverpool [e]University of Surrey, Central research facility.

The characterisation of strained surface channel SiGe nMOSFETs with gate oxide grown by low temperature plasma anodisation is described. We report excellent off-state leakage currents and zero field mobilities of 640 and 420cm^2/Vs for Si and SiGe devices respectively. Reduction in SiGe mobility is primarily attributed to increased interface state scattering in the absence of any increased surface roughness as observed in high resolution microscopy. The oxide compositon is mixed and conservation of crystallinity of SiGe during oxidation is observed via RBS analysis.

1. INTRODUCTION

Plasma anodisation can be undertaken at greatly reduced temperature due to the use in the oxidation environment of highly reactive ions and the abundance of electrons. It has been shown to be a good candidate for the use in SiGe technology to preserve the integrity of the strained layers[1]. In contrast to standard thermal oxidation of SiGe, Ge is incorporated into the growing oxide layer as GeO_2 with the exception of a small percentage of Ge residing as electron trapping centres in the oxide[2]. We report here results of Si and SiGe nMOSFETs fabricated with gate oxide realised by plasma anodisation with a SiGe surface channel. Low off-state leakage and effective electron mobilities about 30% lower than those from Si controls were measured. The reduced mobility is likely to result from the higher density of interface states at the oxide:SiGe interface. This is supported by an increase in the inverse subthreshold slope parameter of the SiGe transfer characteristics and is consistent with results from MOS capacitors[1].

2. FABRICATION

A relatively simple self-aligned process was devised to realise SiGe MOSFETs. The epitaxial structure was grown by gas-source MBE at 650°C. A thick (500nm) Si buffer layer doped with boron to 5x10^{16}cm^{-3} was grown on a low resistivity n-type 100mm Si <100> wafer, followed by a thin (10nm) spacer layer to reduce any dopant chamber memory effects. A 25nm strained, undoped $Si_{0.85}Ge_{0.15}$ layer was then grown followed by a 13nm sacrificial Si capping layer. The above structure is shown schematically in Figure 1. The Si control samples were produced on standard 1-5Ωcm <100> orientated p-type wafers.

| 13 nm i-Si |
| 25 nm i-SiGe |
| 10nm i-Si |
| 500 nm p-Si. |
| |
| n+ Si substrate |
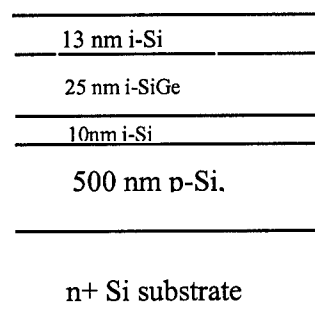
Figure 1. Schematic structure of GS-MBE layers for SiGe nMOSFET fabrication.

Pattern Active Area
LTO mask, KOH anisotropic mesa etch
Remove mask, RCA clean
Plasma anodisation for gate oxide
Pattern a-Si gates (550°C)
Pattern Source/Drain
Self aligned As implantation
Pattern substrate implants BF$_2$
RTA 750°C 30s
LTO deposition.
Contacts windows and back strip
Metallisation.
Forming Gas sinter (420°C)
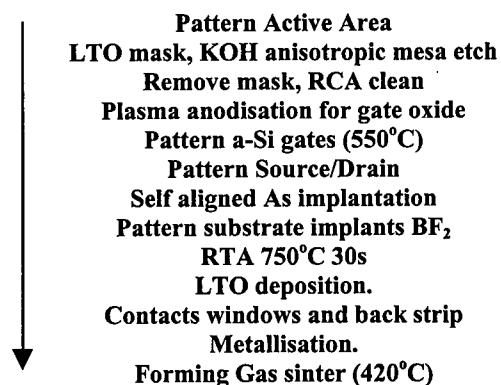
Figure 2. Simplified process flow of test MOSFET fabrication.

Figure 2 shows the process flow leading to the schematic of Figure 3. Great care was taken with the thermal budget to avoid relaxation of the SiGe.
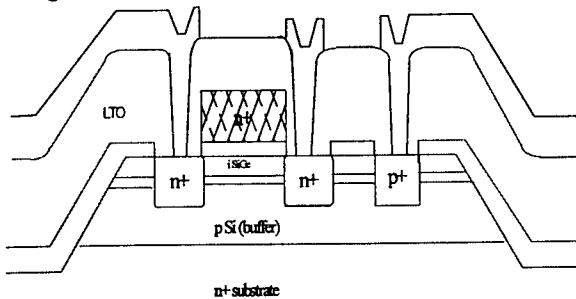
Figure 3. Cross-sectional schematic of final device

3. RESULTS

Subthreshold characteristics of both the Si and SiGe MOSFETs are shown in Figure 4. Excellent off-state leakage currents of circa 1pA were observed indicating high quality layers and good S/D junctions formed by implantation. Inverse subthreshold slopes of 82mV/decade and 200mV/decade were calculated for the respective devices.
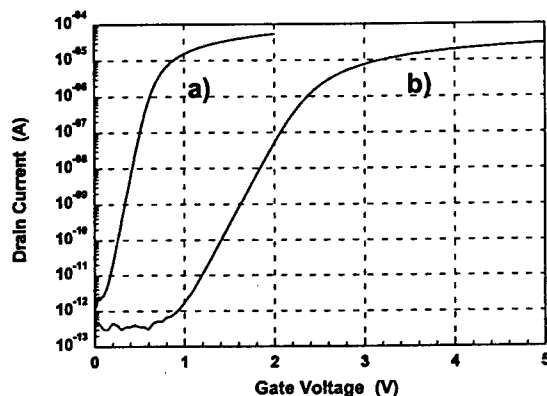
Figure 4. Subthreshold plot for (W/L) 20/1μm nMOSFETs a) Si and b) SiGe, S= 82mV/dec and 200mV/dec respectively.

The increased value for the SiGe was assumed to arise from increased surface state capacitance. The corresponding interface state density, as derived from the subthreshold regime was ~1x10^{12}cm^{-2}eV^{-1} (midgap to E_v). This value is in agreement with typical values observed in M:O:SiGe capacitors formed by plasma anodisation[1]. The shift of the SiGe transfer characteristic to the right is representative of negative fixed oxide charge, estimated as 8.5x10^{11}cm^{-2} from the 1V shift from the Si control case. This has been attributed to an unsaturated O bond at the interface in other SiGe oxide

studies[5]. A standard forming gas anneal only has been performed at this stage. Further treatments have been shown to be effective in reducing the interface state density for oxides produced in similar growth systems[6]. The very low off-state leakage currents attest to the good quality of the epi-layers and the effectiveness of the low temperature schedules employed in fabrication. To investigate the junctions further, subthreshold plots with Vd varied from 0.5 to 5V are shown in figure 5a and b for Si and SiGe 100/20μm devices respectively.
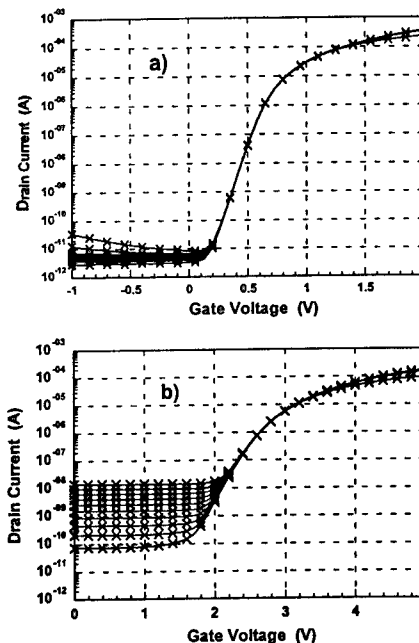
Figure 5. Subthreshold plot of 100/20μm FETs as function of drain voltage Vd = 0.5 to 5V in steps of 0.5V, a) Si control , b) SiGe nMOSFET

The increase in off state current with drain voltage seen in the SiGe case suggests field enhanced generation via traps in the epitaxial layers encompased by the drain depletion region. Measured activation energies were 0.51eV and 0.37eV for the low and high field cases respectively. In contrast, the activation energy of the Si device leakage current was 0.73eV showing minimal field dependance, (between 0 and 5V). The differences in activation energies are representative of the quality of the epitaxial material in comparison to bulk Si.

Figure 6 shows a typical output characteristic of a long channel SiGe device. The relatively low current levels are due to high series resistance most likely due to incomplete activation of the source/drain implants and long junction to channel distances. Junction breakdown typically occurred at approximately 7V.

The derived effective mobilities of both SiGe and Si-control devices are plotted in figure 7 as a function of

V_{gs}-V_t. Mobilities were calculated using dra[i]
conductance after compensating for the effect of seri[es]
resistance at low (0.1V) drain voltage. Estimates of tl[e]
zero field transverse effective electron mobilities of 6[4]
and 420cm^2/Vs are observed for Si and SiGe MOSFE[T]
respectively.
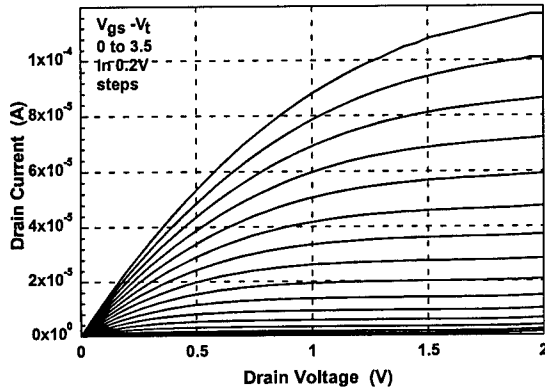
Figure 6. Output Characteristic of 18/6μm SiGe nMOSFET V_{gs}-V_t =0 – 3.5V in steps of 0.2V

The reduced value for the SiGe device is attributed to increased scattering due to the higher density of interface states. Interface roughness could also be a factor; however, figure 8 shows a high resolution cross-sectional electron micrograph of a plasma oxide on SiGe revealing that the interface is very abrupt and well defined and not significantly different from the plasma oxide on Si case (not shown here).
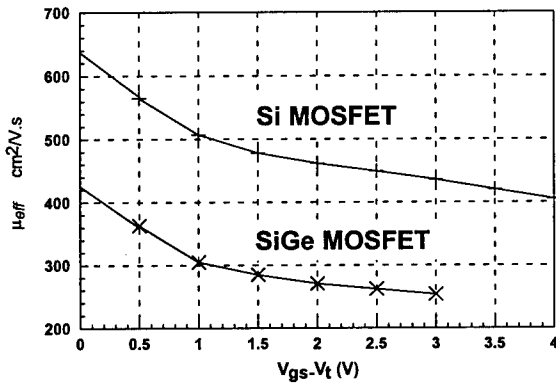
Figure 7. Effective mobilites of Si and SiGe MOSFETs as a function of V_{gs}-V_t exptrapolated zero field values of 640 and 420cm^2/Vs respectively.

The theoretical transverse mobility of electrons in SiGe has been predicted to be lower than in Si due to enhanced alloy scattering[7] and so the reduced mobility measured could be a combination of alloy and interface scattering.

5nm
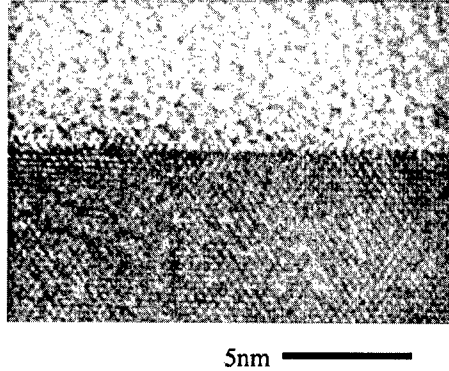
Figure 8. High resolution TEM image of SiGe (bottom layer) interface with mixed oxide.

Turning now to the composition of the oxide grown on SiGe by plasma anodisation, it has been found from a study using XPS[2] that the oxide is mixed, comprising of SiO_2 and GeO_2 with the exception of a small percentage of unoxidised Ge species whose concentration agrees with that found in earlier avalanche injection studies[1]. A complementary study using Rutherford Backscattering Spectrometry on the same sample agrees with the above findings within fitting uncertainties. Simulated annealing analysis of RBS data[8] was performed giving agreement with XPS results showing a mixed oxide consisting of on average 19at.% of Ge and an underlying SiGe layer of approximately 21at.% Ge concentration.
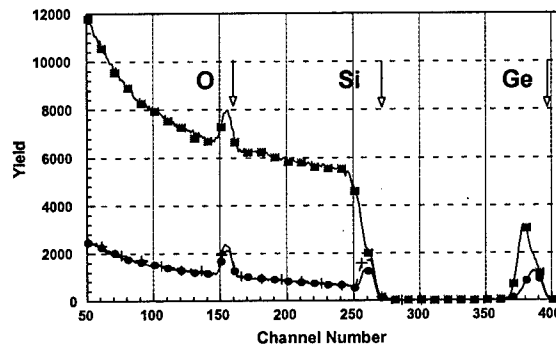
Figure 9. RBS spectra of plasma oxidised SiGe:Si layer, ■ non-channelling spectra, ● channelling spectra and + channelling spectra of plasma oxidised Si control sample.

The RBS spectra identify the existence of Ge in the oxide and from comparison with channelling spectra, crystalline SiGe under the gate as indicated by the reduction of the Ge peak as a consequence of channelling.

The differences of pure Si and SiGe plasma oxide interfaces are now investigated. A sample was grown by gas source MBE at 650^0C and comprised of a Si buffer/100nm strained $Si_{0.84}Ge_{0.16}$ /8nm Si cap. All layers were doped with boron to $10^{17}cm^{-3}$. Plasma oxides were then grown at 100^0C to thicknesses to consume the Si and SiGe layer by varying degrees. MOS capacitors were formed by shadow evaporation of Al followed by a post metal anneal (PMA) in H, at 450^0C for 30 mins. Quasi-static C-V measurements for MOS-Cs with increasing oxide thickness are shown in Figure 10a.

The first and second plots (t_{ox} ~19, 20nm) are judged to correspond to the oxide/semiconductor interface moving from inside the cap to within the SiGe layer. Subsequent plots correspond to consumption of SiGe by the oxide and are characterised by a peak in the fast state distribution, D_{itmax}. (QS plots before PMA showed very high $D_{it} > 10^{13}$ $cm^{-2}eV^{-1}$ and were difficult to analyse). Figure 11 shows the variation of D_{itmax} - D_{itmin} as a function of oxide thickness (and thickness of Ge consumed) where D_{itmin} is the minimum density of the oxide:Si interface; excellent linearity is observed.

indicates a shift in energy of the peak, towards the conduction band with increasing oxide thickness (growth time). This could reflect a monotonic change in the compressive strain as the oxide interface ventures into the SiGe layer. Thicker plasma grown oxides of SiGe have shown fast state peaks at $\sim E_V + 0.75$ eV [1]. It could also be argued that the Ge profile is not abrupt and figure 11 is explained by an increased Ge % with depth.
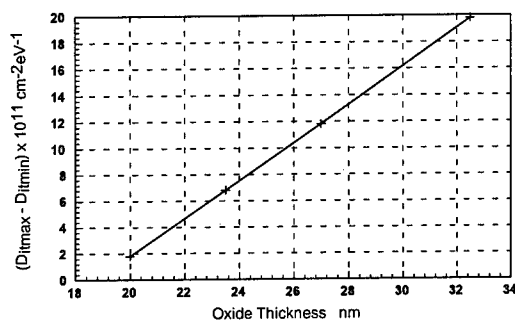
Figure 11. D_{itmax}- D_{itmin} as a function of oxide thickness

4. CONCLUSIONS

SiGe surface channel nMOSFETs have been fabricated using a low temperature process with gate oxides formed by plasma anodisation. Excellent off-state drain leakage and well-behaved characteristics have been observed. The influence on the interface state density is likely to be the primary cause of the degradation of the subthreshold characteristic and mobility in comparison to Si control devices. Material analysis has identified the nature of the oxide and interface showing a mixed composition and formation of an interface state peak and negative oxide charge as SiGe is oxidised.

Figure 10. a) QS-CVs of plasma oxides as the oxide front moves further into the SiGe. and b) corresponding Dit as a function of surface potential. + 19.0nm MOSi:SiGe:Si Capacitor, x 20.0nm, ■ 23.5nm, ●27.0nm and □32.5nm oxide thickness.

Two possible explanations for the D_{it} peak and its linear dependence are plasma/radiation damage during the oxidation (linear growth regime) giving rise to Si and oxygen (plus Ge) dangling oxygen bonds which could also be responsible for excess fixed charge. The C-V plots indicate a switch in the measured fixed oxide charge from positive to negative with Ge consumption, in agreement with analysis of figure 4. Figure 10b) also

REFERENCES

1. I.S.Goh, J.F. Zhang, S.Hall, W. Eccleston and K.Werner, Semicond. Sci. and Tech., **10**, p818 (1995)
2. L.S.Riley, S.Hall, J.of Appl. Phys, **85** (9) p6828 (1999)
3. S.Taylor, J.F. Zhang and W. Eccleston, Semicond. Sci. and Tech, **8**, p1426 (1993).
4. L.S.Riley and S. Hall to be published
5. P.W. Li and E.S. Yang, Appl. Phys. Lett. **63**, p2938 (1993).
6. D.Tchikatilov, Y.F.Yang and E.S.Yang , Appl. Phys.Lett, **69** (17), 21 October 1996.
7. L.E.Kay, T.-W.Tang, J.of Appl. Phys, **70** (3) p1483 (1991)
8. N.P.Barradas, C. Jeynes, R.P. Webb Appl. Phys. Lett. **71** p291-293 (1997).

Ion beam synthesis of narrow Ge nanocluster bands in thin SiO$_2$ films

J. von Borany[a1], K.-H. Heinig[a], R. Grötzschel[a], M. Klimenkov[a], M. Strobel[a], K.-H. Stegemann[b], and H.-J. Thees[b]

[a] Forschungszentrum Rossendorf, Institut für Ionenstrahlphysik und Materialforschung,
P.O. Box 510 119, D-01314 Dresden, Germany

[b] Zentrum Mikroelektronik Dresden GmbH, Grenzstr. 28, D-01109 Dresden, Germany

This paper reports on self-organization of narrow bands of Ge nanoclusters in thin thermally grown SiO$_2$ layers by means of ion beam synthesis. Although the implanted Ge profile is distributed over almost the whole SiO$_2$, a δ-like nanocluster band very close to, but well separated from the Si/SiO$_2$ interface is formed under specific implantation and annealing conditions. The evolution of this band can be explained by a model taking into account collisional ion beam mixing and reactions near the Si/SiO$_2$ interface, which is in good agreement with the experimental results.

1. INTRODUCTION

One of the most promising approaches towards new structures suitable for future microelectronics is the nanocrystal memory, which recently has been proposed by the IBM group [1]. This memory is based on the effect that electrons can be reliably stored even at room temperature in small semiconductor quantum dots ($d_{nc} \leq 10$ nm, $C \approx 10$ aF) embedded within the gate oxide of the field effect transistor. The charge exchange between the cluster and the substrate occurs via direct tunneling.

The main technological challenge consists in the fabrication of a band of small nanoclusters (mean size ≈ 5 nm) with a density of about 10^{12} cm^{-2} very close (≤ 5nm) to the Si/SiO$_2$ interface. CVD and PVD techniques have been applied to form Si or Ge clusters on ultrathin (~ 3 nm) thermally grown SiO$_2$ layers [3, 4]. Alternatively, it has been shown that ion beam synthesis (IBS) is a well suited method to fabricate semiconductor nanoclusters in SiO$_2$ layers [4]. The IBS process for memory applications is performed by ion implantation

of Si or Ge of about 10^{16}cm^{-2} (exceeding the solubility limit in the SiO$_2$ matrix) followed by subsequent annealing in order to form clusters. Following this route Normand et al. [5] applied very-low energy (1-5 keV), high-fluence (> 10^{16}cm^{-2}) Si implantation into 11 nm thin SiO$_2$ films. After annealing they found a band of clusters around the position of the implantation peak.

In this paper we propose another ion beam based approach for the formation of Ge nanoclusters in thin SiO$_2$ layers. The ion implantation is performed at larger ion energies and lower fluences than in ref. [5], which is more comfortable for industrial purposes. Under special implantation conditions we found a very sharp (δ-like) near-interface nanocluster band, which position differs to the implantation profile maximum. The formation of the cluster band as a result of a self-organising process is explained by a model taking into account ion beam interaction in the SiO$_2$ matrix and interface reactions.

---

[1] Corresponding author: Phone: +49 351 260 3378 ; Fax: +49 351 260 3285 ; e-mail: j.v.borany@fz-rossendorf.de

2. EXRERIMENTAL

Thin SiO_2 films of 30 nm thickness have been prepared on (100) Si wafers by thermal oxidation. Afterwards, the wafers were implanted with [74]Ge ions at 12 or 20 keV to vary the position of the implantation profile with respect to the Si/SiO_2 interface. The fluence ranged from $3\text{-}5 \times 10^{15} cm^{-2}$, which results in a Ge peak concentration of about 3 at.%. After a cleaning procedure identically prepared samples were annealed using rapid thermal (950 °C, 30 sec, N_2) or furnace (900°C, 30 min, N_2) annealing (RTA or FA), respectively.

Transmission electron microscopy (TEM) using a Philips CM300-TEM has been performed at cross sectional prepared specimens to investigate the nanocluster formation in the SiO_2 films. The detection limit of the cluster size is about 2 nm. In addition, Rutherford backscattering spectrometry (RBS) with [4]He ions of 1.7 MeV was applied to measure the Ge distribution in the SiO_2 layer. Using an incidence of 70° perpendicular to the surface a depth resolution of < 2 nm could be realized.

3. RESULTS AND DISCUSSION

Fig. 1 shows the Ge depth distribution for an implantation of 20 keV / 5×10^{15} cm^{-2} for the as-implanted state and after different thermal treatments. The as-implanted profile is characterised by a single peak located at 16 nm, which corresponds to the expected value from TRIM-calculations (R_p=18 nm). However, after RTA processing the situation has completely changed. The Ge profile is now characterised by a double peak distribution, with a peak at about 12 nm and a second Ge peak, which is very close to the Si/SiO_2 interface. After FA at 900°C for 30 min a broad Ge distribution with only a weak Ge decoration of the Si/SiO_2 can be observed.

In TEM micrographs no nanoclusters can be detected in the as-implanted state. After RTA
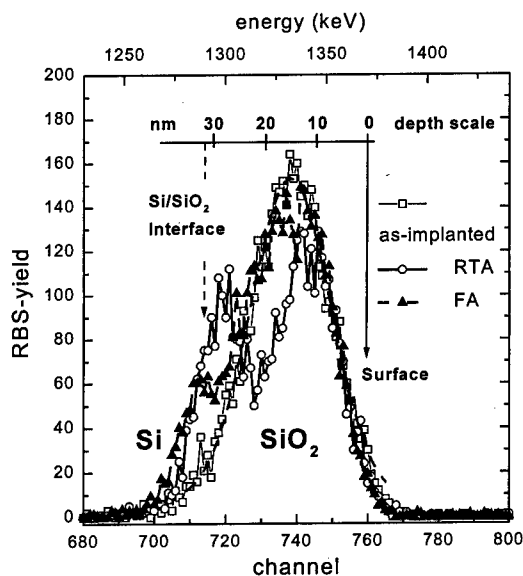
Fig. 1: RBS-spectra showing the Ge depth distribution after implantation ([74]Ge / 20 keV / 5×10^{15} cm^{-2}) in 30 (32) nm SiO_2 films for the as-implanted state and after RTA (950°C, 30 sec) or FA (900°C, 30 min), respectively.

processing we found two different cluster bands, which are separated by a zone free of clusters (Fig. 2). A relatively broad "bulk" cluster band of about 10-12 nm width is located around the peak position of the initial implantation profile (slightly shifted towards the surface). The clusters have a mean size of $(3,5\pm1,0)$ nm and the cluster density can be estimated to about $8 \times 10^{11} cm^{-2}$. The second cluster band is located in a distance of only 2-3 nm from the Si/SiO_2 interface. This cluster band is very narrow (δ-like) and consists of a plane of single, clearly separated nanoclusters parallel to the Si/SiO_2 interface. The mean size of these spherical clusters is slightly smaller $(3,0 \pm1,0)$ nm as compared to the bulk clusters. The cluster density within this band is about $5 \times 10^{11} cm^{-2}$ ($\pm50\%$). All clusters have been found to be in the amorphous state. The position of the cluster bands corresponds to the Ge distribution measured by RBS. That means, that the implanted Ge is mainly within the clusters.
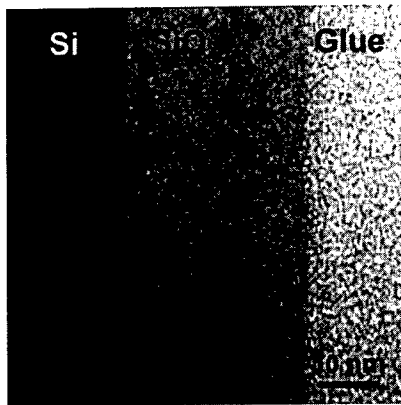
Fig.2 TEM micrograph showing ion beam synthesised Ge (or oxidised Ge) nanoclusters in a 32 nm SiO₂ film after implantation and annealing at 950°C, 30 sec, N₂. The different cluster bands in the bulk and very close to the Si/SiO₂ can be clearly distinguished.

In the case of FA (900°C, 30 min) TEM micrographs show no evidence for Ge clusters. In particular, as expected from the RBS spectrum (Fig. 1) no near-interface cluster band is observed. A careful analysis of the RBS spectrum shows, that a small amount of Ge is now incorporated into the first monolayers of the Si substrate. Additional investigations (not presented here) give rise to the conclusion, that nearly all Ge is now in the oxidized state (GeO$_x$), where the oxygen comes from residual moisture present even in the high purity (5.0) annealing atmosphere [6].

For comparison the experiment has been repeated at an implantation energy of 12 keV (R_p= 12 nm). No significant Ge redistribution with respect to the as-implanted profile can be established by RBS for all annealing procedures. TEM analysis reveal Ge nanocluster formation only as a bulk cluster band around the peak of the implantation profile after RTA processing. The size and density of the clusters are similar to the values mentioned above. No near-interface cluster band was found.

4. MODEL FOR THE NEAR-INTERFACE CLUSTER BAND

It has been found by several experiments [7, 8] that under specific implantation conditions the formation of an interface nanocluster band can be observed independent of the implanted species (Ge, Sn, Sb etc.) or the thickness of the SiO₂ layer. Therefore, this effect is assumed to reflect intrinsic mechanisms of the ion implantation process.
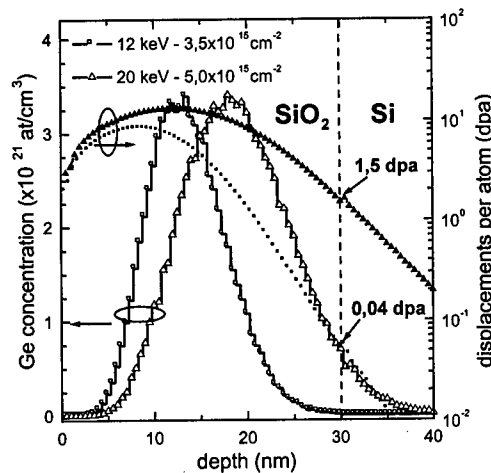
Fig. 3 Ge distribution (left) and number of displacements (right) of 12 and 20 keV ^{74}Ge implantation into 30 nm SiO₂ films calculated by the TRIM code (TRIM-96).

Our model of the formation of a near-interface cluster band is based on the dynamics of irradiation defects. TRIM calculations (see Fig. 3) show for the 20 keV implantation an average value of 1.5 dpa at the SiO₂/Si interface. This implies, that due to collisional mixing (mainly by the recoil atoms) every atom of the SiO₂ network is displaced at least once, which means a dissociation of the SiO₂ network into the elemental components Si and O. Whereas within the SiO₂ layer displaced Si and O recombine, the SiO₂/Si interface acts as a strong sink mainly for oxygen due to its relatively large diffusion coefficient during irradiation. As a result the SiO₂ region close to the Si substrate becomes depleted from oxygen or, conversely, Si accumulates above the stoichiometric value. This phenomenon has been pro-

Si   **SiO₂**   **Surface**
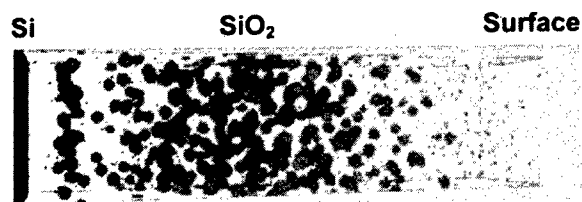
Fig. 4 Result of the 3D kinetic MC-simulation of the evolution of an interface nanocluster band

ven by STM-EDX measurements at appropriately irradiated 500 nm thick SiO$_2$ layers on Si [9]. Rate-equation studies assuming realistic diffusion coefficients reveal, that the maximum value of Si excess is realised in a distance of about 3-4 nm from the Si/SiO$_2$ interface. During subsequent thermal treatment the excess Si atoms form precipitates onto which Ge diffusing from the implanted region condensed.

A corresponding kinetic 3D lattice Monte Carlo (MC) simulation [10] starting with a Gaussian-like Ge profile and taking into account the excess Si close to the SiO$_2$/Si interface qualitatively describes the experimental results (Fig. 4). During the simulated annealing a near-interface band of Ge nanoclusters (containing Si) and additional Ge clusters around Rp are formed, which are separated by a zone nearly free of clusters.

Since in the case of the 12 keV implantation almost no dpa are created close to the SiO$_2$/Si interface (see Fig. 3), no accumulation of excess Si is obtained. As the Si precipitates are a prerequisite of our model, no near-interface Ge nanoclusters can be expected in this case.

5. CONCLUSIONS

The presented model gives an explanation for the IBS based formation of δ-like nanocluster bands near the Si/SiO$_2$ interface due to a self-organising process, which agree with the experimental results. It should be pointed out, that this nanocluster band is metastable. Due to evaporation of Ge from the cluster surface

or (in most cases) due to chemical reactions with residual contaminations (oxygen and/or hydrogen) within the annealing atmosphere the cluster band may disappear, especially at higher temperatures (> 900°C) and/or longer annealing times (> 15 min) [6]. Much care has to be spent to realise reproducibly clean and dry annealing conditions.

REFERENCES

[1] S. Tiwari, F. Rana, K. Chan, L. Shi, H. Hanafi, Appl. Phys. Lett. 69 (1996) 1232.

[2] A. Kohno, H. Marakami, M. Ikeda, S. Miyazaki, M. Hirose, Ext. Abstract of the Intern. Conf. on Solid State Devices and Materials (Hamamatsu 1997), 566

[3] H. Seifarth, R. Grötzschel, A. Markwitz, W. Matz, P. Nitzsche, L. Rebohle, Thin Solid Films 330 (1998) 202

[4] H.A. Atwater, K.V. Shcheglov, S.S. Wong, K.J. Vahala, R.C. Flagan, M.L. Brongersma, A. Polman, Mat. Res. Soc. Symp. Proc. 316 (1994) 409

[5] P. Normand, D. Tsoukalas, E. Kapetanakis, J.A. van den Berg, D.G. Armour, J. Stoemenos, Microelectr. Engineering. 16 (1997) 79

[6] K.H. Heinig, B. Schmidt, A. Markwitz, R. Grötzschel, M. Strobel, S. Oswald, Nucl. Instr. Meth. B148 (1999) 969

[7] A. Nakajima, T. Futatsugi, H. Nakao, T. Usuki, N. Horiguchi, N. Yokayama, J. Appl. Phys. 84 (3) (1998) 1316

[8] A. Markwitz, L. Rebohle, H. Hofmeister, W. Skorupa, Nucl. Instr. Meth. B147 (1999) 361

[9] V. Klimenkov, K.-H. Heinig, unpublished

[10] M. Strobel, K.-H. Heinig, W. Möller, A. Meldrum, D.S. Zhou, C.W. White, R.A. Zuhr, Nucl. Instr. Meth. B147 (1999) 343

Characteristics of Novel Polysilicon Oxide by Anodic Oxidation

C. F. Yeh[a], J. S. Liu[b], and M. C. Chiang[b]

[a]Department of Electronics Engineering & Institute of Electronics, National Chiao-Tung University, 1001 Ta-Hsueh Road, Hsinchu, Taiwan, R.O.C.

[b]National Nano Device Laboratories,
1001-1 Ta-Hsueh Road, Hsinchu 30050, Taiwan, R.O.C.

Abstract

For nonvolatile memory applications, a novel oxide grown on polysilicon by anodic oxidation (anodic polyoxide) is first investigated. In this work, the electrical characteristics of anodic polyoxide is discussed and compared with the conventional thermal polyoxide. The results show that the anodic polyoxide exhibits considerably excellent characteristics, i.e., low leakage current, high breakdown electric field, and high reliabilities.

1. Introduction

For nonvolatile memory applications, the scaling down of device geometry requires high quality thin polyoxide in order to keep a high gate-coupling ratio and improve the data retention characteristics. However, the polyoxide films formed by high temperature oxidation still suffer from high leakage current and low dielectric strength, compared with the oxide grown on single-crystalline silicon [1-3]. These can be attributed to the non-uniform thickness of the polyoxide film and the asperity at polysilicon/polyoxide interface, which can lead to a serious enhancement of localized electric field [4-5]. In this work, anodic oxidation of polysilicon is first proposed, and anodic polyoxide is investigated and compared with conventional polyoxide prepared by furnace.

2. Experiment

40 nm-thick oxide was prepared on P-type wafers with thermal oxidation. Then, 300-nm thick in-situ doped polysilicon layer (poly1) with sheet resistance of 35 / was prepared at 550 . Anodic polyoxide was grown in 0.01wt% citric acid solution at 35 V for 3 min. For comparison, polyoxide grown by furnace at 850 was also prepared to serve as control sample (thermal polyoxide). Subsequently, a second layer of in-situ doped polysilicon (poly2) of 300-nm thick was deposited. Some anodic polyoxide samples received a post-poly2 annealing with RTP at 700 for 30 sec in N_2 ambient (i.e., post-anneal samples). After defining poly2, all samples received a 500nm-thick oxide via PECVD as a passivation oxide. Contact holes were then opened, and 500nm-thick TiN/Al/TiN/Ti electrode was deposited and patterned. As shown in Fig. 1, poly2/polyoxide/poly1 capacitors were completed. Finally, all samples were sintered at 350 for 20 min in N_2 ambient. From C-V measurement, the thickness of anodic and thermal polyoxide is 7.0 nm.

3. Results and Discussion

Figure 2 (a) and 2 (b) compare the current density vs electric field (J-E) characteristics under positive (+Vg) and negative (-Vg) bias for as-grown, post-annealed anodic-polyoxide samples and thermal polyoxide sample. The
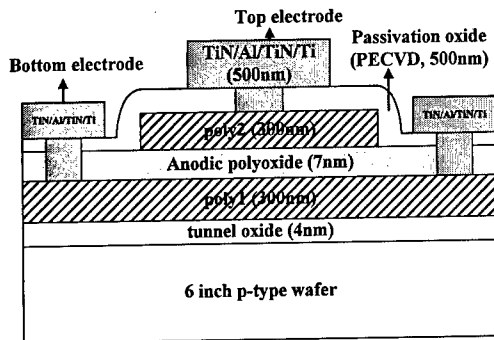
Fig. 1. Cross-sectional view of n^+-polysilicon/ polyoxide/n^+-polysilicon capacitors.
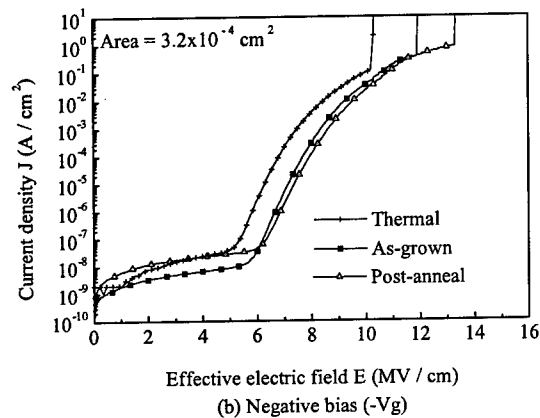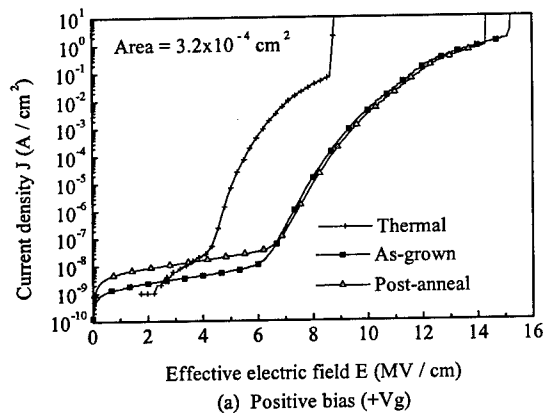
Fig. 2. J-E characteristics of anodic polyoxide compared with thermal polyoxide under (a) positive bias (+Vg) and (b) negative bias (-Vg).

anodic polyoxide depicts smaller leakage current and higher onset of tunneling field than those of thermal polyoxide for both bias polarities. In addition, for anodic polyoxide the phenomenon that the onset tunneling field under bottom injection (+Vg) is lower than that under top injection (-Vg) does not appear. The higher onset of tunneling fields indicate that the enhancement of localized electric field is less severe. It implies that for anodic polyoxide the roughness at polysilicon/polyoxide interface has been reduced, compared to that of thermal polyoxide. In fact, atomic force microscopy (AFM) has verified that the roughness at polysilicon/anodic polyoxide is indeed better than that for thermal polyoxide.

Furthermore, anodic polyoxide can sustain a much higher current density near dielectric breakdown. The effective electron barrier heights under positive bias for thermal, as-grown, and post-anneal samples are 2.1, 2.7, and 2.7 ev, respectively; while those under negative bias are 2.3, 2.6, and 2.6 ev, respectively.

Figure 3 (a) and 3 (b) compare the Weibull breakdown field (E_{bd}) plots for as-grown, post-annealed anodic-polyoxide samples and thermal polyoxide sample under positive and negative biases, respectively. For thermal, as-grown, and post-anneal samples, the breakdown electric fields at 50% cumulative failure under positive bias are 9.1, 15.2, and 15.4 MV/cm, respectively; while under negative bias, they are 10.2, 13.9, and 14.4 MV/cm, respectively. Clearly, anodic polyoxide samples show higher breakdown field than thermal polyoxide sample under both bias polarities. This results show the anodic polyoxide has more densified film quality and more stable Si-O bond structure than thermal polyoxide.

(a) Positive bias (+Vg)
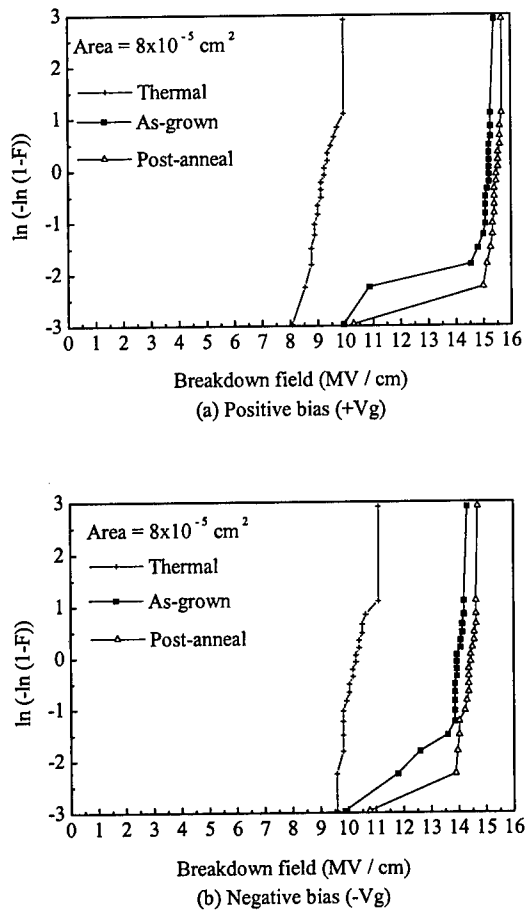
(b) Negative bias (-Vg)

Fig. 3. Weibull breakdown field plots of anodic polyoxide compared with thermal polyoxide under (a) positive bias and (b) negative bias.

It is well known that the higher quality oxide the less electron trapping. The Weibull charge-to-breakdown (Q_{bd}) plots under +40 mA/cm^2 and -40 mA/cm^2 stress are shown in Figure 4 (a) and 4 (b), respectively. For thermal, as-grown, and post-anneal samples, the Q_{bd} at 50% cumulative failure under +40 mA/cm^2 stress are 0.4, 2.1, and 2.4 C/cm^2, respectively; while under -40 mA/cm^2 stress, they are 0.3, 1.6, and 2.4 C/cm^2, respectively. Anodic polyoxide samples show larger Q_{bd} than thermal polyoxide sample under both current stresses. The anodic

polyoxide samples indeed suffer from less electron trapping and thus have larger Q_{bd} than thermal polyoxide sample.

(a) +40 mA/cm^2 stress
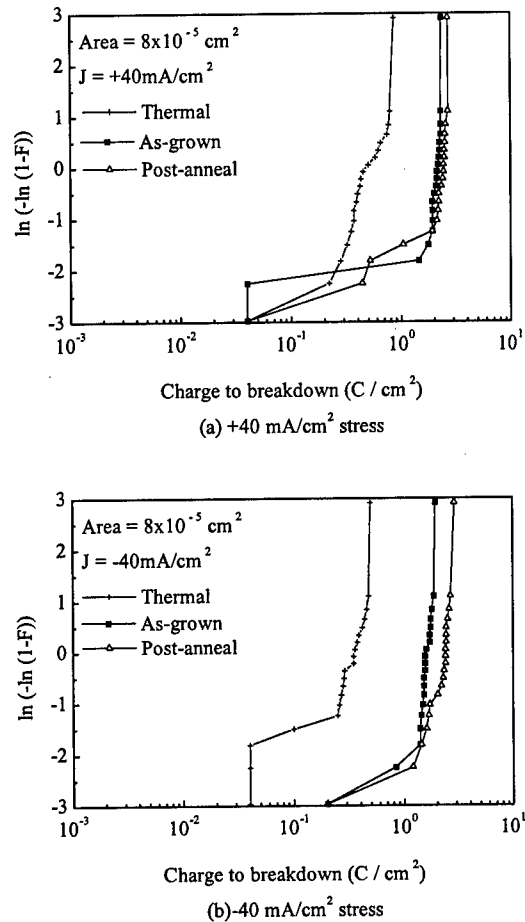
(b)-40 mA/cm^2 stress

Fig. 4 Weibull charge-to-breakdown plots of anodic polyoxide compared with thermal polyoxide under (a) +40 mA/cm^2 stress and (b) -40 mA/cm^2 stress.

4. Conclusion

In conclusion, application of anodic oxidation

on polysilicon is first proposed. The anodic polyoxide has higher effective electron barrier height, higher electric breakdown field and higher charge-to-breakdown than the conventional thermal polyoxide. A post-poly2 RTA treatment can further improve the electrical characteristics of anodic polyoxide. From these excellent characteristics, we believe that anodic polyoxide will be a highly potential candidate as high quality thin polyoxide for nonvolatile memory applications.

References

[1] C. Cobianu, O. Popa, and D. Dascalu, "On the electrical conduction in the dielectric layers," IEEE Electron Device Lett., vol. 14, p. 213, 1993.

[2] R. E. Shiner, N. R. Mielke, and R. Hag, "Characterization and screening of SiO2 defects in EEPROM structures," Int. Reliability Phys. Symp. IEEE Proc. pp. 248-256, 1983.

[3] T. Ono, T. Mori, T. Ajilka, and T. Takayashiki, "Studies of thin poly-Si oxides for E and E^2PROM," IEDM Tech. Dig., pp. 380-383, 1985.

[4] R. Anderson and D. Kerr, "Evidence for surface asperity mechanism of conductivity in oxides grown on polycrystalline silicon," J. Appl. Phys., Vol. 48, No. 11, p. 4834, Nov. 1977.

[5] H. Huff, R. Halvorson, T. Chiu, and D. Guterman, "Experimental observations on conduction through polysilicon oxide," J. Electrochem. Soc., Vol. 127, No. 11, p.2482, Nov. 1980.

SILICON CARBIDE

Electronic properties of SiO$_2$/SiC interfaces

V.V. Afanas'ev

Laboratory of Semiconductor Physics, Department of Physics, University of Leuven, 3001 Leuven, BELGIUM

Oxidation of silicon carbide (SiC) provides a possibility to combine the unique electronic and thermal properties of this semiconductor with the surface passivation attained by the natural insulating oxide SiO$_2$, in a similar way as it is done for silicon. However, the electron transport at the oxidized SiC surfaces is deteriorated by the SiC/oxide interface imperfections much more strongly than in the Si/SiO$_2$ structures. Two main contributions to the enhanced SiC/SiO$_2$ defect densities are revealed, including the carbon-related states and the near-interfacial oxide defects with energy levels in the SiC bandgap. These defects can be formed not only during the oxide formation or annealing, but, also, as a result of charge injection at the SiC/oxide interface which stresses the importance of the degradation effects in the SiC electronic devices.

1. INTRODUCTION

Due to their physical and, in particular, electronic properties various crystallographic modifications (polytypes) of silicon carbide (SiC) have attracted great interest as materials alternative to silicon in electronic device manufacturing [1,2]. The most important properties of SiC are the large bandgap width (from 2.38 eV for 3C-SiC to 3.26 eV for 4H-SiC [3]), the high critical electric field, the high saturation electron velocity, and the high thermal conductivity. These properties promise superior, as compared to Si, performance of SiC devices in extreme conditions of operation such as at high voltages, high power levels, high temperatures, and high frequencies. Moreover, SiC has a significant advantage over other wide-gap semiconductors related to the possibility to grow insulating SiO$_2$ overlayers by simple thermal oxidation. Taking into account the achievements in the SiC crystal growth, the latter was long considered as a guarantee of rapid success in the SiC metal-oxide-semiconductor (MOS) device development [1,2,4,5].

However, the SiC MOS devices, albeit fabricated in recent years by many groups, have not yet met the expectations. Their failure is largely related to the greatly enhanced density of imperfections at the SiO$_2$/SiC interface, which not only degrades the device performance but, also, may cause device reliability problems in relationship with anticipated extreme operation conditions. Thus, reduction of the SiO$_2$/SiC interface defect density becomes a key issue in the development of SiC MOS technology. Experience of the past few years indicates that the standard ways, known from the Si MOS processing, to reduce the density of electrically harmful defects fail for SiO$_2$/SiC structures. Clearly then, a better understanding of the defect nature in the oxidized SiC is needed in order to find technologically feasible methods to eliminate and/or passivate the defect sites.

Another aspect of the problem is related to the supposed use of the SiO$_2$/semiconductor interface in SiC MOS devices in much more harsh conditions (high electric field, high temperature, hydrogen exposure) than in Si MOS structures. The oxide degradation may pose limits to the range of device operation parameters (see, e.g., Ref. 6) reducing the gains provided by intrinsic SiC properties and forcing us to search for an alternative insulating material for SiC [7,8].

Primarily, two aspects of the defect problem will be addressed in the present paper: First, the possible reasons for the enhanced defect density in the as-prepared SiO$_2$/SiC structures as compared to oxidized Si will be discussed. Second, a highly efficient SiC/oxide interface defect generation by injection of electrons and holes into the oxide layer will be demonstrated. The latter suggests that there are significant hot-carrier reliability problems to be addressed in the SiC MOS electronics.

Table 1. Energy band diagram parameters of the interfaces of SiO₂ with Si and SiC

	Si	3C-SiC	15R-SiC	6H-SiC	4H-SiC
E_g (eV)	1.12	2.38	2.96	3.02	3.26
ΔE_c(eV)	3.15	3.6	3.0	2.95	2.7
ΔE_v(eV)	4.6	2.9	2.9	2.9	2.9
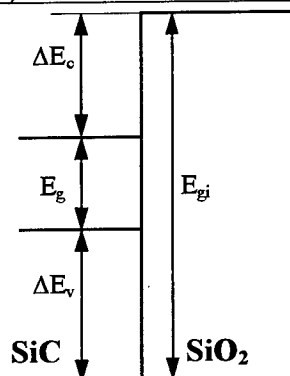
Fig. 1. Schematic energy band diagram of the SiC/SiO₂ interface indicating the principal parameters listed in Table 1.

2. BAND DIAGRAM OF SiO₂/SiC INTERFACE

The primary requirement to the gate or field dielectric consists in an efficient blockage of the charge carrier transport between transistor channel and gate. Thus, the energy barriers for electrons and holes in semiconductor have to be high enough in order to prevent thermally or field stimulated injection. With the semiconductor (E_g) and insulator (E_{gi}) bandgap widths fixed, the energy diagram would be optimal from the point of view of injection reduction if the conduction band offset (ΔE_c) would be equal to the valence band offset (ΔE_v) at the interface (Fig. 1). The band diagram parameters for the interfaces of SiO₂ with Si and four SiC polytypes are listed in Table 1 [9], in which the valence band offset is calculated using E_{gi} = 8.9 eV for thermal SiO₂ [10, 11]. It is seen from the listed data that the band diagram of SiO₂/SiC interfaces is nearly symmetric. Because of this, the leakage current in the SiC MOS structures remains acceptably low, even at elevated temperatures [8,12,13], while other insulators (nitrides, metal oxides) fail [8]. High barriers for electron and hole injection and good oxide quality also account for the high oxide breakdown field (~10 MV/cm at 300 K [8,14,15,16]). Good insulating properties of SiO₂
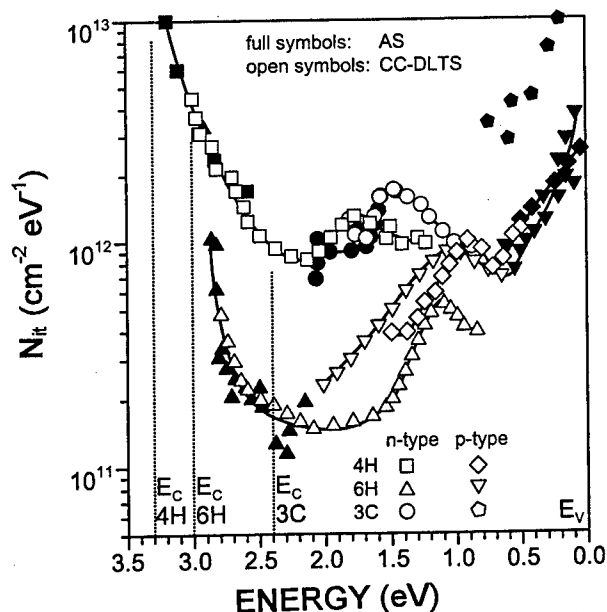
Fig. 2. Energy distribution of interface state density N_{it} for 3C-, 4H-, and 6H-SiC MOS structures measured using AS (full symbols) and CC-DLTS (open symbols). The origin of energy scale is the top of the SiC valence band; the dotted lines mark the conduction band energies in various SiC polytypes.

combined with its technological feasibility make it most suitable for the gate insulation of SiC [17].

3. INTERFACE DEFECTS IN AS-OXIDIZED SiO₂/SiC

3.1 Energy distribution of the interface states

Despite the high quality of the thermally grown SiO₂ (comparable to the oxide grown on Si), the oxidized SiC, particularly crystals of p-type conductivity, exhibit a much less perfect interface than the oxidized Si. This is exemplified in Fig. 2 by the energy distribution of SiO₂/SiC interface states (N_{it}) measured for three SiC polytypes (3C, 4H, 6H) using the admittance (AS) and the high-temperature constant capacitance deep-level transient (CC-DLTS) spectroscopies [18]. The increase of N_{it} from approximately midgap towards valence band is seen for all the SiC polytypes, indicating a common type of interface states. These states are donor type and largely account for the positive charge commonly observed in the p-type SiC MOS structures. In the n-type samples they are filled with electrons and remain neutral.
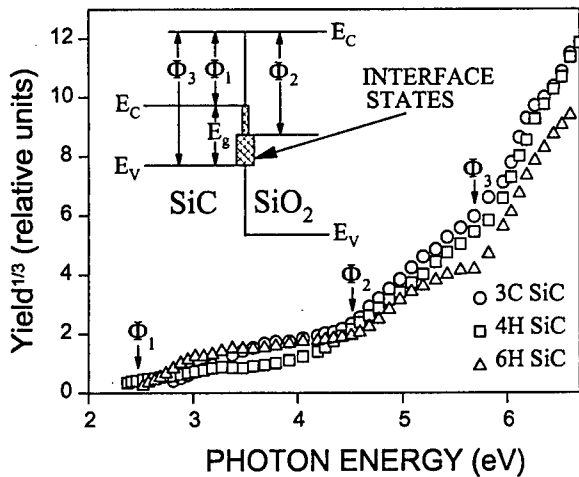
Fig. 3. Spectral dependences of the IPE yield for n-type 3C- (O), 4H- (□), and 6H-SiC (△) MOS structures measured with an electric field strength in SiO₂ of 3 MV/cm. The insert shows a schematic band diagram and the spectral thresholds of electron transitions.

Fig. 4. 1 kHz C-V curves for an n-type 4H-SiC MOS capacitor with 100-nm thick oxide measured at room (solid line) and low (dash-dotted line) temperatures at the voltage sweep rate of 0.4 V/s. The arrows indicate the direction of voltage sweep.

A high density of the interface states in the lower part of the SiC bandgap is also observed in the internal electron photoemission experiments (IPE), in which electrons are optically excited from the states at the semiconductor surface and emitted into the oxide. As can be seen from the spectral curves of the IPE yield (Fig. 3), in addition to the fundamental spectral thresholds Φ_1 and Φ_3 corresponding to the excitation of electrons from the conduction and valence bands of SiC, respectively, there is an additional emission with spectral threshold Φ_2 corresponding to the excitation of electrons from states with energy within the SiC bandgap [9,18]. These states are observed in all the studied SiC polytypes (also in 15R-SiC [9]) which make us believe that the electrical techniques (Fig. 2) and the IPE (Fig. 3) reveal the same defects.

The density of the SiO₂/SiC interface states appears to be highly sensitive to the orientation of the substrate crystal: both the electrical data and IPE indicate that N_{it} is much higher on the carbon face of hexagonal SiC [(000$\bar{1}$) plane] than on the silicon face [(0001) plane] [9,18,19]. Oxidation of other SiC planes leads to higher defect densities than at the Si-face SiC/oxide interface [18,20]. An amorphization of the SiC surface results in a great N_{it} increase, leading to Fermi level pinning [21].

Importantly, the general type of the interface state distribution – higher density of states in the lower half of the SiC bandgap as compared to the upper part –, is observed irrespectively of particular plane chosen for oxidation [18]. This observation shows that the SiC substrate orientation influences the number of the defects but not their type.

Another component of the SiO₂/SiC interface spectrum is seen close to the conduction band edge in the 6H- and 4H-SiC MOS structures (Fig. 2). It can also be directly observed by capacitance-voltage (C-V) measurements at low temperature. In Fig. 4 are shown the C-V curves measured in a n-type 4H-SiC MOS structure at 300 and at 80 K, indicating that the temperature lowering leads to additional electron trapping. With the accumulation voltage increasing, the number of trapped electrons increases suggesting their tunneling to the traps. The trapped electron density in the 6H-SiC samples is considerably lower than in the 4H-SiC ones, in agreement with data shown in Fig. 2. As no additional charge trapping is observed neither in the n-type 3C-SiC samples nor in the p-type SiC MOS structures, it is likely that the shallow traps are present only in the energy range close to the conduction bands of 4H- and 6H-SiC. These shallow traps may be associated with oxide defects

revealed at both SiO_2/Si and SiO_2/SiC interfaces by using the photon-stimulated tunneling spectroscopy [18,22]. Their energy level is at 2.8 eV below the oxide conduction band edge, i.e., at 0.1 eV below the conduction band of 4H-SiC at the interface with the oxide (cf. Table 1).

Trapping of electrons by shallow traps suggests that these defects may affect the effective channel mobility in the 4H-SiC MOS devices. Indeed, this trend is observed experimentally: While the lateral 6H-SiC n-channel transistors show a mobility of 35-50 cm^2/Vs, similar 4H-SiC devices have a mobility of only 4-5 cm^2/Vs [23]. These values can be compared with the Hall electron mobilities in the 4H- and 6H-SiC of ≥ 100 cm^2/Vs measured at 300 K in the direction normal to the c-axis [24]. Apparently, electrons are largely transferred from the inversion channel to the oxide traps and do not contribute to the channel current. Also, these states may account for the thermally activated electron transport along the oxidized SiC surfaces [25].
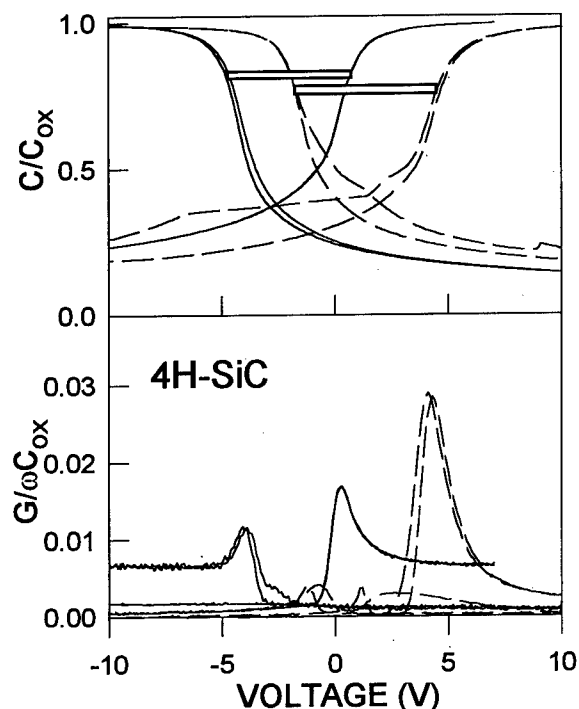
Fig. 5. Normalized 1 kHz capacitance (top) and conductance (bottom) as a function of voltage for the n- and p-type 4H-SiC MOS structures with 50-nm thick SiO_2 layers grown at 1120 °C in dry O_2 (solid lines) or in $Ar+H_2O$ (dashed lines).

3.2 Effects of processing on the interface states

Observation of higher defect density in the p-type SiC MOS structures as compared to the n-type ones led initially to suggestion that the defects are related to Al (used as acceptor dopant in SiC) incorporation into the oxide [26]. However, similarities observed between Al- and B-doped p-type SiC [27], and the absence of an adverse influence of Al on n-type SiC/SiO_2 samples prepared by over-compensation of the Al-doped p-type SiC by nitrogen implantation [28], allows to exclude Al as the origin of poor quality of SiC/oxide interfaces. Obviously then, the interface defects result from some process(es) intrinsic for the SiO_2/SiC system.

Technological experiments reveal that the density of the SiO_2/SiC interface states can be reduced if the pre-oxidation surface cleaning is performed using an aggressively oxidizing ambient like UV-ozone [16] and oxygen plasma [4]. Preservation of the oxidizing ambient during wafer loading was also found beneficial [4,27]. These observations suggest that the generation of the SiO_2/SiC interface states is related to the interface species unstable against oxidation, but resistant to the standard wet cleaning.

Another way to reduce the interface defect density in p-type SiC/SiO_2 structures is an additional re-oxidation at a lower temperature in presence of water [29-31]. As can be seen from the MOS capacitance and conductance curves shown in Fig. 5, the densities of charged defects responsible for the shift of the C-V curve and of the fast interface states accounted for the bell-shaped conductance curves are strongly reduced if oxidation is done with slow (0.1 °C/min) cooling in H_2O-containing ambient as compared to the standard dry oxidation. However, it appeared that the *decrease* in the defect density in p-type structures is accompanied by the *increase* of defects number in n-type samples [31], as also shown in Fig. 5. The difference in the flatband voltage between the n- and p-type samples which has to be close to the SiC bandgap width in an ideal MOS structure, is barely affected by the wet process. Moreover, the negative charge of the defects observed in the n-type samples seems to degrade surface electron mobility [31]. Taking into account that the wet oxidation leads to additional oxide conduction (see, e. g., Fig. 2 in Ref. 15) and related reliability problems, one cannot say that the problem of SiO_2/SiC interface defects is solved.
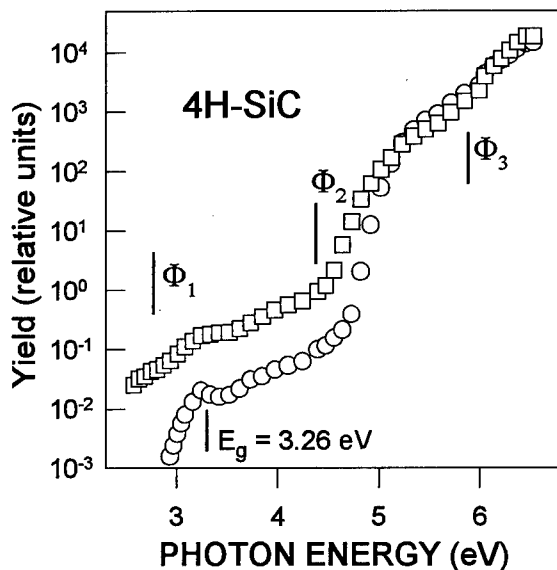
Fig. 6. Spectral dependences of the IPE yield for n-type 4H-SiC MOS structures with 50-nm thick oxide grown at 1120 °C in dry O_2 (O) or in Ar+H_2O (□). The curves were measured with an electric field strength in SiO_2 of 2 MV/cm. The SiC bandgap value and the spectral thresholds of electron transitions from various states are indicated.

The reason for the opposite action of wet oxidation on the p- and n-type SiC/SiO_2 samples is revealed by the IPE spectra shown in Fig. 6. The major impact of wet oxidation is ~10-fold increase of the IPE yield in the low-energy emission band Φ_1, which corresponds to excitation of electrons from the SiC conduction band (cf. Fig. 3). As the total electron density at the SiC surface remains the same (it is determined by the applied field), the yield rise suggests an increase in the probability of electron escape, which means that electrons are excited not from the SiC conduction band, but from some other states. The absence of the yield drop at $h\nu = E_g$ due to onset of electron-electron scattering in SiC [9] suggests that in the wet-oxidized samples electrons are emitted from the oxide traps close to the SiC/oxide interface. Apparently then, the wet oxidation results in generation of oxide acceptor - type defects which compensate positive charge of the donor states (which remain untouched: IPE in the defect-related band Φ_2 is unchanged, Fig. 6) and, because of Coulomb attraction, shift the energy levels of the donors deep into SiC bandgap.

4. DEGRADATION OF SiO_2/SiC INTERFACES

4.1 Annealing-induced SiO_2/SiC degradation

In the course of device processing, the thermal oxidation represents one of the first steps. After it, the SiO_2/SiC structure may be subjected to various supplemental thermal treatments aimed on the passivation or annealing of defects, activation of doping impurities, formation of ohmic contacts, etc. This additional processing may also have a negative effect on the electrical properties of the SiC/oxide interface and result in generation of new defects. As the first indication of a problem, the annealing of the poly-Si gate electrode deposited on p-type 6H-SiC/SiO_2 at T >800 °C was found to increase the interface state density, which was ascribed to the influence of thermally-induced strain [32]. Very similar effect, however, was reported for the bare oxide samples annealed at 1000 °C in order to obtain ohmic contacts [30], suggesting involvement of some chemical factor in the defect generation.

Annealing in hydrogen, typically used in the Si MOS technology for defect passivation, appears to be marginally efficient in SiO_2/SiC structures. So far, the only improvement observed is passivation of interface states produced by radiation in a sputtering metallization process [8,33,34]. At T >500 °C anneal in hydrogen leads to positive charge buildup at the interfaces of both n- and p-type SiC with SiO_2 [35]. As the similar effect is observed in the SiO_2/Si structures, it was ascribed to interaction of H with the near-interfacial oxide leading to formation of a over-coordinated oxygen center [36]. The density of positive charge shows a trend to increase with increase of the conduction band offset at the interface indicating that electron emission from an H-related center into semiconductor (the interfacial ionization of hydrogen) is in the core of the charging effect. Moreover, in the wide-gap SiC polytypes (4H-, 6H-) the H-induced degradation is found to be substantially enhanced in the p-type as compared to the n-type samples. Apparently, the trapping of holes from SiC plays significant role in the interface degradation [37]. Taking into account that formation of the H-induced positively charged state may be involved into the interfacial bond rupture process [37], it is well possible that the presence of H may also account for the SiO_2/SiC interface state generation upon anneal.

4.2. Injection-induced SiO₂/SiC degradation

Electron injection in the SiC MOS oxide is another source of degradation. Such injection may be even more efficient than the hot-electron injection in Si/SiO₂ structures because, first, the electric fields in the SiC devices are much higher [1,13], and, second, the electrons with energy of about 3 eV (sufficient for injection into the oxide, cf. Table 1) will have large free path with respect to electron-electron scattering because of wide SiC bandgap. The presence of high-energy electrons in SiC is affirmed by efficient field emission of electrons from SiC p-n junctions into vacuum [38]. Moreover, as the device operation temperature increases, the interfacial barrier becomes lower [12,13] leading to further increase of the electron emission rate at elevated temperatures.

As revealed by the early studies, the electron injection into the oxide grown on SiC results in enhanced generation of interface states [33,39]. Noteworthy here is that the charge trapping in the oxide itself remains similar to that in the SiO₂ layers grown on Si [39-41], indicating an interfacial degradation process. Electron injection results in the generation of both donors and acceptors, which

leads to a net positive charge in the p-type samples and to a negative one in the n-type samples. The interface state charge generation kinetics shown in Fig. 7 reveals that a) the density of interfacial acceptors is higher than that of donors and b) there is a strong influence of the SiC polytype on the degradation process: 4H-SiC samples show much more defects than 6H- or 3C-SiC MOS structures. Generation of interface acceptors correlates with the presence of C at the interface [41].

The injection-induced SiC MOS degradation is not only significantly enhanced as compared to the Si MOS samples but also results in defects more stable than Si/SiO₂ interface states [39,41]. This is seen in Fig. 8 which demonstrates the effect of anneal on the injection-degraded 4H-SiC/SiO₂ interfaces. Importantly, an efficient anneal of the defects at 400 °C cannot be reached neither in vacuum nor in hydrogen, suggesting that the generated states are likely related to some stable bonding configurations of interfacial atoms. These results explain why an enhanced SiO₂/SiC interface state density is observed when e-beam or sputtering metallization is used [34]. Clearly, generation of such stable defects poses a serious threat to the device reliability.

Fig. 7. Density of charge trapped at the interface states derived from the hysteresis of C-V curves as a function of the injected electron flux in p- (empty symbols) and n-type (filled symbols) SiC MOS structures with ~100 nm-thick oxides grown on 4H- (circles), 6H- (squares) and 3C-SiC (triangles).
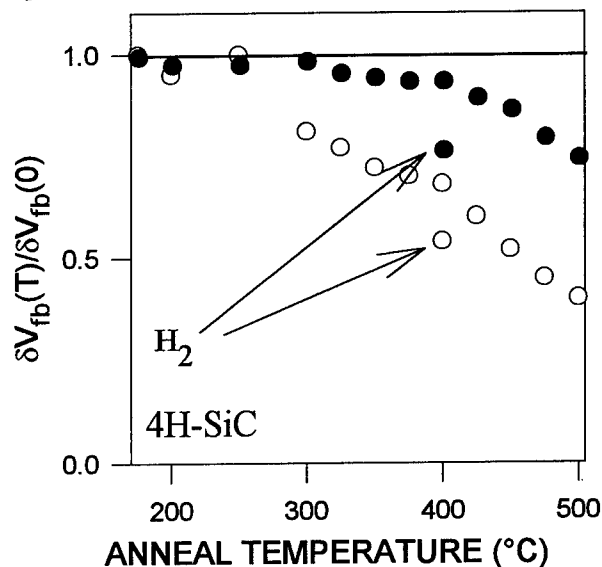
Fig. 8. Fraction of interface states in p- (○) and n-type (●) 4H-SiC MOS structures remaining at the interfaces (degraded by injection of 10^{17} e/cm²) after 30 min annealing in vacuum at the indicated temperature. The arrows indicate the effect of 30 min anneal in pure H₂ at 400 °C.

5. ORIGIN OF SiO_2/SiC IMPERFECTIONS

Dangling bonds of surface Si atoms (P_b centers) represent the main interface state source at the (111)Si/SiO_2 interface [42], which is structurally isomorphic to the oxidized Si-faces of hexagonal SiC. As these defects are introduced to account for the structural mismatch between Si and the oxide [43], it is logical to expect their presence at the SiO_2/SiC interfaces as well. However, the interface states observed in the as-oxidized SiC appear to be different from the well-studied P_b-s at least in two principal aspects. First, they cannot be passivated with hydrogen [18,35]. Second, their electrical behavior is dissimilar to well known amphoteric character of the P_b centers at the (111)Si/SiO_2 interface [18]. Probably, the dangling bond defects are present at the SiO_2/SiC interfaces, but, like in the SiO_2/Si structures, they are well passivated with H during processing.

In an attempt to explain the energy distribution and stability of the SiO_2/SiC interface traps, it was suggested that they are related to the carbon clusters with sp^2-bonded fragments [18,44]. Indeed, the presence of excessive carbon at the SiC/oxide interfaces was revealed by electron spectroscopy [45-47] and by electron spin resonance [40]. The carbon cluster model succeeds in explaining not only the stability of interface defects against passivation with H, but, also, the correlation of the interface state density with C presence at the interface, modification of their spectrum at high C contents (graphitization), and the inhomogeneous distribution of the interfacial electrostatic potential [48]. The strong sensitivity of the interface state density to the pre-oxidation SiC surface treatments [16,49] points towards surface defects and/or carbon contaminants as possible nucleation sites for carbon during oxidation. Atomic force microscopy of SiC surfaces after oxide removal shows that they contain imperfections in a density much higher than (111)Si/SiO_2 (see, e.g., Fig. 8 in Ref. 47). Moreover, in the oxidized 4H-SiC the interface state density was found to correlate with the presence of clusters which chemically resemble carbon [35].

Another possible source of the interface states at the SiO_2/SiC interfaces are the oxide defects with energy levels inside the SiC bandgap. In addition to the already mentioned 2.8-eV deep acceptor centers

[18], one may recall the H-decorated oxygen vacancy as operating electron trap with energy level at 3.1 eV below the SiO_2 conduction band [50]. Donor-type states may also originate from oxide defects: The energy of the 0/+ levels of oxygen vacancy centers are predicted to be in the range of 4.3-5 eV below the oxide conduction band [51,52]. Taking into account that the E'$_\gamma$-centers related to the O-vacancy defects have been observed in the SiC/SiO_2 structures [40], contribution of these oxide defects to the N_{it} distribution cannot be excluded.

It may appear difficult to distinguish between the two latter contributions to the interface states density, because, as known from SiO_2/Si structures, the presence of carbon promotes thermally-induced oxide degradation [53,54]. Thus, in the vicinity of C clusters, carbon can promote SiO_2 reduction [53]:

$$SiO_2 + 2C \rightarrow Si + 2CO\uparrow, \qquad (1)$$
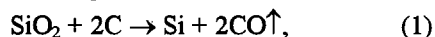
where symbol Si should be understood as formation of an oxygen-deficient defect in the oxide. According to this scenario, any presence of carbon at the SiC/SiO_2 interface in a bonding configuration different than that in SiC may have a potentially deteriorating effect on the electronic properties of the interface and of the oxide.

6. CONCLUSIONS

The high density of the interface defects represents the major obstacle for a wide SiO_2/SiC application in solid state electronics. With the two main sources of excessive (with respect to the oxidized Si) defects isolated as interfacial carbon and the oxide defects, one can point to possible ways to improve the situation. First, pre-oxidation preparation of the SiC substrate aiming surface defect elimination and carbon removal, may be useful to suppress formation of extra defects. Second, as the intrinsic SiO_2 defects are naturally incorporated into the oxide and cannot be avoided, the SiC polytype is to be chosen to minimize their influence. As it has become clear now, 4H-SiC seems not to be the material of choice for MOS applications.

REFERENCES

1. M. Ruff, H. Mitlehner, and R. Helbig, IEEE Trans. Elecron Dev., 41 (1994) 1040.
2. H. Morkoc, S. Strite, G. B. Gao, M. E. Lin, B. Sverdlov, and M. Burns, J. Appl. Phys., 76 (1994) 1363.
3. W. J. Choyke, In: The physics and chemistry of carbides, nitrides and borides, R. Freer (ed.), Kluwer, Dordrecht, 1990, p. 653.
4. J. A. Cooper, Jr., Phys. Stat. Sol. (a), 162 (1997) 305.
5. M. R. Melloch and J. A. Cooper, Jr., MRS Bull., 22 (1997) 42.
6. M. Mathur Maranowski and J. A. Cooper, Jr., IEEE Trans. Electron Dev., 46 (1999) 520.
7. C.-M. Zetterling, M. Östling, C. I. Harris, N. Nordell, K. Wongchoutigul, and M. G. Spencer, Mater. Sci. Forum, 264-268 (1998) 877.
8. L. A. Lipkin and J. W. Palmour, IEEE Trans. Electron Dev., 46 (1999) 525.
9. V. V. Afanas'ev, M. Bassler, G. Pensl, M. J. Schulz, E. Stein von Kamienski, J. Appl. Phys., 79 (1996) 3108.
10. T. H. DiStefano and D. E. Eastman, Solid State Commun., 9 (1971) 2259.
11. V. K. Adamchuk and V. V. Afanas'ev, Sov. Phys.: Solid State, 26 (1984) 1519.
12. E. Bano, T. Ouisse, P. Lassagne, T. Billon, C. Jaussand, Inst. Phys. Conf. Ser. No 142 (1996) 733.
13. A. K. Agarwal, S. Seshardi, and L. B. Rowland, IEEE Electron Dev. Lett., 18 (1997) 592.
14. D. Alok, P. K. McLarty, and J. Baliga, Appl. Phys. Lett., 64 (1994) 2845.
15. P. Friedrichs, E. P. Burte, and R. Schörner, Appl. Phys. Lett., 65 (1994) 1665.
16. V. V. Afanas'ev, A. Stesmans, M. Bassler, G. Pensl, M. J. Schulz, and C. I. Harris, Appl. Phys. Lett., 68 (1996) 2141.
17. C. I. Harris and V. V. Afanas'ev, Microelectronic Eng., 36 (1997) 167.
18. V. V. Afanas'ev, M. Bassler, G. Pensl, and M. J. Schulz, Phys. Stat. Sol. (a), 162 (1997) 321.
19. C.-M. Zetterling and M. Östling, Phys. Scripta, T54 (1994) 291.
20. J. N. Shenoy, M. K. Das, J. A. Cooper, Jr., M. R. Melloch, J. W. Palmour, J. Appl. Phys., 79 (1996) 3042.
21. D. Alok, B. J. Baliga, and P. K. McLarty, IEEE Electron Dev. Lett., 15 (1994) 424.
22. V. V. Afanas'ev and A. Stesmans, Phys. Rev. Lett., 78 (1997) 2437.
23. J. Spitz, M. R. Melloch, J. A. Cooper, Jr., and M. A. Capano, IEEE Electron Dev. Lett., 19 (1998) 100.
24. T. Kinoshita, K. M. Itoh, J. Muto, M. Schadt, G. Pensl, K. Takeda, Mater.Sci.Forum, 264-268 (1998) 295.
25. T. Ouisse, Phys. Stat. Sol. (a), 162 (1997) 339.
26. M. Shinohara, M. Yamanaka, S. Misawa, H. Okumura, S. Yoshida, Jpn. J. Appl. Phys., 30 (1991) 240.
27. J. N. Shenoy, G. L. Chindalore, M. R. Melloch, J. A. Cooper, Jr., J. W. Palmour, and K. G. Irvine, J. Electron. Mater., 24 (1995) 303.
28. M. Bassler, V. V. Afanas'ev, and G. Pensl, Mater. Sci. Forum, 264-268 (1998) 861.
29. L. A. Lipkin and J. W. Palmour, J. Electron. Mater., 25 (1996) 909.
30. K. Ueno, Mater.Sci.Forum, 264-268 (1998) 845.
31. H. Yano, F. Katafuchi, T. Kimoto, and H. Matsunami, IEEE Trans. Electron Dev., 46 (1999) 504.
32. W. Xie, J. N. Shenoy, S. T. Sheppard, M. R. Melloch, and J. A. Cooper, Jr., Appl. Phys. Lett. 68 (1996) 2231.
33. E. G. Stein von Kamienski, F. Portheine, J. Stein, A. Gölz, and H. Kurz, J. Appl. Phys. 79 (1996) 2529.
34. J. Campi, Y. Shi, Y. Luo, F. Yan, J. H. Zhao, IEEE Trans. Electron Dev., 46 (1999) 511.
35. V. V. Afanas'ev, A. Stesmans, and C. I. Harris, Mater. Sci. Forum, 264-268 (1998) 857.
36. V. V. Afanas'ev and A. Stesmans, Phys. Rev. Lett. 80 (1998) 5176.
37. V. V. Afanas'ev and A. Stesmans, Mater. Sci. & Eng. B, 58 (1999) 56.
38. L. Patrick and W. J. Choyke, Phys. Rev. Lett., 2 (1959) 48.
39. V. V. Afanas'ev, M. Bassler, G. Pensl, and M. J. Schulz, Microelectron. Eng., 28 (1995) 197.
40. V. V. Afanas'ev and A. Stesmans, Appl. Phys. Lett., 69 (1996) 2252.
41. V. V. Afanas'ev, A. Stesmans, M. Bassler, G. Pensl, M. J. Schulz, C. I. Harris, J. Appl. Phys., 55 (1999) 8292.
42. E. Poindexter, Semicond. Sci.Technol., 4 (1989) 961.
43. A. Stesmans, Phys. Rev. B, 48 (1993) 2418.
44. G. Pensl, V. V. Afanas'ev, M. Bassler, M. Schadt, T. Troffer, J. Heindl, H. Strunk, M. Maier, and W. J. Choyke, Inst. Phys. Conf. Ser. No. 142 (1996) 275.
45. B. Hornetz, H.-J. Michel, and J. Halbritter, J. Mater. Res., 9 (1994) 3088.
46. V. R. Vathulya, D. H. Wang, and M. H. White, Appl. Phys. Lett., 73 (1998) 2161.
47. P. Måretensson, F. Owman, and L. I. Johansson, Phys. Stat. Sol. (b), 202 (1997) 501.
48. E. Bano, T. Ouisse, L. DiCioccio, and S. Karmann, Appl. Phys. Lett., 65 (1994) 2723.
49. K. Ueno, R. Asai, and T. Tsuji, IEEE Electron Dev. Lett., 19 (1998) 244.
50. V. V. Afanas'ev and A. Stesmans, Appl. Phys. Lett. 71 (1997) 3844.
51. A. Chu, W. B. Fowler, Phys. Rev. B, 41 (1990) 5061.
52. K. C. Snyder and W. B. Fowler, Phys. Rev. B, 48, (1993) 13238.
53. A. G. Revesz, G. A. Brown, and H. L. Hughes, Mat. Res. Soc. Symp. Proc., 244 (1993) 555.
54. S. I. Raider, Microelectronic Eng., 22 (1993) 29.

Effect of Post Oxidation Processing on Dry Oxides on N-Type 4H-SiC

C.J. Anthony and M.J. Uren

Defence Evaluation and Research Agency,
St. Andrews Road, Malvern, Worcestershire, WR14 3PS, England.
Telephone: +44-1684-894404, Facsimile: +44-1684-895113
Email: canthony@dera.gov.uk

We report a study of the effect of post oxidation processing on the interface quality of dry oxides on n-type 4H SiC. We have found that the processing required to make a polysilicon gated SiC capacitor with full ohmic back contact did not affect the interface state density of the oxide, but rather improved the fixed oxide charge and slow state density, with both effects increasing the further through the process the oxides were taken.

1. INTRODUCTION

Early work on thermal oxides on SiC showed a relatively poor quality oxide on p-type material as opposed to n-type [1]. Following intensive study, the quality of oxides on p-type has improved rapidly. Silicon Carbide (SiC) now shows considerable promise for high power electronic device applications utilizing MOS devices. Most researchers investigating the SiO_2/SiC interface have chosen the MOS capacitor with a large area front substrate contact as their characterization structure. However, the simple and fast fabrication process for this device does not subject the oxide to the subsequent processing steps required for actual power device production.

Previously we have reported [2] that wet and dry oxides on 4H SiC, which had both received a full fabrication process, showed similar interface state densities. However, the dry oxide exhibited a superior oxide reliability as seen by Friedrichs et al [3]. Here we report on the effect of our post oxidation processing, on the quality of our thermally grown dry oxides on n-type 4H-SiC. We report the first direct measurement of slow state densities on SiC.

2. EXPERIMENT

The oxide samples described here have used a production grade 4H-SiC substrate purchased from Cree Research Inc. The wafer was the vendor's standard product and was 35mm in diameter, had an n-type doping of about $9 \times 10^{18} cm^{-3}$, with a $3\mu m$ thick $2 \times 10^{17} cm^{-3}$ n-type epitaxial layer. The wafer was first diced into quarters. The dicing was followed by both solvent and chemical cleaning. The quarters were then oxidized to a thickness of 21nm by dry oxidation at 1050°C for 14hrs. From this identical starting point four Al gate capacitor samples were fabricated, each experiencing additional processing steps. (1) Oxide only - Al gate & back contact. (2) Poly - PolySi deposited/ stripped then Al gate & back contact. (3) Poly/LTO - PolySi & LTO (Low temperature oxide) deposited/ stripped then Al gate & back contact. (4) Full process - PolySi & LTO deposited, sample back stripped, Ni/Au ohmic back contact e-beam evaporated and annealed at 950°C in N_2, LTO & PolySi stripped from fronts and Al gates formed.

All measurements reported here were carried out on wafer with a screened hot

chuck. High frequency capacitance-voltage (HFCV) measurements using a HP4284A precision LCR meter were performed at room temperature. The AC conductance technique [4] at 25°C & 200°C was used to accurately extract the interface state density (D_{IT}) close to the conduction band edge. Here a wide frequency range, high accuracy, bridge was used in conjunction with a HP4284A impedance analyzer [5].

3. RESULTS
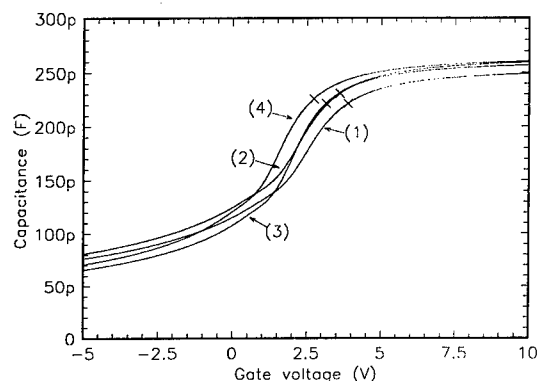
Figure 1. High frequency CV curves @25°C

Figure 1 shows the 100kHz CV curves evaluated at 25°C for each sample and the net fixed charge at flat-band is shown in table 1. A reduction in net fixed oxide charge at flat-band occurred with additional processing.

Table 1
Net fixed oxide charge (m^{-2})

(1) – Oxide only	-3.4×10^{16}
(2) – Poly only	-3.3×10^{16}
(3) – Poly/LTO	-2.8×10^{16}
(4) – Full process	-2.4×10^{16}

Figure 2 shows typical equivalent parallel conductance data measured at 200°C over a particularly wide frequency range. It is apparent that in addition to the peak due to the fast states, there is a

considerable density of slow states, as evidenced by the plateau on the low frequency side of the fast-state peak [5]. To our knowledge, this is the first direct observation of slow states at the SiC/SiO_2 interface.
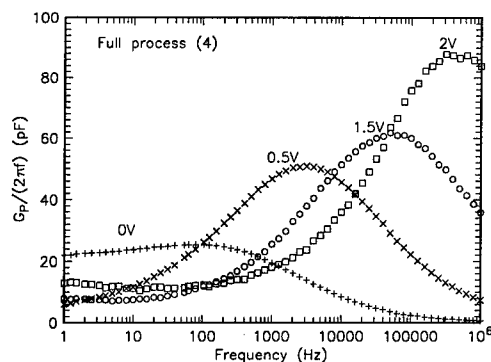
Figure 2. Conductance curves for sample (4) @200°C

The conductances measured at 200°C have been fitted with a 4 parameter model. The fast states are modelled using the normal Brews potential fluctuation model [4] (parameters are: peak frequency, amplitude and width), and the slow states by the McWhorter tunnelling model [6] i.e. traps uniformly distributed in energy and depth into the oxide (plateau height and maximum frequency equal to the fast state peak). A good fit was obtained to the peaks, but the assumption of a uniform slow trap density was fairly crude. Conductance measurements at 25°C were fitted with fast states only (3 parameters) since the peaks were so broad that no slow state plateau could be reliably distinguished. Since reliable LF CV curves could not be obtained, the surface band bending was obtained from the fast state peak frequency by assuming a capture cross-section of $10^{-19} m^2$.

Figure 3 shows the fast state density, where the rapid rise towards the band edge can be seen. This could be the carbon related level on 4H-SiC described by Bassler [7]. (It should be noted that the rapid rise may partly be an artefact associated with the

assumption of a constant capture cross-section.) No strong dependence on processing was seen.
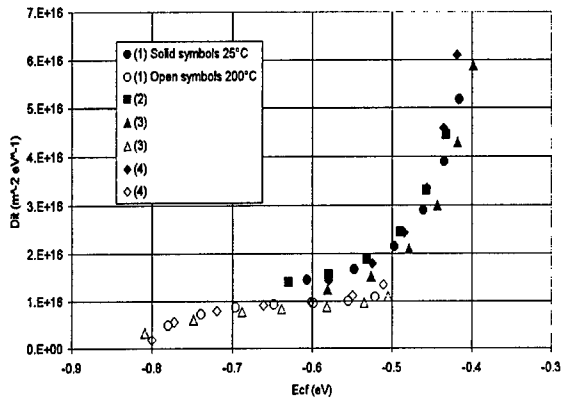
Figure 3. Fast interface state density

The slow state density is shown in figure 4 and is very high by comparison with values inferred for silicon where typical values are below $10^{23} m^{-3} eV^{-1}$. It is clear that the assumption of a uniform density within the oxide is too simple an assumption to explain this data, with a U-shape density in energy observed. Similarly, the curvature in the height of the plateau seen in figure 2 suggests a strong dependence of density on depth into the oxide (or in activation energy
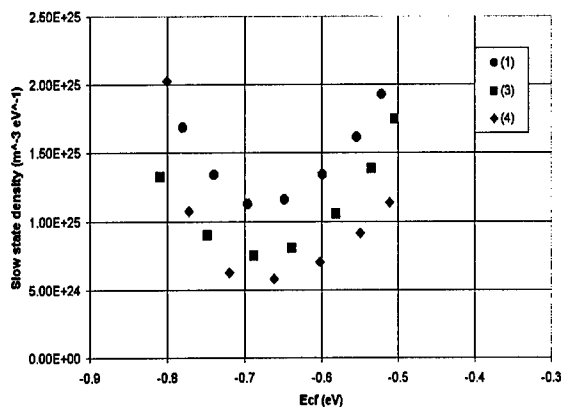
Figure 4. Slow state density

for capture[5]). However, it is clear that processing has produced a significant reduction in the slow state density.

Figure 5 shows the surface potential fluctuations inferred from the peak width, large potential fluctuations are typically seen [8]. The lines show the result of fitting Brews potential fluctuation model with a screening parameter, $\lambda = 15nm$ and no compensation. The model predicts a fall in potential fluctuations towards the band edge as a result of the screening by the rapidly increasing interface state capacitance, however, experimentally this is not seen in the data measured at 25°C which probes closer to the conduction band edge. This could be due to distortion of the curves by a high density of slow states (ignored in the fitting of this data as noted earlier), or it could be due to an intrinsic distribution of capture cross-sections (chemical inhomogeneity or multiple defect types) [9].
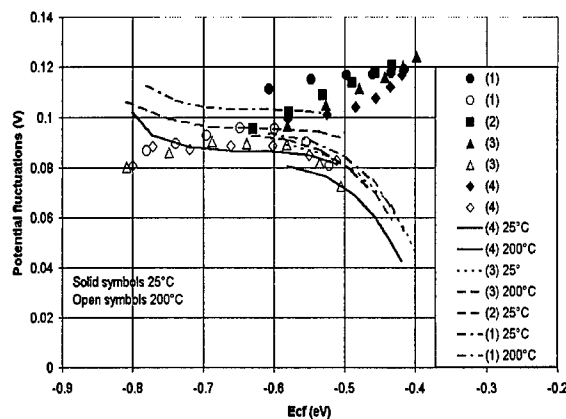
Figure 5. Potential fluctuations

4. CONCLUSIONS

We have presented results on the effect of post oxidation processing on the interface quality of dry thermal oxides grown on n-type 4H-SiC. For the first time, the slow state density was measured and found to be reduced by processing. By contrast, the fast

state density was broadly unaffected by processing. The slow states are important technologically since they will give rise to threshold voltage instability and 1/f noise in MOSFETs.

Our results did not show any particular process step as being responsible for a dramatic change in interface properties, rather, there was a gradual change in properties with the deposition of polysilicon and LTO and the annealing of ohmic contacts. These results indicate the desirability of carrying out SiC/SiO_2 research in a whole process context.

REFERENCES

[1] D.M. Brown, M. Ghezzo, J. Kretchmer, E.Downey, J.Pimbley and J.W. Palmour, IEEE Trans on Electron Dev., 41(4) (1994) 618

[2] C.J. Anthony, A.J. Jones and M.J. Uren, presented at ECSCRM98, to be published in Mat. Sci. Eng. B.

[3] P. Friedrichs, E.P. Burte and R. Schorner, J. Appl. Phys., 79 (10), (1996) 7814

[4] E.H. Nicollian and J.R. Brews, MOS (Metal Oxide Semiconductor) Physics and Technology, Wiley, New York, 1994

[5] M.J. Kirton and M.J. Uren, Adv. in Phys., 38 (1989) 367

[6] A.L. McWhorter, in Semiconductor Surface Physics, edited by R.H. Kingston, Uni. Of Philadelphia Press, (1956) 207

[7] M. Bassler, G. Pensl and V. Afanas'ev, Diamond and Rel. Mat., 6 (1997) 1472

[8] T. Ouisse, Phys. Stat. Sol. (a) Appl. Res., 162 (1997) 339

[9] M.J.Uren, J.H. Stathis and E. Cartier, J. Appl. Phys., 80 (7) (1996) 3915

Reliability of Metal-Oxide-Semiconductor Capacitors on 6H-Silicon Carbide

M. Treu[*,a], R. Schörner[b], P. Friedrichs[b], R. Rupp[b], A. Wiedenhofer[b], D. Stephani[b], and H. Ryssel[a,c]

[a] Fraunhofer-Institut für Integrierte Schaltungen, Abteilung Bauelementetechnologie IIS-B, Schottkystrasse 10, D-91058 Erlangen, Germany

[b] Siemens AG, Corporate Technology, P.O. Box 3220, D-91050 Erlangen, Germany

[c] Lehrstuhl für Elektronische Bauelemente, Cauerstraße 6, D-91058 Erlangen, Germany

1. Introduction

Silicon carbide (SiC) offers the capability to form silicon dioxide (SiO_2) by thermal oxidation analogously to silicon. For thermal oxides on silicon carbide (SiC) breakdown field strengths above 10 MV/cm and interface state densities around 10^{12} cm^{-2} have already been achieved [1, 2, 3, 4, 5]. However, concerning the long term reliability of the oxide, a crucial parameter for the commercial use of MOS devices, experimental data are still scarce. In this paper the long term stability of polysilicon gated MOS capacitors on 6H-SiC is characterized by voltage ramping, constant current stress and the corresponding charge-to-breakdown values (Q_{bd}).

2. Experimental

The substrate used in this work was n-type 3.5 ° off-axis 6H-SiC supplied by Cree Research. Before each high temperature process, the wafers were cleaned in a mixture of H_2SO_4 and H_2O_2 at 140 °C for 15 minutes followed by a 2% HF dip for 30 seconds. Two n-type epitaxial layers were grown homoepitaxially with significantly different process parameters on the Si face of the wafers with a net donor concentration in the lower 10^{16} cm^{-3} range by a low pressure chemical vapor deposition process described elsewhere [6]. The gate oxide of the capacitors was formed at 1100 °C by dry thermal oxidation of the two epitaxial layers. In addition one substrate without epitaxial layer was oxidized in the same way.

* on leave to Infineon Technologies AG i. Gr., Technology Development Power Semiconductors c/o Siemens AG, P.O. Box 3220, D-91050 Erlangen, Germany
Tel: ++49 (0)9131 7-32072, Fax: ++49 (0)9131 7-23046, E-Mail: michael.treu@erls.siemens.de

The resulting oxide thicknesses were about 26 nm. For comparison capacitors on silicon were co-processed with an oxidation temperature of 1000 °C resulting in an oxide thickness of 31 nm. After the oxidation, a 350 nm thick polysilicon film was deposited in a low pressure chemical vapor deposition reactor and doped by a $POCl_3$ process at 950 °C for 30 minutes. Finally, circular gate dots with gate areas ranging from 1.7×10^{-4} cm^2 to 9.5×10^{-4} cm^2 were defined by photolithography and wet etching. The oxide quality of the MOS capacitors was characterized electrically by voltage ramping and constant current injection, both performed with positive gate voltage. For each test about 40 MOS capacitors were used. To check the reproducibility of the measurements all the capacitors on the SiC wafers were etched off after the first examinations and MOS capacitors were manufactured again on the same wafers with the same process parameters. The measurements for these capacitors revealed that all the presented results are dominated by the characteristics of the semiconductor material and not by contamination during processing.

3. Results and Discussion

The voltage ramping tests showed that the oxide current for all capacitors is dominated by Fowler-Nordheim tunneling. The mean breakdown field strengths of the capacitors on the different SiC wafers varied from 10.3 to 12.2 MV/cm. The corresponding current densities just before breakdown ranged from 35 to 400 mA/cm^2 for the capacitors on SiC. For the capacitors on silicon this current density was about 2.7 A/cm^2.
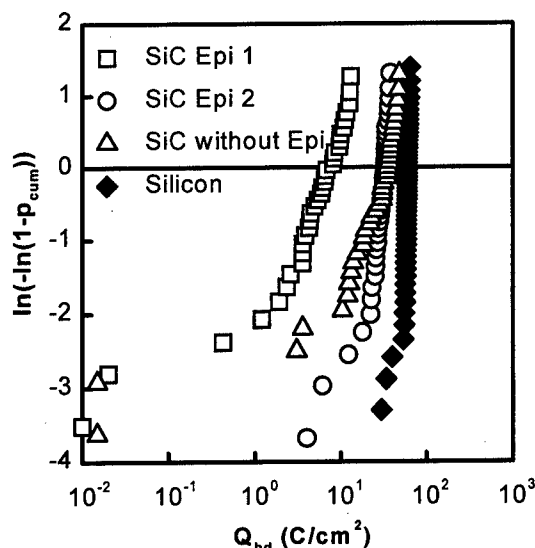
Fig. 1 Weibull plot of the charge-to-breakdown Q_{bd} for MOS capacitors on 6H-SiC and silicon. The injection current density was 5 mA/cm^2 with an area of 1.7×10^{-4} cm^2 for the capacitors on SiC and 10 mA/cm^2 with an area of 3×10^{-4} cm^2 for the capacitors on silicon. p_{cum} is the cumulative failure density.

Fig. 1 shows the Weibull plots of the charge-to-breakdown measured with constant current stress. The comparison of Epi 1 and the SiC without epitaxial layer shows that the epitaxial process can influence the quality of the later grown thermal oxide. This effect is most probably a consequence of the generation of additional crystal defects. The degrading effect of crystal defects on the oxide quality is a commonly accepted fact in the silicon technology [7].

The comparison of Epi 2 and the SiC without epitaxial layer shows that the epitaxial growth can be optimized to achieve as good results as without epitaxial layers. These results compare favorably with the data of the MOS capacitors on silicon.

The charge-to-breakdown for different gate areas is printed in Fig. 2. The Q_{bd} values of
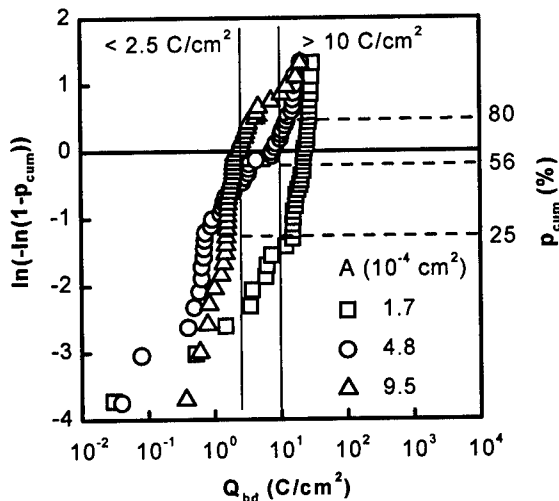
Fig. 2 Weibull plot of the charge-to-breakdown Q_{bd} for the MOS capacitors on Epi 2 with the gate area as a parameter. The injection current density was 10 mA/cm^2. The shift in the Weibull plots is caused by a defect density of 1.7×10^3 cm^{-2}.

Fig. 3 Q_{bd63} (63 % quantile of the cumulative failure densitiy of the Q_{bd} values) as a function of the injection current density J_{inj} for MOS capacitors on 6H-SiC and silicon.

these Weibull plots can be divided into two groups with a transition region in between: the first group contain values lower than 2.5 C/cm^2 and the second group contains values higher than 10 C/cm^2. Each group is characterized by a rather steep increase of p_{cum} in the Weibull plot. For increasing gate area the number of values below 2.5 C/cm^2 augments. From this behavior it can be concluded that:

1. There is a high density of defects which can be calculated according to Wolters et al. [8] by the relation

$$- \ln(1 - p_{cum,t}) = DA$$

where $p_{cum,\,t}$ is the value of p_{cum} in the transition region between the two groups, A is the gate area, and D is the defect density. The values of $p_{cum,\,t}$ noted on the right axis in Fig. 2 and the corresponding values of A result in a defect density of $D = 1.7 \times 10^3$ cm^{-2}.

2. There are defects which are correlated with one distinct value of Q_{bd}, leading to

steep segments in the Weibull plot. This means also that steep Weibull plots are not always a result of intrinsic oxide behavior.

3. The Q_{bd} values lower than 2.5 C/cm^2 are most probably related to crystal defects since the Q_{bd} values of the MOS capacitors on Epi 2 in Fig. 1 are nearly in the same range.

Fig. 3 shows the Q_{bd63} values (63 % quantile of the Q_{bd} values) as a function of the injection current density. For low injection current densities the MOS capacitors on the optimized epitaxial layer Epi 2 exhibit nearly identical long term reliability as the MOS capacitors on silicon. In contrast to this the higher density of crystal defects in Epi 1 obviously degrades the long term stability of the oxide also for lower injection current densities.

The Q_{bd63} values of the MOS capacitors on SiC decrease significantly if J_{inj} approaches a critical value, which is close to the current density where breakdown occurs in the E-

ramp test. Obviously the breakdown mechanisms change drastically for higher field strengths. Consequently it would not be serious to extrapolate the experimental data to lower injection current densities and field strengths in order to get information about the lifetime of thermal oxides on SiC under device operating conditions. It has to be noted that in our case such an extrapolation should be possible for injection current densities lower than 1 mA/cm^2, since in this regime the charge-to-breakdown is a weaker function of the injection current density and the gradient is comparable to the one observed for MOS capacitors on silicon. This means that tests providing exact values for the lifetime of the oxide under device operating conditions have to be performed at even lower values of the injection current densities than used in this work. However, these current densities will result in extremely long measurement times.

4. Conclusions

The presented data demonstrate that for small gate areas and electrical field strengths acting during practical use of MOS devices the long term reliability of thermal oxides on SiC and on silicon compare favorably. Indeed the reliability of the oxide is limited by the relatively high density of defects. These oxide defects are most probably related with crystal defects originating from the SiC substrate.

These data demonstrate the potential of SiC for the use as a semiconductor for commercial MOS devices. For practical use, however, the crystal quality has to be improved to achieve a good reliability for MOS devices with larger gate areas.

References

1. H.-F. Li, S. Dimitrijev, H. B. Harrison, and D. Sweatman, Appl. Phys. Lett., 70 (15), 2028 (1997)
2. L. A. Lipkin, and J. W. Palmour, J. Electr. Mat., 25 (5), 909 (1996)
3. J. W. Palmour, H. S. Kong, and R. F. Davis, J. Appl. Phys., 64, 2168 (1988)
4. P. Friedrichs, E. P. Burte, and R. Schörner, Appl. Phys. Lett., 65, 1665 (1994)
5. M. Treu, E. P. Burte, R. Schörner, P. Friedrichs, D. Stephani, and H. Ryssel, J. Appl. Phys., 84, 2943 (1998)
6. R. Rupp, P. Lanig, J. Völkl, and D. Stephani, J. Crystal Growth 146, 37 (1995)
7. Y. Takano, and N. Fuma, Electrochem. Soc. Proc., 94, 44 (1994)
8. D. R. Wolters, and J.J. van der Schoot, Philips J. Res. 40, 115 (1985)

Degradation of 6H-SiC MOS Capacitors Operated at High Temperatures

M. Bassler[a], V.V. Afanas'ev[b], G. Pensl[a], and M. Schulz[a].

[a] Institute of Applied Physics, University. of Erlangen-Nürnberg, Staudtstr. 7, D-91058 Erlangen, Germany
[b] KU Leuven, Lab. for Semicond. Phys., Celestijnenlaan 200D, B-3001 Leuven, Belgium

6H-SiC MOS capacitors were operated at temperatures above 600K under negative bias. Enhancement of energetically shallow and deep interface states at n/p-type SiC/SiO_2 structures and of a fixed charge are observed, which can partially be passivated by a hydrogen treatment. The generation and passivation of the fixed charge is explained in the framework of the "negative-bias-temperature instability" originally proposed for Si-based MOS capacitors.

1. INTRODUCTION

It has been demonstrated that silicon carbide (SiC)-based MOS FETs result in promising properties (see e.g. [1,2]), although a comparable quality as for Si/SiO_2 interfaces is not yet reached. Bano et al. [3] and Afanas'ev et al. [4] have shown that the electronic properties of the SiC/SiO_2 interface and the oxide layer can be degraded by Fowler-Nordheim injection and by UV illumination, respectively. The crucial open question is whether SiC-based MOS transistors will withstand elevated temperatures. We have, therefore, investigated 6H-SiC MOS capacitors, which are operated at temperatures above 600K. It turns out that charged defect centers are additionally formed by these heat treatments. The charge can partially be removed by a hydrogen passivation at 450°C for 20min. In analogy to silicon [5], we have proposed a „negative-bias-temperature instability (NBTI) model" for SiC, which explains the experimental results obtained for n/p-type MOS capacitors.

2. EXPERIMENTAL

MOS structures were prepared on n/p-type 6H-SiC epilayers (Si-face) with a net concentration of $N_N-N_{Comp}=2x10^{16}cm^{-3}$ and $N_{Al}-N_{Comp}=1x10^{16}cm^{-3}$, respectively. The samples were subjected to a pre-oxidation surface cleaning described elsewhere [6]. Wet oxidations were conducted at 1120 °C between 12h and 24h, leading to an oxide thickness in the range from 54nm to 110nm. The samples were subsequently exposed to a post-oxidation anneal in Ar at 1120 °C for 1 h. Gold contacts ($\varnothing = 0.85$ mm) were prepared by thermal evaporation. For the degradation, the samples were heated up to temperatures T_a ranging from 550K to 670K in a He ambient and were simultaneously biased at -20V for 10min/20min. Subsequent to each individual heat treatment, the samples were cooled down to room temperature and capacitance-voltage (C-V) and conductance-voltage (G-V) measurements were performed. Subsequent to this analysis, the samples were hydrogen-passivated in a rapid isothermal annealing system operated by tungsten halogen lamps in a $H_2(5\%)/N_2(95\%)$ ambient. The interface state density D_{it} and the generated fixed charge ΔQ_f are evaluated from the G-V/C-V characteristics.

3. RESULTS

C-V and G-V characteristics of an n-type 6H SiC MOS capacitor are shown in Fig. 1. The characteristics are taken subsequent to degradation steps conducted sequentially at different temperatures T_a (600K (curve a) to 650K (curve f)) under bias of V_i=-20V (oxide field: E=1.6MV/cm); they are measured at room temperature (300K), each curve is started at negative bias (accumulation condition). Two features are observed: Firstly, the C-V characteristics shift to positive voltages, which reflects an increase of negative charges either trapped in the SiO_2 layer or at the SiC/SiO_2 interface.
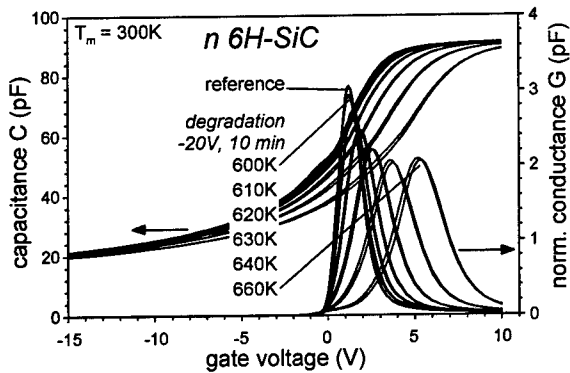
Figure 1. C-V and G-V characteristics of an n-type 6H SiC MOS capacitor taken each at room temperature subsequent to the degradation process. The degradation process was caused under depletion bias of $V_i=-20V$ for 10min at temperatures T_a ranging from 600K to 660K.

Secondly with increasing heat treatment, the conductance peaks (see G-V characteristics) smear out due to increasing surface potential fluctuations suggesting that degradation occurs in a laterally non-uniform manner possibly with participation of interface imperfections. In addition, the amplitude of the conductance peaks decreases. In analogy, degradation of p-type 6H-SiC MOS capacitors

results in an additional positive charge (not shown).

The interface state density D_{it} generated in an n-type MOS capacitor as a function of energy is shown in Fig. 2 (open squares); it is compared with the non-degraded reference sample (full circles). In the investigated energy range close below the conduction band edge, the interface state density D_{it} increases with increasing temperature. Corresponding results are obtained for the p-type MOS capacitor, there a strong increase of D_{it} is observed close above the valence band edge (not shown).

Fig. 3 reveals the additional negative/positive fixed charge as a function of T_a generated in an n/p-type MOS capacitor by the degradation process. A detectable increase of the fixed charge starts at a temperature of 600K.

We like to point out that degradation does not occur in neighboring MOS capacitors, which are exposed to identical temperatures T_a, however, which are not stressed by a depletion bias.

In order to trace the effect of hydrogen on the additionally generated fixed charge, degraded n/p-type MOS capacitors were exposed to a hydrogen passivation process conducted at 450°C for 20min. In Figs. 4 a)/b) C-V characteristics of an n/p-type MOS capacitor are displayed for three experimental

Figure 2. Interface state density D_{it} as a function of energy E_{it} for n-type SiC: D_{it} is determined for the reference sample (full circles) and subsequent to sequentially repeated degradation steps (open squares).
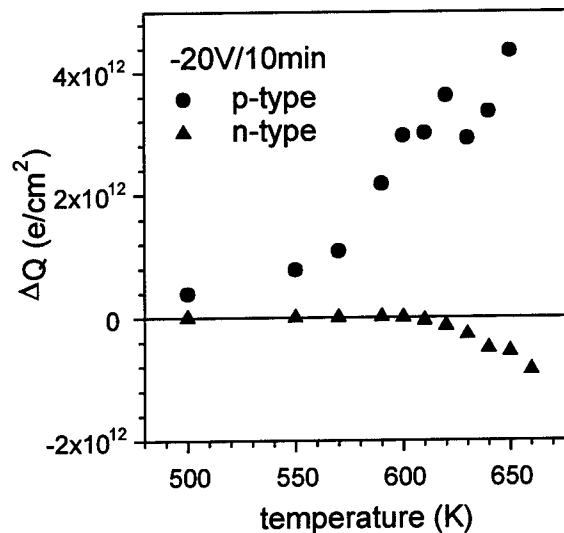
Figure 3. Comparison of the additionally generated fixed charge ΔQ_f for an n/p-type SiC MOS capacitor as a function of T_a for sequentially repeated degradation steps; ΔQ_f is obtained from the flatband voltage.

Table I. Reduction of the additionally generated fixed charge ΔQ_f in a degraded n/p-type MOS capacitor by subsequent hydrogen passivation; ΔQ_f is evaluated from the flatband voltage.

	ΔQ_f(bias) (e/cm^2) after 10min, -20V, 650K	ΔQ_f(H$_2$) (e/cm^2) after 20min, H$_2$, 450°C	$\dfrac{\Delta Q_f(H_2)}{\Delta Q_f(bias)}$
n-type	-2.0×10^{12}	$+1.1 \times 10^{12}$	0.55
p-type	$+1.2 \times 10^{12}$	-0.6×10^{12}	0.50
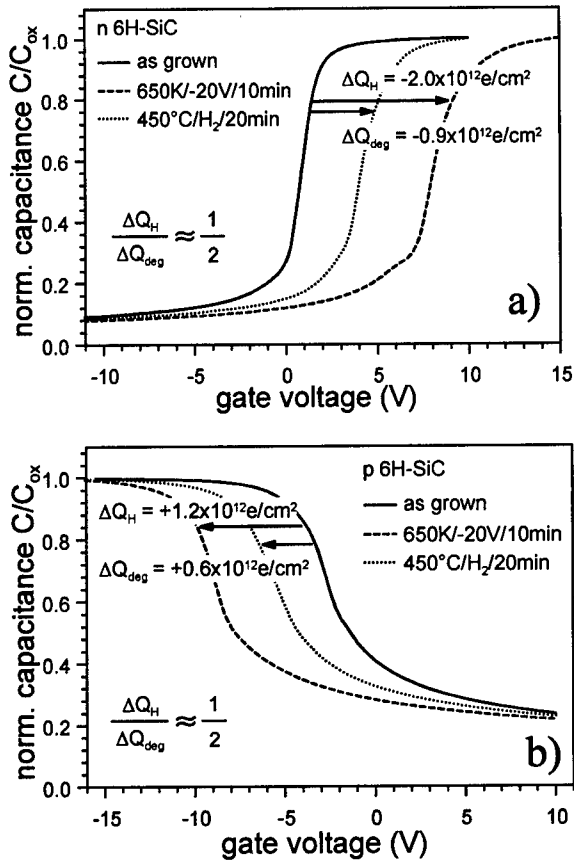
Figure 4. C-V characteristics for (a) an n-type and (b) a p-type 6H-SiC MOS capacitor. Three experimental situations are monitored: as-grown (solid curve), degraded (dashed) and hydrogen passivated (dotted). As indicated in the figure approximately half of the fixed charge is passivated independent of the sign of the charge.

situations: as-grown (solid curve), degraded (dashed curve) and hydrogen passivated (dotted curve). The passivation step leads to a reduction of approximately half the generated fixed charge independent of its sign. The passivation of the charge occurs rapidly, almost the entire passivation process is completed after 2min. The experimental results of the passivation process are summarized in Tab. I. From conductance spectra taken on passivated n/p-type MOS structures (not shown), information is obtained that energetically shallow interface states are only little affected by the hydrogen passivation. The origin of these shallow

states is supposed to consist of an unknown carbon species.

4. NEGATIVE-BIAS-TEMPERATURE INSTABILITY FOR SiC

The observed degradation of the electronic properties of SiC/SiO$_2$ capacitors and the passivation of fixed charges are consistent with a model proposed for silicon. Blat et el. have demonstrated a "negative-bias-temperature instability (NBTI)" for n/p-type Si-based MOS capacitors [5]. These authors have stressed MOS capacitors by heat treatments at 600K under negative bias and have generated equal amounts of dangling bonds at the interface and of a fixed positive charge in the oxide layer introduced by a water-related species. By a subsequent hydrogenation step, the dangling bonds could completely be passivated while the fixed charge remained unaffected.

A quite similar experimental situation is observed for SiC-based MOS capacitors. The following findings are available: (1) degradation occurs by heat treatments above 600K under negative bias (see e.g. Fig.1), (2) a fixed negative charge is generated in n-type MOS capacitors, (3) a fixed positive charge is generated in p-type MOS capacitors, (4) hydrogen passivation removes half of the fixed charge independent of the sign of the charge (see Fig. 4).

In order to explain our results in the framework of the NBTI model, two additional assumptions are required for SiC-based MOS capacitors:

1) Dangling bonds at the interface of SiC/SiO$_2$ are energetically deep and can not follow the voltage sweep; as a consequence these states contribute to the fixed charge. These states can be passivated by
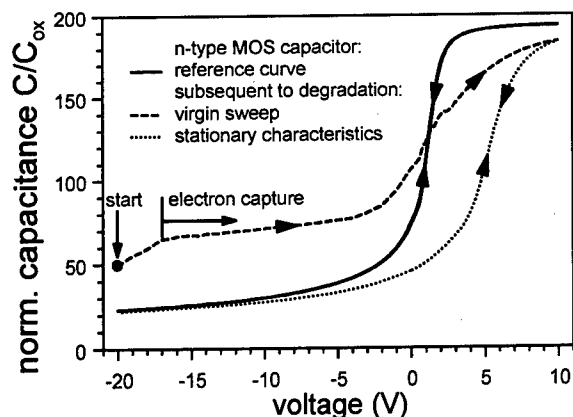
Figure 5. C-V characteristics of an n-type SiC MOS capacitor. The solid curve corresponds to the as-grown reference, the virgin curve (dashed) is taken on the degraded MOS capacitor starting at deep depletion and the dotted curve displays its stationary characteristic.

hydrogen, which leads to the reduction of half of the observed fixed charge as demonstrated in Figs.4a),b). The absolute energy position of dangling bonds at SiC/SiO₂ interfaces is not determined yet, however, our experiments indicate that these states are located deep in the band gap of SiC (E_{gap}=3eV). The sign of the fixed charge is negative in n-type SiC MOS capacitors and is positive in p-type SiC MOS capacitors while it is always positive in n/p-type Si-based MOS capacitors. It is well established [5] that trivalent Si acts as a dangling bond ($\equiv$Si') at the Si/SiO₂ interface and is amphoteric; depending on the type of majority carriers, it can be occupied by an electron or a hole.

2) In the case of SiC, we assume that the unknown water-related species located in the oxide, which is responsible for one half of the fixed charge, is also amphoteric.

We demonstrate this behavior in Fig. 5. The solid C-V curve is taken on the as-grown n-type SiC MOS capacitor. The dashed curve reveals the virgin curve of the degraded MOS capacitor, which was negatively charged during ramping down the temperature. In this special case, the measurement was started in deep depletion were no free electrons were available. As can be seen the increase of the C-V curve starts at negative bias (approximately -20V) indicating a large positive fixed charge. As soon as free electrons are available the deep dangling bonds

as well as the fixed oxide states trap electrons and turn to a negative charge state, which is stable under the conditions applied. The dotted characteristic corresponds to the stationary curve of degraded MOS capacitors. The NBTI model for SiC can be described in two steps:

$$\equiv\text{Si-H} + X + 2 \text{ holes} \leftrightarrow \equiv\text{Si}^+ + Y^+ \qquad (1)$$

$$\text{and } \equiv\text{Si}^+ + Y^+ + 4 \text{ electrons} \leftrightarrow \equiv\text{Si}^- + Y^- \qquad (2)$$
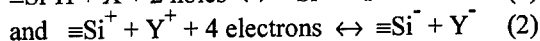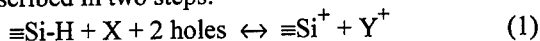
X, Y = unknown, water-related species in the oxide layer.

5. SUMMARY

We have demonstrated that heat treatments of SiC/SiO₂ MOS capacitors above 600K under negative bias generate energetically shallow and deep interface states and a fixed charge in the oxide layer. The degradation and passivation behavior of the fixed charge is consistent with the NBTI model originally proposed for Si/SiO₂ MOS capacitors.

6. ACKNOWLEDGEMENT

References

[1] J.A. Cooper, Jr., phys. stat. sol. (a) **162**, 305 (1997)

[2] S. Onda, R. Kumar, and K. Hara, phys. stat. sol. (a) **162**, 396 (1997)

[3] E. Bano, T. Ouisse, C. Leonhard, A. Gölz, E.G. Stein von Kamienski, Diamond and Related Materials **6**, 1489-1493 (1997)

[4] V.V. Afanas'ev, A. Stesmans, M. Bassler, G. Pensl, M. Schulz, and C.I. Harris, J. Appl. Phys. **85**, 8292 (1999)

[5] C.E. Blat, E.H. Nicollian, and E.H. Pointdexter, J. Appl. Phys. **69** (3), 1712 (1991)

[6] V.V. Afanas'ev, A. Stesmans, M. Bassler, G. Pensl, M. Schulz and C. I. Harris, Appl. Phys. Lett. **68**, 2141 (1996)

Oxidation of 6H-SiC(0001)

*L. Simon[a-b], L. Kubler[a], A. Ermolieff[b] and T. Billon[c]

[a]Laboratoire de Physique et de Spectroscopie Electronique, UPRES. A CNRS 7014
4 rue des Frères Lumière 68093 Mulhouse Cedex, France

[b]LETI (CEA-Technologies Avancées), SCPM/DOPT-CEA/G
17, rue des Martyrs, 38054 Grenoble cedex 9, France

[c]LETI (CEA-Technologies Avancées), DMITEC-CEA/G
17, rue des Martyrs, 38054 Grenoble cedex 9, France

In this work we intend to determine whether the SiO_2/6H-SiC(0001) interface is really carbon enriched. Among numerous publications dealing with 6H-SiC oxidation, our specificity resides in a simultaneous use of ultra-high vacuum surface cleaning and characterization methods (X-ray spectroscopy) and standard technological conditions for the oxidation i.e. dry oxidation at 1 atm of O_2 and a temperature of 1000°C. The oxidation of 3x3 Si-rich and $6\sqrt{3}x6\sqrt{3}R30°$ C-rich reconstructed surfaces is compared, taking care to transfer the samples without uncontrolled air exposures. No C-C or C-O bonds are observed at the SiO_2/SiC interface by C 1s core line analyses whatever the initial surface is. We conclude that there is no contraindication to obtain nearly ideal SiO_2/SiC interfaces on flat SiC terraces.

1. Introduction

Silicon carbide provides a lot of interesting properties such as a wide band gap, a large electron mobility as well as high physical and chemical stability. The material is therefore promised to be widely used in high temperature and high power applications [1]. High quality silicon carbide single crystal have been grown for several years, nevertheless research in SiC based devices have been slowed down by difficulties in obtaining good MOS (metal-oxide semiconductors) structures. Silicon carbide MOS capacitors have experimentally demonstrated that both n- and p-type structures exhibit a strong near-interfacial oxide charge. This high charge rate added to a high interface state density and interface disorder result in the existence of numerous localized states and in a thermally activated transport, even at room temperature [2].

A lot of authors [3-5] suspected a carbon enrichment at the SiO_2/SiC interface to be at the origin of the high interface state density. In all these papers, either the oxidation conditions (low oxidizing pressure) are far from those used in the

technology or, when real technological conditions are employed, the surface preparation procedures and the initial surface characterizations before oxidation are not clearly defined.

In this study we present XPS characterizations of the SiO_2/SiC interfaces realized in technological conditions i.e. using dry O_2 oxidation at a pressure of 1 atm and 1000°C. Particular attention is paid to control and characterize the surfaces before oxidation. We compare the results of the oxidation of two types of initial surfaces prepared in ultra-high vacuum conditions: a 3x3 reconstructed Si-rich surface and a $6\sqrt{3}x6\sqrt{3}R30°$ graphitized C-rich surface. The surface characterization and subsequent oxide growth is performed in situ without any uncontrolled air exposure. The growth kinetics observed as a function of O_2 exposure time are discussed in terms of enhanced carbon exodiffusion.

2. Experimental procedures

6H-SiC(0001) single-crystals purchased from Cree Inc. and capped with a n-doped epilayer, are cleaned with acetone and alcohol in an ultrasonic bath. Then they are loaded in an ultra-high vacuum

preparation chamber (10^{-10} mbar). To remove the surface native oxide and hydrocarbide impurities, we use a cleaning technique by which the samples are annealed at 800°C during a few minutes under a silicon flux. Samples are resistively heated. This cleaning procedure removes the impurities and, by LEED, we observe a 3x3 surface reconstruction. The other 6√3x6√3R30° reconstructed surface is obtained by holding the samples at 1250°C during several minutes without silicon flux.

All XPS spectra are acquired using an MgKα radiation. Each reconstructed surface is analyzed before the oxidation. Afterwards, they are moved under ultra-high vacuum in another preparation chamber and isolated. The samples are kept at room temperature during the introduction of oxygen (O_2) though a leak valve. When the gas pressure is about 10^{-2}mbar the sample is quickly heated to 700°C. During the time necessary to stabilize the O_2 pressure at 1 atm (30 seconds), the temperature is increased to 1000°C. At the end of the oxidation, the temperature is very rapidly reduced and the gas is pumped down. Finally a base pressure of 10^{-9} mbar is restored after 6 minutes. The sample is then moved to the analysis chamber without disrupting the ultra-high vacuum environment.

3. Results

3.1. Initial surfaces

Fig. 1 shows the C 1s and Si 2p spectra recorded on the two different reconstructed surfaces at the polar angle θ = 50° (mainly sensitive to the surface states). The peaks are fitted by a standard Voigt function.

For the 3x3 reconstructed surface, the Si 2p peak presents two components (Fig.1.c). The dominant one is attributed to the SiC substrate. The other one located at 1.2 eV towards the higher kinetic energy, is attributed to Si-Si overlayer bonds and reveals a Si-enrichment of the surface. The C 1s peak only presents one single component (Fig.1.a). This Si-rich 3x3 reconstructed surface has been extensively studied in the past. STM and holographic LEED studies have shown that this reconstruction corresponds to a surface completely covered by Si clusters extending over three monolayers of silicon atoms[6]. After annealing at 1250°C during several minutes without Si flux, a 6√3x6√3R30° LEED pattern is obtained. The Si 2p peak (Fig.1.d),

presents only one component and the Si surface enrichment is suppressed. The C 1s peak (Fig.1.b) now shows three components: one named B_1 corresponds to the SiC substrate, and two others named S_1 and S_2 correspond to C-C bonds. Two similar surface components were observed by Johansson et al[7] in C 1s core-level peak using synchrotron radiation. They assign these components to two types of C-C bonds in a C-rich environment defining a fingerprint of a graphitic surface layer.
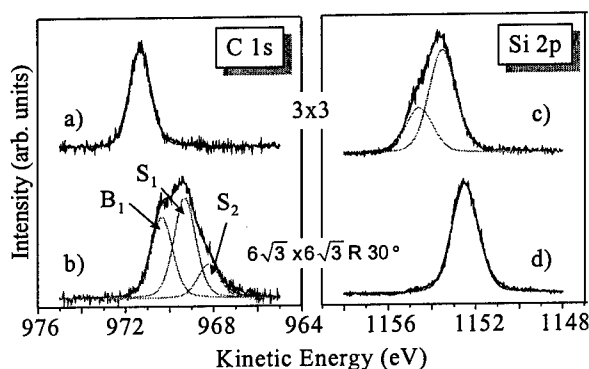
Fig. 1. C 1s and Si 2p core-level peaks recorded on the initially Si-rich (a and c) and C-rich (b and d) surfaces at the polar angles θ = 50°.

3. 2. Thermal oxide growth

Fig. 2 shows the evolution of the C 1s XPS peaks recorded at normal emission as a function of oxygen exposure time at a temperature of 1000°C.

The initial surface is the preceding graphitized one and the C 1s peak presents the three components previously discussed. In spite of this initial highly C-enriched surface, the C 1s core-level peak is reduced during the first minute of the oxidation to a unique component with a half-maximum (FWHM) of 1.1 eV. Meanwhile the intensity of the C 1s peak decreases. The signal intensity is divided by four for a longer exposure time (14.5 minutes). At the same oxidation stage, at grazing incidence and using a pass energy of 20 eV, the C 1s photoemission intensity becomes lower than the detection limit. This indicates an homogeneous formation of the oxide on the overall surface and a lack of carbon atom in the oxide layer.

Before oxidation, the Si 2p peak exhibits only one component. In the first minute of oxidation, in

addition to the main substrate component, a second component associated with the Si-O bonds in the oxide layer starts increasing. For larger oxidation times the latter becomes dominant and a final oxide thickness of 5 nm is estimated. For this largest oxide thickness the Si 2p component in SiC has an FWHM of 1.2 eV, comparable to that of the clean surface. The FWHM of the oxide component does not exceed 1.8 eV. This probably rules out possible interfacial suboxides like those observed on native oxide by Onneby et al[8]. The Si 2p binding energy shifts between the Si-O and Si-C (or Si-Si) components are 2.5 eV (3.7), respectively. The latter value is close to that obtained in the case of the oxidation of silicon substrates [9].
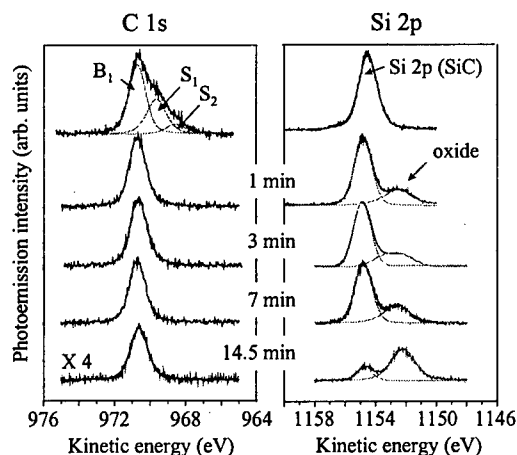
Fig. 2. Evolution of the C 1s and Si 2p core level peaks as function of the oxygen exposure time of the initially graphitized surface.

All these XPS features are unchanged for the oxidation of the 3x3 reconstructed surface and are therefore not repeated.

3. 3. Growth kinetics

Fig. 3 shows the evolution of the oxide growth kinetics on three different samples, corresponding to a graphitized and two Si-rich initial surfaces. The kinetics are reflected by the plot of the intensity ratio (I-ratio) between the respective Si 2p components in SiO_2 (I (Si 2p/SiO_2)) and in SiC (I(Si 2p/SiC)). The I-ratio remains nearly constant during the first 7 oxidation minutes where the calculated oxide thickness is around 0.7 nm. The kinetics are identical

whatever the initial surface is, the graphitic bonds of the C-rich surface being removed during the first oxidation minute. The exodiffusion of the carbon atoms is thus very fast. After this first stage of the oxidation, nearly linear kinetics are taking place. The change of the oxide growth rate is probably due to a change of the mechanism of the oxidation. A similar kinetic shape was previously observed by Muehlhoff et al[10], and the retarded oxidation was attributed to the presence of C-C bonds on the initial surface. Our control of the surface before and after oxidation shows that the carbon enrichment of the initial surface is by no means the reason. It is probably connected to the mechanism of the initial oxidation stage. Growth morphology effects such as an inhomogeneous nucleation of oxide patches up to their coalescence within the first 7 oxidation minutes might also explain this behavior. The linear growth regime would only start after a complete covering of the overall surface by a dense oxide layer. The latter growth rate may be attributed to a Mott-Cabrera mechanism [11] that will be discussed below.
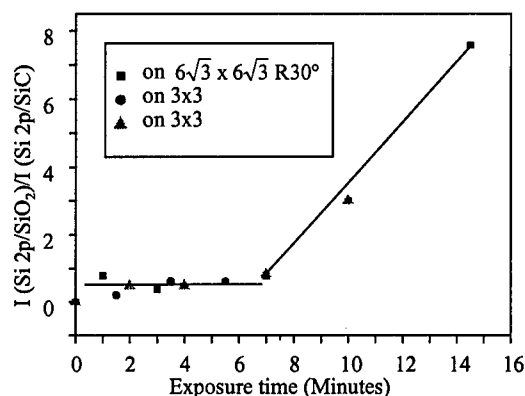
Fig. 3. Evolution of the Si 2p ratio (I(Si 2p/SiO_2)/(I (Si 2p/SiC)) in oxide and carbide respectively, as a function of O_2 exposure time.

4. Discussion

The observation of a single, narrow (1.1 eV) and non convoluted C 1s component at the SiO_2/SiC interface is a new result in the literature. It implies a complete absence of C-C or C-O bonds and characterizes a prominently simplified and abrupt interface. This is in contradiction with Hornetz et al

[4] who, in spite of comparable energy spectrometer resolutions, have found five components in their C 1s peak and explained their results by the presence of sp^2 bonds in a complicated interface layer implying silicon, carbon and oxygen atoms. We can assert that our C 1s line is really a homogeneous one which only reveals the substrate contribution. An ideal SiO_2/SiC interface should be reducible to the last Si layer on top of the SiC substrate connected to the first oxygen atoms of the oxide like SiO_2 onto Si(111). As suggested by Onneby et al [8] and Hornetz et al [4] such an interface would be formed by $Si-(OC_3)$ species linking the first oxygen atoms of the oxyde layer with silicon atoms themselves back-bonded with three carbon atoms beneath the top substrate layer. According to these authors, these particular silicon atoms would provide a weak additional Si 2p component located between the Si 2p (SiC) and Si 2p (SiO_2) components. Our single C 1s feature, excludes connections of the oxide with any C substrate terminations. Concerning the Si 2p feature, within our detection limit, we did not observe possible suboxide components in addition to the two main SiC and SiO_2 components for sample thicknesses in the linear growth regime (Fig. 3). With similar analyzer energy resolution than these authors, as for C 1s, the absence, or at least, the presence of weaker suboxide components may be related to a more abrupt SiO_2/SiC interface in our case. We think that the interface we have obtained is reducible to an ideal monoatomic connexion layer for which the residual Si-O interfacial contribution to the Si 2p peak stays below our detection limit.

Concerning the growth kinetic, drawing a parallel with Si oxidation, the linear regime we observe in Fig. 3 may speculatively be compared with the well known Mott-Cabrera mechanism [11]. The latter also concerns dry O_2 oxidation and involves interfacial space charge effects resulting in a modified diffusion of charged oxidizing species within a given Debye oxide thickness ranging from 100 to 200 Å. We can assume that the thin oxides we have grown here are similarly submitted to a charge injection disappearing for higher oxide layers. Then the nature of the oxidation agent may thus be thickness dependent and possibly, in the case of SiC, determining in the formation of C-O bonds able to exodiffuse. The carbon enrichment of the oxide/SiC interface could be oxide thickness dependant.

5. Conclusion

The initial stage of oxide growth using technological conditions (dry oxidation at 1 atm and 1000°C) has been probed by X-ray Photelectron Spectroscopy (XPS) measurements. We compare the results of the oxidation of two types of surfaces i.e., reconstructed 3x3 Si-rich and $6\sqrt{3}x6\sqrt{3}R30°$ graphitized C-rich surfaces. In both cases neither C-C bonds nor C-enrichment could be detected after oxidation at the SiO_2/SiC interface. With the initially C-rich graphitized surface the C-C bonds are exodiffused during the first minute of the oxidation. These results associated with the absence of any discernible suboxide component in the Si 2p signal are indicative of a first elaboration of abrupt, nearly ideal, SiO_2/SiC interfaces. Possible causes of the previously observed C-enrichments are discussed.

REFERENCES

1. J. A. Powell, D. J. Larkin, P. B. Abel, L. Zhou and P. Pirouz, Inst. Phys. Conf. Ser. **142**, 77 (1996).
2. E. Bano, T. Ouisse, P. Lassagne, T. Billon and C. Jaussaud, Inst. Phys. Conf. Ser. **142**, 729 (1995).
3. V. M. Bermudez, Appl. Surf. Sci. **84**, 45 (1995).
4. B. Hornetz, H. J. Michel and J. Halbritter, J. Mater. Res. **9**, 3088 (1994).
5. V. V. Afanasev, M. Basler, G. Pensl and M. Schulz, Phys. Stat. Sol. A. **162**, 321 (1997).
6. U. Starke, J. Schardt, J. Bernhardt, M. Franke, K. Reuter, H. Welder, K. Heinz, J. Furthmüller, P. Käckell and F. Bechstedt, Phys. Rev. Lett. **80**, 758 (1998).
7. L. I. Johansson, F. Owman and P. Martensson, Phys. Rev. B, **53**,13793 (1996).
8. C. Onneby and C. G. Pantano, J. Vac. Sci. Technol. A **15**, 1597 (1997).
9. G. Hollinger, Appl. Surf. Sci. **6**, 318 (1981).
10. L. Muehlhoff, M.J. Bozack, W. J. Choyke T. John and Jr. Yates, J. Appl. Phys. **60**, 2258 (1986).
11. *The Si/SiO2 system*, Edited by Peter Balk, Elsevier, Mat. Sci. Monograph, **32**, p39.

Electrical characterization of the amorphous SiC-pSi structure

V.S.Lysenko, I.P.Tyagulski, Y.V.Gomeniuk, I.N.Osiyuk, and A.K.Mikhnov

Institute of Semiconductor Physics, National Academy of Sciences of Ukraine,
45 Prospekt Nauki, 252650 Kiev, Ukraine

Forward and reverse currents in the amorphous SiC/pSi structure were studied in the temperature range from 30 to 80 K. Charging of electron traps located in the SiC layer increases the forward current due to reduction of the potential barrier for holes. The current was found to be controlled by the Poole-Frenkel emission via the system of levels in the SiC layer located at 0.048 eV and 0.065 eV above the edge of the valence band.

1. INTRODUCTION

Silicon-on-Diamond (SOD) material is the promising one for some SOI applications due to an outstanding heat conductivity and a high resistivity of diamond [1]. However, known techniques for deposition of diamond films on the silicon substrate do not provide a perfect Si/C interface [2]. The silicon/diamond interface is characterized by the presence of the highly defective transition layer which is believed to be composed, at least partially, of amorphous SiC (a-SiC) and is known to dramatically affect long-term reliability and noise characteristics of SOI-based devices.

Investigations of the a-SiC-pSi heterostructure using thermally stimulated charge release (TSCR) technique [3] have shown that the SiC layer contains the system of shallow levels near the bottom of the conduction band with activation energies of 0.08, 0.16 and 0.22 eV.

In this paper results of measurements of forward and reverse current-voltage ($I–V$) and current-temperature ($I–T$) characteristics in a-SiC-pSi heterojunctions in the temperature range from 30 to 80 K are presented, and the effect of the charge state of electron centres in the SiC layer on the transport current is studied.

2. EXPERIMENTAL

The SiC layer was deposited onto the silicon wafer using a low-pressure chemical vapour deposition (LPCVD) technique in an atmosphere of methyltrichlorosilane at temperature 1000°C. The thickness of the SiC layer was 1.1 μm. Measurements were carried out on the Al-SiC-pSi structure. The aluminium capacitor's area was 0.001 cm². The character of the current transport was studied either by measuring the set of $I-V$ characteristics at several values of temperature, or by measuring the set of temperature dependencies of the current ($I–T$ characteristics) at a fixed bias voltage applied to the structure [4, 5].

Filling of the system of electron traps located in the silicon carbide layer was performed by applying a positive bias to the aluminium electrode at a fixed temperature (about 120 K). Cooling the sample to a lower temperature and switching off the filling voltage turns the system into the non-equilibrium state with a fraction of charge being trapped in the centres of the SiC layer.

3. RESULTS AND DISCUSSION

3.1. Forward I-V and I-T characteristics

Figure 1 shows I-V characteristics of the forward-biased heterostructure measured at 40 and 80 K with and without filling the electron system of centres in the SiC layer. It can be seen in the figure that each curve consists of three sections with the substantially different character of the dependence of current on

bias and temperature. In the low-voltage region, the current rises rather slowly (section I), then a sharp increase of the current is observed (section II), and in the higher voltage region of section III a smaller slope of the $\log I$ vs. V dependence is recorded. Such a behaviour reflects the existence of different mechanisms of the transport current control with increasing forward bias.
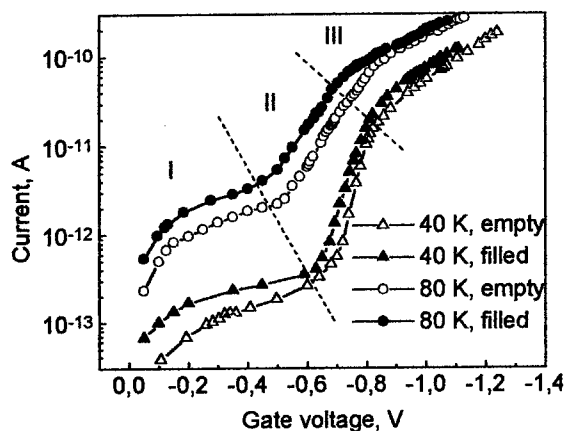
Figure 1. Forward I–V characteristics at the temperatures 40 K (Δ) and 80 K (O) for empty (open symbols) and filled (full symbols) electron states.

Filling of the electron traps results in an increase of the current in the whole investigated voltage range. It was shown in [3] that at temperatures below 80 K the major part of the electron charge remains trapped. The fact that the forward current increases after filling of the traps proves that the dominant component of the forward current is the hole current. In this case the negative charge of electrons trapped in the SiC layer lowers the potential barrier for the hole transport from the silicon substrate through the heterointerface and the silicon carbide layer into the metallic contact.

If the absolute magnitude of the negative bias applied to the aluminium electrode increases, different situations will eventually appear. At low forward bias the band diagram does not differ significantly from the classical Anderson energy band diagram for heterojunctions. At some bias $V=V_{FB}$ the negative voltage at the gate electrode compensates the built-in diffusion potential and the flat band conditions for the semiconductor are reached.

Since doping of the wide-bandgap semiconductor

is low, its resistivity is high at temperatures below the liquid nitrogen temperature, therefore, an amorphous silicon carbide layer can be considered as a dielectric with leakage. As the forward voltage is further increased, the voltage drop across the silicon part of the heterojunction remains practically unchanged, because the excess of holes in the accumulation layer near the interface is removed to the amorphous layer. The portion of the bias exceeding the value which is spread over the silicon substrate appears to be applied to the silicon carbide layer and when the magnitude of the negative gate bias increases, the voltage drop across the amorphous film increases. In this case the single-crystalline silicon plays a role of a source of charge carriers and the current is determined uniquely by the transport processes in the amorphous material.

Thus, as the forward bias increases, different situations may occur, so that different theoretical models of transport should be drawn to explain the observed behaviour.

The fact that the slope of logarithm of the current vs. voltage dependence at low forward biases, is practically independent of temperature (see figure 1) can be considered as an evidence of the tunnelling mechanism of the transport current. As was shown in [6], the most probable transport mechanism for such heterojunctions is the multi-step tunnelling process via the localized states in the semiconductor bandgap. According to this mechanism the current is described by the relation: $J=J_0\exp(V/V_0)\exp(T/T_0)$, where V_0 and T_0 are constant parameters.
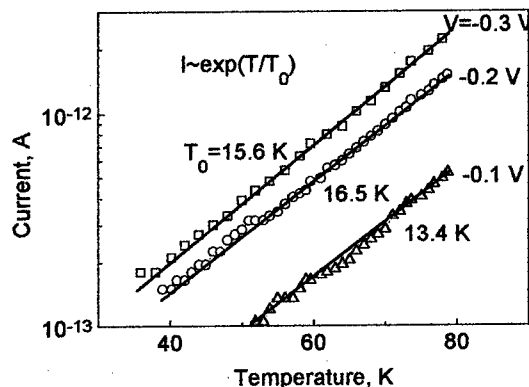
Figure 2. Temperature dependencies of forward currents at small biases.

Figure 2 shows the temperature dependencies of the current at low forward biases. One can notice that an exponential dependence of the current on temperature is really observed in this part of I–V curves. The characteristic temperature T0 ranges between 13 and 16 K.

If the forward bias exceeds the flat-band voltage V_{FB}, the onset of an additional transport mechanism for holes takes place. One of such mechanisms is the conduction via the system of centres located near the edge of the bandgap by the thermally activated capture-emission process with account of barrier lowering in the electric field. This mechanism is known as the Poole-Frenkel emission and was initially suggested to explain the forward current-voltage characteristics of amorphous solids. According to this mechanism the current must obey the law [7]:

$$J \propto E \exp\left(\frac{-q(\varphi_B - \sqrt{qE/\pi\varepsilon_i})}{kT}\right),$$

here φ_B is the depth of the trap's potential well, E is the electric field in the amorphous material, ε_i is the static dielectric permittivity of the material.

Assuming that the electric field in the amorphous material is constant, $E = |V - V_P|/d$, where d is the thickness of the silicon carbide layer, V_P is the voltage of the onset of the Poole-Frenkel transport mechanism. Then the plot of the $ln\{J/|V-V_P|\}$ vs. $1/T$ dependence is a straight line, its slope gives an effective activation energy E_A. The activation energy, in turn, will be a linear function of $|V-V_P|^{1/2}$, giving the value of the trap depth φ_B at the intercept with the ordinate axis. The magnitude of V_P should depend on a number of factors, for example, on the fraction of the diffusion potential over the amorphous part of the heterostructure, density and energy distribution of surface electron states at the amorphous-crystalline interface, the magnitude of the built-in charge, etc.

Figure 3 shows the temperature dependence of the forward current for several values of the voltage corresponding to section II of the current-voltage

characteristic. The dependencies were measured at empty centres. Two temperature regions, firstly, from 40 to 60 K, and, secondly, from 60 to 80 K, were considered, within which the fitting could be performed with a reasonable accuracy. From the slope of the dependencies the activation energy was determined, which then was plotted as a function of $|V-V_P|^{1/2}$ (V_P was taken to be 0.59 V). The linear dependence of the calculated activation energy on the square root of the electric field indicates that the assumption about the Poole-Frenkel transport mechanism is valid. The respective values of trap well-depths are 48 and 65 meV.
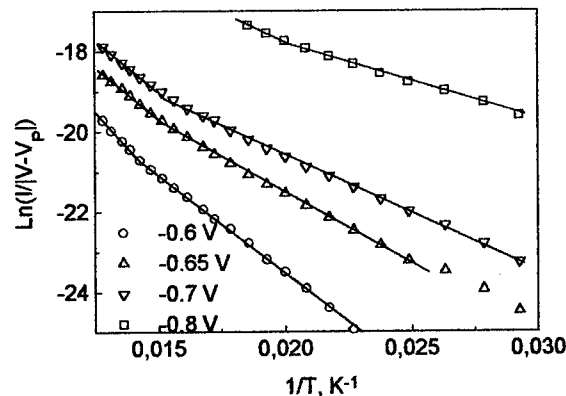
Figure 3. Dependence of $ln(I/|V-V_P|)$ vs. reciprocal temperature for different forward biases.

Our previous investigations [3] have shown that near the bottom of the conduction band of SiC the levels with activation energies of 0.08, 0.16 and 0.22 eV are situated. Investigations of the lower half of the bandgap were not performed because it is impossible to fill hole states by a forward bias due to a high transport current. The above presented results indicate that near the top of the valence band the levels are located with activation energies 0.048 and 0.065 eV.

Consider part III of the current-voltage characteristics (figure 1), where the rise of current becomes more gradual. In this part of the curve the slope of $I-V$ characteristic, in semi-log scale, is independent of temperature. Therefore, we can suggest that the main transport current mechanism here is tunnelling. It should be mentioned, that at temperature 80 K the dependencies for empty and filled states coincide. At such temperatures the electrons trapped in the SiC layer are released into

the conduction band and are swept into the silicon substrate, because sufficiently large forward bias reduces the effective barrier for their release. As a result, the structure turns to its initial condition with empty electron states.

3.2. Reverse I-V characteristics

If a reverse bias is applied to the structure ($V_g > 0$), then at low temperature in the p-type silicon substrate the non-equilibrium depletion is formed and the major part of the applied voltage drops at the silicon substrate.

Measurements of currents at a reverse bias have shown that if the electron centres are initially empty, the magnitude of the reverse current varies during the measurement, that is, at a constant temperature and bias, the current slowly (with a time constant of around several minutes) decreases approaching to some steady-state value. If the trap filling is performed before the measurement, the change in the current is not observed. This effect was attributed to the fact that the negative charge of electron centres in silicon carbide creates a barrier for carrier transport and causes the reduction of the current. This reduction of current proves that the reverse current is the current of electrons from the silicon substrate to the silicon carbide layer and further to the metal electrode.
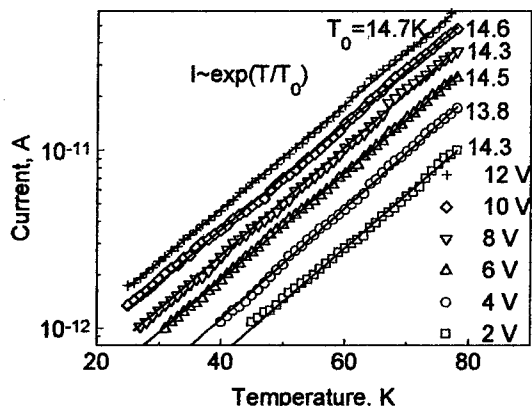
Figure 4. Temperature dependence of the reverse current at different bias.

Figure 4 shows temperature dependencies of the reverse currents measured at different reverse biases for empty traps. It is seen that the reverse currents can be described by the same relation as the forward

currents at low voltages (see figure 2) which is the signature of a multitunnelling character of the current. The characteristic temperature T_0 is practically the same as that obtained for the forward currents.

4. CONCLUSIONS

From the study of the forward and reverse currents in the a-SiC/pSi heterojunction in the temperature range from 30 to 80 K a system of electrically active centres has been found possessing the levels near the valence band with activation energies 0.048 and 0.065 eV. At a reverse bias and a low (less than 0.3 V) forward bias transport is determined by the process of multi-step tunnelling of electrons and holes, respectively, via the states in the bandgap of amorphous silicon carbide. The transport current in the heterojunction at forward biases ranging from 0.5 to 0.8 V is controlled by the Poole-Frenkel emission of holes. At higher voltages the transport mechanism changes to the tunnelling one.

Filling of the system of electron traps located in the silicon carbide layer causes a reduction of the potential barrier for holes and, as a result, leads to the increase of the forward current. The reverse current determined by the transport of electrons in this case decreases.

REFERENCES

1. B. Edholm, A. Soderbarg, J. Olsson and E. Johansson, Jpn. J. Appl. Phys., 34 (1995) 4706.
2. B. Edholm, J. Olsson, N. Keskitalo and L. Westling, Microelectronic Engineering, 36 (1997) 245.
3. V.S. Lysenko, I.P. Tyagulski, Y.V. Gomeniuk, I.N. Osiyuk, and I.I. Tkach, Journ. of Phys. D: Applied Physics, 31 (1998) 1499–1503.
4. H. Chaabane, J.C. Bourgoin, Appl. Phys. Letters, 64 (1993) 1006–1008.
5. H. Chaabane, J.C. Bourgoin, J. Appl. Phys., 76 (1994) 315–318.
6. A.R. Riben, D.L. Feught, Int. J. Electronics, 2 (1966) 583-599.
7. S.M. Sze, Physics of Semiconductor Devices, 2nd Ed., Wiley, New York, 1981.

Dangling Bond Defects in SiC: the Dependence on Oxidation Time

P.J. Macfarlane and M.E. Zvanut
University of Alabama at Birmingham, Dept of Physics, Birmingham, AL, 35294-1170

Electron paramagnetic resonance is used to study a near surface defect in oxidized 4H/6H SiC substrates and 3C SiC epitaxial layers. The defect is observed after wet oxidation and a 900° C dry thermal treatment, but the defect is not located in the oxide. The heat treatment activates the EPR center by removing a hydrogen-related species from a C bond. Oxidation studies suggest that reaction of H_2O with SiC creates the defect. However, an alternative theory in which the defect is intrinsic to SiC cannot be disregarded.

1. INTRODUCTION

SiC remains a focus of much research because it could replace Si in specialty electronic applications involving high temperature and high power [1]. For any application, surface quality is critical for successful device operation. This paper addresses a near-surface defect in SiC that we suggest is created by oxidation. The defect is observed using room temperature electron paramagnetic resonance (EPR) spectroscopy on SiC substrates that have been oxidized and thermally treated.

EPR is one of the few spectroscopies that can measure defect densities on the order of parts per million, the level critical to microelectronic devices. The technique depends on the absorption of microwave energy at a defect that has been slightly perturbed by the application of a magnetic field [2]. The magnetic field at which absorption occurs is related to the g-value of the defect through the Zeeman interaction, $h\nu = g\mu_B B$. Figure 1 illustrates an X-band EPR spectrum obtained from an oxidized double-sided polished 6H substrate that was heat treated at 900° C for 200 min in dry N_2. The g-value, which is calculated from the point where the spectral amplitude crosses zero, is isotropic and has a value of 2.0026+/- 0.0001. The wealth of structural information typical of EPR spectra is obtained from nuclear hyperfine interaction which, for these samples, is below the detection limit. To study the center more thoroughly we have executed a series of annealing and oxidation studies [3]. The defect

characteristics may be summarized as follows:

1) The g-value is typical of C-related defects in polycrystalline diamond films, abraded SiC, and 3C epitaxial SiC films [3].

2) The signal is observed in oxidized SiC after dry heat treatment ($H_2O < 1$ ppm) in either an inert (N_2) or oxidizing (O_2) ambient. The temperature dependence for generation of the signal reveals an Arrhenius behavior and gives an activation energy of 4.0 eV. We suggest that the activated reaction is the removal of H from a C-dangling bond.
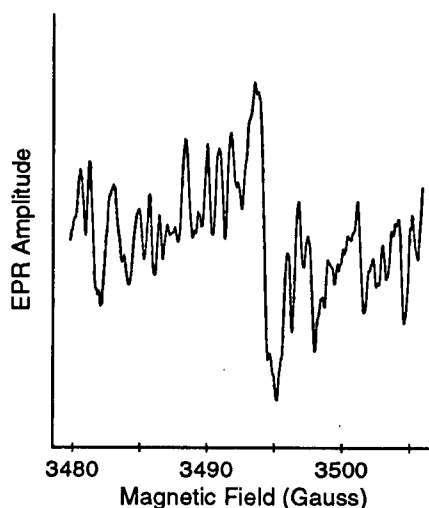
Figure 1. EPR spectra of oxidized double-sided polished (0001) 6H SiC after 900° C 200 min treatment in dry N_2 ($H_2O < 1$ ppm).

3) The signal is not generated after heat treatment in an ambient containing H_2O.

4) The concentration of the center does not change after the SiO_2 layer is removed using an $HF:H_2O$ solution; however steam oxidation decreases the density at least an order of magnitude.

Inherent to the description of the defect given above is the understanding that an EPR signal can result from either generation of a new defect or activation of a pre-existing defect. For example, we assume that the spectrum shown in figure 1 results from an activation process consisting of the depassivation of hydrogen or a hydrogen-related species from a C-dangling bond during dry heat treatment. The model is consistent with several results. The signal is obtained after a dry heat treatment, but is not observed after annealing in an ambient containing moisture [3]. The activation energy, 4.0 eV, is reasonable for dissociation of a CH bond [3]. An alternative idea, in which the heat treatment generates the defect, would be consistent with the process of Si evaporation known to occur at the surface of SiC under certain high temperature conditions [4,5]. However, we discount the Si evaporation model because ultra high vacuum conditions are necessary to observe the loss of Si, and there is no indication of evaporation when as little as 10 Torr of O_2 is added at $1000°$ C [4,5]. In our work, adding dry O_2 (H_2O < 1 ppm) to the heat treatment ambient does not alter the intensity of the EPR signal [6]. Having established the role of the dry heat treatment as an activation process, the work presented here examines the role of oxidation in generating the defect.

2. EXPERIMENTAL DETAILS

The types of SiC studied include double-sided polished (0001) hexagonal polytypes and 1 µm thick 3C epitaxial layers grown on a (100) Si substrate. All were oxidized in steam at $1150°$ C, and the oxide thickness was measured by profilometry. Dry heat treatment was performed by placing the samples in a double walled quartz tube at $900°$ C with flowing 0.99998 N_2 or 0.99996 O_2 for 200 min. The moisture reading at the output of the furnace tube used for the dry heat

treatment was always less than 1 ppm. After heat treatment, the oxides were etched from the substrate at room temperature using a 9:1 $H_2O:HF$ (50%) solution. The same substrates were used for the subsequent oxidations and dry heat treatments.

X-band EPR measurements were performed at room temperature after oxidation, after thermal treatment, and after oxide removal. The EPR signal of the defect discussed here is induced by the thermal treatment. A spectrum with the same g-value as the heat treatment induced signal persists in the 3C epilayers after oxidation. To obtain the density of centers which were activated by the thermal treatment, the EPR spectrum after oxidation was subtracted from that measured after the dry treatment step. No signal remained after oxidation of the hexagonal polytypes, so the density of the center was calculated directly from the EPR signal measured after heat treatment. The number of centers was obtained from double integration of the spectrum and comparison to a standard, weak pitch. To minimize error due to the low signal-noise ratio, the signal with the largest intensity was used to obtain the absolute number of defects. The number of centers in the remaining samples was estimated by comparing the EPR signal amplitudes. Areal density was determined by dividing the number of centers by the area of one side of the sample. Error was based on the amplitude of the noise in each spectrum. Because many samples of the 3C polytype could be measured at one time, data obtained from them exhibit a greater signal-noise ratio than data from hexagonal polytpes. Therefore, the 3C data has the smallest error bars.

3. RESULTS

Figure 2 shows the concentration of centers measured after dry heat treatment of samples that were oxidized for varying lengths of time. Each data point was obtained after oxidation for a time shown on the x-axis, t_{ox}, and 200 min thermal treatment. The 4H (▲), 6H (O), and 3C (□) data exhibit similar trends. A peak is observed at 2 hr for the hexagonal polytypes and 4 hr for the 3C epilayers. After 32 hr, when the 3C epilayer should be totally consumed by oxidation, the EPR signal is below the detection limit
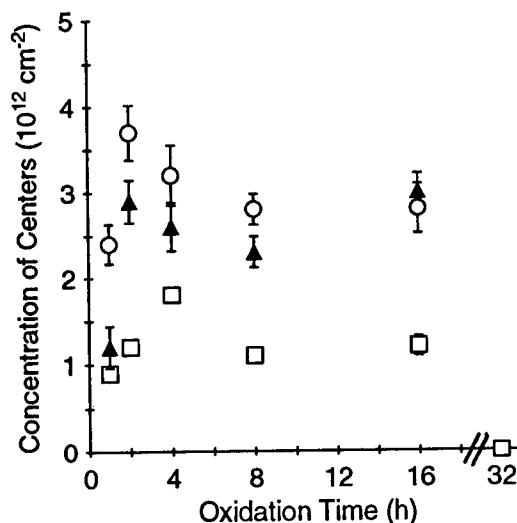
Figure 2. Density of EPR centers induced by dry heat treatment at 900° C in oxidized 3C epitaxial layers (□); oxidized 6H substrates (O); oxidized 4H substrates (▲). Samples were oxidized in steam at 1150° C for the time indicated on the x-axis.

Figure 3. Density of EPR centers measured after layers of SiC were removed by oxidation. Data are the same as those shown in figure 2. Depth is measured below the original, unoxidized surface. For the hexagonal polytypes the depth below the C-face is shown. Depth below the Si-face may be approximated by multiplying by 0.2.

Like Si oxidation, SiC oxidation occurs at the surface of the substrate and consumes the substrate as the SiO_2 grows [7,8]. Thus, oxidation may be thought of as an etching process where the etched material is transformed into SiO_2 rather than dissolved in solution. Since we know that the defect is not located in the SiO_2 layer, the change in the EPR signal intensity with oxidation 'etching' time may be interpreted in terms of a concentration variation with depth into the SiC. The amount of SiC removed during t_{ox} was calculated through the relationship, $d = 0.46 d_{ox}$, where d is the thickness of the SiC consumed by oxidation and d_{ox} is the thickness of the SiO_2 layer grown during t_{ox}. In figure 3, we plot the concentration data of figure 2 vs. the total depth below the original, unoxidized SiC surface. The total depth was obtained by adding the d_{ox} calculated at each t_{ox} to the depths found for all previous times. The 4H (▲) and 6H (O) data shown in figure 3 are plotted vs. distance beneath the C-face surface. Because the Si face and C face of the hexagonal polytypes oxidize at a different rates, depth beneath the surface of the Si face must be obtained by multiplying distance by about 0.2 [7].

4. DISCUSSION

The data of figures 2 and 3 offer two possible ways in which oxidation may contribute to the formation of this center. The dependence on oxidation time shown in figure 2 suggests that oxidation generates the defect by breaking Si-C bonds. On the other hand, the depth profile of figure 3 indicates that oxidation brings a pre-existing defect close to the SiC surface and allows the passivating species to out-diffuse. We discuss each interpretation below.

Oxidation of SiC is thought to create solid SiO_2 and gaseous CO that diffuses from the growing oxide [7]. We speculate that oxidation also creates C-dangling bonds which combine with H atoms during wet oxidation. Some of the CH bonds remain in the substrate as the oxide forms. These yield the defects observed in the EPR spectrum after a dry heat treatment drives off the hydrogen. Other CH bonds may react with H_2O or O_2 to form a gaseous species that leaves the sample, taking with it the potential C dangling

bond. The dependence of the defect density on oxidation time should be determined by the relative reaction rates for CH bond formation and its removal by oxidation. Presumably, CH bond formation dominates initially. As the density of CH in the substrate increases, the defect density measured after dry heat treatment increases. After longer oxidation times, when the CH bonds begin to be removed by reaction with steam or oxygen, the defect density decreases. This is seen in the data at times greater than 2 hr in the hexagonal polytypes and 4 hr in 3C material. Eventually, when equilibrium is reached between CH production and its removal by oxidation, the EPR signal should saturate as seen in the data of figure 2 at time greater than 4 hr. Because the oxidation rate for our cubic and hexagonal samples differs, the transition time between CH formation and removal should not be the same. This is consistent with the differences in the peak times for the hexagonal and cubic polytypes shown in figure 2. The 4H and 6H polytypes have the same oxidation kinetics; consistently, as shown, the transition time from CH formation to CO formation is the same. The different faces of the hexagonal polytypes also exhibit very different oxidation rates. However, because the faces are measured simultaneously it is impossible to separate the contributions of each from the present data.

The depth profile of figure 3 offers a different scenario. Assuming that the EPR center is related to intrinsic defects such as Si vacancies or micropipes, Figure 2 illustrates that the defects exist throughout the 3C epilayer or, in the hexagonal polytypes, to a depth of at least 1 μm below the C-face or 200 nm below the Si face. In this interpretation, we assume that oxidation simply removes material and brings the defect closer to the surface of the SiC. Once the surface is within the diffusion length of the hydrogen-related species that is released, the defect may be revealed through the dry heat treatment. This presents a simple interpretation of the experimental data. However, it is somewhat surprising that materials made by such drastically different means would contain defects that have similar depth profiles. The distribution of defects in the hexagonal and cubic polytypes is the same, and the concentration peak occurs at approximately the same depth for the C-face and 3C samples. In

addition, the defect densities differ by no more than a factor of two. In fact, assuming that the defects arise from both faces of the double-sided polished samples, the hexagonal polytype data should be halved. In this case, the density of defects in the different polytypes is almost identical. Although there does not appear to be any reason why the defects in these materials should be distributed in a like manner, the coincidence is not sufficient basis to eliminate the interpretation of the data in terms of depth dependence of intrinsic defects.

In summary, we observe a defect in several different polytypes of SiC subjected to oxidation and dry heat treatment. A theory in which the defect is intrinsic to SiC cannot be disregarded. However, we believe that the defect is generated by oxidation, and the EPR center is activated by removal of hydrogen from a CH bond.

REFERENCES

1. D.M. Brown, E. Downey, M. Ghezzo, J. Kretchmer, V. Krishnamurthy, W. Hennessy, and G. Michon, Solid-State Electronics 39, 1531 (1996).
2. J. A. Weil, J. R. Bolton, and J.E. Wertz "Electron Paramagnetic Resonance" John Wiley & Sons, Inc, N.Y., 1994.
3. P.J. Macfarlane and M.E. Zvanut, J. Electron. Materials 24, 144 (1999).
4. L. Muehlhoff, M.J. Bozack, W.J. Choyke, and J. T. Yates, Jr., J. Appl. Phys. 60, 2558 (1986).
5. J.M. Powers and G.A. Somorjai, Surface Science 244, 39 (1991).
6. P.J. Macfarlane and M.E. Zvanut, J. Vac. Sci. Tech., accepted for publication.
7. Akira Suzuki, H. Ashida, N. Furui, K. Mameno, and H. Matsunami, Jap. J. Appl. Phys. 21, 579 (1982).
8. C.D. Fung and J.J. Kopanski, Appl. Phys. Lett. 45, 757 (1984).

A simple method for the evaluation of the recombination parameters in SiC MOS structures

T. Rudenko[1], E. Sveinbjörnsson[2] and M. Sadeghi[2]

[1]Institute of Semiconductor Physics NASU, Kyiv, Ukraine
[2]Department of Microelectronics ED, Chalmers University of Technology, Göteborg, Sweden

This paper presents an analysis of non-equilibrium behavior of the room-temperature inversion layer in a SiC MOS structure. Different mechanisms responsible for recombination of the excess inversion carriers are discussed. A simple method for the evaluation of recombination lifetime in the subsurface region of SiC is proposed. The method is based on the room-temperature high-frequency CV-measurements. The limitations for the applicability of the proposed method are outlined.

1. INTRODUCTION

A large attention has been paid in the last few years on the fabrication and investigation of SiC MOS structures due to their potential for high-power and high-temperature devices [1,2]. For design and application of SiC MOS devices, it is important to have a clear understanding of different physical processes in SiC MOS structures. In this work, we discuss non-equilibrium behavior of the room-temperature inversion layer in a SiC MOS structure, which can be created by an external minority carrier generation or injection. Based on the analysis of different recombination current components, we propose a simple method for the extraction of the carrier recombination lifetime in the subsurface region of SiC using room-temperature high-frequency CV-measurements. The limitations for the applicability of the proposed method are analyzed.

2. NON-EQUILIBRIUM ROOM TEMPERATURE INVERSION LAYER

The minority carrier generation rate in SiC at room temperature is known to be extremely low due to the wide band gap of SiC. So, if a SiC MOS capacitor is driven in deep depletion, very long times are required to form an inversion layer and to restore thermal equilibrium. The inversion layer can be formed using an external carrier generation, for example, by illumination. The source of minority carriers in a SiC MOS structure can be also an inversion layer beyond the gate electrode caused by the oxide charge, similar to that in MOS structures on p-type Si with a large positive oxide charge. Once the inversion layer has been formed in a SiC MOS structure, its disappearance by recombination at room temperature is also very slow. The reason is very similar to that in Si MOS structures at low temperatures [3]. It lies in the fact that due to a large band gap of SiC at strong inversion potential the majority carrier concentration at the SiC surface and in the depletion region is very low, and the potential barrier for injection of minority carriers into SiC volume is too high. As a result, a large deviation from the equilibrium state is needed in order for recombination to become noticeable. In other words, high forward biases should be developed across the inversion layer (being in fact gate-induced p-n junction) before a measurable recombination current can flow at room temperature in a SiC MOS structure.

Therefore, if the inversion layer has been formed in a SiC MOS structure, for practical sweep rates, the system is in non-equilibrium condition, when the gate voltage is swept in a direction of decreasing inversion. This results in non-equilibrium C-V-curves in the inversion region. Below it will be shown that at proper conditions the carrier recombination lifetime in the surface region of SiC can be extracted from the non-equilibrium steady-state inversion capacitance measured with the linear return sweep of the gate voltage.

3. EXPERIMENTAL

Measurements presented in this work have been performed on n-type SiC MOS structures obtained by a thermal oxidation of N-doped 6H-SiC wafers in dry O_2 at 1100^0C. The doping of the epitaxial layer was 1.3×10^{16} cm^{-3}. The oxide thickness was 1380Å.

Fig.1 shows the high-frequency C-V-curve measured in a SiC MOS structure at room temperature, when inversion layer is obtained by illumination at the turning point. Physical processes responsible for different regions of the C-V-curve are, in fact, the same as in a Si MOS structure at low temperature [3]. Recombination processes appear on the return sweep. A steady-state non-equilibrium inversion capacitance is observed when the recombination current is $I_{rec} = C_{ox} dV / dt$, with C_{ox} the oxide capacitance and dV / dt the voltage sweep rate.
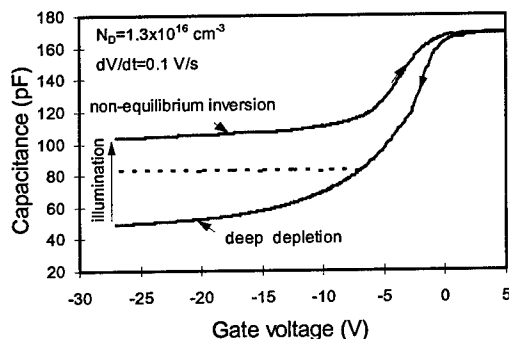
Fig.1. Non-equilibrium HF C-V curve measured on the n-type 6H-SiC MOS capacitor at room temperature ((f=1MHz, A_g=7.85x10^{-3} cm^2). Dashed line indicates an equilibrium inversion capacitance.

4. RECOMBINATION CURRENT COMPONENTS

In general, there are the following components of recombination of the excess inversion carriers in a MOS structure: (1) surface recombination under the gate, (2) recombination in the depletion region under the gate, or volume recombination component, (3) recombination in the quasi-neutral region surrounding the depletion region under the gate, or volume diffusion component, (4) surface recombination in the lateral gate-induced depletion region, (5) recombination at the surface outside of

the gate-induced depletion region, or the surface diffusion component.

Generally, the recombination current follows the law $I_{rec} = I_0 \exp(qV_F / nkT)$, where V_F is the forward bias, n=2 when recombination components are dominant, and n=1 when the diffusion mechanism is prevailing.

Let us consider each of the components. Surface recombination under the gate should be suppressed for strong inversion. The estimation of the second and third components shows that for low current densities (10^{-8}-10^{-10} A/cm^2), which correspond to practical sweep rates, and typical lifetime values in SiC (10^{-7}-10^{-10} s), recombination in the depletion region is expected to be several orders of magnitude greater than the diffusion component.

The contribution from surface recombination in the lateral gate-induced depletion region (the perimeter surface recombination) depends on the surface recombination velocity and area-to-perimeter ratio. In practice, this component usually will be of a minor importance compared to volume component. It can be shown that for the case of surface recombination velocity S=1000 cm/s and recombination lifetime in SiC τ_v=10^{-7} s, in the capacitor with the gate radius of 0.01 cm the perimeter effect should be about 2% of the volume recombination contribution.

Recombination at the surface outside the gate-induced depletion region depends critically on the surface conditions. For the case of accumulation, this component appears to be suppressed, however, for the case of depletion or inversion, it may be significant. Besides, if an equilibrium surface band bending and depletion width outside the electrode are higher than their non-equilibrium values under the gate, the minority carriers will sink out of the inversion layer by drift-diffusion mechanism.

5. A SIMPLE MODEL FOR RECOMBINATION IN THE DEPLETION REGION

The above analysis shows that if the surface diffusion (or drift-diffusion) component out of the gate is avoided (for example, due to accumulation at the surface or by using a guard ring biased in accumulation), recombination in the depletion region under the gate is expected to be prevailing. For this case, the carrier recombination lifetime in the surface

region of SiC can be evaluated from the C-V measurements.

An observation of a constant non-equilibrium inversion capacitance, or constant recombination current, means that a steady-state analysis can be applicable. With the assumption of a single-level bulk trap with a uniform spatial distribution, a steady-state recombination rate can be given by the usual SRH expression [4]:

$$R_V = \frac{pn - n_i^2}{\tau_p(n + n_1) + \tau_n(p + p_1)} , \qquad (1)$$

where symbols have their usual meaning. For a low carrier injection conditions, quasi-Fermi levels for electron and holes are assumed to be constant: $E_{Fn} - E_{Fp} = qV_F$, so that throughout the depletion layer $pn = n_i^2 \exp(qV_F / kT)$. The recombination rate is non-uniform across the depletion layer with a maximum value in the point where $\tau_p n = \tau_n p$. From (1) for $E_t = E_i$, one can obtain:

$$R_{V\max} = \frac{n_i \cdot \exp\left(\dfrac{qV_F}{2kT}\right)}{\tau_v} , \qquad (2)$$

where $\tau_v = 2\sqrt{\tau_n \cdot \tau_p}$ is the carrier recombination lifetime,

In a MOS capacitor, the value of the "forward bias" V_F developed across the inversion layer, is set by the balance between the rate of recombination processes and the charge decrease at the gate. We suggest that V_F can be determined from the difference of an equilibrium strong inversion potential $2\varphi_F$ and non-equilibrium inversion potential φ_n resulting in steady-state balance:

$$|V_F| = 2 \cdot \left(\frac{kT}{q}\right) \ln\left(\frac{N}{n_i}\right) - |\varphi_n| . \qquad (3)$$

where φ_n is calculated from the measured steady-state non-equilibrium inversion capacitance

Fig. 2 shows the potential and recombination rate distributions in the depleted region calculated from the Poisson equation for non-equilibrium

conditions for various gate voltages. It is evident that the potential and recombination rate distribution in the depletion layer remains in fact unchanged for a wide range of V_g. Thin lines in Fig.2 present the case of the threshold conditions, when $|\varphi_s| = 2\varphi_F - V_F$. It can be seen that for the threshold point, the recombination rate distribution has the same shape, as for the high density of inversion carriers. This reflects the fact that in strong inversion, the free carrier concentration distribution in the depletion layer is independent of the gate voltage. Therefore, recombination current will be independent of the gate voltage, as long as strong inversion is maintained.
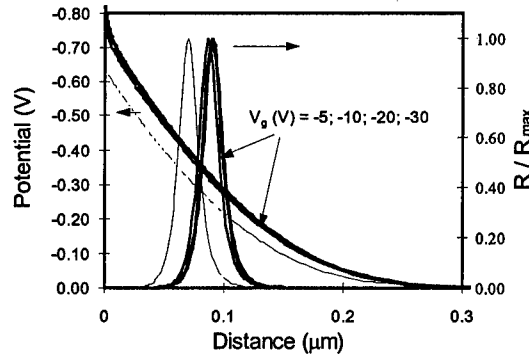
Fig.2. Potential and recombination rate distributions in the depletion layer calculated for various V_g (d_{ox}=140 nm, N_D=1.3x10^{16}cm^{-3}, V_F=2 V, V_{FB}=0).

As can be seen from Fig.2, recombination has to be efficient only in a narrow region of the depletion width where the electron and hole capture rates are comparable. This active width X_{rec} corresponds to the region within the potential variation of $2kT/q$ from the point of the maximum recombination rate. That is, X_{rec} can be determined as:

$$X_{rec} = \frac{2kT}{q}\left(\frac{d\varphi}{dx}\right)^{-1} \approx \frac{2kT\varepsilon_s}{q^2 N_D w} , \qquad (4)$$

where w is the depletion layer width, N_D and ε_s are respectively the doping concentration and dielectric permittivity of SiC. (For the simplicity, the electric field $d\varphi / dx$ in eqn.(4) is approximated by its surface value given by Gauss' law). Therefore, the recombination current per unit area in the depletion region can be expressed as follows:

$$I_R = q \int_0^w R_V(x)dx \approx qR_{V\max}X_{rec} . \qquad (5)$$

The validity of the approximation given by (5) is demonstrated in Table (τ_r=300ns). The difference in values of I_R obtained from the present simple model and numerical integration is about 13%.

Table
Recombination Current in Depletion Region

V_F (V)	Numer. integration I_R (A/cm^2)	I_R=q$R_{V\max}X_{rec}$ (A/cm^2)
1.8	6.5×10^{-10}	5.64×10^{-10}
1.9	4.82×10^{-9}	4.15×10^{-9}
2.0	3.61×10^{-8}	3.12×10^{-8}
2.1	2.75×10^{-7}	2.32×10^{-7}

Therefore, with a reasonable accuracy, the recombination lifetime can be extracted from the experimental data using the following expression:

$$\tau_r = \frac{q \cdot n_i \cdot A_g \cdot X_{rec}}{C_{ox} \cdot \dfrac{dV}{dt}} \cdot e^{\frac{qV_F}{2kT}}, \qquad (6)$$

where A_g is the gate area, V_F and X_{rec} are given respectively by (3) and (4).

6. RESULTS

Fig. 3 presents the $I_R(V_F)$-dependence obtained from the experimental C-V curves measured with various sweep rates ranging from 10^{-2} to 10 V/s on the capacitor with the gate radius of 500 μm.
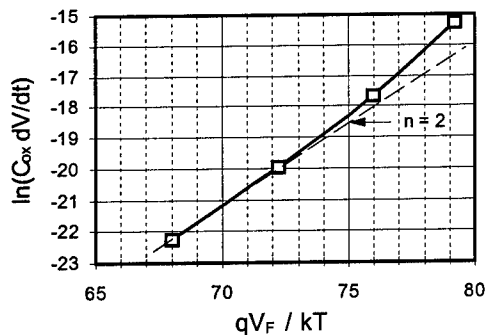
Fig. 3. $I_R(V_F)$-dependence extracted from the experimental C-V curves (A_g=7.85x10^{-3}cm^{-2})

It is seen that in a given device, recombination current can be fairly well fitted by $\exp(qV_F / 2kT)$, which indicates the predominance of the recombination component over the diffusion component. Thus the above simple analysis seems to be applicable. (If n tends to 1, or there is no dependence of the inversion capacitance on the sweep rate, the dominant mechanism is most likely diffusion, or drift-diffusion at the surface out of the gate. For this case, Eqn.(6) is not valid). For the example presented in Fig.3, the extracted lifetime is 290 ns.

7. CONCLUSIONS

If the inversion layer has been formed in a SiC MOS structure by external generation, its disappearance by recombination at room temperature should be very slow due to a wide band gap and low intrinsic concentration in SiC. A large deviation from equilibrium state is needed to ensure a measurable recombination current at room-temperature. This results in highly non-equilibrium C-V curves in inversion region.

At room temperature for practical voltage sweep rates, the most important mechanisms responsible for disappearance of the inversion layer are shown to be recombination in the depletion region under the gate and diffusion, or drift-diffusion at the surface outside the gate. The latter mechanism, which strongly depends on the surface conditions beyond the electrode, seems to be avoided for the case of accumulation, or by using a guard ring biased in accumulation. For this case, the carrier recombination lifetime in the surface region of SiC can be extracted from room-temperature C-V measurements.

REFERENCES

1. J.A. Edmond, H.-S. Kong, A. Suvorov, D. Waltz, C. Carter Jr., Phys. Stat. Sol. (a) 162 (1997) 481.
2. M. Bhatnagar and B.J. Baliga, IEEE Trans. Electr. Dev., 40 (1993) 645.
3. A. Göetzberger, IEEE Trans. Electr. Dev., 14 (1967) 787.
4. W. Shockley and W.T. Read, Jr., Phys. Rev., 87 (1952) 835.

Defect Studies in Epitaxial SiC-6H Layers on Insulator (SiCOI)

E. Hugonnard-Bruyère[b&c&d], J.L.Cantin[a], H.J.von Bardeleben[a], F. Letertre[e], L. DiCioccio[b], T. Ouisse[c]

[a] Groupe de Physique des Solides,Universités Paris 6&7, UMR 75-88 au CNRS, F-75251 Paris
[b] LETI-CEA, Département de Microtechnologies, 17 rue des Martyrs, F-38054 Grenoble
[c] LPCS (UMR CNRS 5531), ENSERG, 23 rue des Martyrs, F-38016 Grenoble
[d] LPM (UMR CNRS 5511), INSA, bât. 502,20 av. A. Einstein, F-69121 Villeurbanne
[e] SOITEC S.A., Parc Technologique des Fontaines, F-38190 Bernin

Nitrogen doped N-type 6H-SiCOI structures have been obtained by the Smart Cut® technique. The process induced defects and their electrical activity have been analysed by electron paramagnetic resonance (EPR) spectroscopy, photoluminescence and Hall measurements. The influence of thermal annealing at temperatures between 800 and 1300°C has been studied. The 1300°C annealing step is required to recover the N-type conductivity, even though compensating defects at concentrations of $1x10^{18}cm^{-3}$ are still present.

1. INTRODUCTION

The Smart Cut® technique [1], allowing the separation of a thin semiconductor layer from the substrate, is based on hydrogen implantation and wafer bonding. A key challenge for the successful preparation of $SiC/SiO_2/Si$ structures is the control of the electrically active defects induced by the implantation and the reduction of their concentration below the initial shallow donor doping level.

The principal information, about the microscopic structure of implantation induced defects has been obtained from positron lifetime studies which have evidenced silicon vacancy and divacancy centres [2]. In a recent DLTS study [3], it has been shown that 2 MeV electron irradiation in N-type SiC layers leads to the formation of the same defects as 300 keV deuterium implantation. These defects have been tentatively attributed to negatively charged carbon vacancies related defects (Ec-0.34eV, Ec-0.41eV), carbon vacancy related (0.51eV) and a divacancy $Vc-V_{Si}$ center (Ec-0.62/0.64eV) respectively. Mono-Vacancy defects have been identified recently by EPR in electron irradiated SiC-6H bulk samples: the negatively charged Vsi defect has been identified in N-type SiC (electron spin S=3/2, $g_{isotropic}$ =2.0030) [4] and the positively charged carbon vacancy Vc has been observed in p-type B doped bulk samples (electron spin S=1/2, triclinic symmetry, g_1=1.9962, g_2=2.0019, g_3=2.0015)[5]. Analogous EPR studies of the vacancy related defects have not yet been reported for ion implanted SiC. In one previous EPR study of 300 keV Ge implanted bulk samples [6], only carbon dangling bond centres ($g_{isotropic}$=2.0028) have been observed at a high concentration ($10^{19}cm^{-3}$).

We report here the results of an EPR study on the nitrogen dopant and the process induced point defects in transferred N-type SiC-6H layers and correlate them with DLTS, Hall and photoluminescence measurements.

2. EXPERIMENTAL DETAILS

N-type SiC-6H epitaxial layers were grown on bulk SiC-6H substrates and transferred by the Smart Cut® technique on SiO_2/Si. The N-type doping concentration was varied between 10^{16} and $6x10^{18}cm^{-3}$. The detailed process conditions are given in ref [1,7] and will only be resumed here: the layers were implanted at room temperature with 120 keV hydrogen atoms at a dose of $7x10^{16}cm^{-2}$, bonded on SiO_2/Si and split from the SiC substrate by an anneal at 800°C. For these conditions the

thickness of the transferred layer is 5800Å. The SiCOI structures were annealed up to 1300°C. The EPR measurements were performed with an X-band spectrometer at room temperature and in the 4-77K range. Absolute spin concentrations were determined with a calibrated Ruby standard sample. Hall effect measurements were carried out in the 40 K- 420 K range. The Van Der Pauw test patterns were realised with a MESA structure. The photoluminescence spectra were measured at 4K under 354nm excitation.

3. EXPERIMENTAL RESULTS AND DISCUSSION

After the transfer, all SiC layers are found to be strongly compensated and highly resistive at 300K. Even after 1300°C annealing, the resistivity values are still in the range of $10^6\Omega.cm$ for a sample with an initial doping level of $1x10^{17}cm^{-3}$ (0.2Ω.cm). For lower annealing temperatures, ohmic contacts cannot be realised, preventing electrical characterisation. The EPR technique allows a quantitative assessment of both the compensating (paramagnetic) defects and the nitrogen donors. The room temperature EPR spectrum of the as transferred layers is characterised by a nearly isotropic line with a g-factor of 2.0027±0.0001 and a linewidth of 2G (fig.1). The volume averaged spin concentration is $\approx 1x10^{18}cm^{-3}$. The g-factor and linewidth of the EPR spectrum are very close to those of the dominant paramagnetic centres in a-C, a-C:H [8] and in a-SiC [9]. These defects are generally attributed to carbon dangling bond centres and we will use the same model here. Whereas these defects are not expected to exist in crystalline material, implantation can generate such defects in the ion track regions and/or in the stopping range. Higher resolution measurements at low temperature show however the EPR spectrum to contain various slightly anisotropic, non resolved centres, none of which corresponds to the negatively charged silicon vacancy defect.

The EPR observation of the neutral N donor spectrum requires low temperature (T<80K) to avoid thermal ionisation. In the temperature range (4-77K), irrespective of the N doping level, no EPR spectrum of the neutral N° donor on the hexagonal or quasi-cubic sites is

observed. From the electrical measurements it is clear that this nonobservation is due to electrical compensation by acceptor traps with a global concentration above $6x10^{18}cm^{-3}$; a passivation of the nitrogen donors by hydrogen can be excluded for the 1300°C annealed layers [7].
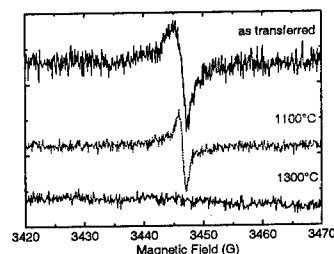
Figure 1. Room temperature EPR spectra of the as-transferred and annealed layers

To reduce the defect concentration, thermal annealing at 1100°C, 1200°C and 1300°C were performed. At each temperature, an additional fraction of the g=2.0027 paramagnetic centres, observable at room temperature, is eliminated. After the 1300°C anneal their spectrum is no longer observable at room temperature (fig. 1).

The annealing introduces new paramagnetic defects: after the 1100°C annealing step, an electron spin S=1 centre is observed at 4K under in situ optical excitation. The EPR spectrum is characterised by an emission high field line typical for optically populated S=1 centres. The EPR spectrum shows axial symmetry around the [0001] axis; its spin Hamiltonian parameters are $g_{//}=2.0020$, $g_{\perp}=2.0052$, $|D|=187x10^{-4}cm^{-1}$. Various spin S=1 centres characterised by D values between $551x10^{-4}cm^{-1}$ and $9x10^{-4}cm^{-1}$ have been reported in electron and neutron irradiated and high temperature annealed SiC-6H [5,10,11], but the spectrum observed here, has not been reported before. Different models for the spin 1 centres can be considered. Whereas, the neutral Si vacancy defect, in an optically excited triplet state, could be evoked as a possible defect, the variety of spin 1 defects observed in SiC-6H is clearly in favour of an associated defect model. From the studies on thermally induced defects [10], these centres have been attributed to divacancy centres with different vacancy separations. Within this model, assuming a simple dipole approximation to explain the D

values, the spin 1 centre observed in the 1100°C annealed layers corresponds to an inter-centre separation of 5.2Å. The second nearest neighbour distances along the c-axis for two carbon (silicon) vacancies depend on the site symmetry: they are 5.0 and 10.0 Å for the two quasi cubic sites and 7.5 Å for the hexagonal site. The defect should thus be assigned to a nearest neighbour carbon-carbon (or silicon-silicon) vacancy. This defect anneals out at higher temperature and can no longer be observed after the 1300°C annealing stage.
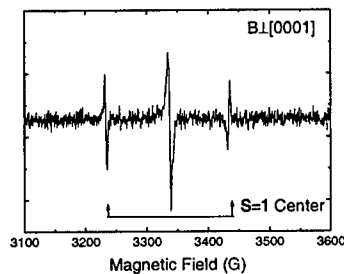
Figure 2. EPR spectrum of the S=1 Centre at 4K

In the 1300°C annealed samples a new single line EPR spectrum is observed (fig.3) under thermal equilibrium conditions at 4K. It is characterised by a spin S=1/2 and axial symmetry along the c-axis with $g_{//}=1.99952$ and $g_{\perp}=1.99905$; it shows an angular dependent linewidth with $B_{//}=0.62G$ and $B_{\perp}=0.78G$. From the g-values and the absence of resolvable superhyperfine interaction the spectrum is attributed to delocalised electrons in Si clusters. Similar centres have been previously observed in thermally annealed silicon SIMOX layers [12]. The EPR spectrum can be bleached by optical excitation, which demonstrates the electrically active character of this centre. The concentration of this defect is estimated to $1 \times 10^{17} cm^{-3}$.
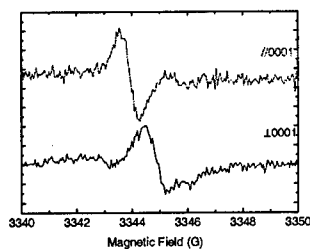
Figure 3. EPR spectra at 4K of the Si cluster related defect for two orientations

Whereas the 1200°C annealing step is still insufficient to reduce the compensating centres such as to allow the observation of the neutral N donor, its EPR signal is observed under thermal equilibrium conditions at 4K in the highly doped 1300°C annealed samples. A minimum N doping of $2 \times 10^{18} cm^{-3}$ is necessary to overcome the residual compensation (fig.4). For the $[N]=3 \times 10^{18} cm^{-3}$ doping level, the EPR spectra of the isolated N donor on the quasi-cubic sites is seen. The N donors on the hexagonal site are still ionised due to the electrical compensation. For higher doping levels of $6 \times 10^{18} cm^{-3}$, the neutral donor from all three sites are observed and the formation of exchange coupled donor centres takes place.
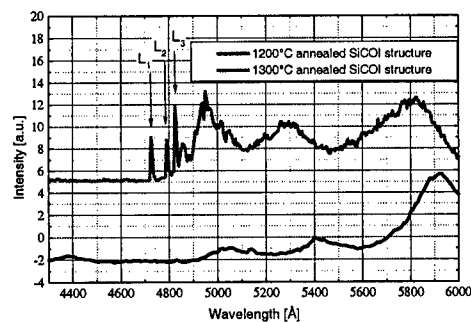
Figure 4. 7K PL spectra of the as-transferred and annealed layers. D_1-centre is characterised by the three lines:L_1, L_2 and L_3.

In the 1300°C annealed layers, the PL spectrum of the D_1 centre is observed irrespective of the initial N doping level [7]. This centre is known to be stable to at least 1700°C [13,14]. Whereas it usually appears in the PL spectrum after a 1000°C sample treatment [13], here, the D_1-centre is not observed in the 1200°C annealed SiCOI structure (fig. 4). It is probable that the high concentration of deep defects prevents its observation. These layers show nevertheless different broader emission bands in the 5200-6000Å range, not assigned to a particular defect.

DLTS measurements on $6 \times 10^{18} cm^{-3}$ doped 1300° annealed SiCOI layers show the presence of the Z1/Z2 centre at Ec-0,62/0,64 eV [7] which might correspond to the D_1-center.

The remaining compensation in a 1300°C annealed sample, with an initial doping level equal

to $3x10^{18}cm^{-3}$ has been evaluated by the EPR and Hall effect techniques.
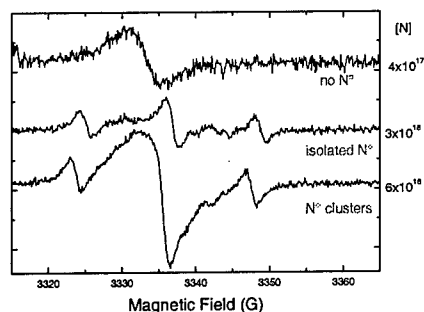
Figure 5. EPR spectra of SiCOI layers annealed at 1300°C for [N] = $4x10^{17}$, $3x10^{18}$, $6x10^{18} cm^{-3}$
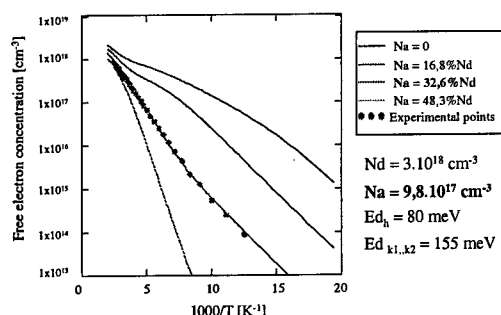
Figure 6. Free electron concentration of SiCOI layer annealed at 1300°C for[N]=$3x10^{18} cm^{-3}$

The analysis of the EPR spectrum – in which only N donors at the quasi-cubic sites are observed (fig.5)-, indicates a remaining compensation equal or higher than the nitrogen concentration on hexagonal sites, i.e. typically 1/3 of the total donor concentration ($\sim 1x10^{18}cm^{-3}$). This is confirmed by Hall measurements between 40K and 420K on differently doped layers. For an initial doping concentration Nd of $3x10^{18}cm^{-3}$ donors the concentration of compensating defects Na corresponds to 32% of Nd and for a sample with Nd = $1.3x10^{18}cm^{-3}$, we deduce a Na value of 92%. Thus the concentration of the compensating centres Na ($1x10^{18}cm^{-3}$) does not depend on the initial doping level of the material. On the other hand, the mobility is not too much affected by the defects. The $3x10^{18}cm^{-3}$ doped sample shows a mobility of $80cm^2.V^{-1}.s^{-1}$.a value close to current mobility

values (around 130 $V^{-1}.s^{-1}$ for this doping range) [15]. Further recent improvements have drastically decreased this compensation below some $1x10^{16} cm^{-3}$ (not shown here).

4. CONCLUSION

As transferred SiCOI layers are electrically compensated and characterised by a high concentration of carbon dangling bond defects. A thermal annealing at 1300°C leads to a strong reduction of these defects and a recovery of the N-type character for layers doped above $2x10^{18}cm^{-3}$. At intermediate annealing temperatures a divacancy defect is observed.

RERERENCES

1. L.DiCioccio, Y.LeTiec, F.Letertre, C.Jaussaud, M.Bruel, Elect.Lett.32,1144(1996)
2. G.Brauer, W.Anwand, P.G.Coleman et al., P.Willutzki, Phys.RevB54,3084(1996)
3. M.O.Aboelfotoh, J.P.Doyle, Phys.Rev.B59, 10823 (1999)
4. J.Schneider and K.Maier, Physica B185,199 (1993)
5. N.T.Son,W.M.Chen,J.L.Lindström et al., Proc.of the 2nd Europ.Conf on SiC and Related Mat. (ECSCRM98), Montpellier (1998), to be published
6. R.C.Barklie, M.Collins B.Holm, Y.Pacaud, W.Skorupa, J.of Elect.Mat.26, 137 (1997)
7. E.Hugonnard-Bruyère et al., Proc. of the 2nd Euorp. Conf on SiC and Related Mat. (ECSCRM98), Montpellier (1998), to be published
8. L.Watanabe, S.Hasegawa, Y.Kurata, Jap.J.Appl. Phys.21,856 (1982)
9. T.Shimizu, M.Kumeda, Y.Kiriyama, Solid State Commun.37,699(1981)
10. N.M.Pavlov, M.I.Iglitsyn, M.G.Kosaganova, V.N.Solomatin, Sov.Phys.Semicond.9,845 (1975)
11. V.S.Vainer, V.A.Il'in, Sov.Phys.Solid State 23,2126 (1981)
12. A.Stesmans, A.G.Revesz, H.L.Hughes, J.Appl.Phys.69,175(1991)
13. Lyle Patrick and W. J. Choyke, Physical Review B, Vol. 5, No 8, 3253-3258, 1972
14. Ch. Haberstroh,, Appl. Phys. 76 (1), 509-513, 1994
15. H. Matsunami, T. Kimoto, Materials Science and Engineering, R20, 125-166, 1997

ALTERNATIVE DIELECTRICS

Characteristics of tantalum pentoxide dielectric films deposited on silicon by excimer-lamp assisted photo-induced CVD using an injection liquid source.

M.B. Mooney[a], P.K. Hurley[a], B.J. O'Sullivan[a], J.T. Beechinor[a], J.-Y. Zhang[b], I.W. Boyd[b], P.V. Kelly[a], J.-P. Sénateur[c], G.M. Crean[a], C. Jimenez[c] and M. Paillous[c]

[a]National Microelectronics Research Centre, Lee Maltings, Prospect Row, Cork, Ireland;

[b]Department of Electronic & Electrical Engineering, University College London, Torrington Place, London WC1E 7JE, United Kingdom;

[c]Laboratoire des Matériaux et du Génie Physique,
Institut National Polytechnique de Grenoble, St. Martin d'Heres, Grenoble, France.

We report the physical and electrical characteristics of the first tantalum pentoxide dielectric films as deposited by the new technique of low pressure ultraviolet-assisted injection liquid source (UVILS) chemical vapour deposition (UVILS-CVD) using the precursor tantalum tetraethoxy dimethylaminoethoxide ($Ta(OEt)_4(dmae)$). The films as deposited exhibit high leakage currents due to carbon impurities. Significant porosity is found at deposition temperatures below $350^\circ C$. Conventional C-V characteristics are exhibited by thick (200 Å to 1000 Å) as-deposited films, with dielectric constants of 17.4 to 24.

1. FORMAT

Tantalum pentoxide (Ta_2O_5) is of interest for optical waveguiding, sensors, on-chip capacitors, as capacitor insulators in high density dynamic random access memories (DRAMs) and in ultra-large-scale-integrated gate dielectrics due to the high dielectric constant (about 25). Ta_2O_5 films have been previously deposited by methods including thermal chemical vapour deposition (CVD) [1], low pressure CVD [2-5], plasma-enhanced CVD [6], photo-induced CVD (photo-CVD) [7, 8] using lasers and a carbon-free thermal CVD process [5]. Recently excimer lamp assisted sol-gel [9], and thermal injection liquid source CVD deposited tantalum pentoxide were reported [10].

In this paper we report the physical and electrical characteristics of the first tantalum pentoxide dielectric films deposited by the new technique of low pressure (2 to 15 mbar) ultraviolet-assisted injection liquid source (UVILS) chemical vapour deposition (UVILS-CVD). The physical properties were studied using phase modulated spectroscopic ellipsometry (SE) , Fourier transform infrared spectroscopy (FTIR) and dynamic secondary ion mass spectroscopy depth profiling (SIMS). The electrical properties of the as-deposited films were determined by capacitance-voltage (C-V), and current-voltage (I-V) measurements, on Al/ Ta_2O_5/Si capacitor structures.

2. EXPERIMENTAL

Si(100) wafers (Wacker Siltronic) were RCA cleaned [11] in an FSI spray post cleaner. UVILS-CVD combines the intense radiation of a KrCl* excimer lamp source (wavelength 222 ± 6 nm) with a novel injection liquid source capable of delivering precise controllable quantities of a liquid metalorganic precursor into the cold-wall CVD chamber, through a showerhead maintained at temperatures between $80^\circ C$ and $110^\circ C$. The films were deposited on p-type Si(100) wafers using tantalum tetraethoxy dimethylaminoethoxide (Ta-T-DMAE), a novel high-volatility tantalum precursor, dissolved in cyclohexane.

The injection liquid source [10] with an excimer lamp photo-CVD reactor [9], have both been described elsewhere. An optical power of approximately 10 $mW \cdot cm^{-2}$ at 6 cm from the lamp outer surface was employed. The injection liquid source was filled with a 15% cyclohexane solution of tantalum tetraethoxy dimethylaminoethoxide ($Ta(OEt)_4(dmae)$), a high volatility liquid tantalum precursor intended for low temperature deposition processes recently used in a thermal MOCVD process for tantalum pentoxide [12]. This solution was injected pulsewise into the reactor under conditions of constant drop mass and pulse rate (1 Hz). An argon gas flow carried the precursor through the showerhead at $110^{\circ}C$ to the wafer. Nitrous oxide (N_2O) was used as oxidising agent.

Infrared absorbance spectra were recorded in transmission mode using a Bio-Rad FTS-40APC FTIR Spectrometer in the spectral range 400 - 4000 cm^{-1}, with the spectrum of the as-cleaned chemically oxidised control wafer subtracted as background. Spectroscopic ellipsometry was performed to measure the film thickness and refractive index, using a Jobin Yvon UVISEL™ variable angle spectroscopic ellipsometer in the spectral range 1.5 to 5.0 eV (wavelength 827 nm to 248 nm) at 70 degrees angle of incidence. Secondary Ion Mass Spectrometry (SIMS) depth profiles were performed using a CAMECA IMS 3f Ion Microscope.

3. RESULTS AND DISCUSSION

SIMS depth profiling (Figure 1) shows the presence of significant carbon and silicon within the unannealed as-deposited dielectric layers. The source of the carbon is probably the precursor, but the solvent cannot be ruled out as a source. The presence of silicon is tentatively either to tantalum silicide formation and oxidation during deposition, or to diffusion from the substrate into the slightly porous tantalum pentoxide film.

The porosity is indicated by refractive index measurements (to a precision of ± 0.02 at 633 nm wavelength (1.96 eV photon energy)) of significantly less than the values of 2.2 or greater which are characteristic of densified (annealed) tantalum pentoxide. The refractive index values were derived

by fitting ellipsometric (psi, delta) spectra to an optical model in which a single dielectric layer, represented by a modified expression of the amorphous dispersion relation model of Forouhi and Bloomer [13], is present on a Si(100) substrate represented by the reference data of Jellison [14].
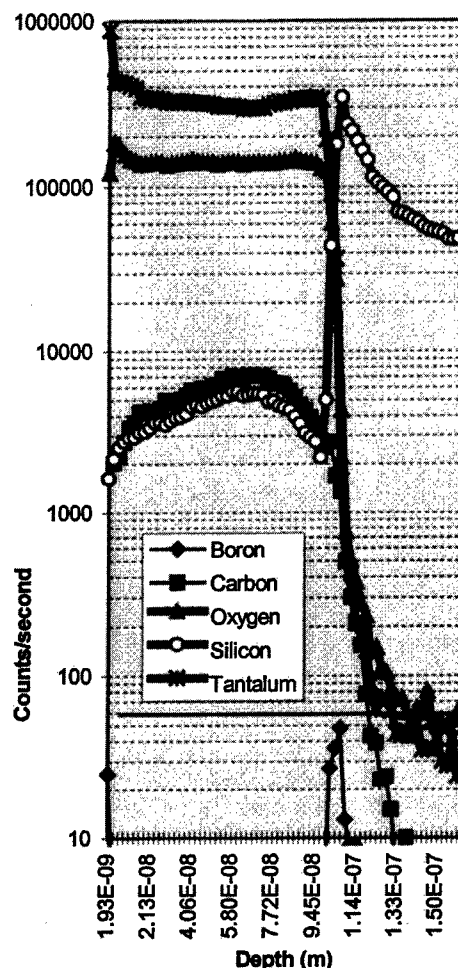
Figure 1: SIMS depth profile through UVILS-CVD tantalum pentoxide film. Significant carbon content is evident in the dielectric layer, but negligible alkali metals are observed.

Figure 2 shows the relationship of refractive index to the substrate temperature during deposition. Lower refractive index is an indication of porosity in the as-deposited material. The dependence of refractive index on other process parameters at $350^{\circ}C$ substrate temperature is much less pronounced, and

not fully understood. The mean refractive index of films deposited at 350°C under a variety of other process conditions was 2.09, with a spread of ± 0.07. The highest refractive index observed in the unannealed films was 2.16.
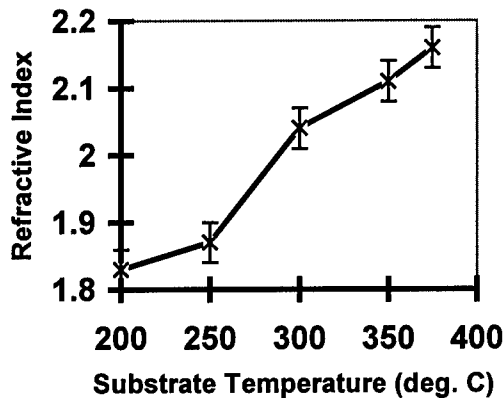
Figure 2 Refractive index of UVILS-CVD tantalum pentoxide films as a function of substrate temperature during deposition.

It is found that the precursor bears sufficient oxygen to deposit a tantalum pentoxide layer in the absence of an oxidising gas flow, although the suboxide content indicated by FTIR spectra decreases on depositing the layer in a nitrous oxide flow (Figure 3).
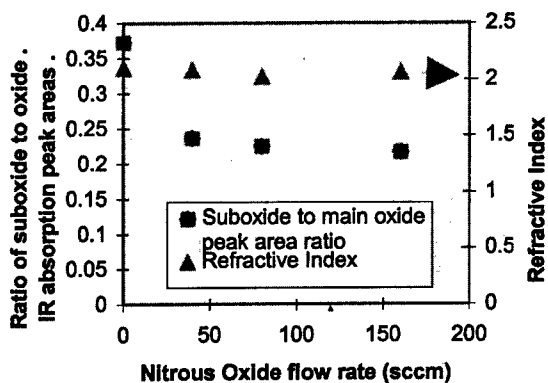
Figure 3 Ratio of the area under the Ta=O peak to that under the Ta-O and Ta-O-Ta peak/shoulder in infrared absorption spectra of tantalum pentoxide layers as-deposited by UVILS-CVD, as a function of the nitrous oxide flow. Refractive index also shown.

Breakdown fields higher than 2 MV/cm, dielectric constants of 17.4 to 24 and fixed oxide charge values of $< 5 \times 10^{10}$ cm^{-2} have been obtained in the as-deposited films. Conventional C-V behaviour was observed (Figure 4) for the as-deposited films.
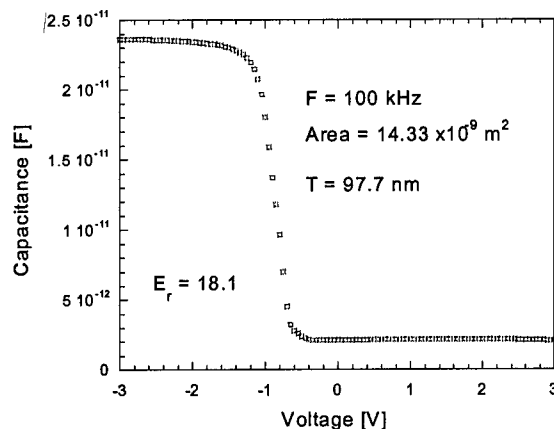
Figure 4 High frequency (100 kHz) CV characteristics of an aluminium - UVILS-CVD tantalum pentoxide - silicon test capacitor. Note the flat band voltage offset of 0.9 eV as a result of the Al gate to p-type Si workfunction difference.

The leakage current densities exhibited by the unannealed Ta$_2$O$_5$ films (Figure 5) are comparable to results presented in other works [2,3] for unannealed films and metal-organic precursors (typically $J = 10^{-3}$ to 10^{-1} A.cm^{-2} at $E = 1$ MV.cm^{-1} where E is the real electric field across the dielectric).

Preliminary work on post deposition annealing (both furnace annealing at 780°C and ultraviolet assisted 400°C with oxidising gas) has reduced the leakage currents by several orders of magnitude, but also reduces the dielectric constant (typically to 10-12). The refractive index of annealed films has been found to exceed 2.2. Nevertheless, effective silicon dioxide thicknesses approaching 60 Å have now been achieved with layers deposited using the UVILS-CVD method. This new technique offers significant potential for low temperature thin dielectric film preparation applied to ultra-large-scale-integrated devices.
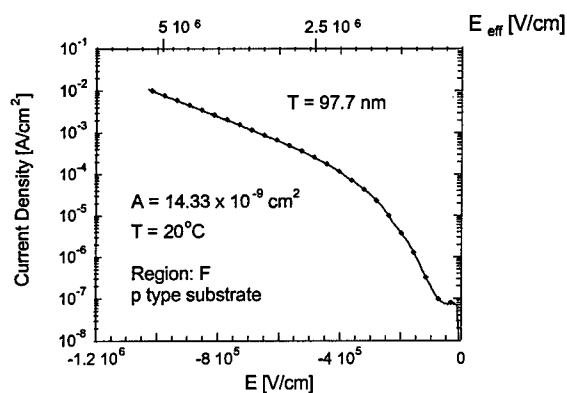
Figure 5 Current density versus electric field. The top y-axis is the effective electric field E $_{eff}$ (i.e., applied voltage divided by the equivalent thickness of SiO$_2$).

4. CONCLUSIONS

Tantalum pentoxide dielectric films have been deposited by the new technique of low pressure ultraviolet-assisted injection liquid source (UVILS) chemical vapour deposition (UVILS-CVD) using the tantalum tetraethoxy dimethylaminoethoxide (Ta(OEt)$_4$(dmae)) precursor. Conventional C-V characteristics are exhibited by thick (200 Å to 1000 Å) as-deposited films, with dielectric constants of 17.4 to 24. The films as deposited exhibit high leakage currents due to carbon impurities, which preliminary annealing studies show can be improved by several orders of magnitude, but at the cost of a lower effective dielectric constant. Significant porosity is found at deposition temperatures below 350°C.

ACKNOWLEDGEMENTS

REFERENCES

1. K.F. Jensen and W. Kern., Ch IV-1 in Thin Film Processes II, (J.L. Vossen and W. Kern eds.) Academic Press, New York (1991) 283
2. S.Kamiyama, P.Lesaicherre, H.Suzuki, A.Sakai, I.Nishiyama and A.Ishitani, J. Electrochem Soc. **140** (1993) 1617
3. S.C.Sun and T.F.Chen, Proc. of IEDM, p.687 (1996)
4. L.K.Han, G.W.Yoon, D.L.Kwong, V.K.Mathews, and P.C.Fazan, IEEE Elec. Device Lett., 15, p. 280 (1994).
5. R.A.B. Devine, C.Chaneliere, J.L.Autran, B.Balland, P.Paillet and J.L.Leray, Proc. of INFOS, p.61 (1997).
6. R.A.B.Devine, L.Vallier, J.L.Autran, P.Paillet and J.L.Leray, Appl. Phys. Lett. **68** (1996) 1775
7. W.H. Lan, S.T. Tu, S.J. Yang and K.F. Huang, Jpn. J. Appl. Phys. **29** (1990) 997.
8. F. Rochet, G. Dufour, H. Roulet, B. Pelloie and M. Froment, Phys. Rev. **B37** (1988) 6468.
9. J.-Y. Zhang, L.-J. Bie and I.W. Boyd, Jpn. J. Appl. Phys. **37** (1998) L27-L29
10. F. Felten, J.P. Senateur, M. Labeau, K. Yu-Zhang and A. Abrutis, Thin Solid Films **296** (1997) 79-81
11. W. Kern, RCA Review (June 1970) 187-206
12. I.C. Kizilyalli, R.Y.S. Huang and P.K. Roy, IEEE Electron Device Letters **19**, 423-5 (1998)
13. A.R. Forouhi and I. Bloomer, Phys. Rev. **B34**, 7018, (1986)
14. G.E. Jellison Jr., Optical Materials **1**, 41-47 (1992)

Structural Properties of Thin Films of High Dielectric Constant Materials on Silicon

H.C. Lu[a], N. Yasuda[a,1], E. Garfunkel[a], T. Gustafsson[a], J.P. Chang[b], R.L. Opila[b], and G. Alers[b]

[a]Departments of Chemistry and Physics, and Laboratory for Surface Modification, Rutgers University, 136 Frelinghuysen Road, Piscataway, NJ 08854-8019, USA[2]

[b]Lucent Technologies-Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974, USA

We have used Medium Energy Ion Scattering (MEIS) and other techniques to investigate the structure and formation mechanisms of ultrathin (less than 10 nm) layers of Ta_2O_5 on Si. We find that a compositionally graded oxide with < 2 nm effective thickness can be formed. The film degenerates at high annealing temperatures both by roughening at the outer surface, and reacting at the interface, but a buffer layer of Si_3N_4 can prevent the latter effect to a certain extent. Introducing a TiN/Ti layer between Ta_2O_5 and Si (which may be desirable for DRAM applications) has an adverse effect on the thermal stability of the Ta_2O_5 overlayer due to migration and subsequent reaction of oxygen with titanium.

1. INTRODUCTION

As the dimensions of microelectronic devices are scaled down, the high electron tunneling rates in ultrathin gate oxides is becoming an increasingly critical problem. The materials properties of alternative higher-K dielectrics are therefore attracting increasing attention. Metal-oxides with high dielectric constants have the potential to extend scaling of transistor gate capacitance beyond that of silicon dioxide. Ta_2O_5, TiO_2, ZrO_2, HfO_2 and Al_2O_3 are some of the materials that are under active consideration in different laboratories, and are now attracting the attention of the device community [1-8]. These materials offer dielectric constants approximately an order of magnitude higher than that of SiO_2.

Another area where high-K materials will become of use is as dielectric films in cell capacitors of dynamic random access memory (DRAM). There, polysilicon has been used as a storage electrode, forming a capacitor with a metal-insulator-silicon (MIS) structure. The use of a metal for the storage electrode has been proposed to eliminate interfacial oxide growth. Structures such as $W/Ta_2O_5/W$ [9, 10], $TiN/Ta_2O_5/W$ [11], and $TiN/Ta_2O_5/Ru$ [12] have been studied, and in some cases have shown beneficial properties.

Problems that may prevent the use of metal oxides on Si are excessive electrical defects, metal silicide formation or the formation of interfacial SiO_2. Indeed, as simple equilibrium thermodynamics dictates [13], the deposition of Ta_2O_5 on silicon results in the formation of SiO_2. An interfacial SiO_2 layer as thin as 2 nm is enough to negate the benefits of the high-K material in gate applications [14].

We have used Medium Energy Ion Scattering (MEIS), Temperature Programmed Desorption (TPD) and other surface and interface sensitive spectroscopies to investigate the structure and formation mechanisms of ultrathin layers of Ta_2O_5 on Si. Below, we summarize some of our recent results on Ta_2O_5 films on Si [15] and provide a first report on our work on such films with TiN/Ti layers; more extensive discussions can be found elsewhere.

2. EXPERIMENTAL

MEIS can be thought of as a high-resolution, low-energy version of conventional Rutherford

---

backscattering (RBS). Instead of operating at MeV energies, MEIS experiments are performed in the 100 keV range [16]. The lower energy allows the use of higher resolving power ion detection equipment, in our case an electrostatic ion energy analyzer. An additional benefit is that the ion energy loss per unit distance traveled (the stopping power) has its maximum in this range.

Depth distributions of the target elements are obtained by performing a simple computer simulation of the backscattered ion energy distributions [17]. In this way, it is possible to measure quantitative depth profiles for different species with a resolution as high as 3 – 5 Å in the near surface region. (Due to the statistical nature of the ion-solid interaction ("ion straggling"), the depth resolution becomes less good for deeper lying layers.) An advantage relative to certain other profiling methods is that MEIS does not rely on back-etching or sputtering of the sample. By using a low ion dose and averaging over several sample spots, one can avoid distortion of the data due to ion beam induced damage. One should keep in mind that the data obtained represent (incoherent) averages over the entire sampled target area, so it is difficult to distinguish between a large scale compositional gradient and an atomic scale variation, such as interface roughness.

Prior to deposition, the silicon surfaces were cleaned to remove any native oxide material. For some samples, Ti and TiN were deposited by sputtering a titanium target at a substrate temperature of 300 °C. The tantalum pentoxide films were deposited by chemical vapor deposition at temperatures less than 400 °C.

3. RESULTS

3.1. Ta₂O₅ on Si(100)

In Fig. 1 we show a typical MEIS spectrum from a ~70 Å film of Ta_2O_5 on Si(100). We observe clearly resolved peaks from Ta, Si and O. After annealing to 800 °C, the amount of interfacial Si increases and the film becomes denser, indicated by a narrower and higher Ta peak. Our TPD data showed H_2O and $C_xH_yO_z$ desorption consistent with the formation of a denser film.

From these data, we extract depth profiles, shown in Fig. 2. The transition between the silicon and the tantalum oxide is not abrupt and that tantalum extends ~ 1.5 nm into the interfacial
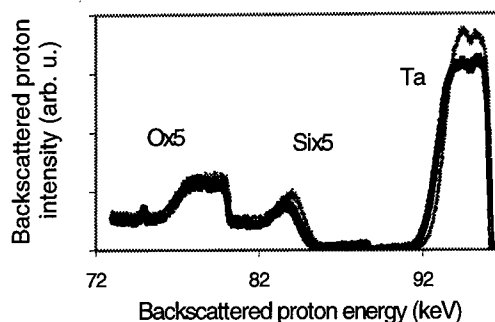
Fig. 1 MEIS spectra of the as-deposited (Thick line) and 800 °C annealed (thin line) Ta_2O_5 film on Si.

region. According to TEM data, film thickness non-uniformity is too small to account for this at our growth temperatures. Our results, further supported by other data, imply that the majority of the interfacial region is not pure SiO_x but a Si-Ta-O mixture. The dielectric constant is significantly higher than that of pure SiO_2. We note in passing also, as have others earlier [18], that up to a monolayer of Si is observed at the outer surface of the as deposited Ta_2O_5 film.
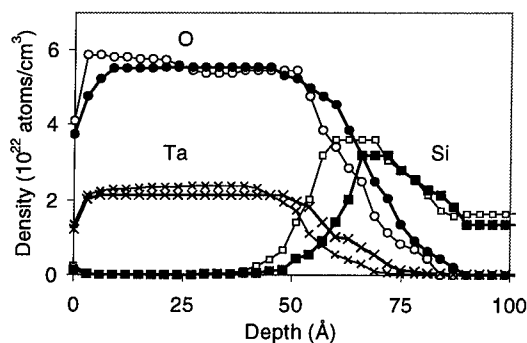
Fig. 2 Depth profiles (as-deposited: thick lines) of Ta, Si, and O from the data in Fig. 1.

Annealing changes the near-interfacial composition substantially. When post anneal temperatures are kept low, stable composite oxide structures (with physical thickness greater than 7 nm) can be obtained that demonstrate good electrical properties and an effective SiO_2 thickness of less than 2 nm. The capacitance is much lower after a rapid thermal anneal to 800 °C due to the growth of the interfacial region. MEIS depth profiles (Fig. 2) show that the interfacial SiO_x region becomes thicker, while the width of the tantalum distribution

actually narrows for a sample annealed in UHV at the same temperature. Annealing to 550 °C is enough to grow the interfacial region, an unexpectedly low temperature. The Ta_2O_5 film is reduced by reacting with the interfacial Si above 800 °C. Above 900 °C, there is substantial oxygen loss (Ta_2O_5 reduction), which probably results from SiO desorption, since some silicon always exist on the surface of the sample (but not in the middle of the Ta_2O_5 film). This indicates that silicon diffusion in Ta_2O_5 can occur even below 900 °C. (SiO desorption is known to decompose the SiO_2 films below this temperature.) The concentrations of Ta, Si and O in the interface region show a non-monotonic dependence on temperature, demonstrating that several different reactions occur. At 930 °C annealing temperature, 2/3 of the oxygen has left the film. There is substantial Si diffusion into the Ta_2O_5 region, and Ta diffusion into the Si. The film eventually is destroyed at high temperatures, and fairly complete intermixing occurs. Significant roughening is also observed using TEM and AFM.
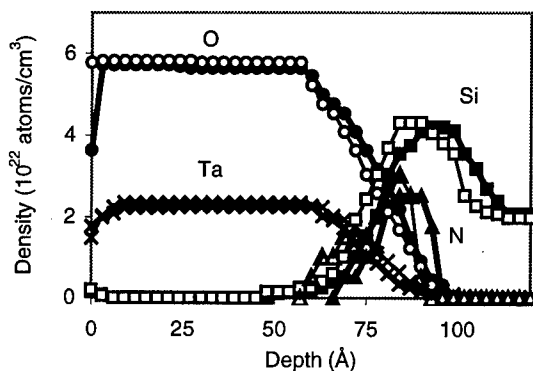
Fig. 3 Depth profiles before and after annealing a $Ta_2O_5/Si_3N_4/Si$ structure.

With a ~ 2.0 nm Si_3N_4 buffer layer, the interfacial region does not grow (no increase in the interfacial Si) when annealed to 800 °C (Fig. 3). Above 900 °C, reduction of the Ta_2O_5 film occurs and TaO_2 can be formed after annealing to 970 °C. While some SiO desorption may occur, a substantial cause of this depletion is oxygen diffusion to the silicon substrate, forming silicon oxide (depth profiles not shown). However, the breakdown of the film is significantly less drastic, and there is, at a given temperature, less oxygen loss.

3.2. Memory cell structures
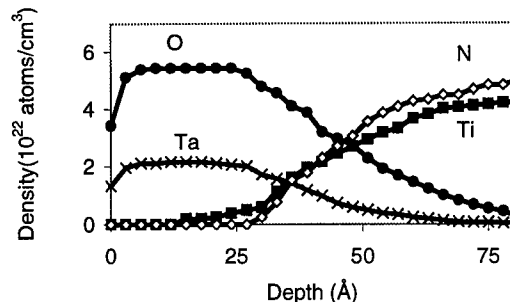
Fig. 4 Depth profiles of the $Ta_2O_5/TiN/Ti/Si$ film.

We have also studied the thermal stability of $Ta_2O_5/TiN/Ti/Si$-type structures, a candidate class of systems for use in DRAMs. The purpose of the Ti layer is to serve as an ohmic contact, while the TiN acts as a buffer layer, as it is thermodynamically stable in contact with Ta_2O_5 [19]. Figure 4 shows MEIS depth profiles for such a sample. The interfacial region is considerably thicker than for the Ta_2O_5/Si case, partly due to a rougher starting TiN surface. Most of the oxygen, however, disappears from the Ta_2O_5 layer after vacuum annealing at 500 °C with the decrease in the O/Ta ratio beginning already at 450 °C. TDS results show no evidence for release of oxygen outside the sample in the form of O_2 gas. For comparison, we have also studied a $Ta_2O_5/TiN/SiO_2/Si$ sample, and found that oxygen does not disappear even at 800 °C. These results suggest that disappearance of oxygen from the Ta_2O_5 layer occurs only in the presence of the Ti layer. TEM and XPS sputter depth profiling experiments show directly an interfacial layer containing oxygen between the TiN and Ti layers [20]. All of these results then support the idea of oxygen migration through the TiN layer. The driving force for the oxygen migration is clearly the thermodynamics between Ta_2O_5 and Ti. A postulated coupled oxidation / reduction reaction between Ta_2O_5 and Ti of the type

$$Ta_2O_5 + 5/x\ Ti \rightarrow 2\ Ta + 5/x\ TiO_x$$
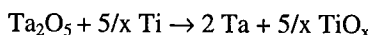
is strongly exothermic. There is thus a strong thermodynamic driving force to reduce Ta_2O_5. What may be surprising is that the reduction actually occurs through a thick TiN layer (100 nm) and at such a low temperature (450 °C). One can speculate about the reasons for the ease of oxygen diffusion through the TiN layer. It is known that TiN films

form columnar grain structures [21]. Enhanced oxygen diffusion may occur through the grain boundaries. Additionally, the TiN films in the present study are non-stoichiometric with excess nitrogen. The low density in this layer [22] may also enhance the oxygen diffusion. It is also possible that the hydrogen originating from the Ti layer may help the diffusion of oxygen through the TiN layer. This argument is supported by the fact that Ta_2O_5 reduction ceases after H_2 desorbs.

We also note that the N/Ti ratio in the MEIS depth profile near the outer surface of the TiN layer increases further after annealing for both of these samples. To better understand this behavior, we have performed TPD experiments on samples of the type Ta_2O_5/TiN/Si and TiN/Si, which are identical except for the Ta_2O_5 overlayer. Two peaks of N_2 desorption (around 500 °C and above 800 °C) are observed for the latter sample, while only one large N_2 desorption peak above 900 °C is found for the former. We infer that molecular nitrogen is generated from the bulk of the TiN layer and is confined near the Ta_2O_5/TiN interface at a moderate temperature (below 900 °C) when a Ta_2O_5 overlayer is present.

4. CONCLUSIONS

We find that, when Ta_2O_5 is deposited on Si, a compositionally graded oxide with beneficial electrical properties can be formed. This film breaks up at high annealing temperatures due to interdiffusion and recrystallization induced roughening. Introducing a buffer layer of Si_3N_4 can to a certain extent prevent the interdiffusion. The reduction of Ta_2O_5 during annealing in Ti containing structures is correlated with migration of oxygen from Ta_2O_5 through TiN and reaction with Ti at the TiN/Ti interface. MEIS and TDS results also indicate a pileup of molecular nitrogen near the Ta_2O_5/TiN interface for as-deposited TiN films with excess nitrogen. All of the results are consistent with known thermodynamics.

REFERENCES

1. A. Chatterjee, *et al.*, in: *IEDM Tech. Dig.*, (1998) 777.
2. J.-L. Autran, R. Devine, C. Chaneliere, and B. Balland, IEEE Electron Device Lett. 18 (1997) 447.
3. D. Park, Y. King, Q. Lu, T.-J. King, C. Hu, A. Kalnitsky, S.-P. Tay, and C.-C. Cheng, IEEE Electron Device Lett. 19 (1998) 441.
4. H. F. Luan, B. Z. Wu, L. G. Kang, B. Y. Kim, R. Vrtis, D. Roberts, and D. L. Kwong, in: *IEDM Tech. Dig.*, (1998) 609.
5. B. He, T. Ma, S. A. Campbell, and W. L. Gladfelter, in: *IEDM Tech. Dig.*, (1998) 1038.
6. S.A. Campbell, D.C. Gilmer, X.C. Wang, M.T. Hsieh, H.S. Kim, W. Gladfelter, and J. Yan, IEEE Trans. Electron Devices 44 (1997) 104.
7. G. D. Wilk and R. M. Wallace, Appl. Phys. Lett. 74 (1999) 2854.
8. L. Manchanda, *et al.*, in: *IEDM Tech. Dig.*, (1998) 605.
9. Y. Nishioka, H. Shinriki, and K. Mukai, J. Appl. Phys. 61 (1987) 2335.
10. H. Shinriki, T. Kisu, S. Kimura, Y. Nishioka, Y. Kawamoto, and K. Mukai, IEEE Trans. Electron Devices 37 (1990) 1939.
11. S. Kamiyama, H. Suzuki, H. Watanabe, A. Sakai, M. Oshida, T. Tatsumi, T. Tanigawa, N. Kasai, and A. Ishitani, in: *IEDM Tech. Dig.*, (1993) 49.
12. T. Aoyama, S. Yamazaki, and K. Imai, J. Electrochem. Soc. 145 (1998) 2961.
13. R. Beyers, J. Appl. Phys. 56 (1984) 147.
14. Y. Nishioka, S. Kimura, H. Shinriki, and K. Mukai, J. Electrochem. Soc. 134 (1987) 410.
15. G. B. Alers, D. J. Werder, Y. Chabal, H. C. Lu, E. P. Gusev, E. Garfunkel, T. Gustafsson, and R. S. Urdahl, Appl. Phys. Lett. 73 (1998) 1517.
16. J.F. van der Veen, Surf. Sci. Rept. 5 (1985) 199.
17. E.P. Gusev, H.C. Lu, T. Gustafsson, and E. Garfunkel, Phys. Rev. B 52 (1995) 1759.
18. S. Tanimoto, Y. Shichi, K. Kuroiwa, and Y. Tarui, in: *Ext. Abs. 1993 Int. Conf. Solid State Devices & Materials*, (1993. Makuhari, Chiba, Japan) 859.
19. H.-J. Lee, R. Sinclair, M.-B. Lee, and H.-D. Lee, J. Appl. Phys. 83 (1998) 139.
20. J. P. Chang, R. L. Opila, G. B. Alers, M. L. Steigerwald, H.C. Lu, E. Garfunkel, and T. Gustafsson, Solid State Technol. 42 (1999) 43.
21. S. R. Ryu, *et al.*, Appl. Phys. Lett. 62 (1993) 579.
22. J.-E. Sundgren, Thin Solid Films 128 (1985) 21.

The Effects of Chemical Bonding and Band Offset Constraints at Si-Dielectric Interfaces on the Integration of Alternative High-K Dielectrics into Aggressively-Scaled CMOS Si Devices

Gerald Lucovsky* and James C. Phillips**
*Departments of Physics, Materials Science and Engineering, and Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27695-8202
**Lucent Bell Laboratories, Murray Hill, NJ 07974, USA

This paper identifies three aspects of the chemical bonding at Si-dielectric interfaces that play crucial roles in the implementation of alternative gate dielectrics for advanced CMOS Si devices: i) the character of the interface bonds, either isovalent with bond and nuclear charge balanced as in Si-SiO$_2$, or heterovalent, with an inherent *mismatch* between bond and nuclear charge, ii) physical bonding constraints related to the average number of bonds/atom, N_{av}, and iii) reduced conduction band offset energies that are result because of increased ionic bonding and d-state derived conduction bands in transition metal oxides.

1. INTRODUCTION

As in-plane device dimensions are scaled to <100 nm, there must also be decreases in the oxide-equivalent thickness of dielectrics, t_{ox-eq}, to <2 nm. Direct tunneling current increases exponentially with decreasing thickness thereby setting a limitation for using SiO$_2$ as a gate dielectric. A practical limiting thickness for SiO$_2$ is t_{ox-eq} ~1.6 to 1.7 nm, where the direct tunneling current density reaches about 1 A-cm^{-2} at a bias of 1 V. A proposed solution is to increase physical thickness to reduce tunneling by using alternative insulators such as transition metal oxides with dielectric constants higher than SiO$_2$ to reduce t_{ox-eq}. This raises three interface chemical bonding issues: i) the balance between electron and nuclear charge [1], ii) physical bonding constraints related average interfacial coordination, N_{av} [2], and iii) reduced band offsets arising from d-state conduction bands [3].

2. CHEMICAL BONDING CONSTRAINTS

2.1 Charge balance in interface bonding

The bonding between Si *surface atoms* and O-atoms in SiO$_2$ at Si-SiO$_2$ interfaces is isovalent as defined in Ref. 1. Interface bonds are partially-ionic (~50%), two-electron pair bonds in which one of the electrons comes from interfacial Si-atoms, and the other from O-atoms of the SiO$_2$ network. Electrons in these bonds are exactly balanced by positive charge on the respective Si- and O-nuclei. Similar considerations apply to bonding of interfacial N-atoms. However, because the transition metal (T-M) oxides are significantly more ionic than SiO$_2$, the bonding between Si and these proposed alternative dielectrics will not involve partially covalent two electron pair bonds.

Moreover, the bonding coordination in non-crystalline T-M oxides such as Ta$_2$O$_5$ is not well characterized; Raman measurements indicate that Ta atoms are six-fold coordinated so that Ta-O bonds have considerable ionic character [4]. This suggests that interface bonding may require formation of Si-Ta bonds to balance nuclear and bond charge. These aspects of interface bonding are not easily resolved due to interfacial Si-SiO$_2$ bonding and Ta$_2$O$_5$-SiO$_2$ alloying [5,6], which occur during thermal CVD deposition, or during the post deposition oxidizing anneals needed to fully oxidize deposited dielectrics.

However, the interfacial bonding at interfaces between Si and group IV(Ti, Zr, and Hf) T-M oxide-SiO$_2$ alloy or silicate compounds can be by isovalent two-electron pair bonds for low alloy atom concentrations of Ti, and for Zr(Hf)O$_2$-SiO$_2$ alloys with Zr(Hf) compositions up to the respective compound silicate compositions. These are regimes in which interface bonding can be through Si-O-Si bonding groups. The limitation on the Ti-content in the TiO$_2$-SiO$_2$ alloy system is associated with bonding interactions between neighbor Ti-O bonds that promote formation of a second phase, e.g., c-TiO$_2$, in which the coordination of Ti is increased to six and the O coordination to three.

The interfacial bonding between Si and Al$_2$O$_3$ can be based two electron pair bonds, but this interface requires additional charged bonding arrangements to balance the more ionic bonding of Al$_2$O$_3$ [7] The atomic structure of non-crystalline Al$_2$O$_3$ dielectrics has two charged bonding components in a 3:1 ratio: *three* tetrahedral

$[AlO_{4/2}]^-$ groups that form a covalently-bonded network for *each* Al^{+3} ion that occupies an approximately octahedral network modifier site within that network. As in the group IV T-M oxides, interfacial Si-SiO$_2$ bonding, and/or interfacial mixed oxide Al$_2$O$_3$-SiO$_2$ bonding are also likely to occur in thermal CVD deposition, or during post deposition anneals. Bonding at internal Al$_2$O$_3$-SiO$_2$ interfaces or within Al$_2$O$_3$-SiO$_2$ alloys will be isovalent, with neutral $[SiO_{4/2}]^0$ and charged $[AlO_{4/2}]^-$ groups forming interface bonds analogous to $[SiO_{4/2}]^0$–$[AlO_{4/2}]^-$ bonding in zeolites, with charge compensation at interfaces and in alloys by the interstitial Al^{+3} ions.

Based on charge balance alone, interfacial Si-SiO$_2$ or nitrided Si-SiO$_2$ may be be required for use with i) transition metal oxides such as Ta$_2$O$_5$, iii) Al$_2$O$_3$ and additionally iii) other group III oxides such as Y$_2$O$_3$ and/or Y$_2$O$_3$-SiO$_2$ alloys.

2.2 Physical bonding constraints

Constraint theory is based on the idea that the bonding forces in a covalently bonded network can be arranged in a hierarchy from strong to weak [8-12]. The constraining effects of these forces are a linear function of the average coordination number, N_{av}. The optimal average coordination number, N_{av}^*, that matches constraints to degrees of freedom is 2.4 as in As$_2$S(Se)$_3$, however, for SiO$_2$, $N_{av}^* = 2.67$ is optimal because bending forces at O atoms are too weak to function as significant constraints at growth or annealing temperatures [8-11]. For over-constrained networks such as Si$_3$N$_4$ (N_{av} = 3.43), Si-atom stretching constraints are stronger than bending constraints and strain energy accumulates along the bending constraints. The average bond angle θ_{ij} is distorted from the local value θ_{ij}^* by an amount $\delta\theta \propto \delta N_{av}^* = N_{av} - N_{av}^*$. Since total strain energy is proportional to $(\delta\theta)^2$ [12], defect creation will be proportional to $\{N_{av} - N_{av}^*\}^2$.

A model for the extension of constraint theory to crystalline-Si dielectric-interfaces is illustrated in Fig. 1 and identifies three interfacial contributions to N_{av}: i) the Si substrate atoms, ii) an ultrathin oxide or nitride interfacial layer, and iii) the bulk dielectric [2,11]. Experiments performed on Si-Si$_3$N$_4$, Si-Si$_3$N$_4$-SiO$_2$ and Si-SiO$_2$-Si$_3$N$_4$ stacks have demonstrated that $N_{defect} \sim \{N_{av} - N_{av}^*\}^2$, where N_{av}^* is the average coordination for an ideal Si-SiO$_2$ interface (see Fig. 2) [4], and $N_{av} \sim 3$ separates device quality from defective interfaces. The model *confirms* that Si-SiO$_2$ interfaces are device-quality ($N_{av}\sim2.8$), whereas Si-Si$_3$N$_4$

interfaces are not ($N_{av}\sim3.5$) [13]. The calculations demonstrate that interposition of ultrathin SiO$_2$ layers between Si and Si$_3$N$_4$ results in values of $N_{av} \leq 3$, whereas interposition of ultrathin Si$_3$N$_4$ layers between Si and SiO$_2$ results in $N_{av} > 3$ (see Table I).

Table I includes calculated values of N_{av} for the interfaces discussed above, as well as interfaces with *emerging* high-K candidate materials such as TiO$_2$, Ta$_2$O$_5$ and Al$_2$O$_3$. N_{av} values for SiO$_2$-group IV T-M oxide alloys and compounds are not inherently over constrained; however, Si-SiO$_2$ interfacial regions may be a natural consequence of the deposition and/or processing used for forming these dielectrics.
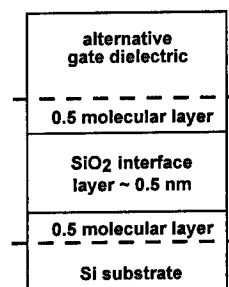
Figure 1. Interface bonding constraint model
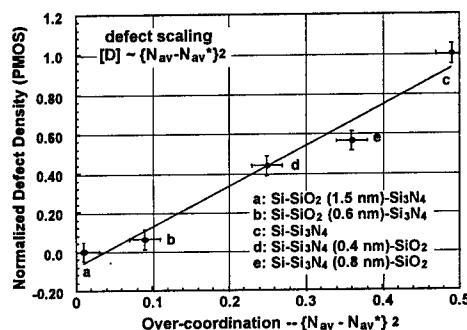
Figure 2. Defect scaling for Si-dielectric interfaces

2.3 Band offset energies

The valence and conduction band offset energies for Si with respect to both SiO$_2$ and Si$_3$N$_4$ are positive and each energy is greater than 2 eV; e.g., the Si-SiO$_2$ conduction band offset is ~3.1 eV and the Si-Si$_3$N$_4$ is ~2.1 eV [14]. The conduction band offset energies between Si and elemental and binary T-M oxides are expected to be less because of two

Table I Interfacial bonding coordination, N_{av}

Interface	N_{av}
Device-quality interfaces	
Si-SiO$_2$	2.8
Si-N-SiO$_2$	2.8
Si-SiO$_2$(0.5 nm)-Si$_3$N$_4$	3.0
Defective interfaces	
Si-Si$_3$N$_4$	3.5
Si-Si$_3$N$_4$(0.4nm)-SiO$_2$	3.3
Alternative-dielectric interfaces	
Si-TiO$_2$(N_{av} = 4)	4.0
Si-Ta$_2$O$_5$ (N_{av} = 3.4)	3.5
Si-Al$_2$O$_3$(N_{av} = 3.6)	3.6

contributing factors: i) increased ionic bonding which moves the valence band to lower energies than in SiO$_2$, and ii) reduced band-gaps due to the lowest conduction band states being derived from unoccupied d-states, rather than s-states of the transition metals [3]. This is shown schematically in Fig. 3. Based on the calculations of Ref. 2, as well as much of the discussion of sections 2.1 and 2.2, it will be necessary to bridge the crystalline Si substrate and T-M oxide dielectrics, with thin (~ 0.5 nm) nitrided SiO$_2$ interface layers (see Fig. 3). Based on comparisons between Si-Al and Si-SiO$_2$-Al interfacial electronic structures, the insertion of SiO$_2$ interface layers should promote an increase in

In addition to the T-M oxides, other studies have addressed non-transition metal oxides. In contrast to interfaces involving the group IV and group V T-M oxides, the band offset energies at Si-Al$_2$O$_3$ interfaces are expected to be comparable to those at Si-SiO$_2$ interfaces because the valence and conduction band states of Al$_2$O$_3$ derived from respectively from O-atom 2p states and Al-atom 3s states. However, the conduction band offset energies at Si-Y$_2$O$_3$ interfaces are expected to be reduced because, as in the group IV T-M oxides, the conduction band states are derived from Y-atom d-states, which are lower than the atomic s-states of the Y-atoms.

Finally, the band offset energies for Si-SiO$_2$-(T-M oxide-SiO$_2$ alloy) heterostructures will depend on alloy composition as well as the bonding coordination of the T-M atoms. To a zeroth order approximation, the bandgaps of the these alloys are expected to scale linearly between the end member values. This is expected because the four-fold coordinated Si atoms sp^3 can mix with the T-M d-states to form the conduction band. This would predict band-gaps of at least 6-7 eV for the compound silicates, ZrSiO$_4$ and HfSiO$_4$, with conduction and valence band offset energies of approximately 2 eV.
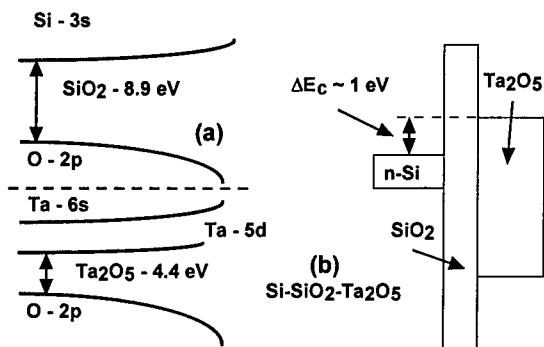
Figure 3. (a) Band gaps in SiO$_2$ and a representative transition metal oxide. (b) Band offset energies in Si-SiO$_2$-Ta$_2$O$_5$ stacked dielectric heterostructures.

conduction band offset energies from those calculated in Ref. 3. Taking this into account the effective offset between Si and Ta$_2$O$_5$, with an SiO$_2$ interface layer is still expected to be ~ 1eV, and in a range where leakage current may still involve Poole-Frenkel transport through Ta$_2$O$_5$ films [15].
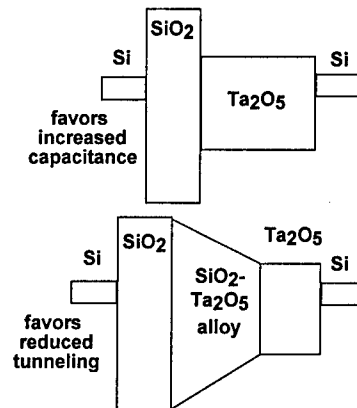
Figure 4. Potential profiles at abrupt and graded interfaces

3. DISCUSSION

Based on considerations of i) electron/nuclear charge balance, ii) physical bonding constraints, and iii) estimates of band offset energies, three different classes of alternative dielectric structures have been identified: (a) group III and V T-M oxides, (b) group IV T-M oxide alloys and silicate compounds, and (c) non-T-M oxides, e.g., Al$_2$O$_3$ [16].

Consider first group V T-M oxides. Figure 4 includes two different types of heterostructures with (a) an abrupt interface between the SiO_2 and Ta_2O_5 layers, and (b) a graded interface between the SiO_2 and Ta_2O_5 layers. The structure in (a) is more effective in reducing $t_{ox,eq}$ and increasing capacitance, whereas the structure in (b) is more effective in reducing i) direct tunneling and ii) any temperature dependent contributions to that leakage current that involve hopping transport (e.g. Poole-Frenkel mechanism). [15]. Comparing experimental results for devices with Ta_2O_5 suggests that the devices of Ref. 5 which have a temperature independent leakage current may have a band diagram corresponding to (b), and those that display a temperature dependent leakage current in Ref. 16 may have a band diagram corresponding to (a). Clearly additional characterizations of these devices are needed to confirm this proposed explanation for the different leakage current results.

Since band offset energies in Al_2O_3, and other non-T-M oxides, are expected to be comparable to those in SiO_2, the band offset energy profiles are expected to similar to those in Si-SiO_2 structures. Neglecting the more ionic bonding, tunneling would be the dominant leakage mechanism; however it more likely that the transport will be modified by the presence of the charged species, particularly the Al^{+3} interstitials which may act as trapping states in a two step tunneling process, or transport states in a hopping conduction, depending on their electronic energies relative to Si [16].

Si-SiO_2-group IV TM oxide alloys are expected to have band structures intermediate between Si-SiO_2 and the abrupt interfaces in Fig. 4(a). As such they represent a compromise between promoting reduced tunneling and increased capacitance. Based on two assumptions which must be born out be experiment, a linear variation of band energies and an electron tunneling mass essentially equal to that of SiO_2, these materials will have effective dielectrics constants of the order of 10-15, and function as gate dielectrics down to $t_{ox,eq} \sim 1$ nm or less with tunneling leakage below 10^{-2} A-cm^{-2}.

Table II Estimated values of $t_{ox,eq}$ for stacked dielectrics: 0.5 nm SiO_2 and 2.0 (1.5) nm high-k

2.0 nm high-k: k	8	10	15	20	25
$t_{ox,eq}$ (nm)	1.45	1.26	1.02	0.88	0.80
1.5 nm high-k: k	8	10	15	20	25
$t_{ox,eq}$ (nm)	1.21	1.07	0.88	0.79	0.73

4. SUMMARY

A consideration of chemical bonding constraints involving the substitution of alternative high-K gate dielectrics for SiO_2 in aggressively scaled CMOS devices identifies the need for interfacial SiO_2 or nitrided SiO_2 interface layers. These clearly impact of the anticipated reductions in $t_{ox,eq}$ limiting reductions to just below 1 nm (see Table I). In addition, inherent differences between the nature of the conduction band states in transition metal oxides and SiO_2 (and Al_2O_3) place additional restrictions on composition profiles that must address increases in capacitance and reductions in leakage current.

REFERENCES

1. W.A. Harrison et al., Phys. Rev. **B 18**, 1402 (1978).
2. G. Lucovsky et al., Appl. Phys. Lett. **74**, 2005 (1999).
3. J. Robertson and C.W. Chen, Appl. Phys. Lett. **74**, 1168 (1999).
4. F.L. Galeener, et al., in *The Physics of MOS Insulators*, ed by G. Lucovsky et al, (Pergamon, NY, 1980), p. 77.
5. H.F. Luan et al., IEDM Tech. Dig. 609 (1998).
6. G.B. Alers et al., Appl. Phys. Lett. **73**, 1517 (1998).
7. G. Lucovsky et al., in *The Structure of Non-Crystalline Materials 1982*, edited by P.H. Gaskell et al.,(Taylor and Francis, London, 1983), p. 193.
8. J.C. Phillips, J. Non-Cryst. Solids **34**, 153 (1979); J. Non-Cryst. Solids **47**, 203 (1983).
9. H. He and M. F. Thorpe, Phys. Rev. Lett. **54**, 2107 (1985).
10. J.C. Phillips, in *Rigidity Theory and Applications*, ed. by M.F. Thorpe and P. Duxbury, (Michigan State University Press, East Lansing, 1999) to be published.
11. G. Lucovsky and J.C. Phillips, J. Non-Cryst. Solids **227**, 1221 (1998).
12. J. H. Van der Merwe, J. Appl. Phys. **34**, 123 (1963).
13. V. Misra et al., submitted to IEEE Electron Device Trans. (1999).
14. L. Manchanda, IEDM Tech Dig. 605 (1998).
15. H.Y. Yang, H. Niimi and G. Lucovsky, J. Appl. Phys. **83**, 2327 (1998).
16. A. Chatterjee et al, IEDM Tech. Dig. 779 (1998).

Modeling the Trends in Valence-Band Electron Tunneling in NMOSFETs with Ultrathin SiO_2 and SiO_2/Ta_2O_5 Dielectrics with Oxide Scaling*

A. Shanware,[1] H. Z. Massoud,[1] E. Vogel,[2] K. Henson,[2] J. R. Hauser,[2] and J. J. Wortman[2]

[1]Semiconductor Research Laboratory, Department of Electrical and Computer Engineering,
Duke University, Durham, North Carolina 27708-0291, U.S.A.

[2]Department of Electrical and Computer Engineering,
North Carolina State University, Raleigh, NC 27695, U.S.A.

Gate oxide scaling in NMOSFETs causes electrons to tunnel from the conduction and valence bands of the silicon substrate in the direct-tunneling regime. In NMOSFETs, the tunneling of electrons from the substrate's valence band is a source of the substrate current I_B and contributes to the gate current I_G. Oxide thickness scaling leads to an increase in the substrate current I_B and in the ratio I_B/I_G of substrate to gate current. In this paper, we report the trends in the I_B/I_G ratio due to oxide thickness scaling in ultrathin SiO_2 and SiO_2/Ta_2O_5 composite gate dielectrics.

1. INTRODUCTION

The gate current of NMOSFETs with ultra-thin gate dielectrics consists of tunneling contributions from the conduction and valence bands of the substrate. Electrons in the valence band of the substrate of NMOSFETs biased in inversion tunnel through the gate dielectric and lead to the generation of holes in the substrate. These holes are collected at the substrate contact and contribute to the substrate current provided the device is biased such that the valence-band edge in the substrate is at an energy level higher than the conduction-band edge in the gate [1–5]. In this study, we find that the contribution of this tunneling mechanism to the total gate current increases with oxide scaling, and compare the trends in tunneling-induced substrate currents in NMOSFETs with SiO_2 gate oxides and SiO_2/Ta_2O_5 composite dielectrics with the same oxide-equivalent thickness. We show that the benefits obtained from composite dielectrics depend on the device bias and the dielectric-stack composition.

2. EXPERIMENTAL RESULTS

In the characterization of tunneling-induced substrate currents it is important to separate this current component from other sources of substrate current such as hot carriers or junction leakage by proper selection of the applied bias. The substrate current was measured at zero or low drain bias to avoid the generation of hot carriers. Experiments were carried out on NMOS-FETs with gate oxide thicknesses ranging from 1.5 to 3.5 nm. The oxide thickness was obtained from $C(V)$ characteristics while taking into account quantum-mechanical effects in the substrate [6].

Figure 1 shows the dependence of the substrate current density in NMOSFETs with gate oxide thicknesses ranging from 1.5 to 3.5 nm. The substrate current data exhibits the usual transition in the tunneling characteristics from direct to Fowler-Nordheim tunneling at a gate oxide voltage of approximately 4.2 V. Valence-band electron tunneling increases exponentially as the gate oxide thickness is scaled. The substrate current due to the contribution of valence-band electron tunneling is compared to the gate current. Figure 2 shows the experimental trends in the I_B/I_G ratio with gate oxide scaling. These results sug-

gest that the scaling of the gate oxide thickness increases the contribution of valence-band electron tunneling in the gate current. In NMOSFETs with 1.5 nm-thick gate oxides, the current carried by electrons tunneling from the valence band is found to be as much as 20% of the gate current.
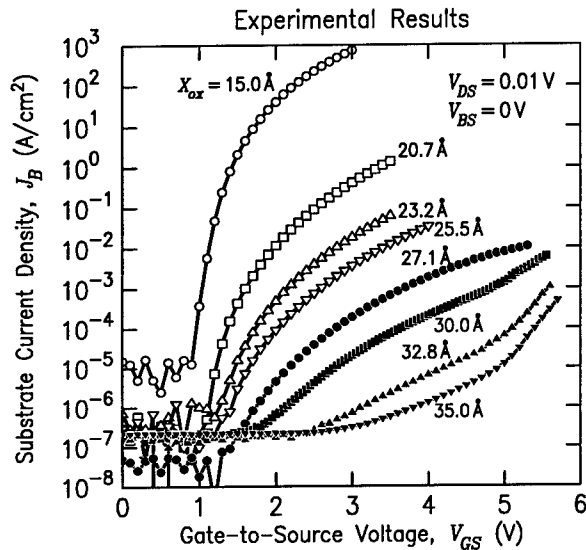
Fig. 1. Experimental results of the dependence of the substrate current density on the gate voltage for oxides ranging from 1.5 to 3.5 nm.
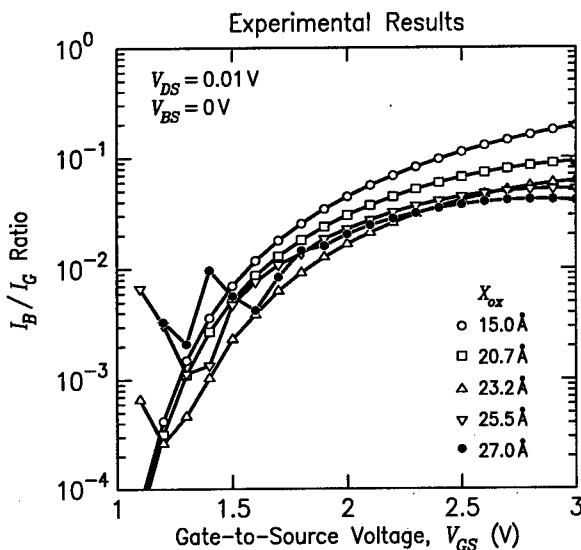
Fig. 2. Effect of the oxide thickness on the dependence of the I_B/I_G ratio on V_{GS} for $V_{DS} = 0.01$ V and $V_{BS} = 0$ V.

3. DISCUSSION AND SIMULATION RESULTS

The trend shown in Fig. 2 for the I_B/I_G ratio can be explained by calculating the ratio of the probability of electron tunneling from the valence band to that from the conduction band, as shown in Fig. 3, while assuming the WKB approximation and a trapezoidal barrier. The barrier heights were assumed to be 3.15 eV for electron tunneling from the conduction band and 4.20 eV for electron tunneling from the valence band. This result clearly suggests that the tunneling probability is responsible for the trends in Fig. 2.
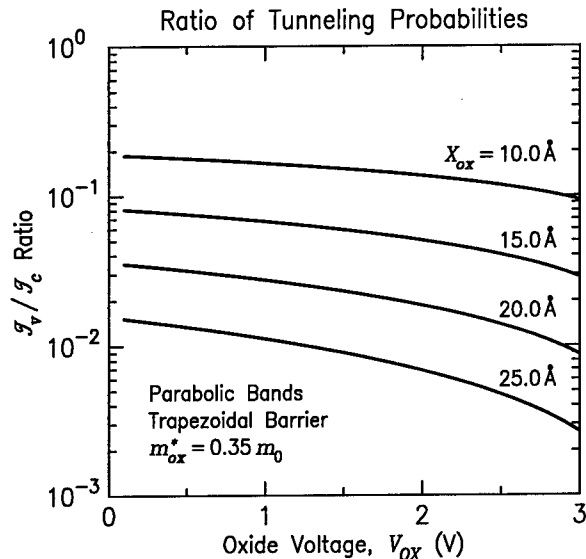
Fig. 3. Dependence of the ratio of valence-band to conduction-band tunneling probabilities on oxide voltage and thickness.

The use of composite dielectrics is widely considered as an alternative gate-oxide engineering solution for the reduction of the gate current of MOSFETs with ultrathin dielectrics. We now extend the discussion of composite dielectrics to the contribution of valence-band electron tunneling to the gate current. The trends in the I_B/I_G ratio due to gate oxide scaling are compared with devices with composite gate dielectrics. Simulations were carried out to study the impact of scaling on valence-band electron tunneling in NMOSFETs with pure and composite gate oxides. The composite oxide structure consists of a

thin layer of SiO_2 on the substrate followed by a layer of Ta_2O_5. The band diagram of the composite dielectric shown in Fig. 4 is based on values reported recently for the electron affinity and bandgap of Ta_2O_5 [7]. The dielectric stack is assumed to have the same capacitance as that of a 1.0 nm-thick layer of SiO_2. The thickness of the buffer SiO_2 layer was varied in steps. The dielectric constant of tantalum oxide was assumed to be 25 [8] .

Simulations were carried out for $V_{DS}=0$ V for a 50 nm gate length MOSFET, and the dependence of the gate and substrate currents are shown in Figs. 5 and 6, respectively.
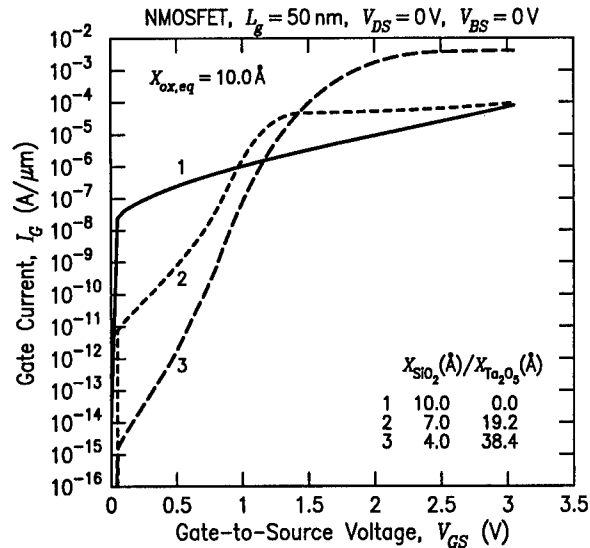
Fig. 5. Simulation of the dependence of the gate current I_G on V_{GS} in SiO_2 and SiO_2/Ta_2O_5 dielectrics.

Fig. 4. Energy-band diagram of silicon, silicon dioxide, and tantalum pentoxide and the composite-dielectric MOSFET.
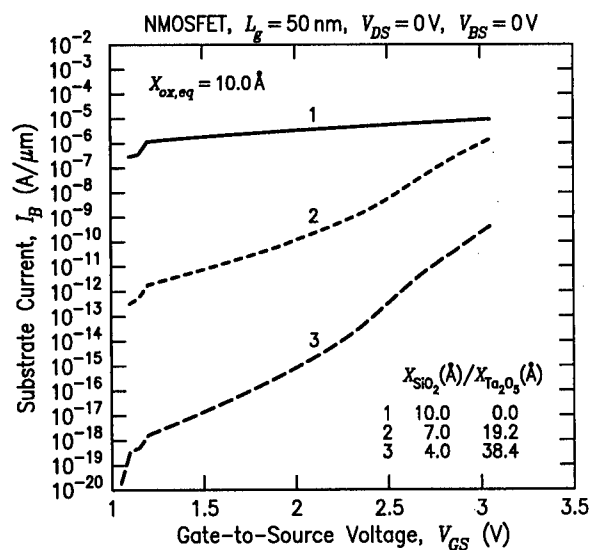
Fig. 6. Simulation of the dependence of the substrate current I_B on V_{GS} in SiO_2 and SiO_2/Ta_2O_5 dielectrics.

Details of the simulations are reported elsewhere [9]. The electron effective mass in the tantalum oxide was assumed to be $0.5\,m_0$ [7]. Three MOSFET structures with the same dielectric capacitance were simulated. The gate dielectric structures were 1.0 nm layer of SiO_2, a composite dielectric consisting of 7 Å of SiO_2 and 19.2 Å of Ta_2O_5, and a composite dielectric consisting of 4 Å of SiO_2 and 38.4 Å of Ta_2O_5. The tunneling current was calculated using the independent-electron tunneling model [10]. It is important to note that the benefit realized by substituting a composite silicon-dioxide/tantalum-oxide dielectric for a pure silicon dioxide layer in reducing the gate current is limited to low gate biases of less than 0.8 V. This limitation is a direct result of the small barrier height of tantalum oxide with respect to silicon. At larger gate biases, the tantalum pentoxide does not contribute to the tunneling barrier of electrons tunneling from the conduction band, and one is left with a thinner layer of silicon dioxide resulting in a larger gate tunneling current flowing through the composite. By comparison, the substrate current resulting from electron tunneling from the valence band is greatly reduced for all gate biases of interest because of the contribution of the tantalum pentoxide to the barrier for electrons tunneling from the valence band.

Our simulations show that the ratio I_B/I_G is a strong function of the bias and the dielectric-stack composition. We found that for an equivalent SiO_2 thickness of 10 Å at a gate bias of 1.5 V, this ratio was reduced from 0.56 for a pure SiO_2 gate dielectric to 1.59×10^{-7} for a composite gate dielectric consisting of 7 Å of SiO_2 and 19.2 Å of Ta_2O_5, as shown in Fig. 7. The substantial reduction in the contribution of the substrate current to the gate current is largely due to the considerable reduction in the substrate current.

4. CONCLUSIONS

Experimental results indicate that the contribution of valence-band electron tunneling to the gate current of NMOSFETs increases with the scaling of the gate oxide thickness. This scaling increases the I_B/I_G ratio of the electrons tunneling from the valence band to the gate current. Simulations show that the I_B/I_G is a strong

function of the oxide composition and is greatly reduced by decreasing the thickness of the SiO_2 layer in the composite gate dielectric.
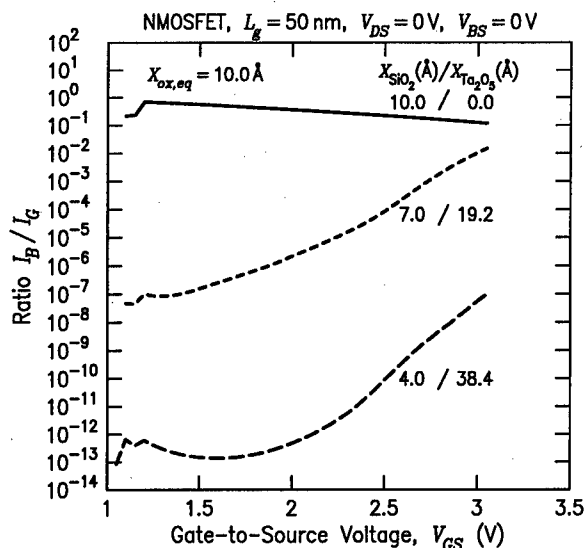
Fig. 7. Simulation results of the dependence of the I_B/I_G ratio on V_{GS} in SiO_2 oxides and SiO_2/Ta_2O_5 dielectric.

REFERENCES

1. A. Modeli, Appl. Surf. Sci., 30, 298 (1987).
2. B. Eitan, et al., Appl. Phys. Lett., 43, 106 (1983).
3. B. Majkusiak, IEEE Trans. on Elec. Dev., 37, 1087 (1990).
4. C. Chang, et al., J. Appl. Phys., 57, 302 (1985).
5. C. Chang, et al., p. 194, IEDM Digest, 1983.
6. J. R. Hauser et al., Characterization and Metrology for ULSI Technology Symposium, APS, 1998.
7. B. C.-M. Lai and J. Y.-M. Lee, J. Electrochem. Soc., 146, 266 (1999).
8. P. K. Roy et. al., Appl. Phys. Lett., 72, 22 (1998).
9. A. Shanware et al., to be published.
10. J. Simmons , J. Appl. Phys., 34, 1793 (1963).

A Novel Low-Temperature (Ba,Sr)TiO$_3$ (BST) Process with Ti/TiN Barrier for Gbit DRAM Applications

G. Beitel, H. Wendt, E. Fritsch, V. Weinrich, M. Engelhardt, B. Hasler, T. Röhr, R. Bergmann, U. Scheler, K.-H. Malek, N. Nagel, A. Gschwandtner, W. Pamler, W. Hönlein, C. Dehm and C. Mazuré

Infineon Technologies AG, Otto-Hahn-Ring 6, D-81739 Munich, Germany,
Tel. +49 89 234 50853, Fax +49 89 234 53263

A new, low temperature (Ba,Sr)TiO$_3$ (BST) MOCVD process has been established at 580°C deposition temperature which can be used for Gbit DRAM applications using Ti/TiN as barrier material. The process window for BST deposition was investigated in terms of deposition temperature, stoichiometry, film thickness, post annealing treatment and variation of the underlying electrode/barrier layer. Electrical characterization revealed specific capacitance values of 45 fF/µm^2 for 25 –30 nm film thickness and 75 fF/µm^2 for 10 nm film thickness which is close to the target value for GBit of 80 – 100 fF/µm^2. Oxidation resistance of the Ti/TiN barrier could be shown up to 600°C. Feasibility of this low temperature BST process has been successfully demonstrated using a 4 Mbit test vehicle.

1. Introduction

For Gbit DRAM applications, (Ba,Sr)TiO$_3$ (BST) is a promising material since it has a high dielectric constant which enables the use of this material for several generations. A typical schematic cross-section of a stacked capacitor cell is shown in Fig. 1. A barrier layer between Pt bottom electrode and poly silicon plug inhibits diffusion of Si, Pt and oxygen. The standard process for BST MOCVD deposition requires high deposition temperatures of 640°C followed by a post anneal at or above 550°C, both performed in oxygen [1]. Due to this considerable thermal budget in oxygen atmosphere, a well characterized standard Ti/TiN barrier will be oxidized and therefore cannot be used for BST integration. On the other hand other three-component barrier materials reported like TiAlN or TaSiN [2] require precise control of stoichiometry as well as further tools and therefore increasing manufacturing costs. In this work, for the first time we describe a low temperature BST process with an optimized Ti/TiN barrier which was successfully integrated into a 4Mbit test vehicle.
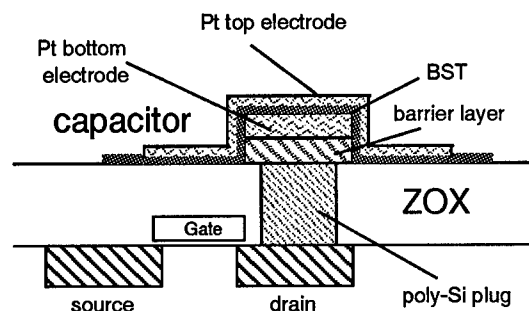
Fig. 1: Schematic cross-section of a stacked capacitor cell. Between the Pt bottom electrode and the poly silicon plug a barrier layer is necessary to inhibit diffusion of Si, Pt and O$_2$.

2. BST deposition process

BST films were deposited by liquid source MOCVD. For process development 100 nm Pt bottom electrodes were sputtered on thermally

oxidized Si or on Ti/TiN coated Si. 50 nm Pt top electrodes were evaporated through a shadow mask. Reproducibility and homogeneity of the BST films concerning stoichiometry and film thickness were within the target value of 5%. The influence of BST stoichiomety (i.e., Ti concentration) on specific capacitance and leakage current was investigated in a range between 35 and 60 at%. The best electrical properties were obtained for a Ti concentration of 51 - 52 at%. After BST deposition a post annealing procedure of BST films in oxygen at temperatures between 400°C and 600°C is essential for reduction of the leakage current.

The BST film composition and deposition rate are stable for deposition temperatures between 520 and 670°C. BST morphology, however, was found to be strongly dependent on deposition temperature. While BST deposited at 640°C showed a crystalline columnar structure, films become amorphous at lower temperatures (Fig. 2).
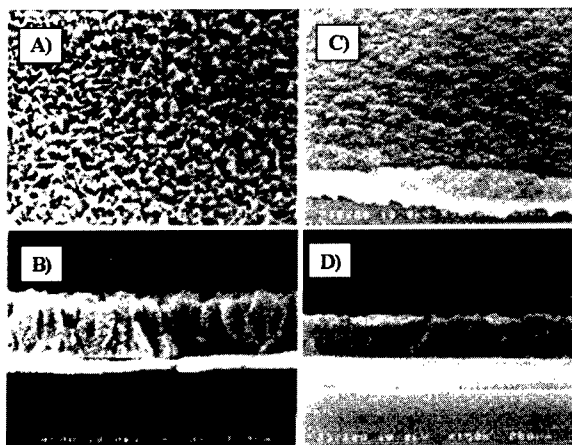
Fig. 2: SEM analysis of BST films, 100 nm BST on Pt/SiO$_2$ (top view (A, C) and cross section (B, D)), BST deposition temperature: A) and B): 640°C; C) and D) 540°C.

Most probably due to these morphology changes the specific capacitance of BST films also is strongly temperature dependent. The dependence of specific capacitance and leakage current on BST deposition temperature is shown in Fig. 3. For 25 nm BST films specific capacitance decreases from 70 fF/μm^2

(640°C dep. temp.) to 20 fF/μm^2 (520°C dep. temp.). BST deposited at lower temperatures shows excellent leakage currents. j$_L$ values of 10^{-8}A/cm^2 (U = ±1V) are obtained for dep. temp. > 560°C and BST thicknesses down to 8 nm. Furthermore, dielectric constant and specific capacitance are strongly dependent on BST film thickness (Fig. 4): while the dielectric constant decreases, a strong increase in specific capacitance can be observed for a decreasing film thickness.
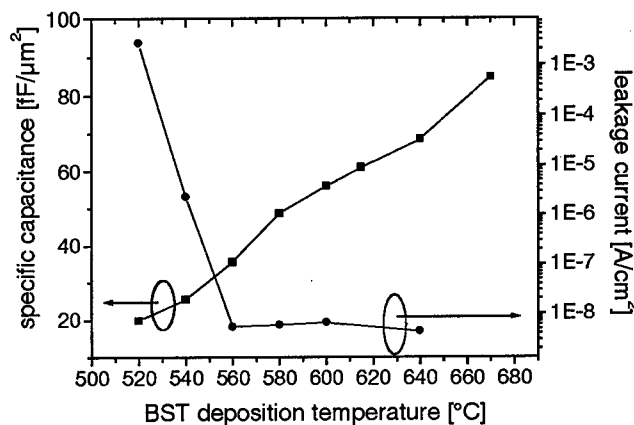
Fig. 3: Specific capacitance and leakage current vs. BST deposition temperature (BST film thickness 25 nm, electrical characterization at RT).
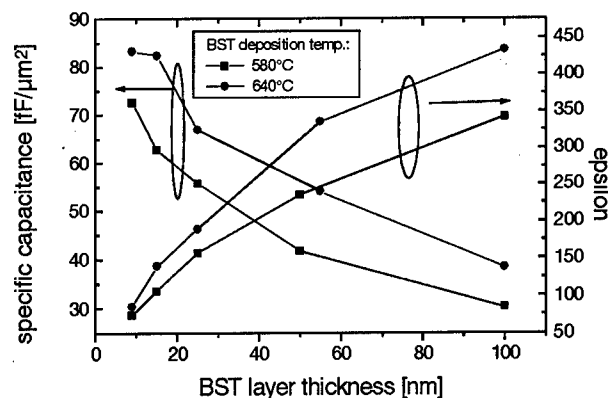
Fig. 4: Specific capacitance and epsilon vs. BST film thickness for films deposited at 640°C and 580°C (electrical characterization at RT).

For a deposition temperature of 580°C specific capacitance values are around 55 fF/μm^2 for 25 - 30 nm films and 75 fF/μm^2 for 10 nm films. This is only 10 - 20% lower than the values obtained for BST deposited at 640°C (standard deposition temperature) which shows the potential of the low temperature BST process for the fabrication of Gbit storage capacitors.

3. Barrier/electrode/BST Module Formation

To achieve high oxidation resistance of the Ti/TiN barrier, the barrier and Pt deposition process had to be optimized. The 20 nm Ti/100 nm TiN barrier was deposited in a way that the O_2 incorporation and the mechanical stress of the layers are low. Pt was sputtered after air exposure of the barrier at 0.5 KW, 3 mbar at temperatures between 100 and 500°C. No blistering or bubble formation occured during subsequent BST-deposition and O_2 post anneal. The Ti/TiN/Pt barrier/electrode stack is oxidation resistant up to 600°C. The complete BST/Pt/barrier stack utilizing 25 nm BST film deposited at 580°C on Pt/TiN/Ti has been investigated by AES. The AES depth profile reveals sharp interfaces between BST and Pt and between Pt and TiN (Fig. 5).
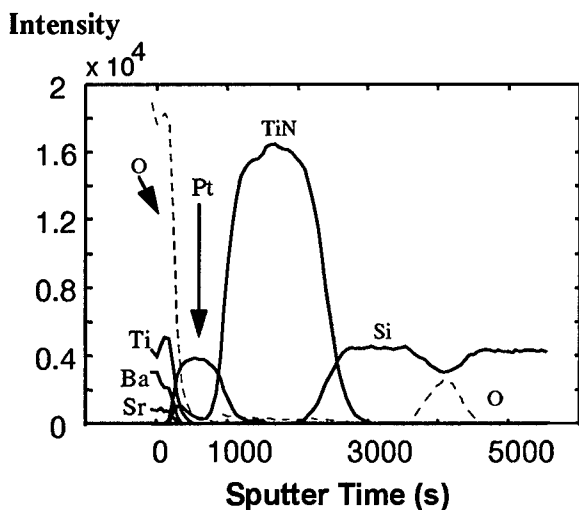
Fig. 5: AES depth profile of BST/Pt/Ti/TiN/poly-Si/SiO$_2$ stack.

This indicates no oxidation of the barrier during BST formation. Finally a typical cross section of the 4Mbit stacked capacitor integrated with the 580°C low temperature BST process is shown in Fig. 6. A homogeneous BST film has been formed between Pt bottom and top electrode and successfully integrated into a 4Mbit stacked capacitor cell.
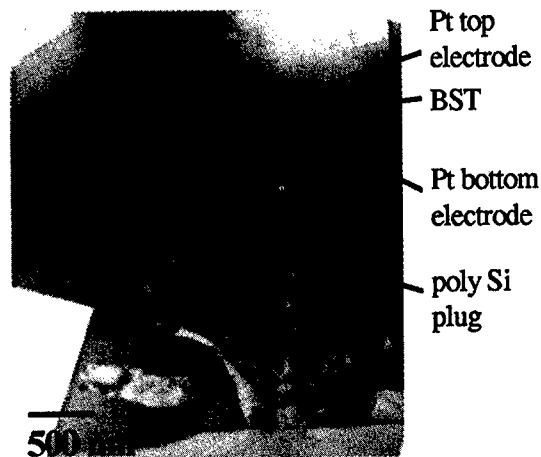
Fig. 6: TEM cross section of a 4Mbit-cell.

A closer cross-section TEM analysis showed clearly no interfacial oxide formation between the Ti/TiN barrier and the poly silicon plug. Standard DRAM product tests showed correct functionality, and, therefore, feasibility of our newly developed stacked capacitor process with low temperature BST.

4. Conclusions

A new low temperature MOCVD BST process as well as an optimized Ti/TiN barrier have been developed. A decrease of BST deposition temperature has a significant effect on BST morphology and dielectric properties. For thin BST films (10 nm) deposited at low temperature specific capacitance values as high as 75 fF/cm^2 were obtained. This shows that BST is able to meet Gbit requirements. Feasibility of this new stacked capacitor process has been successfully demonstrated by integration of BST into a 4Mbit test vehicle.

Acknowledgements

References

[1] S. Bilodeau et al.. Solid State Technology 7, 1997, p.235.

[2] R.B.Khamankar et al., IEDM, Washington, DC (1997).

ELSEVIER

Charge Redistribution at GaN-Ga2O3 Interfaces: A Microscopic Mechanism for Low Defect Density Interfaces in Remote Plasma Processed MOS Devices Prepared on Polar GaN Faces

R. Therrien, H Niimi, T. Gehrke, G. Lucovsky and R. Davis

Departments of Physics and Materials Science and Engineering, North Carolina State University, Raleigh, NC 27695-8202

Interfacial defect densities are typically two orders of magnitude higher at [III-V]-dielectric interfaces than at Si-SiO$_2$ interfaces. This paper demonstrates GaN devices with significantly reduced interfacial defect densities using a two step remote plasma process to form the GaN-dielectric interface and then deposit the dielectric film. Separate plasma oxidation and deposition steps have previously been used for fabrication of aggressively scaled Si devices. Essentially the same 300°C remote plasma processing has been applied to GaN metal-oxide semiconductor (MOS) capacitors and field effect transistors (FETs). This paper i) discusses the low temperature plasma process for GaN device fabrication, ii) briefly reviews GaN device performance, and then iii) presents a chemical bonding model that provides a basis for the improved interface electrical properties.

1. INTRODUCTION

The potential of GaN for use in high frequency, high temperature and high power applications has led to extensive research in the development of GaN microelectronic devices. Studies of metal-insulator-semiconductor (MIS) structures using GaN have been performed with deposited SiO$_2$ and Si$_3$N$_4$ acting as the gate insulator [1-3]. Casey et al. [1] produced a metal-oxide-semiconductor capacitor using remote plasma enhanced chemical vapor deposited (RPECVD) SiO$_2$ on n-GaN. The study indicated a flat-band voltage of -2.35V, which corresponds to a fixed positive charge of 10^{12} cm^{-2}. There was no hysteresis present in the measured C-V which these authors attributed to an absence of interface traps.

MIS capacitors were also fabricated by Arulkumaren et al. [2] using electron beam (EB) evaporated and plasma enhanced chemical vapor deposited (PECVD) SiO$_2$ and PECVD Si$_3$N$_4$ on n-type GaN. A minimum interface state density of 2.5 × 10^{11} eV^{-1}·cm^{-2} was measured in the PECVD deposited SiO$_2$/GaN capacitor. Measurements on the EB evaporated SiO$_2$/GaN and the Si$_3$N$_4$/GaN capacitors yielded interface trap densities of the 5.3 × 10^{11} eV^{-1}·cm^{-2} and 6.5 × 10^{11} eV^{-1}·cm^{-2}, respectively.

Binari et al. [3] fabricated depletion mode MISFETs using PECVD deposited Si$_3$N$_4$ as the gate insulator. The devices displayed a maximum transconductance of 16 mS/mm and a pinchoff at −50V. The author attributed the low transconductance to a high density of interface s traps apparent from the large hysteresis in the capcitance-voltage (C-V) measurements of the structure.

Ren et. al. [4,5] have demonstrated MIS devices on GaN using a deposited Ga$_2$O$_3$(Gd$_2$O$_3$) as the gate insulator. The C-V measurements on theses devices are consistent with a large number of interface traps, perhaps as high as 10^{12} eV^{-1}·cm^{-2}.

In this paper, we report on the fabrication of MOS capacitors and depletion mode FETs using RPECVD SiO$_2$ on n-GaN.

2. DEVICE FABRICATION

The GaN used for this study was grown via metal-organic chemical vapor deposition (MOCVD) on 6H-SiC (0001) substrates using a AIN buffer layer. The films were intentionally doped n-typed to 1-3 × 10^{17} cm^{-3}.

MOS devices were formed on the epitaxial GaN (0001) layers with Ga-atom terminated faces. The devices were fabricated using using a two step remote plasma process to first form the GaN-dielectric interface and then deposit the dielectric film. Separate plasma oxidation and deposition steps have previously been used for fabrication of aggressively scaled Si devices [6]. Prior to in-situ remote plasma-assisted oxidation (RPAO), the GaN surface was cleaned by exposure to reactive species

from a remote N N_2/He discharge. On-line analysis by Auger electron spectroscopy (AES) demonstrated the exposure to products of the N_2/He discharge resulted in significant reductions in C impurities to below AES detection limits, but residual sub-monolayer O remained (see Figs. 1 and 2). RPAO was then performed using an O_2/He plasma at 300°C followed by the 300°C remote plasma-enhanced CVD of SiO_2 using plasma excited O_2/He and downstream SiH_4. After SiO_2 film deposition, a 900°C anneal in an inert ambient was done for interface and bulk film relaxation. Device structures were fabricated by conventional lithography, metallization, and post metallization annealing.
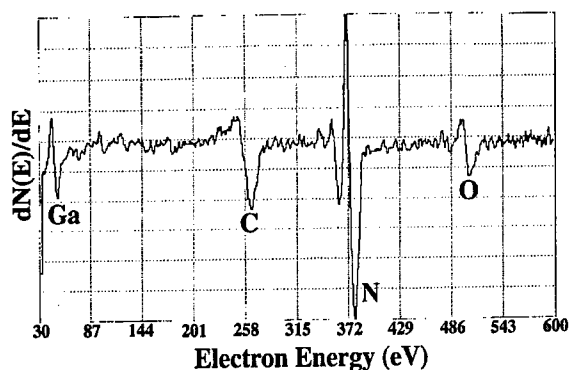
Figure 1. AES of as-loaded GaN surface
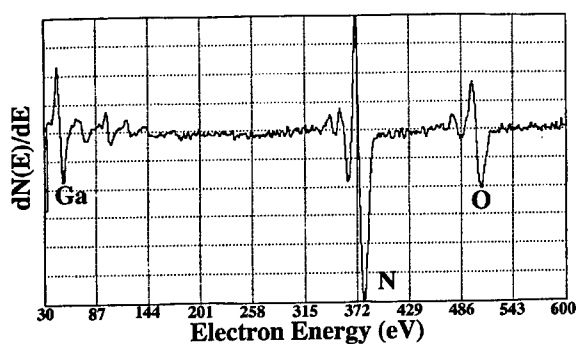
Figure 2. AES of remote N_2 plasma cleaned GaN

The structures of the MOS capacitors and FETs can be seen in Figs. 3 and 4, respectively. The MOS capacitors required a top side ohmic contact due to the insulating AlN buffer layer. Al was used for all the contacts on the devices in this study.

Capacitance-voltage measurements were performed using a HP 4284A variable frequency LCR meter at

frequencies ranging from 100 Hz to 1 MHz. The curves were obtained by sweeping the voltage from depeltion (- voltage) to accumulation (+ voltage) and back to depletion at a rate of 0.1 V/s with 0.1 V steps. Electrical measurements on the FETs were performed with a Keithly 236 source-measure unit and a Keithly 237 high voltage source measure unit.
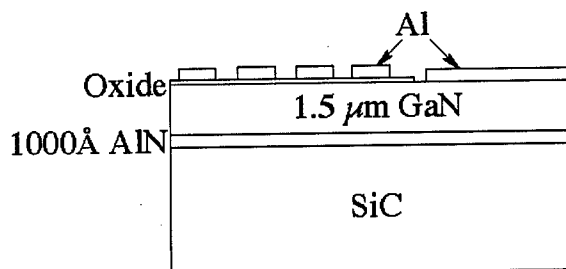
Figure 3. SiO_2/GaN capacitor cross-section
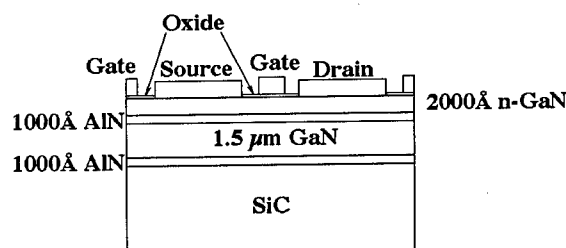
Figure 4. GaN MOSFET cross-section

3. RESULTS AND DISCUSSION

Figure 5 shows the C-V behavior of a SiO_2/GaN MOS capacitor fabricated with a RPAO step.
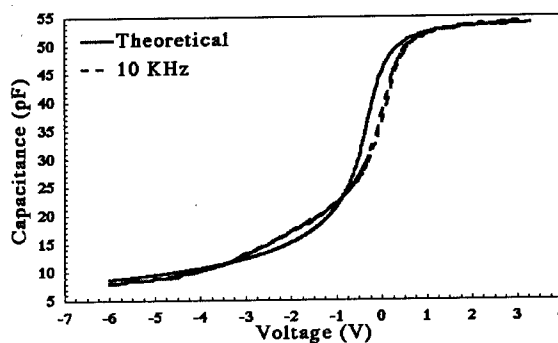
Figure 5. C-V of SiO_2/GaN MIS capacitor

An estimate of the number of traps per cm^2 was calculated from the hysteresis at flatband of the measurement taken at 10 kHz, on a capacitor with an area of 4.9 × 10^{-4} cm^2 where ΔV_{FB} 0.025 V to be N = 1.7 × 10^{10} cm^{-2}. A comparison of the ideal flatband voltage of -0.14 V for an Al gate and GaN with a n-doping level of 3 × 10^{16} cm^{-3} to the value of the flatband voltage of the curve swept in the forward direction yields the total amount of charge per cm^2 for this capacitor to be 1.6 x 10^{11} cm^{-2}.

In comparison the MOS capacitors fabricated without the use of the RPAO step demonstrated significantly increased interface defects (See Fig. 6). As the measurement frequency is lowered an increased number of traps are able to respond to the applied gate voltage. In contrast, the C-V behavior of the RPAO treated capacitors did not change with frequency consistent with significantly reduced trap densities.

The DC drain characteristics of a GaN MISFET with a gate length of 75 μm and a source to drain separation of 125 μm is shown in Fig. 7. The device showed a maximum g_m of 0.3 mS/mm and the gate current was typically below.1 nA for the gate voltages used in these experiments. The small
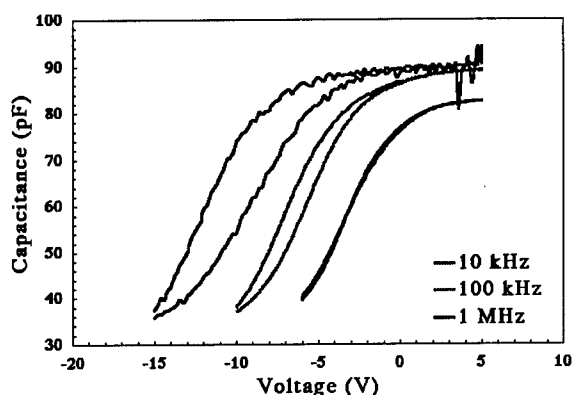
Figure 6. C-V of SiO$_2$/GaN MIS capacitor fabricated without a RPAO step

drain currents are due to the low doping in the n-type GaN layer and the poor ohmic behavior of the source and drain contacts, shown in Fig. 8. Figure 7 shows how increases in the gate voltage above 0 V resulted in increases in the drain current as more carriers (i.e. electrons) are attracted to the GaN/SiO$_2$ interface. Pinchoff occurred at a gate bias of 1.5 V, which corresponded well with the calculated value of -0.8V.

There are significant differences between GaN-dielectric and Si-SiO$_2$ interfaces: i) there is no GaN-oxide analog of SiO$_2$, ii) plasma or thermal oxidation of GaN results in formation of Ga$_2$O$_3$ and evolution of gaseous nitrogen oxides, and iii) the electron occupation of Ga dangling bonds on polar
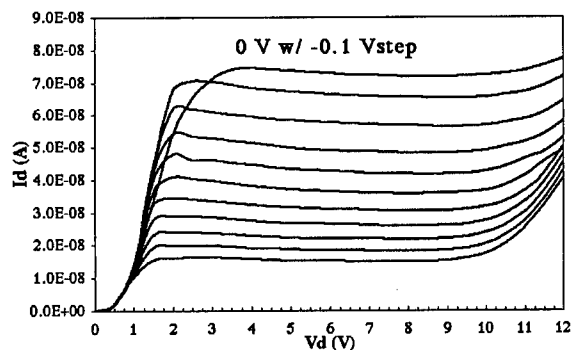
Figure 7. DC drain characteristics of GaN MISFET
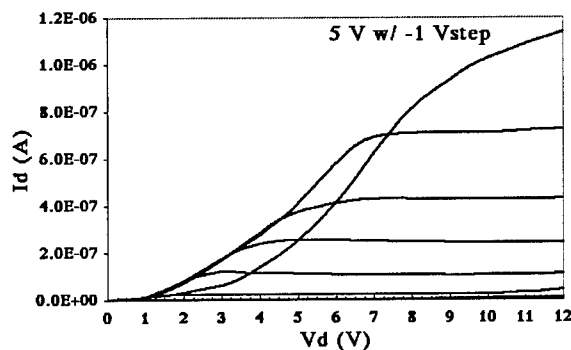
Figure 8. DC drain characteristics of GaN MISFET

aces of GaN is formally fractional rather than integral. Fractional occupancy means that there is an inherent mismatch between electronic charge and nuclear charge for interface bonding with deposited dielectrics such as SiO$_2$ or Si$_3$N$_4$; i.e. the bonding is heterovalent [7]. To form device quality GaN MOS devices, there must then be a redistribution of electronic charge in the Ga atom dangling bonds that is integral; this takes place during the plasma oxidation that forms the ultra thin Ga$_2$O$_3$ layer. The atomic structure of Ga$_2$O$_3$ has two bonding components in a 3:1 ratio: three tetrahedral [GaO$_{4/2}$]- groups that form a covalently-bonded network for each Ga^{+3} ion that occupies an octahedral interstitial or network modifier site within that network [8]. The improved performance in plasma-processed GaN MOS devices is consistent

with a redistribution of charge in Ga dangling bonds during the plasma oxidation step that matches the 3:1 distribution in Ga_2O_3 forming a isovalently bonded interface: i.e., 4 Ga(3/4e)o --> 3Ga(1e)-1/4 + Ga(0e)+3/4. Bonding at the internal Ga_2O_3-SiO_2 interface within the composite gate stack is also isovalent, with bonding connections between neutral $[SiO_{4/2}]$o and charged $[GaO_{4/2}]$- which is analogous to the bonding between $[AlO_{4/2}]$- and $[SiO_{4/2}]$o in zeolites

4. CONCLUSIONS

Experimental data has demonstrated that SiO_2/GaN capacitors and FETs with low interfacial defect densities are produced when the GaN surface is treated with the RPAO prior to the SiO_2 film deposition. We propose that the Ga_2O_3 layer formed during the RPAO provides an isovalent bonding arrangement for the SiO_2 as opposed to the heterovalent that occurs when SiO_2 is deposited directly on a polar surface of GaN. The isovalent bonding allows for the reduction of dangling bonds at the interface thus reducing the interfacial defect density.

Finally, the same bonding model has been extended to other [III-V]-dielectric stacks such as GaAs-Ga_2O_3-SiO_2.

ACKNOWLEDGEMENTS

REFERENCES

1. H. C. Casey, Jr., G.G. Fountain R.G. Alley, B . P. Keller, S. P. Denbaars, Appl. Phys. Lett. **68**, 1850 (1996).
2. S Arulkumaran, T. Egawa, H.Ishikawa, T. Jimbo, M. Umeno, Appl. Phys. Lett. **73**,   809 (1998).
3. S.C. Binari. L.B. Rowland, G. Kelner, W. Kruppa, H.B. Dietrich, K.Doverspike, D.K. Gaskill, Proc. 21ˢᵗ Int. Symp. On Compound Semicond., San Diego CA, Aug. 1994
4. F. Ren, C.R. Abernathy, J.D. MacKenzie, B.P. Gila, S.J. Pearton, M. Hong, M. Macos, M.J. Schurman, A.G. Baca, R.J. Shul, Mat. Res. Soc. Symp. Proc. **483**, 443 (1998).
5. F. Ren, M. Hong, S.N.G. Chu, M.A. Marcus, M.J. Schurman, A. Baca, S.J. Pearton, C.R. Abernathy, Appl. Phys. Lett. **73**, 3893 (1998).
6. T. Yasuda, Y. Ma, S. Habermehl and G. Lucovsky, Appl. Phys. Lett. **60**, 434 (1992).
7. W.A. Harrison, E.A. Kraut, J.R. Waldrop and R.W. Grant, Phys. Rev. **B 18**, 4402 (1978).
8. G. Lucovsky, A. Rozaj-Brvar and R.F. Davis, in *The Structure of Non-Crystalline Materials 1982*, ed. by P.H. Gaskell, J.M. Parker and E.A. Davis (Taylor and Francis, London, 1982), p. 193, and Refs. therein.

The effects of interfacial suboxide transition regions on direct tunneling in oxide and stacked oxide-nitride gate dielectrics

H. Yang, H Niimi, Y. Wu, G. Lucovsky, J.W. Keister and J.E. Rowe

Departments of Physics, Materials Science and Engineering, and Electrical and Computer Engineering,
North Carolina State University, Raleigh, NC 27695-8202

This paper builds on previous work that has demonstrated the effects of interfacial suboxide transition regions at Si-SiO$_2$ interfaces on tunneling oscillations in the Fowler-Nordheim regime. This paper extends thes effects to the direct tunneling regime and focuses on differences in interfacial transition regions between Si-SiO$_2$ interfaces with, and without monolayer level interface nitridation. Tunneling currents in devices with the same oxide-equivalent thickness are reduced by monolayer level interfacial nitrogen with respect to devices without interface nitridation for i) substrate and gate injection and ii) in both the direct and Fowler-Nordheim tunneling regimes.

1. INTRODUCTION

Exponential increases in tunneling current with decreasing gate oxide thickness, t_{ox}, present a significant limitation in aggressive scaling of Si CMOS devices for $t_{ox} < 2.0$ nm. The 1997 SIA NTRS roadmap [1] requires an oxide-equivalent thickness, $t_{ox,eq}$, < 1 nm by 2012, stimulating research on alternative insulators with dielectric constants, k, higher than SiO$_2$. This in turn has generated near-term interest in nitrided oxides, oxide/nitride stacks [2-5] and oxynitride alloys [6].
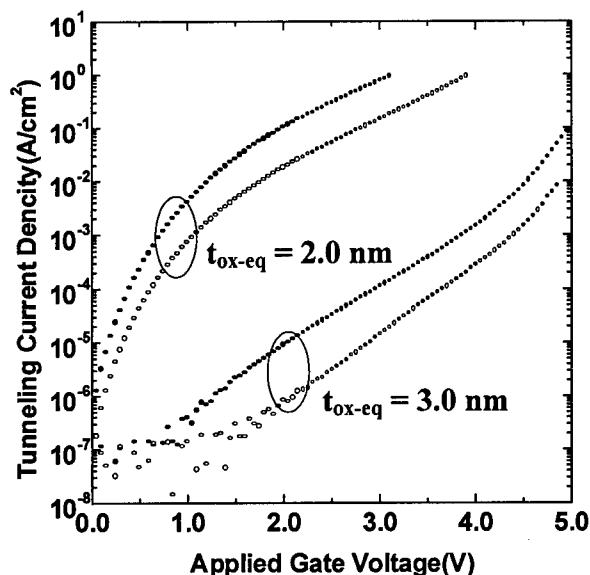
Figure 1. J-V Characteristics of capacitors with ultra thin gate oxides in the direct tunneling regime.

Studies by our group [3,7] have demonstrated that tunneling is reduced in devices with monolayer plasma-assisted Si-SiO$_2$ interface nitridation as compared with devices with non-nitrided interfaces and the same $t_{ox,eq}$. Direct tunneling is reduced by factors of about 8±2 for both substrate and gate injection for t_{ox-eq} in the range of 1.8 to 3 nm. Reductions are not due to i) flatband voltage shifts [7], ii) increases in physical thickness [7], and/or iii) decreases in interface roughness [8]. Maserjian et al. were first to recognize interfacial transition regions with suboxide bonding modified Fowler-Nordheim tunneling [9]. This paper extends their approach into the direct tunneling regime, using X-ray photoelectron spectroscopy (XPS) to estimate widths of interfacial transition regions [10].

2. EXPERIMENTAL RESULTS

Figure 1 displays J-V plots in the direct-tunneling regime for devices prepared by combined remote plasma-assisted and rapid thermal processing [7,11]. Characterizations used to establish interfacial nitridation, and to quantify interfacial N content are detailed in Refs 7, 11 and 12. Etch-back combined with Auger electron spectroscopy, angle-resolved XPS (ARXPS) and non-linear optical second harmonic generation (SHG) demonstrated interfacial N-localization; secondary ion mass spectrometry (SIMS) and nuclear reaction analysis (NRA) were used to establish the N content, identifying process conditions for achieving monolayer N-concentrations of $7\pm1\times10^{14}$ cm^{-2} at Si(100)-SiO$_2$ interfaces [7,11].

XPS studies were performed on Si(111) and Si(100) substrates with similar results: i) interfaces showed integrated spectral content in the suboxide bonding regime between the 2p Si substrate feature at ~ 99 eV and the SiO_2 bulk oxide feature at ~ 103 eV in excess of the one monolayer defining the interface, ii) integrated intensities in this region decreased by 15-20 % between as-deposited (300°C) and annealed conditions (900°C) consistent with other studies [13], and iii) decreases in suboxide bonding occurred mostly in bonding arrangements not intrinsic to a particular surface; e.g., the Si^{2+} feature decreased markedly on Si(111) after the 900°C anneal [10]. Other studies have demonstrated that interfaces formed by thermal oxidation at 900°C display additional changes in interface morphology after 900°C non-oxidizing anneals [14], emphasizing *fundamental differences* in interface morphology for growth and annealing at the same temperatures.
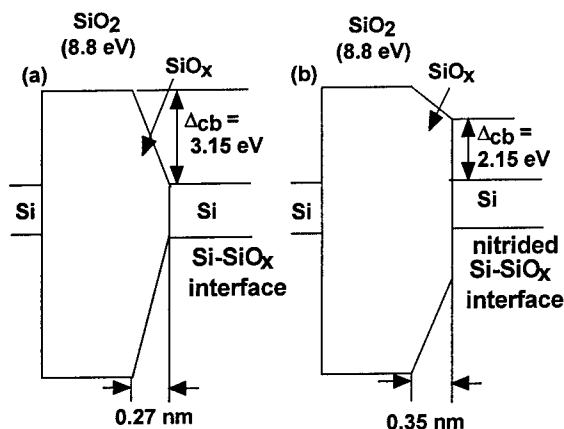
Figure 2 Oxide tunneling barriers with (a) non-nitrided and (b) nitride suboxide interfacial regions.

3. MODIFIED BARRIER LAYER MODEL

Figure 2 indicates modified interface band structure proposed by us for 900°C annealed (a) non-nitrided and (b) nitrided Si-SiO_2 interfaces. The suboxide bonding in (a) defines a transition region with an average SiO composition between the Si substrate and the SiO_2 dielectric. Following Ref. 9, a linear variation in the conduction band offset energy has been assumed. The width of this region was estimated from XPS results of Ref. 10. The areal density of Si atoms in a *mono*-molecular layer of SiO is estimated to be ~5.2±0.1x10^{14} cm^{-2} by averaging Si-atom areal densities in c-Si, ~7.8x10^{14} cm^{-2}, and SiO_2, ~2.6x10^{14} cm^{-2}. The thickness of this

molecular layer, 0.32±0.05 nm, was obtained by similar averaging. For non-nitrided interfaces, XPS data indicated an excess Si areal density of 4.4±0.4x10^{14} cm^{-2}, or approximately 0.85±0.1 of a molecular layer, corresponding to a transition region width of ~0.27±0.03 nm. For the nitrided interfaces the corresponding excess areal density increased to ~1.1 ML with a thickness ~0.35±0.03 nm. The observance of different resonance energies for nitrided and non-nitrided Si-SiO_2 interfaces in optical SHG measurements [12] establishes that conduction band offset profiles at these interfaces must be different. A potential step of 2.15 eV has been assumed at the nitrided Si-dielectric interface, and following Ref. 9, the remainder of transition region is represented in Fig. 2(b) by a linear variation between the top of the interface step and the SiO_2 conduction band edge.
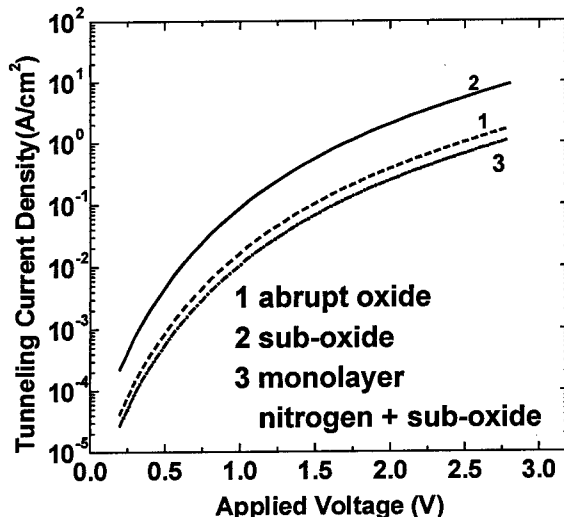
Figure 3 Calculated direct tunneling J-V curves for devices with n$^+$(5x10^{19} cm^{-3})Si–SiO_2–n(5x10^{17} cm^{-3})Si stacks with t_{ox-eq} ~ 2.0 nm; (1) suboxide transition region, (2) ideal oxide, and (3) nitrided suboxide transition region.

Figure 3 compares calculations of direct tunneling based on the WKB approximation [15,16] for three different dielectrics with t_{ox-eq} = 2.0 nm. These are i) an *ideal* oxide with no interfacial transition region (1), ii) an oxide with interfacial suboxide bonding (2), (t_{ox} = 1.77 nm), and iii) an oxide with interfacial suboxide bonding and a nitrided interface (3), (t_{ox} = 1.85 nm). The tunneling barrier is 3.15 eV, and the tunneling mass is 0.5 m_0 [15], where m_0 is the free electron mass. The determinations for t_{ox} for ii) and iii) set t_{ox-eq} = 2.0

nm, and assumed k in the suboxide region in ii) increased from 3.8 for SiO_2 to 6.7, and in iii) from 3.8 to 5.7. The calculations demonstrate that direct tunneling is reduced by interfacial nitridation ~8-10 with respect to devices with suboxide bonding. Additionally, the calculated tunneling in the device with the nitrided interface is nearly equal to that of the device with *ideal* abrupt interface; i.e. in the context of the WKB approximation, they both have approximately the same thickness-(tunneling mass-average barrier height)$^{0.5}$ products [15].

The *approximate* reduction in direct tunneling, R_{D-T}, is estimated using an analytic expression based on the exact solution of tunneling through an equivalent rectangular barrier, t_o. The transmission probability through such a barrier is given by:

$$T = \exp(-4\pi h^{-1} t_o (2m^*_i E_i)^{0.5}), \qquad (1)$$

where m^*_i is the tunneling mass and E_i and the effective barrier height. The *approximate* reduction in direct tunneling current, R_{D-T}, is then given by:

$$R_{D-T} \sim \exp \{1.02(t_n(m^*_{ox} \cdot E_n)^{0.5} - t_o(m^*_{ox} \cdot E_o)^{0.5})\}, \qquad (2)$$

where t_n and t_o are the widths of suboxide regions (in Å) with and without interface nitridation, E_n and E_o are average conduction band offset energies (in eV), and m^*_{ox} is the electron tunneling mass in units of the electron mass, m^*_o. From Fig. 2, $t_n = 3.5Å$, $t_o = 2.7Å$, $E_n \sim 2.7$ eV, $E_o \sim 1.6$ eV, and $m^*_{ox} \sim 0.5m^*_o$, so that $R_{D-T} \sim 6.5$, consistent with decreases in the calculated J-V curves of Fig. 3. Since the integrand in the WKB integral includes interface terms in the same way independent of oxide thickness [15], the *same reductions* in direct tunneling are expected for $t_{ox,eq} \sim 1.5$ to 3 nm in agreement with experiment [7,12]. Additionally, since the tunneling probability is independent of current direction, R_{D-T} is the same for substrate and gate injection consistent with results in Fig. 4.

An estimate of tunneling reductions, R_{F-N}, in the Fowler-Nordheim tunneling regime is not straight-forward, since the integrand in the WKB integral for gate injection does not explicitly include the interface region, but for substrate injection does. As shown in Fig. 5, experiments indicate a reduction, R_{F-N}, of approximately 10-12 for gate injection, which is increased to ~50-60 for substrate injection. Based on the discussion presented above, the difference in R_{F-N} for substrate and gate injection should be approximately a factor of 6, whereas the experiments yield an average R_{F-N} of about 5. Stated differently, R_{F-N} for substrate injection is approximated by multiplying the value for gate

injection by the *relative interface transmission* factor, R_{D-T}, of ~6 as calculated from Eqn. (1). The combined effect of these two contributions is 66±12, in good agreement with the experimentally observed value of ~50-60 from the data in Fig. 5. The value of $R_{F-N} \sim 11$, for gate injection is assumed to be related to an increase in the electric field in the triangular region of the tunneling barrier for with the physically thicker oxide the non-nitrided interface.
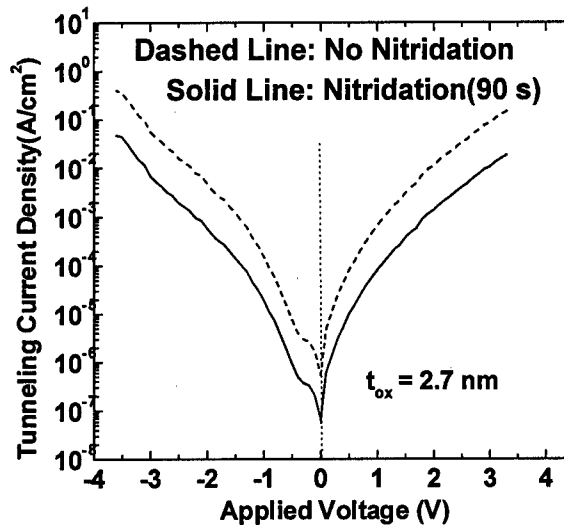
Figure 4 J-V Characteristics for substrate and gate injection for an NMOS FET in which the source, drain and substrate are grounded.

4. CONCLUSIONS

Experimental results have demonstrated that monolayer interface nitridation reduces tunneling currents in the direct tunneling regime by a factor of 8±2 compared to devices with non-nitrided interfaces [7,11], consistent with calculations. Decreases in direct tunneling due to interface nitridation must be considered in evaluating performance of aggressively-scaled devices with O/N stacks and oxynitride alloys [2-6], since *interfacial* nitridation can result from nitrogen transport during processing, including post-deposition anneals [3]. A breakthrough in device fabrication and performance has just been achieved by our group by combining monolayer interface nitridation with an aggressively scaled O/N stack in a pMOS FET. The O-layer is ~ 0.7 nm thick, and the RPECVD nitride layer is ~1.8 nm thick. $T_{ox,eq}$ is ~ 1.6 nm as determined from analysis of C-V data that included quantum mechanical corrections. For gate injection, the current at 1 volt applied bias was <

5×10^{-3} A/cm^2 compared to a tunneling current of ~ 1 A/cm^2 for an oxide dielectric.

The experimental results for the Fowler-Nordheim regime support the interfacial model developed for the direct tunneling regime. As expected on the basis of the analytical expressions for the WKB approximation, the reduction in tunneling current for interfacial nitridation for substrate injection should be larger than for gate injection because the integral in the exponent
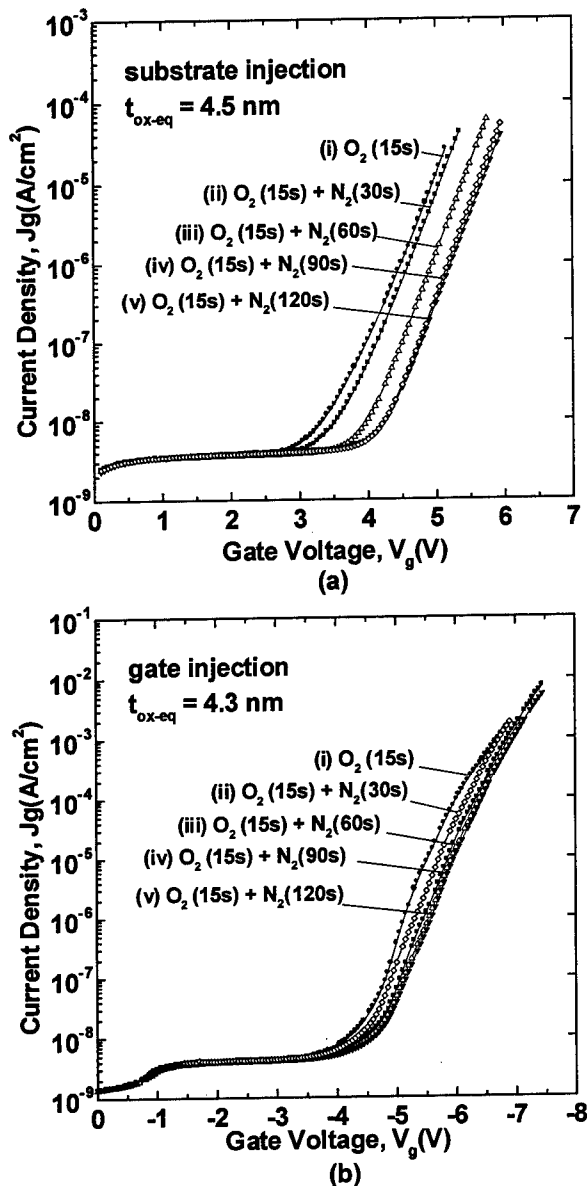
includes the interface region for substrate injection, but does not for gate injection.

Finally, the comparisons in Fig. 3 between calculated direct tunneling currents for *ideal abrupt* Si-SiO$_2$ interfaces, and those with suboxide transition regions demonstrate that transition regions must be considered in fitting experimental data to model calculations even for non-nitrided dielectrics. For example, determinations of fit parameters based on ideal interface models can lead to values of tunneling masses and/or band offset energies which are *effective values* due to neglect of the interfacial suboxide transition regions.

ACKNOWLEDGEMENTS

REFERENCES

[1] National Roadmap for Semiconductor Technology, (SIA, Santa Clara, CA, 1997).

[2] C. Parker et al., IEEE Electron Device Letters **19**, 106 (1998).

[3] Y. Wu and G. Lucovsky, IEEE Electron Device Letters **19**, 367 (1998).

[4] Y. Shi et al., IEEE Trans. on Electron Devices **46**, 362 (1999).

[5] S.C. Song et al., *IEDM Tech. Dig.* 375, (1998).

[6] X. Guo and T.P. Ma, IEEE Electron Device Letters **19**, 207 (1998).

[7] H. Niimi et al.,, in *Characterization and Metrology for ULSI Technology: 1998 International Conference*, ed. by D.G. Seiler et al., (AIP, Woodbury, NY 1998), p. 273.

[8] H.C. Lin et al., J. Vac. Sci.Tech. **A 15**, 790 (1997).

[9] J. Maserjian and N. Zamani, J. Appl. Phys. **53**, 559 (1982).

[10] J.W. Keister et al., submitted to J. Vac. Sci. Technol. **A 17** (1999).

[11] H. Niimi and G. Lucovsky, submitted to J. Vac. Sci. Technol. **A 17** (1999).

[12] G. Lucovsky et al. *The Physics of SiO$_2$ and Si-SiO$_2$ Interfaces - 3*, ed. H.Z. Massoud, et al., (ECS, Pennington, NJ, 1996), p. 441.

[13] G. Lucovsky et al., J. Vac. Sci. Tech. **B 15**, 1075 (1997).

[14] X. Chen and J.M. Gibson, Appl. Phys. Lett. **70**, 1462 (1997).

[15] H. Y. Yang et al., J. Appl. Phys. **83**, 2327 (1998).

Figure 5. J-V characteristics for tunneling in the Fowler-Nordheim regime for (a) substrate and (b) gate injection.

SILICON ON INSULATOR

Recent Advances in Wafer Bonding of Silicon and Alternative Materials

Dr. Scott Blackstone

BCO Technologies, 5 Hannahstown Hill, Belfast, UK
Scott.Blackstone@BCO-Technologies.com

Wafer bonding technology is rapidly gaining acceptance as the technology of choice for an ever widening range of applications. In this paper I review the fundamental bonding mechanisms along with new applications. These include, pattern buried layers, high thermal conductivity buried insulators such as diamond, hetero-bonding of silicides and compound semiconductors.

1. Introduction

Wafer bonding is rapidly becoming an accepted technology for the manufacture of silicon based products. Already a large number of thick film SOI products, which incorporate wafer bonding, are commercially available (3,4,5,11,12). Many predict that in the future, thin film SOI/CMOS will also become a significant technology.

Wafer bonding has migrated to include silicon compatible materials, such as glass, quartz, diamond, alumina and sapphire. It has also been extended to non-silicon materials such as GaAs, InP and SiC among others.

Wafer bonding technology, as understood in its present form, was first proposed by RCA Laboratories in a series of patents filed in the 1960's. However, 1960's standards of surface roughness and particles prevented commercial production of bonded wafers. This technology was re-evaluated in the 1980's, principally in Japan, with successful results. Commercial production of bonded wafer pressure sensors began in the mid 80's and bonded wafer electronic devices in the late 80's (12).

2. Bonding Mechanisms

2.1 Joining

Wafer bonding works by a two step process, joining and bonding. In the joining operation, two atomically flat surfaces are brought into intimate contact. At some points, the space between the two surfaces will be sufficiently small to achieve VanDerWaal bonding. This point bonding will pull the two surfaces together allowing the surrounding areas to also VanDerWaal bond, creating a zipper like effect. This propagates a joining wave, which sweeps across the surface leaving the two surfaces tightly coupled. This is shown in SAM (scanning acoustic microscope) images of Fig 1. The image on the top left is only joined in the bottom, in the top right image, the joining wave has moved half way across the wafer and in bottom image on the right the joining wave is complete.
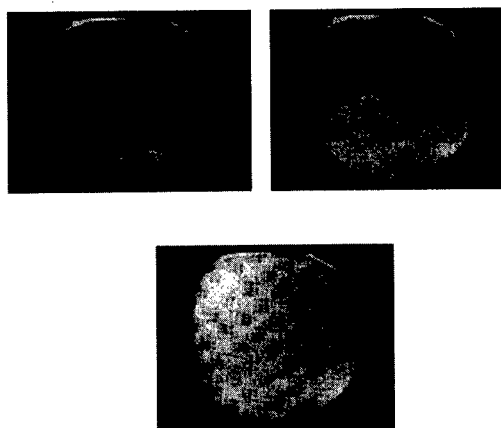
Fig 1 Progression of Joining Wave Across Wafer

This wave releases a great deal of energy, which is sufficient to "snowplow" particles, adsorbed gases

and other foreign material off the wafer edge. An example of the strength of this wave can be seen in the joining of thin wafers, about half normal thickness. In this case the wave can be clearly seen on the back of the thin wafer as a distortion moving across the surface.

This joining wave is the reason why wafer bonding works. Completely particle free surfaces are beyond current technology and if no particles were allowed for good bonding, then it could not be achieved. The joining wave is sufficiently strong such that it can push off relatively large numbers of particles, 100's, and relatively large particles, approx. 0.5um. These are well within current cleaning technology.

If there are large particles on the surface, or an accumulation of smaller particles, the joining wave will stop and sweep around the particle, hermetically sealing it into the joined interface. This is known as a void. While voids do not exist on commercially produced bonded wafers, it is interesting to understand them for reliability purposes. Silicon is very stiff material and will not conform to the particle entrapped at the interface, but will "tent-over" the particle creating a void, which is three or four orders of magnitude greater than the particle size. In other words, a 1um particle creates a void, which is about 1mm in diameter.

2.2　　Bonding

After the two surfaces are joined, they are heated to allow covalent bonding of the two surface materials resulting in a permanent, void free interface. As the two surfaces are heated and chemically bond, the water molecules, which were holding the two surfaces together, coalesce into water droplets and migrate away from the interface.

There are basically three types of bonding within the silicon system. These are oxide bonding, in which one or both surfaces have a thermally grown oxide, hydrophilic, in which both wafers have bare hydrophilic surfaces and hydrophobic in which both wafers have bare hydrophobic surfaces.

2.2.1　　Oxide Bonding

If one of the surfaces is an oxide, the water molecules diffuse into the oxide allowing the silicon and oxide surfaces to bond. In this case, a voidless

wafer, as joined, will remain voidless during and after bond.

2.2.2　　Hydrophilic Bonding

If the surfaces are hydrophilic, the wafers will join with a strong joining wave. However, because the surfaces are not water permeable, the water molecules at the interface must diffuse across the interface and exit at the edge of the wafer.

Even if the wafer is voidless as joined, at about 400C, a dense array of "voids" form across the wafer. These voids are two dimensional water droplets which have agglomerated and which then migrate across the wafer exiting from the edge. These voids clear first in the middle and then the voidfree area spreads toward the edge of the wafer. The temperature at which the wafers become void free is a strong function of wafer surface roughness.

After the wafer expels the interfacial water and again become voidless, the interface is composed of a thin oxide, which was the native oxide on the wafers prior to joining. At higher temperatures, this oxide also agglomerates, forming into islands of oxide. Outside these islands, the two silicon layers are directly bonded together.

An example of this is shown in the following photoluminescense plan view image of the bonded interface (fig 2) courtesy of BioRad, UK.
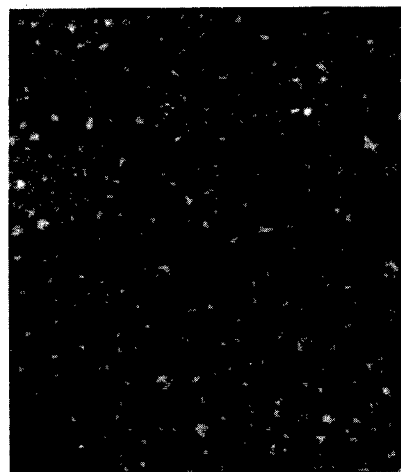
Fig 2 Photoluminescense Plan View of Hydrophilically Bonded Wafer.

In this image the thick oxide islands are shown as white in a dark background. After a sufficient anneal the entire interface becomes oxide free.

2.2.3   Hydrophobic Bonding

If the surfaces are hydrophobic, the wave, if it exists, will be weaker. The existence of a joining wave, in the absence of water vapor at the surface, is a matter of debate. Like the hydrophilic case, heating the wafer results in water agglomeration and "voids". These voids clear at a lower temperature than the hydrophilic case and higher temperature does not improve the bond as in the case of the hydrophilic case. This is probably because of the lack of a native oxide.

3.   Pre-Processed Wafers

There are a number of applications in which it is desired to process the device wafer prior to bonding. The simplest example of this is to implant the wafer and thus create a buried layer. However this can also include deposited films, cavities and other processing.

The bonding of doped layers requires special care as many standard cleaning solutions interact with the diffused layer resulting in voids. This is function of specie and dose.

Patterned buried layers present additional challenges of surface topology. The implantation doping step results in a volume expansion and surface step. With careful attention to the surface, it is possible to routinely achieve successful bonding on even extensively pre-processed wafers.

Fig 3 is a plan view SEM of a pre-processed wafer which has been polished back almost to the buried oxide showing a void free bond. This wafer had a deep N- and P- diffusion, followed by a deep N+ and P+ diffusion followed by a shallow N+ and P+ implant. This wafer was processed through 7 photo steps, and two high temperature diffusions. In the photo, the dark lines are from a trench, the clear
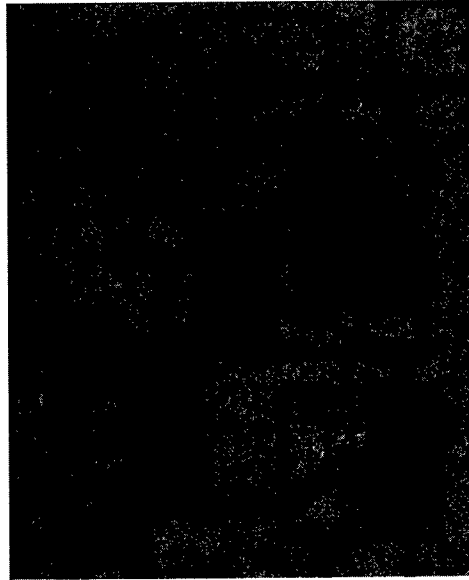
Fig 3 Preprocessed Wafer Thinned to Sub-Micron Thickness Showing Voidless Bonding and Doping Profiles.

areas are the N areas, the light gray areas are the shallow P+ and the dark areas are the deep P+ areas. The polishing process decorates these P-type areas.

4. Buried Thin Film Materials

There are a number of potential improvements on the SOI structure. These include higher conductivity, both electrical and thermal, as well as improved optical properties.

4.1   Buried Silicides
Forming a silicide layer between the top silicon layer and the buried oxide can have a number of interesting applications(1-7). This layer can provide a low electrical resistance path for vertical current flow, high optical reflectivity and a "channel" for doping.

A TEM cross-section of a bonded Silicon-Silicide-SOI (S^2OI), courtesy of DERA, Malvern, UK is shown in Fig 4.
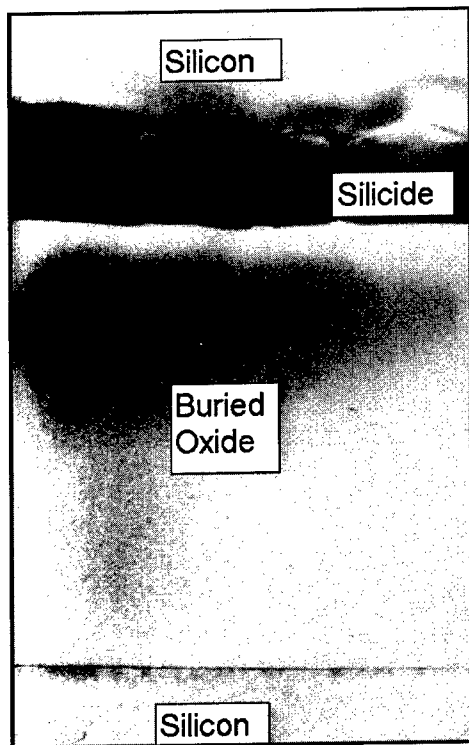
Fig 4   Cross-Sectional TEM of Bonded Silicon-Silicide SOI Structure.

The silicide in this case is tungsten silicide, WSi. WSi silicide was chosen because it is refractory, stable and can be deposited from commercially available CVD equipment. The silicide was first deposited onto the device wafer and then the silicide layer was wafer bonded to the oxidized handle wafer.

4.2   Improved Dielectric

One of the limitations of the SOI structure is the very limited thermal conductivity of the buried thermal silicon dioxide layer. Alternative materials such as diamond, sapphire (Al_2O_3) and aluminum nitride (AlN) have been investigated. Diamond offers a 1000 times improvement over silicon dioxide in terms of thermal conductivity. AlN is 100 and sapphire is 20 times better.

Diamond appears to offer the most promising alternative to silicon dioxide and CVD techniques for depositing diamond are now becoming available. However, the surfaces are very rough, in the order of 10nm, and thus will not bond. To overcome this

problem, a layer of polysilicon is deposited on to the diamond and then polished producing a bondable interface.
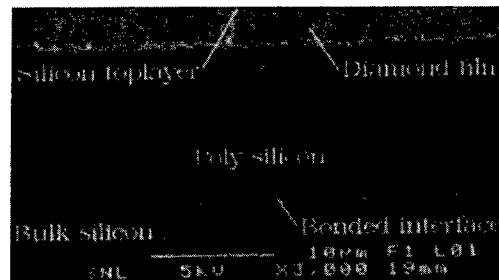
Fig 5 Cross Sectional SEM of a SOD, Silicon On Diamond Bonded Wafer.

This technique is demonstrated in fig 5, courtesy of Chamlers University, Sweden. The rough diamond surface is smoothed by the polysilicon, which is then polished and bonded to the handle wafer.

Complete MOS circuits have been fabricated on this SOD (Silicon On Diamond) with excellent results. The higher thermal conductivity of the SOD device was clearly demonstrated in the device characteristics.

5.   Non-Silicon Semiconductors

In addition to the bonding of non-silicon based materials such as tungsten silicide and diamond, other semiconductor are also bondable. These are most notably the III-V and II-VI materials such as GaAs, InP and their compounds.

Currently an area of great interest is not in bonding compound semiconductors to each other but to silicon. This allows the unique features of the compound semiconductors to be combined with the economy and scale of silicon technology.
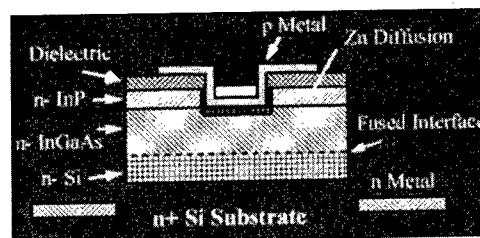
Fig 6 Schematic of PIN Detector.
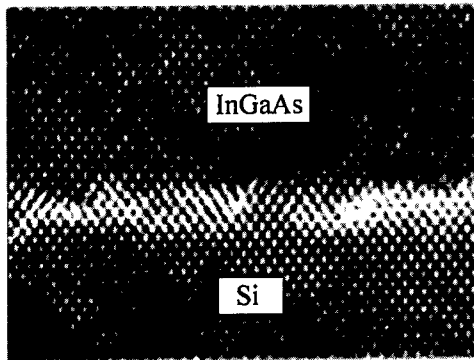
Fig 7　Cross Sectional TEM of PIN Detector
Bonded Interface

A simple example of this is the PIN detector shown
in Fig 6, courtesy of Ciena Corp.

The bonding interface is between the InGaAs layer
and the silicon substrate. As seen in the cross-
sectional TEM, Fig 7, this is a voidless bond without
an interfacial layer.

6.　Epoxy Bonding

While room temperature gluing of wafers is not
considered a form of bonding, I mention it because
wafer bonding has resulted in an interest in "layer
transfer" which is assisted by epoxy bonding. It has
only been recently possible to form a thin, voidless,
continuous adhesive layer between the finished SOI
wafer and a substrate such as glass.
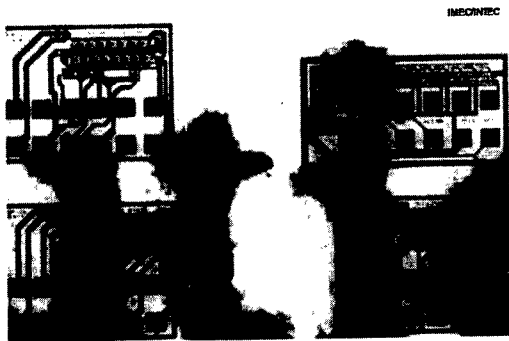
Fig 8　SOI Layer Transferred Onto Glass Wafer and
Laid on Newsprint.

After a circuit is build on an SOI wafer, it may be
desirable to "detach" it from the handle wafer, using

the buried oxide as an etch stop. In order to
accomplish this, the front side of the wafer must
first be glued onto a suitable substrate.

An example of an SOI circuit which was glued onto
a glass substrate and then had the handle wafer
removed is shown in Fig 8. This photo, courtesy of
IMEC, Belgium, shows the circuit sitting on a
printed page. This technique shows great promise
for optical devices, microwave devices and high
density packaging.

7.　MST/MEMS

No discussion of bonded wafer thick film SOI
would be complete with a discussion of micro-
machining. Single crystal silicon is an excellent
mechanical material and the buried oxide provides
controlled undercutting of beams, etc. Bonded
wafer MST devices are used in applications ranging
from high performance gyroscopes for space
applications to airbag sensors in automotive
products.
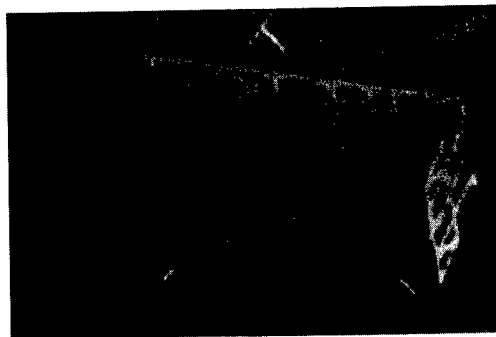
Fig 9　Self Assembled 3-D Device

An advanced version of MST design is shown in Fig
9, courtesy of Imperial College London, UK. In this
structure, 2-D plates are formed in the thick film
bonded SOI and then lifted vertically using specially
designed hinges. It is projected that complete
micro-machines will be constructed using these
techniques.

8.　Conclusion

In conclusion, wafer bonding technology has been
applied to an ever widening variety of applications
and materials that go far beyond the original ideas of
the inventors. It is rapidly becoming just one more

process in the "tool box", similar to oxidation, diffusion and implantation. This widespread acceptance of the technology will mean that it will continue to grow and evolve.

References

1. R. Wilson, C. Quinn, B. McDonnell, S. Blackstone and K. Yallup, "Bonded and Trenched SOI with Buried Silicide Layers" in Electrochem. Soc Proc. Vol 95-7, p535
2. K. Yallup, R. Wilson, C. Quinn, B. McDonnell and S. Blackstone, "Buried WSi SOI Structure", 1995 IEEE Int'l SOI Conf. P 137
3. V. Nayar et al, "Bonded Silicon Silicide on Insulator (S2OI): A Study of Chemical, Physical and Optical Properties for Advanced Device Fabrication", In 4th Semiconductor Bonding Symp. 192 Electrochem, Soc Conf, Paris, Sept 97
4. S. Lizotte, W. Cantarini, C. Quinn, A. Nevin and S. Blackstone "Silicided Buried Layer Smart Power Process" 4th Semiconductor Bonding Symposium, 192 Electrochem, Soc Conf. Paris Sept 97
5. S.B. Goody, P.H. Saul, "A fully Oxide Isolated Bipolar Transistor Integrated Circuit Process", Electrochem Soc Vol 95-1, May 95
6. W.L. Goh, J.H. Montgomery, S.H. Raza, B.M. Armstrong and H.S. Gamble "Electrical Characterization of Dielectrically Isolated Silicon Substrates Containing Buried Metallic Layers", IEEE Electron Devices
7. W.L. Goh, D.L. Campbell, B.M. Armstrong and H.S. Gamble, "Buried Metallic Layers with Silicon Direct Bonding", Electrochem Soc 95-7,'95
8. K. Hermansson and S. Blackstone "Bonded Wafers for Use in MEMS Manufacturing", Micro Structure Workshop '98
9. M. DeCourcy, J.C Alderman, S. Blackstone, H. Gamble, "A Comparison of Bipolar Devices on Bonded and Trenched Wafers with Junction and Trench Isolated Bulk Wafers", Electrochem Soc 3rd Semiconductor Bonding Symposium '95
10. S. Blackstone "Mechanical Thinning for SOI" 3rd Semiconductor Bonding Symposium '95
11. K. McStay, F. Smoot, D. Hariton, K Egan, D. Irvine, G. Brown. P. Edwards, S. Blackstone and H. Gamble, "A Commodity Analog Process Based on Bonded Wafers", Electrochem Soc 3rd Semiconductor Bonding Symposium.
12. Electronic Design, July 13, 1989 "Direct Wafer Bonding Builds Rugged Rectifiers"

Physics and characterization of transient effects in SOI transistors

A.L. Lacaita[a] and L.M. Perron[b]

[a]DEI, Politecnico di Milano, Piazza L. Da Vinci 32 – 20133 Milano (Italy)

[b]STMicroelectronics, Via Olivetti, 2 - 20041 Agrate Brianza (Italy)

The work reviews the physics of transient floating-body effects in SOI devices and gives some examples of their impact on circuit performance.

1. INTRODUCTION

As CMOS technology advances beyond sub-quarter micron, the performance gain achievable from scaling tapers off. Short channel effects and limitations in thinning the gate oxides below 3nm narrow down the improvements for each new generation step [1]. On the other side, the reduction of the parasitic capacitances achievable with the adoption of SOI technology and the improved conductivity of copper interconnections makes it possible to attain a 20-35% performance benefit over bulk CMOS technologies with the same minimum feature size. Pushed up by these figures, in the last years major efforts have been devoted to the demonstration of SOI high performance circuits. Recently, microprocessors have been fabricated on SOI substrates by IBM and Samsung [2-4], 16Mb SOI-DRAM's have been reported [5,6] and operation has been demonstrated below 1.0V [6].

Due to their better threshold control and scalability properties, partially depleted (PD) SOI MOSFET's have attracted more interest than fully depleted (FD) devices. However, as these devices are usually operated with the body floating, peculiar effects arise [7]. The substrate current flows through the source/body diode, leading to a forward bias of the source/body junction and, via the body effect, to a

reduction of the MOSFET threshold voltage, V_T. Sect.2 is devoted to a brief review of these DC effects.

The floating body is also capacitively coupled to the MOSFET terminals. During circuit operation, any voltage transient causes a fluctuation of the body potential, making the threshold of digital gates dependent on the specific signal waveform (rise and fall times, duty cycle, etc...) and on whether the circuit was quiescent for a long time or running for a while. The prediction of the circuit operation becomes therefore more complex. The most important features of the body transients are discussed in Sect.3, while Sect.4 deals with the transients in the basic MOS inverter. The history dependence of the threshold voltage and of the propagation delay together with the increased leakage of the pass-transistor in dynamic circuits are addressed in Sect. 5.

Due to these effects the use of PD-SOI devices in ULSI applications has been questioned for a while and the addition of body ties to the source terminal has been proposed and studied [8-11]. The adoption of body ties is certainly needed for analog applications that require high gain and are very sensitive to disturbances collected by the body and to harmonic distortion. However, in digital circuits the body ties consume extra area and add complexity to layout and fabrication. The alternative to the
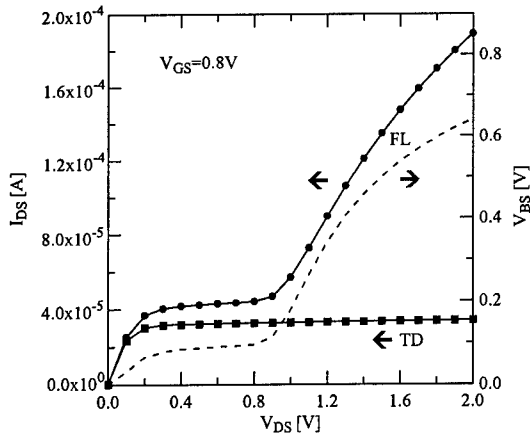
Figure 1. Output characteristics of a PD-SOI MOSFET with L=1μm measured leaving the body floating (FL) and tied to the source (TD). The dashed line, referred to the scale on the right vertical axis, shows the body potential in the floating body configuration.

Figure 2. Subthreshold characteristics of a 0.24μm PD-SOI MOSFET measured in the tied-body (TB) and floating body (FB) configurations [1].

extensive use of body ties is to modify the circuit in order to cope with the floating body effects. Nowadays, advanced SPICE-like models of SOI devices are available [12,13] and circuit operation can be reliably simulated. In addition, floating body effects can be minimized by properly tailoring the device structure. Technological details like the silicon film thickness, halo doping, source/drain extensions and so forth, can help to alleviate parasitic bipolar effects and reduce the impact of threshold and propagation delay uncertainties within the tolerances usually adopted to accommodate process, bias and temperature variations [14]. Today, the use of body contacts is therefore limited only to a few critical circuits, like sense amplifiers.

2. DC FLOATING BODY EFFECTS

Figure 1 shows the typical I_{DS}-V_{DS} curve of an SOI device with V_T =0.6V. When the body is externally contacted and tied to the source (tied body configuration - TD) the transistor behaves like a bulk device. When the body is left floating (FL) the drain current increases, also at low V_{DS}. The effect is due to holes thermally generated at the body-to-drain junction and collected in the body. Due to this excess of carriers, the body potential rises thus lowering the V_T. The dashed line shows the V_{BS} value as directly measured on the sample.

For $V_{DS}>0.9V$ impact ionization at the drain end begins to take place. More holes are generated, V_{BS} increases further, V_T decreases and the drain current swiftly rises. The so-called current kink shows up. Note that in conventional MOSFET's impact ionization becomes evident when the potential drop at the drain end (i.e. V_{DS}-V_{DS}^{sat}) is at least equal to 1.1V, corresponding to an energy drop equal to the silicon band-gap. From the curve in Figure 1, measured at $V_{GS}=0.8V$, $V_{DS}^{sat}=0.2V$. Impact ionization is therefore expected to become evident for $V_{DS}=1.3V$. The kink instead occurs at $V_{DS}=0.9V$. The presence of an internal amplification mechanism makes the floating body devices very sensitive to avalanche multiplication. At $V_{DS}=1.0V$ the substrate current is about 100fA, corresponding to a multiplication coefficient of the transistor current less than 10^{-8}. Even if extremely low, this current is sufficient to cause a 0.1V change of the V_{BS} value, that via the body coefficient leads to a 30% increase in the drain current.

Figure 2 shows the effect of the floating body on the subthreshold characteristics [1]. Again, when the body is tied to source, the device behaves as a bulk MOSFET. As V_{DS} increases from 50mV to 2.5V Drain Induced Barrier Lowering (DIBL) makes the subthreshold characteristic shift to the left. If the body is left floating, the V_T reduction due to the body effect adds up to DIBL and the off-state current increases by almost two orders of magnitude.

In 1997 a study from Intel [15] concluded that, due to the floating body effect, no speed and insignificant power advantage over bulk can be expected for PD-SOI devices as the channel length is scaled below 0.25µm. The authors argued that the PD/MOSFET requires a V_T higher than bulk devices in order to comply with the 1nA/µm I_{off} limit, thus impairing the circuit performance. However a major effect was neglected in that analysis. The temperature coefficient for the V_{BS} is negative, as typically happens for the voltage drop across any forward biased junction [16]. It follows that, even if at room temperature PD-SOI may have I_{off} values higher than bulk devices, at the operating temperature (between 55 and 85°C) the I_{off} values can be much closer. Table 1 shows the predicted performance of a 0.11µm n-MOSFET with a V_T of 0.4V [16]. Increasing the temperature, the ratio between I_{off} in the FB and TD configurations decreases. Moreover, the gap can be further narrowed down well below one order of magnitude by increasing the recombination current at the body/source junction, via defects created by ion implantation [17].

T(°C)	ΔV_T^{FB}(mV)	$I_{off(FB)}/I_{off(TD)}$
27	219	232
55	168	43
85	105	9
100	80	5

Table 1. Dependence on temperature of the additional V_T lowering due to the floating body effect (ΔV_T^{FB}) and of the ratio between the off-state currents in the floating body and in the tied-body configurations. The values refers to a 0.11µm n-MOSFET with a V_T=0.4V [16].

In summary, careful device engineering can reduce the I_{off} penalty of PD-SOI devices. Optimized PD-SOI sub-quarter micron technologies have been already developed, featuring I_{off} values very close to bulk performance [18].

3. FLOATING BODY TRANSIENTS

During circuit operation, the body potential follows, by capacitive coupling, the voltage transients at the device terminals. The most important features of the transients can be described
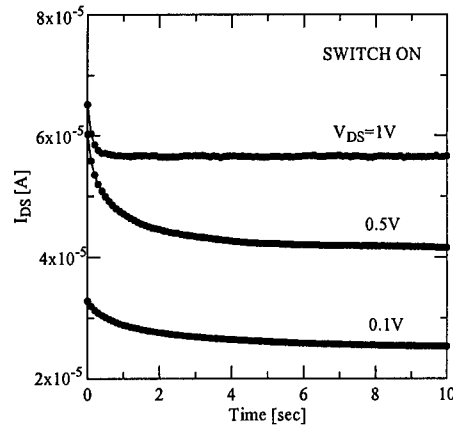
Figure 3. Transients of the drain current measured by switching on the gate from 0V to 0.8V on a PD-SOI n-MOSFET. The drain bias is labeled for each curve [19].
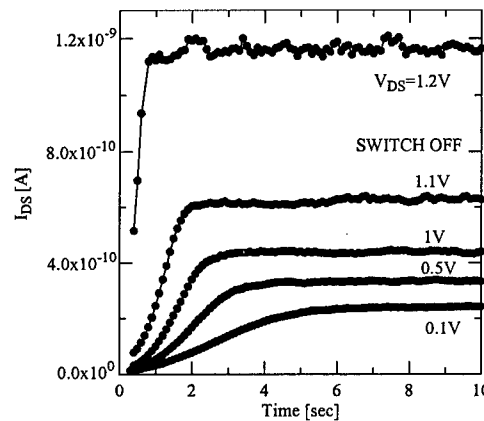
Figure 4. Transients of the drain current in a PD-SOI n-MOSFET when the gate is switched off from 0.8V to 0.2V. The drain bias is labeled for each curve [19].

referring to a simple case: the switching of a PD-SOI MOSFET with the source grounded and the drain held at a constant bias [19,20]. When a turn-on voltage step is applied to the gate, the body potential rises, capacitively following the gate signal, until the inversion channel builds up. The V_T value lowers correspondingly, thus increasing the device current capability.

Figure 3 shows the transient of the drain current induced by switching the gate on from 0V to 0.8V. The drain bias is labeled for each curve. The excess holes, stored in the body soon after the turn-on, is removed in several seconds by recombination

through the source/body diode. As V_{DS} increases, the V_{BS} value reached at steady-state increases and the transient shortens progressively. Figure 4 shows the corresponding switch-off transients. In this case the body potential decreases, following the gate signal. The V_T value increases and the drain current undershoots, well below the steady state off-state value. The lack of holes generated in the body is then recovered by carrier generation. In the measurements in Figure 4 the gate signal was switched from 0.8 to 0.2V and not to ground in order to have a measurable I_{off}. We may therefore conclude that the capacitive coupling between the body and the gate is beneficial, since it dynamically increases the on-state current and suppresses, at the switch-off, the off-state current, pulling down the body potential.

When the device is continuously switched the average V_{BS} slowly approaches a steady-state value set by the balance between generation and recombination processes (Figure 5). When the device is turned on, the gate pulls up the body potential. However carrier recombination during the on state is more effective than generation during the off state and the average body potential, progressively decreases.

The transient changes when V_{DS} is biased into the kink region and impact ionization takes place. The corresponding V_{BS} transient is schematically shown in Figure 6. Following the gate turn-on, V_{BS} rises and more holes are generated at the drain side by impact ionization. Note the rise of the body charge when the device is turned on (dashed-line).
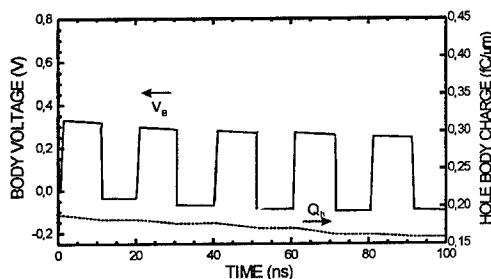
At the switch-off, V_{BS} is pulled down, following the gate waveform. However it does not reach the previous value. An excess of holes progressively builds up in the body until a steady state condition is reached, when the forward current through the source/body diode balances impact ionization.

Even if the curves in Figures 5 and 6 refer to a quite simple case, they highlight some general features of the floating body effects: i) The capacitive coupling to the gate is beneficial. It lowers the V_T value at the switch-on, thus increasing the current delivered to the load. ii) The average body potential depends on the switching history of the device, therefore the V_T value is not precisely defined. This effect causes variations of both the switching threshold and the pulse propagation delay in digital circuits. iii) When impact ionization takes place, hole charge builds up in the body and the steady state V_{BS} reached in switching operation is high. The I_{off} current can be significantly increased with respect to the DC I_{off} quoted in Figure 2 and Table 1.
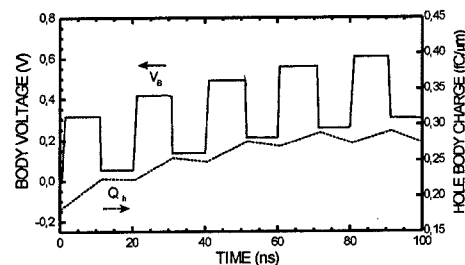
Figure 6. Schematic transients of the hole charge in the body (dashed line) and of the body potential (solid line) in a PD-SOI n-MOSFET under continuous switching operation. The device is biased in a region where impact ionization takes place.

4. THE INVERTER OPERATION

Similar transients take place in the inverter stage, the only difference being the additional presence of the coupling between the body and the voltage swing at the drain [21]. Figure 7 schematically shows the resulting transients of the body potential in the n-MOSFET of a CMOS SOI inverter biased at 1.8V. The highest peaks correspond to the leading edge of the gate waveform. As the gate is switched

Figure 5. Schematic transients of the hole charge in the body (dashed line) and of the body potential (solid line) in a PD-SOI n-MOSFET under continuous switching operation. The device is biased in a region where impact ionization is negligible.
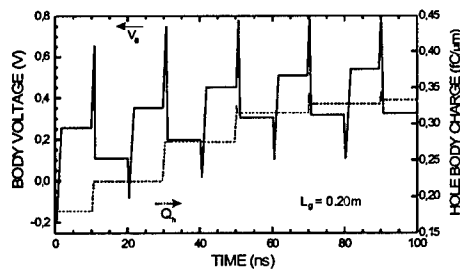
Figure 7. Transients of the hole charge in the body (dashed line) and of the body potential (solid line) in the n-MOSFET of a CMOS PD-SOI inverter [21].

on the body potential rises. However, after a few hundreds of picoseconds the drain potential falls and V_{BS} is pulled down. In Figure 7, the supply voltage is supposed to be high enough for impact ionization to take place. Therefore while the drain voltage is high the hole charge stored in the body steeply increases (dashed curve in Figure 7).

Also in this case the steady-state body potential depends on the details of the voltage waveform (i.e. rise and fall times, duty cycle), however the results in Figure 7 highlight the effects of the coupling between the body and the drain. If the capacitive coupling to the drain is strong (as in Figure 7) the V_{BS} swings driven by the drain transients exceed those due to the gate waveform. It follows that, when the device is off, the V_{BS} value is pulled up by the high drain bias. The I_{off} value is therefore increased, and V_T decreases. The opposite happens if the gate coupling is dominant and the V_{BS} transients approach the waveform in Figure 5.

The ratio between the gate/body and the gate/source capacitance mainly depends on the channel length and the silicon film thickness. By increasing the latter, the V_{BS} swings due to the gate signal decrease and the V_{BS} transient becomes similar to the one in Figure 6. Figure 8, taken from Ref.[22], shows the V_T dependence in a 0.2μm PD-SOI device. By changing the silicon thickness, the channel doping and the DC V_T value were chosen to have a DC I_{off} of 1nA/μm at V_{DS}=1.8V. For a thickness larger than 50nm, the device is partially depleted. The solid line shows the V_T value of the devices, when they are held off for a long time. In this case V_{BS} is 0.2-0.3V, since the body is charged by avalanche multiplication of the weak off-state

current. The dashed line is instead the V_T when V_{DS}=0 for a long time and the body reaches the equilibrium with the source side. The dotted line shows the V_T value reached at steady state when the inverter is continuously switched by a square wave. As expected the steady state V_T decreases with increasing silicon thickness. As a general rule, in order to minimize threshold variations and the related circuit problems, the device should be designed in order to keep as close as possible the three V_T values.

As the film thickness is reduced below 50nm the SOI MOSFET becomes fully depleted. Note that also FD-SOI suffers from carrier storage in the body and corresponding floating body effects. These effects disappear only in FD-SOI with very narrow silicon films [22].

It is evident how the floating body effects make complex the optimization of an SOI technology, adding further interactions between the different aspects of the device structure. For example the increase of the halo doping does not only affect the short channel performance, but also the capacitive coupling of the body to the source and the drain, as well as the IV characteristic of the source/body junction, with a direct impact on the V_T variations.
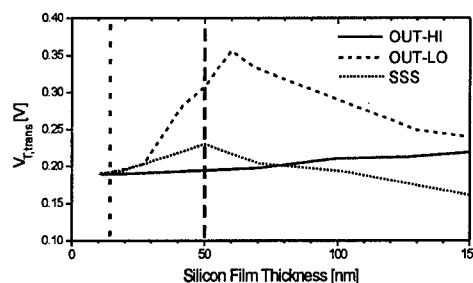
Figure 8. The threshold voltage dependence on the silicon film thickness in a 0.2μm PD-SOI device. The solid line represents the V_T values when the device is held off for a long time with the V_{DS} at 1.8V, the dashed line is instead the V_T reached when V_{DS}=0 for a long time. The dotted line shows the V_T value at steady state when the inverter is continuously switched by a square wave [22].
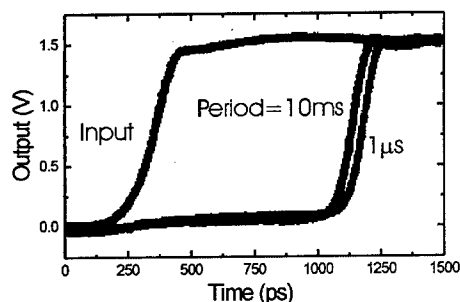
Figure 9. Leading edge of a voltage pulse as measured at the input and at the output of an inverter chain. The period values refer to the time interval between two subsequent pulses [14].

5. IMPACT ON CIRCUITS

In digital circuit, the history dependence of the threshold voltage, may lead to racing, instabilities and logic errors. Let us consider, for example, the impact on the operation of a simple inverter chain [23,24]. If the input to the inverter chain has been idle for a long time at the low level, the first stage nMOS (input low, output high) has a V_{BS} value of 0.2-0.3V, and therefore low V_T. The same happens for all the n-MOS devices of the odd-stages and in a complementary fashion for the p-MOS of the second stage and for the p-MOS of all the even-stages. On the contrary, the devices in the on state (n-MOS of the even-stages and p-MOS of the odd-stages) have $V_{DS}=0$ and therefore $V_{BS}=0V$. In summary the devices in the off state have lower threshold values. When the input pulse rises, the rising edge of the input pulse propagates down the chain faster. Figure 9 shows experimental evidence of such an effect that may be also summarized by saying that idle circuits switch faster [14]. If the period between two subsequent pulses is long enough (10ms) the body potential of the off-state devices reaches the idle steady state conditions. As the period decreases, the difference between the body potential of the on and off devices is reduced and the leading edge propagates slowly.

The opposite happens to the trailing edge of the pulse. If the inverter chain is idle for a long time, the devices contributing to the propagation of the falling edge are those with the higher threshold.
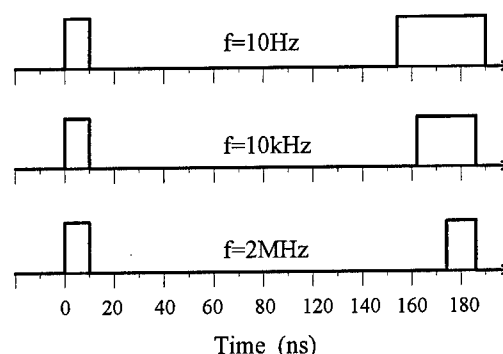
Figure 10. Stretching of a pulse propagating through an inverter chain. The input pulse width is 10ns [23], the output width depends on pulse frequency.

It follows that the falling propagates slower than the leading edge and the pulse stretches (Figure 10). Also this pulse effect depends on the input frequency. As the input frequency increases, the pulse stretching decreases, as there is no time for devices to recover the initial quiescent condition.

However, in optimized PD-SOI the history dependence of the propagation delay is minimal. The switching delay measured on open chain structures shows less than 5% variation in inverter, NAND and NOR circuits [14] and is not the main source of delay uncertainty in complex circuits like microprocessors. Table 2 shows that across chip line width variation and changes in supply voltage and temperature can cause a delay variation of 15-20%, larger than that due to floating body [14].

Major causes of delay uncertainty	% Delay variation
Floating-body effects	8%
Line width variations	15-20%
10% change in V_{dd}	10-20%
Temperature (25-85°C)	15-20%

Table 2. Sources of delay uncertainties in digital logic circuits [14].

Delay variations due to the floating body can be therefore accommodated within the same design margins already adopted to account for other sources of delay uncertainties. This does not mean that a bulk circuit design can be directly transferred to SOI. There are circuits, like the sense amplifiers, where the floating body effects can cause logic errors and, in some cases, body ties must be adopted to guarantee the proper circuit operation.

The transient activation of the parasitic bipolar transistor is another peculiar effect that has to be managed in PD-SOI circuits. Figure 11 schematically shows the PD-SOI pass-transistor of a DRAM cell. The source, corresponding to the DRAM bit line, is pulsed from V_{DD} to ground while the gate (word line) is off. The V_{BS} value increases, the source/body junction is turned on and the parasitic bipolar may be activated. Two main current components contribute therefore to the transient drain-to-source leakage: the off-state channel current, increased by the V_T lowering, and the parasitic bipolar current [23]. Since the body is open, the parasitic is switched on only for a few ns (Figure 12) until the body is depleted by holes and its potential returns less than one diode cut-in voltage below the source potential. This effect causes therefore extra current leakage that may degrade the data retention of DRAM's and the operation of dynamic logic circuits. The p-MOS devices suffer in principle by the same effect but the β of the p-n-p is lower and the leakage current is smaller. If the bit-line is repeatedly pulled down, the body is progressively discharged and V_{BS} follows a transient similar the one in Figure 5. In this case the most significant leakage takes place during the first cycles. As the body bias decreases the bipolar leakage is quenched off.

Figure 12. Pass-gate leakage in a PD-SOI n-MOSFET due to parasitic bipolar current [14].

To induce significant transient leakage current from this point on, the bit line would have to stay high for a time long enough (10ms) for substantial body charging to recur. The worst case for dynamic retention would be therefore a sequence of bit line pulses with a long quiescent time in between. The problem can be alleviated through silicon process/device optimization. By properly tailoring the source-drain extension implant and by adopting retrograde n-well, the β current gain of the parasitic bipolar can be reduced considerably and as V_{DD} is scaled down to 1.2V, the effect becomes almost negligible (Figure 13).

For DRAM's Ge-implantation into the source has also been proposed and studied [25]. The resulting band-gap narrowing increases in fact the potential barrier for the electrons and lowers the β current gain of the BJT.

Figure 11. Schematic view of a PD-SOI pass transistor in a DRAM cell, with the corresponding parasitic bipolar.
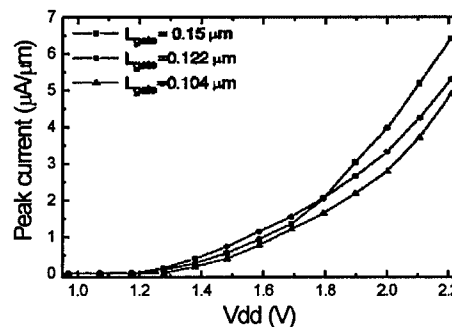
Figure 13. Experimental dependence of the pass-gate leakage on the supply voltage for a PD-SOI n-MOSFET [18].

However, even without this process-step, fully working low-voltage 16Mb SOI-DRAM has been demonstrated and a 50ns access time has been obtained down to 1V power supply [6]. In this design body-tied transistors have been used for the boosted voltage generator, the sense amplifier and word line drivers. The cell transistor has instead the body floating. A threshold voltage of 0.8V is sufficient to suppress the off-state leakage.

6. CONCLUSIONS

The paper presents a discussion of the unique transient floating body effects in PD-SOI devices. Hysteretic V_T variations and their impact on the operation of some simple circuits have been addressed. To fully exploit the potential benefit of this new technology, these floating body effects must be fully understood and the designers must be trained to play with them. Even the choice of where dropping a body contact requires insight of the device behavior, of circuit topology, as well as of their switching pattern dependency. It is however clear that SOI does provide superior performance with respect to bulk. Advanced PD-SOI circuits have been already demonstrated and the SOI application to mainstream high-performance microprocessors is now under way. By taking advantage of the unique SOI device structure, the operating voltage, power and performance will be extended to values unattainable with bulk CMOS technology.

REFERENCES

1. Y. Taur et al., Proceedings of the IEEE, (1997) 486.
2. M. Canada et al. ISSCC Tech. Dig., (1999) 430.
3. D. H. Allen et al., ISSCC Tech. Dig., (1999) 438.
4. Y.W. Kim et al. ISSCC Tech. Dig., (1999) 432.
5. J.W. Park et al. ISSCC Tech. Dig., (1999) 434.
6. T. Eimori, IEEE Trans. Electron Devices, 45 (1998) 1000.
7. S. Krishnan and J.G. Fossum, IEEE Circuits and Devices, (1998) 32.
8. J. Sleight and K. Minstry, IEEE IEDM Tech. Dig., (1997) 419.
9. M. Horiuchi and M. Tamura, IEEE Trans. Electron Devices, 45 (1998) 1077.
10. S. C. Kuehne et al., IEEE Trans. Electron Devices, 45 (1998) 1084.
11. S. Maeda et al., IEEE Trans. Electron Devices, 46 (1999) 151.
12. D. Suh and J. G. Fossum, IEEE Trans. Electron Devices, 42 (1995) 728.
13. G. O. Workman, et al., IEEE Trans Electron Devices, 45 (1998) 125.
14. G.G. Shahidi et al. ISSCC Tech. Dig., (1999) 426.
15. R. Chau et al., IEEE IEDM Tech. Dig., (1997) 591.
16. J. G. Fossum et al., IEEE Electron Device Lett., (1998) 414.
17. T. Ohno et al., IEEE Trans. Electron Devices, 45 (1998) 1071.
18. E. Leobandung et al., IEEE IEDM Tech. Dig., (1998) 403.
19. L.M. Perron et al., Microelectronics Reliability, 38 (1998) 759.
20. D. Munteanu et al., IEEE Trans. Electron Devices, 45 (1998) 1678.
21. J. Gautier et al. IEEE IEDM Tech. Dig., (1997) 407.
22. A. Wei et al., IEEE Trans. Electron Devices, 45 (1998) 430.
23. C. Chuang et al. Proceedings of the IEEE, (1998) 689.
24. P. Lu et al., IEEE J. Solid-State Circuits, 32 (1997) 1241.
25. J. G. Fossum et al., IEEE Trans. Electron Devices, 45 (1998) 1055.
26. Y. Koh et al., IEEE Trans Electron Devices, 45 (1998) 1063.

CMOS/SOI Technologies for Low-Power and Low-Voltage Circuits

J. L. Pelloie, C. Raynaud, O. Faynot, A. Grouillet, J. Du Port de Pontcharra

LETI (CEA/Grenoble), 17 rue des Martyrs, 38054 Grenoble Cedex 9, France
Phone: +33 4 76 88 40 05 Fax: +33 4 76 88 94 57 e-mail: jlpelloie@cea.fr

This paper provides the fundamental basis for adressing current and future CMOS/SOI technologies, which will be mainly dedicated to low-voltage and low-power applications. This review focuses on SOI substrates fabrication, SOI devices (fully-depleted, partially-depleted, DTMOS) and their electrical behavior, and SOI transistor modeling.

1. INTRODUCTION

Power consumption is now a major concern for high-performance digital systems and portable applications. The most efficient technological approach for reducing power consumption is power-supply voltage scaling. Threshold voltage must consequently be reduced to maintain speed but its lowest value is set by the maximum tolerable off-current. Furthermore the off-current also depends on temperature increase during circuit operation and this increase needs to be minimized. For this purpose SOI devices bring their unique inherent advantages over bulk devices: lower junction capacitance, lower junction leakage, no latch-up, lower sensitivity to temperature variation, no substrate reverse-bias effect, and full dielectric isolation. Various SOI device architectures (fully-depleted, partially or non-fully depleted) are suggested and used either for high-speed [1] or low-power applications [2]. Emergence of CMOS/SOI technologies was not possible in the past because of the lack of a reliable SOI substrate in large quantities. SIMOX substrate quality has improved, and new SOI substrates are now available using different wafer bonding techniques. From a technology point of view, SOI is now mature enough to be used in commercial products. The other obstacle was the availability of tools for SOI circuit design: today several SOI MOSFET models (SOISPICE, BSIM3SOI, STAG) are implemented in commercial electrical circuit simulators, so circuit designers only need to get experience. Since SOI devices exhibit specific characteristics such as related floating-body effects, efforts must now focus on device modeling coupled to circuit simulation in order to build SOI library cells.

2. SOI SUBSTRATE FABRICATION

SIMOX is the most common technique for fabricating SOI substrates: a high-energy/high-dose of oxygen is implanted into a bulk silicon substrate, and the buried oxide layer is formed after a high temperature anneal (1350°C). The anneal also recrystallizes the upper silicon layer.

The Smart-Cut process [3] is a recent technique for fabricating SOI substrates. It combines hydrogen implantation, wafer bonding and splitting (figure 1). While a SIMOX buried oxide is obtained by high-dose oxygen implantation and high temperature annealing, the Smart-Cut buried oxide results from thermal oxidation. Silicon thickness is controlled by the hydrogen implant. Due to its good uniformity ultra-thin silicon layers are achieved. In contrast to other wafer bonding techniques, one wafer is saved at the end of the process and may be used again. Smart-Cut is particularly suited for large volume production. From a cost point of view, the Smart-Cut process will compete with low-dose SIMOX processes, and wafer cost is expected to be strongly reduced for large production volume.

Whatever the fabrication technique used for SOI substrates, the requirements for advanced

CMOS/SOI technologies dedicated to low-voltage applications are good silicon thickness uniformity (± 5 nm) across a wafer and from wafer to wafer, and crystalline quality sufficient to grow a high-quality gate oxide. An advantage is the ability to vary the buried oxide thickness for device optimization.
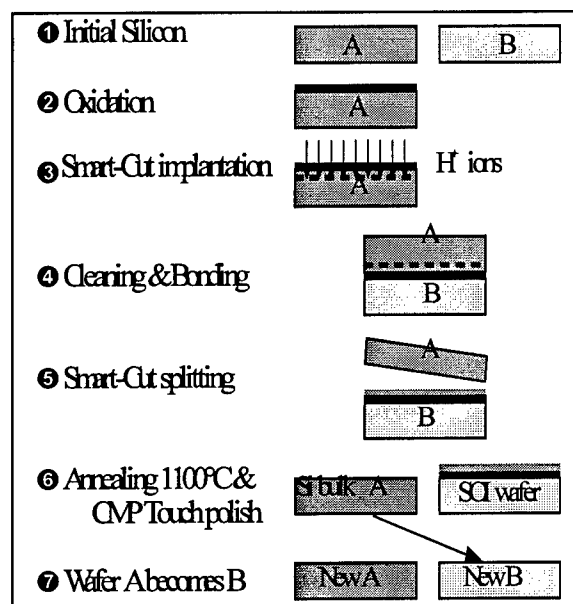
Figure 1. Smart-Cut process description

3. SOI DEVICES

Full or partial depletion of SOI MOSFET's depends on several technological parameters: doping level and thickness of the silicon film, gate oxide and buried oxide thicknesses. Furthermore the depletion condition varies with gate bias and is also related to two-dimensional effects as devices become shorter. Full depletion condition is reached when the total depletion charge, including lateral and back depletion, exceeds the possible depletion charge limited by the silicon thickness (figure 2). For a given device, transition from partial to full depletion may be obtained by varying front and back gate biases but also by varying the drain bias for short-channel devices. SEM cross-sectional views of fabricated fully-depleted and partially-depleted devices are shown in figure 3. Design optimization of 0.18 μm gate length fully-depleted devices

requires an ultra-thin silicon layer (35 nm) while only a thin silicon layer (> 100 nm) is needed for partially-depleted devices. Manufacturability of ultra-thin silicon films is still a question, especially because the electrical characteristics of fully-depleted devices depend on silicon thickness variation. Using SIMOX and UNIBOND wafers, we measured a 10 mV/nm threshold-voltage variation with silicon thickness for a long-channel fully-depleted device on an 8" wafer (0.25 μm CMOS/SOI technology). This demonstrates that good control of threshold voltage may be achieved but will depend on silicon thickness wafer to wafer variation.
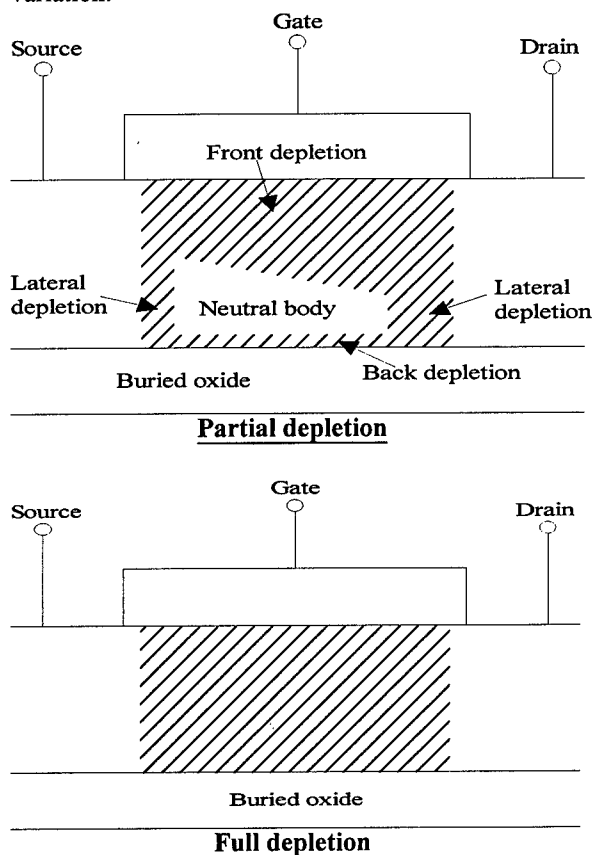
Figure 2. Schematic depletion description.

Due to severe two-dimensional effects, thinner silicon layers (< 35 nm) [4] will be required to achieve optimized sub-0.18 μm fully-depleted devices, silicon thickness control will become more and more critical. Contacting source and drain

regions on such thin silicon films is a challenge. A simple way to overcome it is to use a recessed channel structure (as shown in fig. 3) [5]. As this structure is not self-aligned, a very good lithography matching is required. These considerations indicate that manufacturing an advanced (0.18 μm and below) fully-depleted SOI technology will not be straightforward. Elevating source/drain regions by selective silicon epitaxy or metal deposition could be a potential solution in the future.
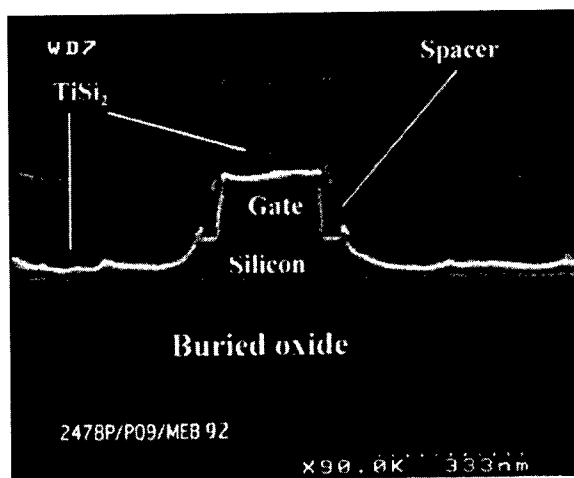
0.25 μm partially-depleted SOI device
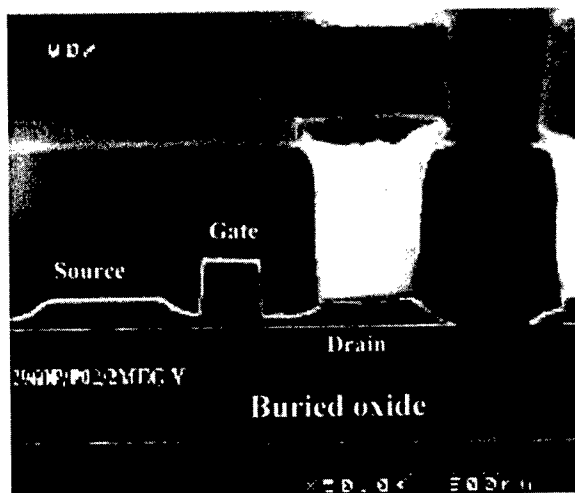
0.25 μm fully-depleted SOI MOSFET

Figure 3. SEM cross-sections of SOI devices

Partially-depleted devices are less sensitive to silicon thickness fluctuations, so threshold voltage control is identical to bulk technologies. A partially-depleted CMOS/SOI technology is quite similar to a bulk technology. Only channel and source-drain implants need to be optimized due to the limited silicon thickness and dopant interactions at the buried oxide interface. Junction depth is automatically limited by the silicon thickness, a good device design is obtained using a retrograde channel profile combined with shallow source-drain extensions and halo implants. Starting from a bulk process, a partially-depleted SOI technology needs minor changes.

A comparison of static $I_d(V_{gs})$ (figure 4) and $I_d(V_{ds})$ (figure 5) curves measured on fully-depleted and partially-depleted devices emphasizes the main differences: a lower subthreshold swing and an attenuated kink effect are observed for fully-depleted devices in DC condition.
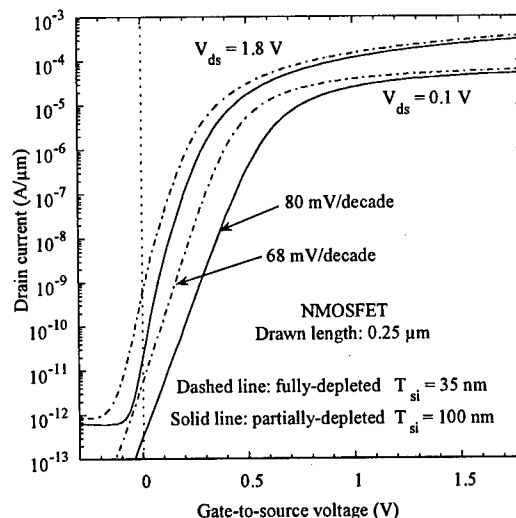
Figure 4. Measured static $I_d(V_{gs})$ curves

Consequently, a lower threshold voltage may be achieved using fully-depleted devices for a given off-current, yielding more margin for low-voltage design. As the body is floating its potential is determined by the different competing physical mechanisms: at high V_{ds} impact ionization effect is counter-balanced by body-source junction self-biasing. A change in the body charge results in a

threshold voltage variation. As majority carriers are in excess in the body, threshold voltage is reduced. This explains the anomalous substhreshold swing at high V_{ds} and the $I_d(V_{ds})$ kink effect. The body charge increase due to generated majority carriers may suppress the full-depletion condition, a kink effect is therefore observed even on fully-depleted devices. Also a fully-depleted device does not remain fully-depleted in the accumulation regime ($V_{gs}<0$ for NMOS). Consequently, defining what is a fully-depleted device is not that obvious. However the full depletion condition leads to a coupling between the front and the back gates, resulting in a variation of the electrical characteristics with the back gate bias. Therefore a device may be considered operating in a fully-depleted mode when its electrical characteristics are influenced by the back gate bias. Again it is important to keep in mind that the full depletion condition may result from the combination of the potential applied at the different nodes.

normalized to the equilibrium value ($t \rightarrow \infty$), is shown for different gate lengths. This variation is explained as follows: when V_{gs} switches from 0 to 1.8V, the surface potential is increased from its initial value to two times the Fermi potential ($2\Phi_f$). Excess majority carriers cannot be instantaneously removed from the body (neutral region) and the body potential (V_b) consequently increases leading to a threshold voltage (V_t) lowering. The initial drain current value is higher than the equilibrium value as V_t is reduced. V_b enhancement involves a forward biasing of body-source and body-drain junctions (depending on drain potential), the excess majority carriers are removed through a thermal recombination process which is a slow process at room temperature. As V_b slowly decreases, V_t increases and the drain current decreases to its equilibrium value. The overshoot duration depends on the gate length. The longer the device, the greater the density of excess majority carriers which need to be eliminated through the recombination process.

Figure 5. Measured static $I_d(V_{ds})$ curves.
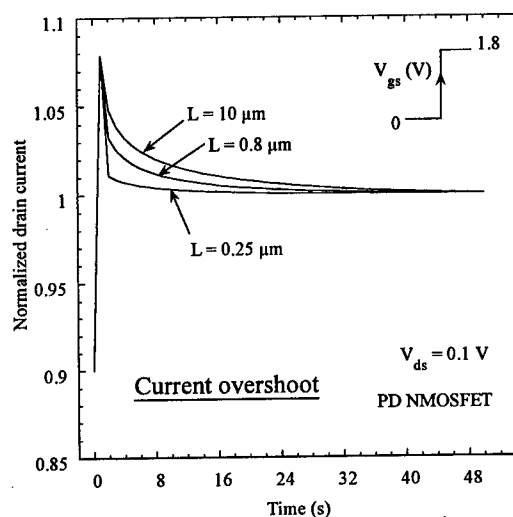
Figure 6. Measured recombination transient.

4. TRANSIENT EFFECTS IN SOI DEVICES

The transient behavior of partially-depleted devices is described in this section. At low V_{ds}, when the gate is switched from off-to-on state, V_{gs} varying from 0 to 1.8V, a drain current overshoot is observed (figure 6). The measured current variation,

This current overshoot is beneficial in circuit operation. V_t is lowered when the transistor is turned from off to on, a behavior which is similar to the Dynamic Threshold MOSFET (DTMOS) operation but is intrinsic to the SOI device. If the gate is switched from V_{gs1} to V_{gs2}, ($V_{gs2}>V_{gs1}>V_t$), no current overshoot is observed since there is no

depletion region variation above threshold. No current overshoot is observed for fully-depleted devices because the body charge remains constant as long as the gate potential is high enough to insure full depletion, a condition determined by the silicon thickness and the doping level.
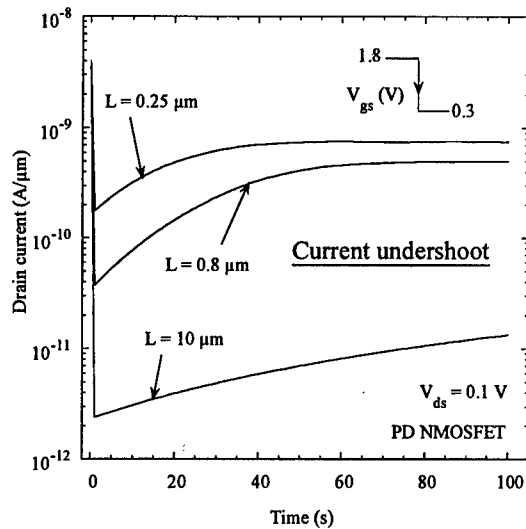
Figure 7. Measured generation transient.

At low V_{ds}, a current undershoot is measured when the transistor is turned from on-to-off state (figure 7). When V_{gs} switches from 1.8 to 0.3V (instead of 0V to allow an easy measurement of the drain current), the surface potential decreases from $2\Phi_f$ to its final value (corresponding to V_{gs}=0.3V). V_b instantaneously decreases to a negative value, body-source and drain junctions are reverse biased, V_t is increased, and the transistor is turned off more efficiently. Majority carriers needed to replenish the depletion region are slowly generated through thermal generation mainly in the space charge region below the gate. The body potential slowly increases and V_t decreases, producing a slow current transient. The current undershoot duration also depends on gate length. For shorter gate lengths the junction contribution becomes significant as the space charge region below the gate is reduced. Again the current undershoot is not observed for fully-depleted devices if the full depletion condition is maintained in the whole V_{gs} operating range. A

simple analytical model describing these slow transient phenomena is given in [6].
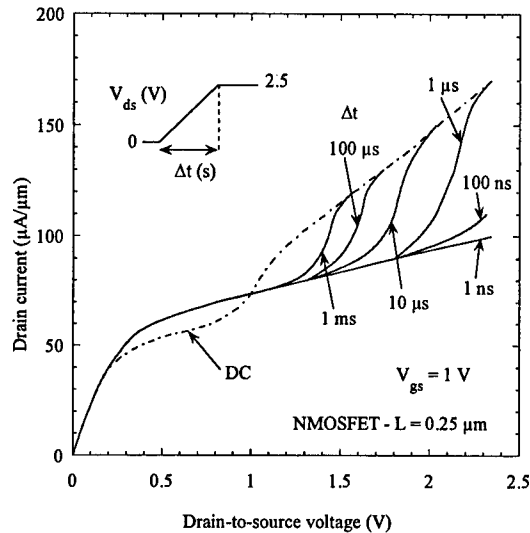
Figure 8. Simulated DC and transient $I_d(V_{ds})$ curves.

At high V_{ds}, partially-depleted transistor behavior is mainly governed by the impact ionization mechanism which is a generation mechanism. Electron-hole pairs are created and holes are stored in the neutral body where the potential takes a minimum value. The body potential is increased due to positive charge accumulation, and the body-source junction is forward biased. An equilibrium body potential is reached when the junction current counter-balances the impact ionization current (neglecting thermal generation in the body-drain junction). If V_b is raised sufficiently, an additional bipolar effect must be taken into account. For modeling, DC characteristics are obtained by equating the generation current (impact ionization) to the recombination current and extracting the body potential. The drain current is dependent on the body potential. The transient characteristics depend on the switching rate: holes may be considered to be instantaneously generated by impact ionization, a positive body charge builds up involving a V_t decrease until the body potential is high enough for recombination to balance the generation mechanism. Our home-made compact analytical model (LETISOI) has been used to simulate transient $I_d(V_{ds})$ curves (figure 8). After the gate voltage is

fixed to 1V, the drain voltage is switched from 0 to 2.5V with different rates, and the duration of the applied ramp varying from 1 ns to 1 ms. The DC curve is shown for reference. For a 1 ns duration, no kink effect is observed, the hole charge build-up in the body due to impact ionization is not sufficient for lowering V_t. The kink appearance is shifted towards the DC case when the ramp duration is increased. As already pointed out by several authors, the kink effect is time dependent and an accurate device model is necessary for predicting the kink effect impact in circuit simulation.

If undesired transient effects are encountered or if a stand-by current reduction is necessary, body contacts may be used in the circuit layout, thus suppressing V_b transient variations. Addition of body-contacts will change the circuit performance, the corresponding specific layout involves a stronger equivalent capacitance due to an area increase of the device. The dynamic performance degradation is shown in figure 9. Off-current is reduced when body-contacts are used (figure 10) thus confirming the suppression of the floating-body effect. However the stand-by current increase due to the floating-body effect is strongly reduced at low supply-voltage. Use of body-contacts is then limited to the management of transient effects.

The intrinsic dynamic V_t lowering effect in partially-depleted devices is an advantage for low supply-voltage applications. The only requirement is an accurate description of this effect in a model.
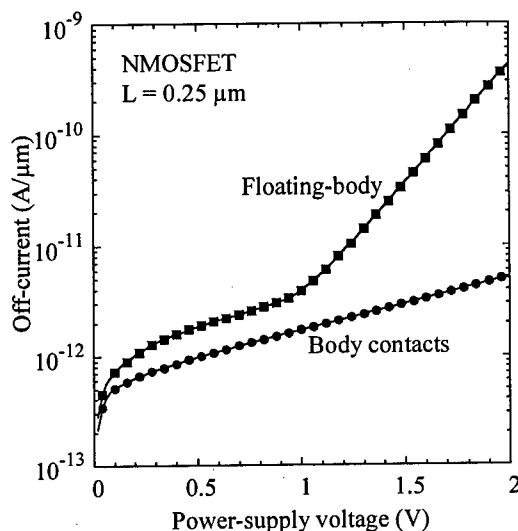
Figure 10. Measured off-current comparison.

5. SOI TRANSISTOR MODELING AND CIRCUIT OPERATION

SOI circuit design is achievable if a circuit simulator is available, this simulator must include an accurate SOI device model. Some of the published models are SOISPICE (University of Florida), BSIM3SOI (University of California), and STAG (University of Southampton). Any model must accurately reproduce the dynamic behavior of the SOI devices which is the most important in circuit operation. Static characteristics are only useful for extracting the model parameters and determining the off-current responsible for the stand-by current. Considering only the static behavior of the devices leads to a misunderstanding in circuit operation. A first example is the kink effect, as shown in the previous section, which is time dependent. Due to the impact ionization mechanism, a very low subthreshold swing ($<$ 60 mV/decade) is deduced from DC $I_d(V_{gs})$ curves at high V_{ds}, but this reduction is not usable, as the DC case does not correspond to circuit operation. Therefore this

Figure 9. Measured dynamic performances.

subthreshold swing reduction is not beneficial. On the contrary, a higher off-current arises from V_t reduction due to forward body self-bias which is generally considered to be a drawback for SOI. This is true at room temperature, but a circuit must operate in an extended temperature range (up to about 100°C). The off-current is minimized at high temperature because SOI offers low junction (or diode) current and the impact-ionization mechanism is more strongly counter-balanced by the thermal recombination mechanism.
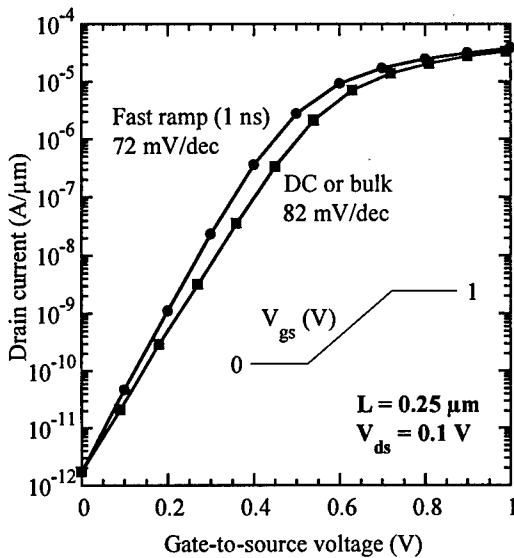
Figure 11. Simulated transient $I_d(V_{gs})$.

The electrical behavior of partially-depleted and bulk devices is generally claimed to be similar, apart from the kink effect. Again this equivalence is only true from a DC point of view. As the body is floating, its potential is related to the potential of the other terminals of the transistor via a capacitive coupling. In dynamic operation, the $I_d(V_{gs})$ curve differs from the DC curve (figure 11), even at low V_{ds}. The body potential increase (figure 12) due to the capacitive coupling, between gate and body in this case, leads to a reduction of the subthreshold swing.

A general SOI MOSFET model must account for all the physical effects occuring in the device and allow operation in regimes seldom considered for bulk MOSFET's. Most importantly the forward biasing of the body must be accurately described

through $V_t(V_{bs})$ variation. A partially-depleted SOI model is quite similar to a bulk model. The main difference is the addition of a fifth node and a floating-body node (figure 13).
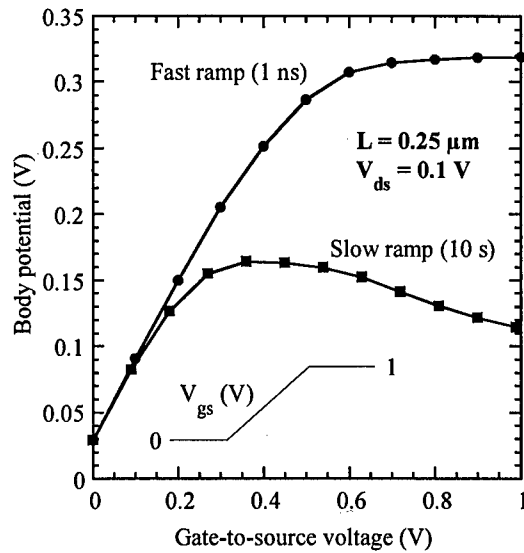
Figure 12. Simulated body potential variation.
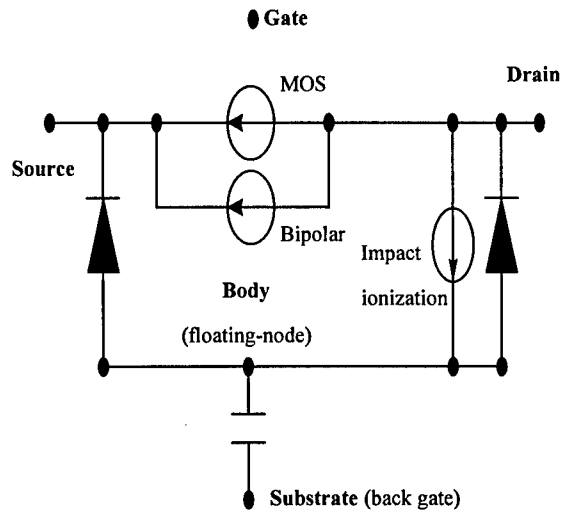
Figure 13. SOI MOSFET equivalent circuit.

Modeling fully-depleted SOI transistors requires describing the transition between full and partial depletion and accurately describing the coupling effect between front and back interfaces when the

silicon film is fully depleted, including two-dimensional effects. The control of the body depletion charge is shared between both interfaces. The model must also predict that the full depletion condition may be suppressed when majority carrier generation takes place, a situation encountered when a kink effect is observed.

Finally, in contrast with bulk devices, SOI devices do not suffer from the substrate reverse-bias effect. The depletion charge is not increased when a source potential is applied (NAND gate for instance where several NMOS are connected in series). This is a significant advantage for low power-supply voltage operation.

6. DTMOS OPERATION

SOI offers a simple implementation of dynamic threshold voltage MOSFET's (DTMOS) [7], where the gate is connected to the body, allowing very low power-supply voltage operation (0.5V) [8]. DTMOS offers an ideal 60 mV/decade subthreshold swing and is therefore a promising candidate for low-voltage applications. In DTMOS operation, V_t lowering is obtained by forward biasing of the body (or substrate for bulk), which has two limitations: the forward junction current contributes to the stand-by current, and the junction capacitance increases with the body potential increase. Although the DTMOS may be fabricated with a bulk process, SOI inherently offers a strong reduction of the junction area and therefore of the junction current and capacitance. Moreover, implementation of bulk DTMOS requires a triple-well process in order to manage independently the substrate of each device. SOI brings its intrinsic dielectric isolation between devices. Addition of a body contact is needed for DTMOS operation, a higher drivability is gained at the expense of a loss in integration density. Due to junction leakage current, use of DTMOS is limited to 0.6-0.7V supply-voltage. However, higher supply-voltage may be achieved by adding a current limiter in the path between the body and the gate [9].

DTMOS provides more drivability than bulk or floating-body partially-depleted devices (figure 14). However due to junction capacitance increase its use should be preferred for large capacitive loads.

7. CONCLUSION

CMOS/SOI technologies offer a large panel of devices: fully-depleted, partially-depleted and DTMOS. These devices may be combined to achieve the best trade-off when designing a circuit, they open the way towards new low-voltage and low-power applications.
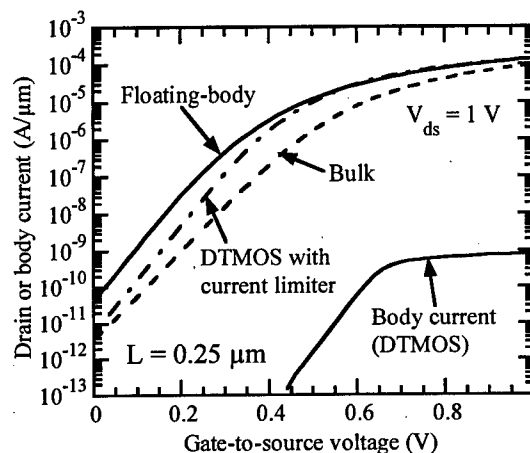
Figure 14. Measured $I_d(V_{gs})$ curves and limited DTMOS body current.

REFERENCES

[1] G. Shahidi et al., Tech. Dig. of ISSCC, San Francisco, p. 426, 1999.

[2] M. Ino et al., Tech. Dig. of ISSCC, San Francisco, p. 86, 1996.

[3] M. Bruel, Elec. Lett., vol. 31, n° 14, p. 1201, 1995.

[4] L. T. Su et al., IEEE Int. SOI Conf. Proc., Palm Springs, p. 112, 1993.

[5] M. Chan et al., IEEE Int. SOI Conf. Proc., Palm Springs, p. 172, 1993.

[6] H. C. Shin et al., IEEE Trans. Elec. Dev., vol. 43, n° 2, p. 318, 1996.

[7] F. Assaderaghi et al., IEEE Trans. Elec. Dev., vol. 44, n°3, p. 414, 1997.

[8] T. Douseki et al., Tech. Dig. of ISSCC, San Francisco, p. 84, 1996.

[9] J. L. Pelloie et al., Proc. ESSDERC, Bordeaux, France, p. 280, 1998.

Photoluminescence spectra of SIMOX buried oxide layers prepared under various conditions

L. Rebohle[a], A. G. Revesz[b], W. Skorupa[a] and H. L. Hughes[c]

[a] Forschungszentrum Rossendorf, POB 510119, D-01314 Dresden, Germany; tel.: +49-351-2603612, fax: +49-351-2603411, e-mail: skorupa@fz-rossendorf.de

[b] Revesz Associates, 7910 Park Overlook Drive, Bethesda, MD 20817, USA; tel: +1-301-469-6675, fax: +1-301-469-7638, e-mail: revesz@boo.net

[c] Naval Research Laboratory, Washington, DC 20375, USA; tel.: +1-202-767-2429, fax: +1-202-404-7194, e-mail: hughes@estd.nrl.navy.mil

The intensity of photoluminescence (PL) peaks at 4.1 eV (300 nm) and 3.1 eV (400 nm) for the buried oxide (BOX) layer of SIMOX structures is less for triple-implant than for single implant samples and is further reduced in single-implant sample by supplemental oxygen implantation. Heat treatment of samples from which the top Si layer was removed decreases the PL intensity in the ultraviolet (UV) region even further. The behavior of the PL is mainly attributed to oxygen deficiency centers in the surface region of a-Si nanoclusters in the BOX layer.

1. INTRODUCTION

The properties of the buried oxide layer of SIMOX structures prepared by implanting oxygen ions into silicon are significantly different from those of thermally grown SiO_2 films [1]. The reason is that the oxide network in BOX layers is similar to that of densified (compressed) with concomitant excess silicon ranging from oxygen vacancies to amorphous silicon (aSi) nanoclusters, as well as even crystalline Si islands in seperate phases. These features depend very strongly on the preparation conditions.

It was found previously that the PL spectrum of the BOX layer of a single-implanted SIMOX sample prepared with 1.7×10^{18} O^+/cm^2 dose is different from that of thermally grown oxide as it exhibits two peaks, one at 4.3 eV (290 nm) and the other at 2.7 eV (460 nm) [2]. It was suggested that the oxygen vacancies are responsible for them. In [3] it was shown, that the UV emission around 280 nm of a BOX layer can be reduced by a supplemental oxygen implant.

Recently, a lot of activity to investigate Si- and Ge-rich SiO_2 layers produced by ion implantation was started. In these studies the PL spectra were also explained with neutral oxygen vacancies or other oxygen deficiency centers [4,5].

The purpose of this work was to study the effects of oxygen implant conditions and post-implant processes on the PL behavior of the BOX layer of SIMOX structures. The effect of heat treatment on pseudo-SIMOX structures (top Si layer removed) is also reported; this point is relevant to the defect structure of BOX layers. An important aspect of this work is that the samples used in this work have been extensively studied by various electrical and other techniques so that the PL spectra could be correlated with the results of those studies.

L. Rebohle et al. / Microelectronic Engineering 48 (1999) 335–338

2. EXPERIMENTAL

We have studied single-implant $(1.5 \times 10^{18}$ $O^+/cm^2)$ and multiple-implant $[(0.5+0.5+0.8)$ $\times 10^{18}$ $O^+/cm^2]$ SIMOX samples annealed at 1320 °C in Ar as well as these samples with supplemental implantation of 0.1×10^{18} O^+/cm^2 followed by 1 hour heat treatment in Ar at 1000 °C. In addition, we have also measured SiO_2 films thermally grown in dry O_2 at 1000 °C in a poly-Si tube. The thickness range of all the oxide layers is 300 to 350 nm. Two pieces of each SIMOX sample were prepared in the following manner: whereas one remained untreated, the other piece was heat treated in O_2 at 1000 °C for 1 hr before the PL measurement.

The top Si layer (when present) was removed from the samples before the PL measurements. These measurements were performed at room temperature in a Spex Fluoromax spectrometer with a R928 Hamamatsu Photomultiplier using an excitation wavelength of 250 nm.

3. RESULTS

The results are shown in Figs. 1 and 2 where the curves labelled "a" refer to samples without the oxygen post-anneal treatment and those labelled "b" refer to samples with such a treatment.

Thermally grown oxide did not show any PL effect. The predominant features of the PL spectra are the peaks around 300 nm (4.1 eV) and 400 nm (3.1 eV). A comparison of curves 1a and 2a shows that both peaks are smaller for the multiple-implant sample than the single-implant one. Curve 3a shows that the supplemental O^+ implantation resulted in a reduction of the height of both peaks. A comparison of curves a and b shows that the heat treatment of the pseudo-SIMOX samples reduced the UV peaks for all the samples.
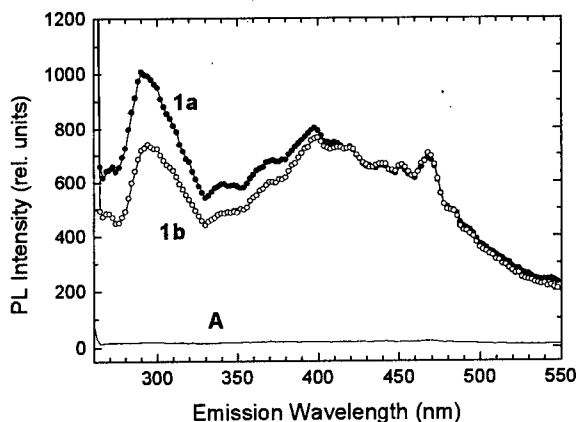
Fig. 1. PL spectrum of a single-implant SIMOX sample without heat treatment at 1000°C (curve 1a) and with heat treatment (curve 1b); also shown is the lack of PL activity of thermally grown oxide (curve A).

Furthermore, the supplemental O^+ implantation causes a shift of the peak positions towards longer emission wavelengths. In contrast to the curves 2a and 2b the UV peak now exhibits a bimodal behavior: In the case of the supplemental O^+ implantation it has a main component at 310 nm and a shoulder around 290 nm.

It was found that the use of additional interference filters in the excitation beam can suppress the blue PL (above 400 nm), whereas this is not possible in the case of UV emission. This implies that the blue PL features are mainly due to scattering in the samples, whereas the UV emission is caused by real PL.

The PL spectra shown in Figs. 1 and 2 can be correlated with various other results, primarily those obtained with photo-injection of electrons into the BOX; the results of these studies are summarized in a review paper [1]. The most pertinent result is that SIMOX-BOX layers usually contain amorphous silicon clusters ranging in size from about 1 to 4 nm. The a-Si clusters are photo-active amphoteric defects that under high electron injection level act as electron traps with a large capture cross sec-
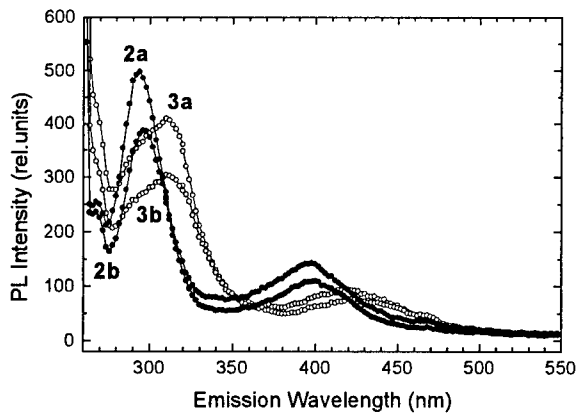
Fig. 2 PL spectra of multiple-implant SIMOX sample (curves 2a and 2b) and multiple-implant sample with supplemental O-implant (curves 3a and 3b; curves a and b: without and with heat treatment at 1000°C, respectively.

tion. At low injection level electrons can be removed from them by photo-depopulation so that they become positively charged. Qualitatively, the order of increasing height of both PL peaks is the same as the extent of positive charging of the BOX induced by photo-injection: multiple-implant with supplemental O-implant, multiple-implant, single-implant samples. However, the quantitative aspects are different, as shown, for example, by the much larger difference between single-implant and multiple-implant samples regarding the density of positive charge (about 1.2×10^{12} |e| cm^{-2} vs. $(0.01-0.05) \times 10^{12}$ |e| cm^{-2}) than the difference in height of either of the 300 nm or 400 nm peaks.

A comparison of curves a and b for all the samples indicates that the heat treatment in oxygen decreased the density of the entities responsible for the PL peaks but did not eliminate them. On the other hand, this heat treatment practically eliminates the positive charging effect, provided the heat treatment was performed without the top Si layer being present, i.e. the oxide layer was un-confined.

4. DISCUSSION

Based on these considerations we suggest that the UV emission is associated with oxygen deficiency centers in the surface region of a-Si clusters.

It seems unlikely, that the UV peaks are due to quantum confinement effects in Si nanoclusters. For very small nanoclusters the bandgap should strongly depend on their size, and according to the size range of the a-Si clusters (1 to 4 nm) a very broad PL distribution is expected.

Previously it was demonstrated, that the strong short wavelength PL of Si$^+$ implanted thermally grown SiO$_2$ layers is not associated with intrinsic radiation-induced defects but is due to excess silicon in the oxide. The neutral oxygen vacancy was assumed to be responsible for the observed PL. This molecular defect can exist in the network in isolated form or in the close surrounding of a-Si clusters. In each case there is no PL dependence on the cluster size [4].

However, there is another point which is not yet fully clarified. It is known that the density of hole traps (which are usually assumed to be the E' centers acting as precursors for the neutral oxygen vacancy) in SIMOX BOX layers is practically independent of the preparation conditions, even including the additional increase of the BOX thickness by supplemental O$^+$ implantation. This behavior is in contrast to the density of photo-active and PL-active defects.

Based on these considerations we believe that the UV emission is caused by oxygen deficiency centers located at the surface or in the close surrounding of the a-Si clusters.

The observation, that the density of a-Si clusters is reduced by a post-anneal oxygen heat treatment at 1000 °C but not completely eliminated, is different from the earlier result which indicated that practically no positive charging

occurs after such a heat treatment. It is possible that the heat treatment changed the cluster/oxide interface of some clusters preventing the photodepoulation of electrons but maintaining their PL-activity.

5. CONCLUSION

This study demonstrates that the PL spectra of BOX layers of SIMOX samples depend strongly on their preparation conditions. The intensity of the PL peaks correlate qualitatively with the density of a-Si nanoclusters as derived from photo-injection studies. The PL behavior reveals an additional complexity in the defect structure of SIMOX BOX layers as the interface between the a-Si nanoclusters may play a role in their electron depopulation behavior. The potential technological implication is that a-Si clusters may be present even in those samples which do not show the positive charging effect.

REFERENCES

1. A. G. Revesz and H. L. Hughes, Microelectronic Eng. **36** (1997) 343.
2. K. S. Seol, A. Ieki, and Y. Ohki, J. Appl. Phys. **79** (1996) 412.
3. H. Nishikawa, J.H. Stathis, and R.E. Stahlbush, to be published.
4. L. Rebohle, I.E. Tyschenko, H. Fröb, R.A. Yankov, G.A. Katschurin, and W. Skorupa, Microelectronic Eng. 36 (1997) 107.
5. N. Nishikawa, Y. Miyake, E. Watanabe, D. Ito, K.S. Seol, Y. Ohki, K. Ishii, Y. Sakurai, and K. Nagasawa, J. non-cryst. solids 222 (1997) 221.

Investigation of SOI MOSFETs with Ultimate Thickness

T. Ernst[a,c], D. Munteanu[a], S. Cristoloveanu[a], T. Ouisse[a], S. Horiguchi[b], Y. Ono[b],
Y. Takahashi[b] and K. Murase[b]

[a]LPCS/ENSERG, 23 Rue des Martyrs, BP 257, 38016 Grenoble Cedex France

[b]NTT Basic Research Laboratories, 3-1, Morinosato Wakamiya, Atsugi-shi, Kanagawa Pref., 243-0198 Japan

[c]STMicroelectronics, 38920 Crolles, France

Ultra-thin SOI MOSFETs with 1-5nm thick SOI film, are experimentally and theoretically investigated. Single- and double-gate configurations are compared; the double-gate MOSFET exhibits a substantial increase in transconductance, presumably resulting from volume inversion. Most of the experimental data can be explained by combining classical models with self-consistent quantum calculations. The characteristics are well-behaved and reveal unique "ultra-thin" film properties: enhanced interface coupling and body-substrate coupling, degraded mobility, increased threshold voltage.

1. INTRODUCTION

Ultra-thin Silicon On Insulator (SOI) MOSFETs are very attractive in terms of high performance and attenuated short-channel effects, being capable of extending the frontiers of the silicon-based microelectronics. "Ultra-thin" is a generic definition for Si films ~ 50 nm thick. This work is focused on much thinner films, in the ultimate range of 1-5 nm. N-channel MOSFETs have been fabricated at NTT (Japan) on low-dose SIMOX substrates with 62 nm thick buried oxide.

The transistor body has been thinned by carefully-controlled sacrificial oxidation. The devices have elevated source and drain (unthinned regions) and natural (residual) body doping. The gate oxide is 50 nm and the channel is long (30 μm) in order to attenuate the parasitic influence of series resistances and device topology. We discuss the room temperature characteristics of single- and double-gate MOSFETs as well as the influence of quantum confinement in ultra-thin Si films.

2. OPERATION OF ULTRA-THIN MOSFETS

2.1 Single-gate MOSFETs

Typical curves are shown, for a 3 nm thick MOSFET, in Figs. 1 and 2. The front-channel

$I_D(V_{G1})$ characteristics are plotted in Fig. 1 for V_D=50mV and various back-gate biases V_{G2}. The lateral shift illustrates the coupling between the front threshold voltage V_{T1} and V_{G2} which is remarkably strong in thin films. The subthreshold swing is between 70 and 100 mV/dec, depending on V_{G2}.
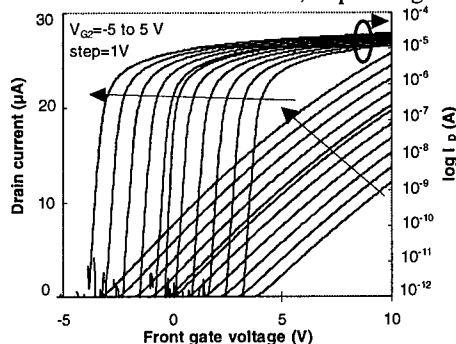
Figure 1. Drain current versus front gate bias in a 3nm thick transistor with L=30μm and W=100μm.

These characteristics are very similar to those currently observed in much thicker fully-depleted MOSFETs [1], except that the interface coupling and substrate-body coupling are reinforced. Rather symmetrical curves are obtained by probing the back channel with V_{G1} as a parameter (Fig. 2). The only difference is due to the onset of substrate depletion which increases the apparent thickness of the buried

oxide and causes a hump in the back channel transconductance (for $V_{G1} \cong 0V$).
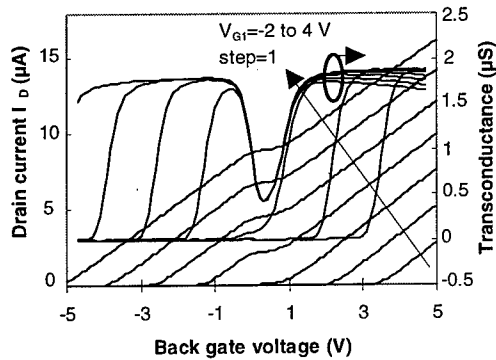
Figure 2. Drain current and transconductance versus back gate bias in a 3nm thick transistor.
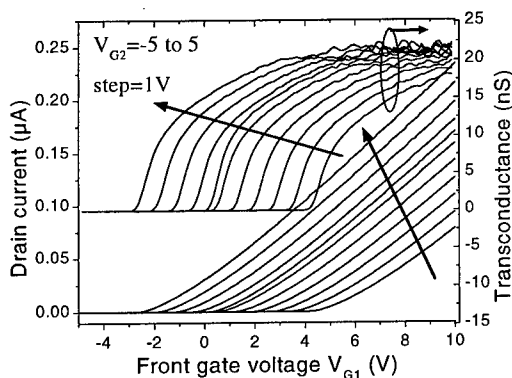
Figure 3. Drain current and transconductance versus front gate bias in a 1nm thick transistor.

Since the $I_D(V_{G1,2})$ characteristics are not "unusual", they serve for parameter extraction with conventional methods [1]. It is found that the front-channel and back-channel field effect mobilities are comparable and reach maximum values when the vertical field is minimized. The electron mobility is reasonable (210 cm^2/Vs) in 3 nm thick MOSFETs, suggesting that the thinning process did not dramatically degrade the crystal quality of the Si film. However, the mobility does increase with thickness: 260 cm^2/Vs in 5 nm-thick and 650 cm^2/Vs in 45 nm-thick MOSFETs.

The characteristics of a 1 nm thick MOSFET are reproduced in Fig. 3. Surprisingly, they are still "MOS-like" and well-behaved. Of course, the drain current and the mobility are very low but this is due to a comprehensive imperfect control of the thick-

ness: the body now contains zero-thick regions and therefore the effective width of the transistor is drastically reduced.
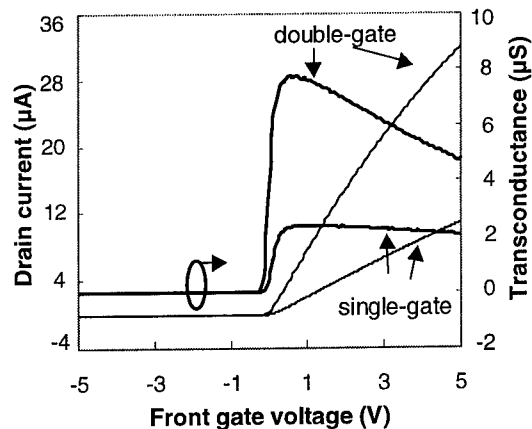
Figure 4. Drain current and transconductance vs. front gate bias of a 3nm thick transistor with L=30μm and W=3μm. In double gate regime, $V_{G2}=V_{G1}\times1.24$.

2.2 Double-gate MOSFETs

These extremely thin MOSFETs are also ideal candidates for operation in volume inversion mode [2]. Double-gate operation has been achieved by simultaneously biasing the front and back gates with $V_{G2}=t_{ox2}/t_{ox1}=1.24\times V_{G1}$. The experimental curves, illustrated in Fig. 4, show a clear increase in transconductance and current as compared to single-gate transistor operation.

3. QUANTIZATION EFFECTS

The analysis of carrier transport mechanisms in ultra-thin SOI MOSFETs proceeds from the comparison between (i) classical analytical models, (ii) advanced classical numerical simulations (Silvaco) and (iii) quantum simulations. The Poisson and Schrödinger equations are solved self-consistently, as described earlier [3,4,5] for the representative case of 3nm thick single and double-gated quantum wells.

3.1 Potential profile and subbands energy splitting

Carrier confinement results from the modification of subband wave functions and energy levels in very narrow potential wells. As the film becomes thinner, the energy levels and their separation increases, making them more difficult to

populate. This explains the sudden increase in threshold voltage observed in sub-10nm SOI films [6]. The difference between single-(SG) and double-gated (DG) wells is illustrated in Fig. 5. In DG-MOSFETs, the in-depth potential distribution is symmetrical, allowing the vertical electric field to cancel in the middle of the film. The quantization effects are essentially thickness-defined, and therefore lesser than in SG-MOSFETs, where the potential well is more "triangular". Two different regions of operation are predicted by quantum calculations in 3nm films:

* At low and moderate field (corresponding to weak/moderate inversion and transconductance peak), the electrons are confined mainly by the infinite rectangular potential well; little additional contribution arises from the potential profile (quasi-flat) in the film. Consequently, both the energy levels and the wave functions of single- and double-gate modes are very close to each others. Moreover, the potential in the film is not high enough to populate the second energy level and only the ground level has to be considered in charge transport.

* At high electric fields (strong inversion), additional confinement is introduced by the triangular shape of the film potential. This results in a higher confinement in SG mode than in DG mode. In double-gate mode, subbands (in particular the second energy level of lateral valleys) begin to be populated.
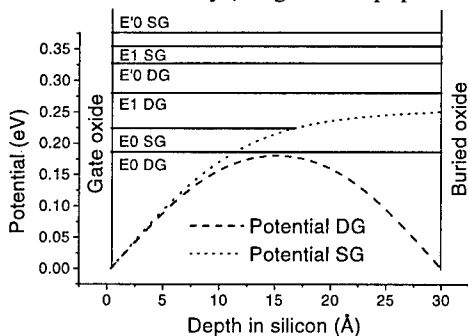
Figure 5. Comparison between SG and DG energy levels and potentials at high gate voltage (2.10^6V/cm surface field).

3.2 Carriers distribution in ultra-thin devices

The spatial distribution of minority carriers is always symmetrical in DG-MOSFETs, with more carriers flowing in the middle of the film. This means that the volume inversion is more pronounced from the quantum viewpoint as compared to the

classical distribution solely defined by the Poisson equation. In SG-MOSFETs, the carrier distribution becomes more and more asymmetrical in strong inversion when V_G increases.
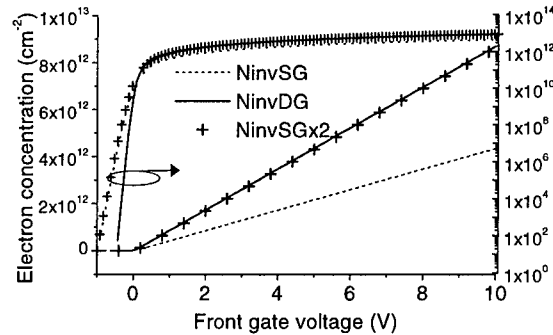
Figure 6. Comparison between double-gate and single-gate inversion charge vs. gate voltage.

The total electron charge in strong inversion is predicted by Gauss' theorem: $Q_{inv}=C_{ox}(V_G-V_T)$ in SG mode, or twice as much in DG mode. The charge control for the two modes are compared in Fig. 6. An interesting point is that in the double-gate regime, the total inversion charge is marginally higher than twice the inversion charge in single-gate regime. This means that, except for the natural 200% gain there is no distinct advantage of volume inversion in terms of total charge. But, the charge distribution, the subband occupancy and the related mobility behavior may bring a supplementary gain.·

3.3 Carrier mobility behavior in ultra-thin films

The impact of silicon layer thinning on carrier mobility was studied for the 20-80 nm range which corresponds to the present thickness for fully-depleted SOI technologies. In this range, the long channel mobility behavior can be explained by universal bulk-Si mobility curves vs. effective electric field [7]. For thinner Si films (<10nm), the electron confinement induces a significant modification in electron wave function and energy which may have a strong effect on the apparent mobility behavior.

Electron-phonon scattering rates increase as $1/t_{si}$ where t_{si} is the silicon film thickness [8]. This scattering mechanism was identified as the most limiting factor of low-field mobility in "ideal" ultra-thin structures. It was shown that the mobility increases in thinner films down to 3.5nm, beyond which a dramatic degradation is observed [9]. We

believe that surface roughness scattering may strongly affect the motion of carriers, essentially near the Si-SiO₂ interface, and should be included in the simulations.

The channel conductivity is given by:

$$g_m = a\frac{\mu_{eff}W}{L}C_{ox}(V_G - V_T)$$

where a=1 for single-gate and 2 for double gate operation. The effective mobility μ_{eff} decreases with increasing the transversal electrical field.
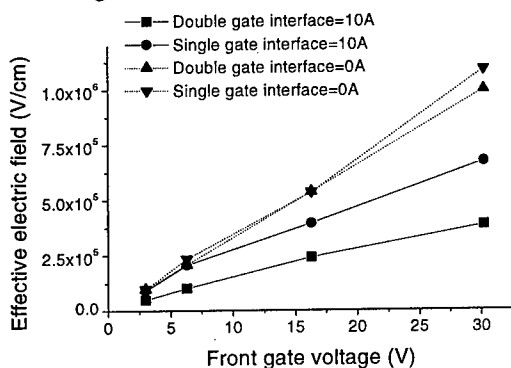
Figure 7. Double-gate and single-gate effective electric field vs. gate voltage for a 3nm film.

The outstanding enhancement of the effective mobility measured in DG-MOSFETs can be qualitatively explained by electrostatic arguments: the electric field is negligible in the middle of the film where most of the inversion charge flows. For thicker films, double-gate regime implies a general diminution of the electric field and consequently a mobility enhancement. But the center of the film (i.e zero electric field) no longer contains the maximum of the total inversion charge.

For 3nm double-gate devices, the zero electric field region corresponds to the maximum probability of electrons presence. As electron scattering via acoustic phonons or surface roughness is strongly dependent on the transverse field, the mobility is most likely enhanced in the center of the film.

A precise mobility law is not easy to formulate because of (i) the effective mass and lattice vibration mode that are not trivial in ultra-thin constrained films, (ii) the contribution of parasitic resistances, and (iii) the thickness fluctuations in such films.

Before such a sophisticated modeling becomes mature enough, we propose a first-order model based on the empirical influence of the vertical electric field. Our approach consists in calculating

the in-depth average mobility, weighed by the transverse field and carrier distribution. Surface roughness effects are arbitrarily assumed to highly degrade the mobility over 1nm near each interface. Integration of the field in the central area of the film leads to a much lower effective field in DG MOSFETs than in SG MOSFETs (Fig. 7) because of the spatial distribution of carriers (Fig. 8). This may offer a basis for understanding the mobility enhancement in volume inversion DG transistors.
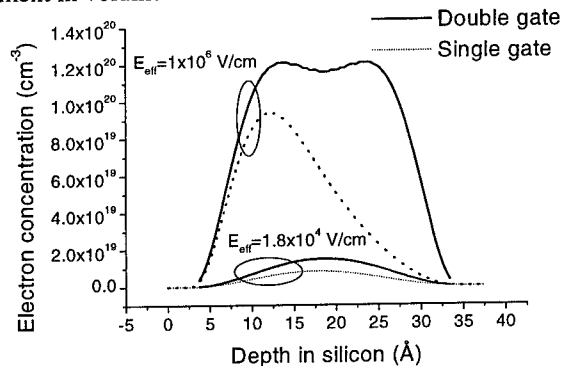
Figure 8. Carrier distributions in DG and SG MOSFETs

4. CONCLUSION

The feasibility and proper operation of ultimately thin transistors has been demonstrated and used to analyze thickness-related mechanisms. This opens new perspectives for the fabrication of advanced SOI devices. Yet, many questions remain open concerning the quantum transport in ultrathin SOI films.

REFERENCES

1. S. Cristoloveanu and S.S. Li, Electrical Characterization of Silicon-On-Insulator Materials and Devices, Kluwer Academic Publishers (1995)
2. F. Balestra, S. Cristoloveanu, M. Benachir, J. Brini and T. Elewa, IEEE EDL., (1987) 410
3. X. Baie, J. P. Colinge, Solid-State Electr., 42 (1998) 499.
4. C. Fiegna and A. Abramo, IEDM (1996) 815.
5. T. Ouisse, J. Appl. Phys., 76 (1994) 5989.
6. Y. Omura, S. Horiguchi, M. Tabe, K. Kishi, IEEE EDL, 14 (1993) 569.
7. J.-H. Choi, IEEE EDL, 16 (1995) 527.
8. B. K. Ridley, J. Phys., C-15 (1982) 5899.
9. M. Shoji, S. Horiguchi, J. Appl. Phys., 82 (1997) 6097.

High-Amplitude and High–Frequency Oscillations of Temperature and Current in SOI Structure

V.N. Dobrovolsky[a], L.V. Ishchuk[a], G.K. Ninidze[a], M. Balucani[b] and A. Ferrari[b]

[a]Department of Radiophysics, Kiev University, 2 Glushkov Av., build. 5, 252127 Kiev, Ukraine,
e-mail: dob@uninet.kiev.ua (Dobrovolsky).

[b]INFM Unit E6 - Department of Electronic Engineering, Rome University "La Sapienza",
Via Eudossiana 18, 00184 Rome, Italy, e-mail: ferrari@die.ing.uniroma1.it

Interconnected oscillations of current, lattice temperature and electron-hole pair concentration were observed in SOI structures upon heating with current at extremely high power. They occur because of the joint action of two competing mechanisms: temperature dependent thermal generation of electron-hole pairs and pair concentration decreasing by current flowing in silicon film through a non-uniform temperature field.

INTRODUCTION

Known current oscillations in bulk semiconductor are not accompanied by essential changes of lattice temperature. In contrast, upon heating of the silicon films of SOI structures by current with power greater than 3.5 GW/cm^3, we have discovered interconnected oscillations, with frequency up to 3 MHz, of current, lattice temperature and concentration of thermally generated pairs. The current and concentration changes were more than an order of magnitude, and temperature change was approximately from 700 K to 1200 K.

1. EXPERIMENTAL RESULTS

The SOI structure is represented in Fig.1. The film of n-silicon has length a and width w of 45 and 10 μm. The thicknesses of the silicon film d_{Si}, buried oxide d_{ox} and substrate d_{sub} are 0.4, 1 and 400 μm respectively. There were two probes in the middle of the film.
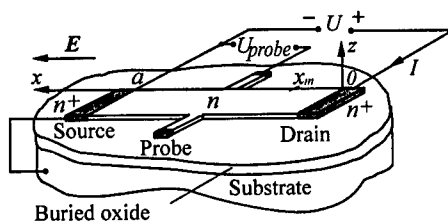
Rectangular pulses of voltage U were applied to the silicon film and oscillograms of current I were recorded. They are shown with solid lines in Fig.2. The values of U in volts are shown near the oscillograms.

Fig.1. SOI structure.
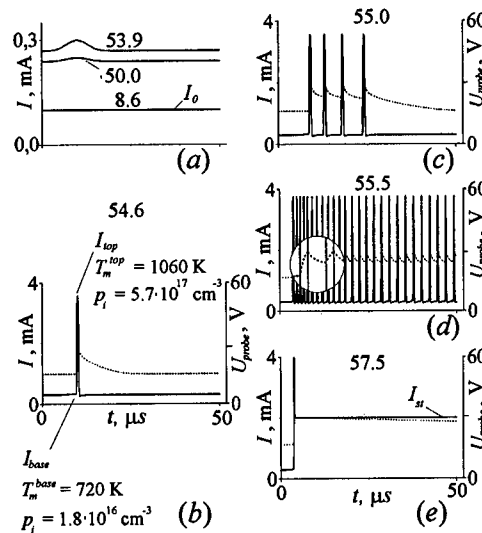
Fig.2. Oscillograms of I and U_{probe}.

Up to a voltage U_i=50 V the current pulses were rectangular with amplitude I_o (Fig2.*a*; U=8.6 V). At higher voltages there was a broad maximum on a pulse (Fig.2,*a*). At a threshold voltage U_{peak}=54,6 V a narrow ($\leq 1 \mu S$) peak appeared at the maximum (2,*b*). The ratio of current values at its top I_{top} and at base I_{base} exceeded an order of magnitude. At U higher U_{osc1}=54.9 V there were several peaks (2,*c*), and then continuos current oscillations occurred (2,*d*). The oscillation frequency observed was up to (1...3) MHz for different structures and was decreasing in the course of time (2,*d*). The oscillations lasted from several seconds up to about 1-2 minutes.

At connection of a film to transistor collector circuit the oscillations were not broken.

The stationary current I_{st} have been settled after peak at U higher U_{osc2}= 57.5 V (Fig.2,*e*).

In Fig.3 the curves for currents $I_o(\square)$, $I_{base}(\circ)$, $I_{top}(\bullet)$ and $I_{st}(\blacksquare)$ vs U are shown. Peak width decreased with increasing voltage and at U>U_{osc2} it was so small that the bandwidth of used oscillograph did not allow us to measure precisely I_{top} values. The appropiate part of $I_{top}(U)$ curve is drawn by dashed line.
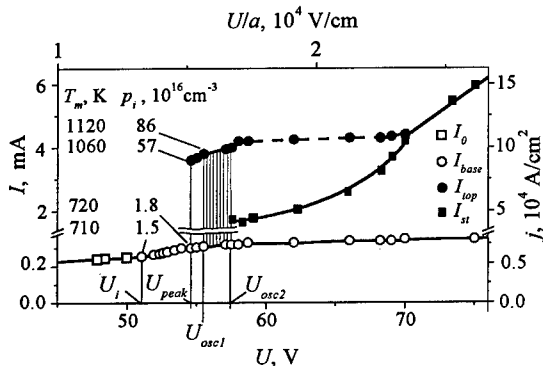
Fig.3. *I-V* curve.

The dotted lines in Fig.2,*b-e* are oscillograms of voltage U_{probe} between probe and source (Fig.1). It follows from them, that source-probe and probe-drain voltages are approximately equalized during the peak.

A yellow glow of the region near the drain with length of (5...7) μm was observed at $U \geq U_i$.

2. DISCUSSION

Up to voltage U_i the *I-V* curve (Fig.3) is a typical drain characteristic of the field-effect transistor with only back gate [1,2], and we need to explain the current behavior only at U>U_i. Both great power $P=IU$, dissipated in a film at such voltages (average power $\overline{P}$ =P/awd_{Si} for U=U_i is equal 0.1 GW/cm³), and glowing of the drain region, allow us to assume, that the observed effects are caused by film heating. Key in their explanation is a mechanism of pair concentration change in a non-uniformly heated semiconductor named *b*-drift [3,4].

2.1. *b*-drift

Suppose there are electrons and holes with concentration and mobility n, μ_n and p, μ_p respectively in a silicon film (Fig.1), n>p and the concentration of electron-hole pairs is equal p, $\mu_n \propto T^{-2.42}$, $\mu_p \propto T^{-2.2}$, b=$\mu_n/\mu_p \propto T^{-0.22}$ [5]. If T=$T(x)$, then a current with density j changes concentration of pairs p with rate g, which (for p>>n-p) is given by $g = -\dfrac{0.22b}{(1+b)^2}\dfrac{j}{e}\dfrac{1}{T}\dfrac{\partial T}{\partial x}$ [3,4]. This phenomenon of pair concentration change is named *b*-drift. The signs of g and $\left(-j\partial T/\partial x\right)$ are the same. If $\left(-j\partial T/\partial x\right)$<0, *b*-drift decreases p.

2.2. Occurrence of peak and oscillations

Heat flows to the substrate, source and drain resulting in silicon film cooling. As its ends are cooled most intensively [6], for a fixed moment of time $T(x)$ is a curve with maximum T_m at a point x_m (curve 1 in Fig.4). The point x_m is nearer to the drain, because the high-resistance area near it is heated strongly with current.
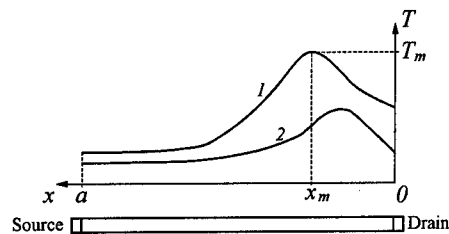
Fig.4. Schematically shown distribution of $T(x)$.

After applying of $U>U_i$ the current growth (Fig.2) after T_m reached the temperature T_i of intrinsic conductivity beginning and an intensive thermal generation of electron-hole pairs begins. After pair appearance the b-drift occurs. It increases both pair concentration and fim conductivity in the source-x_m area, where $\partial T/\partial x<0$ and $g>0$, and decreases them in x_m-drain area, where $\partial T/\partial x>0$ and $g<0$. Such film conductivity change decreases in the course of time the current I and temperature T_m from their maximal values I_{top} and T_m^{top}, and moves the greatest field strength area, and therefore the maximum of $T(x)$ distribution, to a drain (curve 2 in Fig.4). Upon that, increased cooling of the maximum temperature area by the drain, as x_m approaches it, additionally reduces current and temperature. The similar change of $E(x)$ and $T(x)$ by b-drift also was observed in injected electron-hole plasma of germanium [3] and silicon [4] diodes and was accompanied by the thermal glowing [7].

In a case of $U>U_{osc2}$ both stationary current I_{st} (Fig.2,e) and temperature distribution are established after the current fall. The voltage less than U_{osc2} does not provide sufficient heating for maintenance of a stationary current. In this case current and temperature fall down to initial values (a-d), then the recurrence of current peaks (c,d) is possible.

The probe-drain voltage is composed from probe-x_m and x_m-drain voltages. Beginning with heating, the b-drift increases conductivity of probe-x_m area, where $\partial T/\partial x<0$ and $g\propto\left(-T^{-1}\,\partial T/\partial x\right)>0$, and reduces the voltage across this area. Voltage equalization on two half of film during a peak (Fig.2,b-e) is associated with this reason.

2.3. Heating temperature

At a point x_m the rate $g=0$ and pair concentration p is changed only by their generation (recombination) and diffusion, and also by bipolar drift at small p. If the first mechanism changes p more strongly than others and pair lifetime τ is much less than characteristic time of T_m changing, then the thermodynamic equilibrium n and p values are realized at point x_m. Thus for $T_m>T_i$ there are $n=p=p_i$, where p_i is a thermodynamic equilibrium concentration of electron-hole pairs in i-silicon. The current through the film quasi-stationary follows temperature T_m and can be written as

$$I=wd_{Si}ep_i(T_m)\{v_n[E(x_m),T]+v_p[E(x_m),T]\},$$

where v_n and v_p are electron and hole drift velocities [5].

We have replaced $E(x_m)$ by the average field strength value U/a and have constructed $I(T_m,U)$ curves. T_m were found from comparison of experimental $I(U)$ values with those curves for various I at $U>U_i$. The above replacement did not introduce an essential error to the calculated T_m value, because the current depends strongly on T_m (approximately exponentially) and linearly on (v_n+v_p). Therefore the dependence of T_m from (v_n+v_p) is weak. Besides, according to U_{probe} measurements (Fig.2) the heating up makes the E distribution along the film more smooth.

Some values of film maximal temperature T_m and appropriate pair concentration are given in Figs.2,b and 3.

2.4. Thermal resistance

Temperature values T_m in a film at currents I_{base}, I_{top} and I_{st} and power values P, equal to $I_{base}U$, $I_{top}U$ and $I_{st}U$, we shall denote T_m^{base}, T_m^{top} and T_m^{st} respectively. They were calculated from the data of Fig.3. In all changing range of T_m from 720 to 1100 K and power from 0,01 to 0,25 W (of $\overline{P}$ from 0.05 to 1.5 GW/cm³), values T_m^{base} (o) for $U<U_{osc2}$ and T_m^{top} (●) converge on strait line $T_m-T'=R_hP$ with $T'=T'_{base,top}$ =690 K and thermal resistance $R_h=R_h^{base,top}$ =1830 K/W. T_m^{st} (■) values have the similar behavior. In this case $T'=T'_{st}$ =914 K, $R_h=R_h^{st}$ = 590 K/W, and changing ranges of T_m, P and $\overline{P}$ are 900...1200 K, 0.10...0.65 W and 0.5...3.6 GW/cm³.

Calculated temperature dependence on power is characteristic for stationary heat flow through thermal resistance R_h. If it is supposed, that released heat basically leaves in a drain, then T' value is close to the drain temperature. According to reasons of section 2.2, at I_{st} the temperature maximum point x_m is located closer to the drain, than at I_{base} and I_{top}. This explains inequalities $T'_{st}>T'_{base,top}$ and $R_h^{base,top}>R_h^{st}$.

The stationary kind of dependence T_m on P is in accordance with the assumptions made at temperature T_m definition.

3. ESTIMATION OF CHARACTERISTIC VALUES

Let's confirm the above interpretation by an estimation of characteristic values. From quasi-stationary relation between current and temperature T_m it follows that pair lifetime τ is less than the peak duration. From the data of Fig.2 we have obtained $\tau \leq 1\mu s$. This value is in agreement with the lower limit value of $\tau \approx 0.1$ μs, given in Refs.[1,2].

Let's denote the thermal conductivity, heat capacity and density by χ, c and ρ respectively, then coefficient of thermal diffusion $D^h = \chi/c\rho$. Values of quantities in silicon and oxide we shall mark with subscripts "*Si*" and "*ox*", χ_{ox}=const, $\chi_{Si} = \chi_{Si}(T)$. $\tau^h_{Si(ox)} = d^2_{Si(ox)}/D^h_{Si(ox)}$ are times of thermal diffusion through silicon and oxide films, $\tau^h_{ox} \approx 0.01$ μS, $\tau^h_{Si}(1000K) \approx 1\mu S$ and $\tau^h_{ox} >> \tau^h_{Si}(T)$. This inequality allow the heat diffusion in silicon film to be treated as one-dimensional process with leakage of heat in an isothermal substrate and to introduce length [6] and time of cooling

$$\lambda_{cool}(T) = \sqrt{D^h_{Si}(T)\tau_{cool}} \, , \tau_{cool} = \left(c_{Si}\rho_{Si}/\chi_{ox}\right)d_{Si}d_{ox}.$$

The time $\tau_{cool}(1000K)$ is equal 0.5 μS, that is in agreement with peak duration (Fig.2).

The voltage U_{probe} increasing during the peak (Fig.2,*b-e*) is due to the temperature gradient with $\partial T/\partial x < 0$ and $g \propto \left(-T^{-1}\partial T/\partial x\right) > 0$ occurrence in probe-point x_m area. At film cooling after the current peak, $\left(-T^{-1}\partial T/\partial x\right)$ decreases more slowly than T_m. Therefore U_{probe} falls down more slowly, than the current in peak (Fig.2,*b-d*).

At $U > U_{osc2}$ a temperature $T^{st}_m > T^{base}_m$ is established after peak, and there is significant temperature gradient in the probe-point x_m area. According to that the U_{probe} voltage does not decrease after the peak (Fig.2,*e*).

The length $\lambda_{cool}(1000K)$=3 μm and is about half of the glowing area length. Therefore it is assumed that the glowing of a film is associated with its heating.

From $P \approx wd_{Si}\chi_{Si}(T_m)(T_m - T')/[\lambda(T_m)]$ one can easily estimate $R^{est}_h \approx \lambda(T_m)/wd_{Si}\chi_{Si}(T_m)$.

Resistance $R^{est}_h(1000K) \approx 5 \cdot 10^3$ K/W, that is at the same order as $R^{base,top}_h$.

The above explanation of peak occurrence assumes that the pair concentration change rate g reaches the value of thermal generation rate, whose maximum is p_i/τ. Let's compare these values.

Using the formula for g (section 2.1), we shall write such expressions

$$|g| \approx \frac{0.22b}{(1+b)^2} \frac{(\mu_n + \mu_p)p_i U}{a} \frac{1}{T} \frac{T - T_i}{\lambda} , \ |g| = \frac{p_i}{\tau_b}$$

for introduction of τ_b. At T=1000 K and U=55 V the time τ_b= 0.1 μS $\approx \tau$ and $|g| \approx p_i/\tau$.

REFERENCES

1. J.P.Colinge, *Silicon-On-Insulator Technology : Materials to VLSI*, Kluwer Academic Publishers (1991).
2. S.Cristoloveanu and S.S.Li, *Electrical Characterization of Silicon-on-Insulator Materials and Devices*, Kluwer Academic Publishers (1995).
3. V.N.Dobrovolsky and S.P.Pavlyuk, Fiz. Tekn. Poluprovodn., **15**, 120 (1981).
4. V.N.Dobrovolsky, S.P.Pavlyuk and A.V.Romanov, Fiz. Tekn. Poluprovodn., **15**, 411 (1981).
5. C.Canali, G.Majni, R.Minder and G.Ottaviani, IEEE Trans. Electron Devices, **ED-22**, 1045 (1975).
6. M.Berger and Z.Chai, IEEE Trans. Electron Devices, **ED-38**, 871 (1991).
7. V.N.Dobrovolsky, S.P.Pavlyuk, I.E.Paltsev and A.V.Romanov, Fiz. Tekn. Poluprovodn., **19**, 1513 (1985).

The effect of confinement and temperature on the initial hole trapping efficiency of buried oxide films

M.E. Zvanut[a], Wang Chun[a], and R.E. Stahlbush[b]

[a]University of Alabama at Birmingham, Dept of Physics, Birmingham, AL, 35294-1170

[b]Naval Research Laboratory, Washington, DC 20375

Vacuum ultraviolet hole injection was used to study the initial trapping properties of three different types of buried oxides, SIMOX, Unibond, and polySi/SiO$_2$/Si samples. After accounting for the charge centroid, the initial hole trapping efficiency for SIMOX samples is shown to be at least a factor of four greater than that measured for Unibond buried oxides. We suggest that the difference is due to the post fabrication anneal temperature and the lack of moisture in the ambient during the anneal.

1. INTRODUCTION

Many investigations have attributed the dissimilarities between a gate oxide and buried oxide (BOX) to the encapsulation and high temperature anneal of the BOX [1-5]. Using vacuum ultraviolet induced hole injection, others have investigated the hole trapping characteristic of various types of BOX [6].

Focusing on the trapping of positive charge during VUV illumination, we have used capacitance-voltage (CV) and photocurrent-voltage (PIV) measurements to study different types of buried oxides. Unlike earlier work, we observe a dependence of the initial trapping rate on anneal temperature. We also explore the effect of confinement by comparing charge trapping in nonburied oxides annealed in a dry environment with buried oxides treated at the same temperature.

Experimental Details

Three different types of buried oxides were examined. Each oxide was approximately 370 nm thick and annealed in an Ar + 1% O$_2$ ambient. One wafer was formed by a standard Unibond or "Smart Cut" process where the oxide is subjected to hydrogen implantation, a step necessary to form the thin top Si layer of the Si/SiO$_2$/Si structure, and annealed at approximately 1100° C [7]. The SIMOX wafers were formed by either a single implant process with O dose of 1.5×10^{18} cm^{-2} or triple implant process with two doses of 0.5×10^{18} cm^2 and one dose of 0.8×10^{18} cm^{-2}. Both wafers were annealed at 1310° C for 5-6 hr after each implant dose. Several oxides were thermally grown in steam at 850° C on (100) Si. A 450 nm thick polysilicon layer was deposited onto two of the steam grown oxides. These will be referred to as buried thermal oxides. A third steam-grown oxide remained uncoated and will be referred to as the non-buried oxide. The buried thermal oxides were annealed at 1100° C or 1250° C for 5 hr, and part of the non-buried oxide sample was annealed in dry N$_2$ (< 1 ppm H$_2$O) at 1000° C for 5 hr. After formation of the Si/SiO$_2$/Si structures, the top silicon layer was stripped from each of the BOX samples, and Cr/Au contacts were deposited on the oxide surface. Hole injection was performed using 10.2 eV light from a 30 W D$_2$ lamp while a field of 1 MV/cm was applied across the MOS capacitor. One MHz CV measurements made before and after hole injection indicate several different types of distortion; however, the most prominent effect for all the buried oxides is the shift along the voltage axis, ΔV_{cv}. This total shift will be interpreted in terms of trapped hole density through the equation, $x_1 Q = C_{ox} \Delta V_{cv}$ where x_1 is the charge centroid, Q is the total trapped

charge density, and C_{ox} is the oxide capacitance/contact area. The charge is often separated into that which can be discharged by a reverse bias and that which is stable with respect to bias [8]. Here, no distinction is made between stable and unstable trapped positive charge. Effects from electron trapping during VUV illumination due to 4-5 eV light were investigated and found to be negligible. However, electron injection caused by 5-10 eV photons has not yet been considered.

To separately determine x_1 and Q, we have performed photocurrent-voltage (PIV) measurements on samples before and after VUV exposure. Photocurrent was produced using a 100W Hg lamp/monochrometer. The VUV-induced hole trapping can cause shifts in the photocurrent along the positive and negative gate voltage axis. These shifts may be combined to provide the total amount of charge trapped farther than about 3 nm from the interfaces and the centroid for that charge [9]. PIV was measured between 0.5-1 MV/cm to minimize perturbation in ΔV_{cv} [9].

2. RESULTS

Figure 1 illustrates the time dependence of hole trapping measured on triple implant SIMOX (■), single implant SIMOX (▲), and Unibond (●) wafers. The results obtained on an unannealed, non-buried thermal oxide (◊) are shown for comparison. The initial trapping efficiency, ξ, is defined as the ratio of the trapped hole density to VUV exposure time near t=0 (the initial slope of the curve). ξ for the SIMOX data is a factor of four greater than that measured for the Unibond. Both are at least an order of magnitude greater than ξ measured from the thermal oxide results. The extended time behavior of the hole trap density of single implant SIMOX is different from that of the triple implant material, but more complete data must be obtained before the detailed kinetics can be determined. It is tempting to attribute the differences illustrated in figure 1 to fabrication technique, but the results shown in figure 2 suggest that the true source of the difference is anneal temperature. Figure 2 compares the results of the triple implant SIMOX (■) and Unibond (●) with data obtained from buried thermal oxides annealed
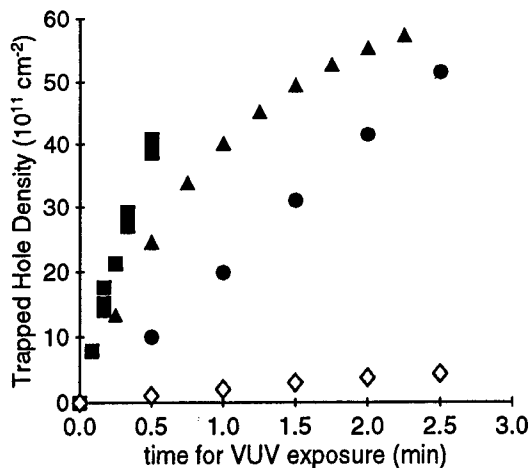
Figure 1. Density of trapped positive charge calculated from CV curve shifts during VUV illumination with applied oxide field of 1 MV/cm. ■ Triple Implant SIMOX; ▲ Single Implant SIMOX; ● Unibond; ◊ unannealed non-buried thermal oxide
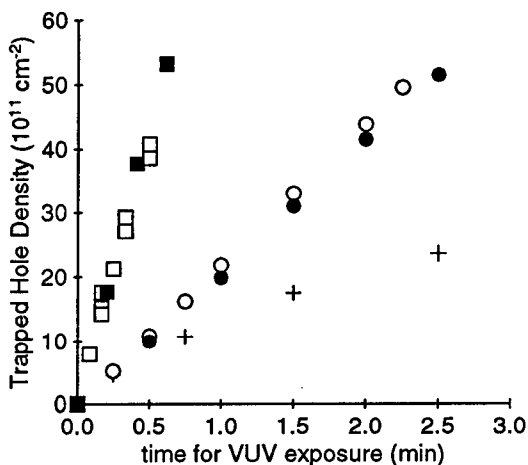
Figure 2. Density of trapped positive charge calculated from CV curve shifts during VUV illumination with applied oxide field of 1 MV/cm. ■ Triple Implant SIMOX; □ 1250° C annealed buried thermal oxide; ● Unibond; O 1100° C annealed buried thermal oxide; + 1000° C dry heat-treated non-buried thermal oxide

at 1250° C (□) and 1100° C (O). The buried oxide data coincide when the anneal temperature

is similar, regardless of fabrication technique. The figure also displays the results obtained from a non-buried oxide (+) annealed at 1000° C in a hydrogen-deficient ambient.   The anneal conditions were chosen to imitate the situation experienced by a buried oxide in which the Si layer prevents most hydrogen-related species from reaching the oxide during the anneal. Although the long time behavior differs, ξ for the thermal oxide is approximately the same as that obtained from the Unibond samples and is a factor of 4 smaller than the data measured on the higher temperature annealed samples.

Although figures 1-2 provide information regarding the shift of the CV curves as a function of   VUV   exposure   time,   a   fundamental understanding of the differences between the buried oxides requires knowledge of the charge location and total trapped charge density, independently. To estimate these, PIV measurements were performed on the capacitors periodically throughout the VUV hole injection process.   Figure 3 depicts results obtained during the first 15 s for a Unibond samples (fig. 3a) and 30 s for the single implant SIMOX (fig. 3b). The Unibond sample indicates a change in the positive gate bias current after the first 15 s of hole injection. On the other hand, the SIMOX data indicate a change in the negative bias photocurrent initially, followed by continued change under negative bias and a small shift under positive bias. The PIV data of the two samples suggest that the centroid for the holes trapped in Unibond oxide is

near the Si/SiO$_2$ interface (x$_1$~1), similar to the situation in a buried thermal oxide. However, fig. 3b shows that the centroid for trapped holes in the SIMOX sample is located close to the Al/SiO$_2$ interface initially (x$_1$~0), and then moves towards the center of the film (x$_1$~0.2). Beyond the time period shown here, the PIV curves do not change.

3. DISCUSSION

The data of figure 2 illustrate that the initial rate of hole trapping depends on the anneal temperature   of   the   buried   oxide   and   is independent of buried oxide fabrication technique. The temperature dependence shown in figure 2 is consistent with x-ray irradiation studies, but contrary   to   previously   reported   VUV   hole injection studies [1,6]. VUV studies similar to those reported here showed that hole trapping is identical for different types of buried oxides including   SIMOX   and   Unibond   [6].   The discrepancy between the earlier work and that presented here may be attributed to different irradiation intensity.   We estimate that the D$_2$ source used for our experiments is three orders of magnitude less intense than the Kr discharge source used by Afanes'ev and coworkers. It is likely that the lower intensity decreases the

Figure 3a. PIV curves measured during VUV illumination on Unibond buried oxides with applied oxide field of 1 MV/cm. (1) before VUV exposure (2) after 15 s VUV exposure
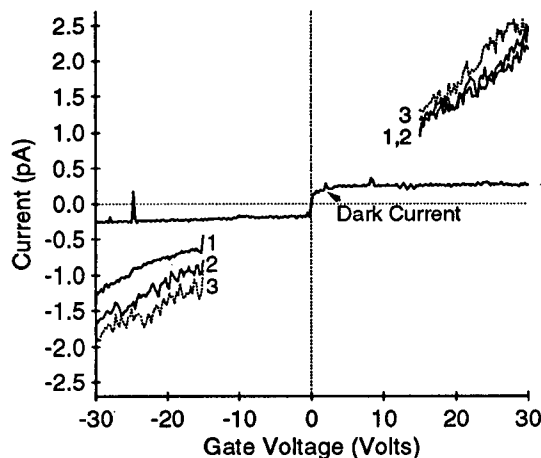
Figure 3b. PIV curves measured during VUV illumination on Single Implant SIMOX with applied oxide field of 1 MV/cm. (1) before VUV exposure (2) after 15 s VUV exposure; (3) after 30 s VUV exposure

probability of trap creation that is known to occur during use of the higher intensity source [6]. Thus, the results of the earlier work may indicate that trap creation by VUV irradiation is the same for the different types of samples, while our work suggests that intrinsic hole trapping differs for buried oxides annealed at different temperatures.

The location of the trapped holes in the Unibond BOX is similar to that observed in non-buried thermal oxides. In this case, holes reside primarily near the Si substrate/SiO$_2$ interface. The SIMOX data present a different situation. For the single implant SIMOX, the centroid for the trapped holes is weighted closer to the Al/SiO$_2$ interface than to the Si/SiO$_2$ interface. Further, the centroid moves toward the Si as charge trapping proceeds. This is consistent with the picture suggested by x-ray irradiation studies of SIMOX BOX [1,10]. The PIV results imply that the charge density reported in figures 1 and 2 must be adjusted. Because the centroid for the trapped holes in SIMOX is less than one, the charge density calculated from the CV curve shift represents less than the total amount of charge in the sample. However the Unibond data, which exhibit a centroid close to 1, represent almost the entire charge density. After accounting for the centroid, ξ for SIMOX will be larger than that calculated from figure 2, and the difference obtained for the initial trapping efficiency between the types of buried oxides will be increased. We have not yet obtained a reliable centroid for the buried thermal oxides, so the adjustment required for these samples is not known, and we cannot form a quantitative relationship between anneal temperature and hole trapping. We can conclude that there exists at least a factor of four difference between the initial trapping efficiencies for SIMOX and Unibond buried oxides.

Finally, consider the non-buried thermal oxide data represented by '+' in figure 2. The similarity between these data and the 1100° C annealed buried oxide data for time less than 0.5 min suggests that the presence of the top Si is not the key feature determining the initial hole trapping characteristics. Rather, the hydrogen-deficient environment experienced by a buried thermal oxide during the post fabrication anneal is responsible for the behavior.

4. SUMMARY

Using VUV hole injection and PIV measurements, we have determined that the hole trapping efficiency for SIMOX samples is at least four times that for Unibond samples. The difference appears to be related to the anneal temperature rather than the fabrication technique. Comparison to a non-buried oxide subjected to a moisture-deficient heat treatment suggests that the source of the difference is the lack of hydrogen-related species reaching the buried oxide during the anneal.

REFERENCES

1. R. E. Stahlbush, G.J. Campisi, J.B. McKitterick, W.P. Maszara, P. Roitman, and G.A. Brown, IEEE Trans. Nucl. Sci. 39, 2086 (1992).
2. O. Gruber, P.Paillet, J.L. Autran, B. Aspar, A.J. Auberton-Herve, Microelectronic Engineering 36, 387 (1997).
3. R.A.B. Devine, W. L Warren, J.B. Xu, I. H. Wilson, P. Paillet and J.-L. Leray, J. Appl. Phys. 77, 175 (1995).
4. M.E. Zvanut, T.L. Chen, R.E. Stahlbush, E.S. Steigerwalt, and G.A. Brown, J. Appl. Phys. 77, 4329 (1995).
5. R.E. Stahlbush, IEEE Trans. Nucl. Sci. 44, 2106 (1997).
6. V.V. Afanas'ev, A. Stesmans, A.G. Revesz, and H.L. Hughes, J. Electrochem. Soc. 144, 749 (1997).
7. M. Bruel, Electronics Lett. 31, 1201, 1995.
8. D.M. Fleetwood, IEEE Trans. Nucl. Sci. 39, 269 (1992).
9. D.J. DiMaria, Z.A. Weinberg, and J.M. Aitken, J. Appl. Phys. 48, 898 (1977).
10. H.E. Boesch Jr., T.L. Taylor, L.R. Hite, W.E. Bailey, IEEE Trans. Nucl. Sci. NS-37, 1982 (1990).

Designing MOS/SOI Transistors for High Frequency and Low Voltage Applications

V. Ferlet-Cavrois[a], A. Bracale[a, b], C. Marcandella[a], O. Musseau[a],
J. L. Pelloie[c], C. Raynaud[c], O. Faynot[c]

[a]CEA/DIF, BP12, 91680 Bruyères-le-Châtel, France

[b]ENSEA, 6 avenue du Ponceau, 95014 Cergy-Pontoise, France

[c]CEA/DTA-LETI, 17 rue des Martyrs, 38054 Grenoble Cedex 9, France

With the development of digital SOI technologies for portable communication systems [1, 2], high frequency modelization and design efforts are necessary to take into account some specific SOI effects. In this paper we evaluate the influence of the transistor design (process and geometry) on its frequency performances. We particularly investigate the influence of gate length and width, the adjunction of body ties to avoid floating body effects, and the DTMOS structure [3-5] for low voltage applications.

1. INTRODUCTION

During the last decade, the growing needs of wireless communications in the GHz range leads to the development of specific high frequency technologies, like GaAs, bipolar SiGe, BICMOS, etc. However, the gate length shrinking of silicon CMOS technologies improves their high frequency performances and make them suitable for GHz operation. In particular, SOI CMOS technologies, developed for digital applications [6, 7], show good frequency results at low supply voltage. SOI performances allow integration of RF parts with digital treatment on a single chip. This reduces both cost and power consumption of portable systems. In this paper, we show the high frequency performances of a partially depleted 0.25μm SOI technology [8]. In a first part, we show the influence the transistor design and process. In a second part, we show the influence of body ties, and the potential of DTMOS for very low voltage (0.6 V) RF applications.

2. TECHNOLOGY AND EXPERIMENT DESCRIPTION

NMOS and PMOS transistors are processed in a **partially depleted 0.25μm SOI technology** (Fig. 1a). The substrates are 8' UNIBOND wafers with a resistivity of 20 Ω.cm. LOCOS dielectrically isolates transistors. After gate oxide growth, a polysilicon gate is deposited and etched. Low Doped Source and Drain (LDS and LDD) allows to improve reliability and short channel effects. After oxide spacer process, the active region is **silicided (TiSi₂)** to reduce the source, drain and gate access resistances. The final thickness of gate oxide, silicon film and buried oxide are respectively 45 Å, 1000 Å, and 4140 Å.

Transistors were measured with a Cascade probe station and an HP8510 network analyzer. They are composed of several gate fingers to reduce the gate access resistance (Fig. 1b). They are embedded in a coplanar waveguide structure, and a TRL calibration is realized with in situ calibration structures.

3. INFLUENCE OF THE TRANSISTOR GEOMETRY AND PROCESS

Fig. 2 shows the transition frequency f_T and the maximum oscillation frequency f_{max} versus transistor gate length L. The tested transistors have Nf = 8 gate fingers of Wf = 12.5 μm width (total width $W = Wf \times Nf = 8 \times 12.5 = 100$ μm). The gate and the drain are biased at 1.5 V. f_T increases inversely to the gate length L, which is consistent with the usual transconductance behavior (gm ∝ 1/L). Moreover the f_T of NMOS is approximately 1.7 time higher than the f_T of PMOS because of the electron and hole mobility ratio. For both NMOS and PMOS devices, f_{max} is higher than f_T for long gate length (0.8 μm), and lower than f_T for short gate length (0.25 μm). f_{max} is affected by parasitic elements (gate

and source resistances), while f_T is not, as shown by their analytic expressions (Fig. 3). The gate resistance R_g (Fig. 3) increases as the gate length is reduced, and contributes to degrade f_{max}.

Transistors with different width of gate fingers Wf have been tested to evaluate the influence of the gate geometry. Fig. 4 shows that f_T remains approximately constant. The variations of gate and source resistances R_g and R_s with Wf does not influence f_T as shown by f_T expression on Fig. 3. The slight decrease of f_T for short Wf, at 3.1 µm, is probably due to the influence of LOCOS edge capacitances. Fig. 4 also shows that f_{max} becomes higher than f_T when Wf decreases. f_{max} reaches a maximum value and saturates when Wf ≤ 6.25 µm. R_g and R_s strongly influence f_{max} (Fig. 3). When Wf decreases, R_g decreases, but R_s increases (Fig. 3). Both the expression of f_{max} (Fig. 3) and the HF measurements (Fig. 4) show that an optimum f_{max} can obtained for 3.1 µm ≤ Wf ≤ 6.25 µm. If the gate sheet resistance was lower (aluminium or T-gate), the maximum f_{max} would have been obtained for wider transistor gate fingers. To summarize, **f_T and f_{max} values higher than respectively 40 GHz and 50 GHz are reached for 0.25 µm NMOS transistors at 1.5 V with optimized geometry (short gate fingers).**

Fig. 5 shows the f_T and f_{max} values versus supply voltage for the 0.25µm NMOS and PMOS transistors with optimized geometry (Wf = 6.25 µm, Nf = 8). **f_T and f_{max} of floating body NMOS and PMOS are constant on a wide range of supply voltage down to 1 V.** The floating body effect of SOI transistors lowers the threshold voltage and allow high frequency operation at low supply voltage.

4. BODY TIES, DTMOS

Adding body ties to a transistor implies a specific gate shape. We tested transistors with an H-shape gate (Fig. 1c) [9] with either a standard or an improved version of design rules. Transistors with grounded body are useful to avoid floating body effects. Another interesting structure, the DTMOS (Dynamic Threshold voltage MOS) [3-5], is obtained by connection of body ties to gate instead of ground. DTMOS can be used for very low supply

voltage applications (0.6 V). Fig. 6 compares the HF performances of NMOS with floating body, grounded body, and DTMOS. The grounded body and the DTMOS transistors are designed with either standard or improved design rules. Fig. 6 shows the higher f_T and f_{max} values obtained with the new version of design rules.

At high supply voltage (≥ 1 V), the DTMOS and the grounded body NMOS with new design rules show good high frequency results, close to the floating body NMOS. The new design rules lower the parasitic gate-source and gate-drain capacitances induced by the specific gate shape.

At low supply voltage (≤ 1 V), the observed differences are due to the variation of the threshold voltage with the transistor structure. The grounded body transistor has a higher threshold voltage (Vt = 0.61 V) than the floating body transistor (Vt = 0.56 V) and shows the lowest f_T and f_{max} values. However, it can be used for 1 V applications without any performance degradation. DTMOS has the lowest threshold voltage (Vt = 0.48 V) and shows the best performances at low supply voltage. **At 0.6 V, the f_T and f_{max} values of the DTMOS with improved body tie design rules are respectively of 16 GHz and 33 GHz.**

REFERENCES

[1] K. Takeya, "Telecommunications LSIs Using 0.25µm CMOS/SIMOX Technology", Electrochemical Society Proceedings Volume 97-23, p. 389-400, 1997.

[2] I. Lagnado, P. de la Houssaye, "TFSOS : Quo Vadis," Electrochemical Society Proceedings Volume 97-23, p. 329-339, 1997.

[3] J. P. Colinge, "An SOI Voltage-Controlled Bipolar-MOS Device," IEEE Trans. Electron Devices, vol.ED-34, No. 4, pp.845-849, April 1987.

[4] F. Assaderaghi, D. Sinitsky, S. Parke, J. Bokor, P. K. Ko, C. Hu, "A Dynamic Threshold Voltage MOSFET (DTMOS) for Ultra-Low Voltage Operation," Tech. Dig. of IEDM, p. 809, 1994

[5] T. Tanaka, Y. Momiyama, T. Sugii, "Fmax Enhancement of Dynamic Threshold-Voltage MOSFET (DTMOS) Under Ultra-Low Supply Voltage," IEDM Technical Digest, pp. 423-426, Dec. 1997.

[6] M. Canada, C. Akrout, D. Cawthron, J. Corr, S. Geissler, R. Houle, P. Kartschoke, D. Kramer,

P. McCormick, N. Rohrer, G. Salem, L. Warriner, "A 580 MHz RISC Microprocessor in SOI," ISSCC proceedings, paper WP 25.3, p. 430, Feb. 1999.

[7] Y. W. Kim, S. B. Park, Y. G. Ko, K. I. Kim, I. K. Kim, K. J. Bae, K. W. Lee, J. O. Yu, U. Chung, K. P. Suh, "A 0.25 µm 600 MHz 1.5 V SOI 64b ALPHA™ Microprocessor," ISSCC proceedings, paper WP 25.4, p. 432, Feb. 1999.

[8] J. L. Pelloie, O. Faynot, C. Raynaud, B. Dunne, F. Martin, S. Tedesco, J. Hartmann, "A Scalable SOI Technology for Three Successive Generations : 0.18, 0.13 and 0.1 µm for Low-Voltage and Low-Power Applications," IEEE International SOI Conference proceedings, Oct. 1996, p. 118.

[9] C. B. Oh, J. Ahn, Y. W. Kim, "The Effect of Body Contact Arrangement on Thin SOI MOSFET Characteristics," INFOS'95, Microelectronic Engineering, Vol. 28, 1-4, p. 467, (1995).

a) Transistor section

b) Top view of a transistor without body ties
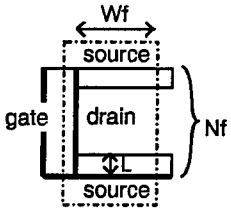
c) Top view of a transistor with body ties
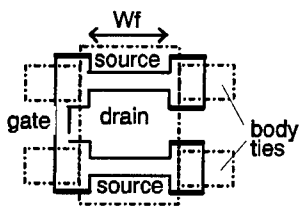
Fig. 1 : Section (a) and top views of SOI transistors without (b) or with (c) body ties. The transistors have Nf gate fingers of Wf width.

Fig. 2 : Transition frequency f_T and maximum oscillation frequency f_{max}, versus gate length L of NMOS and PMOS transistors.
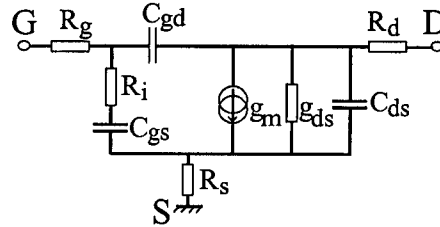
R_g, R_s, R_d	Gate, source and drain resistances
C_{gs}, C_{gd}, C_{ds}	Gate-source, gate-drain, and drain-source capacitances
g_m, g_{ds}	RF transconductance and conductance
R_i	Equivalent charging resistance

$$f_T = \frac{g_m}{2\,\pi\,(C_{gs} + C_{gd})}$$

$$f_{max} = \frac{f_T}{2\,\sqrt{g_{ds}\,(R_g + R_s + R_i) + 2\pi R_g\,f_T\,C_{gd}}}$$

$$R_g \propto R_{\,g}\,\frac{1}{Nf}\frac{Wf}{L}$$

$$R_s \propto R_{\,s}\,\frac{1}{Nd\,Wd}$$

with $R_{\,g}$ and $R_{\,s}$ the gate and source sheet resistances.

Fig. 3 : Simplified small-signal equivalent scheme, and analytic expression of f_T, f_{max}, R_g and R_s [5].

a) Transition frequency f_T
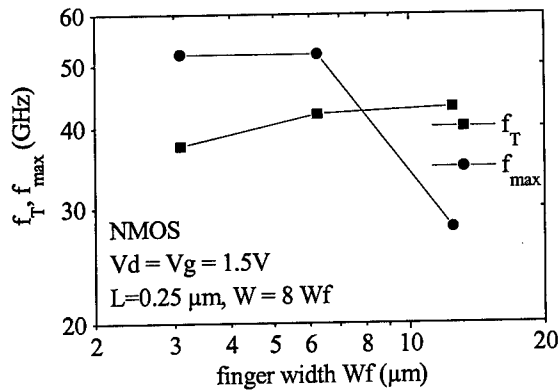
Fig. 4 : Transition frequency f_T and maximum oscillation frequency f_{max}, versus the finger width Wf of NMOS transistors.

b) Maximum oscillation frequency f_{max}
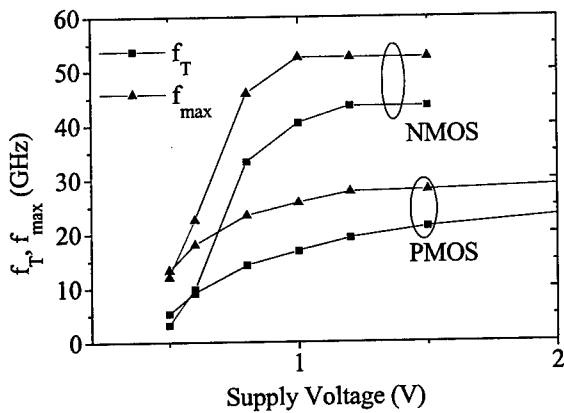
Fig. 5 : Transition frequency f_T and maximum oscillation frequency f_{max}, versus supply voltage for a 0.25µm NMOS transistor with 8 gate fingers of 6.25 µm width (Nf = 8, Wf = 6.25 µm). The gate and the drain are both connected to the supply voltage.
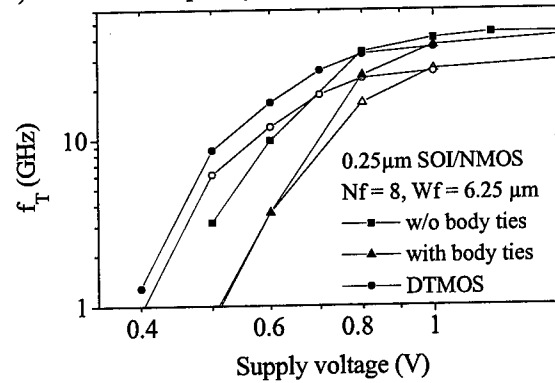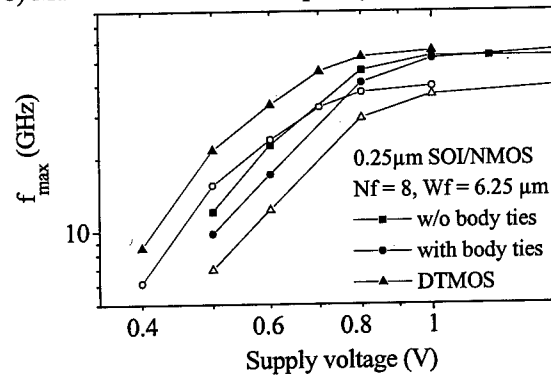
Fig. 6 : Transition frequency f_T and maximum oscillation frequency f_{max} versus the supply voltage of NMOS without body ties (floating body), with body ties, and DTMOS. The gate and the drain are both connected to the supply voltage. The open symbols of NMOS with body ties and DTMOS correspond to the standard version of H-gate design rules. The full symbols correspond to the improved design rules.

Model for the Extraction of the Recombination Lifetime in Partially Depleted SOI MOSFETs

Daniela Munteanu and Sorin Cristoloveanu

LPCS/ENSERG, 23 Rue des Martyrs, BP 257, 38016 Grenoble Cedex 1, France

A new model is proposed for the extraction of the recombination lifetime in partially depleted (PD) SOI MOSFETs. The model uses drain current overshoots induced when turning on the transistor or when pulsing the back gate from strong accumulation to depletion. This model is validated through systematic simulations and is applied to the characterization of recent SOI substrates.

1. INTRODUCTION

The future progress in SOI technology is closely related to the availability of accurate characterization techniques. One of the main parameters for SOI materials (crystal quality) and devices (floating-body effects) is the carrier lifetime [1,2]. Drain current transients (undershoots) in partially depleted (PD) SOI MOSFETs were used for the extraction of the *generation* lifetime [3]. The recombination lifetime has been determined from photoconductivity-decay [1] and pulsed MOS capacitors [4]. This paper presents a new general model for the evaluation of the *recombination* lifetime from the drain current overshoot in both single-gate and dual-gate PD SOI MOSFETs.

2. DUAL-GATE TRANSIENTS

The drain current overshoot is obtained by maintaining the front interface in strong inversion ($V_{G1} > V_T$), while the back-gate is pulsed from accumulation to depletion (Fig. 1). Consequently, the large accumulation charge existing at the back interface is expelled into the neutral body and the depletion region shrinks rapidly. The body potential increases and causes the threshold voltage to drop (i.e. excess current). The carrier recombination process results in the relaxation of the "neutral"-body potential, $V_{BS}(t)$ (Fig. 2), inducing a drain current transient overshoot (Fig. 1). Atlas 2-D numerical simulations fully reproduce the experiment.
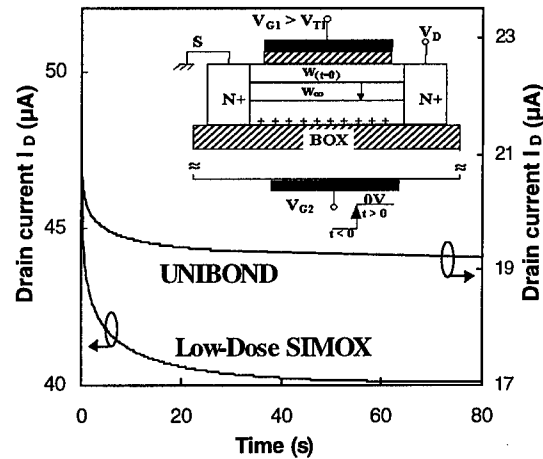
Fig. 1. Experimental setup and current overshoots obtained in PD SOI MOSFETs.

3. MODEL

The charge balance relation describing the current overshoot is:

$$\frac{dQ_D}{dt} + \frac{dQ_a}{dt} = -qR(t) - qn_i s \qquad (1)$$

where Q_D is the depletion charge, s is the effective surface recombination velocity and Q_a is the accumulation charge at the back interface. Q_D and Q_a are correlated with the channel inversion charge and the potential raise, $\Psi_B(t) = V_{BS}(t) - V_{BS}(\infty)$, through

the front-gate and back-gate charge-potential equations:

$$Q_D(t) = -qN_A w(t) = \sqrt{2\varepsilon_{Si} q N_A (\Psi_s(t) - \Psi_B(t))} \quad (2)$$

$$-Q_D(t) - Q_{inv}(t) = -C_{ox1}[V_{G1} - V_{FB1} - \Psi_s(t)] \quad (3)$$

$$Q_a(t) = -C_{ox2}[V_{G2} - V_{FB2} - \Psi_B(t)] \quad (4)$$

where $w(t)$ is the depletion region depth, V_{FB1} is the flat-band voltage at the front interface, $Q_{inv}(t)$ is the inversion channel charge. During the transient, the front interface is permanently in strong inversion, hence the surface potential Ψ_s is constant, $\Psi_s = 2\Psi_F$.
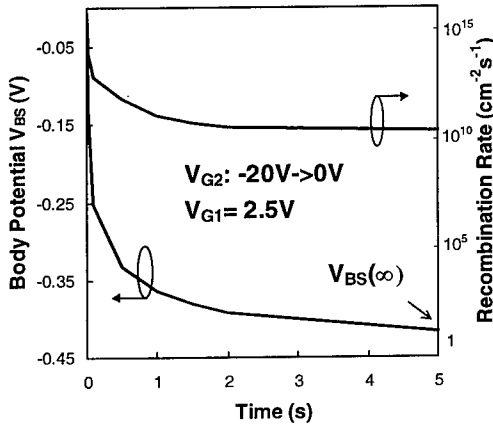
Fig. 2. Transient variation of the body potential and recombination rate after the back-gate pulse.

The volume recombination rate is given by the SRH mechanism:

$$R = (p \cdot n - n_i^2)/[\tau_p(n + n_1) + \tau_n(p + p_1)] \quad (5)$$

$$n_1 = n_i e^{\frac{E_T - E_i}{kT}} \quad \text{and} \quad p_1 = n_i e^{\frac{E_i - E_T}{kT}} \quad (6)$$

where E_i is the middle gap energy, E_T is the trap level energy and $\tau_{n,p}$ are the electron and hole recombination lifetimes. In the following calculations we assume that $\tau_n = \tau_p = \tau_R$ and $E_i = E_T$.

The carrier concentrations n and p can be expressed as:

$$n = n_i e^{\frac{\Psi_i - \Phi_n}{V_{th}}} = n_i e^x, p = n_i e^{\frac{\Phi_p - \Psi_i}{V_{th}}} = n_i e^{a-x} \quad (7)$$

where $\Psi_i = -E_i/q$ is the intrinsic potential, $\phi_n = -E_{Fn}/q$ and $\phi_p = -E_{Fp}/q$ are the quasi-Fermi potentials, $V_{th} = kT/q$, $a = \Psi_B/V_{th}$, $x = (\Psi_i - \Phi_n)/V_{th}$. The increase in the "neutral"-body potential, $\Psi_B(t)$, calculated at the source edge of the body and at the back interface, is expressed as $\Psi_B = \phi_p - \phi_n$.

The product $n \cdot p$ is an in-depth constant, depending only on the transient variation of the potential Ψ_B: $p \cdot n = n_i^2 e^{\frac{\Psi_B}{V_{th}}} = n_i^2 e^a$. Relation (5) becomes:

$$R = \frac{n_i}{\tau_R} \frac{e^a - 1}{e^x + e^{a-x} + 2} \quad (8)$$
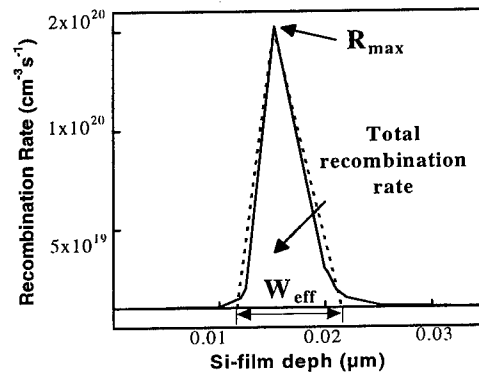
Fig. 3. In-depth variation of the recombination rate 1ms after pulsing the back gate (— 2D simulation, ---- triangular approximation).

The recombination rate, R, is maximum in the depletion region at the film depth where $x = a/2$ (i.e. where the sum of hole and electron concentrations, $n+p$, is minimum and $n=p$):

$$R_{max}(t) = \frac{n_i}{2\tau_R}(e^{\frac{a}{2}} - 1). \quad (9)$$

Figure 3 illustrates the in-depth variation of $R(t)$, suggesting a triangular approximation:

$$R(t) = \frac{1}{2} R_{max}(t) W_{eff}(t) \quad (10)$$

where $W_{eff}(t)$ is an effective recombination region.

The recombination takes place in the depletion region and not in the neutral region as it could have

been expected. The effective recombination width, $W_{eff}(t)$, can be approximated as a function of $\Psi_B(t)$. We define $W_{eff} = w_1-w_2$ (see Fig. 4) as the region where the recombination rate is higher than 5% of the peak value. This condition, replaced in equation (8) gives: $-4 \leq x-a/2 \leq 4$, which indicates the variation of the potential $\Psi_i - \phi_n$ within the effective recombination region W_{eff}: $\Psi_B/2 - 4V_{th} \leq \Psi_i-\phi_n \leq \Psi_B/2 + 4V_{th}$ (Fig. 4).
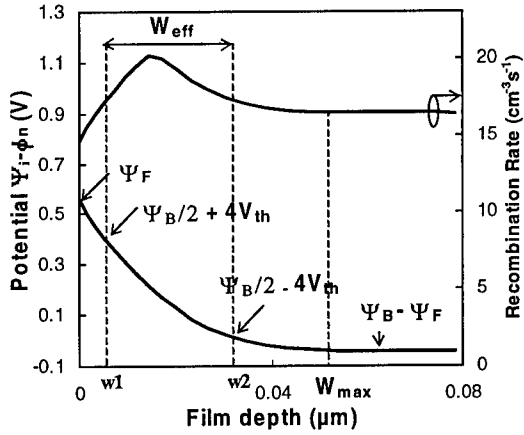
Fig. 4. In-depth variation of the potential $\Psi_i-\phi_n$ and recombination rate.

The potential $\Psi_i-\phi_n$ varies between Ψ_F (at the front interface) and $\Psi_B(t) - \Psi_F$ (in the "neutral" region) and equals $\Psi_B/2$ at the point of maximum recombination rate.

Taking into account the parabolic potential dependence, we calculate the limits w_1 and w_2 and the first-order approximation of the effective recombination region depth:

$$W_{eff}(t) = 4\sqrt{2}\,w(t)\frac{V_{th}}{2\Psi_F - V_B(t)}$$

Finally, the total recombination rate is given by:

$$R(t) = \frac{1}{\tau_R}\frac{1}{A\sqrt{2}}\frac{N_A}{w(t)}(e^{-Aw(t)^2} - \frac{n_i}{N_A}) \qquad (11)$$

with $A=qN_A/(4\varepsilon_{Si}V_{th})$.

Eliminating $\Psi_B(t)$ between equations (1), (2), (3) and using (11) for the recombination rate leads to a generic Zerbst-type relationship:

$$LHS(t) = \frac{RHS(t)}{\tau_R} + S \qquad (12)$$

$$LHS(t) = \frac{dw(t)}{dt}\left[1 + \frac{C_{ox2}}{\varepsilon_{Si}}w(t)\right], S = s\frac{n_i}{N_A} \qquad (13)$$

$$RHS(t) = \frac{1}{A\sqrt{2}}\frac{1}{w(t)}\left[e^{-Aw(t)^2} - \frac{n_i}{N_A}\right] \qquad (14)$$

In the linear region of operation (low V_D), Q_{inv} is given as a function of the drain current $I_D(t)$:

$$Q_{inv}(t) = -\frac{I_D(t)}{V_D\mu Z/L} \qquad (15)$$

where μ is the carrier mobility, Z is the device width and L its length. Combining equations (2)-(3) yields the relation between the variations of w(t) and $I_D(t)$:

$$w(t) = w_\infty + K[I_\infty - I_D(t)] \qquad (16)$$

with $K^{-1} = q\mu N_A V_D(Z/L)$; w_∞ and I_∞ are the steady-state depletion depth and drain current, respectively.
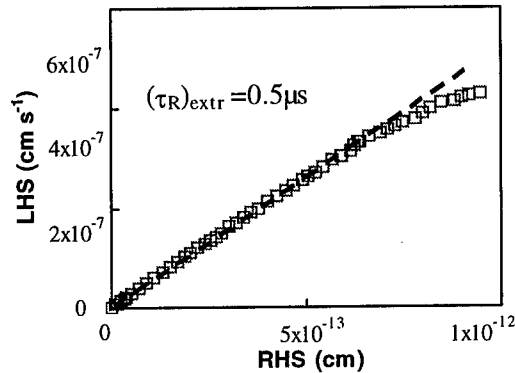
Fig. 5. Extraction of τ_R from the drain current overshoot in a Low-Dose SIMOX MOSFET.

The plot of LHS(t) versus RHS(t) (Fig. 5) gives the recombination lifetime τ_R (slope) and S (the intercept).

4. SIMULATION

The model was validated by a self-consistent procedure: drain current overshoots were simulated by injecting into the simulator a recombination

Table I. Recombination model validation: comparison between carrier recombination lifetimes injected into the simulator and those extracted from simulated single-gate and dual-gate drain current overshoots.

Input τ_{input}, (µs)		0.05	0.1	0.5	1
Extracted τ_{extr} (µs)	Single-gate	0.0538	0.108	0.506	1.05
	Dual-gate	0.054	0.109	0.508	1.06

lifetime and applying the model to extract τ_R from the simulated curves.

Table I shows the excellent matching found between the extracted values and those initially imposed, for a wide range of recombination lifetimes. The validation was conducted for various SOI MOSFET structures (oxide and film thicknesses, film doping) and bias conditions (different front-gate pulse magnitudes).

5. SINGLE-GATE TRANSIENTS

A similar drain current overshoot can be obtained by pulsing the front gate from depletion to strong inversion while the back gate is grounded. The minority carriers needed in the inversion channel are rapidly supplied by the source/drain regions. The positive pulse on the front gate leads to the extension of the depletion region, but the majority carriers expelled from this region cannot be removed instantaneously and are stored in the neutral body. This induces a temporary increase in the potential of the "neutral"-body, $V_{BS}(t)$, and results in an excess drain current. The recombination model is simpler in this case:

$$\frac{dQ_D}{dt} = -qR(t) - qn_i s \qquad (17)$$

In the single-gate transient the expression of R(t) is the same as in the dual-gate case, allowing the extraction of the recombination lifetime from the drain current transient. A Zerbst-type relation similar to relation (12) is obtained, with the same RHS and:

$$LHS(t) = \frac{dw(t)}{dt} \qquad (18)$$

with $w(t) = w_\infty + K[I_\infty - I_D(t)]$.

Extracted values of the recombination lifetime from simulated single-gate transients are presented in

Table I where the same self-consistent procedure was used for validating the model.

6. APPLICATION

This model has been used to characterize two recent SOI materials: UNIBOND ($\tau_R \cong 3$µs) and Low-Dose SIMOX ($\tau_R \cong 0.5$µs). The symmetric experiment performed by pulsing the back-gate in strong inversion allows to estimate the recombination lifetime at the back interface.

The extracted lifetime decreases substantially in short channels, where the source- and drain-body junctions contribute to the evacuation of the majority carriers.

7. CONCLUSION

Single or dual-gate drain current overshoots can be used for estimating the carrier recombination lifetime in partially depleted SOI transistors. Both interfaces can be characterized and compared using symmetric experiments. A simple model, based on a triangular approximation of the recombination rate was proposed for the analysis of all types of overshoot. Optimal extraction of the recombination lifetime requires the use of long channels, the extracted lifetime in short channels being strongly affected by the junctions influence.

REFERENCES

1. S. Cristoloveanu and S.S. Li, Kluwer Academic Publisher, 1995.
2. H.C. Shin et al., IEEE Trans. Electron Devices, vol. 43, p. 318, 1996.
3. D.E. Ioannou et al., IEEE Electron Device Letters, vol. 11, p. 409, 1990.
4. D.K. Schroder et al., IEEE Trans. Electron Devices, vol. ED-31, p. 462, 1984.

Hydrogen Thermal Stability in Buried Oxides of SOI Structures

D. Ballutaud[a], A. Boutry-Forveille[a], and A. Nazarov[b],

[a]Laboratoire de Physique des Solides, Centre national de la Recherche scientifique, 1 Place Aristide Briand, 92195 Meudon cedex (France).

[b]Institute of Semiconductor Physics, National Academy of Science of Ukraine, Prospekt Nauki 45, 25650 Kyiv (Ukraine)

The interactions of hydrogen (deuterium used as tracer) with Si-SiO$_2$-Si buried oxide layers (BOX) prepared by thermal oxidation or by oxygen implantation (SIMOX) are investigated using Secondary Ion Mass Spectrometry (SIMS) measurements combined with effusion experiments and isothermal annealings.

1. INTRODUCTION

Oxygen implantation (SIMOX, for « separation by implantation of oxygen ») is one way to process buried SiO$_2$ layers in Silicon for silicon-on-insulator (SOI) technology [1]. Another way to get buried SiO$_2$ layers is the thermal oxidation of a silicon substrate followed by the deposition of an poly-Si layer recrystallized by laser zone melting (ZMR) [2]. Hydrogenation, typically using a plasma source, is used to improve the electronic properties of poly- or microcrystalline silicon for electronic devices [3]. In general, it is admitted that passivation occurs when hydrogen is trapped on dangling bonds [4], on extended defects as silicon grain boundaries or dislocations. Heterogeneous interfaces such as silicon-silicon-dioxide interfaces also presents electrically active centers able to trap hydrogen atoms [5]. The deuterium permeation through the silicon overlayer and the trapping in the oxide have been previously studied in a high range of temperature (500-1000°C) in SIMOX structures, using a deuterium gas source and nuclear-reaction analysis as characterization technique [6] and the diffusion coefficient of hydrogen has been calculated in thermal SiO$_2$ by capacitance measurements [7].

In the present work, the interactions of hydrogen (deuterium used as a tracer) with SiO$_2$ buried layers (BOX) prepared by thermal oxidation (ZMR) or by oxygen implantation (SIMOX), were investigated using Secondary Ion Mass Spectrometry (SIMS) measurements combined with effusion experiments and isothermal annealing.

2. EXPERIMENTAL

The deuteration of the samples was performed in a deuterium RF plasma reactor (Deuterium plasma pressure: 1mbar) at different temperatures (plasma power density 1W/cm^2) during 30 min. The deuterium concentration profiles were obtained by SIMS measurements with a CAMECA IMS4F apparatus, with a Cs$^+$ primary ion beam (14 keV, 5.10^{-3} A/cm^2). The crater depths were measured with a Tencor stylus profilometer and the absolute concentrations were determined by calibration with deuterium implanted Si standard specimen.

Effusion spectra were measured by a quadrupole mass spectrometer coupled to an evacuated quartz tube (10^{-10} mbar) which contained the deuterated samples, which were

submitted to a linear heating rate (15°C/min) or isothermal annealing. The pumping was controlled so that the mass spectrometer signal was proportional to the deuterium gas flow outgassing from the sample, and was normalized to one square centimeter sample surface [8].

3. RESULTS

3.1. Deuterium diffusion profiles in BOX

Figure 1 shows the influence of the temperature on deuterium diffusion profiles in a SIMOX sample. Permeation through the implanted oxide layer occurs at temperatures higher than 250 °C.
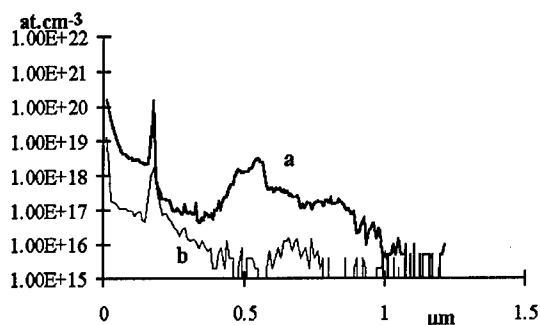
Figure 1. Deuterium concentration profiles in SIMOX samples; after deuteration at 250°C (a) and at 150°C (b)(1W/cm² during 30 min).
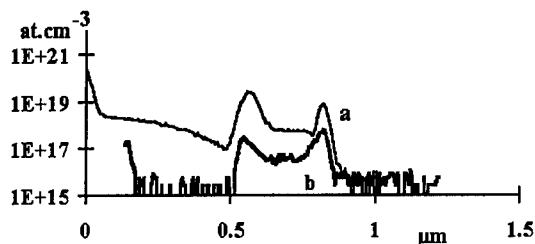
Figure 2. Deuterium concentration profiles in ZMR samples; (a) after deuteration at 250°C; (b) sample (a) after isothermal annealing at 600°C during 2 hours (1W/cm² during 30 min).

The deuterium diffusion profiles observed in the superficial mono-Si layer are in good agreement with the previous works of D. Mathiot et al [9, 10]. The deuterium profiles exhibit undulations showing that some deuterium is trapped on residual reactive sites in the superficial mono-Si and in the implanted oxide layers. The deuterium concentration profile in the BOX and in the silicon substrate evidences some trapping on defects (silicon inclusions or oxide protusions)

The diffusion profile after deuteration at 250°C in ZMR sample is depicted in figure 2 (curve a). The deuterium diffusion profile follows accurately an *erfc function* in the poly-Si superficial layer with an effective diffusion coefficient of 5.5 10^{-13} cm² s⁻¹ at 250 °C (Figure 3), although it is not generally the case in polycrystalline silicon [11, 12]. The large deuterium peak at the poly-Si/SiO₂ interface shows an important deuterium trapping between the thermal oxide and the polycrystalline silicon layer.
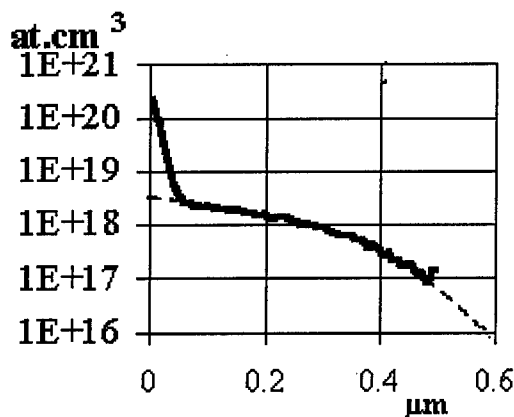
Figure 3. Simulation of the deuterium diffusion profile in the poly-Si corresponding to curve (a) in figure 2.

3.2. Effusion results

The deuterium effusion spectra obtained from SIMOX and ZMR samples after deuteration

($1W/cm^2$ at 250°C during 30 min) are reported in figure 4 (respectively curves a and b). The SIMOX sample effusion spectrum presents three large peaks at respectively 380 °C, 480 °C and 820 °C, with some shoulders at 300 °C and 650 °C. By comparison with previous results obtained on monocrystalline silicon [10], the 300 and 380 °C effusion features may be attributed to deuterium accumulated just beneath the surface where defects and traps due to plasma damage are in a high concentration, while the breaking of isolated Si-D bonds may be responsible for the effusion peak at 480 °C. Then the high temperature peaks at 630 and 820 °C would be due to the effusion of deuterium trapped in the buried oxide layer and/or at the interfaces Si-SiO$_2$.

A deuterated SIMOX sample, with the same deuteration conditions as above, was treated by potassium hydroxide in order to etch off the top silicon layer of the BOX structure, and was submitted to an effusion experiment. The results (Figure 4, curve c) show that actually the four effusion peaks at 300, 380, 480 and 630 °C may be attributed to deuterium coming from the top silicon layer and the top silicon / BOX interface. The peak at 820 °C (Figure 4, curve c) would be due then to the effusion of deuterium trapped in the oxide layer or at the second BOX / silicon interface. It must be pointed out that a continuous deuterium effusion flow is still observed between 400 °C and 630 °C for the etched sample.

The deuterium effusion spectrum of the ZMR sample after deuteration at 250 °C during 30 min, presented on figure 4, curve b, shows two main effusion peaks at respectively 380 and 520 °C, with a shoulder at 620 °C and a small effusion feature at 820 °C. The important peak at 620°C could be explained by the effusion of the deuterium trapped on the poly-Si/SiO$_2$ interface.

After etching, only a high temperature deuterium effusion peak at about 900 °C is detected.

3.3 Thermal annealing

The results of an isothermal annealing at 600°C during 2 hours performed on a deuterated SIMOX sample are presented in figure 5. It may be observed

that mainly deuterium contained in the silicon top layer has outdiffused after this thermal treatment, and that deuterium concentration in the BOX after isothermal annealing is of the same order as before annealing. The deuterium concentration profile in the oxide layer has become constant.
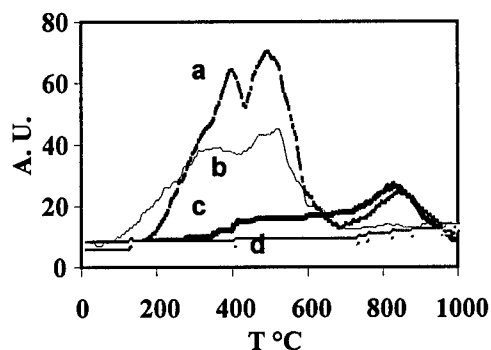
Figure 4. Deuterium effusion spectra from SIMOX sample (curve a), ZMR sample (curve b) after deuteration at 250 °C during 30 min; Deuterium effusion spectra from similar samples SIMOX (curve c) and ZMR (curve d) after etching.
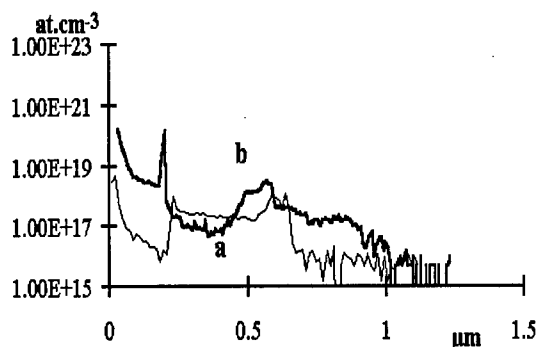
Figure 5 . Deuterium concentration profiles in SIMOX samples; (a) After deuteration at 250°C ($1W/cm^2$ during 30 min) ; (b) Sample (a) after an isothermal annealing at 600°C during 2 hours.

Figure 2, curve b, shows the effects of a 600 °C isothermal annealing on a ZMR sample deuterated with a RF plasma power of $1W/cm^2$ at 250°C during 30 min. As for the SIMOX sample, the deuterium is stable in the oxide layer, where it can still be detected after a thermal annealing at 600°C during 2 hours.

4. CONCLUSIONS

The concentration of deuterium in the buried oxide is governed by two parameters: the diffusivity in the upper silicon layer, which is different in monocrystalline and polycrystalline silicon, and trapping on defects. From this point of view, deuterium diffusion profile analysis is a tool to evidence defects in the different layers. In SIMOX samples, trapping offers evidence of reactive sites probably related to silicon inclusions or oxide protusions with different compositions. However, it must be pointed out that after the isothermal annealings, the deuterium concentration profile in the implanted oxide layer becomes constant.

In ZMR samples, the large deuterium peak at the poly-Si/SiO$_2$ interface shows an important deuterium trapping between the thermal oxide and the polycrystalline silicon layer.

Permeation through the implanted oxide layer occurs at temperatures higher than 250°C, while the deuterium does not diffuse in the silicon substrate through the thermal oxide, for the studied experimental parameters range.

The deuterium is stable in the oxide layers where it can still be detected after a thermal annealing at 600°C during 2 hours. Combining isothermal annealing and effusion results, it may be concluded that the detrapping of hydrogen from the BOX occurs above 800°C.

REFERENCES

1. J.-P. Colinge, *Silicon-on-Insulator Technology: materials to VLSI*, Kluwer Academic Publishers, Dordrecht (1991).

2. E. I. Givargizov, V. A. Loukin and A. B. Limanov in *Physical and Technical Problems of SOI Structures and Devices*, ed. by J.-P. Colinge, V. S. Lysenko and A. N. Nazarov (NATO ASI series, Kluwer Academic Publishers, Dordrecht, 1994) p.27.

3. A. Chari, P. de Mierry, A. Menikh and M. Aucouturier, Rev. Phys. Appl. **22**, 655 (1987).

4. D. Ballutaud, M. Aucouturier and F.Babonneau, Appl. Phys. Lett. **49**, 1622 (1986).

5. K. L. Brower, Appl. Phys. lett. **43**, 1111 (1983).

6. S.M. Myers, G.A. Brown, A.G. Revesz and H.L. Hughes, J.Appl.Phys. **73(5)** 2196 (1993).

7. I.P. Lisovskii, V.G. Litovchenko, G.P. Romanova, P.I. Didenko and E.G. Schmidt, Phys. Stat. Sol. (a) **142**, 107 (1994).

8. D. Ballutaud, P. de Mierry, J.-C. Pesant, R. Rizk, A. Boutry-Forveille and M. Aucouturier, Mat. Science Forum, vol. **83-87**, pp. 45-50 (1992).

9. D. Mathiot, Phys. Rev. B **40**, 5867 (1989).

10. R. Rizk, P. de Mierry, D. Ballutaud and M. Aucouturier, Phys. Rev. B **44**, 6141 (1991).

11. L. Lusson, P. Elkaim, A. Correia and D. Ballutaud, J. Phys. III (France) **5**, 1173 (1995).

12. N. H. Nickel, W. B. Jackson and J. Walker, Phys. Review B **53** (12) 7750 (1996).

ACKNOWLEDGMENTS

Reactions and diffusion during annealing-induced H^+ generation in SOI buried oxides

K. Vanheusden[a,†], D.M. Fleetwood[b,§], R.A.B. Devine[c], and W.L. Warren[b,§]

[a]US Air Force Research Laboratory, 3550 Aberdeen Ave., SE, Kirtland Air Force Base, New Mexico 87117-5776, USA

[b]Sandia National Laboratories, P. O. Box 5800, Albuquerque, New Mexico 87185, USA

[c]France Telecom/CNET, B. P. 98, Meylan Cedex, France

We report experimental results suggesting that mobile protons are generated at strained Si-O-Si bonds near the Si/SiO$_2$ interface during annealing in forming gas. Our data further suggest that the presence of the top Si layer plays a crucial role in the mobile H^+ generation process. Finally, we show that the diffusion of the reactive species (presumably H_2 or H^0) towards the H^+ generation sites occurs laterally along the buried oxide layer, and can be impeded significantly due to the presence of trapping sites in the buried oxide.

1. INTRODUCTION

In a number of recent studies the generation of mobile protons in the buried oxide of SOI materials and in thermal oxide buried underneath a poly-Si layer has been discussed [1-3]. The protons are stable and can be easily rearranged by applying an electric field. The details of the hydrogen reactions leading to the generation of the mobile H^+ are still under investigation. In a recent work [2] a dynamic equilibrium model was proposed:

$$H^0 + X \underset{r}{\overset{f}{\rightleftarrows}} H^+ + X + e^-, \qquad (1)$$
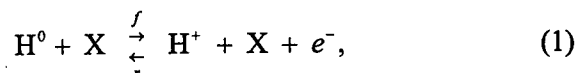
where the catalytic site X is a bridging oxygen at the Si/SiO$_2$ interface. The forward reaction dominates above ~ 500 °C and the resulting H^+ is mobile and entrapped inside the SiO$_2$. The electron is donated to the Si. The H^0 is likely to be formed through $H_2 + K \Leftrightarrow HK + H^0$, where K is a cracking site. In the same work it was shown that the reactive hydrogen species enter the oxide from the device edges [2] and that the amount of the reactive species reaching the oxide by diffusion through the Si overlayer is negligible. These results seem to contradict earlier studies [4,5] where it was shown that hydrogen can easily diffuse through the top Si layer under the given experimental conditions.

We present here new details about the reactions and the diffusion of the various hydrogen species during 500 - 700 °C forming gas annealing. These new findings offer valuable insight into the mechanism responsible for the generation of mobile H^+ at the Si/SiO$_2$ interface.

2. EXPERIMENTAL DETAILS

The Si/SiO$_2$/Si material used for this work was silicon-on-insulator (SOI), in particular standard dose SIMOX with a supplemental oxygen implant, and Unibond material [6], each with a SiO$_2$ layer of ~ 400 nm buried below a 200-nm c-Si layer. The specific capacitor structures used in this study are shown in Figs.1
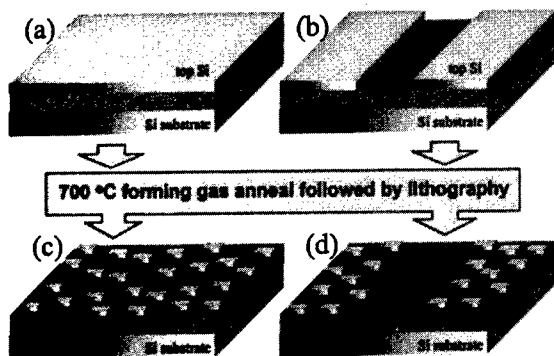
Figure 1. Schematic of SOI samples used in this work. All substrates measure 10 mm × 10 mm.

(a) and (b). These substrates were annealed for 600 s in a flow of forming-gas (N_2:H_2; 95:5) at 700 °C. After the anneal the original large capacitor structures (a) and (b) were transformed into a subset of smaller Si islands using selective etching, as shown in Figs. 1 (c) and (d). Flatband voltage shifts were obtained From pseudo-MOSFET curves measured on these small structures, which allowed us to probe the local concentration of mobile protons in the buried oxide generated during the anneal of the original structures, thus revealing lateral non-uniformities across the substrate.

In addition to the lateral distribution, the initial local distribution of the mobile H^+ over
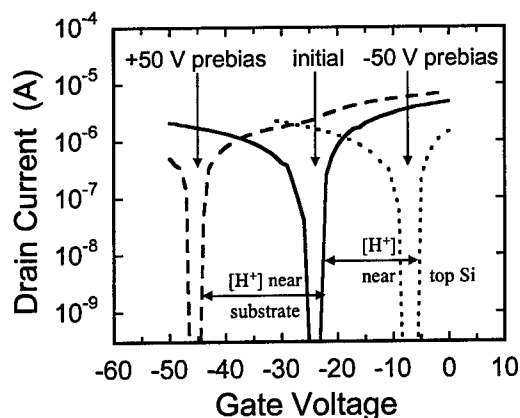
the top and bottom interface, i.e., the H^+ distribution present in the buried oxide right after the forming gas anneal is performed, was measured. This was done by first obtaining the flatband voltage shift on the small capacitors right after the anneal and lithography, without disturbing the initial charge distribution. Subsequently, the flatband voltage shift is measured after first applying a positive (+50 V) prebias to the substrate (gate) for 5 minutes. This will accumulate all the mobile positive species at the top Si interface, resulting in a maximum voltage shift. Finally a measurement after a negative (–50 V) prebias is performed, which will accumulate all the mobile protons at the substrate interface. Figure 2 shows how the initial H^+ distribution can be derived from this set of data.

We also measured the local density of both fixed and mobile charges present at the substrate interface of sample type (b). This was achieved by removing the remaining top Si strips after performing the forming gas anneal, followed by the deposition of a uniform matrix of small metal electrodes for capacitance-voltage characterization. This procedure allows us to probe the local oxide charge density across the entire substrate, and to compare the central area (oxide exposed to the forming-gas ambient) to the Si-covered areas.

3. RESULTS AND DISCUSSION

Figure 3 (a) shows the lateral distribution of the mobile protons in the buried oxide after annealing a large SIMOX substrate as shown in Fig. 1 (a). The mobile proton density is maximum near the edges and rapidly drops to zero towards the center of the capacitor· substrate. The local H^+ density is distributed evenly over the top and bottom interfaces, as can be seen in the lower part of Fig. 3 (a). Figures 3 (b) and (c) show the lateral distribution of the mobile protons in the buried oxide after annealing SIMOX and Unibond substrates like the one shown in Fig. 1 (b). Again, the mobile proton density is maximum near the edges of the substrate and rapidly drops when moving away from the edges. The

Figure 2.   Pseudo-MOSFET IV curves measured on a small Si island [see Fig. 1 (d)] on a Unibond substrate annealed in forming gas at 700 °C.
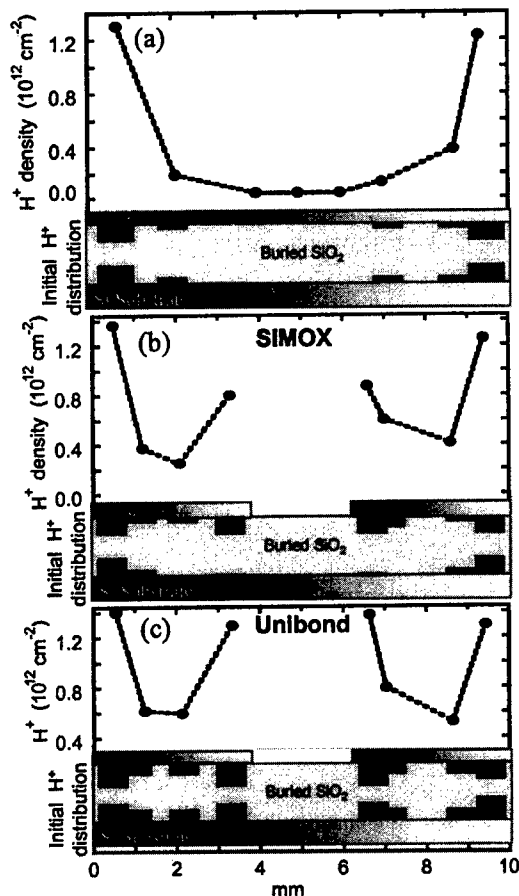
Figure 3. Lateral distribution of mobile proton density over the top and bottom Si/BOX interfaces after annealing in forming gas at 700 °C for 10 min.

local profile bar-charts show that at the outer substrate edges, equal densities of mobile protons ($\sim 6 \times 10^{11}$ cm^{-2}) are observed near both the top and the substrate interfaces for all three samples. However, at the inner edges of the two Si strips on the SIMOX substrate (Fig. 3 b), the total proton density only reaches ~ 65 % of the density that is observed at the outer substrate edges. From the profile bar-chart in Fig 3 (b) it can be seen that the inner edges of this sample show mobile proton densities near the top Si interface, similar to those at the outer edges. However, the proton densities are much smaller here near the substrate interface ($\sim 1 \times 10^{11}$ cm^{-2}). This surprising difference between

the inner and outer edges is not observed in Unibond material [see bar chart in Fig. 3 (c)].

These observations, in particular the differences between the inner and outer edges of the SIMOX substrates of type (b) and the absence of this difference in the Unibond sample allow us to draw a few interesting conclusions. The reactive hydrogen species (presumably H_2 or H^0) enter the SiO$_2$ directly from the ambient (not through the top Si) and then diffuse to the reactive sites at the Si/SiO$_2$ interface where they react to form H^+. For the anneal times studied in this work, diffusion of the reactive species through the BOX is the rate-limiting step for this reaction. This implies that in the initial stages of the forming gas anneal treatment, the H^+ reaction will only occur at those parts of the SiO$_2$/Si interfaces that are closest to the ambient. Hence, in our particular samples, the reaction will first occur at the top and bottom interface right at the outer edges of the substrates, and at the top interface right at the inner edges. After the H^+ is generated at these reaction sites, it will diffuse laterally along the interface (accompanied by its image charge in the Si) away from the edges. The H^+ reaction at the substrate interface below the inner edges of the top Si layer will be retarded as the reactive hydrogen has to diffuse across the buried oxide layer. This particular effect can explain the observations in Fig. 3 (b). The fact that Unibond material does not show this effect of retarded H^+ buildup for the same anneal time suggests a longer effective diffusion length of the reactive species in Unibond buried oxide as compared to SIMOX buried oxide. This may be due to a higher density of empty hydrogen traps [4] in SIMOX buried oxide as compared to Unibond. The very different growth conditions of both buried oxides (thermally grown vs. implanted) and the differences in processing between SIMOX and Unibond material, in particular the hydrogen implantation step used in the Unibond process [6], may account for this difference in hydrogen content.

Figure 4 shows the effect of 30-minute forming gas anneal at 600 °C for a SIMOX substrate with the top Si present during the
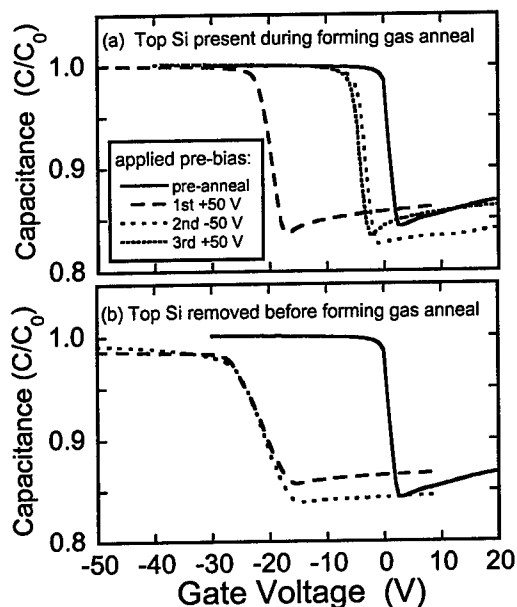
Figure 4. Sequence of CV data before and after forming gas annealing of both exposed and Si-covered buried oxides.

forming gas anneal, and for a substrate with the top Si layer removed prior to the forming gas anneal. After the anneal a +50 V bias is applied to the metal gate for 5 min, followed by a CV scan. Subsequently, this procedure is repeated with a −50 V prebias, and finally again with a +50 V prebias. The data in Fig. 4 (a) show that annealing the closed $Si/SiO_2/Si$ structure introduces positive charge in the buried oxide, most of which is mobile. The small fraction of fixed positive charge observed in this sample was found to be located near the substrate interface. The data further show that after sweeping the mobile charge to the gate it becomes trapped near, or gets absorbed by, the metal gate electrode. The results in Fig. 4 (b) show that a forming-gas anneal performed on the open Si/SiO_2 structure results in about the same amount of positive charge in the oxide. However, unlike in the closed structure, all the positive charge is fixed near the substrate interface. The interesting conclusion from this comparison is that the presence (or absence) of the top Si layer has a drastic impact on the nature of the positive charge being generated at

the substrate interface. We propose that stress at the substrate interface is the key to explain the observed difference. We propose that the generation of fixed vs. mobile protons is related to the level of stress present at the buried oxide interfaces. It is likely that removing the top Si layer and subsequently performing a 600 °C anneal can reduce this stress at the substrate interface. This may then cause trapping of the generated protons near the interface. Quantum chemical calculations are under way relating this stress to the local H^+ binding energy.

4. CONCLUSIONS

We have shown that the proton generation reaction in buried SiO_2 is limited by the rate of diffusion of the reactive species (presumably hydrogen species) through the SiO_2 film. This diffusion can be impeded significantly by the presence of trapping sites inside the buried oxide layer. It is also shown that the presence / absence of the top-Si layer during forming gas annealing results in mobile / fixed protons at the substrate interface. This may be due to a change in Si/SiO_2 interface stress.

REFERENCES

[1] K. Vanheusden, W.L. Warren, D.M. Fleetwood, J.R. Schwank, M.R. Shaneyfelt, P.S. Winokur, R.A.B. Devine, and Z.J. Lemnios, Nature 386, 587 (1997).

[2] K. Vanheusden et al., Appl. Phys. Lett. 73, 674 (1998).

[3] R.E. Stahlbush, R.K. Lawrence, and H.L. Hughes, IEEE Trans. Nuc. Sci. 45, 2398 (1998).

[4] S.M. Myers, G.A. Brown, A.G. Revesz, and H.L. Hughes, J. Appl. Phys 73, 2196 (1993).

[5] H.C. Mogul, L. Cong, R.M. Wallace, P.J. Chen, T.A. Rost, K. Harvey, Appl. Phys. Lett. 72, 1721 (1998).

[6] A.J. Auberton-Herve, T. Barge, F. Metral, M. Bruel, B. Aspar, C. Maleville, H. Moriceau, T. Poumeyrol, in Amorphous and Crystalline Insulating Thin Films, Vol. 446, (Materials Research Society, Pittsburgh, PA, 1997), p. 177.

Direct Sub-μm Lateral Patterning of SOI by Focused Laser Beam Induced Oxidation *

R. A. Deutschmann†, M. Huber, R. Neumann, K. Brunner and G. Abstreiter

Walter Schottky Institut, Technische Universität München, Am Coulombwall, 85748 Garching, Germany

We have developed a method for resistless patterning of Silicon-On-Insulator (SOI) in order to directly achieve lateral dielectric isolation. Our method employs a diffraction limited laser spot of a cw argon laser which is scanned across the SOI surface and directly oxidizes the silicon in ambient atmosphere. We investigate the dependence of the line width on the laser power and laser wavelength and scanning speed by AFM measurements. Line widths as narrow as 200 nm are achieved. We have fabricated in-plane-gate transistors with effective channel widths of 250 nm and excellent gate to source-drain isolation. We further demonstrate how our method can be extended for local simultaneous oxidation and doping of the silicon surface with high spatial resolution. A novel in-plane-FET device has thus been realized.

1. INTRODUCTION

Silicon-on-insulator (SOI) CMOS has several advantages as compared to conventional CMOS, amongst which are the dielectric isolation of the devices, thus preventing latch-up, and a substantial reduction of parasitic capacitance, communication delays and power consumption due to the presence of only vertical junctions. Further advantages are radiation hardness and an improved transconductance and subthreshold slope.

Besides the standard methods for lateral patterning of SOI by means of optical lithography, other methods have been developed for special purposes. Electron beam lithography and subsequent etching is often employed for nanoscale patterning [1]. Crell et.al. have fabricated in-plane-gate transistors on SOI by focused ion beam implantation of Ga+ ions [2]. AFM surface oxidation of SOI was employed by Campbell et.al. to produce a mask for a subsequent etching step [3].

In our work we were driven by the idea to make use of the non linear dependence of the silicon oxidation process on the temperature. Previous work by Micheli and Boyd has shown that oxidation of bulk silicon by a cw argon laser is easily achieved [4]. Whereas in their work no special emphasis on lateral spatial resolution was made, here we aim for the smallest possible line width of the silicon oxide.

2. METHOD AND MATERIAL

We use a cw argon laser at a wavelength $\lambda = 458$ nm (514 nm) which we focus through a high numerical aperture lens system to achieve a diffraction limited laser spot of 315 nm (350 nm) FWHM. A laser power controller stabilizes the laser intensity to a precision of 0.05%. Values given in the paper relate to laser powers on the sample surface. An autofocus system is to ensure a constant spot size during the writing process. The sample is moved computer controlled by a X/Y piezo table under the laser spot, scanning speeds up to 1.7 μm/sec can be achieved. Sample treatment occurs at room temperature and ambient atmosphere.

We have used several different SOI substrates for laser oxidation.

- SIMOX with 370 nm buried oxide (BOX) and nominally undoped 15 nm silicon layer

- BESOI with 195 nm buried oxide and nominally undoped 20 nm silicon layer

- Both SIMOX and BESOI where the silicon
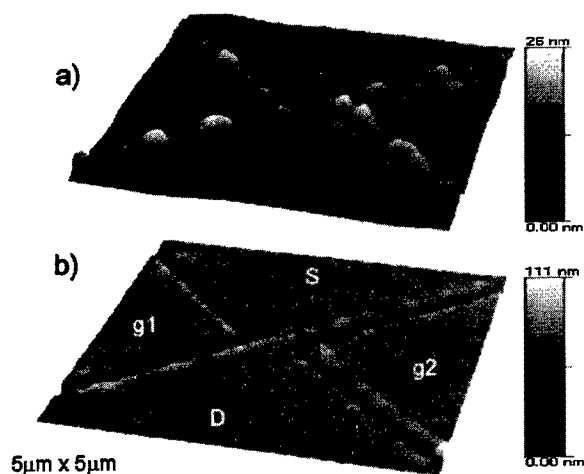
Figure 1. AFM pictures of an in-plane-gate transistor with source S, drain D and gates g1, g2 (a) before and (b) after etching the laser written oxide in diluted HF. The geometrical channel width is 600 nm, the effective channel width is smaller due to the finite line width.

Figure 2. Measured line widths for different laser powers and for laser wavelengths of 458 nm and 514 nm obtained on a 17 nm silicon layer.

layer was doped up to $2 \times 10^{18} \text{cm}^{-3}$.

- Both SIMOX and BESOI which were epitaxially overgrown with up to 100 nm Si or Si/Ge including different bulk or delta doping densities.

3. EXPERIMENTAL RESULTS

In this section first we discuss the surface morphology obtained by focused laser beam oxidation of the SOI surface. Second we characterize the electrical properties of oxidized lines and present two different in-plane-gate transistors.

3.1. Laser oxidation of SOI surface

An AFM picture of two adjacent V shaped lines written using $\lambda = 514$ nm and a laser power of 40 mW into a 20 nm SOI surface is shown in Figure 1a. The silicon surface along the trace of the laser beam appears elevated by about 20 nm in a hill shaped, irregular fashion with a width of about 1 μm. Additionally a region of about 1 μm to both sides of the trace is elevated homogeneously by a few nm. This surface morpho-

logy is common to both wavelengths. Fast scans (1.7 μm/sec) can result in a regular chain of repeated elevations and valleys along the trace with a period of 250 nm.

The hill shaped elevations appear to be thermally oxidized silicon, although there might be a contribution by non thermal, photonic oxidation [4]. The shallow elevations to both sides of the laser trace may be due to the strain induced by the oxidized line within the non oxidized silicon and/or due to debris.

After a 1 min wet etch in diluted HF all oxide is removed, see Figure 1b. The silicon surface next to the laser written line is flat and the line itself becomes apparent as a narrow, rather well defined channel. We find that oxidation is an all-or-nothing process, i.e. when oxidation takes place then the entire silicon layer is affected. Of utmost interest for sub-μm scale patterning are the line widths that can be obtained, see Figure 2. We find that the best lines that are reliably oxidized are as narrow as 200 nm, which is well below the laser spot diameter. Optimal results are obtained using $\lambda = 458$ nm and a laser power just above the threshold power. The threshold power sensitively depends on the silicon layer thickness, but not on the BOX thickness. As an example for $\lambda = 458$ nm and a silicon layer thick-
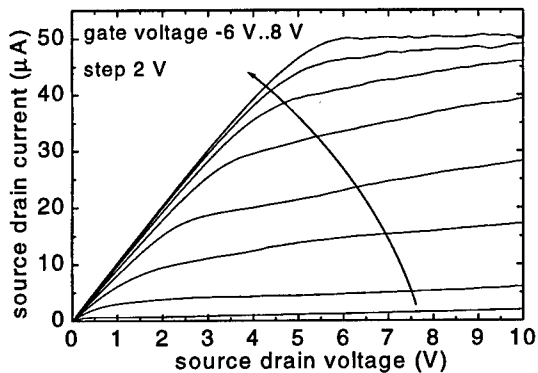
Figure 3. IPG current-voltage relation for different in-plane-gate voltages.

ness of 37 nm we determined a threshold power of 25 mW, whereas for a silicon layer thickness of 100 nm only 10 mW were necessary. The threshold powers for $\lambda = 458$ nm are generally smaller than for $\lambda = 514$ nm which is due to a larger absorption coefficient in conjunction with a smaller laser spot diameter.

From the error bars in Figure 2 it can be seen that the fluctuation of the line width is generally larger close to the threshold power, where the oxidation process is less stable. Smoother lines are achieved for larger laser powers. Line width fluctuations of ±10% can then be achieved.

Further it is interesting to note that no dependence of the line width on the scan speed is found within the confidence level. We explain this observation by the stabilization of the line width by an effective self limiting mechanism. This mechanism is caused by an abrupt drop of absorption when silicon is oxidized.

3.2. In-plane-gate transistor fabricated by focused laser beam oxidation

We have fabricated in-plane-gate (IPG) transistors [2][7] in the geometry of Figure 1 in order to test the dielectric properties of the oxidized lines and to gain information about the silicon-silicon oxide interface. Leakage currents at 100 V across a 15 μm long line are in the pA range, and

electrical breakthrough at higher voltages in general occurred across the BOX first. After etching the oxide to reduce the effect of charge trapping in the side walls [6] and possibly of dangling bonds we have obtained transistor action on all substrates except for highly doped samples. A Hydrogen plasma anneal for 60 min at 150 °C also improved the transconductance. In Figure 3 we show the current-voltage curve of an IPG with a geometrical channel width of 600 nm fabricated on undoped SOI. The rather high saturation voltage is due to the bad quality of the BOX on its upper interface where an electron inversion channel was induced by the back gate. The electron density was 4×10^{12} cm^{-2} at a back gate voltage of 50 V. Hole channels showed a similar behavior. The narrowest IPGs we have been able to reliably fabricate had a geometrical channel width of 450 nm, which leaves an effective channel width of 250 nm assuming a line width of 200 nm. The fact that in even narrower IPGs the source drain channel was closed may be due to a) enhancement of the line width in the vicinity of another line due to a reduced thermal conductivity and b) degradation of the silicon directly adjacent to the silicon oxide. We note that the IPGs showed a slight hysteresis when sweeping the gate voltage which might be due to carrier trapping in the silicon-silicon oxide interface.

3.3. IPFET by combined laser beam oxidation and codoping

We have extended our work to achieve not only local oxidation but also local doping of the silicon surface. For that purpose we deposit by spin-on technique a silicon dioxide film containing appropriate doping atoms onto the silicon surface. Thus when heating the surface by the focused laser beam, doping atoms are diffused into the silicon layer [7]. We observe that undoped SOI becomes conductive along a laser oxidized line even when the line is interrupted by gaps up to 500 nm in length. This means that oxidation is accompanied by codoping in close vicinity with an extent of about 150 nm. In other words electrically isolating lines can be generated while simultaneously a very confined region of only 150 nm to both sides is doped. The doping concentration
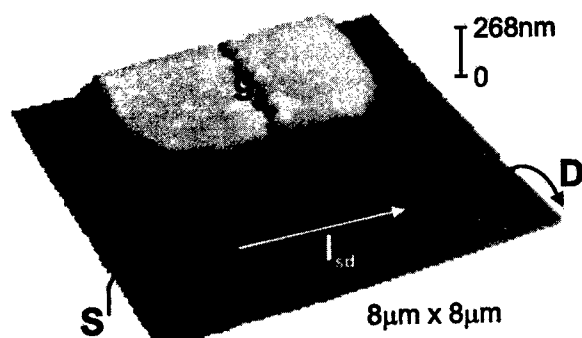
Figure 4. AFM picture of the central part of an IPFET. The gate contact (g) metalization and two orthogonal oxidized lines are visible. Alongside the longer line the source drain current I_{sd} is passed, whereas the other side is tied to the gate potential.

Figure 5. IPFET current-voltage relation for side-gate voltages from 0 to 5 V.

and conductivity increase with decreasing writing speed.

We have fabricated a novel 'In-Plane Field Effect Transistor' (IPFET) by making use of this technique. An AFM picture of part of the device is shown in Figure 4. The source drain current is conducted along one side of the oxidized line, whereas to the other side the gate voltage is applied. We have fabricated normally ON and normally OFF IPFETs by appropriately choosing the laser scan speed. The current voltage relation of a normally OFF IPFET is shown in Figure 5. At zero gate voltage no source drain current is observed, whereas at a gate voltage of 1.5 V the IPFET turns on. Saturation can be observed even though the source drain current is passed in close vicinity to the laser induced oxide.

4. CONCLUSION

In summary we have developed a novel method which allows the direct resistless patterning of thin film SOI. The method employs a focused cw laser beam which locally oxidizes the silicon layer. Line widths as low as 200 nm can reliably be

generated by appropriately choosing wavelength, laser power and scanning speed. Electrical isolation across oxidized lines is excellent. We have demonstrated the method by the fabrication of functional IPG transistors with geometrical channel widths as narrow as 450 nm. Additionally we have extended our work to local oxidation with codopig. A novel in-plane field effect transistor has thus been fabricated. The quality of the laser induced oxide may be improved further by oxidation in pure dry oxygen.

REFERENCES

1. E. Leobandung, L. Guo, Y. Wang and S. Chou. Appl. Phys. Lett. 67, 938 (1995)
2. C. Crell, K. Wieczorek, H.-U. Schreiber, A. D. Wieck. Appl. Phys. Lett. 68, 2538 (1996)
3. P. M. Campbell, E. S. Snow, P. J. McMarr. Appl. Phys. Lett. 66, 1388 (1995)
4. F. Micheli and I. W. Boyd. Appl. Phys. Lett. 15, 1149 (1987)
5. G. J. Willems, J. J. Poortmans and H. E. Maes. J. Appl. Phys. 65, 3408 (1987)
6. P. C. Karulkar in: Silicon-On-Insulator Technology and Devices, Ed. S. Cristoloveanu, The Electrochemical Society 1994, page 209
7. P. Baumgartner et.al. Appl. Phys. Lett. 64, 592 (1994)

Sub-band-gap impact ionization events in transient regimes of floating body SOI devices

A.M. Ionescu[1][2] and A. Chovet[3]

[1]LEG-EPFL, Lausanne, CH-1015, Switzerland
Phone: +41- 21 693 3978, Fax: +41 – 21 693 36 40, E-mail: Adrian.Ionescu@epfl.ch

[3]LPCS, ENSERG, 23, rue des Martyrs, BP 257, 38016 Grenoble Cedex, France

The aim of this paper is to demonstrate via experiments and 2-D numerical simulation that in transient regimes of partially depleted (PD) SOI MOSFETs the sub-band-gap impact ionization becomes more critical compared to the static regime, and can stand for sub-1.1V=E_{GSi}/q drain voltages, in both short and long channels.

1. INTRODUCTION

Most of the conventional Si-bulk models fail to accurately account for SOI specific effects related to the thin silicon film and buried insulator. Thus, specific floating body and impact ionization modeling and simulation are essential to predict SOI MOSFET features in both digital and analog ICs. Superficial similarity with bulk MOSFET supports naive expectations such as in SOI MOSFET the impact ionization can not occur for drain-to-source voltages corresponding to lower energy than Si band gap.

In contrast, recent experiments [1] showed non-conventional kink effects at voltages as low as 0.8V in submicron SOI transistors. Other recent experimental data [2] confirmed this mechanism in PD-SOI MOSFETs as well as our previous experiments and simulations in some special transient regimes [3] which revealed unexpected impact ionization for sub-1V drain-to-source voltages. This new impact ionization phenomenon is called *sub-band-gap impact ionization* and explained by dynamically screened electron-electron scattering. It is worth noting that sub-band-gap impact ionization can be accentuated by the floating body charging in PD-SOI devices [1] and it is a *threshold-type event*. Because carriers created by impact ionization can cause hot electron reliability problems, the study of this kind of events is important.

2. EXPERIMENT

2.1 Static experiments

In typical static experiments, the impact ionization is revealed by the bell-shaped body current, Fig. 1. For long channel devices, no bell shaped I_B versus V_G are observed for V_D<1.1V=E_{GSi}/q. In contrast, for submicron devices, the bell-shaped curve can appear at much lower V_D voltages due to sub-band-gap impact ionization events [1, 2], Fig. 2, explained by lucky electron theory and Coulomb interactions.
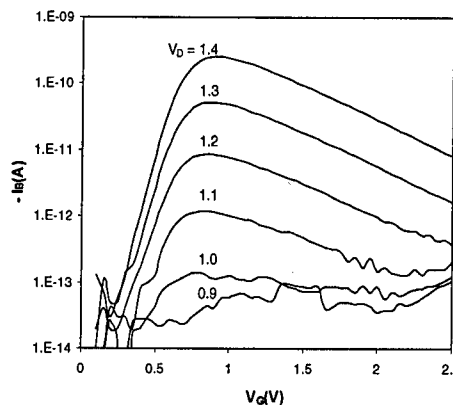
Fig. 1 Typical bell-shaped body current, I_B, in a short channel (L=0.25 μm) SOI n-MOSFET with the drain voltage V_D as a parameter and V_{G2}=0V.

---

[2] On leave from: Center for Integrated Systems, Stanford University, Stanford CA 94305-4070, USA
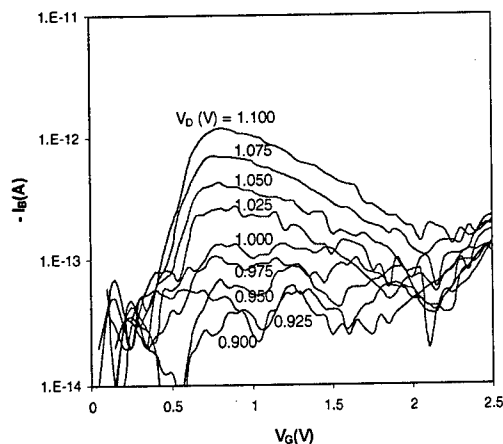
Fig. 2 Same type of measurement as in Fig. 1, highlighting sub-band-gap impact ionization; bell-shaped curves for $V_D < 1.1V$.

2.2 Dynamic experiments

In the adapted-to-SOI Zerbst transient technique, the front gate of the SOI MOSFET is biased in strong inversion while the back gate is switched from depletion to accumulation and a low voltage is applied to the drain. The induced current transient mirrors generation phenomena in the Si-film [4], Figs. 3-7, and is generally used for carrier lifetime extraction [4].

We have fully modeled and simulated drain and source junction contributions to the generation process compared to the bulk one [3], canceling confusion when obtaining shorter relaxation for shorter gate length.

In the following, experiments on both long and short channel PD SOI n-MOSFETs are presented. The SOI substrates are standard: SIMOX or UNIBOND, with Si film and buried oxide thicknesses of 150 nm and 380 nm, respectively. The front oxide thickness is 22 nm in case of long channel devices and 5 nm in case of submicron devices.

Figs. 3 and 4 show that when in a typical experiment, with a small V_{G2} step-voltage, the V_D value is increased, the transient duration is reduced due to carrier supply by conventional impact ionization starting at a *threshold* value of V_D. At low V_{G2} steps, this value is systematically greater than 1.1V for long channel SOI n-MOSFET (Fig. 4). In contrast, a strong transient reduction, is

observed in short channel SOI MOSFETs at V_D close to the lower value for which the I_B bell-shaped peak is observed ($V_D=0.975V$ in Fig. 3). This corresponds to a dynamically induced sub-band-gap impact ionization event.

In a complementary experiment (see Figs. 5 and 6) V_D is kept low and constant, while V_{G2} is successively increased in small steps. A low V_D ensures linear operation for a simpler analysis and is assumed to avoid impact ionization at equilibrium (static regime).

The first observed effect, when V_{G2} amplitude is increased, is a near off-current region, corresponding to a large extension of the space charge region under the front gate. In this case, the duration of current relaxation is essentially the same.

Secondly, an unexpected substantial reduction of transient duration is observed starting with a V_{G2} *threshold value*, which correspond to the on-set of sub-band gap impact ionization (which is a *threshold-type* event). It is worth noting that this phenomenon is *observed in both long and short channels* (Figs. 5 and 6) and has to be related to the strong drop of the floating body potential (after applying a V_{G2} step-voltage) and increased induced electric fields. Note that it was experimentally observed that in ultra-short channels, where longitudinal fields are high, the transient reduction is obtained for lower V_{G2} steps, compared to longer channels.
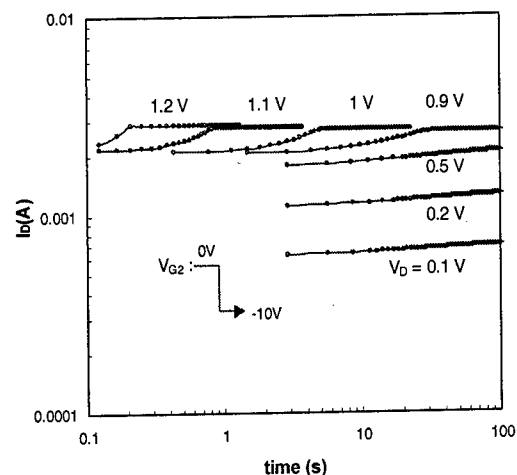
Fig. 3 Drain current transients in a submicron (L=0.25 μm) PD SOI MOSFET at various V_D, induced by a small V_{G2} step-voltage.

In fact, after applying a strong back gate step-voltage, the electrical field distribution is strongly modified. Temporarily increased vertical and lateral electrical fields make possible electron-electron interaction and consequently, impact ionization events even if V_D is kept low.

The impact ionization hypothesis is reinforced by the experimental higher peaks of channel charge time-rate (2-3 orders of magnitude), observed when V_{G2} is substantially increased combined with a low junction leakage (off current), Fig. 7.

A key point of the Zerbst-type transient regime in PD floating-body devices is that it becomes a unique detector of low level ionization events (a 50fA limit has been reported in [2]), generally impossible to be highlighted in any conventional bulk MOSFETs.

Similar transient effects can be obtained by reversing the role of front and back gates or by switching the front gate from *on* to *off* state (closer to real working transistor operation).

Fig. 4 Drain current transients in a long channel (L=1.4μm) PD SOI n-MOSFET at various V_D.
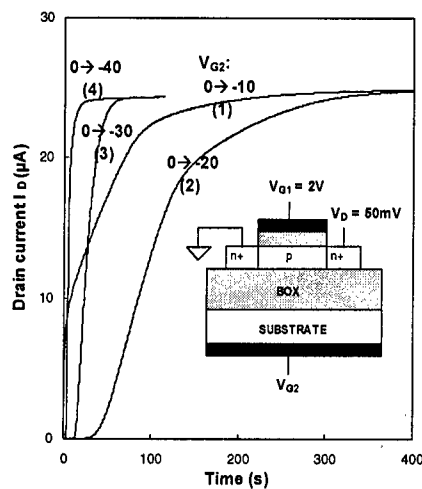
Fig. 6 Drain current transients at low V_D (=50mV) as a function of the back gate step-voltage amplitude in a long channel (L=1.4μm) PD SOI MOSFET. Transient duration is strongly reduced starting with V_{G2}: 0 → –30V.

Fig. 5 Drain current transients in a submicron (L=0.25 μm) PD SOI n-MOSFET at various V_{G2}.
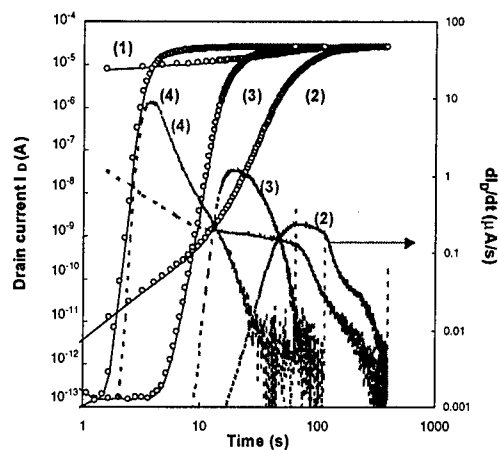
Fig. 7 Log scale of the drain current transient given in Fig. 6 and calculated current rate (peaks). Drain current values in the off-region (time < 10s), after high pulse application, are revealed.

3. 2-D NUMERICAL SIMULATION

2D numerical simulation[#] of Zerbst-type transient regimes has been performed on both long and submicron channel devices in order to reveal the field 2D distribution and impact ionization rates. At very low V_D and/or long channels, sub-band-gap impact ionization events appears to *exist only at the beginning* of the transient, when the body potential is very low. Field and potential time relaxation together with carrier recombination usually inhibits rapidly sub-band-gap impact ionization events. The 2D-simulation of current transient regime allowed us to demonstrate that, Figs. 8-10:

(i) the current transient duration reduction with V_{G2} amplitude, at low V_D, can be reproduced only if the impact ionization models are *enabled* in simulation (Selberherr ionization model),
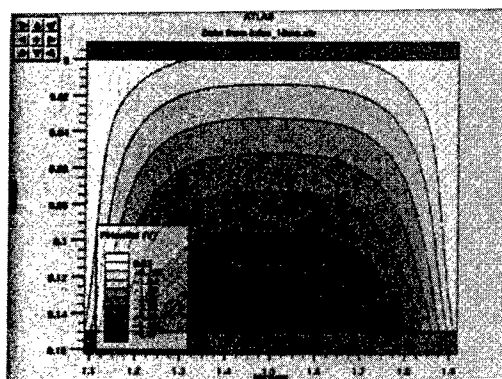
Fig. 8 2D-simulated potential drop after a V_{G2}-step voltage: 0 → –50V, at low V_D and time=1ms.
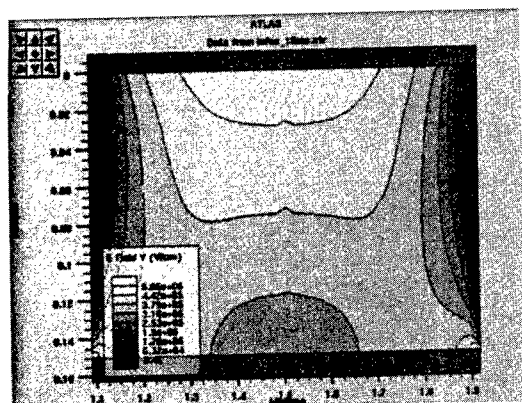
Fig. 9 2D-simulated vertical electrical field corresponding to conditions in Fig. 8, showing a strong electric field under the front gate.

(ii) the potential of the floating body drops severely (few volts for high V_{G2}) at time <100ms and strong electric field peaks (near 7×10^5V/cm) are induced near the Si/SiO_2 interface, which demonstrate enhanced conditions for sub-band-gap impact ionization as suggested by experiments.

Finally, numerically simulated 2D-charts of recombination/generation and impact ionization rates confirmed also the existence of sub-band gap impact ionization, Fig. 10. Moreover, accordingly to experiments, transient simulations also showed that this phenomenon is more accentuated for shorter channels (<0.5μm).
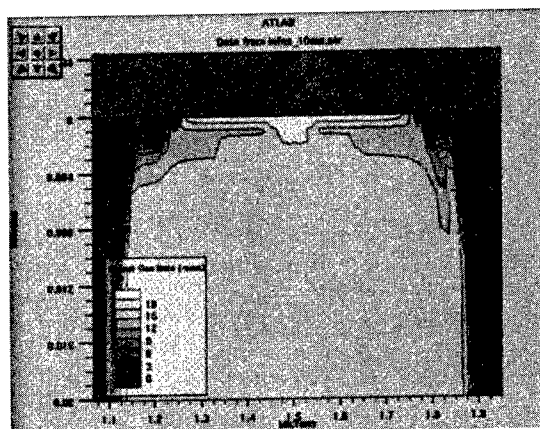
Fig. 10 Simulated impact ionization 2D-chart corresponding to Figs. 8 and 9 showing high ionization rate under the front gate at time=1ms.

4. CONCLUSION

Sub-band-gap impact ionization in Zerbst-type transient experiments using high back gate step voltages, at drain voltages lower than 1.1V, in both long and submicron channel PD SOI MOSFETs, has been demonstrated. These events, producing hot electrons at low drain bias, appear to be a reliability concern for future VLSI SOI ICs in similar working conditions.

REFERENCES
1. M.V. Fischetti and S.E. Laux, IEDM Tech. Dig. (1995) 305.
2. T. Saraya, M. Takamiya, T.N. Duyet and T. Hiramoto, Jpn. J. Appl. Phys, 37 (1998) 1271.
3. A. M. Ionescu, A. Chovet and F. Chaudier, MRS Proc., 490 (1998) 261.
4. D.K. Schroder, Semiconductor Material and Device Characterization, 2nd Ed. Wiley, 1998

Theory of the MOS/SOI Tunnel Diode

M. Badri and B. Majkusiak

Institute of Microelectronics and Optoelectronics, Warsaw University of Technology
Koszykowa 75, 00-662 Warsaw, Poland

Theoretical model of the MOS/SOI tunnel diode is presented. Current-voltage characteristics of the MOS/SOI tunnel diodes with both the p-type and n-type semiconductor films are considered versus the front-gate voltage, back-gate voltage, front-gate oxide thickness, and the semiconductor film thickness. An influence of the neighborhood of the under-gate region is also included in a quasi-two-dimensional model.

1. INTRODUCTION

The silicon-on-insulator (SOI) technology has already reached its maturity and becomes to compete with the bulk MOS technology. The SOI technology potentials are attractive for their tolerance to the scaling rules requirements and soon, according to them, the semiconductor film thickness may reach the range of nanometers. An extremely thin oxide can conduct a considerable tunnel current, which may constitute a parasitic leakage for an MOS/SOI transistor but also may be utilized in MOS/SOI tunnel devices. Many studies concerning tunnel currents in bulk MOS tunnel diodes and transistors can be found in the literature [e.g. 1-4]. On contrary, up to the best of our knowledge, the MOS/SOI tunnel diode has not received any attention in the literature until now. This may be due to seemingly complex electrical behavior of the MOS/SOI system and, to some extent, novelty of SOI structures. One may expect that the tunnel current flowing through the MOS/SOI system can be easily modeled by implementation of the same theoretical models as used for the bulk MOS tunnel diodes. It may be the case provided the semiconductor film is thick enough and the front-gate induced space charge region does not spread on the whole semiconductor region. However, the silicon film thickness can be of the range of nanometers and it must be taken into account in modeling the tunnel currents. It is the purpose of this work to develop a theoretical model of the MOS/SOI tunnel diode and to consider its current - voltage characteristics in dependence on the front-gate voltage, back-gate voltage, front-gate oxide thickness, the semiconductor film thickness, and some other related parameters.

2. THEORETICAL MODEL

2.1. Model assumptions

Fig. 1 shows a schematic diagram of the MOS/SOI tunnel diode with a p-type silicon film.
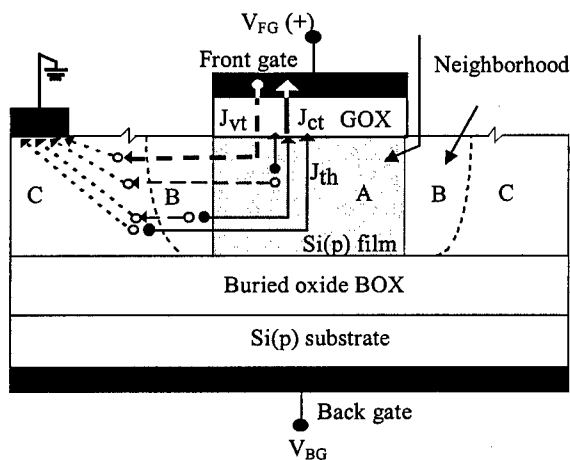
Figure 1.Schematic diagram of the p-type MOS/SOI tunnel diode with the positively biased front gate.

The MOS/SOI tunnel diode is constituted by the following system: front gate electrode - gate oxide - semiconductor film with an ohmic contact. It is assumed that the electrical contact to the silicon film is far enough from the gate area that it does not influence the physical state of the semiconductor under-gate region. The tunnel current through the front gate oxide consists of the gate-semiconductor valence band tunnel current (hole tunnel current J_{vt}) and the gate-semiconductor conduction band tunnel current (electron tunnel current J_{ct}). Due to the flow of these currents the semiconductor region under the gate is in the thermal non-equilibrium state. It is

assumed that the exchange of majority carriers between the under-gate region and the silicon film contact is so effective that the majority carrier quasi-Fermi level is practically flat. The minority carrier quasi-Fermi level changes with a position in the semiconductor and results from a balance between the drift-diffusion transport mechanisms, thermal generation-recombination rates in the semiconductor, and tunneling process through the front gate oxide.

Fig. 1 includes also a contribution of minority carriers coming from the neighboring regions (B and C) of the under-gate area (A) to the whole minority carrier current at the semiconductor - oxide interface J_{th}. This contribution is a source of the peripheral effect if a finite diameter of the front gate electrode is taken into account. In order to isolate this parasitic effect the basic model assumes that minority carriers are generated or recombined only inside the under-gate region. This is equivalent to the assumption that the gate electrode diameter is infinitely large and the minority carrier quasi-Fermi level is practically flat in the lateral direction. Moreover, to investigate the peripheral effect resulting from the influence of the neighborhood a quasi-two-dimensional model is also developed, in which minority carriers coming from the neighborhood are taken into account.

2.2. Calculation procedure

The main condition which must be fulfilled by the computation algorithm is to maintain the continuity of minority carrier current at the gate oxide - semiconductor film interface.

The semiconductor film region of the diode is determined by the following set of equations (assuming the p-type semiconductor):

a) Poisson's equation:

$$\frac{d^2\varphi}{dx^2} = \frac{q}{\varepsilon_s}\left[(n-p)-N_d+N_a\right] \qquad (1)$$

b) The minority carrier current continuity equation in the vertical direction across the structure:

$$\frac{1}{q}\frac{dJ_n}{dx} = U_{th} \qquad (2)$$

c) The minority carrier current:

$$J_n = \mu_n(x)n(x)\frac{dE_{Fn}}{dx} \qquad (3)$$

where E_{Fn} is the electron quasi-Fermi level.

d) Shockley-Read-Hall formula for thermal recombination-generation rate (for $E_{trap}=E_i$, $\tau_{n0}=\tau_{p0}=10^{-5}$s) :

$$U_{th} = \frac{np - n_i^2}{\tau_{p0}\left(n+n_i\right)+\tau_{n0}\left(p+n_i\right)} \qquad (4)$$

The electron J_{ct} and hole J_{vt} tunnel currents are calculated by integration of the product of supply function $N(E_x)$ and tunneling probability $P(E_x)$ over the energy E_x associated with the x-direction:

$$J_{ct} = q\int_{E_c(0)}^{E_{xmax}} P(E_x)N(E_x)dE_x \qquad (5)$$

$$J_{vt} = q\int_{E_v(0)}^{E'_{xmax}} P(E_x')N'(E_x')dE_x' \qquad (6)$$

where $N(E_x)$ and $N'(E_x')$ are given by (7) or (8), respectively:

$$\frac{4\pi m_{ci}}{h^3}\left\{\ln\left[1+\exp\left(\frac{E_{Fn}-E_x}{kT}\right)\right]-\ln\left[1+\exp\left(\frac{E_{Fm}-E_x}{kT}\right)\right]\right\} \qquad (7)$$

$$\frac{4\pi m_{vi}}{h^3}\left\{\ln\left[1+\exp\left(\frac{E_x-E_{Fp}}{kT}\right)\right]-\ln\left[1+\exp\left(\frac{E_x-E_{Fm}}{kT}\right)\right]\right\} \qquad (8)$$

m_{ci} and m_{vi} are the electron and hole effective masses in the oxide, respectively. The tunneling probability is calculated with the use of WKB approximation and the two-band barrier model assuming $m_{ci}=m_{vi}=0.6365m_0$ and $E_{gi}=8.64$ eV):

$$P(E_x) = \exp\left[-2\int_{x1}^{x2}\eta(x)dx\right] \qquad (9)$$

where $\eta(x)$ is the imaginary wave vector [4].

The effect of the neighboring regions (B and C) on the current flowing through the MOS-SOI tunnel diode biased in the range of depletion in the under-gate semiconductor region results from a contribution of minority carriers generated or recombined in these regions to the whole minority carrier current J_{th}. Region B has a depletion width, which varies as function of depth into the semiconductor film as a result of potential distribution in the vertical direction. The contribution of minority carriers coming from the quasi-neutral region C is calculated by integrating the current contributions of small elements of

thickness dx over the whole semiconductor thickness:

$$\Delta J_{diff}(x) = \frac{qD_n n_i^2}{N_a L_n}\left[e^{\left(q\varsigma(x)/kT\right)}-1\right]\frac{2\pi\left(R+X_d(x)\right)}{\pi R^2}dx \quad (10)$$

where D_n and L_n are the diffusion constant and diffusion length of electrons in the semiconductor, respectively. ς is the quasi-Fermi level split, R is the gate electrode radius and $X_d(x)$ is the depletion width. The contribution of minority carriers coming from the depletion region B is calculated by integrating the elementary current contributions:

$$\Delta J_{dep}(x) = \frac{qn_i}{2\tau_0}\left[e^{\left(q\varsigma(x)/2kT\right)}-1\right]\frac{\pi\left[\left(R+X_d(x)\right)^2-R^2\right]}{\pi R^2}dx \quad (11)$$

where τ_0 is the effective minority carrier lifetime.

3. RESULTS AND DISSCUSION

Fig. 2 shows current-voltage characteristics of the p-type (a) and n-type (b) MOS/SOI tunnel diodes with the gate oxide thickness T_{gox} as a parameter. Since we assumed the gate-SiO$_2$ work function of 3.2 eV as for Al or n$^+$-poly Si gate, the gate tunnel current is dominated by the electron current J_{ct}. This means that a p-type diode is a minority carrier tunnel diode [1] and the n-type diode is a majority carrier tunnel diode [2].

In the bias range corresponding to the accumulation under the gate oxide the both tunnel currents depend exponentially on the gate voltage and oxide thickness. In the bias range corresponding to the strong inversion the both tunnel currents saturate with the gate voltage. This saturation results from the non-equilibrium depletion in the semiconductor when the tunneling process is more effective than the thermal generation processes in the silicon.

For dopant concentration of 10^{16}cm^{-3} the maximum depletion depth at the onset of strong inversion is equal to about 300 nm. This means that reducing the semiconductor thickness below 300 nm should directly influence the minority carrier current via the thermal generation-recombination current J_{th} and indirectly the majority tunnel current via the tunnel current multiplication mechanism [2]. The

influence of semiconductor film thickness T_b on the gate tunnel current components is shown in Fig. 3. Due to the strong tunnel current multiplication in the majority carrier tunnel diodes the effect of silicon film thickness on the total tunnel current is much stronger for the n-type MOS/SOI tunnel diodes.

By influencing the potential distribution in the SOI structure the back gate voltage V_{GB} affects the rate of thermal generation-recombination processes in the semiconductor. Fig. 4 shows for the p-type diode that when the back gate voltage increases, it introduces electrostatically the whole silicon film into the strong inversion, reducing the thermal generation rate and, as a result, the J_{ct} current.

(a)

(b)
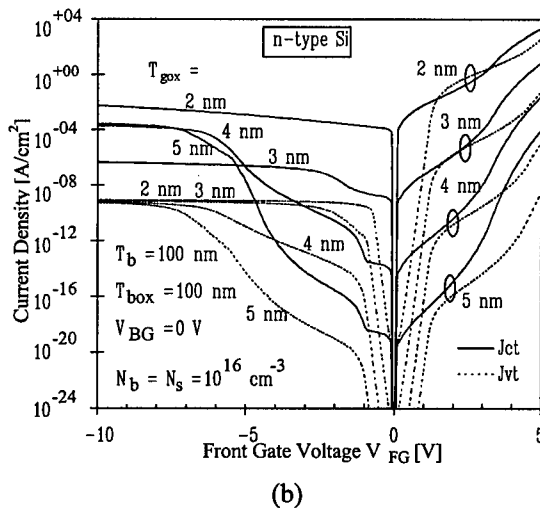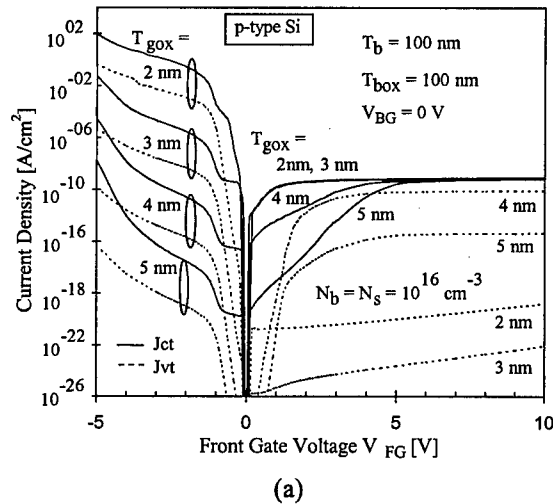
Figure 2. Influence of the gate oxide thickness T_{gox} on I-V characteristics of a p-type (a) and an n-type (b) MOS/SOI tunnel diodes.

M. Badri, B. Majkusiak / Microelectronic Engineering 48 (1999) 375–378
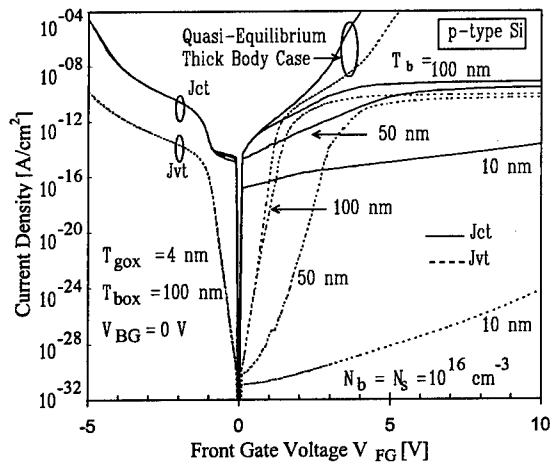
(a)
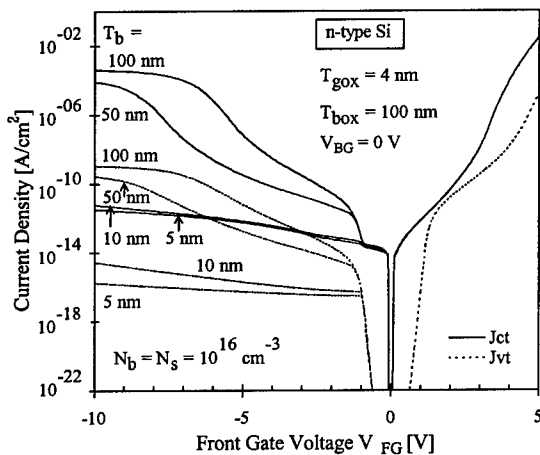
(b)

Figure 3. Influence of the semiconductor film thickness T_b on I-V characteristics of p-type (a) and n-type (b) MOS/SOI tunnel diodes.

If the neighborhood of the under-gate region is included, the carriers generated in the regions B and C are drawn into the under-gate region by a gradient of concentration and the electric field. By enlarging the minority carrier density at the semiconductor surface these carriers affect directly the minority carrier tunnel current and indirectly the majority carrier tunnel current via the tunnel current multiplication mechanism. This effect is illustrated in Fig. 5 for an n-type MOS/SOI diode. The thinner the semiconductor is, the stronger this peripheral effect is.
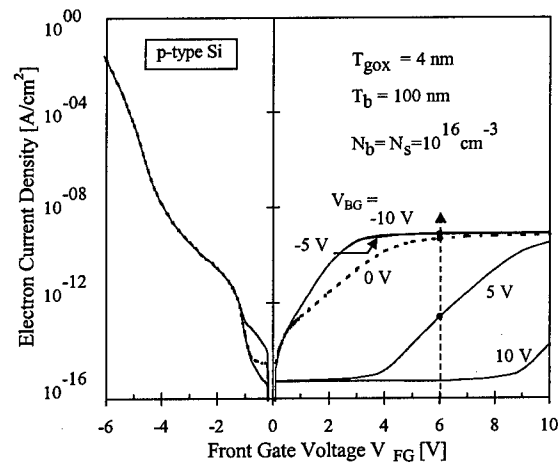
Figure 4. Influence of the back gate voltage on I-V characteristics of the p-type MOS/SOI tunnel diode.
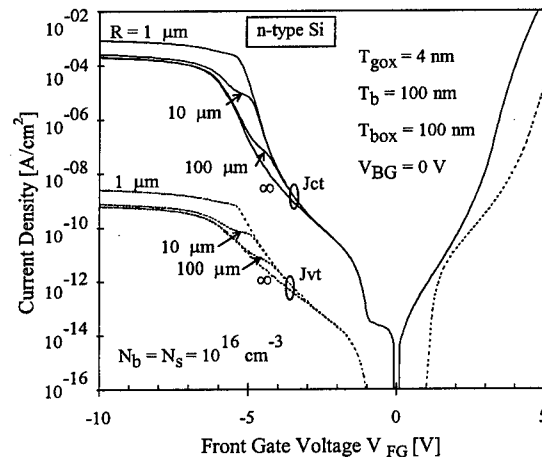
Figure 5. Effect of the gate electrode diameter on the tunnel current density for an n-type MOS/SOI tunnel diode with a 100 nm silicon film.

ACKNOWLEDGMENTS

REFERENCES

1. M. A. Green, F. D. King and J. Shewchun, *Solid-State Electronics*, vol. 17, pp. 551-561, 1974.
2. M.A. Green and J. Shewchun, *Solid-State Electronics*, vol. 17, pp. 349-365, 1974.
3. B. Majkusiak and A. Strojwas, *J. Appl. Phys.*, vol. 74, pp. 5638-5647, 1993.
4. B. Majkusiak, *IEEE Trans. Electron Devices*, vol. 37, pp. 1087-1092, 1990.

Association of high-temperature kink-effect in SIMOX SOI fully depleted n-MOSFET with bias temperature instability of buried oxide

A.N.Nazarov[a], I.P.Barchuk[a], V.S.Lysenko[a] and J.-P.Colinge[b],

[a] Institute of Semiconductor Physics, National Academy of Science of Ukraine,
Prospekt Nauki 43, 252028, Kyiv, Ukraine

[b] Department of ECE, University of California, Davis, CA 95616, USA

A high-temperature drain current "jump" in fully depleted SOI n-channel MOSFET is reported for the first time. The phenomenon appears in the SIMOX SOI MOSFETs at temperatures above 200°C after a negative voltage is applied to the substrate. Direct link of the current "jump" with the positive charge formation in the buried oxide (BOX) has been demonstrated. The current "jump" or a front channel high-temperature kink-effect is explained by electron retrapping in the (BOX) in the vicinity of the BOX-silicon body interface.

1. INTRODUCTION

Fully depleted (FD) silicon-on-insulator (SOI) structures are very attractive for high-temperature CMOS integrated circuit applications [1], and inversion mode (IM) n-MOSFET is one of the main elements of such circuits. However, high-temperature instability processes in the BOX of SOI structures can considerably impact MOSFET operation due to the charge coupling effect between front and back interface in the devices [2]. Charge bias-temperature (BT) instabilities in the BOX of SOI structures have been investigated by several groups [3-6]. It has been shown that BT instability of the BOX can be substantial under some particular conditions [5].

In this work, for the first time, a drain current "jump" in FD SOI MOSFETs is reported, which takes place at high temperature after a negative voltage has been applied to the silicon substrate. The nature of the phenomenon is analyzed.

2. EXPERIMENTAL

FD SOI IM n-MOSFET with L/W=3/3 μm has been fabricated using standard single implanted SIMOX material (the implanted dose was 1.8×10^{18} O^+/cm^2, the energy of implantation was 200keV, and the temperature of implantation was 600°C). Post-implantation annealing was performed at 1320°C in Ar+2%O_2 for 6 hours. After device processing the thickness of the BOX, the silicon layer and the gate oxide were 400, 85 and 35 nm, respectively. The final boron concentration in the channel region is $7 \times 10^{16} cm^{-3}$.

The drain current (I_{sd}) vs. front- and back-gate voltage characteristics of the SOI MOSFETs were measured in the 100 to 320°C temperature range for $V_{sd} = 100$ mV. The voltage applied to the silicon substrate (back-gate voltage, V_{gb}) was usually changed from -30 to +30V and in the some cases was held at different constant values for different times. Prior to every set of measurements all terminals of the MOSFET were grounded and the devices were heated up to 300°C for 300 sec.

3. RESULTS AND DISCUSSION

The sets of $I_{sd}(V_g)$ curves vs. back-gate voltage measured at 300°C are presented in Figure 1. It should be noted that the form of the curves depends considerably on the direction of the back-gate voltage changing. If, at first, the back-gate voltage

was negative, the $I_{sd}(V_g)$ curves show the "jump" of the current in the vicinity of the zero value of gate voltage when the back-gate voltage reaches the positive values (Fig.1a). In another case, when in the beginning of the set of measurements the back-gate voltage was positive, the $I_{sd}(V_g)$ curves have a usual form for all voltages at the substrate (Fig.1b).
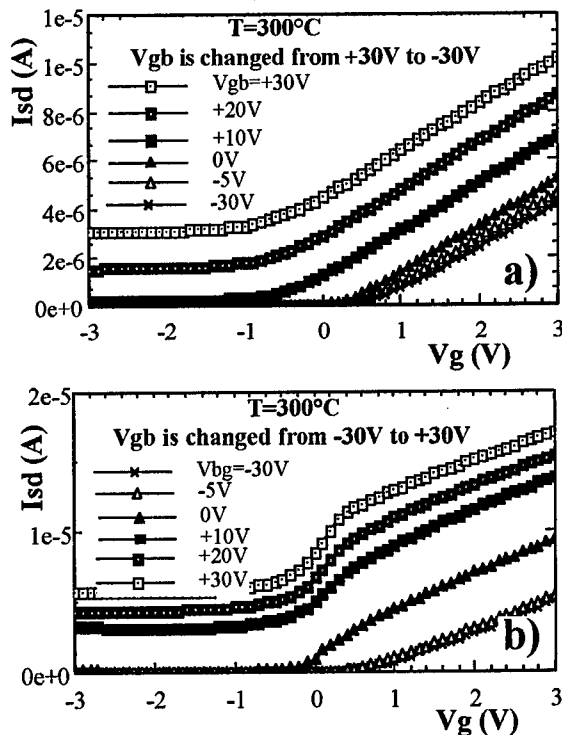
Figure 1. Drain current vs. gate voltage for different substrate voltages (Vgb) at 300°C when substrate voltage is swept from +30V to -30V (a) and from -30V to +30V (b).

It is worthy noting that the current jump originating in the $I_{sd}(V_g)$ curves depends also on the temperature of measurement and is created by the above mentioned conditions at the temperatures higher than 200°C (Fig.2a).

To characterize the $I_{sd}(V_g)$ curves with the current "jump" we determine the value of a threshold voltage for this case as the intersection point of a line, drawn as the extrapolation of the

$I_{sd}(V_g)$ curve to the negative gate voltage, with the voltage axis (see Fig.2a). The variations of the threshold voltages for the front channel transistor (V_{tf}) vs. the back-gate bias (so named charge coupling effect [2]) for different measurement temperatures and for different directions of the back-gate voltage changing are presented in Figure 2b. It is shown that the current jump leads to the more considerable dependence V_{tf} vs. V_{gb} as compared to the theoretical one and this dependence is stronger when the temperature of measurement grows up.
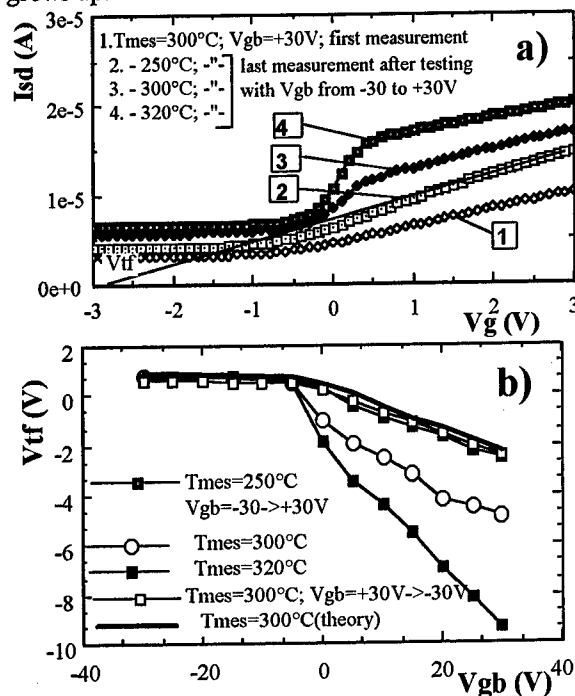
Figure 2. Drain current vs. gate voltage (a) and threshold voltage vs. substrate voltage (b) for different temperature and sequence of measurement.

Because of the high temperature and the use of a low drain voltage, impact-ionization and floating body-related effects [7] are ruled out to explain the current jump. The observed phenomenon can rather be related to the charge transport in the gate oxide

and/or in the BOX and to discharging effects in these dielectric layers.

The important point is that neither the current jump nor hysteresis phenomena are observed when the front gate oxide is subjected to BT stressing at the temperatures up to 320°C and the back-gate bias is constant during the experiments. So, the observed phenomenon is related to processes in the BOX.

In the next experiments a negative voltage is applied for some time to the back gate (silicon substrate). The front gate bias is -3V to avoid the charge coupling effect. If the back-gate voltage is then swept from -15V to +30V, a drain current jump of the back-channel current (back-channel high-temperature kink-effect [6, 8]) is observed in $I_{sd}(V_{gb})$ characteristics (Fig.3a). The presence of this back-channel high-temperature kink-effect suggests that some time- and temperature-dependent generation of positive charge takes place in the BOX when a negative voltage is applied to the substrate (Fig.3b) [8]. The important point is that this positive charge at above mentioned conditions is almost fully compensated by electrons trapped from the substrate. The discompensation and positive charge manifestation can be observed after the back-gate voltage shifts up to zero or to some positive magnitude with a fast sweeping rate (Fig.3c). If, after the negative voltage, we apply to the substrate the positive one, the positive charge in the BOX is neutralized by electrons injected from the silicon body (Fig.3d). This phenomenon has to lead to a decrease of the $I_{sd}V_{gb}$ characteristic slope in comparison with the initial one (without positive charge in the BOX) and to the current relaxation. Actually we observed such effects (Fig.3a).

To estimate the value of BT generated positive charge in the BOX we have determined the threshold voltage of the back-channel MOSFET in the case of the high-temperature kink effect taking into account the discharging effect [8]. For this we draw a line which is parallel to the initial characteristic (without a current jump), from the point of current step beginning to the intersection with the voltage axis. The point of the line intersection with the axis is the threshold voltage (see Fig.3a). Using this method we determined that

the maximum positive charge related to the BOX-silicon body interface can reach 2.5×10^{12} cm^{-2}.
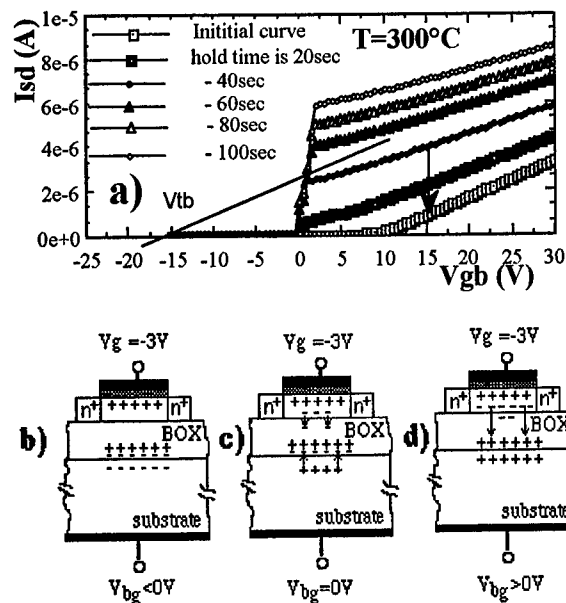
Figure 3. Drain current vs. substrate voltage for different hold time at -15V (measurements were performed from -15V to +30V; arrow indicates the direction of current relaxation) (a); Schematic illustration of the positive charge accumulation in the BOX at $V_g<0$ V (b), of the redistribution of the charge at $V_g=0$ V (c) and the neutralization of the charge at $V_g>0$ V (d).

The link between the drain current jump in the front MOSFET and the positive charge created in the BOX is illustrated in Figure 4a. The application of the negative voltage to the substrate for 150 seconds at a temperature of 250°C gives rise to a small shift of the $I_{sd}(V_g)$ curve towards more negative voltage values, showing that some small positive charge is produced in the BOX (curves 1 and 2 in Figure 4a). Then, the substrate voltage is switched to zero. This causes the discompensation of the positive charge in the BOX and gives rise to a considerable current increase and to formation of a current jump (curve 3 in Figure 4a). After that, if the drain, the source and the substrate are shortened for 150 sec., the values of the drain current and the

magnitude of the current jump decrease (curve 4 in Figure 4a), that is the compensation of the positive charge in the BOX (process shown in Figure 3d) leads to the decrease of the drain current jump in the $I_{sd}(V_g)$ characteristic. So, we can conclude that the drain current jump formation in the MOSFET (which we will call "front-channel high-temperature kink-effect") has a direct link with generation of the positive charge in the BOX and its magnitude.
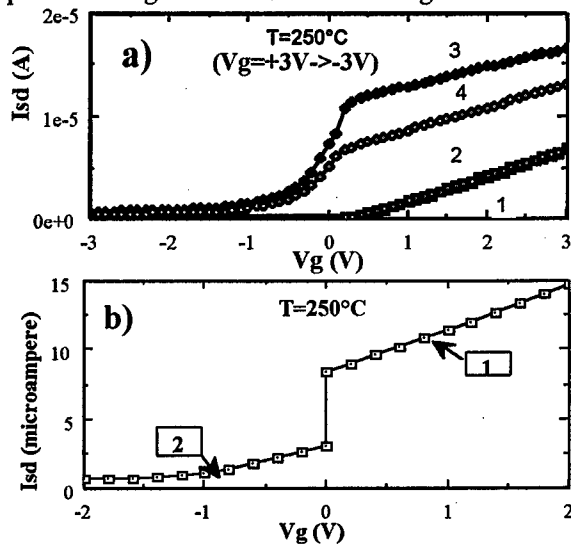
Figure 4. (a) Drain current vs. gate voltage in dependence on conditions of the substrate: 1. V_{gb}=-30V, hold time is 0 sec; 2. V_{gb}=-30V, hold time is 150 sec.; 3. V_{gb}=-0V, hold time is 0 sec after stressing at V_{gb}=-30V for 150 sec.; 4. V_{gb}= 0V, hold time is 150 sec. (b) MEDICI simulation of the drain current "jump".

The phenomenon can be related to trapping of electrons from the back channel of the MOSFET into some metastable states localized in the BOX near the BOX-silicon body interface and created together with the positive charge in the BOX during a negative voltage application to the substrate at high temperature.

The results of 2D-MEDICI simulation [9] are presented in Figure 4b. The part of the curve for V_g>0 was calculated for the case of a uniform positive charge distribution as a function of depth in

the BOX. The positive charge 3.5×10^{17} cm^{-3} is uniformly distributed within a distance of 200 nm from the BOX-silicon substrate interface which compensates the electric field formed by the -30 volts applied to substrate. The part of the curve for Vg<0 in Figure 4b was calculated for the same positive charge distribution, but where 1.05×10^{12} cm^{-2} electron charge is located in the BOX near the BOX-silicon body interface. A good agreement between experimental results (curve 3 Figure 4a) and MEDICI simulations (Figure 4b) can be observed.

ACKNOWLEDGMENT

REFERENCES

1. M.Francis, A.Terao, B.Gentinne, D.Flandre and J.-P.Colinge, IEDM'92 Tech. Dig., (1992) 352.
2. H.K.Lim and J.G.Fossum, IEEE Trans. Electron Devices, 30 (1983) 1244.
3. C.S.Ngwa and S.Hall, Microelectronic Engineering, 28 (1995) 387.
4. P.Dimitrakis, G.J.Papaioannou and S.Cristoloveanu, J.Appl.Phys., 80 (1996) 1605.
5. K.Vanheusden, J.R.Schwank, W.L.Warren, D.M.Fleetwood and R.A.B.Devine, Microelectronic Engineering,, 36 (1997) 241.
6. A.N.Nazarov, J.-P.Colinge and I.P.Barchuk, Microelectronic Engineering, 36 (1997) 363.
7. S.Cristoloveanu and S.S.Li, Electrical Characterization of Silicon-On-Insulator Materials and Devices, Boston: Kluwer, 1995.
8. A.N.Nazarov, I.P.Barchuk, V.S.Lysenko and J.-P.Colinge, Silicon-on-Insulator Technology and Devices IX, PV99-3 (1999) 299.
9. TMA MEDICI, Two-Dimensional Simulation Program. ver.2.2.1.

Splitting and electrical properties of the SOI structure formed from the heavily boron doped silicon with using of the smart-cut technology

I.V. Antonova, V.P. Popov, V.F. Stas, A.K.Gutakovskii, A.E.Plotnikov, V.I.Obodnikov

Institute of Semiconductor Physics Lavrentieva 13, 630090 Novosibirsk, Russia

An increase in boron concentration in hydrogen implanted silicon leads to increase in blister concentration and decrease in defect concentration. Boron-related complexes are supposed to provide nucleation centers for microcavities formed during the ion implantation and following annealing due to effective boron interaction with different types of hydrogen-defect complexes. The high boron concentration in the initial crystal in the combination with high hydrogen concentration in as-bonding SOI leads to n-type of conductivity in the top layer of the SOI structure.

1. INTRODUCTION

Splitting and blistering processes are known to happen rapidly in the heavily boron-doped silicon irradiated by hydrogen ions [1]. In the result a lower hydrogen dose can be used for creation of Silicon-On-Insulator (SOI) structure for initial material with high boron concentration. In order to explain blistering and exfoliation, at least two types of mechanisms have been reported in the literature. The first of them is based on an increase in the H_2 gas pressure inside the microcavities and their growth during annealing [2]. The second connected with the defect stimulated increase in a lattice stress and its release by means of exfoliation [3,4]. The increase in boron concentration in the initial silicon can change the process of H_2 gas accumulation in the microcavities and had to effect on the defect structure of implanted crystal. The aim of the present work was to investigate the influence of high boron concentration on (1) the defect structure, blistering and exfoliation in hydrogen implanted silicon and (2) on the electrical properties of the top silicon layer in the SOI structure.

2. EXPERIMENTAL DETAILS

Czochralski grown silicon wafers with <100> orientation and p-type of conductivity were used for the investigation. The electrically active boron concentration was changed in an interval from 10^{15} cm^{-3} to $10^{20} cm^{-3}$. Samples with boron concentrations of $2x10^{15}$ and $2x10^{18}$ cm^{-3} were doped with boron during crystal growth. Heavily boron doped samples were also prepared by 300 keV B^+ ion implantation (doses are equal to $0.5x10^{16}$, $1x10^{16}$ and $1.5x10^{16}$ cm^{-2}) and annealed at 1075°C for 100 minutes. Average boron concentrations in the implanted layers of these samples are equal to $3x10^{19}$, $6x10^{19}$, $9x10^{19}$ cm^{-3}. Hydrogen implantation was carried out at room temperature by H_2^+ ions with energy of 135 keV in the dose range of $(1.5 - 4.5)x10^{16}$ H/cm^2. The thickness of the oxide layer on the handle wafer was equal to 0.35-0.4 μm and as a rule this oxide was used as a buried oxide for SOI structure. The voltage-capacity (CV) and Hall effect measurements, scanning electron microscopy (SEM), transmission electron microscopy (TEM) and Rutherford back-scattering (RBS), secondary ion mass spectroscopy (SIMS) were used for studying the SOI and implanted layer properties.

3. RESULTS AND DISCUSSION

3.1. Exfoliation process in heavily doped silicon

Figure 1 presents TEM images of samples with different boron concentrations implanted by hydrogen ions and annealed at 1050°C.
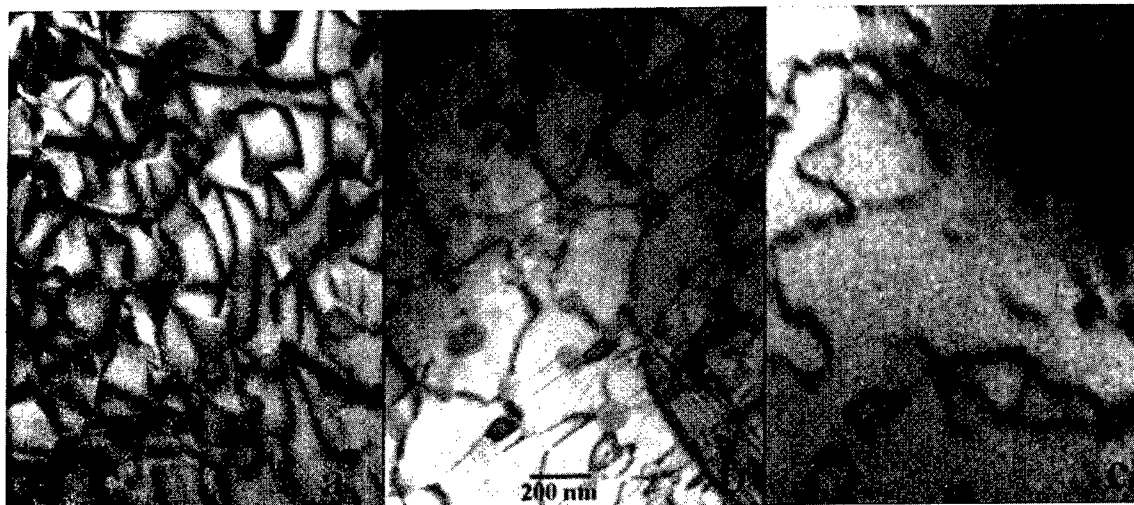
Fig.1 TEM images of the samples with boron concentrations of 1×10^{15} (a) 1×10^{18} (b) 1×10^{20} (c) cm^{-3} implanted by hydrogen ions with dose of $4.5\times10^{16}\,H/cm^2$ and annealed at $1050°C$.

The increase in boron concentration causes the decrease in the concentration of secondary radiation defects (bubbles and dislocation loops for a given annealing temperature) in implanted layer.

Relative changes of the carrier concentration as a function of boron concentration are presented in Fig 2. The increase in hydrogen dose leads to proportional decrease in boron concentration for samples with initial concentration higher than 10^{19} cm^{-3}. Thus, very effective interaction of boron atoms with point defects (V and I) causes an appearance of interstitial boron B_I (or its clusters, complexes), B-H and other boron-hydrogen related complexes in high concentration in silicon lattice and a decrease in concentration of the radiation defects in implanted layer. RBS measurements are really found the decrease in defect concentration near the depth of ion projected range, R_p (0.65 μm) for heavily boron-doped silicon samples in comparison to low-doped samples.

The SEM images of samples with different boron concentration implanted with hydrogen dose of 4.10^{16} cm^{-2} and flaked after annealing at $450°C$ during 1 hour show that the increase in boron concentration leads to increase in blister concentration and decrease in their sizes (Fig.3). Thus, the heavily boron-doped silicon samples have the higher concentration of blister nucleation centers.

Thus, the distinctive characteristics of the

hydrogen implanted heavily doped silicon are high concentration of boron complexes (like B-H, B_I and other) and low concentration of radiation defects.
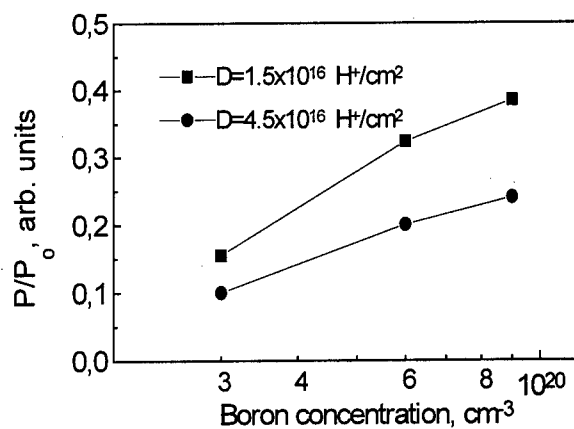
Fig.2. Relative changes in the carrier concentration in as-implanted samples as a function of boron concentration. D is hydrogen dose.

$N_B = 1.10^{15}$ cm^{-3}	$N_B = 5.10^{19}$ cm^{-3}	$N_B = 1.10^{20}$ cm^{-3}
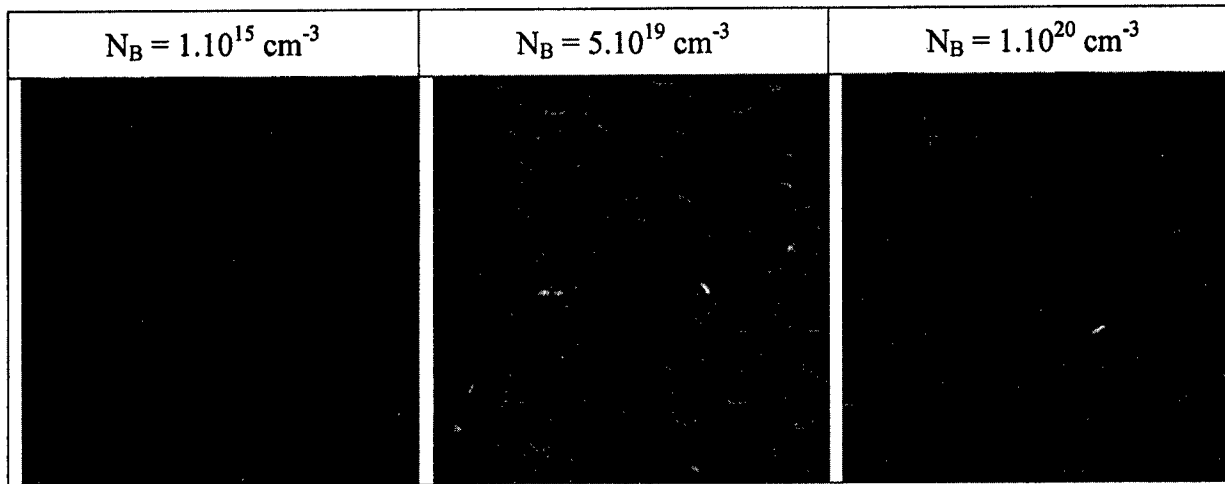
Fig.3. SEM images of surfaces for samples implanted by hydrogen ion with dose of 4.10^{16} cm^{-2} and flaking after annealing at 450°C during1 hour. Image size is equal to 200 μm.
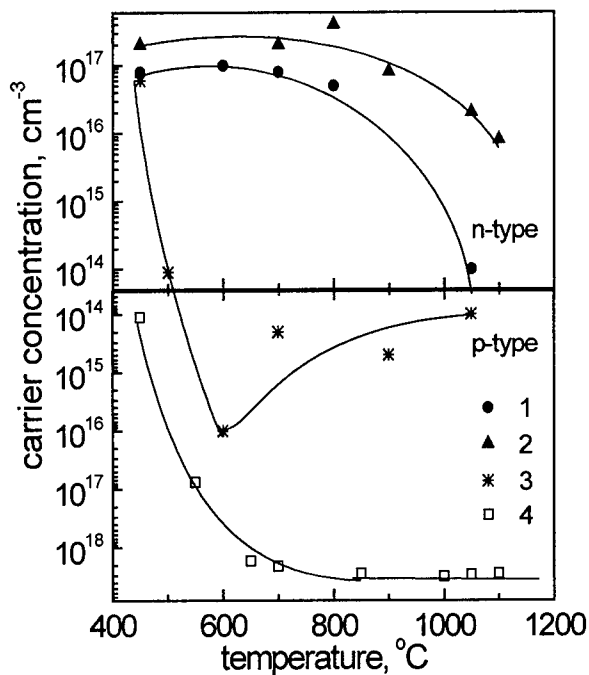
Fig.4 Dependence of carrier concentration in as-bonded SOI top layer on annealing temperature for initial wafers with boron concentration of 2×10^{18} (2,4) and 2×10^{15} cm^{-3} (1,3). Bonding was carried out at temperature 450°C during 0.5 (1,2) and 2 hours (3,4).

These boron-related complexes can provide the nucleation centers for microcavities formed during the ion implantation and following annealing. Moreover B_I can be a catalyst of a H_2 molecule formation due to an indirect recombination of the mobile hydrogen contenting defects (VH$_n$ and IH$_n$) on boron atoms.

3.2. Electrical properties of SOI structures

Two types of silicon wafers were used as initial crystals for fabrication of SOI structures by Smart-Cut technology: the boron concentrations in wafers were of 2×10^{18} and 2×10^{15} cm^{-3}. The temperature of bonding and exfoliation was 450°C. Few minutes are enough for these processes but real annealing time were equal to 0.5 and 2 hours. Let us named this time as bonding time. Commonly the annealing at 1100°C for high quality bonding and defect removal completes the SOI fabrication [5]. In the present work as-bonded wafers were subjected to the isochronal 30min annealing in the temperature range of 450 –1100°C.

Figure 4 presents the changes in carrier concentration as a function of annealing temperatures. P-type of conductivity in the initial crystal in combination with the short bonding time leads to n-type of conductivity in the top layer of the SOI structure. The increase in boron concentration causes the increase in donor

concentration. These donors are stable up to 1100°C. Free carrier mobility in the low doped silicon film is 400 cm^{-2}/V.sec. Prolonged annealing at bonding temperature gives rise to the decrease in donor concentration in the low doped SOI. The heavily doped silicon appears p-type for the 2 h bonded time and attains their initial concentration at 800°C. Carrier mobility in the last case is equal to 80 cm^{-2}/V.sec (mobility in initial wafer was 106 cm^{-2}/V.sec).

The annealing at 450°C is known to lead to the effective removal of hydrogen out of the crystal [6]. Fig. 5 is demonstrated the hydrogen profiles in as – bonded SOI structure obtained by SIMS for bounding time of 2 hours. Hydrogen profiles for SOI annealed at different temperatures are also given in Fig.5. The decrease in bonding time has to lead to increase in the hydrogen concentration in the silicon top layer.

The obtained results allow us to suggest the existence of a high-temperature stable complex of interstitial boron secured by structural defects with the assistance of hydrogen atom(s). Its stability in our opinion is provided by some structural defects (for example, dislocation loops) which as shown the RBS data exist in the structure up to the 1100°C.

The increase in boron concentration higher than 10^{19} cm^{-3} in hydrogen implanted silicon leads to the higher concentration of blister nucleation centers. On the other hand defect concentration in heavily doped silicon is very low due to the effective interaction of boron atoms with deferent types of defects. The boron concentration of 10^{19} cm^{-3} is the minimal one that is needs in the crystal for the complete possible defect annihilation through the boron atoms for hydrogen dose used. Boron-related complexes are supposed to provide nucleation centers for microcavities formed during the ion implantation and following annealing. The high boron concentration in the initial crystal in the combination with the high hydrogen concentration in as-bonding SOI leads to n-type of conductivity in the top layer of the SOI structure. This donor is suggested to be a complex of the interstitial boron with hydrogen atom(s) secured by structural defect. The more prolonged annealing at bonding temperature (450°C) caused decrease in the hydrogen and donor concentration.

Fig.5 Depth hydrogen distribution for SOI structure fabricated with a bonding time of 2 h.

4. SUMMARY

REFERENCES

1. Q.-Y. Tong, R. Scholz, U. Goesele T.-H. Lee, L.-J. Huang, Y.-L. Chao, and T.Y. Tan. Appl. Phys.Lett, 72 (1998) 49.
2. J. H. Evans, J. Nucl. Mater. 68 (1977) 129.
3. M. K. Weldon, V. E. Marsico, Y. J. Chabal, A. Agarwal, D. J. Eaglesham, J. Sapjeta, W. L. Brown, D. C. Jacobson, Y. Caudano, S. B. Christman, E. E. Chaban. Journ. Vac. Scie. & Technol. B15 (1997) 1065.
5. .4. B. M. U. Scherzer, in Sputtering By Particle Bombardment II, Topics in Applied Physics, edited by R. Behrisch (Springer, New York), 52 (1987) 271.
6. Bruel, M. Nucl. Inst. Meths. B, 108, (1996) 313.
7. 6. Hara, T., Kakizaki, Y., Oshima, S., Kitamura, T., Electrochem. Soc. Proceedings, 97-23, (1998) 33.

NON-VOLATILE MEMORY

Low voltage flash memory by use of a substrate bias

Marco Mastrapasqua

Bell Laboratories Lucent Technologies
600 Mountain Ave., Murray Hill, NJ 07974 USA

The enhancement of the gate current in MOSFET devices by use of a substrate bias has been explained by a second impact ionization event at the drain-to-substrate junction. Hot electron luminescence measurements confirm this interpretation showing that only the high-energy tail of the electron distribution is affected by the substrate bias. This phenomenon has been applied to a FLASH memory array to increase the injection efficiency. Low voltage and good control of disturbs has been demonstrated in a 0.5 Mbit embedded FLASH array using substrate bias during programming.

1. INTRODUCTION

Recently there has been a renewed interest in the carrier heating process known as "Secondary Ionization Induced Substrate Hot-Electron" injection [1], by which an enhanced injection gate current in MOSFET devices is observed when applying a substrate bias. This phenomenon, called CHISEL (CHannel Initiated Secondary ELectron injection), has been shown by MonteCarlo simulation [2] to be due to impact ionization feedback at the drain-to-substrate junction.

2. PHYSICAL PHENOMENON

In Figure 1 the CHISEL injection mechanism is sketched: when a channel electron impact ionizes at the pinch-off region it creates an electron-hole pair; the hole, accelerated by the drain to substrate bias, may also undergo a second impact ionization event. The electron, generated by this second impact ionization event, is accelerated

back into the channel contributing to populate the high-energy tail of the electron distribution and has high probability to be injected into the gate. Hot electron luminescence experiments have confirmed this explanation showing that the substrate bias increases the high energy tail of the electron energy distribution [3]. In Fig. 2 the light intensity, normalized to the drain
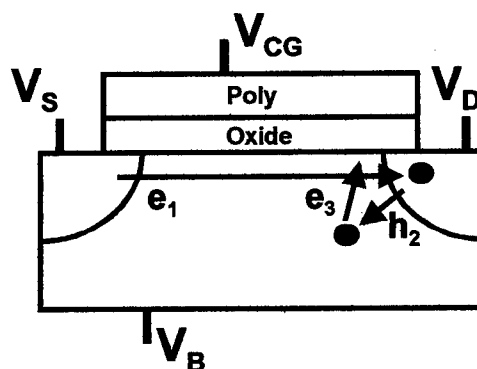
Figure 1: Sketch of the CHannel Initiated Secondary ELectron (CHISEL) injection mechanism.

current, is measured as a function of wavelength for the same device at three different bias condition. Open circle is the light spectrum for channel hot electron condition, $V_{GS}=2.0V$ and $V_{DS}=3.2V$, filled circle correspond to the same gate and drain bias condition but with a negative substrate bias.
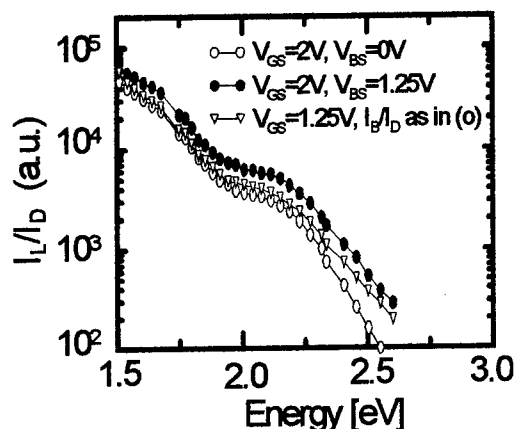
Figure 2: Typical emission spectra at $V_{DS}=2.0V$, $V_{GS} = 2.0V$, $V_{BS} = 0V$ (o) and $V_{BS}=-1.25V$ (•). Open triangles report the emission data at $V_{BS}=-1.25V$ obtained adjusting the gate voltage ($V_{GS} = 2.26V$) in order to keep the same I_B / I_D $\approx 2 \; 10^{-2}$ as in (o).

Because of the body effect, the lateral electric field in the pinch-off region increases and therefore the whole spectrum increases. However, if the gate bias is increased (triangle symbol) to compensate for the body effect, the spectrum differs only in the high energy region. This measurement is the first direct confirmation of the MonteCarlo simulations showing that the effect of substrate bias is to increase the high energy part of the electron distribution in the channel.

3. SINGLE CELL FLASH DATA

The ability to increase the high energy tail of the electron energy distribution has important consequences and application for FLASH memory cells. Due to the increased injection efficiency, with the use of a substrate bias it is possible to obtained comparable programming time with lower drain to source bias. Moreover, the high injection efficiency allow to program the cell in the self-limiting regime (standard channel hot electron injection can not be used in self-limiting regime due to the lower injection efficiency that would make the programming time unacceptably long). Figure 3 shows a typical stacked gate FLASH cell characteristic: the programming proceeds rapidly until the device turns itself off, saturating to a threshold voltage linearly controlled by the word-line bias. As shown in Fig. 3, the value at which the threshold voltage saturated is independent to the initial threshold of the cell and depends only from the word-line bias used during programming.
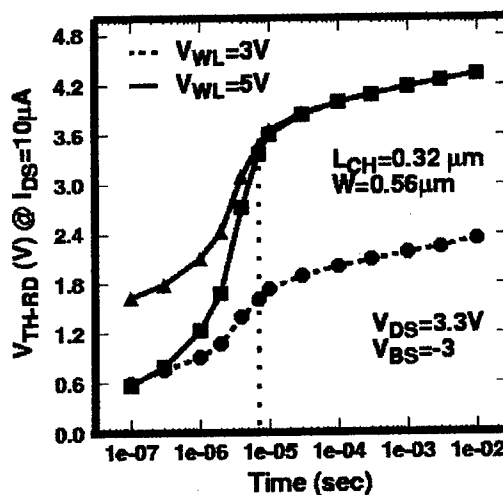
Figure 3: Single cell programming characteristic from two initial states using substrate bias.

4. FLASH ARRAY DATA

We have evaluated the use of a substrate bias in a 0.59 Mbit embedded memory array containing 9 identical 64 kbit blocks (256 rows x 256 columns). The flash NOR array is embedded together with a 0.35 μm CMOS DSP, Fig. 4, using 5 additional lithography steps. The array is placed in a triple well allowing negative substrate biases during writing and positive substrate biases for uniform Fowler-Nordheim erasure.

Figure 4: DSP chip micrograph used to evaluate the CHISEL injection mechanism in a FLASH array.

Figure 5(a) shows the programming time and disturb time for a 64 kbit memory block as a function of the triple well bias, V_{TW}. The programming time was obtained by programming (16 bits a time) a checkerboard pattern from an all zeros state until it was correctly read by on chip sense amplifiers (V_{th}=~2.5 V for a '0' to '1' transition). Therefore the programming pulse width of Fig. 5 corresponds to the longest programming time of 32,000 cells. Continuing the programming of the checkerboard

pattern, the disturb time is defined as the time at which the first of the '0' becomes, as results of any disturb mechanism, a '1'. Because the probability of the second impact ionization phenomena strongly increases with the decrease of V_{TW}, the programming time decreases, approximately four orders of magnitude in going from V_{TW}= 0.0 V to -3.0 V. Simultaneously, the disturb time also decreases. The disturb is due to a parasitic feedback effect initiated by the drain-tub reverse bias junction leakage. Increasing the chuck temperature reduces programming time through increased phonon scattering, but the disturb/program margin is almost insensitive to temperature.

The dependence of the programming /disturb time on word-line bias is shown in Figure 5(b). It is possible to observe an increase in the programming time with increasing V_{WL}. This behavior has the same origin as the well-known decrease of substrate current with increasing gate bias in MOSFETs. Note that this behavior is opposite to channel hot electron injection, where the injection current increases with V_{WL}. The dramatic increase in the programming time for $V_{WL} < 4.0$ V is due to the self-limiting programming characteristic of CHISEL injection. When V_{WL}=3.5 V the cell reaches its programmed threshold voltage in about 3 μs (extrapolating the programming time measured for V_{WL}=4.0 V), but its value it is not sufficient to be read as '1'. Only after the programming time reached 0.2 ms the threshold increases above the level required for reading a '1' state. This slow programming corresponds to the gentle increase of the threshold seen in Figure 3 after the first sharp rise.

Because the array is placed in a triple well allowing negative substrate biases during writing, it is possible to use positive substrate for uniform Fowler-Nordheim erasure. Uniform erase is intrinsically a low current erase method and it has been shown to have better cycling performance [6] than non-uniform erase techniques. Erasure of all 9 memory blocks was achieved in 1 s with 8 V on the triple well (V_{TW}), bit line and source line, and -5 V on the word line (V_{WL}).
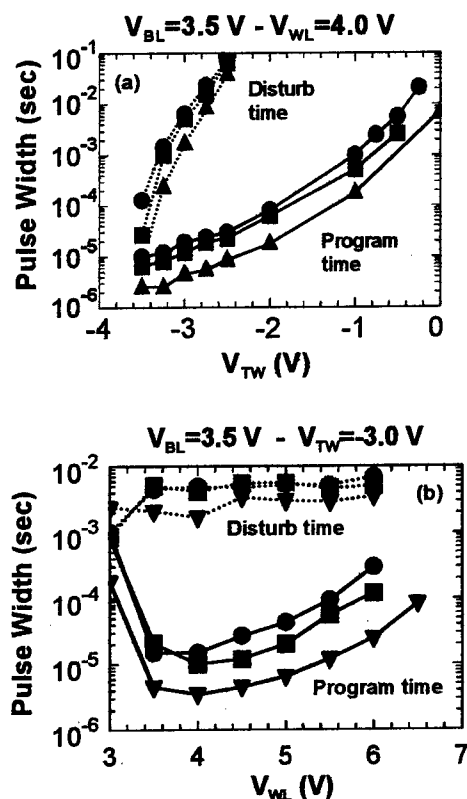
Figure 5: Program and disturb time for 256x256 bit programmed 16 bits a time with a checkerboard pattern. Triangle 25 °C, square 80 °C and circle 110 °C. (a) $V_{BL}=3.5$ V, $V_{WL}=4.0$ V, function of triple well bias. (b) $V_{BL}=3.5$ V, $V_{TW}=-3.0$ V, function of word line bias.

For this test chip, high voltages are supplied externally, but due to the low current required for both erase and writing, a small area on chip charge pump can be easily integrated to supply the required voltages from a single low voltage power supply.

After all 9 memory block are erased, the threshold voltage (V_{th}) distribution is broad (~2V wide) having a maximum near 0.5 V. The array was then written, using a low word line voltage (V_{WL}^0), to a '0' state using a single programming pulse for each cell with no program verify. The convergence property of CHISEL mechanism guarantees that all cells stop programming when they reach a V_{th} of about 1V. The threshold value at which the cell converges, as suggested by Figure 3, is about 1.2 V less than the value of the word line used to write. Figure 6 shows the tight V_{th} distribution obtained after this simple erase algorithm.
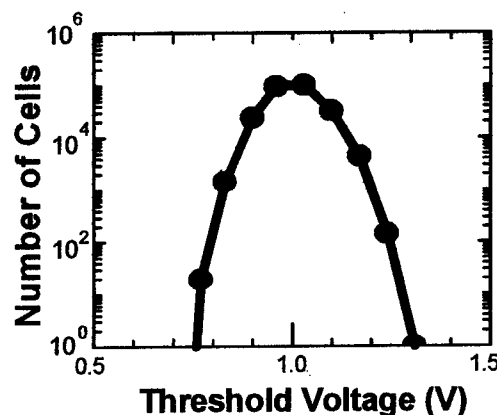
Figure 6: Threshold voltage distribution after the erase algorithm.

The chip read access performance depends on the width of the erase Vth distribution and on its position relative to the threshold of the sense amplifier. Varying V_{WL}^0 during the erase algorithm controls the center of the '0' distribution.

The optimum condition is reached for the lowest V_{WL}^0 that still prevents bit line leakage of unselected cells. It is important to note that in this DSP no word-line boosting is done during reading ($V_{WL}=V_{DD}$) and that the sense amplifier threshold decreases with decreasing V_{DD}. High-speed access requires large cell currents, and hence a lower voltage Vth distribution for a given read V_{WL}. Due to the increased sense amplifier sensitivity at low V_{DD}, bit line leakage becomes a problem, and a higher V_{WL}^0 voltage is required to write the '0' state. For low voltage operation, V_{WL}^0=2.5 V brings the mean of the zero distribution closer to the sense amplifier '0' to '1' transition, and farther away from the over-erase condition (Figure 7). Decreasing V_{WL}^0 produces a zero distribution optimized for fast read access but slightly higher operation voltage.
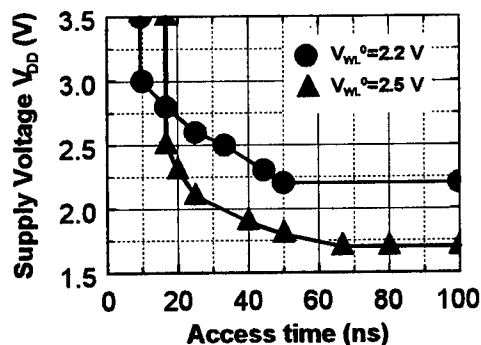
Figure 7: Lower-bound curve of the Schmoo plots at room temperature.

5. CONCLUSION

We have shown that the use of substrate bias increases the high-energy tail of the electron energy distribution because of a second impact ionization event at the drain-to-substrate junction.

This phenomenon has been applied to a flash memory array embedded in a DSP. Fast programming time with good control of disturbs has been demonstrated.

REFERENCES

[1] Y. Nakagome, E. Takeda, H. Kume, and S. Asai, "New observation of hot-carrier injection mechanism," Japanese Journal of Applied Physics, vol. 19, no. Supplement 19-1, p. 85, 1983.

[2] J.D. Bude, "Gate Current by Impact Ionization Feedback in Sub-Micron MOSFET Technologies," in Proc.Symp.on VLSI Technology, p. 101, 1995.

[3] M. Pavesi, L. Selmi, M. Manfredi, E. Sangiorgi, M. Mastrapasqua and J. Bude "Evidence of substrate enhanced high energy tails in the distribution function of deep submicron MOSFETs by light emission measurements" submitted to Electron Device Letters, 1999.

[4] J.D. Bude, M. Mastrapasqua, M.R. Pinto, R.W. Gregor, P.J. Kelley, R.A. Kohler, C.W. Leung, Y. Ma, R.J. McPartland, P.K. Roy, and R. Singh, "Secondary Electron Flash - a High Performance, Low Power Flash Technology for 0.35 μm and Below," in IEDM Technical Digest, p. 279, 1997.

[5] C. Y.Hu, D.L. Kencke and S.K.
Banerjee   "Substrate-Current-Induced
Hot Electron (SCIHE) injection: a new
convergence scheme for flash memory"
1995 IEDM Tech. Digest p. 283, 1995.
[6] J.D. Blauwe, J. Van Houdt, D.
Wellekens,   L.   Haspeslaghg,   G.
Groeseneken and H.E. Maes "Reliable
5.9 nm tunneling-oxide Flash E2PROM
device" 15th annual IEEE Non-Volatile
Semiconductor   Memory   Workshop,
February 9-12, 1997.

Oxide Scaling Limit for Future Logic and Memory Technology

J.H. Stathis and D.J. DiMaria

IBM Research Division, T.J. Watson Research Center, P.O. Box 218, Yorktown Heights NY 10598 USA

The limit of MOSFET oxide scaling is examined from the viewpoint of reliability. Measurements of the voltage dependence of the defect generation rate and the thickness dependence of the critical defect density, together with the breakdown statistics for ultra-thin oxides, are used to provide a general framework for predicting the lifetime of ultra-thin oxides at operating voltage. It is argued that reliability is the limiting factor for oxide thickness reduction.

1. INTRODUCTION

Aggressive scaling of the oxide thickness for future MOS logic and memory technology requires, among other concerns, an assurance that the oxide can meet requirements for reliability [1]. Defects generated by electrons tunneling through the gate oxide ultimately lead to destructive breakdown. We have investigated the relationship between the rate of defect generation and the gate voltage down to voltages where transport occurs by direct tunneling. Next, the critical defect density at breakdown has been measured as a function of oxide thickness in the range 2 to 6 nm. The complete thickness dependence is well described by a percolation model. Finally, the breakdown statistics appropriate for thin oxides have been used to predict time-to-breakdown at low voltage for chips built with ultra thin gate oxide. The result is shown in Figure 1. This model predicts that oxide reliability will limit the oxide scaling to about 2.6 nm (CV extrapolated thickness) or 2.2 nm (QM thickness) for a 1V supply voltage at room temperature and that the current Semiconductor Industry Association (SIA) roadmap will be unattainable for reliability reasons by sometime early next century. In this paper we will describe the experiments that led to the breakdown model predictions of Figure 1. The assumptions of the model are critically examined.

2. SAMPLES AND EXPERIMENTAL TECHNIQUES

A variety of *n*- and *p*-channel FETs were used, with oxide thickness (t_{ox}) ranging from about 1.4 to 5 nm. The gate oxides were grown by a variety of processes including oxidation in O_2, N_2O, and NO,
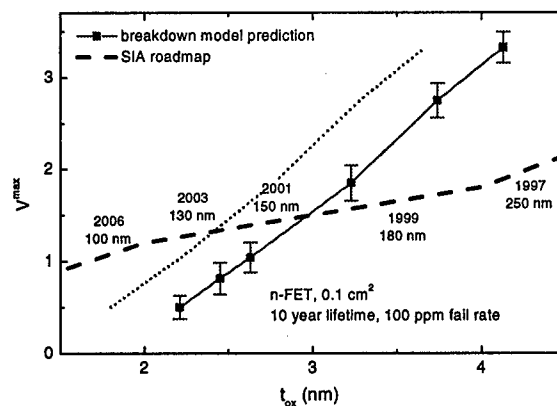
Figure 1. The largest allowable supply (gate) voltage, V^{max}, as a function of oxide thickness, consistent with a failure rate of 100 ppm for a total gate area of 0.1 cm^2 on a chip operating at room temperature for 10 years. Solid and dotted curves correspond to t_{ox} measured by CV extrapolation or a QM method, respectively. The dashed line is the SIA roadmap (1997 version) for t_{ox} and V_{dd}, corresponding to the indicated year of first manufacture and technology generation.

J. Stathis, D. DiMaria / Microelectronic Engineering 48 (1999) 395–401
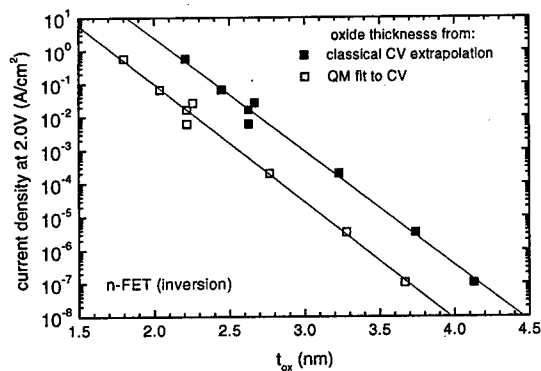
Figure 2. Measured exponential dependence of current density on oxide thickness. Data are for n-FETs in inversion. Solid symbols correspond to oxide thickness determined by a classical CV extrapolation technique, open symbols correspond to thickness determined from a Schrödinger/Poisson fit to the CV curves.

and N_2O oxidation of nitrogen ion-implanted substrates. Unless otherwise stated, the oxide thickness was obtained from the capacitance- voltage (CV) characteristics using a classical extrapolation method. This gives values which are about 0.3 - 0.5 nm larger than those obtained from a self-consistent Schrödinger-Poisson quantum mechanical (QM) analysis. For the thinnest samples ($\lesssim$2nm), where CV measurement was not possible because of gate leakage, t_{ox} was estimated from the gate current in

direct tunneling using the relation shown in Figure 2. The device area was 5×10^{-4} cm^2 for most measurements.

Constant-voltage stress was applied using voltages from 2 to 8 V of either polarity. The current was monitored during stress, and integrated to determine the charge-to-breakdown (Q_{BD}). Breakdown was defined as the first sudden increase in the direct tunneling leakage current. Generated defect densities during stress were determined by interrupting the stress and measuring the stress-induced leakage current (SILC) at a sense voltage of 1.2 - 2.8 V. The SILC is expressed as the relative change in leakage current $\Delta J/J_0$, which can be related to the increase in neutral electron traps generated by hot electrons [2]. J_0 is the initial value of leakage current at the sense voltage. High frequency CV measurements were also used on thicker oxides ($\geq$4.0 nm) to determine trapped charge and interface state densities. The dimensionless SILC can be normalized to the measured interface state density in the region 4-5nm where the CV and SILC techniques overlap, from which we obtain the relation $N_s(cm^{-2}) \approx 3 \times 10^8 \Delta J/J_0$. This is consistent with the value found previously [2] for bulk neutral traps, where $N_n(cm^{-2}) \approx 1.25 \times 10^8 \Delta J/J_0$. Typical data illustrating defect generation as a function of injected charge for various gate voltages are shown in Figures 3-4.
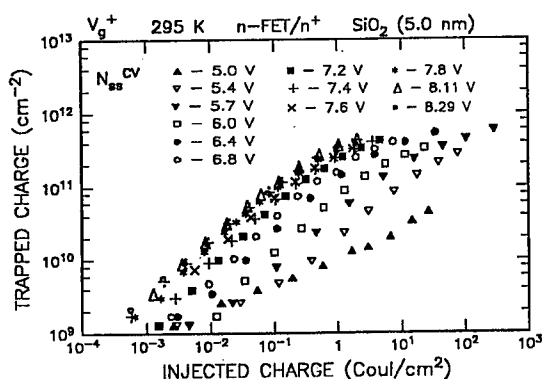
Figure 3. CV measurement of generated interface states as a function of injected charge for various gate biases, for n-FETs with a 5 nm oxide under positive bias at room temperature.
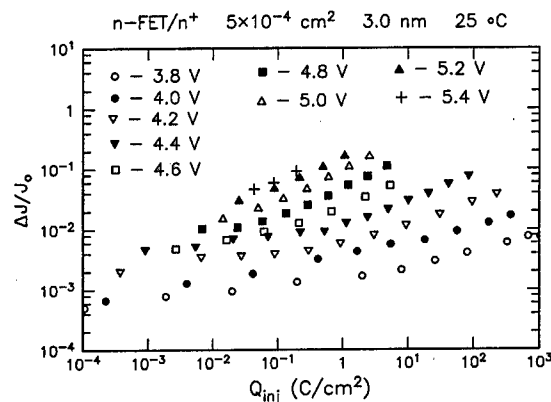
Figure 4. SILC measurement as a function of injected charge for various gate stress voltages, for n-FETs with a 3 nm gate oxide under positive gate bias at room temperature. The sense voltage is 2 V.

3. DEFECT GENERATION RATE

A strong shift of the defect generation to larger fluence with decreasing stress voltage can be seen in Figures 3 and 4. This is quantified in the average defect generation rate, P_g, which is extracted from the slopes in the linear regions of these data or by fitting to a linear function of fluence near breakdown [1,2]. Typically, the defect generation curves exhibit a sub-linear appearance at low fluence, which may be caused by slow charging of interface states. Also, in thicker oxides (>3 nm), or at high fluence when low stress voltage is used, the measured defect density may saturate, possibly because of trap occupancy effects. Because of the strong voltage dependence, the exact fluence dependence of the defect generation has little influence on what follows. Figure 5 shows P_g for a wide variety of samples at room temperature plotted as a function of gate voltage (V_g). P_g exhibits a power law threshold at about 5.6V attributed to hydrogen release from the anode, and an exponential sub-threshold tail extending over 12 orders of magnitude down to at least 2V, which is the lowest stress voltage so far measured using uniform tunneling stress. The slope of the exponential region is 3±0.1 decades/Volt. P_g is independent of gate area. Within the scatter, P_g is independent of t_{ox} and oxidation process. Apparent deviations from the exponential dependence at the higher end of the measured voltage range for some oxide thicknesses are caused by series resistance. P_g does not exhibit a hard threshold or any other features below 5V. In particular, there is no discontinuity or change in slope at the transition from Fowler-Nordheim to direct tunneling at about 3V. At higher voltage, additional thresholds have been observed corresponding to anode hole injection [3] above ~8V and impact ionization [4] at ~12V. These observations suggest that the hydrogen release mechanism of oxide degradation controls all defect generation at low voltage near operating conditions.

4. CRITICAL DEFECT DENSITY AT BREAKDOWN

The critical defect density at breakdown (N^{BD}) is a strong function of t_{ox} [5-8]. Figure 6 shows the measured N^{BD} for each t_{ox} over a range of stress voltages [6-8]. In contrast to P_g, the data for N^{BD} do not show any universal dependence on stress
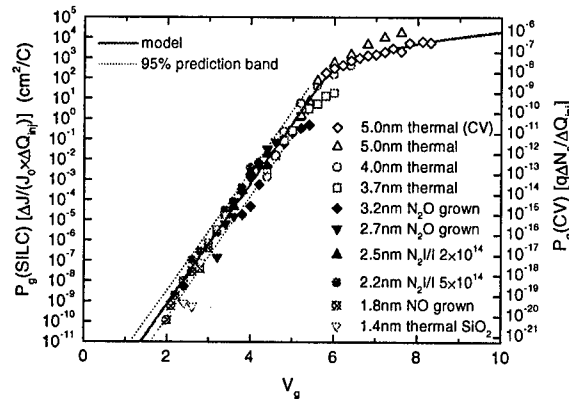
Figure 5. Average defect generation rate as a function of gate voltage. All data were taken at room temperature. Data are from SILC measurements except for the open diamonds which are from CV stretch-out. The relative normalization of CV and SILC data is the same as in Figure 7. The solid line is the model based on an exponential fit to the data below 5.6 V and a power law fit to the CV data.

voltage. Figure 7 shows the dependence on t_{ox} for a gate area of 5×10^{-4} cm². Each point in this figure is the average of several values taken over a range of stress voltages. The normalization of the SILC values and the defect densities from CV is consistent throughout Figures 5-7. The thickness dependence of N^{BD} is quantitatively explained by a percolation model [5,7-9] shown by the solid symbols in Figure 7. The percolation model of breakdown predicts a singular behavior when t_{ox} becomes less than the effective size a_0 of a defect, because in this limit a single defect near the center of the oxide is sufficient
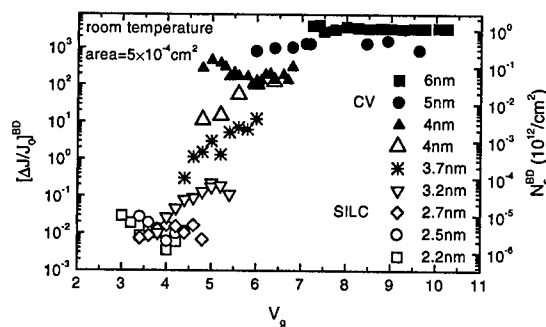
Figure 6. Critical defect density for breakdown over a range of stress voltages, for various oxide thickness. Closed symbols are interface state densities from CV stretch-out, open symbols are SILC measurements of $\Delta J/J_0$ at breakdown.
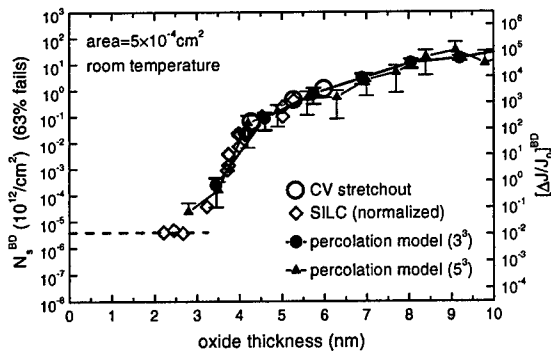
Figure 7. Critical defect density for breakdown as a function of oxide thickness. Open circles are interface state densities from CV stretch-out. Open diamonds are SILC measurements of $\Delta J/J_0$ at breakdown. The data from the two techniques are matched at 5 nm. Solid symbols are from percolation model calculations on a simple cubic lattice with lattice constant $a_0/3$ or $a_0/5$ [9]. The dashed line is the thin-oxide limit predicted from the percolation model [8].

to create a continuous path across the sample. The N^{BD} data are consistent with this model. The critical defect density exhibits a steep drop from about 5 to 3 nm, and becomes constant below this thickness. The plateau value below 3 nm corresponds to a constant number equal to 2000 defects in the 5×10^{-4} cm^2 area of each sample, suggesting that each percolation path has a probability of about 10^{-3} of initiating a destructive breakdown event. Using this value for the fraction of "effective" defects, an excellent fit to

the data is obtained over the entire thickness range with the defect size $a_0 = 3.5$ nm.

5. CALCULATION OF CHARGE-TO-BREAKDOWN AND CHIP LIFETIME

Knowing the critical defect density and the average defect generation rate allows the calculation of Q_{BD} from the relation $Q_{BD}=N^{BD}/P_g$. The result is shown in Figure 8 together with experimental values. The agreement with the measured Q_{BD} is very good with the exception of the thicker oxides where the saturation of the CV measurement causes N^{BD} to be underestimated. Note that Q_{BD} becomes relatively independent of thickness independent below 3 nm, in accordance with the behavior of N^{BD}. T_{BD} is calculated (Figure 9) by dividing Q_{BD} by the measured current density, which is nearly constant during stress for thin oxides (<4 nm). Below 3 nm, T_{BD} decreases exponentially with decreasing t_{ox}, even though Q_{BD} is constant, because of the increasing tunnel current.

The 10 year lifetime indicated in Figure 9 must be projected to the full gate area on the chip and to a specified low cumulative failure rate using the Weibull distribution [10]. The area and failure rate scaling is sensitive to the Weibull slope β, which is thickness dependent [5,11] and goes to unity for ultra thin oxides, as shown in Figure 10. The lifetime scales as $\left[\frac{A}{A'}\right]^{1/\beta}$ where A is the area of the test
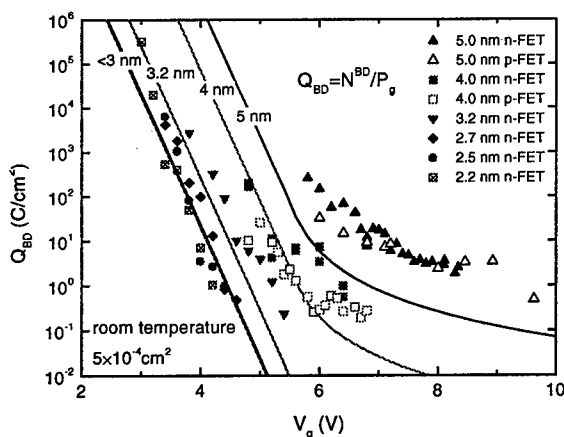
Figure 8. Q_{BD} data and model calculation as a function of gate voltage. Open symbols are p-FETs, filled symbols are n-FETs.
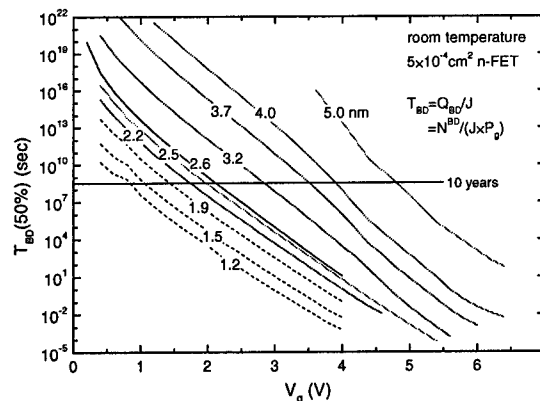
Figure 9. Calculated gate voltage dependence of median time-to-breakdown for various oxide thicknesses. The measured current density is used for $t_{ox}\geq2.2$ nm, and an exponential extrapolation of current density vs. t_{ox} is used for thinner oxides.
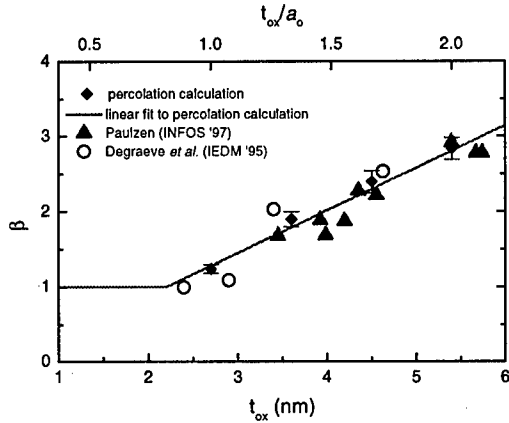
Figure 10. Measured and calculated Weibull slope (β) as a function of oxide thickness.

structure and A' is the total gate area on a chip, and approximately as $\left[\frac{F'}{F}\right]^{1/\beta}$ where F is the measured failure rate (50%) and F' is the desired low failure rate for the product. Figure 1 shows the results for 0.1 cm^2 gate area and 100 ppm failure rate in 10 years at room temperature. Plotted is the maximum gate voltage which can be tolerated for the stated reliability specification. Error bars reflect the measured uncertainty in P_g and N^{BD}. This figure predicts that silicon dioxide cannot be made much
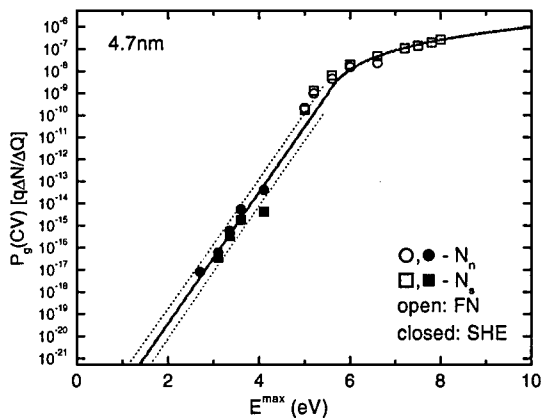
thinner than about 2.6±0.15 nm (2.2 nm according to the QM thickness method) assuming 1 V operation, without becoming unacceptable from a reliability standpoint. For chips operating at elevated temperature, *e.g.* 100-150°C, the minimum t_{ox} will be ~0.2 nm larger because both P_g and N^{BD} are temperature dependent [12].

6. POSSIBLE LIMITATIONS OF THE MODEL

The extrapolation of T_{BD} or Q_{BD} to operating voltage involves an extrapolation beyond the measured range of V_g for both P_g and N^{BD}. It is assumed that the voltage dependence is entirely contained in P_g and that the t_{ox} dependence is entirely contained in N^{BD}. The model predicts that Q_{BD} will have an exponential voltage dependence, the inverse of the voltage dependence of P_g which is approximately 3 decades/Volt.

Since P_g has been measured to 2 V, the extrapolation involves at most 2-3 orders of magnitude in this quantity. However, Figure 7 shows P_g data for each t_{ox} over only a limited range of V_g, leaving the possibility that P_g for a particular t_{ox} might deviate from the exponential slope at lower voltage. To test this, measurements have been performed using substrate hot electron (SHE) injection on injector structures [13]. SHE allows for variation of the electron energy distribution at the substrate/oxide interface independent of the oxide field or electron flux. Figure 11 shows CV data for an *n*-FET with a 4.7 nm oxide for electron energy as low as 2.7 eV. For SHE the energy E^{max} is $q(V_s+\phi_{np})$ where V_s is the voltage applied to the substrate and ϕ_{np} is the contact potential difference (-1.1V) between the n^+ inversion layer and the p^+ contact. The figure shows that defect generation by hot electrons impinging on the substrate/oxide interface follows the same dependence on energy as that from Fowler-Nordheim injection through the oxide. This universal energy dependence, independent of t_{ox}, permits knowledge of P_g to within 1 volt of operating voltage for 1-1.5V technologies.

Also implicit in the voltage extrapolation of T_{BD} or Q_{BD} is the assumption that N^{BD} is independent of V_g [1,14]. There is experimental support for this assumption, at least as a first order approximation [5-8,15,16], but for accurate reliability projections it will be necessary to examine this assumption in greater detail. Published data have examined limited

Figure 11. Defect generation rate from CV measurements on an *n*-FET with a 4.7 nm oxide, comparing Fowler-Nordheim injection above 5 V ($E^{max}=-qV_g$) and SHE ($E^{max}=q(V_s+\phi_{np})$). N_n and N_s are occupied electron traps and interface states, respectively. The line is the model curve derived from the data shown in Figure 5.

ranges of stress voltage [17] or lack sufficient statistical accuracy [6-8]. Figure 6 seems to indicate that substantial voltage-dependent corrections to N^{BD} will be required, although these are far from being understood at this time. The CV and SILC data at 4 nm show different trends, with the SILC at breakdown decreasing as V_g is reduced. This dependence of the saturation value of the SILC on stress voltage has been discussed previously [2]. For the thinnest oxides (<3 nm) studied in Figure 6 there is no clear trend, with N^{BD} decreasing in some cases and increasing in others as V_g is reduced.

A voltage-dependent N^{BD} could arise from several sources. The percolation path, which has been modeled [5,9] in zero field to obtain the fit shown in Figure 7, could in fact be weakly field dependent. The formation of the percolation path, *i.e.* the generation of new defects, could depend on the local field produced by the other defects [18,19]. This would lead to more directed paths at higher voltage, so that the average defect density to form a connecting path across the sample would be reduced. A competing effect of a field-dependent constraint on hopping direction has also been proposed [20]. In addition, the probability that a given percolation path will trigger breakdown (the fraction of "effective" defects [8]) might be voltage dependent. Statistically significant data, on a range of oxide thicknesses and process conditions, will be required to address these issues.

7. CONCLUSIONS

Figure 1 illustrates that at operating conditions gate oxide failure caused by intrinsic degradation and breakdown has not been a significant reliability concern in the past, but that reliability will soon become a limiting factor. If projected scaling trends continue, this limit may be reached earlier than other factors such as chip standby power consumption. If SiO_2 cannot be scaled to meet future demands, another material with higher reliability will be needed. Unfortunately, much work remains before any new dielectric material will reach the level of quality needed to obtain device performance comparable to what can be achieved at present with SiO_2.

With reliability seen as a limiting factor for device scaling, it must be brought to the forefront of concern in FET design and process development.

The influence of oxide breakdown on device and circuit performance will also need to be investigated [21,22]. The general features of the physics of SiO_2 breakdown have been described in this paper, as a voltage dependent defect generation process and a thickness dependent critical defect density. Greater precision in reliability projection will require deeper understanding of these parameters.

REFERENCES

1. J.H. Stathis and D.J. DiMaria, IEDM Tech. Dig., 1998, p. 167.
2. D.J. DiMaria and E. Cartier, J. Appl. Phys. **78**, 3883 (1995).
3. D.J. DiMaria, E. Cartier, and D.A. Buchanan, J. Appl. Phys. **80**, 304 (1996).
4. D.J. DiMaria, E. Cartier, and D. Arnold, J. Appl. Phys. Lett. **73**, 3367 (1993).
5. R. Degraeve, G. Groeseneken, R. Bellens, M. Depas, and H.E. Maes, IEDM Tech. Dig., 1995, p. 863; R. Degraeve, G. Groeseneken, R. Bellens, J.L. Ogier, M. Depas, P.J. Roussel, and H.E. Maes, IEEE Trans. Electron Dev. **45**, 904 (1998).
6. D.J. DiMaria, Microelectron. Eng. **36**, 317 (1997).
7. D.J. DiMaria and J.H. Stathis, Appl. Phys. Lett. **70**, 2708 (1997).
8. D.J. DiMaria and J.H. Stathis, Appl. Phys. Lett. **71**, 3230 (1997).
9. J.H. Stathis, Microelectron. Eng. **36**, 325 (1997).
10. E.Y. Wu, W.W. Abadeer, L.-K. Han, S.-H. Lo, and G. R. Hueckel, Proc. IRPS, 1999, p. 57.
11. G.M. Paulzen, Microelectron. Eng. **36**, 321 (1997).
12. D.J. DiMaria and J.H. Stathis, Appl. Phys. Lett. **74**, 1752 (1999).
13. D.J. DiMaria, to be published, J. Appl. Phys. (August 1999).
14. K. Okada, H. Kubo, A. Ishinaga, and K. Yoneda, Symposium on VLSI Technology Tech. Dig., 1998, p. 158.
15. E. Vincent, S. Bruyere, C. Papadas, and P. Mortini, Microelectron. Reliab. **37**, 1499 (1997).
16. K. Umeda, T. Tomita, and K. Taniguchi, Electron. and Commun. in Japan **2-80**, 11 (1997); K. Umeda and K. Taniguchi, J. Appl. Phys. **82**, 297 (1997); Y. Kamakura, H.

Utsunomiya, T. Tomita, K. Umeda, and K. Taniguchi, IEDM Tech. Dig., 1997, p. 81.

17. T. Nigam, R. Degraeve, G. Groeseneken, M.M. Heyns, and H.E. Maes, Proc. IRPS, 1999, p. 381.

18. L. Niemeyer, L. Pietronero, and H.J. Wiesmann, Phys. Rev. Lett. **52**, 1033 (1984).

19. T. Tanamoto and A. Toriumi, Jpn. J. Appl. Phys. **36**, 1439 (1997).

20. K.P. Cheung, C.T. Liu, C-P Chang, J.I. Colonell, W. Y-C Lai, R. Liu, J.F. Miner, C.S. Pai, H. Vaidya, J.T. Clemens, and E. Hasagawa, Proc. IRPS, 1999, p. 52.

21. E. Wu, E. Nowak, J. Aitken, W. Abadeer, L.K. Han, and S. Lo, IEDM Tech. Dig., 1998, p. 187.

22. T. Pomp, H. Wurzer, M. Kerber, R.C.W. Wilkins, and I. Eisele, Proc. IRPS, 1999, p. 82.

Reliability of thin dielectric for non-volatile applications

A. Modelli

STMicroelectronics
via Olivetti 2, 20041 Agrate Brianza, Italy

With reference to the mainstream flash EEPROM technology, the most relevant failure mechanisms that affect memory reliability are reviewed, showing the primary role played by tunnel oxide defects. The degradation of device performance and single-bit failures induced by program/erase cycling, like the erratic erase phenomenon, are discussed. The impact of stress-induced leakage current on data retention is shown to limit the tunnel oxide scaling.

1. INTRODUCTION

In the most recent years flash EEPROM has become an important portion of the total semiconductor memory market. Flash memory combines the capability of being electrically erased and re-written, featured also by the more expensive EEPROM, with the single transistor structure of EPROM, allowing the production of high density, low cost memories.

Memory reliability, namely program/erase (p/e) endurance and data retention, is a key issue of flash technology. Strong efforts have been devoted to improve the tunnel oxide quality, which is the single most important factor affecting flash memory reliability, both in terms of intrinsic properties and defect density. At the same time design solutions to make devices less sensitive to oxide defects and methodologies for screening of latent defects at wafer sorting have been developed.

Nowadays, flash memory is widely accepted as an established and reliable technology, but the continuous trend to increase the storage density is driving the technology close to its physical limits and new reliability challenges are met. One of the most important concerns the tunnel

oxide scaling, which is essential for memory performance improvement. The slowness of tunnel oxide scaling for flash processes, which is due to reliability reasons, is pointed out in Figure 1 where it is compared with the scaling trend of gate oxide for logic devices.

In the present paper the failure mechanisms limiting memory endurance and data retention will be discussed. Reference will be done to the mainstream flash technology; a floating-gate cell in a NOR-type array is considered; the cell is programmed with channel hot-electron injection and erased with Fowler-Nordheim tunnelling [1]. In Figure 2 a schematic
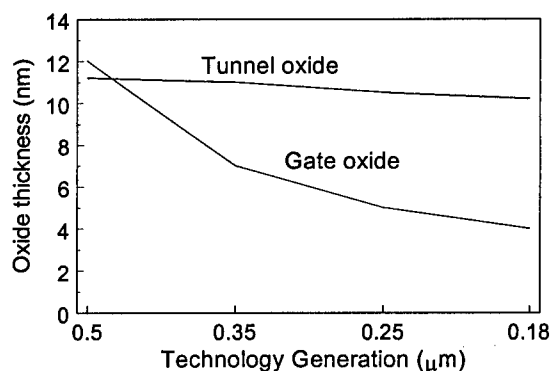
Figure 1. Scaling trends for tunnel oxide of flash and gate oxide of logic devices.
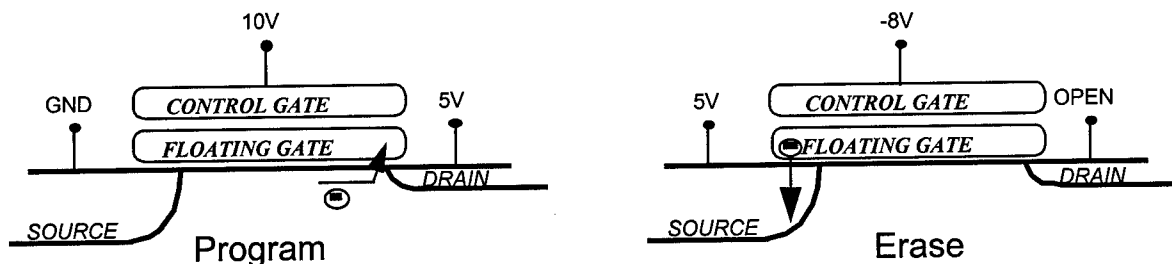
Figure 2. Biasing scheme for programming and erasing a flash memory cell.

drawing of the cell cross section with typical program and erase conditions is reported.

Section 2 is dedicated to failure modes related to the typical cell behavior. The impact on reliability of extrinsic oxide defects, like local oxide thinning or contamination-induced high conductivity spots, will only be mentioned. Sections 3 and 4 deal with single-cell failures, namely erratic erase and single-bit data loss after p/e cycling, which can be both ascribed to a trap-assisted anomalous tunnel oxide conduction.

2. INTRINSIC RELIABILITY

2.1. Program/erase endurance

Endurance of flash memory can be limited by two failure modes: 1) the reduction of p/e efficiency due to oxide aging, which brings to a parametric failure, 2) single-bit failures due to tunnel oxide defects.

The first failure mode is related to a quite uniform and reproducible wear-out of memory cell performance. A typical result is shown in Figure 3. The experiment was performed applying constant amplitude pulses and the variations of programmed and erased V_T levels give a measure of oxide aging.

The evolution of the erased V_T reflects the well-known charge trapping dynamics in the tunnel oxide [2,3]. The initial decrease of erased V_T is due to a pile-up of positive charge, which enhances tunnelling efficiency, while the long term increase is

due to electron trapping and to the generation of negatively charged oxide defects.

The reduction of program efficiency after extended p/e cycling was well explained by S. Yamada et al. [4]; it is attributed to electron trapping and interface state generation at the drain side of the memory cell, a degradation mechanism which is inherent to channel hot electron programming. Interface states reduce electron mean free path, impacting on the hot carrier generation mechanism, and electrons trapped in the tunnel oxide modify the electric field at the injection point, reducing the programming efficiency. A reduction of the cell trasconductance is also observed after extended p/e cycling.

In a flash memory device the V_T window closure is compensated by means of write and verify schemes. The result of a write operation is checked in order to determine whether or not the target V_T is achieved, so as to continue programming or erasing if this is not the case. Erasing and
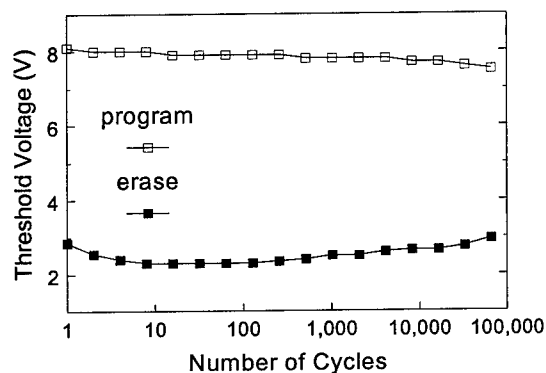
Figure 3. V_T window versus cycle number under constant p/e conditions.

programming times increase with the number of cycles, as is shown in Figure 4, and eventually exceed the specification limits. However, a properly designed memory cell can reach 10^5-10^6 p/e cycles with acceptable performance degradation.

Apart from more sophisticated failure mechanisms like the ones discussed in Section 3, single bit endurance failures can likely be caused by extrinsic defects in tunnel oxide, giving rise to premature oxide breakdown. However, the extreme sensitivity of flash memory to oxide defects makes it possible to detect cells with latent defects at wafer sorting by means of accelerating stress tests.

2.2 Data retention

As any non-volatile memory, flash memories are specified to retain data for over 10 years. The tunnelling current through the oxide sets the fundamental limitation to data retention. The number of electrons stored in the floating gate of a programmed cell is of the order of 10^4 and the floating gate potential in storage condition is between -1 and -2 V. Assuming that the charge lost in 10 years should be less than 20% of the initial charge, the leakage current limit is of the order of 10^{-15} A/cm^2. The minimum oxide thickness is determined by direct tunnelling of electrons and can be estimated to be about 6 nm. In section 4 it will be shown that stress-induced leakage current induced by p/e cycling limits the tunnel oxide scaling to a value well above 6 nm.

Another possible cause of charge loss is mobile ion contamination. Scaling technology and cell size, the demand for high quality processes is becoming more and more severe because of the reduction of cell capacitance and of stored charge. Figure 5 shows a projection to the Gbit generation of stored charge and of the corresponding maximum acceptable concentration of mobile ions in the insulating layers surrounding the memory cell.

Defects in tunnel oxide or in interpoly dielectric could also give rise to excessive leakage current. However, these defects are efficiently screened by an accelerated retention test at 250°C at wafer sorting.

A detailed discussion of data loss mechanisms and process solutions is out of the scope of this presentation; over 20 years of work on EPROM have generated a large literature on the subject and many reviews are available (see ref. 5).

2.3. Read disturb

When a cell is read, a gate voltage and a drain voltage are applied to the selected row and column, respectively. This condition can cause two kind of read disturb on a cell in the erased state.

If we consider the cell that is read, an unwanted programming due to hot electron injection can take place, if the drain voltage is not low enough. For a short channel length and a high channel doping, a

Figure 4. Program and erase times versus the number of cycles for a 512 kbit array.
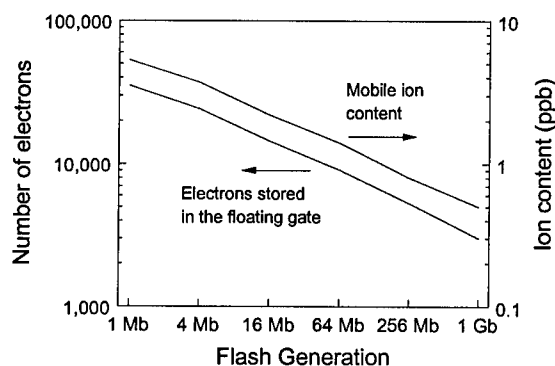
Figure 5. Trend of stored charge and maximum allowable mobile ion density.

significant V_T shift is observed even if the drain voltage is well below the energy barrier for electron injection at the Si/SiO_2 interface, as it is shown in Figure 6.

Moreover, all the cells on the selected row are subjected to a low-voltage gate stress, which can induce electron tunnelling from the channel into the floating gate.

Safety margins can be achieved by a proper cell and circuit designs, mainly in terms of oxide thickness and applied read voltages. A major concern arises when considering the read disturb characteristics after extended p/e cycling, as will be discussed in Section 4.

3. ERRATIC ERASE

3.1. The over-erase problem

In a NOR-type array the source is common to all the cells and the drain of the cells in a column is connected to the same bit line. When reading, the gate of the cells that are not addressed is kept at ground potential and they should be in the off state in order to enable a proper sensing of the addressed cell.

Over-erasing is a potential cause of failure because if the V_T of a cell goes negative all the cells connected to the same bit line would be read as "1" irrespective of their actual content. As all the cells in an array are erased simultaneously, the time required to erase the slowest cell may be

long enough to over-erase the fastest cells. The erase algorithm can only ensure that erasing is stopped as soon as the V_T of all the cells is below a prefixed value, but it cannot prevent fast-erasing cells from being over-erased.

An analysis of the distribution of threshold voltage after electrical erase (Figure 7) shows that the bulk of the distribution, including over 99% of cells, is gaussian; from now on the cells owning to the gaussian distribution will be referred to as "normal" cells. The left side of the distribution is made of an exponential tail of cells that erase faster than the average: these cells will be referred to as "tail" cells.

Understanding the nature of tail cells is of key importance. As these cells erase faster than normal cells with the same applied voltage, one should assume that they are somehow "defective". However they are just too many for being associated to extrinsic defects. Two intrinsic structural imperfections have been invoked to explain the nature of tail cells. Muramatsu and co-workers [6] attributed the presence of a tail in the erase distribution to the polycrystalline structure of the injecting electrode. Dunn and co-workers [7] have modelled the erasing tail as purely due to randomly distributed positive charges in the tunnel oxide. The two models are both probably valid. The impact of polysilicon grain size on erase distribution has been
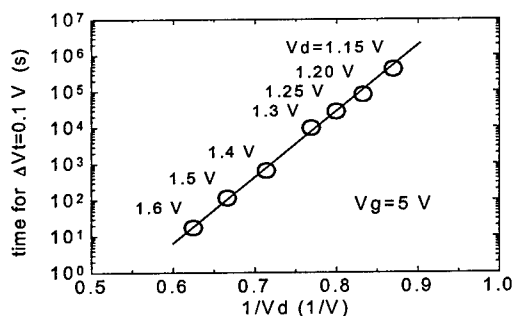
Figure 6. Time for a 0.1 V V_T shift versus $1/V_D$ for a 0.18 μm technology flash cell.
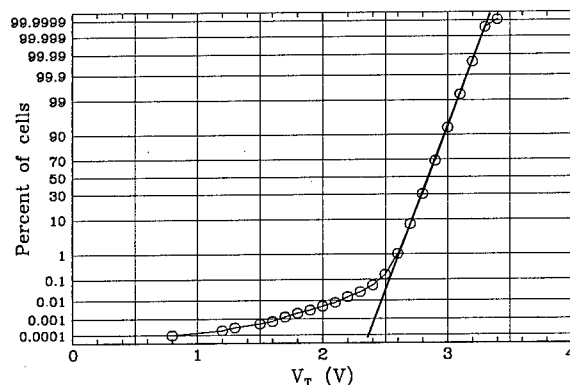
Figure 7. Experimental V_T distribution for a 1 Mbit array (circles) and gaussian fit.

demonstrated, but oxide charges play for sure a role as well. Figure 8 shows the V_T distributions obtained applying a positive voltage to the gate for programming and a negative voltage to the gate for erasing: when the electrons are injected from the substrate the cathode is monocrystalline, but the tail is present as well.

A striking observation concerning this experiment is that cells behaving in an anomalous way for positive gate bias do not show enhanced conduction for negative gate bias and vice versa. With reference to the positive charge model, we need to assume that the charge configuration affecting the electron injection from one interface has no effect on the injection from the other one.

As the exponential tail of the erase distribution is mostly related to structural imperfections, i.e. intrinsic defects, it can be minimized by process optimization but it cannot be eliminated. However, tail cells are as reliable as normal cells for what concern data retention and endurance.

Normal cells and tail cells do not cover the totality of memory cells. Some cells, less than one per million in a "clean" process, erase much faster than normal cells and they stand out also from the tail of the distribution. The physical defects responsible for these "fast-erasing" cells are likely different from the ones of tail cells and they are probably due to extrinsic defects, like, for instance, a local tunnel oxide thinning. Moreover, while tail cells exhibit good data retention, fast-erasing cells often show an enhanced charge loss in data retention tests.

The over-erasing problem should be solved by taking into account the erase distribution in designing the memory and by properly screening at wafer sort the fast erasing cells due to extrinsic defects. Discussing the erratic erasure phenomenon, we will see that this, unfortunately, is not enough.

3.2. The erratic erase phenomenon

One of the most relevant mechanisms of single-bit failure in p/e cycling reported so far is the erratic erase, which has been presented for the first time by T.C. Ong et al. [8] and further investigated by C. Dunn et al. [7]. Erratic cells show an unstable and unpredictable behavior in erasing: their V_T changes randomly from cycle to cycle between two or more distinct values, moving back and forth from the bulk of the V_T distribution to the lowest part of the tail. Erratic cells can cause over-erasing failures in the field. An example of erratic erase is shown in Figure 9.

Such a behavior has been attributed to hole trapping in the tunnel oxide: the statistical distribution of hole traps gives an extremely low but finite probability of having clusters of two or more positive charges, whose electric fields overlap each other to produce a huge local increase of the tunnel current. Trapping/detrapping of an individual positive charge causes, in this condition, a detectable change in erasing

Figure 8. V_T distributions obtained by Fowler-Nordheim program and erase.
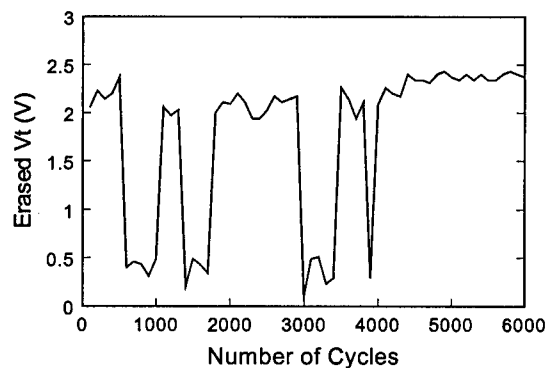
Figure 9. V_T after erase versus the cycle number for a cell exhibiting erratic erase.

speed. The model has been confirmed by tunnelling current calculations [7,8]; the current density increase caused by three elementary positive charges can be up to 5 orders of magnitude, depending on their distance from the injecting electrode. An experimental confirmation of the described model is reported in Figure 10, where the erasing curves, i.e. the cell V_T versus erasing time, are reported for an erratic cell measured in three consecutive p/e cycles. In cycle 3 it behaves as a normal cell and at the end of erasing its V_T reaches a value typical of the bulk of the erased distribution. In cycle 4 the cell starts erasing as in the previous cycle, then the erase rate suddenly increases with an evident discontinuity, implying a strong increase of the tunnelling current; at the end of this cycle the cell is over-erased and its V_T is more than 2 V below that reached at the end of the previous cycle. In the third cycle the cell again is erased at a very fast rate.

As the erratic behavior is due to statistical fluctuations of intrinsic oxide defects, the occurrence of erratic cells can be reduced by process optimization, but cannot be completely avoided. Design solutions have been developed to solve the problem at the circuit level by introducing a reprogramming step at the end of the erase algorithm to recover over-erased cells.

4. DATA RETENTION AFTER CYCLING

A high-field stress on thin oxide is known to increase the low-field leakage level; the excess current component, which causes a significant deviation of current-voltage characteristics from the theoretical tunnelling conduction at low field, is known as stress-induced leakage current (SILC). SILC is clearly related to stress-induced oxide defects and, as far as conduction mechanism, it is attributed to trap-assisted tunnelling [9,10]. For fixed stress conditions the SILC increases strongly with decreasing oxide thickness below 10 nm.

The SILC can be observed in a flash cell as an enhanced sensitivity to low-voltage gate stress after p/e cycling. The major concern does not come from a normal cell but from the tail of the distribution, as is shown in Figure 11, where the effects of a low-voltage gate stress (8 V for 64 hrs) on a 1 Mbit array before and after cycling are compared. The different magnitude of the V_T shift for different cells may be explained again by the random spatial distribution of oxide traps responsible of the tunnelling current enhancement.

The impact of SILC on read disturb and data retention is strongly dependent on tunnel oxide thickness. For very thin tunnel oxide (below 8 nm), SILC is not negligible even at electric field values as low as the ones typical of reading or data storage

Figure 10. V_T of an erratic cell as a function of the erase time during three consecutitive cycles.
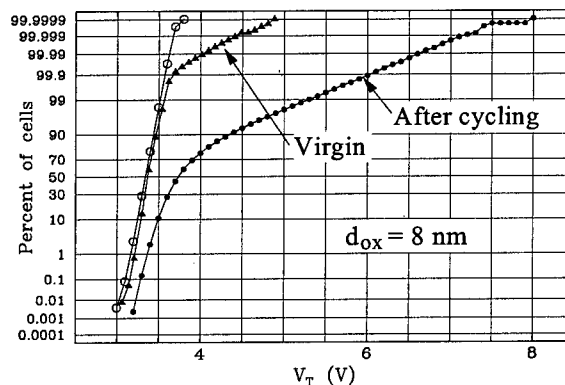
Figure 11. Effect of a gate stress on a virgin sample and after 10^5 cycles. The V_T distribution before stress is also shown.

conditions [11].

From SILC data obtained on capacitors, E.F. Runnion et al. [10] have shown that the leakage current of tunnel oxides in the 5-8 nm range after a stress equivalent to 10^5 cycles largely exceeds the requirements for 10 year data retention. S. Yamada et al. [12] have shown that on a memory device the situation can be worse because of single cells that exhibit a much higher, anomalous, SILC.

For setting the minimum reliable thickness of tunnel oxide, we cannot refer to the typical cell but we must give margin for the tail of the distribution. Figure 12 shows the results of a room temperature retention test on a 1 Mbit array with an 8-nm thick tunnel oxide after 10^5 p/e cycles: while almost the totality of the array does not present any detectable threshold shift, there is a tail of cell that lose charge. For increasing oxide thickness, in the range from 8 to 10 nm, both the number of bits in the tail and the maximum shift decrease. The V_T tail is no more observed for oxides thicker than about 10 nm, even if the retention test is prolonged to the year time scale. In Figure 13 it is shown the result of a test similar to that of Figure 12, but with an oxide thickness of 9 nm: after 250 days only very few cells have a V_T below the bulk of the distribution.

Retention tests performed at different temperatures indicate that the defect responsible of the charge loss is annealed or neutralized in a short time for temperatures higher than 150°C. In this case the tail of the V_T distribution is no more observed and the uniform V_T shift is only slightly larger than that observed on virgin samples. Even at room temperature we observed some leaky cells that suddenly stop to lose charge, a behavior similar to the erratic bit phenomenon. Several examples of erratic behavior of the anomalous SILC in flash cells are reported in the literature [12,13].

An experimental evidence of the role played by trapped holes in the anomalous SILC phenomenon is given by a set of experiments where similar samples were cycled using different erase conditions: the gate and source bias were adjusted in order to have the same erase speed, and thus the same electric field in the tunnel oxide, but with increasing source voltage.

The subsequent retention test showed that samples cycled with high source voltage are more affected by SILC, in agreement with an increased hot hole generation and injection at the source junction [14].

A proper cell structure and erase condition optimization allows to reduce the SILC for a fixed number of p/e cycles and oxide thickness. Moreover, at the device level, the V_T window and the reading condition can be chosen in order to minimize the electric field in the tunnel
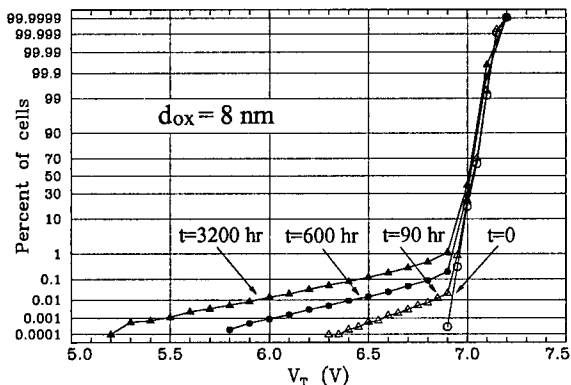
Figure 12. V_T distribution at different room temperature storage times after 10^5 cycles for an 8-nm thick oxide.
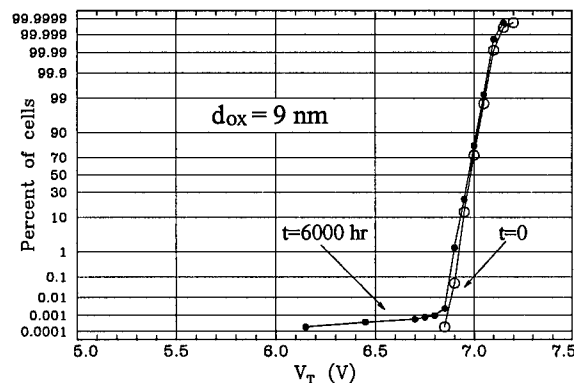
Figure 13. V_T distribution after 250 days at room temperature for an array with a 9-nm thick oxide cycled 10^5 times.

oxide during storage and reading. However, tunnel oxide scaling below 10 nm will be probably possible only in combination with error correction techniques at the device or system level.

5. CONCLUSION

The success of flash memories in the semiconductor market has grown together with the understanding of their reliability issues. Now, flash memory is widely accepted as an established and reliable technology.

Intrinsic degradation mechanisms, responsible for the wear-out of device performances in p/e cycling, are fairly well dominated; through a proper optimization of cell architecture, their impact can be minimized to push the endurance limit into the 10^5-10^6 range. Much more critical are single-bit failures induced by p/e cycling. A quite tricky failure mechanism such as the erratic erase has been identified and solutions have been addressed both at process and at circuit level. Data retention after cycling is the issue that definitely limits the tunnel oxide thickness scaling. For oxides thinner than 8 nm the number of leaky cells becomes so large that even an error correction technique can hardly fix the problem.

REFERENCES

1. P. Pavan, R. Bez, P. Olivo, and E. Zanoni, Proc. IEEE, 85 (1997) 1248.
2. P. Olivo, B. Ricco', and E. Sangiorgi, J. Appl. Phys., 54 (1983) 5267.
3. D.J. DiMaria, E. Cartier, and D. Arnold, J. Appl. Phys., 73 (1993) 3367.
4. S. Yamada, Y. Hiura, T. Yamane, K. Amemiya, Y. Ohshima, and K. Yoshikawa, IEDM Tech. Dig., (1993) 23.
5. G. Crisenza, C. Clementi, G. Ghidini, and M. Tosi, Quality and Reliability International, 8 (1992) 177.
6. S. Muramatsu, T. Kubota, N. Nishio, H. Shirai, M. Matsuo, N. Kodama, M. Horikawa, S. Saito, K. Arai, and T. Okazawa, IEDM Tech. Dig., (1994) 847.
7. C. Dunn, C. Kay, T. Lewis, T. Strauss, J. Schreck, P. Hefley, M. Middendorf, and T. San, Proc. Int. Rel. Phys. Symp., (1994) 299.
8. T.C. Ong, A. Fazio, N. Mielke, S. Pan, N. Righos, G. Atwood, and S. Lai, VLSI Symp. on Tech., 7A-2 (1993) 83.
9. D.J. DiMaria and E. Cartier, J. Appl. Phys., 78 (1995) 3883.
10. E.F. Runnion, S.M. Gladstone, R.S. Scott, D.J. Dumin, L. Lie, and J. Mitros, IEEE Trans. Electron Devices, 44 (1997) 993.
11. M. Kato, N. Miyamoto, H. Kume, A. Satoh, T. Adachi, M. Ushiyama, and K. Kimura, IEDM Tech. Dig. (1994) 45.
12. S. Yamada, K. Amemiya, T. Yamane, H. Hazama, and K. Hashimoto, Proc. Int. Rel. Phys. Symp., (1996) 108.
13. F. Arai, T. Maruyama, and R. Shirota, Proc. Int. Rel. Phys. Symp., (1998) 378.
14. K. Tamer San, Cetin Kaya, and T.P. Ma, IEEE Trans. Electron Devices, 42 (1995) 150.

Charge Trapping in ONO Interpoly Dielectric of FLOTOX EEPROM Cells

E. Brücklmeier, A. Kux, R. Kakoschke, and H. Palm

Infineon Technologies AG
P.O. Box 80 09 49, D-81609 Munich, Germany

The threshold voltage loss of erased FLOTOX EEPROM cells with various ONO interpoly dielectric layers has been studied as a function of bake time and programming pulse amplitude. The results show a displacement of electrons injected into the nitride layer to be the root cause of the threshold voltage lowering. A quantitative model has been used for optimizing the cell layer construction as well as the cell operating conditions with respect to a minimized programming voltage at a maximum data retention.

Introduction

One of the major issues in EEPROM and Flash development is a thorough understanding of the mechanisms of threshold voltage shifts during cell lifetime. Commonly discussed models assume back-tunneling of electrons from the floating gate through the tunnel oxide as the main reason for loss of threshold voltage[1]. In contrast to these back tunneling models our results propose charge trapping in the interpoly dielectric to be the root cause of decreasing threshold voltage observed in our experiments. A detailed understanding of this phenomenon allows to optimize the operation conditions of EEPROM and Flash cells as well as the composition of the ONO interpoly dielectric necessary for scaling minimum feature sizes down to 0.25µm and below.

Experiments

For characterization we use FLOTOX EEPROM cells (Fig. 1) with different layer thicknesses of the bottom oxide and the nitride of the ONO interpoly dielectrics. In our experiments we measure (a) the threshold voltage (V_{TH}) as a function of bake time at fixed 250°C bake temperature (b) the drain current at a fixed gate voltage as a function of time after applying erase pulses with amplitudes up to 30V.

Results

The results of measuring the threshold voltage as function of bake time are: (1) The cells erased with a pulse amplitude of 17V show a much stronger loss in threshold voltage than those erased with 14V. (2) We do not reach a unique threshold voltage level after programming/erasing with different pulse amplitudes of 14V and 17V even after bake times of

more than 100 hours at 250°C (Fig. 2). (3) The thicknesses of the bottom oxide and the nitride layer (not shown) of the ONO layer strongly influences the lowering of the threshold voltage (Fig. 3). From additional experiments varying the bake temperature we know that the V_{TH} loss is strongly temperature dependent[2].

The second type of experiments is using high programming pulse amplitudes up to 30V with the aim to have variable and high electric fields accelerating the threshold voltage loss. Since the control gate voltage during threshold voltage measurement affects the result by an undefined electric field across the ONO we use a non-standard method monitoring the drain current as a function of time at fixed control gate voltage (Fig. 4). From this measurement we can obtain two results: (1) the drain current at a constant gate voltage increases in magnitudes within a few hundred seconds. (2) We see a decrease of V_{TH} as described in the following: The cell programmed with 16V has a threshold voltage far below 7.25V therefore the channel is fully conductive at V_{CG} = 7.25V. For higher programming amplitudes V_{TH} rises and as a consequence the drain current decreases. From $V_{PP} \approx$ 20V on V_{TH} decreases again. By observing the drain current at fixed times we can deduce a threshold voltage as a function of pulse amplitude. It shows a time and pulse amplitude dependent decrease of threshold voltage above 16V pulse amplitude voltage (Fig. 5).

We also have repeated the drain current monitoring with different gate voltages. As a result we see that a higher electric field across the ONO does not slow down the temporal increase of the cell drain current as it would be for back tunneling. If we extrapolate

the 7.25V current-time characteristic into the same time slot as the 6.25V cell we see the same rate. The fact that the drain current is about 7.5 magnitudes higher is just due to the difference in gate voltage of 1V(Fig 6). We can also accelerate the current change rate by elevating the temperature(Fig. 7). This disproves Fowler-Nordheim backtunneling as the reason of the V_{TH} loss.

The results of both types of experiment are consistent in a sense that elevated temperature as well as enhanced electric field across the ONO accelerates V_{TH} loss. Accordingly we are observing the same mechanism in both cases. We note that the programming pulse experiment is working with pulse amplitudes far beyond the usual operating conditions of the cell, however it allows a much faster characterization of the dominant origin of threshold voltage loss than bake experiments do.

Discussion and conclusion

Commonly discussed models of the V_{TH} loss of unstressed cells are based on back tunneling of electrons from the floating gate through the tunnel oxide. In contradiction to these back tunneling models, our temperature and electrical field dependent results indicate charge trapping in the ONO to be the key mechanism of V_{TH} loss[3]. Fig. 8 illustrates the electron injection into the nitride: in the capacitive region (I) V_{FG} follows V_{PP} with the cell capacitive coupling ratio α as proportionality factor and causes a voltage drop across the ONO V_{ONO}. In region (II) Fowler-Nordheim-tunneling gets in and fixes V_{FG} at a certain level and as a consequence the floating gate starts to charge and V_{ONO} raises faster. From a critical programming voltage V_{CRIT} on we additionally inject electrons through the bottom oxide into the nitride of the interpoly dielectric (III), hence the cell charge splits into Q_{ONO} and Q_{FG}. This results in a reduced increase of the floating gate charge from this pulse amplitude on.

As soon as Q_{ONO} starts to displace towards the top oxide under the influence of electric fields and temperature, its contribution to V_{TH} will decrease and lead to threshold voltage loss (Fig. 9).

This critical programming voltage for charge injection into the ONO mainly depends on bottom oxide thickness and capacitive coupling ratio of the cell. It also defines the optimum cell operating condition: All tunneled electrons should be kept on

the floating gate at a maximum possible charge of the floating gate. Based on this model an optimum design of layer thicknesses and operating conditions has been specified and implemented (Fig. 10). Starting from a common V_{TH} before bake we have a negligible V_{TH} - loss for the optimized cell while we observe a V_{TH} - loss for the reference cell. The optimum cell design is defined by proper choice of bottom oxide, nitride and tunnel oxide thickness as well as the cell layout.

References

[1] M. Lanzoni et al., Evaluation of EEPROM data retention by field acceleration, Quality and relyability engineering international Vol. 7, pp 293-297, 1991

[2] U. Feldmann et al., Concurrent Technology and Circuit Development for EEPROMs, Proceedings Asia & S. Pacific Design Automation Conf. 1998

[3] S. Mori et al., Thickness Scaling Limitation Factors of ONO ..., IEEE Trans. Electron Devices, Vol.43, No. 1, Jan 1996, pp. 47 - 53
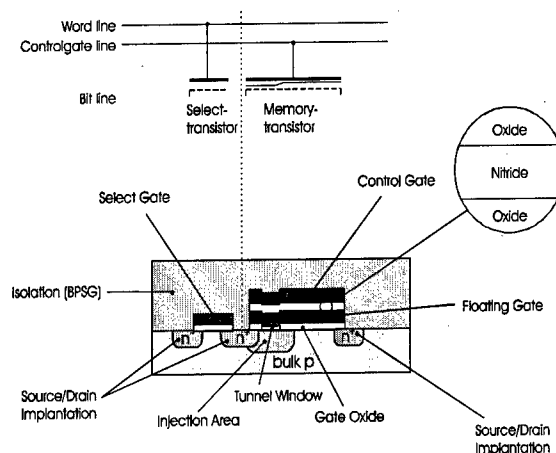
Fig. 1 Schematic cross section of FLOTOX-EEPROM-cell.
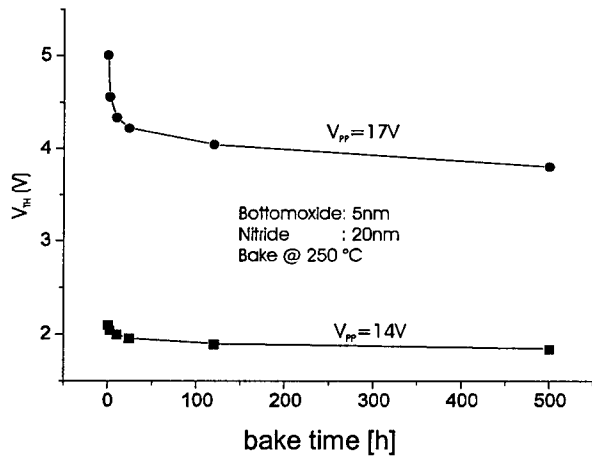
Fig. 2 Decrease of cell threshold voltage V_{TH} as a function of bake time at 250°C for erase voltages of 14V and 17V applied to control gate.

Fig. 3 Decrease of threshold voltage V_{TH} as a function of bake time for oxides thicknesses of 5nm and 6nm. Programming voltage amplitude is 17V on control gate

Fig. 4 Drain current of memory transistor as a function of time for various programming pulse amplitudes at fixed control gate voltage of 7.25V.
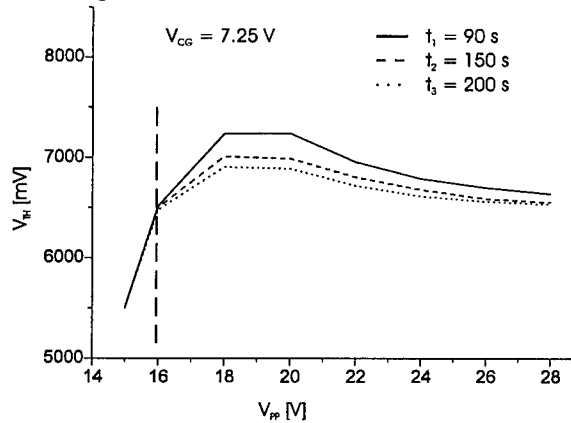
Fig. 5 Threshold voltage as a function of programming pulse amplitude and time at a fixed $V_{CG} = 7.25V$. This function is calculated from the drain current (Fig. 4)

Fig. 6 Drain current of memory transistor as a function of time for control gate voltages of 6.25V and 7.25V after a programming pulse of 30V. $V_{CG} = 7.25V$ corresponds to a larger electrical field across the ONO and a reduced field across the tunnel window.
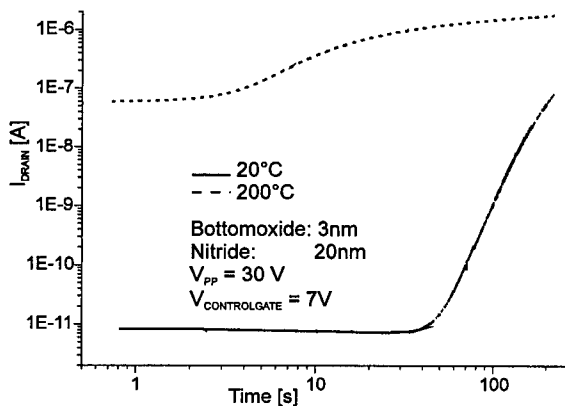
Fig. 7 Drain current of the cell as a function of time for temperatures of 20°C and 200°C after a programming pulse of 30V. At a fixed $V_{CG} = 7V$.
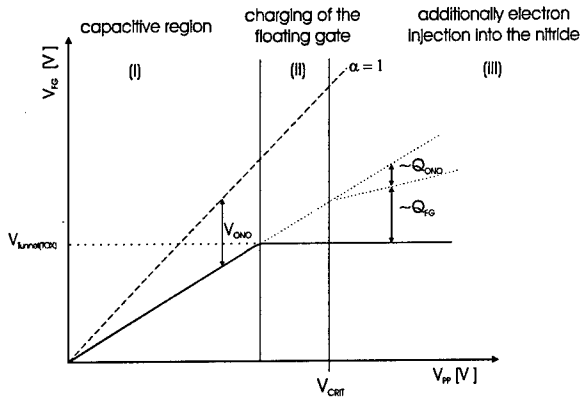
Fig. 8 Schematic behavior of the floating gate potential V_{FG} as a function of the programming pulse amplitude V_{PP}.
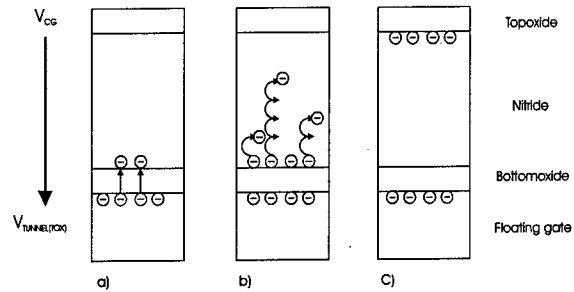
Fig. 9 Injection and displacement of electrons through the ONO interpoly dielectric. a) Fowler-Nordheim-tunneling through the bottom oxide b) beginning of electron movment by Poole-Frenkel-hopping during the bake c) final position of the electrons - the V_{TH} loss stops
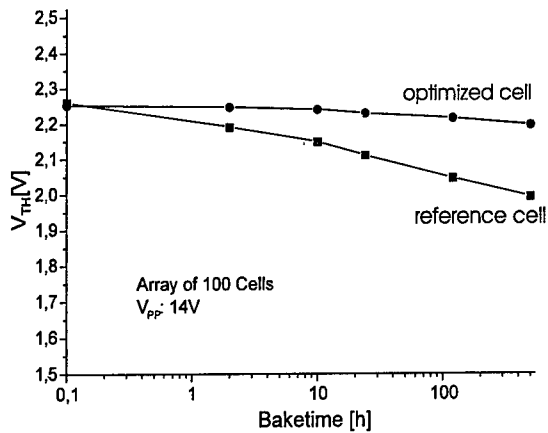
Fig. 10 Threshold voltage loss of optimized and nonoptimized (reference) cell.

SILC in MOS Capacitors with Poly-Si and Poly-Si$_{0.7}$Ge$_{0.3}$ Gate Material

V.E. Houtsma[a], J. Holleman[a], C. Salm[a] and P. H. Woerlee[a,b]

[a]MESA Research Institute, University of Twente, P.O.Box 217, 7500 AE Enschede, The Netherlands.

[b]Philips research Laboratories, Prof. Holstlaan 4, Eindhoven, The Netherlands.

In this paper the DC-SILC characteristics of n$^+$ and p$^+$ poly-Si and poly-SiGe MOS capacitors are studied for substrate($+V_g$) and gate-injection($-V_g$) conditions. P$^+$ and n$^+$-gates with poly silicon (poly-Si) and poly Silicon-Germanium (poly Si$_{-0.7}$Ge$_{0.3}$) were used to study the influence of the gate workfunction on gate current and SILC currents. For n$^+$ poly-SiGe, reduced poly depletion and no significant difference in SILC characteristics compared to n$^+$ poly-Si gate devices is observed. For p$^+$ gate devices asymmetric SILC and reduced SILC for poly-SiGe is observed.

1. Introduction

Stress Induced Leakage Current (SILC) limits the scaling of the gate oxide thickness for non-volatile memories. SILC was studied mainly for n-doped poly silicon (poly-Si) gate material up till now [1,2]. Recently, for p$^+$-gate devices a different conduction mechanism under gate injection conditions ($-V_g$) [3–5], asymmetric SILC (gate bias polarity) and reduced SILC for p$^+$ poly-Si$_{0.7}$Ge$_{0.3}$ was observed [5]. However, the impact of poly-Si$_{1-x}$Ge$_x$ on the SILC characteristics of n$^+$-gate devices has not been studied yet. In this paper the DC-SILC characteristics of n$^+$ and p$^+$ poly-Si and poly-SiGe MOS capacitors are studied for substrate($+V_g$) and gate-injection($-V_g$) conditions and a detailed comparison is made.

2. Experimental Procedures

PMOS and NMOS capacitors were fabricated on 10 Ω-cm n-type and p-type silicon substrates, respectively. A high quality gate oxide with 5.6 nm thickness (ellipsometric) was grown in diluted dry oxygen. Undoped poly-Si and poly-Si$_{0.7}$Ge$_{0.3}$ layers (200 nm thick) were deposited using a LPCVD system. Gate doping was done by a 20 keV, $2.5 \cdot 10^{15}$ cm^{-2} BF$_2^+$ or 40 keV, $5.0 \cdot 10^{15}$ cm^{-2} As$^+$ implant followed by an anneal at 850°C for 30 minutes (p$^+$-gate) or a rapid thermal anneal (RTA) at 1000°C for 20 seconds in N$_2$ ambient (n$^+$-gate). Electrical stress has been applied on A$=4.0 \cdot 10^{-4}$ cm^{-2} (p$^+$-gate) and A$=3.53 \cdot 10^{-3}$ cm^{-2} (n$^+$-gate) MOS capacitors using constant current stress conditions of J$_{stress}=\pm 0.1$ mA/cm^2. DC-SILC was measured in a similar way as in [1]. The oxide electric field was determined by integration of the quasi-static capacitance-voltage curves as in [5].

3. Experimental results

Quasi-Static C-V curves for capacitors with p$^+$-poly Si and poly-SiGe gate are shown in Fig. 1. A shift of the flatband voltage from 0.82V (poly-Si) to 0.60V (poly-SiGe) agrees well with previous data [3,6]. It is caused by a shift in valence band position [6]. Note that the gate depletion is acceptable for both devices. Fig. 2 shows the C-V curves for capacitors with n$^+$-poly Si and poly-SiGe gate material. Note that since for n$^+$ gate devices the workfunction difference is negligible between poly-Si and poly-SiGe, the flatband voltage is the same (-0.98V). Also it is evident that the gate depletion of the poly-SiGe devices is significantly reduced compared to the poly-Si reference devices, as was also found in [7].

Fig. 3 shows the J-E$_{ox}$ characteristics of unstressed poly-Si and poly-SiGe capacitors for substrate- ($+V_g$) and gate-injection ($-V_g$) conditions. For n$^+$-gate devices the J-E$_{ox}$ characteristics are symmetric with -V$_g$ and +V$_g$ injection conditions. However, for p$^+$-gate devices,
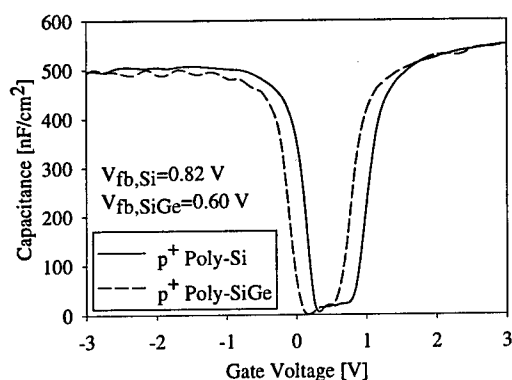
Figure 1. Quasi-static C-V curves for p^+-poly Si and p^+- poly SiGe MOS capacitors. Dashed line poly-SiGe, solid line poly-Si.
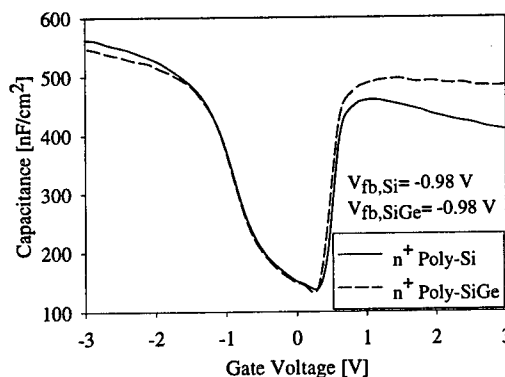
Figure 2. Quasi-static C-V curves for n^+-poly Si and n^+- poly SiGe MOS capacitors. Dashed line poly-SiGe, solid line poly-Si.

the onset of conduction for $-V_g$ occurs at significantly higher oxide field ($\approx 1.5\times$) than for $+V_g$ condition. Note the difference for p^+-gate devices between poly-Si and poly-SiGe at $-V_g$. For $+V_g$ the J-E_{ox} curve can be fitted with the standard Fowler-Nordheim (FN) expression with barrier height of 3.2 eV. For $-V_g$ a fit with a FN expression for p^+-gate devices can not be obtained, unless unphysical fit parameters are used. The gate, substrate and diode currents measured on gate-controlled diodes are shown in Fig. 4.

For $-V_g$ conditions the substrate and diode currents are larger than the gate current (right axis Fig 4, $I_{sub} > 2\times I_{gate}$). The larger substrate current indicates that an avalanche process is taking place. This is not expected when hole tunneling from the substrate dominates the gate current. It appears that gate current at $-V_g$ for p^+ gate devices is caused by electron injection from the gate. There are two possible mechanisms for electron injection from the gate. Firstly, tunneling of minority carriers(MCT) from the conduction band could occur for low active gate doping, since gate depletion increases the electron surface concentration. A second possibility is valence band tunneling (VBT). Gate currents are either influenced by the workfunction for valence band tunneling or

the bandgap of the gate material for MCT mechanism. This will influence the SILC characteristics, stress experiments are discussed below.

In Fig. 5 the J-E_{ox} characteristics of n^+ poly-Si and poly-SiGe gate devices are shown for $-V_g$ and $+V_g$ after various stress intervals. From this it can be observed that no significant difference between n^+ poly-Si and poly-SiGe gate devices is observed for $+V_g$ and $-V_g$ injection conditions. Both SILC currents can be described well by a FN expression with an effective barrier of 0.9 eV [1,2].

In Fig. 6 the J-E_{ox} characteristics of p^+ poly-Si and poly-SiGe gate devices are shown after various stress intervals. For $-V_g$ injection the SILC becomes apparent at much higher E_{ox}, furthermore the field dependence is different from that at $+V_g$. At comparable oxide field the SILC is orders of magnitude smaller for $-V_g$ stress than for $+V_g$ stress. Hence there is a strong asymmetry of the SILC currents for p^+ gate devices with stress polarity. This is different from n^+-poly gate devices where SILC is almost symmetric with gate polarity. For $+V_g$ the SILC current after stress follows a FN dependence with effective barrier height of 0.9 eV, which is similar to the n^+-poly devices. For gate injection SILC can not be described by an FN expression with fixed bar-
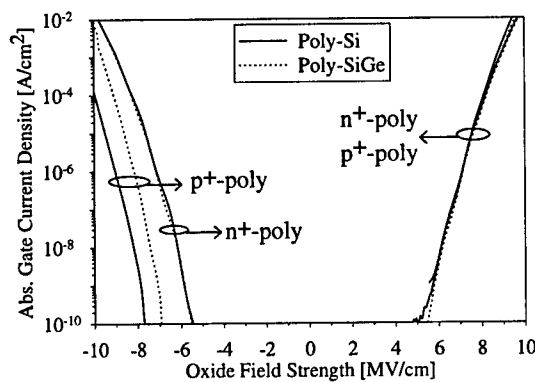
Figure 3. J-E_{ox} measurements of unstressed poly-Si and poly-SiGe gate MOS capacitors. Dashed line poly-SiGe, solid line poly-Si.

Figure 4. Gate, substrate and diode currents of p^+-poly Si gate-controlled diodes with 7nm oxide thickness as a function of gate voltage.

rier height, variations in shape occur with stress time. An (unphysical) FN fit of the SILC current results in a barrier height around 2 eV indicating a much stronger field dependence. It has been observed that SILC is proportional to the neutral trap density created during stress [1,2] and models based on inelastic trap-assisted-tunneling (TAT) are used to describe SILC [8]. The observed bias asymmetry in SILC for p^+-gate devices could be related to a larger tunneling barrier height for TAT or different position of the traps for $-V_g$ stress conditions. For $+V_g$ conditions the SILC current for p^+ poly-SiGe gates is strongly reduced compared to the p^+ poly-Si reference devices. The reduced SILC of poly-SiGe devices for $+V_g$ stress could be the result of a reduced Boron incorporation in the gate oxide. A larger solid solubility and smaller diffusivety of Boron in poly-SiGe lead to a reduced incorporation of Boron in the dielectric [3]. Hence a lower neutral trap density and reduced SILC for poly-SiGe gates is expected [1,2,9].

4. Conclusions

In summary, the DC-SILC characteristics of n^+ and p^+ poly-Si and poly-SiGe MOS capacitors were studied under substrate($+V_g$) and gate-

injection($-V_g$) conditions. For n^+ poly-SiGe no significant difference in SILC characteristics compared to n^+ poly-Si is observed. For p^+ gate devices, both the workfunction of the gate material and Boron penetration into the gate oxide strongly influence the SILC effect. Boron-doped poly-SiGe may be a very interesting gate material due to low SILC at $+V_g$ and better gate oxide quality.

REFERENCES

1. J. De Blauwe *et al.* "A new quantitative model to predict SILC-related disturb characteristics in Flash E^{2}PROM devices", IEDM Tech.Digest, 1996 pp. 343-346.
2. D. J. DiMaria *et al.*, "Mechanism for stress-induced leakage currents in thin silicon dioxide films", J. Appl. Phys., no. 6(78) 1995 pp. 3883-3894.
3. C. Salm *et al.* "Gate Current and Oxide Reliability in p^+ Poly MOS Capacitors with Poly-Si and Poly-Ge$_{0.3}$Si$_{0.7}$ Gate Material", Electron Dev. Lett. vol 19(7) 1998 pp. 213-215.
4. J. -L. Ogier *et al.* "On the Polarity Dependence of Oxide Breakdown in MOS-Devices with n^+ and p^+ Polysilicon Gate", Proc. ESS-DERC'96, 1996 pp. 763-766.
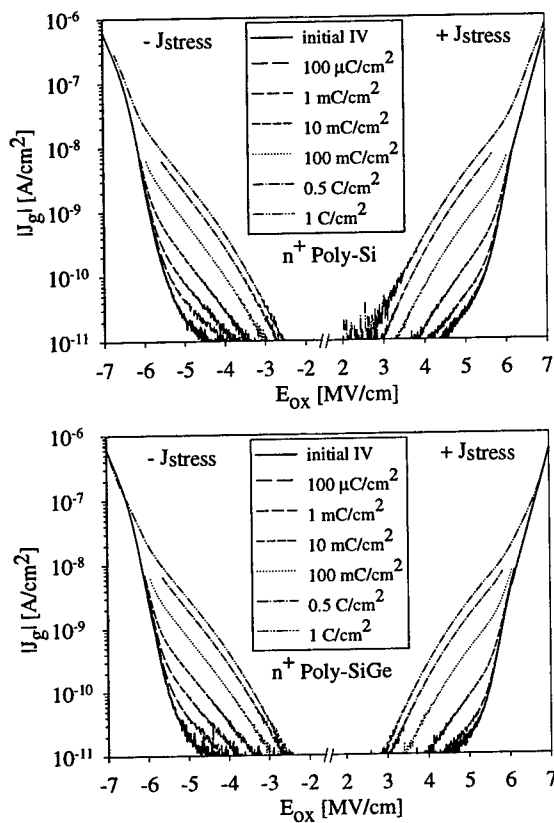
Figure 5. J-E_{ox} curves before and after stress ($100\mu C/cm^2$ to $1C/cm^2$) for n^+ poly-Si and poly-SiGe gate material
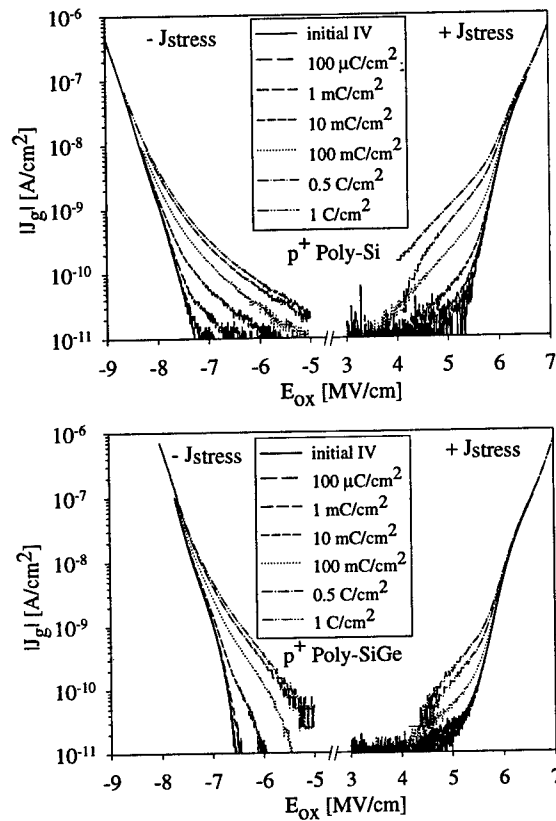
Figure 6. J-E_{ox} curves before and after stress ($100\mu C/cm^2$ to $1C/cm^2$) for p^+ poly-Si and poly-SiGe gate material

5. V.E. Houtsma *et al.* "Stress-Induced Leakage Current in p^+ Poly MOS Capacitors with Poly-Si and Poly-Si$_{0.7}$Ge$_{0.3}$ Gate Material", accepted for publication in Electron Dev. Lett.

6. T. -J. King *et al.* "A Polycrystalline-Si$_{1-x}$Ge$_x$-Gate CMOS Technology", IEDM Tech.Digest, 1990 pp. 253-256.

7. W. -C. Lee *et al.* "Observation of Reduced Poly-Gate Depletion Effect for Poly-Si$_{0.8}$Ge$_{0.2}$-Gated NMOS Devices", Elec. and Solid-State Lett. vol 1(1) 1998 pp. 558-59.

8. A. I. Chou *et al.* "Modeling of stress-induced leakage current in ultrathin oxides with the trap-assisted tunneling mechanism", Appl. Phys. Lett vol 70(25) 1997 pp. 3407-3409.

9. C. Jahan *et al.* "Investigation of Stress-Induced Leakage Current in CMOS Structures with Ultra-Thin Gate Dielectrics", Microelec. Reliab. vol 37(10/11) 1997 pp. 1529-1532.

The impact of SILC to data retention in sub-half-micron Embedded EEPROMs

G. Tao, R. Cöp, J. Dijkstra, F. Kuper and R. Verhaar

Philips semiconductors, Gerstweg 2, 6534 AE Nijmegen, the Netherlands

Tel: +31-24-353-4549, Fax: +31-24-353-2274, Guoqiao.Tao@nym.sc.philips.com

Abstract: Stress-Induced-Leakage-Current (SILC) in sub-half-micron technology embedded EEPROM has been extensively studied by performing gate stress measurements on pre-conditioned EEPROM arrays. This paper reports that the SILC path is very localized; the SILC path can be turned "on" and "off" statistically independent on the electrical field, as shown by the experiments; the SILC caused Vt shift can be fitted into a power-law against the number of programming/erase cycles; the measured current-voltage characteristics of SILC can be well described with Poole-Frenkel model; and EEPROM retention time prediction can be made by extrapolating the gate stress data with fitted Poole-Frenkel parameters.

Keywords: Embedded NVM, SILC, data retention, build-in reliability

1. INTRODUCTION

It has been widely recognized that stress induced leakage current (SILC) is now a major limiting factor for down scaling the tunnel oxide thickness in non-volatile memories[1]. However, it is not clear yet how much SILC can be tolerated without jeopardizing the product data retention. Recently, it has been reported that SILC path is very localized [2] and measurements on capacitors only reveal the average current density. Measurement on large number of non-volatile memory cells gives more information about extreme SILC bits, which can be much larger than the average value as obtained from capacitor measurements. It is these extreme SILC bits which determines the NVM product retention time.
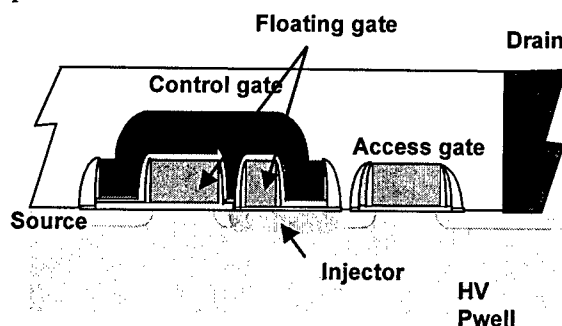
Fig.1 A Schematic cross-section of the EEPROM cell used in this experiment.

After series of experiments carried out on large EEPROM array structures, it has been experimentally observed that SILC has a strong statistical behavior, the leakage path can be "switched" on and off randomly from time to time. For the SILC cells in the "on" state, it has been found that Poole-Frenkel model

fits the experimental data and can be easily applied to predict product retention time caused by SILC in a realistic way. In this paper, experimental results on a large number of EEPROM cells are reported over the SILC dependence on tunnel oxide thickness and stress levels. Data retention time predictions for the worst case cells are also reported based on extrapolation of the Poole-Frenkel model fit.

2. EXPERIMENTAL

256x256 EEPROM cell arrays fully embedded in a 0.35um CMOS process [3] have been employed for studying their SILC behavior. Figure 1 shows a schematic cross section of the cell. During erase, a pulse is applied to the control gate, and electrons tunnel from injector region to the floating gate. During programming, a pulse is applied to the drain via the access gate to the injector and electrons are pulled out of the floating gate. These cell arrays were pre-stressed up to 1 million erase/program cycles, and gate stress measurements have been carried out at several stress levels (2V, 3V and 5V) for tunnel oxide thickness splits ranging from 7.3nm to 8.3nm. The SILC current versus electrical field relation was extracted from the gate stress results and fitted to single trap-assisted tunneling (TAT) [4], Poole-Frenkel (PF) [5] and Fowler-Nordheim (FN) [5] models. The distribution of these SILC bits has been determined directly from the measurements. The fitted parameters for some of the worst-case cells have been used to perform retention time predictions for products with a certain memory size without redundancy. Such data is important for the choice of effective error correction scheme.

G. Tao et al. / Microelectronic Engineering 48 (1999) 419–422

3. RESULTS

3.1. Cell Vt evolution under gate stress

Figure 2 shows a typical Vt envelope evolution of 65536 EEPROM cells in an array under gate stress. The gate stress condition in this case was 3V, the EEPROM array has seen 1 million erase/programming cycles before this gate stress measurement. From the big difference between the maximum value and median value, it can be seen there is a big difference between the average SILC cells (median) and the extreme SILC cells (Maximum).
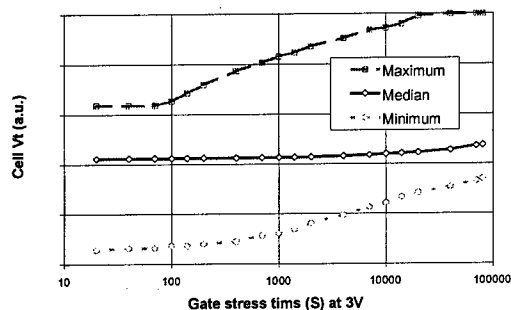
Figure 2. A typical Vt envelope evolution of 65536 cells under gate stress.

3.2. The SILC dependence on the number of erase/program cycles

Figure 3 shows a typical median Vt (of 65536 cells) shift as a function of pre-cycle numbers. The cells had a tunnel oxide thickness of 73Å. This data can be easily fitted with a power law. For worst case bits and/or samples with other tunnel oxide thickness, the same method can be applied, but the fit parameters become different.
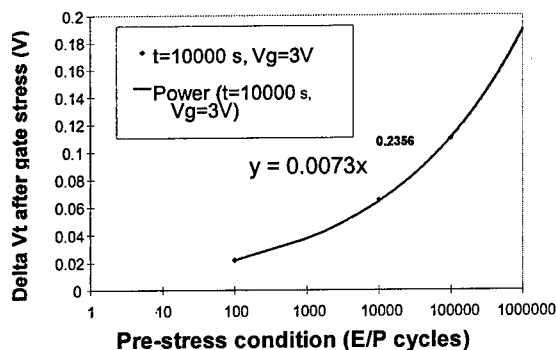
Fig. 3. Vt shift after 10k seconds gate stress at 3V for a typical cell (median cell in an array) as a function of number of erase/programming cycles.
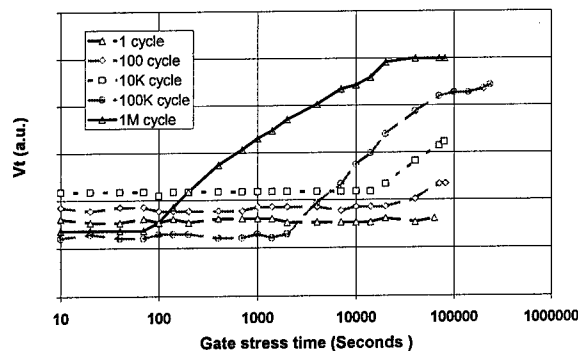
Figure 4. The higher the number of pre-erase/programming cycles, the faster the worst case cell Vt (maximum of 65536 cells) changes. Shown in graph, the maximum cell Vt under 3V gate stress after 1, 100, 10K, 100K and 1M erase/programming cycles.

As for the worst case cell in 65536 cells, a similar relation can be obtained in principle. However, as shown in figure 4, there is a very wide range of Vt shift which strongly depends both on the number of pre-cycling and on the stress time. It is difficult to set a common criterion on time scale because of measurement range and resolution limit.
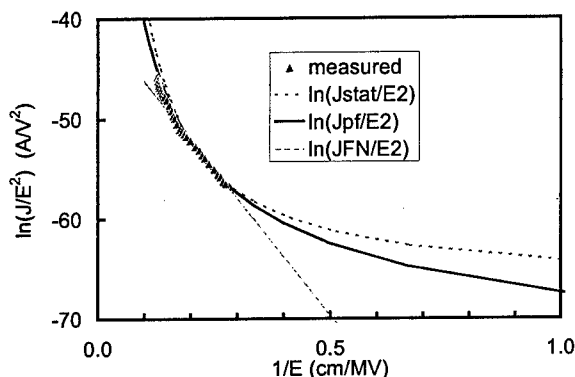
Figure 5. Poole-Frenkel and single Trap-assisted-tunelling models fit better to the measured data than Fowler-Nordheim model.

3.3 The electrical field dependence of SILC current density

In figure 5, the SILC current density as a function of the tunnel oxide electrical field as extracted from gate stress data (worst case cell, Tox=78A, after 100E/P cycles) is plotted on a Fowler-Nordheim plot. The SILC current has been calculated from the cell Vt shift during a certain time interval. The electrical field across the tunnel oxide has been determined from the cell Vt value and the applied gate stress voltage. This method has been widely used in literature [6].

Equations 1 and 2 are the basic equations needed for such data extraction. Here, Q_{FG} is the floating gate charge, the C_{PP} is the capacitance between floating gate and control gate, C_{FG} is the total capacitance seen from the floating gate, the Vt is the threshold voltage shift within period t, and A_{TUN} and X_{TUN} are the tunnel window area and thickness.

$$J_{SILC} = \frac{\Delta Q_{FG}}{A_{TUN} \Delta t} = -\frac{C_{PP} \Delta V_t}{A_{TUN} \Delta t} \qquad (1)$$

$$E_{TUN} = \frac{|V_{TUN}|}{X_{TUN}} = \frac{V_{CG} C_{PP} + Q_{FG}}{C_{FG}} \qquad (2)$$
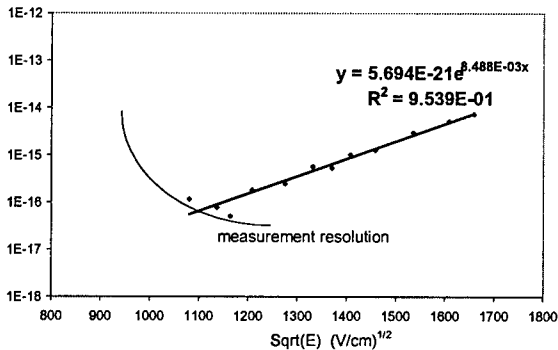
Figure 6. A Poole-Frenkel plot of the extracted SILC of a worst case cell.
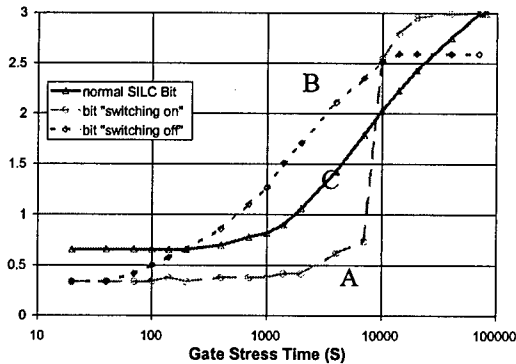
Figure 7. Cell A "developed" SILC, cell B "stopped" SILC, and cell C continued SILC during gate stress.

The extracted electrical current density vs. field characteristics have been fitted with 3 models: namely FN, PF and single TAT. More detailed analysis shows that PF and TAT models fit better than the FN model, especially in the low field, low current region. Since TAT model is quite cumbersome to be applied, PF model is used further for analysis and lifetime prediction. Figure 6 shows a fit with PF model of the

measured data. For the clearness, the Poole-Frenkel model is shown in equation 3:

$$J_{PF} = \alpha_{PF} \exp(\beta_{PF} \sqrt{E})E \qquad (3)$$
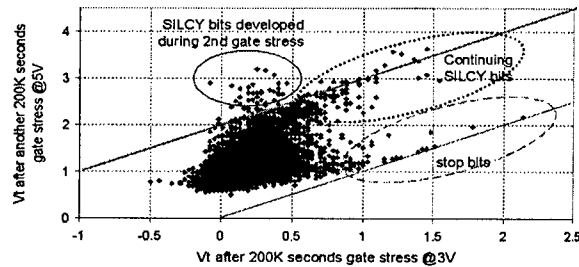
Figure 8. XY plot identifies the "continuing bits", "stop bits" and "start bits".

SILC has a strong statistical character, not only in the sense that it is localized randomly, but also in the sense that it appears and disappears randomly on the time scale, as reported in literature [4]. It has been experimentally observed that the switching does not show any dependency on the electrical field. In figure 7, the Vt development of some typical cells under 3V gate stress are shown: cell A "started" to show SILC after about 3000 seconds gate stress; cell B "stopped" at about 10000 second gate stress, and cell C continued showing SILC. In order to study whether these "starting" and "stopping" have any electrical field dependency, a 5V gate stress was applied directly after this 3V gate stress. The "stop" bits did not show up again in the subsequent 5V gate stress while the "start" bits continued to show SILC. An XY plot of final Vt after the "5V gate stress" versus that after the previous "3V gate stress" is shown in figure 8 where the "stop bits" and "start bits" are indicated. The total number of cells in this plot is about 32000. The two straight help lines show the cells which continued showing SILC during the second stress(upper line) and those did not show up again during the second stress (lower line).

3.4. The distribution of SILC cells

Figure 9 shows a typical threshold voltage distribution of 65536 cells (Tox=78A) in a die before, during and after the gate stress (of 2V). This die has been heavily pre-stressed by 140K erase/program cycles at high voltage (12V) without pulse shaping. The longer the gate stress time, the more cells enter the upper-tail distribution. However, the slope of the upper tail remains roughly the same (0.5V/decade).

4. RETENTION TIME PREDICTION

A simple way to translate gate stress data into retention characteristics is to shift the gate stress Vt versus time curve by −Vcg [7]. However, since gate stress measurement is performed in a limited time period and data retention requirement is much longer than that, some extrapolation is needed. With the knowledge of electrical field dependence of SILC (PF model, as shown in equation 3), the basic equations (1 and 2) and the tail distribution, it is possible to perform "product" lifetime prediction caused by SILC. Figure 10 shows such a lifetime prediction based on the data from worst case cell from figure 9.
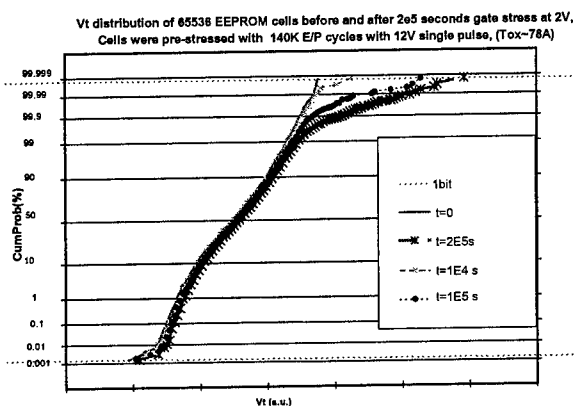
Figure 9. Cumulative Vt distribution as a function of stress time.
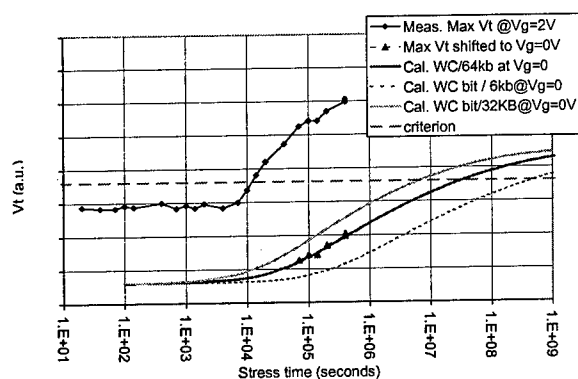
Figure 10. Room temperature storage lifetime prediction based on PF fitted parameters to the worst case cell (Tox=78A, after 140K programming/erase cycles with 12V square pulses): 15 years for a 6kb, 1 year for 64kb, and 2.5 month for a 256Kb "product", if without error correction.

5. DISCUSSIONS

One remark should made that, in a real product, where the pulse is nicely shaped, and the endurance spec is mostly about 100K cycles, the lifetime is much better than the calculated figures presented above because of less effective stress [8]. Quite often, there are also error corrections implemented in real products. Given the size of a product, the applied error correction scheme and the required data retention failure rate specification, a maximum tolerable number of SILC bits can be calculated. The gate stress method described in this paper can be used to check if the product specs can be met. On the other hand, given the data collected from the gate stress measurements as described in this paper, the error correction scheme can be optimized for a certain (large) product to satisfy its reliability spec.

6. CONCLUSIONS

SILC characteristics in embedded EEPROM process and its impact on data retention after full number cycles have been studied. It has experimental observed that the SILC path is very localized and random. The SILC leakage path can be turned "on" and "off" statistically and the switching does not depend on the electrical field. The SILC caused Vt shift can be fitted into a power-law against the number of program/erase cycles. The measured current-voltage characteristics of SILC can be well described with Poole-Frenkel model. EEPROM retention time prediction can be made by extrapolating the gate stress data with fitted Poole-Frenkel parameters. Pulse shaping helps lowering the effective stress level and error correction helps the product data retention.

7. REFERENCES
1. K. Naruke et al., IEDM Tech. Dig. 1998, pp 424.
2. C. de Graaf et al., Non-Volatile Semiconductor Memory workshop 1998, pp. 101–104.
3. C. de Graaf et al., NVSMW 1998, pp 68–72
4. S. Kamohara et al., IRPS 1998, pp 57.
5. P. Hesto, in Instabilities in silicon devices G. Barbottin and A. Vapaille ed.
6. S. Satoh et al. IEEE trans. ED Vol 45, No 2, 1998, pp482-486
7. B. Euzent et al. IRPS 1981, pp 11-16
8. Unpublished data

OFF-STOCHIOMETRIC SILICON OXIDES FOR APPLICATIONS IN LOW-VOLTAGE FLASH MEMORIES

Fernanda Irrera[a], Felice Russo[b]

[a]Dipartimento di Ingegneria Elettronica, Università "La Sapienza",
Via Eudossiana 18, 00184 ROMA, Italy

[b]Texas Instruments, Via Pacinotti, Avezzano (Aq), Italy (Present address: Micron Technology, Via Pacinotti, Avezzano (Aq), Italy)

In this paper we systematically investigate the mechanism of enhanced Fowler-Nordheim injection in a MOS capacitor under accumulation when a film of off-stochiometric silicon oxide is deposited between the poly-gate and the thermal silicon dioxide. Experimental results prompt for application in low-voltage Flash memories.

1. INTRODUCTION

Starting from previous works, this paper investigates plasma enhanced chemical vapour deposition (PECVD) SiO_x to be used as dielectric over the thermal stochiometric silicon dioxide in a non-volatile flash memory cell [1-5]. The PECVD SiO_x films under study are silicon-rich off-stochiometric oxides and will be also called SRO along the text. These films have been characterized from a structural and electrical point of view, through TEM and XRD techniques, current-voltage and capacitance measurements. We will show that n^{++}-poly/SiO_x/SiO_2/p-sub MOS capacitors (SRO-MOS) can be employed in low-voltage non-volatile memories. In fact, an enhanced Fowler-Nordheim (FN) injection has been measured with respect to the conventional n^{++}-poly/SiO_2/p-sub capacitors, thanks to the presence of the SRO: the threshold voltage for the FN regime decreases by about 20% with a two times-thicker dielectric. A transport model is also proposed, in order to explain why the SRO enhances electron injection under substrate accumulation, relying on the bandgap diagram and assuming a trap assisted tunnel through the SRO.

2. DEVICE CHARACTERIZATION

The devices have been fabricated in the laboratories of Texas Instruments Italia (T.I.) at Avezzano, in the production line of the 0.4 μm flash memory. The technology chosen in T.I. for the depostion of the SiO_x films was the PECVD in order to exploit the same chamber used for the growth of stochiometric silicon oxide. Nitrogen was used as the carrier gas, nitrogen oxide as silane diluent, and we define the dilution parameter R as $[N_2O]/[SiH_4]$. The temperature was fixed at 320°C on the wafer surface. No thermal annealing was performed on samples soon after the deposition. We investigated the effect of the radiofrequency power (RFP), the N_2 flux ($[N_2]$), the dilution parameter R, and the chamber pressure (P) on the thickness uniformity and refractive index. As a result, decreasing R down to 1.3, the refractive index increases over the value 2 since the material becomes silicon richer. As expected, the depo-rate and thickness disuniformity are greatly affected by RFP, which prompted us not to exceed the value of 80 W. Also P has a critical effect on the thickness uniformity, and its value was finally chosen around 1 Torr. Structural characterization of R=2 and R=6 films of SiO_x deposited on crystalline silicon have been performed. TEM did not reveal the presence of any crystalline island and the SiO_x appeared quite homogenous. X-ray diffraction was also performed and no signal was detected in correspondance to the crystalline silicon phase. In conclusion, to the extent of the technique sensitivity, we do not have crystalline islands in our SiO_x films.

Two different families of MOS capacitors have been fabricated: 1) conventional n^{++}-poly/SiO_2/p-sub and 2) non-conventional n^{++}-poly/SiO_x/SiO_2/p-sub capacitors (SRO-MOS capacitors).

In Fig.1, the cross section of the SRO-MOS capacitor is sketched.
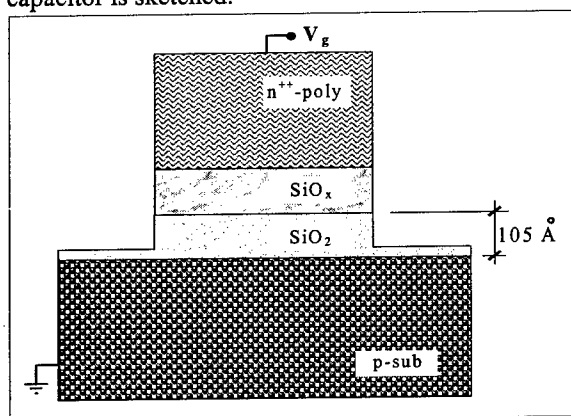
Fig.1 Sketch of the SRO-MOS capacitor

In all the devices, substrate, thermal oxide, and poly-gate were the same: the substrate was [100] oriented, the poly-gate was 3500Å thick deposited by PECVD with ρ=2000Ω/square, the 105Å-thick film of SiO_2 was thermally grown in H_2 and O_2 atmosphere. Devices had an area of 7.07×10^{-2} cm^2. The substrate was contacted by a polysilicon film deposited on the back side of the wafer.

In Fig.2 the I-V curves of conventional poly-n^{++}/SiO_2/p-sub and non-conventional poly-n^{++}/SiO_x/SiO_2/p-sub capacitors are displayed. As one can see, the n^{++}/SiO_x/SiO_2/p-sub capacitor exhibits a lower value of the Fowler-Nordheim threshold V_{th} than the n^{++}/SiO_2/p-sub ones. In that capacitor the SiO_x film was 185 Å thick and had R=2. This is an experimental evidence of enhanced electron injection in the capacitors with SiO_x. The very encouraging result is that V_{th} improves by 20% with a dielectric which is almost three times thicker than in the conventional capacitor.

The role of the dilution parameter R has been investigated in the range 1-10. We obtained that the value of V_{th} systematically shifts with R indicating that when increasing the silicon content in the SiO_x the electron injection enhances. In particular, V_{th}~8.5V for R=1.3 and V_{th}~22V for R=8.3. The role of the SiO_x thickness has been also investigated and we found that its effect on V_{th} is negligible. On

the contrary, the capacitance value under accumulation varies with the thickness confirming that the film is a dielectric and contributes to the total series capacitance of the MOS device. The C-V measurements were performed at 1 MHz.

We focus on the C-V measurements on samples with 372Å and 218Å-thick SiO_x films and measure: V_{FB1}=-1.2V, C_{max_1}=141pF, and V_{FB2}=-1.4V, C_{max_2}=212pF, respectively. Now, since the interfaces and the SiO_2 films are equal in the two capacitors, the quantity V_{FB1}-V_{FB2} is directly related to the volume charge trapped (Q_V) in the two films of SiO_x. We have: V_{FB1}-V_{FB2}=0.2 V. The sign of the flatband variation means that the contribution from trapped electrons exceeds the total contributions from positive charge (ions and trapped holes) in the bulk of the dielectric. The trapped charge distribution has a maximum in the vicinity of the SiO_x/SiO_2 interface since the SiO_2 surface was damaged by plasma during the growth of the SiO_x film. We can depict the defect distribution, following the steps: 1) adopt the one-energy level defect approximation for simplicity 2) assume an exponential trap profile as: $\rho(x)=\rho_0\exp(-x/x')$, where ρ_0 is the value at the SiO_x/SiO_2 interface and x' the decay parameter -which we set at 40 Å-, 3) ascribe
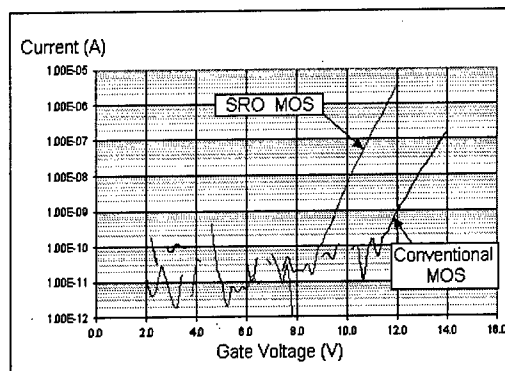
Fig. 2 Comparison between FN injection in conventional MOS and SRO-MOS

the quantity V_{FB1}-V_{FB2}=0.2 V entirely to the charge trapped in the thickness difference between the two SRO films. The obtained values of ρ_0 and Q_V are about 0.02 C/m and 80 pC, respectively. These values of x', ρ_0 and Q_V are merely indicative, since energy integration along the defect distribution in the gap should be performed in a more realistic picture.

3. CONSTANT CURRENT STRESS

An experiment of electrical stress under constant current was performed on both n^{++}-poly/SiO_x/SiO_2/p-sub and n^{++}-poly/SiO_2/p-sub capacitors. During the experiment the negative voltage applied to the gate was varied in order to keep the current constant at the value of 700 nA, corresponding to a current density of about 1mA/cm². Figure 3 shows the change in gate voltage as a function of stress time under substrate accumulation (inversion).

We measured a change in gate voltage indicating that the oxide exhibits significant trap generation and electron (hole) trapping at the gate/ SiO_x (SiO_x/SiO_2) interface, which lowers (increases) the cathode electric field. Results are in agreement with literature [6]. C-V curves during stress were also monitored on the same two samples and we noticed two important facts: 1) after the stress the C-V curves shift by about the same amount, independently of the SRO thickness; 2) δV_{FB}<0. As well known, the capacitance under accumulation is due to charge variation at the farthest interface from the gate. Thus the dominant mechanism of electrical

As shown in Fig.4, data of time evolution of $|\delta V_{FB}|$ during the stress are best-fitted by the equation of voltage drop of a charging capacitance with timescales of trapping-detrapping phenomena of deep traps. An experiment of electrical stress of conventional n^{++}-poly/SiO_2/p-sub capacitors performed in the same conditions showed a δV_{FB} of about 0.45V after 1000 seconds stress time. Since the conventional cell has an erase time of 10ms, this corresponds to 10^5 erase cycles. Now, the erase time of the n^{++}-poly/SiO_x/SiO_2/p-sub cell is lower since the erase current is higher. Actually we are not able to predict how long the erase time in a Flash cell using the n^{++}-poly/SiO_x/SiO_2/p-sub capacitor would be, since the F-N current varies during the erase. As an example, if we consider an extreme case and assume the erase time reduction factor equal to the F-N current enhancement factor just above threshold (100 times), then 10^3 cycles correspond to 10 seconds stress time on the time axis of Fig.4. In this case, δV_{FB} is about 0.15V, largely lower than in the conventional case. Actually, the erase time reduction factor is lower than 100 for most of the erase time duration and the example above is not realistic. Neverthless, it indicates that the n^{++}-poly/SiO_x/SiO_2/p-sub capacitor is likely to be compatible with the endurance requirements of practical applications in non-volatile memories.

Fig.3 Time evolution of δV_g during the stress

degradation of SiO_x is the creation of defects at the substrate interface where holes can be trapped, giving rise to a negative shift of the flat-band voltage. According to literature [6-8] electron stressing can produce several effects in silicon involving release and migration of hydrogen, breaking of weak bonds and formation of traps.

Fig.4 Time evolution of $|\delta V_{FB}|$ during the stress

4. TRANSPORT MODEL

We investigated the temperature dependance of current in conventional MOS and SRO-MOS in the range: 30-110°C. Results indicate that two different mechanisms rule the electron transport through SiO_x

and SiO_2, namely a temperature dependent trap assisted tunnel through the SiO_x and that the FN tunnel of the SiO_2 and the latter is the limiting one.

In particular, the Poole-Frenkel (PF) tunnel fits the voltage dependance once the temperature is fixed, and the multi-step-tunnel (mst) fits the temperature dependance once the voltage is fixed. Thus, transport through the SiO_x can be written as: $I(V,T) \propto V exp(\beta'V^{1/2})* exp(\gamma T)$. Once electrons have passed the SiO_x film by a trap-assisted-tunnel (t.a.t.) mechanism, they must tunnel the second barrier, i.e. the SiO_x/SiO_2 one, by a FN mechanism. Now the point is how two barriers ($\Phi_{poly/SiOx}$ and $\Phi_{SiOx/SiO2}$) enhance electron injection with respect to case single n^{++}-poly/SiO_2 barrier ($\Phi_{poly/SiO2}$).

To this aim, a picture of the band bending of the complete n^{++}-poly/SiO_x/SiO_2/p-sub capacitor under accumulation is sketched in Fig.5, where two conditions must be satisfied: $E_g^{Si}<E_g^{SiOx}<E_g^{SiO2}$ and $\Phi_{poly/SiOx}+\Phi_{SiOx/SiO2}= \Phi_{poly/SiO2} =3.2$ eV.
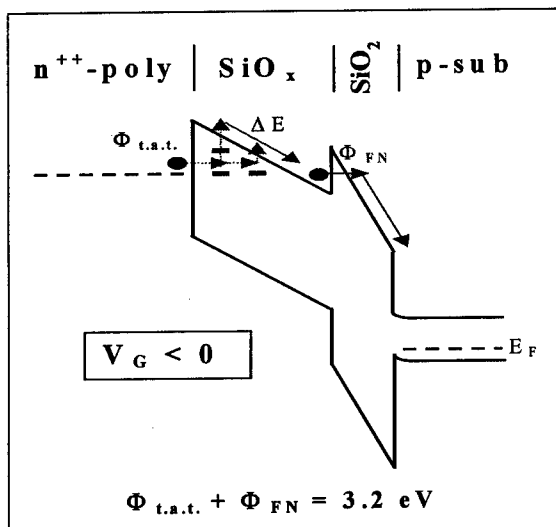
Fig.5 Band diagram of the SRO-MOS

In that figure, we notice that heigth (Φ) and depth (L) of the FN barrier are reduced with respect to the conventional cell and ΔE is an excess energy, accelerating electrons toward the SiO_x/SiO_2 interface. In general, in the FN regime it holds $I=qNAv\Theta$, where Θ is the tunnel probability and $qNAv$ is referred to as the supply function, q is the electron charge, N is the interface density of states in the conduction band, A is the device area, v the medium electron velocity. The tunnel probability Θ is related to L and Φ: $\Theta = exp(-\gamma\Phi^{1/2}L)$ where

$\gamma=4(2m*)^{1/2}/3h$. The supply function is increased by the quantity ΔE through the electron velocity. Values of ΔE depend on V_G and range from 1 to 5 eV, above the FN threshold, even if most of this excess energy is lost at the SiO_x/SiO_2 interface due to direct weak Si-H bonds breaking and rapid thermalization with the shallow defects. Calculations based on this model are in excellent agreement with experiment for an assumed value of the FN barrier ($+\Phi_{SiOx/SiO2}$) around 1.3 eV.

This confirms that the presence of a ΔE in the SiO_x and a smaller FN barrier account for the enhancement injection.

5. CONCLUSIONS

In this paper, n^{++}-poly/SiO_x/SiO_2/p-sub capacitors for applications in low-voltage non-volatile memories have been extensively studied. First, a systematic characterization of the PECVD SiO_x films has been performed. Then, the enhanced electron injection under substarte accumulation has been demonstrated and monitored under constant current stress. As a result, the n^{++}-poly/SiO_x/SiO_2/p-sub capacitor is compatible with endurance requirements of actual Flash memories, and the erase time lower than in the conventional case. We proposed a transport model consisting on two different mechanisms: a trap-assisted-tunnel through SiO_x, and a FN tunnel through SiO_2. The presence of two "small" barriers and an accelerating electric field in the SiO_x causes the injection enhancement respect to the case of a single poly/SiO_2 barrier.

REFERENCES

1. H. A. R.Wegner in IEDM Tech. Dig. 1984,.480
2. C. J. Lin, C. C. H. Hsu, H. H. Chen, G. Hong, L.S.Lu,IEEE Trans.Electron Devices,43,1996, 2021
3. L. Dori, A. Acovic, J. D. DiMaria C. H. Hsu, IEEE Trans. Electron Devices, 14, 1993, 283
4. B. Maiti, C. J. Lee, IEEE Electron Device Letters, 13, 1992, 624
5. S. Yokoyama, D. W. Dong, D. J. DiMaria, S. K. Lai, J. Appl. Phys. 54, 1983, 7058
6. N. Bhat, P. P. Apte, K. C. Saraswat, IEEE Trans. on Electron Devices, 43, 1996, 554
7. D.J. Di Maria, Solid-St. Electronics,41,1997,957
8. T. P. Chen, S. Li, S. Fung, K. F. Lo, IEEE Trans. Electron   Device,   45,   1998,   1920

Degradation and Breakdown of SiO₂-Layers due to Hot and Ballistic Electron Transport

H.-J. Fitting, E. Schreiber, T. Barfels, A. von Czarnowski

Rostock University, Physics Department, Universitätsplatz 3, D-18051 Rostock, Germany

External electron irradiation and internal ballistic electron transport have been used to investigate the induced electronical and optical modification of SiO₂ layers. The dose behaviors of cathodoluminescence (CL), of charge injection I(V) and trapping, of secondary electron field emission (SEFE), and of exoelectron emission (EEE) follow an exponential saturation up to doses of 0.01 – 0.1 As/cm². The respective defect creation is attributed to twofold coordinated silicon =Si: centers and their conglomeration under high electric fields to tips and channels accross the SiO₂ layers. The latter ones allow an almost ballistic free flight of electrons as shown by high energy vacuum emission.

1. INTRODUCTION

Extrinsic and intrinsic defects in SiO₂ have a decisive influence on the material quality and its application. Several defects are optically active and can be studied by luminescence spectroscopy [1]. Cathodoluminescence (CL), the emission of light as a result of electron irradiation, is particulary suitable for investigation of thin films, because large excitation doses are obtained within a small probe volume.

However, ionizing radiation in the dielectric and optically transparent material SiO₂ produces defect luminescence, charge trapping, IR mode softening as well as exoelectron emission after excitation and during thermal annealing. These irradiation and relaxation phenomena have been investigated in context in [2]. Under cathodo-excitation, e.g. the main blue luminescence band B at 460 nm (2.7 eV) is appearing and then strongly growing with an exponential saturation [3], see Fig. 1.
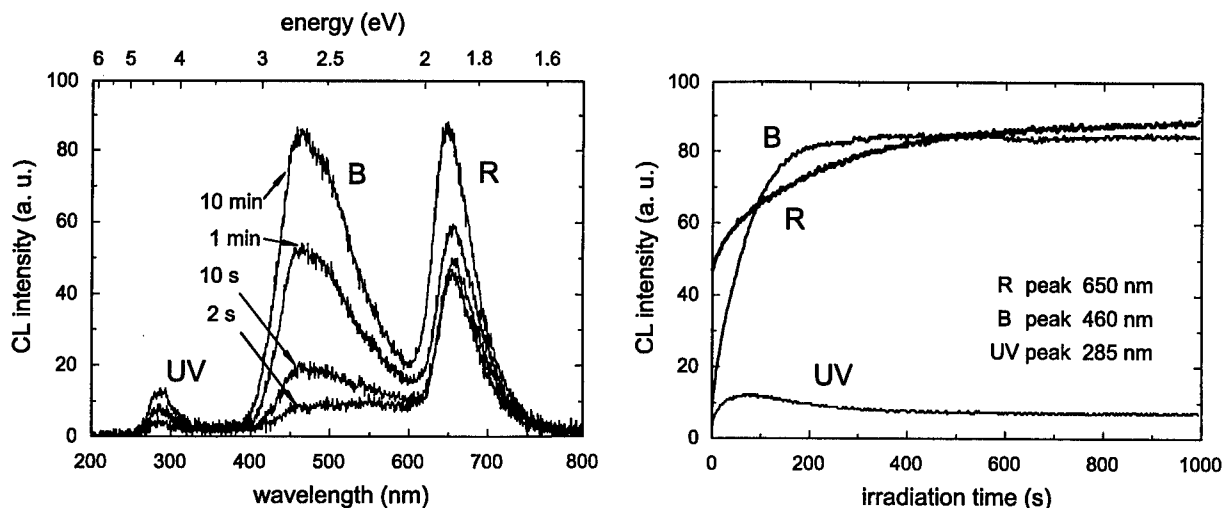
Fig. 1 Evolution of the SiO₂ cathodoluminescence spectrum with electron irradiation time; $d_{ox} = 250$ nm, $E_0 = 5$ keV; $j_0 = 8 \cdot 10^{-4}$ A/cm², RT
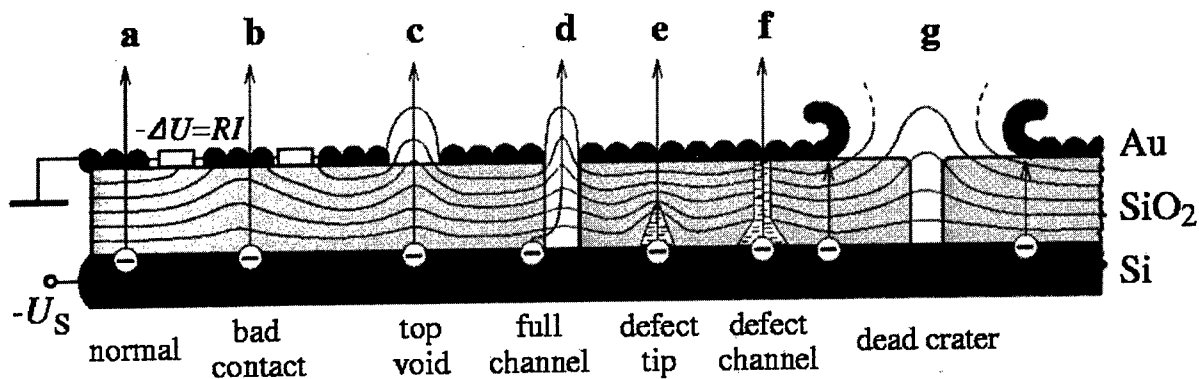
Fig. 2 Normal vacuum emission of electrons (a) and typical (b)(c)(g) and possible (d)(e)(f) damages and degradation effects of MOS stacks

The respective luminescence center is correlated with oxygen deficiency and twofold coordinated silicon =Si:, according to the Skuja model, [4]. A simple oxygen vacancy may present the precursor for this kind of defect transformation, [2, 3].

2. ELECTRONIC AND OPTICAL MODIFICATIONS DUE TO IRRADIATION

The comparison of different electronic excitation-relaxation processes like
- cathodoluminescence (CL),
- charge injection (IV) and charge trapping,
- secondary electron field emission (SEFE),
- exoelectron emission (EEE)

leads to a generally similar excitation-dose behaviour described by an electron beam saturation dose of 0.01-0.1 As/cm^2. This suggests a correlation of these four ballistic electron excitation mechanisms likely related to the same kind of defect in glassy SiO$_2$.

Defect conglomeration and consequently defect tip formation under electric field has been proposed and discussed already by Stoneham et al. [5]. The results are based on I(V) diode characteristics and I$_{VE}$(V) electron vacuum emission curves of MIS structures [6]. Defect tips and channels produce hysteresis and do affect the I(V) characteristics irreversibly. They should be considered as

precursors of the macroscopic breakdown and burn-out.

Defect conglomeration has been detected also by luminescence (PL, CL) measurements in SiO$_2$, [7]. The twofold coordinated silicon centers =Si: with the typical singlet-singlet and singlet-triplet transitions, associated with the ultraviolet UV and the blue B luminescence bands of SiO$_2$ (Fig. 1), may conglomerate to defect complexes [7].

3. HIGH FIELD VACUUM EMISSION

In a second experimental setup we have investigated high field vacuum emission (VE) of ballistic electrons from Au-SiO$_2$-Si-structures. The VE from these MOS structures is recorded by means of a spherical grid retarding field analyzer in combination with a secondary electron multiplier, [8]. Thereby the top semitransparent gold electrode of the samples is grounded and the substrate is connected to the drive voltage U$_S$, as can be seen in Fig. 2.

The MOS samples consist of n-Si substrate with (10 – 500) nm SiO$_2$ layers grown by thermal dry oxidation at about 1200 °C. Very thin (8 to 12) nm sputtered Au layers are deposited on the top. The n-Si substrate enables tunnel injection of electrons from the conduction band of Si into SiO$_2$ at fields of several MV/cm, thus the investigation can be extended to this field range.
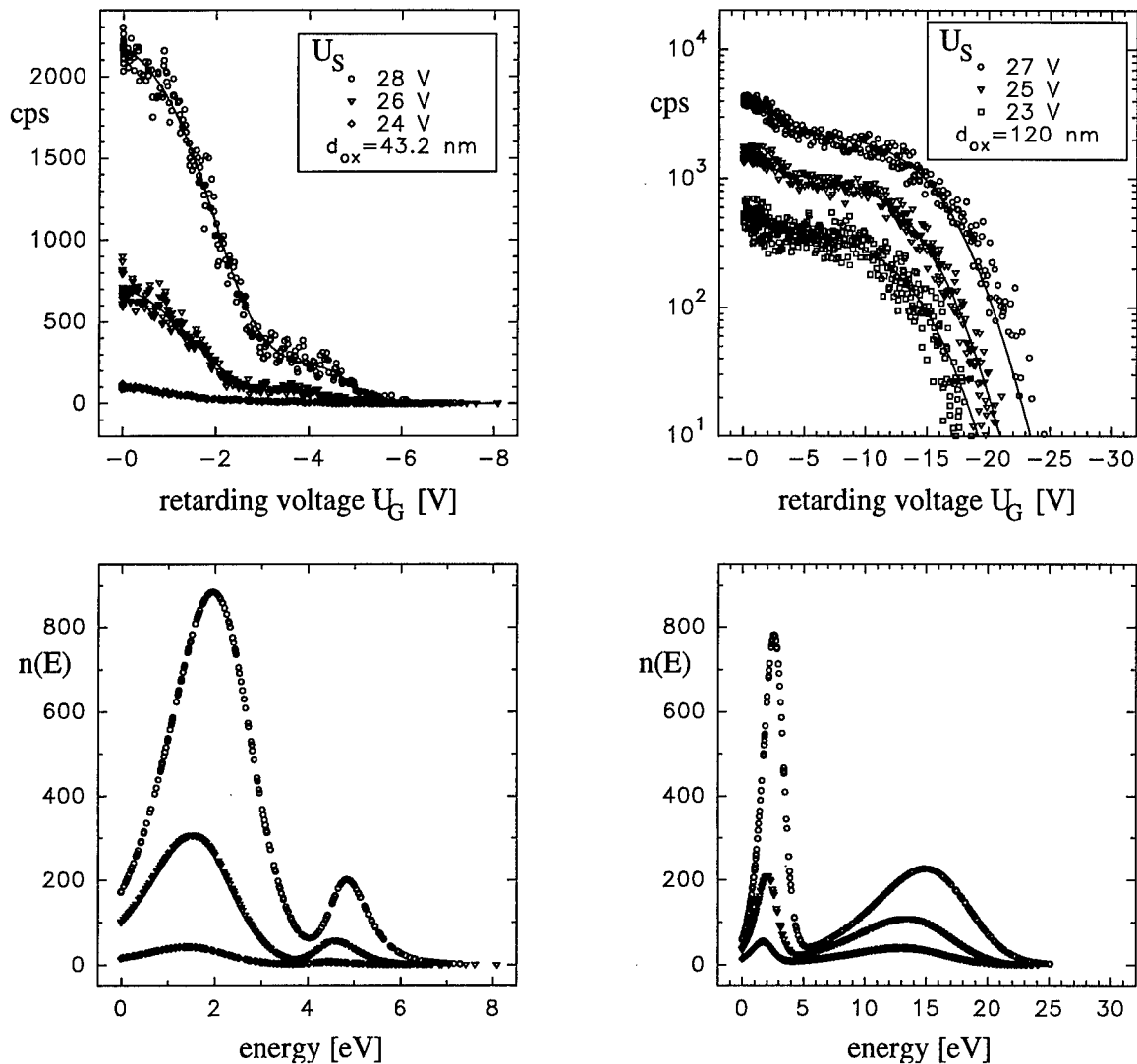
Fig. 3 Retarding field curves (above) and corresponding energy distributions (below) with a beginning double peak structure and wide spreaded distributions for higher fields and greater oxide thicknesses d_{ox} ; U_s - driving voltage.

The injected electrons will undergo scattering processes in the SiO_2 layer and under normal conditions a certain part of them (up to 10^{-3}) will be emitted through the thin top Au-electrode into vacuum, as shown in Fig. 2a. The retarding fields curves in Fig. 3 and their derivatives, i.e. the energy distributions of emitted electrons, demonstrate the energy distribution as a function of the path length across the oxide layer. A series of layer thicknesses

d_{ox} = (13 - 500) nm has been investigated. This experiment has to be carried out very carefully because only at certain flight distances and field strengths the beginning of avalanching will appear as it is indicated by a double peak structure consisting of exciting and excited electrons like in Fig. 3.

For very thin layers (d_{ox} = 13 nm) we observe electron energies below 2eV (Fig. 4) increasing to almost 4 eV at roughly double layer thickness (d_{ox} = 21 nm).
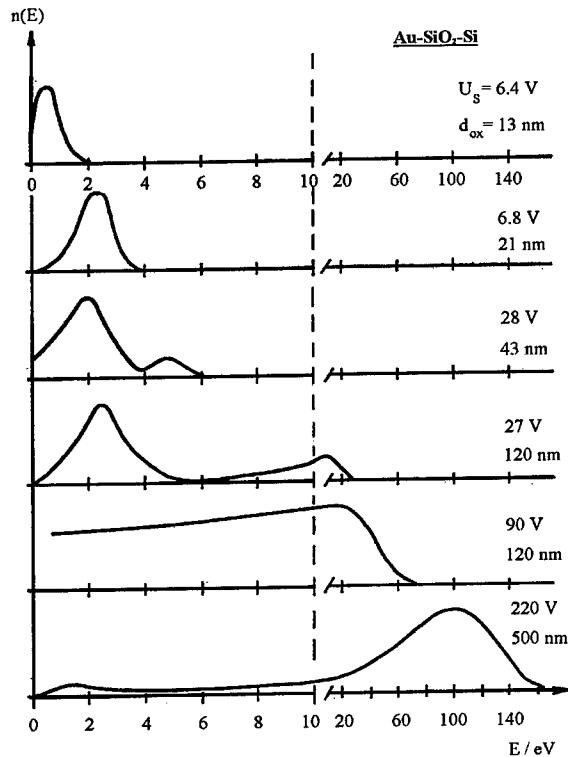
Fig. 4
Evolution of the electron vacuum emission
(VE) energy distributions with increasing oxide
thickness d_{ox} and drive voltage U_S;
the energy scale is interrupted at 10 eV,
slightly above the forbidden gap $E_g \cong 9$ eV
in SiO_2.

This is still below the energy gap of SiO_2
with $E_g \cong 9$ eV. With further increase of
thickness d_{ox} and drive voltage U_S, to 20 V
and higher, we observe the double peak
structure of the energy distributions as
demonstrated in Fig. 3 and 4.
Especially for very thick layers $d_{ox} = 500$nm
and high field strength $F \approx 5$ MV/cm we
measure electron energies up to 200 eV, i.e.
80 % of the maximum ballistic drive energy
$E_{ball} = eU_S$, as shown in Fig. 4 and already
reported in [8].

4. CONCLUSIONS

Obviously, no energy stabilization of emitted
electrons is obtained, very likely caused by
defect conglomeration and forming of defect
tips and defect channels accross the SiO_2
layers. We assume that in a first step these
defects form injection tips at the interface to
the Si substrate, see Fig. 2e. At the interface
a certain oxygen deficit is given presenting an
essential precursor for the electronically active
defect. Then these injection tips may grow to
defect channels as shown in Fig. 2f.
The non-bond free luminescent electron pairs
of the defect complexes **=Si:** may form
electron free flight channels enabling
electrons an almost ballistic emission into the
vacuum.
This hypothesis should be proved more
comprehensively in future. However, the
following burn-out of such a channel by local
heat-up, is probably, leading to the
macroscopic "dead craters" as presented in
Fig. 2g and more thoroughly investigated in
[8]. It is the final state of burn-out as has
been observed by many other authors, too.

REFERENCES

1. D. L. Griscom, J. Ceram. Soc. Jpn. 99
   (1991) 899
2. H.-J.Fitting, A. v..Czarnowski, A.N.Trukhin,
   M. Goldberg, T. Barfels, Solid State
   Phenomena 63-64 (1998) 333
3. M. Goldberg, H.-J. Fitting, A. Trukhin,
   J. Non-Cryst. Solids 20 (1997) 69
4. L. Skuja, J. Non. Cryst. Solids 167 (1994)
   229
5. G.Dearnaley, A.M.Stoneham, D.V. Morgan,
   Rep. Prog. Phys. 33 (1970) 1129
6. T. W. Hickmott, J. Appl. Phys. 33 (1962)
   2669 and 37 (1966) 3480
7. A. N. Trukhin, M. Goldberg, J. Jansons,
   H.-J. Fitting, I. Tale, J. Non-Cryst. Sol. 223
   (1998) 114
8. H.-J. Fitting, Th. Hingst, E. Schreiber,
   CSC'3 – Electric Charge in Solid Insulators,
   1998, Tours/France, Suppl. „Le Vide"
   287 (1998) 636

Growth of Si nuclei on SiO_2 for quantum dot memory applications

J. Vizoso[a], F. Martín[a], X. Martínez[b], M. Garriga[b] and X. Aymerich[a]

[a]Departament d'Enginyeria Electrònica. Universitat Autònoma de Barcelona
[b]Institut de Ciència de Materials de Barcelona (CSIC)
Universitat Autònoma de Barcelona. 08193 BELLATERRA (Barcelona), Spain.

By means of a Rapid Thermal Chemical Vapor Deposition (RTCVD) reactor, we have deposited Si nuclei on SiO_2 with appropriate characteristics for the fabrication of quantum dot based memories. These characteristics have been obtained from AFM measurements. The effects of temperature and deposition time on nuclei density, size and uniformity have been analyzed and discussed.

1. INTRODUCTION

One of the main applications of tunnel oxides can be found in the field of non-volatile memories. Among them, floating gate EEPROM's are the most popular. In these memories, a threshold voltage shift is caused by charging the floating gate via electron injection through the tunnel oxide. The magnitude of this voltage shift as well as device performance (write and read times, charge retention) are related to the thickness and quality of the tunnel oxide. In traditional designs, tunnel oxides as thick as 7nm are necessary to achieve non-volatility. This requires high programming voltages and long writing times. However, thinner tunnel oxides (less than 5nm) can be used if the floating gate is replaced by an array of silicon nano-islands, since in this case, the charge loss through lateral paths to the contacting regions is suppressed. In this way, fast write and read times, small degradation, long retention, operation at smaller voltages and low power consumption can be achieved. It is clear that device performance has to be very dependent on the characteristics of silicon islands, or nuclei, embedded in the insulator (dimensions, density). However, although nuclei growth with application to memory devices has been reported [1-3], additional effort is required to gain insight on how the characteristics of these silicon nuclei grown on SiO_2 can be controlled. In this work, we have studied the initial stages of silicon deposition on SiO_2 by Rapid Thermal Chemical Vapor Deposition (RTCVD) with the aim of obtaining the dependence of nuclei characteristics (size distribution and density) on process parameters (namely temperature and time). The main advantage of such a system is that very fast temperature transitions (up to 50K/s) can be achieved, so that accurate temperature control at relatively low deposition times is possible. This is an advantage compared with other CVD-based techniques, since, as will be demonstrated, deposition times of the order of 1min or less are necessary to grow Si nuclei with suitable properties for long retention time memory applications.

2. EXPERIMENTAL

The reactor is a RTCVD system with air-cooled tungsten lamps that allow temperature transitions on the surface of the wafers as fast as 50K/s, maintaining the walls of the deposition chamber at much lower temperatures. In such a way, the deposition processes are initiated and terminated by temperature switching, thus minimizing the thermal budget. This is important in applications requiring the growth of very thin films or nuclei. The system is equipped with a turbomolecular pump able to give an ultimate vacuum of 10^{-6}mbar before the initiation of the depositions. This provides the clean growth environment, essential to obtain good quality structures. We have performed Si depositions on SiO_2 with layer thickness ranging from several hundreds of nanometers down to silicon island formation. The deposition temperatures and

pressures have been considered within the intervals 520°C-600°C and 0.2mbar-1.5mbar, respectively, while the flow rate of the source gas (undiluted SiH_4) has been established in 50sccm.

The samples used in this study are 4in Si(100) wafers (resistivity 10Ωcm) with a thermal SiO_2 of 100nm thickness. Before depositions, oxide surfaces were cleaned by introducing the wafers in a H_2O_2/H_2SO_4 (3:1) bath during 15min, followed by a 5 min rinse in DI water and a spin dry. Subsequently, the wafers were loaded in the reacting chamber and the deposition process was allowed to start.

3. RESULTS AND DISCUSSION

Our aim was to find a process parameter window suitable to obtain Si nuclei on SiO_2 with appropriate characteristics to develop memory devices based on them, and to study the effects of these parameters (specifically temperature and time) on nuclei properties. To this end, it has been previously necessary to grow bulk films in order to analyze the effects of process parameters on growth kinetics and surface roughness. The latter parameter is important since it is related to the initial deposition stages, i.e. to the properties of the grown nuclei [4]. The growth rate (GR) is important. To have accurate control on nuclei size it is convenient not to exceed few nm/min of bulk film growth. We have obtained the GR dependence on temperature under fixed process pressure (0.2mb) and gas flow (50sccm of undiluted SiH_4). The Arrhenius plot (shown in figure 1) reveals that deposition temperatures in the range considered are suitable for our purposes since the growth rates (3.5nm/min-16nm/min) lie within the above cited specifications. Spectroscopic ellipsometry has been used to determine the thickness of the deposited layers, from which the GR have been obtained. The intercept of the thickness-vs-time linear fit to the time axis has been considered as an estimate of the time scale necessary to obtain stable nuclei. We have called this intercept incubation time. However, other alternatives to measure the time necessary for the onset of nuclei formation have been reported [5].

By analyzing the effects of process pressure, we have concluded that increasing pressure from 0.2mb up to 1.5mb, surface roughness, inferred from AFM measurements, decreases (figure 2). Deposition times have been decreased with increasing pressures to compensate for the increasing GR, with the aim to

obtain Si layers of similar thickness in all cases. The dependence of surface roughness on pressure is attributed to the larger silicon flux to the deposition surface at higher pressures, which enhances the density of stable clusters and hence reduces the grain sizes, providing smoother surfaces. This means that smaller and more dense nuclei are expected in the initial deposition stages under high pressures. One important parameter for most applications is nuclei uniformity. From AFM images obtained under various process pressures, it is seen how grain size uniformity in bulk films does not appreciably vary with this parameter. Therefore, we expect the same behaviour for nuclei growth. However, if SiH_4 is diluted in H_2, grain size uniformity is lost. This suggest to discard the dilution of SiH_4 in H_2 for the growth of Si islands in SiO_2.
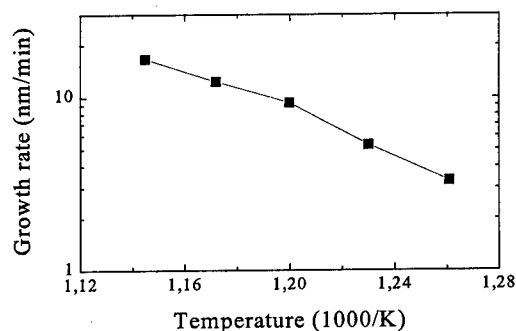
Figure 1. Arrhenius plot for Si films grown on SiO_2 under 0.2mb and 50sccm of undiluted SiH_4
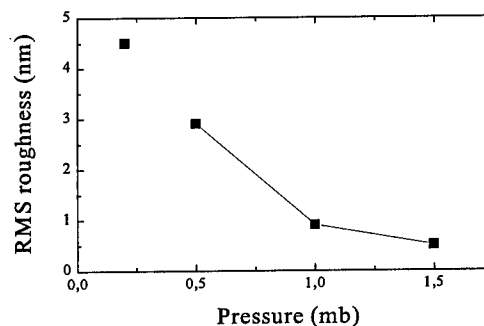
Figure 2.

Figure 2. Dependence of RMS surface roughness on deposition pressure for samples grown at 540°C and 50sccm of undiluted SiH$_4$.

According to bulk film characterization presented so far, and assuming that a correlation exists between final film morphology and the initial nucleation stages, undiluted SiH$_4$ at temperatures between 520°C and 600°C and pressures ranging between 0.2mb and 1.5mb are appropriate process conditions to the formation of nanoscale Si islands on SiO$_2$. In this study, we will analyze the effects of temperature and time on nuclei formation (process pressure and SiH$_4$ gas flow have been fixed in 0.2mb and 50sccm, respectively). In particular, we have grown Si nuclei at three different temperatures (520°C, 540°C and 600°C) and times. Nuclei of nanometer scale dimensions are obtained when deposition times are of the order of the incubation time (figure 3).
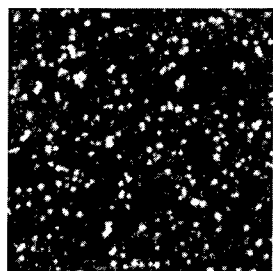
Figure 3. Si nuclei grown on SiO$_2$ at 600°C and 0.2mb during 20s with 50sccm of undiluted SiH$_4$. Nuclei heights are around 3nm and the image range is 1μmx1μm.
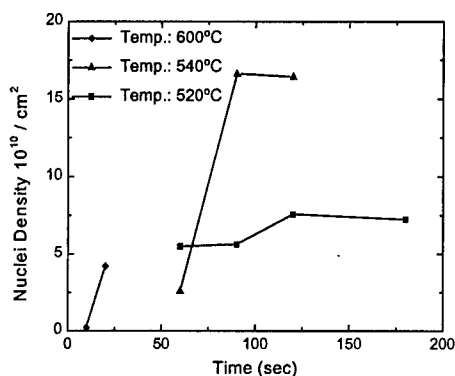
Figure 4. Dependence of nuclei density with time at the temperatures indicated and 0.2mb, 50sccm of undiluted SiH$_4$.

In figure 4, we represent the nuclei density variation with time for the temperatures considered (only two data point are represented at 600°C). As corresponds to a low temperature regime, small nuclei and high nuclei density have been obtained, in comparison to previous results obtained by Basa and co-workers at higher temperatures [6]. This is in accordance to the accepted idea that the higher adatom mobility that results at higher temperatures leads to sparse and large nuclei. However, if we compare the saturation density at 520 and 540°C, a higher value is obtained at the higher temperature, contrary to the previous idea. In our opinion, the different behaviour between our results (showing that saturation density increases with temperature) and those of Basa et al. [6] may probably be explained as a consequence of a different mechanism governing film growth at the first stages, under each experimental conditions. Namely, nuclei growth and nuclei formation both contribute to surface coverage. If the former proceeds at high rate as compared to the latter, large and sparse nuclei are expected before coalescence. Under the conditions of Basa, it is likely that nuclei growth rate is more sensitive to temperature than the rate of nuclei formation, so that an increase in temperature may lead to a lower saturation density. The opposite trend would explain why in our case saturation density increases with temperature. Whether this different behaviour at high [6] and low (present work) temperatures is due to temperature effects or is affected by other changes in experimental conditions (such as precursor gas dilution) requires further test beyond the scope of the present work.

A key property for memory applications is nuclei uniformity. This has been estimated from the standard deviation of nuclei heights and diameters relative to their mean values (figure 5). Nuclei uniformity clearly improves with temperature and deposition time. The variation nuclei in size is due to the fact that nucleation events are time distributed. At high nucleation rates, the time interval between the formation of the first nuclei and last nucleation events is low, and sharp size distribution are expected. This explains why improved uniformity is achieved at the highest temperature. On the other hand, the nucleation rate decreases with increasing time. The effect is related to a displacement of the size distributions towards higher values, which

results in a decrease in the relative standard
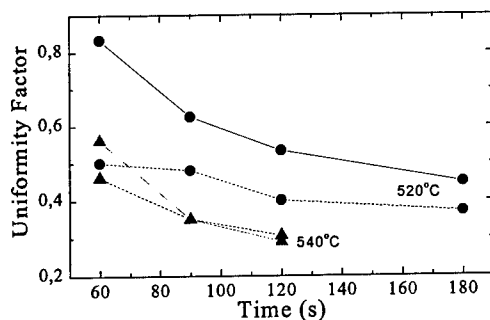
Figure 5. Uniformity factor for size distribution, i.e. heights (continuous line) and diameters (dashed lines). This has been estimated from the standard deviation relative to mean values.

deviation.

Finally, in figure 6 is shown the variation of nuclei heights with time for the lower temperatures considered. The radii (not shown) are larger than the heights, being the shape factor approximately two in all cases. It is obvious that the characteristics of the obtained nuclei are suitable for the fabrication of memories based on self-assembled quantum dots in the tunnel oxide. Our results point out that nuclei uniformity is improved with temperature. However, as temperature increases, a lower control on nuclei density and dimensions is expected since the sensitivity of these parameters with time also increases. These results seem to indicate that uniformity as well as control on nuclei density and dimensions can both be optimized by carrying out temperature varying processes -higher temperatures at the former deposition stages to shorter the period of nuclei formation, i.e. to achieve a high degree of uniformity, followed by a lower temperature stage to obtain a better control on nuclei dimensions and density. We left this as a future work.

4. CONCLUSION

In this work, appropriate process parameters for the formation of Si nanodots on SiO_2 by means of a RTCVD reactor have been obtained. The resulting nuclei have been found to be suitable for the fabrication of quantum dot based memories. We have analyzed the effects of temperature and time on nuclei characteristics. The following conclusions have been derived: (i) contrary to previous experiments carried out at higher temperatures, we have found that the saturation nuclei density increases with temperature; (ii) nuclei uniformity improves with time and temperature; (iii) at lower temperatures, the sensitivity of nuclei parameters (density and size) with time decreases, therefore it is expected a better controllability on nuclei growth at lower temperatures.
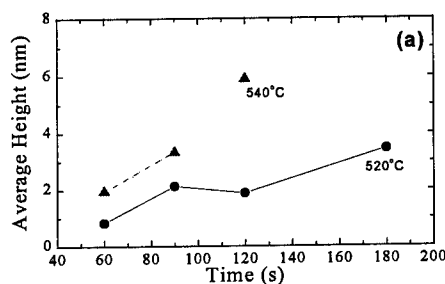
Figure 6. Average heights for the nuclei distributions of figure 4.

REFERENCES

1. H.I. Hanafi, S. Tiwari and I. Khan, IEEE Trans. Elect. Dev., **43**, 1553 (1996).
2. Y. Shi, K. Saito, H. Ishikuro and T. Hiramoto, J. Appl. Phys., **84**, 2358 (1998).
3. I. Kim, H. Kim, J. Lee and H. Shin, Extended Abst. of the 1998 Int. Conf. Sol. Stat. Dev. Mat. (SSDM'98), Hiroshima (Japan), p. 170.
4. K.E. Violette, M.C. Öztürk, K.M. Christensen and D.M. Maher, J. Electrochem. Soc., **143**, 649 (1996).
5. J.b. Rem, J. Holleman and J.F. Verweij, J. Electrochem. Soc., **144**, 2101 (1997).
6. C. Basa, M. Tinani and E.A. Irene, J. Vac. Sci. Technol. A, **16**, 2466 (1998).

Author Index Volume 48

LDSD '99

The Third International Conference on Low Dimensional Structures and Devices

15 –17 September 1999

Renaissance Antalya Resort,
Antalya, Turkey

➡ For the latest advances in...

Materials growth technology
Physics
Devices
Device technology
Microsystems technology, production and manufacturability
Systems integration of low dimensional structures and devices

ELSEVIER
SCIENCE

http://www.elsevier.nl/locate/ldsd99

Be a part of the third LDSD conference and find out about the latest advances in a broad range of low dimensional structures and devices (LDSD) technologies, and their integration for a variety of applications...

LDSD99 will provide a forum for researchers, engineers, and product developers engaged in the full spectrum of research and development, from fundamental research to product manufacturing to systems applications. In addition, the conference will provide an opportunity to get a broad overview of the field by covering a wide range of materials growth techniques, materials, characterisation, processing, device fabrication, microsystems, applications, and other related topics. This diversity of topics has been chosen to encourage the exchange of ideas between the broad range of physics and engineering disciplines.

PRESENT YOUR LATEST RESEARCH FINDINGS...
Invited talks will be presented by technical leaders in each of these key areas to describe the current state of the art of the LDSD field and to stimulate technical discussions. Papers on both fundamental studies and practical issues will be included.

Papers are invited for oral and poster presentation on the themes of the conference. Abstracts should be submitted to the conference secretariat by 5 March 1999. For guidelines on preparing an abstract see the conference website or contact the conference secretariat for a copy of the Call for Papers.

CONFERENCE WEBSITE:
http://www.elsevier.nl/locate/ldsd99

For full details contact:
Sue Stewart
LDSD99 Conference Secretariat
4 Manor Farm Barns, Church Lane
Charlton-on-Otmoor
Kidlington, Oxon OX5 2UA, UK
tel +44 (0) 1865 331040
fax +44 (0) 1865 331125
email suestewart@compuserve.com
web http://www.elsevier.nl/locate/ldsd99

Submission of papers. Papers to be submitted for publication should be sent to one of the Editors-in-Chief. Each paper should be provided with the following: (a) A short title of the paper (for running heads). (b) An abstract not exceeding 150 words and preferably not containing long or complicated formulae. (c) A list of keywords. (d) Original drawings of the figures, suitable for photographic reproduction (drawn in large size, on separate sheets, with wide margins and using India ink). Lettering must be of uniform size (preferably striplettering): typewritten or handwritten lettering is not acceptable.
Upon acceptance of an article, author(s) will be asked to transfer copyright of the article to the publisher. This transfer will ensure the widest possible dissemination of information.
The corresponding author receives **proofs,** which should be corrected and returned by airmail to the publisher within three days of receipt.
No page charge is made. 50 reprints of each paper will be provided free of charge. Additional reprints may be obtained at cost. A reprint order form will be sent to the corresponding author together with the proofs.

Preparation of manuscripts. Manuscripts should be written in English. They should be typewritten originals (plus 4 copies) with wide margins and double space between the lines. Please submit the manuscript in corrected, legible and final form. Corrections in the proof stage other than of printer's errors should be avoided; costs arising from such extra corrections will be charged to the author. Sections must be numbered and must have a title not consisting of formulae only. Long formulae must be displayed and numbered. All symbols which are not typewritten must be listed separately.
Footnotes, which should be kept to a minimum and should be brief, must be numbered consecutively and typed on separate sheets in the same format as the main text.
Figures must be numbered; they must be referred to in the text. The captions must be typed on separate sheets in the same format as the main text.
Tables must be numbered, and typed on separate sheets in ample spacing.
References must be numbered alphabetically. In the text they should be referred to by bracketed numbers. The list of references must be typed on separate sheets, in the same format as the main text, and ordered consecutively, according to the following examples for papers [1], books [2], contributed volumes [3] and unpublished papers [4].
[1] C.K. Cheong and J.L. Teugels, On a generalization of the random walk, *Stochastic Process. Appl.* 1 (1973) 53–66.
[2] S.M. Sze, *Physics of Semiconductor Devices*, Wiley, New York, 1969.
[3] J.M. Abadie, Nonlinear programming, in: K.B. Haley, ed., *Operational Research '78*, North-Holland, Amsterdam, 1979, pp. 900–930.
[4] S.I. Resnick, Almost sure limit points or record values, Tech. Rept. No. 39, Stanford Univ., Stanford, CA, 1972.