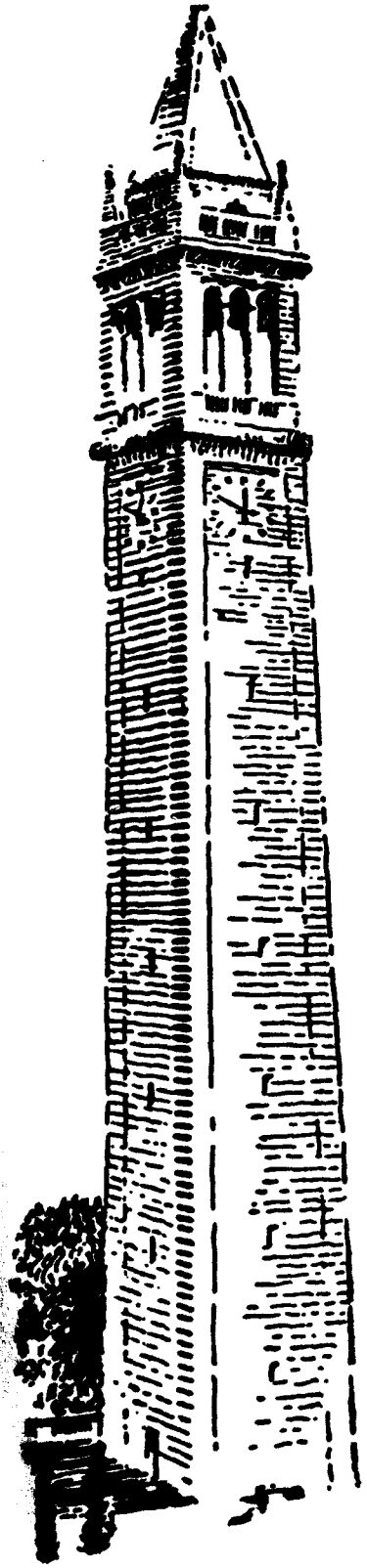


Report No. UCB/ERL 90/2

JOINT SERVICES ELECTRONICS PROGRAM

ANNUAL PROGRESS REPORT
(Contract F49620-90-C-0029)
(1 March 1991 — 29 February 1992)

Chenming Hu and Michael A. Lieberman

5 March 1992

Prepared for:

Air Force Office of Scientific Research (NE)
Bolling Air Force Base
Washington, DC 20332

ELECTRONICS RESEARCH LABORATORY
College of Engineering
University of California, Berkeley, CA 94720

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 5, 1992	Annual 1 MAR 91 - 29 FEB 92

4. TITLE AND SUBTITLE
Annual Report No. UCB/ERL 90/1

5. FUNDING NUMBERS
F49620-90-C-0029
Project/Task: 2305/A9
Program Element: 61102F

6. AUTHOR(S)
Chenming Hu and Michael Lieberman

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Electronics Research Laboratory
253 Cory Hall
University of California at Berekeley
Berkeley, California 94720

8. PERFORMING ORGANIZATION REPORT NUMBER
UCB/ERL 90/2

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research
Building 410
Bolling AFB DC 20332-6448

Program Manager: Dr. Gerald L. Witt, AFOSR/NE

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE,
DISTRIBUTION UNLIMITED

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This document is the 1991 annual report of research conducted in the Electronics Research Laboratory at University of California at Berkeley under the sponsorship of the Joint Services Electronics Program. The research covers the following project areas: Quantum Electronic Devices: (1-A)Nonlinear Optics in Compound Semiconductors(1-B)Ultrafast Optical Techniques(1-C)Optical Probing of Semiconductor Devices and Interfaces by Electro-Optic and Photo-Elastic Effects; Electronic Devices: (2-A).1 Micron BiCMOS Devices in Bulk and SOI Substrates(2-B)Conductive Oxides and Ferroelectrics for Programmable Devices and(2-C)Insulated-Gate GaAs Field Effect Transistors; Neural Networks and Parallel Computation:(3-A)Stochastic Neural Networks and Application to Signal Processing (3-B) Learning and Generalization by Neural Networks (3-C) Reconfigurable Analog Elements for Neural Nets and (3-D) Architectural Issues in Parallel Computation. There is also a section concerning significant accomplishments.

An Appendix of all publications, papers and presentations forms a second part of this report. It is bound under separate cover.

14. SUBJECT TERMS	15. NUMBER OF PAGES
	16. PRICE CODE

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

TABLE OF CONTENTS

PART A - DIRECTOR'S OVERVIEW

JSEP continues to play an important and unique role in electronics research at the University of California, Berkeley. Its emphasis on science and relatively stable funding provides an increasingly rare environment for conducting the more basic research and exploring promising new areas. It also provides a unique opportunity for encouraging collaborative research involving multiple principal investigators or new faculty members. Currently, JSEP at U.C. Berkeley partially supports the research of 12 faculty and 28 graduate students.

The responsiveness of JSEP was demonstrated again last year by its acceptance of our newest faculty member, Prof. Kam Lau, as a principal investigator. Prof. Lau will investigate high speed diode lasers. Also, JSEP accepted the proposal of Prof. Leon Chua to investigate his particularly simple neural network concept -- cellular neural network. Prof. Chua plans to fabricate neural network circuits in silicon and will strengthen the integration between theory and hardware, which is a prominent feature of our neural network program. Since the programs of Prof. Chua and Prof. Lau will not start until March 1992 (as their funding has not yet arrived) they are not represented in this Annual Progress Report.

Progress is reported here for Quantum Electronic Devices, Electronic Devices, and Neural Networks and Application to Signal Processing. Under Quantum Electronic Devices, a surface emitting second-harmonic generator using asymmetric quantum wells for phase matching is investigated. Phase locked epitaxy was used to fabricate vertical cavity surface emitting lasers, which produced up to 10 mW output -- the highest power demonstrated to date for these types of lasers. Band structure engineering is studied for the purpose of minimizing the series resistance in multi-heterostructure devices. Under Electronic Devices, we have fabricated and studied MOSFET with channel lengths down to 0.12 μm to understand the limiting phenomena in these devices. Deep submicron SOI (silicon on insulator) MOSFETS were thoroughly investigated because they offer the best opportunity for achieving functional devices below the 0.1 μm barrier.

Conductive oxides and ferroelectric film PZT were studied for application in new devices. AlN_x insulator on AlGaAs/GaAs substrate was investigated for understanding of insulator -GaAs interfaces. Neural Network and Application integrates various aspects of artificial neural network research into one program. EEPROM devices were fabricated and studied as variable weight device for neural network. The difficulty of training large neural networks and the unavailability of inexpensive weighting devices are addressed by the novel approach of using logic synthesizers to generate a network that produces the correct output for the data in the training set using either traditional gates or threshold gates with small integer weights. Another approach to training large-scale network studied is to partition the network into input domain with subdomains and associated separate classification networks. Fault tolerant ANN training was also investigated. The design of networks of large numbers of interconnecting processors was studied. Classification of sonar returns from different objects and different angles of observation was investigated. A Wigner-Ville transform time-frequency technique achieved 95% success rate at 0dB of signal to noise ratio, outperforming all other known classification algorithms.

Several significant accomplishments are highlighted in Part B of this report. A special effort was made to present a relatively large number of significant accomplishments in this section as the format makes them easy to read and easy to explain.

The enclosed Appendix includes copies of 22 published articles, 5 conference papers, and 7 papers submitted for publication. One paper on MOSFET reliability under AC stress of very thin gate oxide won the Best Student Paper Award at the 1991 International Electron Device Meeting (IEDM).

PART B - SIGNIFICANT ACCOMPLISHMENTS

Vertical Cavity Lasers Fabricated Using Phase-Locked Epitaxy

Professor J. Stephen Smith with Hong Lin, Jeffrey Walker, Gordon Wilson and Dan Kutchta

Our vertical cavity surface emitting lasers include devices which produced maximum output powers of up to 10 mW cw, which is the highest power demonstrated to date for these types of lasers. They also exhibit low series resistance, with turn on voltage of 2.6 volts, also the lowest reported, and unparalleled wafer to wafer yield. This project has established phase locked epitaxy (PLE) as a valuable technique for the fabrication of surface-emitting lasers and other semiconductor structures which have precision materials requirements. These advances in the ability to fabricate these lasers reliably, with virtually every wafer grown lasing, should be very important for further successes in these areas. Our PLE technique, using short period superlattices with growth interruptions for all of the device layers, has applications for other precision layered structures as well.

Publications

[1] Hong Lin, and J. S. Smith, "Compensating Group Velocity Dispersion in Quantum Well Lasers," presented at the Engineering Foundation Conferences, High Speed/High Frequency Optoelectronics, Palm Coast, Florida, March, 1991 and in February, 1992, submitted for publication in *Appl. Phys. Lett.*

[2] S. P. Dijaili, J. M. Wiesenfeld, G. Raybon, C. A. Burrus, A. Dienes, J. S. Smith and J. R. Whinnery, "Cross Phase Modulation in a Semiconductor Laser Amplifier Determined by a Dispersive Technique," *IEEE Journal of Quantum Electronics*, Vol. 28, No. 1, 141-150, January, 1992.

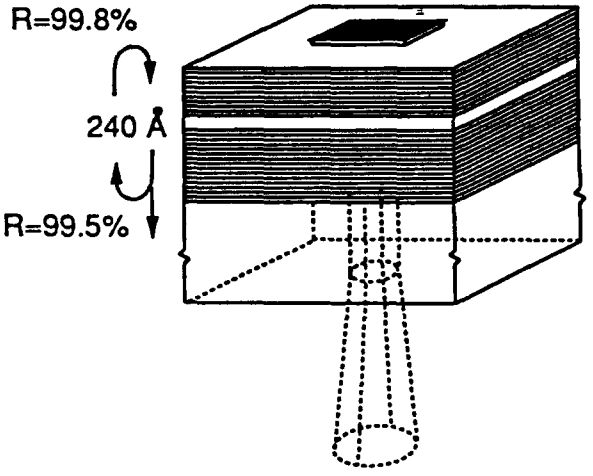
[3] J. D. Walker, Dan Kutchta, and J. S. Smith, "Vertical-Cavity Surface-Emitting Laser Diodes Fabricated by Phase-Locked Epitaxy," *Appl. Phys. Lett.*, Vol. 59, No. 17, pp. 2079-81, October, 1991.

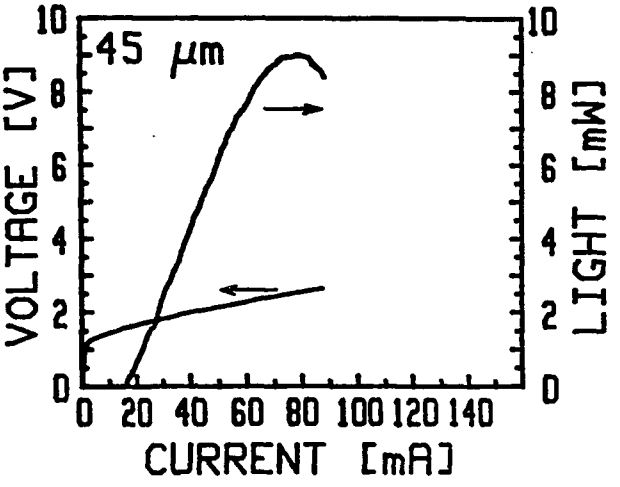
[4] Hong Lin and J. S. Smith, "Optical Time Division Demultiplexing Using Second Order Optical Nonlinear Effects," *Appl. Phys. Lett.*, Vol. 59, No. 22, November, 1991.

JOINT SERVICES ELECTRONICS PROGRAM
AT U. C. BERKELEY, EECS DEPARTMENT

PROF. J. STEPHEN SMITH
J. D. WALKER
D. M. KUCHTA

VERTICAL-CAVITY SURFACE-EMITTING LASER DIODES FABRICATED BY PHASE-LOCKED EPITAXY





- 10 mW CW output power – highest reported for vertical cavity lasers
- Low series resistance – 2.6 V turn-on voltage, also best reported
- High wafer-to-wafer yield
- Powerful MBE technique which can be applied to other devices as well

Asymmetric AlGaAs quantum wells for second-harmonic generation and quasi-phase matching of visible light in surface emitting waveguides

P. J. Harshman and Professor S. Wang

Large second-order susceptibilities can be associated with asymmetric quantum wells. Previous experimental work has concentrated on near resonance intraband processes and corresponding mid-infrared radiation. Of more practical interest is nonlinear optics at the near-infrared wavelengths of most semiconductor lasers, where visible light may be obtained by efficiently doubling the laser output.

We have derived an analytical expression for the second-order susceptibility for second-harmonic generation of visible radiation in asymmetric AlGaAs quantum wells [1]. The calculation suggests that a second-order susceptibility magnitude equal to that of bulk GaAs may be achieved in optimally designed wells. The sign of the second-order susceptibility may be changed by inverting the well asymmetry so that quasi-phase matching of the second-harmonic generation is possible. Simulation of a multiquantum well surface emitting waveguide designed for quasi-phase-matched second-harmonic generation indicates that second-harmonic conversion efficiency can be improved by over a factor of 100.

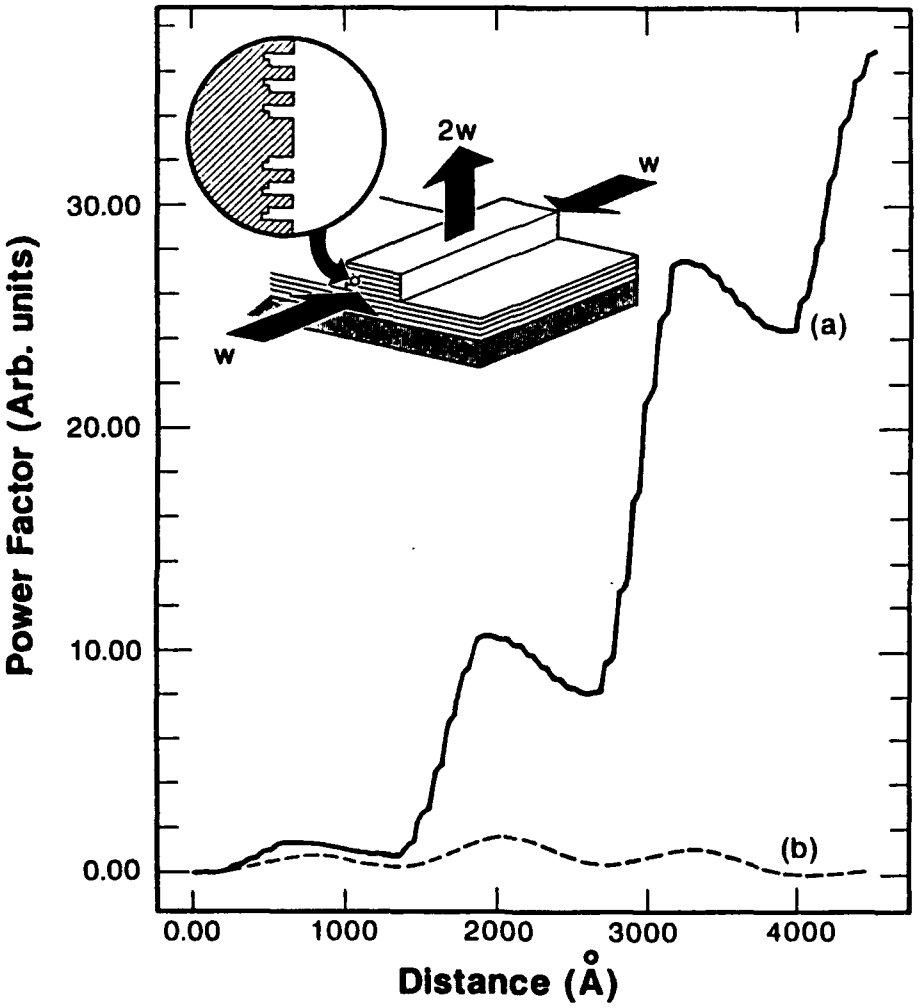
Ongoing work is concerned with experimental verification of the passive waveguide structure and integration of the concept with active edge-emitting semiconductor lasers.

Publications

[1] P. J. Harshman and S. Wang, " Asymmetric AlGaAs quantum wells for second-harmonic generation and quasi-phase matching of visible light in surface emitting waveguides," submitted to *Applied Physics Letters*.

SECOND-HARMONIC GENERATION OF VISIBLE LIGHT IN SURFACE EMITTING WAVEGUIDES

Joint Services Electronics Program
at U.C. Berkeley, EECS Department
P.J. Harshman and Prof. S. Wang



Second-harmonic ($\lambda^{2w} = 540nm$) power factor versus vertical distance in surface emitting waveguide

(a) guiding region consists of phase-matched asymmetric quantum well domains
(b) bulk $Al_{0.5}Ga_{0.5}As$ guiding region

Inset is a diagram of the multiquantum well waveguide geometry with a blow-up of the conduction band edge.

Effect of Bipolarity Oxide Stress on MOSFET Reliability

Professor C. Hu with E. Rosenbaum and Z. Liu

Even though electrical stressing of thin oxides is known to cause oxide breaking and MOSFET mobility and transconductance degradations due to the generation of interface traps, high frequency pulse stress or bipolarity AC stress of thin oxides has not been investigated. We use three different waveforms DC, pulsed DC and bipolarity AC to study the behavior of n-MOSFETs after dynamic oxide stress. Oxide thickness is 11 nm and 8.5 nm. Time to breakdown t_{BD} was found to be a strong function of frequency under bipolar stress and its value is larger than that under unipolar to DC stress (Fig. 1). Note that at low frequencies, bipolar t_{BD} approaches the unipolar value, while at high frequencies, bipolar t_{BD} saturates. A study of this phenomenon at a variety of electric fields shows that the saturation frequency is a strong function of electric field. Interface trap density saturates quickly in DC and unipolar stressed devices; however, it continues to increase in bipolar stressed devices as can be seen from Figure 2. As a result much larger degradation of MOSFET mobility and transconductance was created by bipolarity stressing. There is evidence that hole trapping in oxide precipitates breakdown (see, for example, [1]). Holes are generated near the anode in thin oxides; some of these may become trapped. To explain the observation that bipolar t_{BD} is longer than unipolar t_{BD}, we hypothesize that the field reversal impedes movement of holes toward the bulk and that the field reversal enhances detrapping. Hole transport in SiO_2 has been shown to be quite dispersive in time and strongly electric field activated [2]; this is the same behavior displayed by bipolar t_{BD}. If one interprets time to breakdown as inversely proportional to the number of holes trapped in the bulk per unit time, then our bipolar stress t_{BD} data can be explained by hole transport. Trapped holes which hop toward the cathode under the influence of the electric field are less likely to be detrapped when the field reverses, since the probability of detrapping decreases with increasing distance from the interface. The scenario of hole generation near the anode and detrapping when the field reverses

can also explain why we observe large interface trap generation under bipolar stress. Interface traps are known to result when electrons recombine with holes trapped near the interface.

The high rate of interface trap generation and long t_{BD} under bipolarity stress suggest that interface trap generation is not responsible for oxide breakdown. Since the gate-drain overlap region of MOSFETs in circuits is subject to bipolar stress, DC transistor stress tests may underestimate the MOSFET degradation rate.

This research won the Best Student Paper Award at the 1991 International Electron Device Meeting (IEDM) [3].

References

[1] I.C. Chen, et al., *IEEE Trans. on Electron Devices*, Vol. 32, no. 2, p. 413, 1985.

[2] F.B. McLean, et al., "Electron-Hole Generation, Transport and Trapping," in *Ionizing Radiation Effects in MOS Devices and Circuits*, ed. T.P. Ma and P.V. Dressendorfer, 1989.

[3] E. Rosenbaum, et al., "The Effects of Oxide Stress Waveform on MOSFET Performance," IEDM, p. 720, 1991.

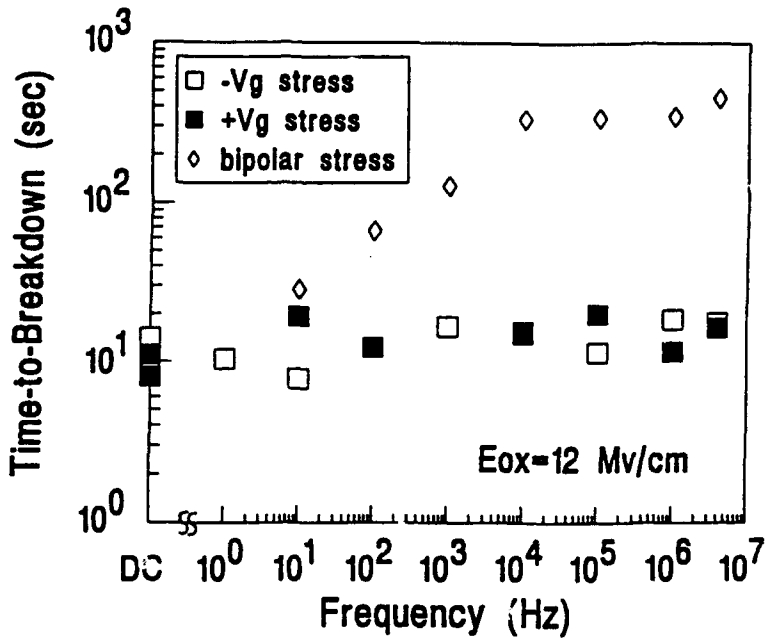


Fig. 1 t_{TD} under DC and unipolar dynamic stress conditions is similar but is longer and frequency dependent under bipolar stress.

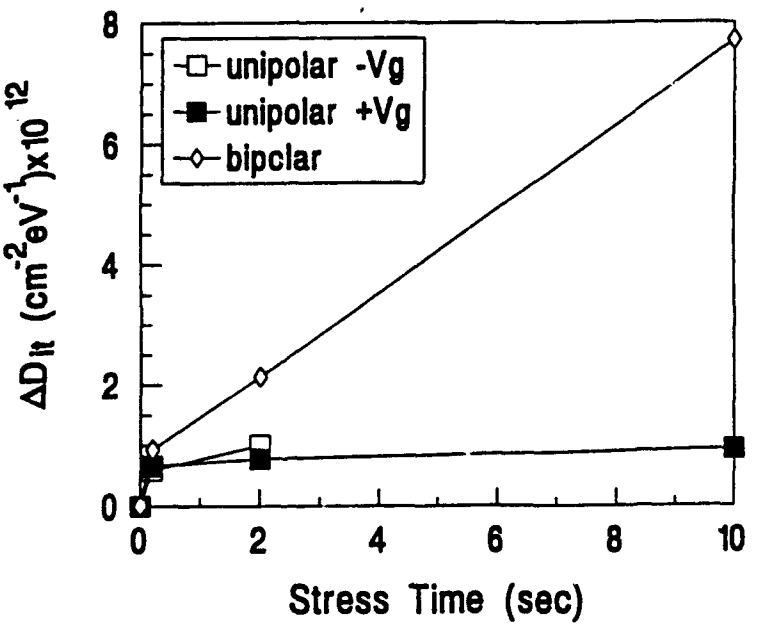


Fig. 2 Interface trap generation quickly levels off in unipolar stressed devices but not in bipolar stressed devices.

Insulated-Gate GaAs Field Effect Transistors

Professors N.Cheung, S.Wang, C.Hu and W.Oldham with James Chan

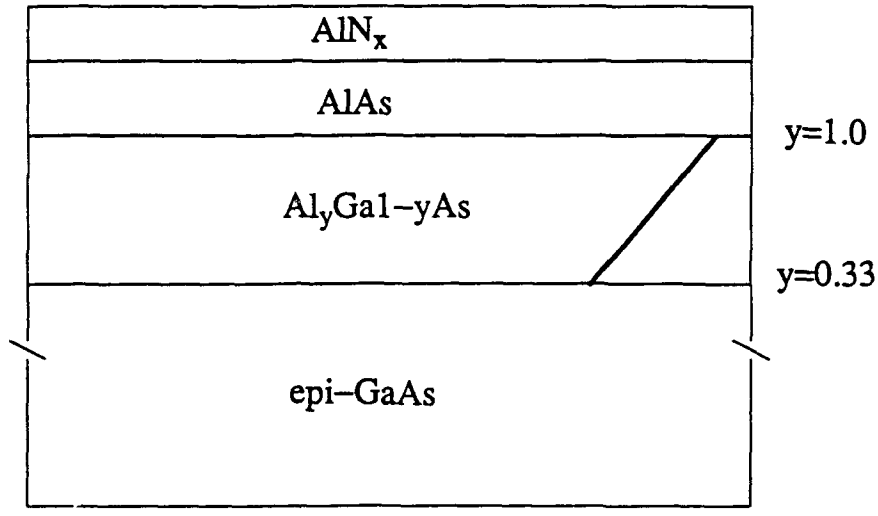
The viability of an insulated-gate material for GaAs metal-insulator-semiconductor FETs (MISFETs) has been actively pursued for many years. The main advantage of MISFETs over conventional GaAs MESFETs is the presence of a much higher barrier height, which gives two benefits: a decrease of the gate leakage current due thermionic emission of electrons across the barrier, and an increased voltage swing, which compared to the small 0.7-0.8V present in the MESFETs, makes GaAs MISFETs extremely attractive for digital applications.

We have successfully synthesized AlN_x by reactive sputtering of an Al target in a nitrogen plasma. In addition, the molecular beam epitaxial substrate (MBE) for the AlN_x thin film has also been fabricated. This MBE substrate employed a 200Å AlAs layer which was epitaxially grown on a 400Å of graded $Al_yGa_{1-y}As$ layer originating from y=1.0 at the $AlAs/Al_yGa_{1-y}As$ interface to y=0.33 at the surface of the 5000Å GaAs buffer layer which in turn was grown on top of the GaAs substrate. This MBE growth was carried out in collaboration with Professor Shyh Wang's MBE research activities.

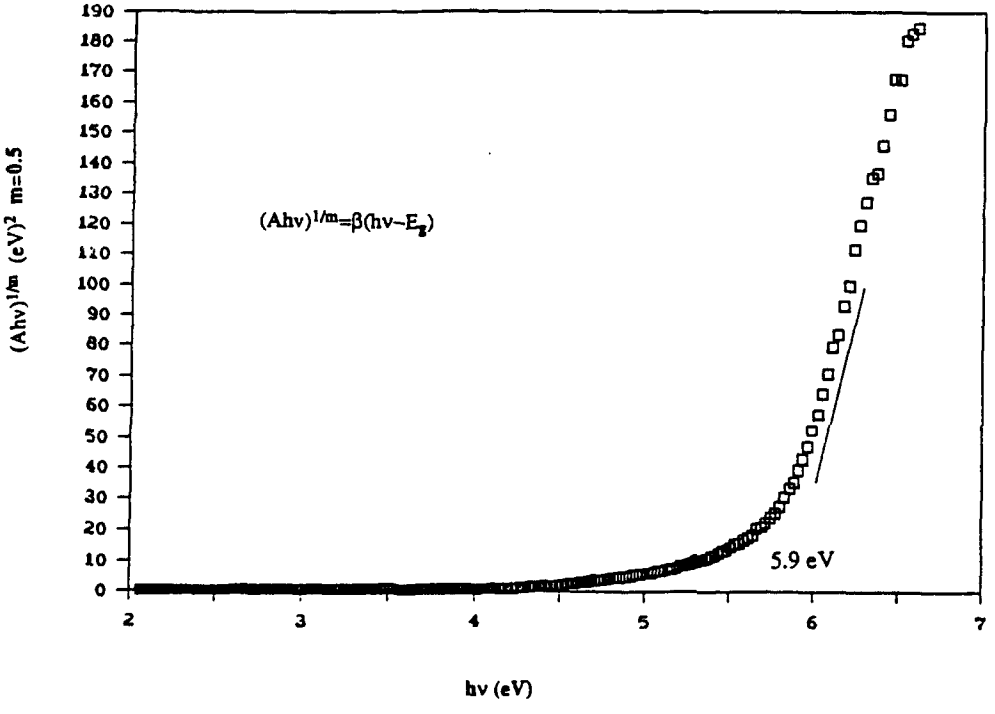
Furthermore, optical absorption measurements were performed on the AlN_x samples, from which an energy gap value of 5.9eV was obtained. In addition, an average index of refraction value of 2.064 was obtained. Electrical measurements also revealed that the breakdown electric field of the AlN_x material was around 10^6 V/cm.

Publications

[1] James Chan, "Low Energy Ion Beam Modification of AlN_x Thin Film for Insulated Gate Field Effect Transistors," to be presented at the MRS Spring Meeting, San Francisco, April 27 - May 1, 1992.



MISFET Structure with graded $Al_yGa_{1-y}As$ grown by Molecular Beam Epitaxy and gate insulating AlN_x by reactive sputtering.



Optical Absorption measurement to determine the bandgap of AlN_x (= 5.9eV).

Automatic Classification of Active Sonar Returns Using Time-Frequency Techniques

Professor A. Zakhor with F. Lari

We address the classification of one dimensional signals using Time-Frequency(TF) techniques. TF techniques have the basic advantage of a physically meaningful interpretation that can not be claimed by classical detection methods such as matched filtering; thus human observers can interpret and classify the return signal from a target in the TF domain and not in the time domain.

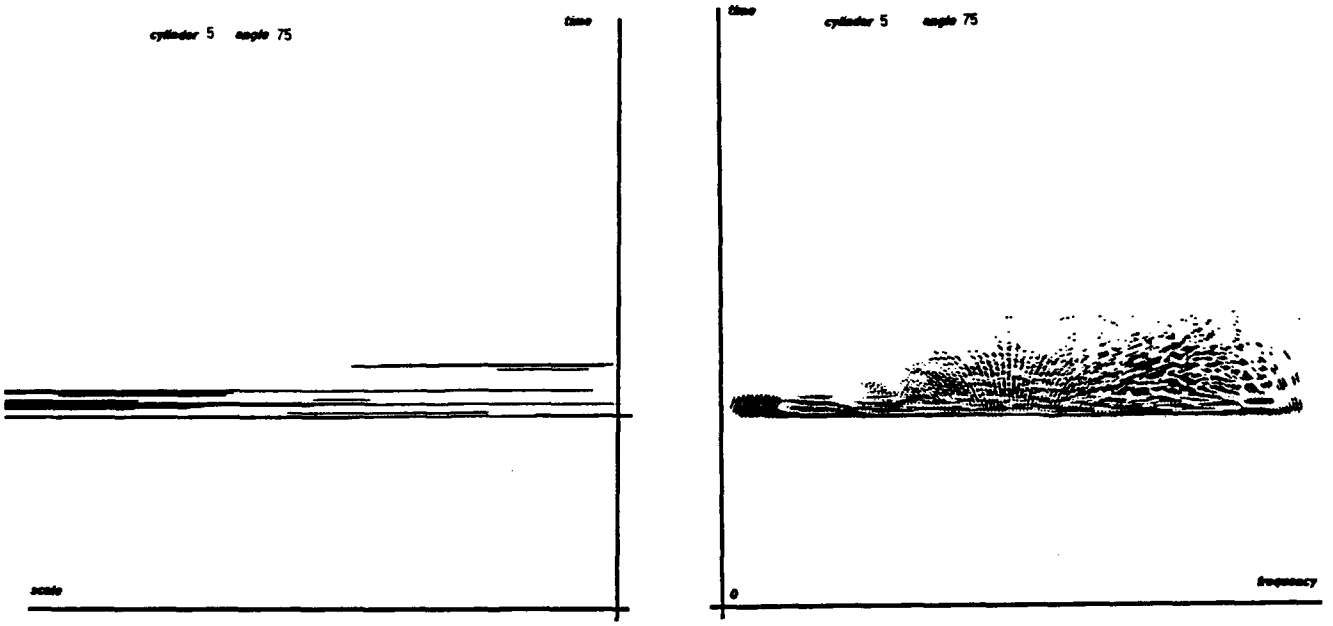
Based on a series of considerations, we have chosen as TF Transforms the Wigner-Ville Transform and the Wavelet Transform. From these Transforms, we extract features by means of integration over appropriate areas of the TF plane. The features are then classified by a decision tree.

Our data consists of computer simulated sonar returns from different objects and different angles of observation, resulting in a total of 6 different classes and 18 different category(class and angle of observation). We add random White Gaussian Noise of up to -4dB of Signal-to-Noise Ratio(SNR) to the original data.

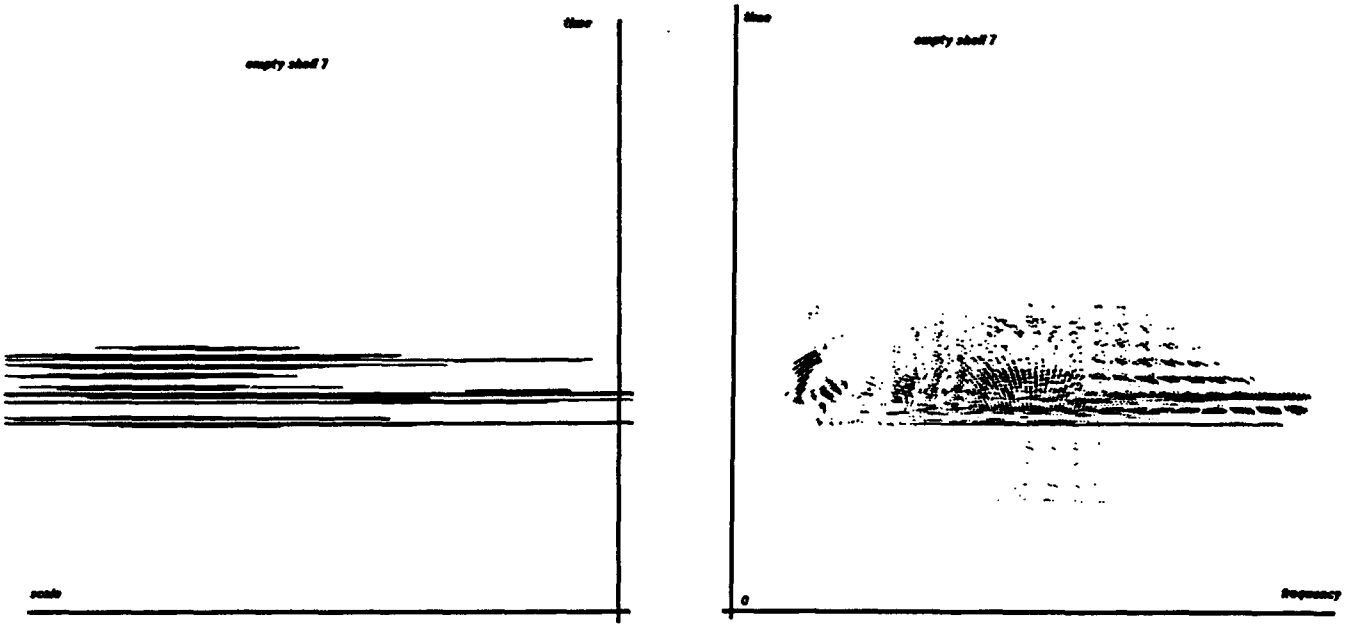
When considering the Wigner-Ville Transform, correct classifications rate are about 95% even with less than 0dB of SNR, outperforming any known classification algorithm.

Publications

[1] F. Lari and A. Zakhor "Automatic Classification of Active Sonar Returns using Time-Frequency Techniques", submitted to SPIE's 1992 International Symposium on Optical Applied Science and Engineering, San Diego, July 1992.



TF Transforms of a 5:1 finite steel cylinder with hemispherical endcaps for a 75^0 off axis incidence



TF Transforms of an empty steel shell

Wigner-Ville Transform(right) and Wavelet Transform(left)

Optimal Adaptive K-Means Algorithm with Dynamic Adjustment of Learning Rate

Chedsada Chinrungrueng and Professor Carlo H. Séquin)

Many competitive learning procedures have been developed for unsupervised learning in artificial neural networks. The most commonly used competitive learning algorithm is the k-means clustering algorithm. It has been used for processing the input data of complicated classification tasks, e.g., in radial basis function networks, or in feature-map classifiers. We have developed an adaptive k-means clustering algorithm that provides an optimal or near-optimal clustering solution with rapid convergence rate, and which can also be used in situations where the statistics of the problem task are varying slowly with time.

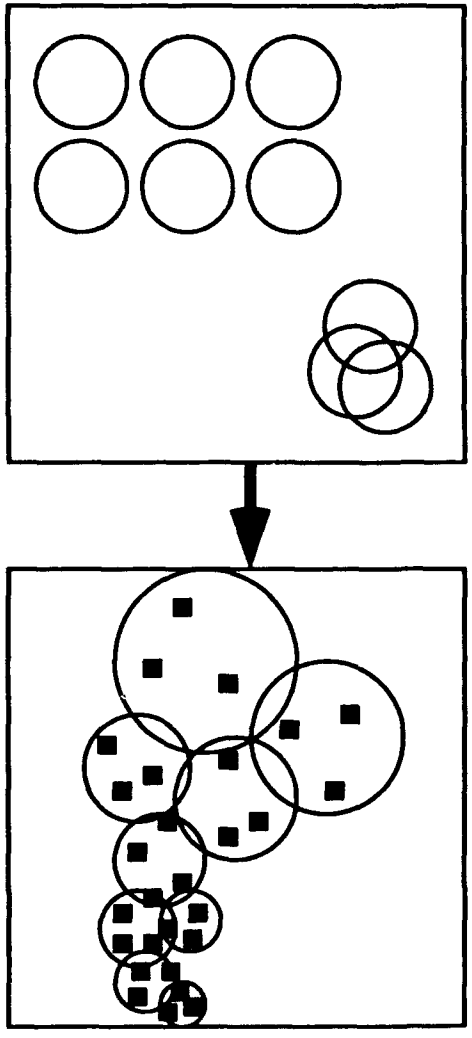
We have added to the traditional k-means algorithm two mechanisms. The first mechanism is to bias the clustering process towards an optimal solution, and the second mechanism is to dynamically adjust the learning rate based on the quality of the clustering. These two mechanisms are based on the necessary condition for optimality of the k-means partition which states that, when the cluster regions in the partition is large, all of the regions in an optimal k-means partition have the same *within-cluster variation*. This within-cluster variation of any cluster is defined as the sum of the squared Euclidean distances between the pattern vectors in that cluster and the center of that cluster. Both, the biasing of new data points towards clusters with a smaller variation and the estimation of how close the current clustering is to optimal are based on a time averaged estimation of this within-cluster variation and require no parameters to be set by the user.

Publications

- C. Chinrungrueng and C. H. Séquin, "Optimal Adaptive K-Means Algorithm with Dynamic Adjustment of Learning Rate," Proc. IEEE IJCNN-91, pp I-855-862, Seattle, July 1991.

JOINT SERVICES ELECTRONICS PROGRAM
AT U.C.BERKELEY, EECS DEPARTMENT
PROF. CARLO H. SÉQUIN

AN OPTIMAL ADAPTIVE K-MEANS ALGORITHM WITH DYNAMIC ADJUSTMENT OF LEARNING RATE



ALGORITHM OVERVIEW:

- Start available partition-clusters anywhere in input space (at origin, in grid, or random).
- For each input pattern presentation:
  - Move closest, least loaded cluster towards data point.
- Keep a running average of the **variation** in each cluster.
- Adjust extent of receptor and scale its distance metric to keep the **variations** of all receptors about the same.
- Use **variation** variances of the clusters to adjust the step size of cluster movement (=learning rate).

APPLICATION:

- Automatic input space partitioning for learning problems.
- Very general ! - User need not adjust any parameters !

Interconnection Network Design Based on Packaging Considerations

M. T. Raghunath (Professor Abhiram Ranade)

A central issue in massively parallel computation is the design of the interconnection network between processors. The choice of the network significantly affects the total cost, as well as the ease of programming such machines.

The goal of our research is to develop high performance, cost effective networks for interconnecting processors. While network design depends upon many factors, our primary concern is packaging technology. Large networks need to be partitioned and packaged in a hierarchical manner into chips, boards, and racks. We are exploring hybrid designs in which the lower levels of the hierarchy use a denser network than the one used at the higher levels. While it is obvious that such networks improve local communication performance (e.g. within a board), we have also observed that they reduce the latency for long distance communication (e.g. between racks). We have evaluated several alternative network designs using detailed simulation. Our results indicate that a significant improvement in network performance is achievable for a small increase in total cost. We have found that organizing a given number of wires as a single wide channel rather a number of dilated channels, provides better network performance. Our simulations also indicate that utilization is improved if we make channels wider by reducing the total number of channels leaving a board or rack. We have also observed that providing the processors with the ability to run multiple threads of execution is successful in hiding the network latency. We can improve processor utilization by almost a factor of 3 by combining all of the above techniques.

Our current research is directed towards continued exploration of the network design space. We also plan to model inter-processor synchronization in greater detail. We will validate our results by simulating real parallel programs on promising network architectures.

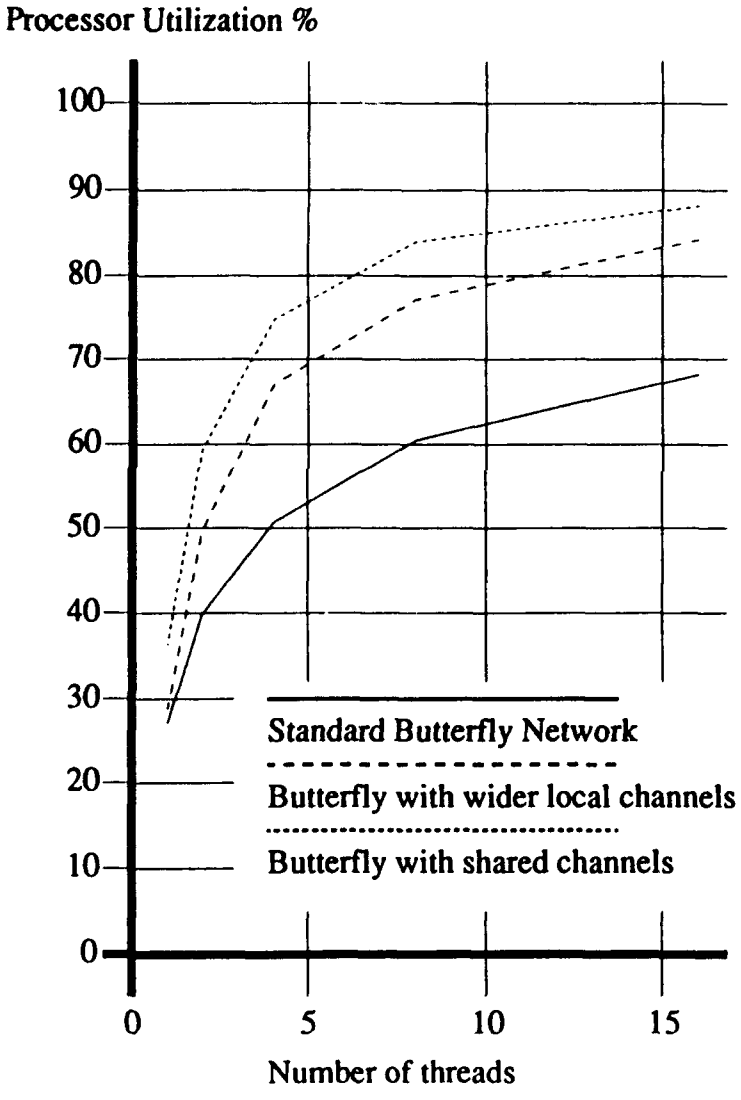
Publications:

[1] M. T. Raghunath and A. G. Ranade, A Simulation-Based Comparison of Interconnection Networks. In *IEEE Symposium on Parallel and Distributed Processing*, pp 98–103, December 1990.

JOINT SERVICES ELECTRONICS PROGRAM
AT U. C. BERKELEY, EECS DEPARTMENT

PROF. ABHIRAM RANADE

INTERCONNECTION NETWORK DESIGN BASED ON PACKAGING CONSIDERATIONS



RESULTS:

Network Bandwidth and Latency can be improved significantly by tailoring the network architecture to the available packaging technology

Local interconnections can usually be denser than remote interconnections. Exploiting the increased density by using wider channels is better than increased dilation

The remote interconnections should be organized as fewer number of wider channels by sharing these channels among the processors on a board.

Multi-threaded processors can successfully hide the network latency, improving processor utilization

By using a combination of all the above, processor utilization can be increased from about 30% to almost 90%.

PART C - INDIVIDUAL WORK UNITS

I-A. Nonlinear Optics in Compound Semiconductors

Professor Shyh Wang with Patrick Harshman

Phase matching is of fundamental importance in nonlinear optics. Our work during the past year has been focused on theoretical study of asymmetric quantum wells for phase-matching in surface-emitting second-harmonic generators. This novel scheme offers the best hope for cumulative build-up of second-harmonic power.

In the original surface-emitting second-harmonic generator, developed previously under JSEP support, two oppositely propagating guided waves interact through the nonlinearity of the bulk semiconductor guide material to generate sum-frequency waves which propagate normal to the waveguide surface. The second-harmonic power out from this device is limited by the fact that a nonlinear wave can not be phase-matched in a bulk III-V semiconductor material. Our present work concerns the development of new, higher output power devices which overcome the phase-matching problem.

We have approached the problem of phase-matching the second-harmonic signal by using asymmetric quantum wells as sources of the nonlinear response in the waveguide devices [1]. The previous experimental work has concentrated on near resonance intersubband processes within a quantum well conduction band and corresponding mid-infrared radiation. Of more practical interest for the surface-emitting second-harmonic generator is the nonlinear optical conversion of near-infrared wavelengths available from existing semiconductor laser technology to visible wavelengths. For the AlGaAs material system this necessitates the use of interband transitions. Specifically, we consider the nonlinear optical process which involves one valence and two conduction subbands. We have calculated the second-order susceptibility for this process to be:

$$\chi^{(2)}_{xzx}(\omega) = \sum_{g,n',n} \frac{e^3 \mu_{n'g} \langle \psi_n | z | \psi_{n'} \rangle \mu_{gn}}{\epsilon_0 \hbar^2 \pi L_z} \frac{m_r}{\hbar(\omega - \omega_{nn'} + i(\Gamma_{ng} - \Gamma_{n'g}))}$$
$$\times \ln \left[\frac{(2\omega - \omega_{ng} - \hbar k_0^2/2m_r + i\Gamma_{ng})(\omega - \omega_{n'g} + i\Gamma_{n'g})}{(2\omega - \omega_{ng} + i\Gamma_{ng})(\omega - \omega_{n'g} - \hbar k_0^2/2m_r + i\Gamma_{n'g})} \right]$$

The two important results of our calculation are as follows. First, there is a sign reversal between $\chi^{(2)}_{xzw}$ of an asymmetric of an asymmetric quantum well and that of the inverse quantum well. Second, the peak magnitude of the calculated $\chi^{(2)}_{xzw}$ is 4 X 10^{-10} m/V, which is approximately equal to that of bulk GaAs.

The fact that $\chi^{(2)}$ for the interband quantum well structure can have a magnitude comparable to that of bulk GaAs, coupled with its ability to reverse sign, makes the interband scheme very attractive for integration into the surface-emitting second-harmonic generation device. Specifically, since the nonlinear wave propagates normal to the surface, asymmetric quantum wells in the guiding region can be used to phase match the nonlinear signal. We have modeled second harmonic generation for the multiquantum well waveguide depicted in Fig. 1. The second-harmonic power at λ^{2w}=540nm for both the multiquantum well waveguide and a waveguide with a bulk $Al_{0.5}Ga_{0.5}As$ guiding region is plotted versus vertical distance in curves (a) and (b), respectively. As Fig. 1 indicates, the second-harmonic signal generated via the quantum well interaction can be effectively phase matched by selectively inverting the asymmetry of those wells which would otherwise contribute a destructively interfering second-harmonic signal. For a guide thickness of 4500 Angstrom (approx 3.5 coherence lengths), the second-harmonic power radiated from the surface of the multiquantum well waveguide is calculated to be over a factor of 100 times greater than the power radiated from a bulk $Al_{0.5}Ga_{0.5}As$ guide. The fine structure in the plot Fig. 1(a) arises from the difference between the nonlinear response of the bulk barriers and the asymmetric quantum wells. We are currently engaged in the experimental evaluation of the proposed device. Also of interest to us in an analgous device based on (111)

self-biased strained quantum wells, where the asymmetry is achieved via the self bias effect [2].

A major limitation of the original device is that of the limited fundamental power which can be coupled into the guiding region from an external light source. To overcome this limitation, we are developing an active device in which the surface second-harmonic is emitted from the active layer of a modified edge-emitting semiconductor laser. Improved performance can be expected from such a device because of the large intensity of the fundamental laser cavity oscillation. Also, the device will be more compact since no external input light source is required. Our final goal is to build an integrated laser-second harmonic-generator capable of delivering visible coherent radiation in the mW power range.

Publications

[1] P. J. Harshman and S. Wang, "Asymmetric AlGaAs Quantum Wells for Second-Harmonic Generation and Quasi-Phase Matching of Visible Light in Surface Emitting Waveguides," to appear in *Appl. Phys. Lett.*, March, 1992.

[2] P. J. Harshman and S. Wang, "Investigation of (111) Strained-Layers: Growth, Photoluminescence, and Internal Electric Fields," accepted for publication in *J. Appl. Phys.*

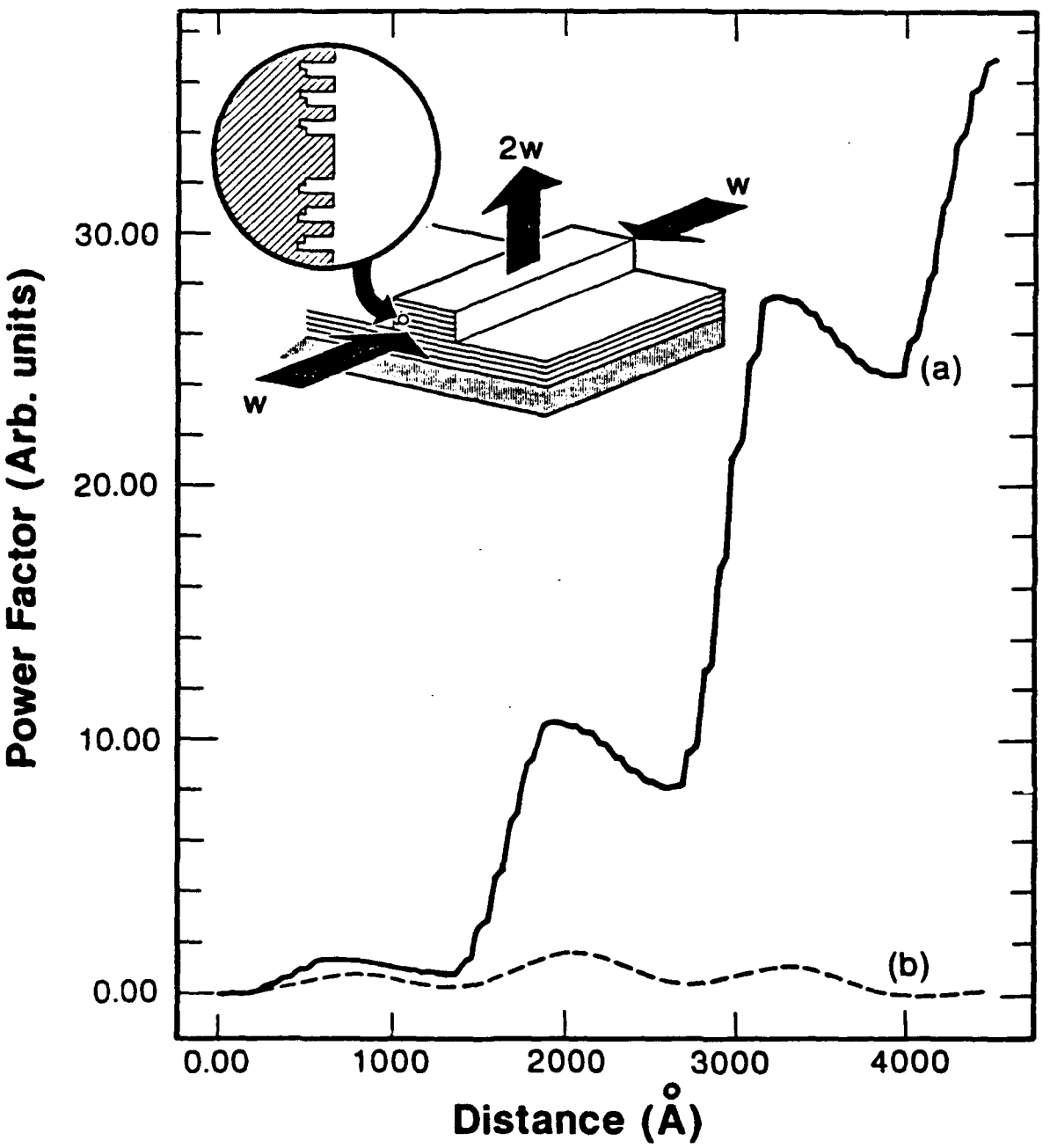


Figure 1

[a] Groups of asymmetric quantum wells in antisymmetric arrangement for every λ/4.

[b] Bulk GaAs guide.

I-B. Ultrafast Optical Techniques

Professor J. Stephen Smith with Hong Lin, Jeffrey Walker, Gordon Wilson and Dan Kutchta

Vertical Cavity Lasers Fabricated Using Phase Locked Epitaxy

Our vertical cavity surface emitting lasers include devices which produced maximum output powers of up to 10 mW cw, which is the highest power demonstrated to date for these types of lasers. They also exhibit low series resistance, with turn on voltage of 2.6 volts, also the lowest reported, and unparalled wafer to wafer yield. This project has established phase locked epitaxy (PLE) as a valuable technique for the fabrication of surface-emitting lasers and other semiconductor structures which have precision materials requirements. These advances in the ability to fabricate these lasers reliably, with virtually every wafer grown lasing, should be very important for further successes in these areas. Our PLE technique, using short period superlattices with growth interruptions for all of the device layers, has applications for other precision layered structures as well.

Group Velocity Dispersion in Quantum Well Lasers

Monolithic mode-locked semiconductor lasers are able to produce short optical pulses at extremely high repetition rates. They have potential for a very wide range of practical applications due to the their compactness and low energy consumption. Group velocity dispersion in a laser cavity has a significant influence on the mode-locking process. The effects of the dispersion are particularly strong for a monolithic mode-locked laser, because both the gain and waveguide regions consist of semiconductor materials which have large group velocity dispersion. In many practical applications, the repetition rate of a mode-locked laser is required to be less than 10 GHz, therefore the cavity length of a laser must be larger than several millimeters.

We have developed a technique using a density matrix formalism to calculate the effective index of a quantum well laser under current excitation, which agrees with experimental observa-

tions by others. Applications of this calculation to various multiple quantum well laser structures demonstrate that the dispersion can be reduced significantly with proper design.

Our results indicate that the quantum-well gain region has opposite dispersion to the dispersion in the cladding region. Therefore with proper design, the quantum well laser can have significant less dispersion.

We will be making a comparative measurement of the dispersion of a single quantum well laser and a multiple quantum well laser with the same cladding region, which should give a good measure of the dispersion of quantum wells under current excitation.

Publications

[1] Hong Lin, and J. S. Smith, "Compensating Group Velocity Dispersion in Quantum Well Lasers," presented at the Engineering Foundation Conferences, High Speed/High Frequency Optoelectronics, Palm Coast, Florida, March, 1991 and in February, 1992, submitted for publication in *Appl. Phys. Lett.*

[2] S. P. Dijaili, J. M. Wiesenfeld, G. Raybon, C. A. Burrus, A. Dienes, J. S. Smith and J. R. Whinnery, "Cross Phase Modulation in a Semiconductor Laser Amplifier Determined by a Dispersive Technique," *IEEE Journal of Quantum Electronics*, Vol. 28, No. 1, 141-150, January, 1992.

[3] J. D. Walker, Dan Kutchta, and J. S. Smith, "Vertical-Cavity Surface-Emitting Laser Diodes Fabricated by Phase-Locked Epitaxy," *Appl. Phys. Lett.*, Vol. 59, No. 17, pp. 2079-81, October, 1991.

[4] Hong Lin and J. S. Smith, "Optical Time Division Demultiplexing Using Second Order Optical Nonlinear Effects," *Appl. Phys. Lett.*, Vol. 59, No. 22, November, 1991.

I-C. Optical Probing of Semiconductor Devices and Interfaces by Electro-Optic and Photo-Elastic Effects

Professor S. Wang with Mark Hadley

Built-in electric field plays an important role in the transport of carriers in semiconductor heterostructure devices. The so-called bandgap engineering derives its principal benefit from our ability to aid or impede the transport of electrons or holes separately by properly varying the spatial distribution of built-in electric field in a heterostructure device. Here we present a new approach to band structure engineering which can lead to new concepts. One fundamental problem of current interest and practical importance is the minimum resistance of a multi-heterostructure device. It is known that the heating effect in surface-emitting laser, a multi-heterostructure device, severely limits the laser performance.

The series resistance of a heterojunction device is caused by the potential barrier existing at an abrupt heterojunction due to band-edge discontinuity and the resultant space-charge field associated with majority-carrier depletion on one side and majority-carrier accumulation on the other sides of the junction. To eliminate this series resistance, we must use a heterostructure with a continuously varying composition. If x denotes the composition of a heterostructure, for example, made of $Al_xGa_{1-x}As$, and z denotes the direction of the heterostructure, then x must be a continuous function of z. However, this step alone is insufficient. Next, we propose to use graded doping, for example, acceptor concentration $N_a(z)$ in a p-type multi-heterostructure, to generate a space-charge potential $\Phi(z)$ such that

$$E_v(z) = \Delta_v(z) + \Phi(z) = constant \tag{1}$$

where $\Delta_v(z) = E_v(z) - E_v(O)$ is the valence-band change with composition. Physically, Equation (1) says that the combined effects of the space charge and the band-edge change are to produce a flat profile for the potential energy of holes. Under the circumstance, the resistance of a multi-heterostructure is reduced to that due to ohmic conduction alone, and a minimum resistance is

reached.

The equation governing the space charge density $P(z)$ is the generalized Poisson's equation:

$$-\frac{\partial}{\partial z}\left[\varepsilon\frac{\partial\Phi(z)}{\partial z}\right]=\rho(z)=-e\left[p-N_a^-(z)\right] \tag{2}$$

Eliminating $\Phi(z)$ from Equations (1) and (2), we find

$$N_a^-(z)=p(z)+\frac{1}{e}\frac{\partial}{\partial z}\left[\varepsilon(z)\frac{\partial\Delta_v(z)}{\partial z}\right] \tag{3}$$

where ε is the permittivity of $Al_xGa_{1-x}As$ and p is the hole density. Under thermal equilibrium, the Fermi level E_f remains constant throughout the multi-heterostructure. Therefore, for flat profile of hole potential energy,

$$p(z)=\frac{p(o)N_v(z)}{N_v(o)} \tag{4}$$

where N_v is the effective density of valence-band states.

Equations (3) and (4) provide the fundamental rule for the design of a multi-heterostructure with minimum resistance, that is, with ohmic resistance only. For $Al_xGa_{1-x}As$ heterostructures, the various quantities depend on x as follows:[1,2]

$$N_v(x)=\left[1+0.4x\right]N_v(o)$$

$$\varepsilon(x)=(13.13-3.12x)\varepsilon_o$$

$$\Delta_v(x)=0.55x$$

For a given composition variation of x with z, the doping variation $N_a(z)$ can be found from Equation (3). As an example, we choose $p_o=4\times10^{18}cm^{-3}$ in GaAs and $x=(1-\cos kz)/2$ with $k=2\pi/400\text{Å}$. Figure 1 shows (a) the required acceptor concentration calculated from Equation (3) and (b) the resultant band structure. The solid curves are for the segment between $z=200\text{Å}$ and 400Å. The solid curve in Figure 1b is an extended structure. The physical parameters $p_o=4\times10^{18}cm^{-3}$ and period $=2\pi/k=400\text{Å}$ are well within the range achievable in practice. For

the band structure of Figure 1b, the resistance of the multi-heterostructure is determined entirely by the ohmic resistance.

Two important points regarding our approach to band structure engineering is worth noting. First, the conduction band edge is affected by the acceptor distribution and not by the donor distribution which is the usual case. Second, the composition change and hence the impurity distribution are not limited to the sinusoidal variation. Therefore, we can design any spatial profile of the conduction band edge by properly choosing the functional dependence $x(z)$ and $N_a(z)$. The sinusoidal variation is of special interest because for small x, $x(z)$ can be approximated by

$$x(z) = \left\{\frac{1-\cos kz}{2}\right\} = \frac{k^2z^2}{2}$$

which is the potential well for a harmonic oscillator. Because the energy levels in a parabolic well are all equally spaced, resonance energy transfer between wells can take place simultaneously between many pairs of well energies. Furthermore, if we place n^+ and n contacts to the multi-heterostructure, the electrons will all emit photons of the same energy as they cascade down from higher to lower well-energy levels. These novel device concepts will be further studied for the invention of new resonant tunneling device and monochromatic infrared generator using intraband transitions.

References

[1] J. Batey and S.L. Wright, *J. Appl. Phys.*, Vol. 59, p. 201, 1986.

[2] S. Adachi, *J. Appl. Phys.*, Vol. 58, p. R1, August, 1985.

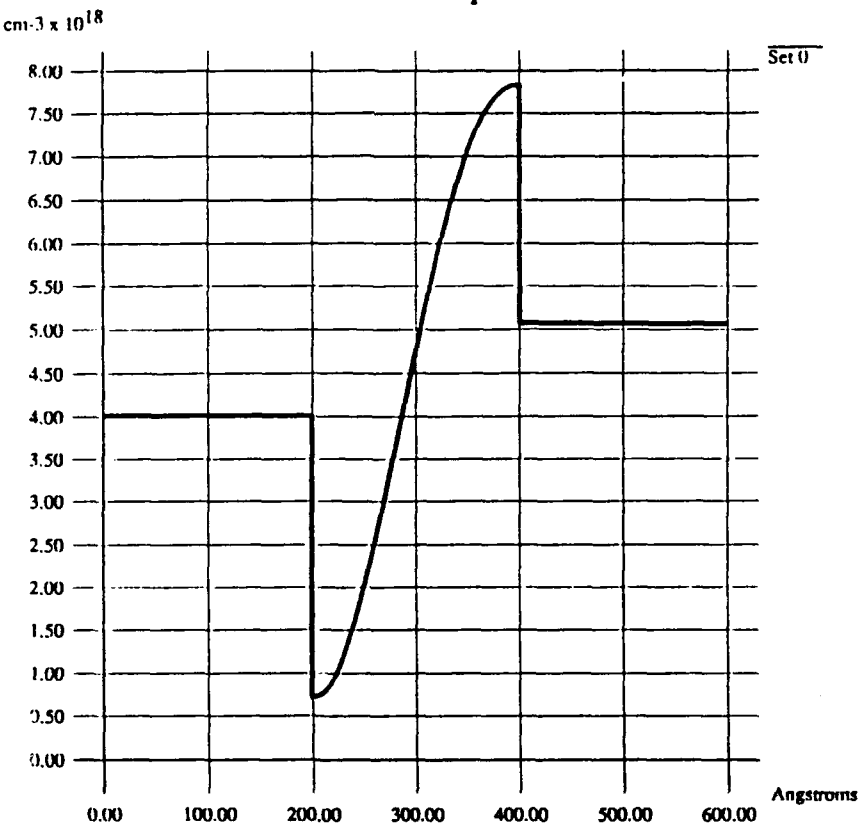


Figure 1(a)

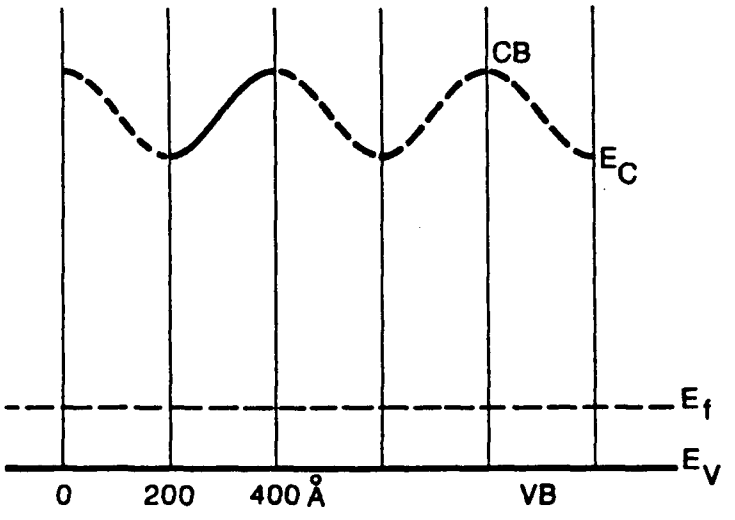


Figure 1(b)

II-A. 0.1 μm BICMOS Devices in Bulk and SOI Substrates

Professors C. Hu and P.K. Ko with F. Assaderaghi, K. Quader, K. Schuegraf, M. Chan, T. Hanif, J. Krick

In this period, we have successfully fabricated the world's first bipolar transistors in SIMOX (separation by implantation of oxygen) substrates. SIMOX substrate is nearly the universal choice for SOI CMOS device research because of its excellent thickness uniformity and the possibility of eliminating punchthrough and drain-induced barrier- lowering so that 0.1 μm CMOS device dimensions can be achieved. SOI bipolar transistors, however, have previously been made in thicker "wafer bonding" substrates in the normal *vertical* fashion. We have designed several lateral bipolar transistor structures that can be fabricated in thin SOI silicon films. These structures can easily be made in the complementary fashion, i.e. npn and pnp transistors. Such complementary devices are believed to be essential for low voltage operation compatible with 0.1 μm BICMOS or bipolar technologies. These SOI bipolar transistors have gains higher than 100 over 5 orders of magnitude of collector current. This indicates that we have achieved an even lower non-ideal base-emitter current in these SOI materials than even in bulk silicon. The high current gain is partially due to the 0.1 μm base width in these lateral transistors accomplished with our unique photoresist ashing technique. Full characterization of these devices is underway.

We have continued to improve the photoresist ashing technique and can now consistently fabricate MOSFETs of 0.2 μm channel length. We are not satisfied with the 0.1 μm MOSFETs fabricated to date, although they have already provided valuable information about 0.1 μm devices. In addition we are developing a technique of defining the gate length not by lithography but by film thickness. This maskless technique of defining 0.1 μm gate will take advantage of the "spacer" technology perfected for LDD structures in the past five years. Initial results are encouraging.

Fifteen papers were published in this period [1-15]. The performance and reliability issues of MOSFET with channel lengths down to 0.1 μm were examined [1]. It was concluded that adequate oxide lifetime, off-state leakage current, hot carrier lifetime, short-channel effect can be accomplished in MOSFET with 0.1 μm channel length. This work heightens the need to understand the limits and characteristics of MOSFETs below 0.1 μm. The most serious limitation of 0.1 μm MOSFETs is punchthrough [12]. We have demonstrated a novel MOSFET structure with source-drain regions positioned above the channel plane, rather than below it. This is done with selective epitaxial growth of silicon [2]. A near zero, or even negative junction depth is thus possible to achieve superior punchthrough characteristics.

Even better punchthrough performance can be obtained from SOI MOSFETs. SOI technology also eliminates the bird's beak problem of isolation and the latch-up problem. We have found, however, a problem due to the lower threshold voltage at the field edge [3]. This led to a solution -- using pt polysilicon as the gate material. We have proven this solution in recent experiments and will publish the results soon. We have developed the first non-destructive technique of measuring the SOI silicon film thickness using finished fully depleted MOSFETs [6] and have shown that most of the observed large V distribution is due to the variation in SOI film thickness. We have found that the bottom silicon/oxide interface quality is surprisingly good, comparable to the usual Si/SiO_2 inferaces on bulk substrates. This was proven using both the noise measurements and the mobility measurements [9]. The floating body of the SOI transistor exhibits some peculiar transient behavior. This can be seen as a time dependent threshold voltage with many seconds of transient time when a step function is applied to the back gate. The phenomenon is similar to the "back-gating" effect of GaAs FET's. We investigated this and explained it [10] by the finite time required to form the inversion layer at the bottom Si/SiO_2 interface through thermal generation. The solution is to choose doping concentration and type of substrate (not the silicon film) so that the bottom is not inverted under all biasing conditions [8].

Low breakdown voltage is a known problem of SOI MOSFETs. The related hot carrier reliability issue is largely ignored so far. One major problem is that the floating body does not allow the measurement of the substrate current, which has yielded most of the information on hot carriers and channel field in bulk MOSFETs. We have demonstrated that the gate current due to channel hot electron injection can yield the same information [7] and used the gate current to develop a hot electron model for SOI devices. Furthermore, we have developed the first quantitative model of the SOI MOSFET breakdown voltage [13]. The voltage capability for SOI devices for present future channel length, silicon film thickness, oxide thickness, etc., can be predicted. This set the benchmark again which new device structures will be tested and indeed the need for new device structures will be determined.

At the same time, we have isolated a single interface trap generated by hot electrons and fully characterized its time constants, energy, physical location, and effect on carrier mobility [5]. These basic data are needed for a full understanding of the complex phenomenon of hot electron induced device degradation. We have developed a unified model that can model both the hot electron effect (channel electric field) and also the short channel effect [12]. As a result of this in-depth research we have made a projection of the MOSFET device scaling to just below 0.1 μm [11], which will probably become a commercially viable technology in the year 2003. One surprising conclusion is that oxide direct tunnel at 40Å will set the limit for MOSFET scaling at 0.09 μm unless some new materials or approaches are introduced.

Publications

[1] J.E. Chung, M.C. Jeng, J.E. Moon, P.K. Ko, and C. Hu, "Performance and Reliability Design Issues for Deep-Submicron MOSFET's," *IEEE Trans. on Electron Devices*, Vol. 38, No. 3, March 1991, pp. 545-554.

[2] J.E. Moon, C. Galewski, M. Wong, W.G. Oldham, P.K. Ko, and C. Hu, "A Deep-Submicrometer Elevated Source/Drain LDD Structure Fabricated Using- Hot-Wall Epitaxy," *Proceedings Technical Papers, 1991 International Symposium on VLSI Technology, Systems, and Applications*, Taipei, Taiwan, May 1991, pp. 117-121.

[3] J. Chen, P.K. Ko, and C. Hu, "Effects of the Field-Edge Transistor on SOI MOSFETS," *Proceedings of Technical Papers, 1991 International Symposium on VLSI Technology, Systems, and Applications*, Taipei, Taiwan, May 1991, pp. 219-223.

[4] S. Parke, J. Moon, P. Nee, J. Huang, C. Hu, and P.K. Ko, "Gate-Induced Drain Leakage in LDD and Fully-Overlapped LDD MOSFETs," *1991 Symposium on VLSI Technology Digest of Technical Papers*, Oiso, Japan, May 1991, pp. 49-50.

[5] P. Fang, K.K. Hang, P.K. Ko, and C. Hu, "Hot-Electron Induced Traps Studied Through the Random Telegraph Noise," *IEEE Electron Device Letts.*, Vol. 12, No. 6, June 1991, pp. 273-275.

[6] J. Chen, R. Solomon, T.Y. Chan, P.K. Ko, and C. Hu, "A CV Technique for Measuring Thin SOI Film Thickness," *IEEE Electron Device Letters*, Vol. 12, No. 8, August 1991, pp. 453-455.

[7] J. Chen, K. Quader, P. Ko, and C. Hu, "Hot Electron Gate Current and Degradation in P-Channel SOI MOSFETs," *IEEE International SOI Conference Proceedings*, Vail, Colorado, October 1991, pp. 8-9.

[8] F. Assadaraghi, J. Chen, P. Ko, and C. Hu, "Time Dependence of Fully Depleted SOI MOSFET's Subthreshold Current," *IEEE International SOI Conference Proceedings*, Vail, Colorado, October 1991, pp. 32-33.

[9] J. Chen, A. Lee, P. Fang, P. Ko, C. Hu, R. Solomon, and T. Chan, "Interface Quality of SOI MOSFET's Reflected in Noise and Mobility, *IEEE International SOI Conference Proceedings*, Vail Colorado, October 1991, pp. 100-101.

[10] F. Assaderaghi, J. Chen, R. Solomon, T.Y. Chan, P.K. Ko, and C. Hu, "Transient Behavior of Subthreshold Characteristics of Fully Depleted SOI MOSFET's," *IEEE Electron Device Letters*, Vol. 12, No. 10, October 1991.

[11] C. Hu, "A Migration Path of Future Silicon MOSFET," Keynote Address, *10th Symposium on Future Electron Devices* (MITI, Japan), Tokyo, Japan, October 1991, pp. 185-191.

[12] Z.H. Liu, J.H. Huang, J. Duster, P.K. Ko, C. Hu, M.C. Jeng, and Y.C. Cheng, "Threshold Voltage Modeling for Deep-Submicrometer Conventional LDD MOSFETs at 300K and 85K," *Proceedings of 1991 International Semiconductor Device Research* Symposium, Charlottesville, Virginia, December 4-6, 1991, pp. 411-414.

[13] J. Chen, F. Assaderaghi, H. Wann, P. Ko, and C. Hu, "An Accurate Model of Thin Film SOI MOSFET Breakdown Voltage," *Technical Digest IEEE International Electron Devices Meeting*, Washington, D.C., December 1991, pp. 671-674.

[14] J.E. Chung, P.K. Ko, and C. Hu, "A Model for Hot-Electron-Induced MOSFET Linear-Current Degradation Based on Mobility Reduction Due to Interface-State Generation, *IEEE Transactions on Electron Devices*, Vol. 38, No. 6, June 1991, pp. 1362-1370.

[15] E. Rosenbaum, Z. Liu, and C. Hu, "The Effects of Oxide Stress Waveform on MOSFET Performance," *IEEE Transactions on Electron Devices*, 1991, pp. 28.1.1-4.

II-B. Conductive Oxides and Ferroelectrics for Programmable Devices

Professor C. Hu with Reza Moazzami and H. Shin

Ferroelectric materials such as lead zirconate titanate (PZT) possess several characteristics which are desirable for high density volatile and nonvolatile memories including a high dielectric constant, high speed read/write capability, large remanent polarization, and long term data retention. In order to develop a viable ferroelectric memory, the properties of these materials need to be characterized to obtain predictable memory operation. Previous work in this area describes reliability characteristics such as fatigue (the gradual loss of detectable ferroelectric polarization following repeated polarization switching) which limits read/write endurance, ageing (the loss of polarizability over long periods of time), data retention, and dielectric breakdown) which is critical for high density DRAM applications. However, little emphasis has been placed on the use of high-speed pulse switching to distinguish the individual polarization components of the films and their behavior during memory operation. We have recently described the basic polarization components which exist during high-speed memory operation.[1] A novel high-speed multiple-pulse testing technique is used to reveal a polarization relaxation mechanism in the microsecond regime. The implications of this relaxation and its behavior during memory operation are discussed.

High-speed testing (2.5 μs pulses) of 400 μm^2, 4000Å Pt/PZT/PT capacitors is performed using a conventional Sawyer-Tower circuit (capacitor divider). A pulse sequence shows the logic 0 and 1 signals as would be observed in a nonvolatile ferroelectric memory. The voltage remaining on the capacitor divider at the end of the applied pulse is the memory signal and is expected to be proportional to the nonvolatile remanent polarization of the ferroelectric film. A logic 0 is not expected to cause remanent polarization switching and therefore should induce no voltage.

We found that neither of these expectations are true: The remanent polarization relaxes within microseconds of the end of the applied pulse for reading a logic 1 and the logic 0 signal is

indeed causing remanent polarization switching. This behavior is not unique to one specific fabrication procedure but exists in a wide variety of capacitor structures and under all test conditions. The nonremanent polarization consists primarily of the nonferroelectric (ionic, electronic, and dipolar) polarizability and is significant for DRAM applications. The remanent polarization which is necessary for nonvolatile memory applications consists of the "volatile" component which relaxes within microseconds and a "nonvolatile" component which remains after the complete relaxation.

The relaxation of the remanent polarization is consistent with the concept of the reversal of domains with very low effective coercive fields after the applied pulse. Once these domains relax, a subsequent pulse in the same polarity causes domain switching and a nonzero logic O signal (i.e., the volatile remanent polarization). The relaxation process follows a logarithmic time law rather than a single exponential decay which supports the concept of a domain population with a broad range of effective coercive fields (i.e., relaxation times).

The relaxation of the remanent polarization does not pose as serious a problem for DRAM applications since any relaxation will only add to the available nonremanent polarization. However, the existence of the volatile remanent polarization not only reduces the initial signal margin for nonvolatile memory applications but may also reduce the endurance and retention specifications for nonvolatile memory applications. Figure 1 shows such an example where the remanent polarization degrades rapidly due to cycling. A conventional measurement technique would give a nonzero remanent polarization even beyond 10^{10} cycles. However, the high-speed polarization measurement clearly shows that the non-volatile remanent has totally disappeared at 10^{10} cycles and the remanent polarization consists only of volatile polarization. In May 1991, we filed a U.S. patent application [2] on a ferroelectric nonvolatile RAM (FNVRAM) which normally operates as a conventional DRAM yet also exploits the hysteresis loop of ferroelectric materials for occasional nonvolatile operation. In this case, the ferroelectric is not cycled

between the two polarization states during read/write operation thus avoiding significant fatigue. Yet, data can be stored nonvolatiled, as in magnetic storage, when desired, e.g., at power-down, or upon a user command. The principle of operation of this was described in last year's report.

We have also completed a study of the reported very high conductivity of oxides thermally grown on heavily silicon substrate. This phenomenon is a serious problem for some microelectronics applications such as the case of capacitors in analog IC's and the field oxides in radiation-hard IC's which have high field doping concentration. On the other hand, this phenomenon has been exploited to fabricate tunneling oxides for EEPROMs without resorting to very thin oxides which are prone to defects. Alternatively, these conductive oxides can be used to lower the program/erase voltage of EEPROMs to make EEPROMs more compatible with advanced CMOS technologies which cannot withstand the 18V or 20V typically required today.

We have fabricated a series of oxides either thermally grown of deposits by CVD on silicon substrates implanted with As, P, B or Si in doses ranging from $5x10^{14}$ cm^{-2} to $5x10^{15}$ cm^{-2}. We have examined the IV characteristics, dependence area and peripheral length, SEM and TEM photographs, annealing effects, defect densities, etc. We have ruled out explanations in terms of surface asperities and lower oxide potential barrier due to chemical impurities in the oxide. We concluded, somewhat surprisingly, that the observed high conductivity may simply be attributed to a band of thinner oxide along the edge of the field oxide. Apparently, the mechanical stress at the field edge significantly reduces the oxidization rate of heavily doped silicon such that a thin band of oxide is formed.[3] This insight immediately suggests a way of minimizing the oxide thinning at the field edge -- growing oxide in dry ambient at high temperature, including rapid thermal oxidation. On the other hand, low temperature wet oxidation is expected to enhance the high conductivity.

Publications/Patents

[1] R. Moazzami, N. Akt, Y. Nissan-Cohen, W.H. Shepherd, M.P. Brassington, and C. Hu, "Impact of Polarization Relaxation on Ferroelectric Memory Performance," *1991 Symposium on VLSI Technology Digest of Technical Papers*, pp. 61-62, Oiso, Japan, May 1991.

[2] C. Hu, and R. Moazzami, "Ferroelectric Dynamic Random Access Memory Device," U.S. Patent Application, May, 1991.

[3] C. Hegarty, J. Lee, and C.Hu, "The Effect of Heavily-Implanted Surface Layers on Fowler-Nordheim Tunneling Currents," accepted by *Solid State Electronics*.

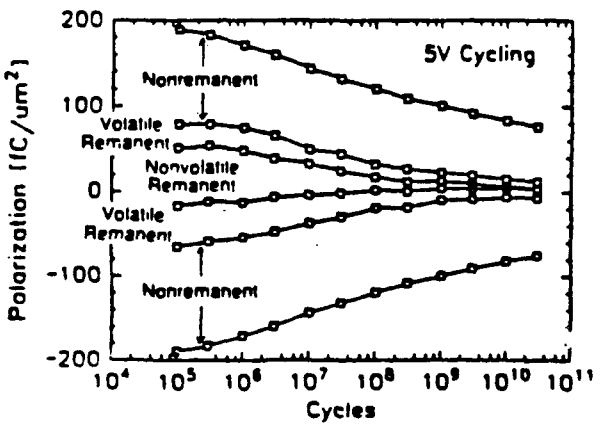


Figure 1. A standard polarization measurement which does not distinguish between the volatile and nonvolatile remanent can produce overly optimistic reliability projections as shown here. In this case, the nonvolatile remanent is completely lost after 10^{10} cycles but the volatile remanent still exists and would be considered the available signal in a conventional measurement procedure.

II-C. Insulated-Gate GaAs Field Effect Transistors

Professors N.Cheung, S.Wang, C.Hu and W.Oldham with James Chan

Introduction

Conventional GaAs MESFET technologies suffered from low voltage swings and high gate leakage current as a direct result of the low barrier height across the metal-semiconductor interface. To circumvent these two limitations, it is natural to investigate new gate technologies which can sustain large gate voltage swings as well as high channel currents. A survey of past attempts to fabricate insulated gates for III-V FETs led us to the concept of a gate structure in the form of insulator/$Al_yGa_{1-y}As$/GaAs. This structure will preserve the high interface quality of the $Al_yGa_{1-y}As$/GaAs interface. The "insulator" will be a high bandgap semiconductor which prevents thermionic or tunneling current through the stacked gate structure. This high bandgap material, if proved viable, will lead to other applications such as deep UV optoelectronics and high temperature electronics, in addition to digital electronics.

The usage of AlN_x as a high bandgap insulator material for GaAs-based Metal-Insulator-Semiconductor Field Effect Transistors (MISFETs) is being investigated. The insulated-gate structure being studied was a capacitor consisting of Al contact on reactive-sputtered AlN_x on top of a linearly graded n-type $Al_xGa_{1-x}As$ layer which varied from x=0.33 at the GaAs substrate surface to the top surface. The motivation for using this structure was to improve upon the AlN/$Al_xGa_{1-x}As$ interface condition through a solid-phase epitaxial reaction. Another set of AlN_x samples sputtered on Si substrates was fabricated also for material properties and for X-ray analyses of AlN_x.

Summary of Experimental Results

AlN_x samples were prepared under various sputtering conditions. Numerous material, optical and electrical characterizations were performed, indicating that AlN_x is indeed an insulator material.

We used a RF reactive sputterer supported by an Al target and a mixture of Ar/N_2 gases to prepare the AlN_x samples. Different sputtering conditions with varying plasma power and Ar/N_2 gas mixtures have been utilized and compared. Rutherford Backscattering composition tests were employed to analyze the composition of the dielectric films. With an Al/N ratio of 1:1 (AlN_x , with x=1.0) as the ideal goal, results show that a 115 sccm:40 sccm Ar/N_2 mixture at a plasma power of 300 watts gave a maximum nitrogen content of x=0.96, with minimal oxygen concentration (See Figure 1). Comparable compositions were achieved for the 80 sccm:80 sccm Ar/N_2 gas mixture at 300W.

The refractive indices of the samples were also extracted. Although values were consistent within each run, they varied from run to run. They ranged from 1.982 to 2.265. The average index for AlN_x samples prepared under the 115 sccm :40 sccm Ar/N_2 gas mixture was 2.064. This compared favorably with data reported in the literature ($n_f = 2.152$). [1]

Current-voltage tests performed on AlN_x capacitors sputtered on the graded-AlGaAs substrate measured leakage current density on the order of $0.1nA/cm^2$. From the I-V measurements, a breakdown field on the order of $10^6V/cm$ was calculated (See Figure 2). Capacitance-voltage measurements of these MBE samples indicated fixed charge density at the AlN_x/AlGaAs interface to be around $10^{12}/cm^2$ for unannealed samples. See Figure 3 for a typical C-V plot of the MIS structure.

In addition to these electrical characterizations, optical transmission measurements employing a monochrometer were carried out to measure the optical bandgap properties of the AlN_x film. From the optical absorption data, a plot of absorption versus energy (in the form of $(Ahv)^{1/m}$ vs. hv, where A is the absorption coefficient) was extracted to give the asymptotic value of the energy gap (E_g). For a direct gap material such as AlN_x, m=0.5, the highest experimental value was about 5.9eV (See Figure 4). [2] This result also agreed with literature. [3]

Conclusion & Future Work

We have confirmed that AlN_x is an insulator material which has a direct bandgap value of 5.9eV. Furthermore, electrical measurements of AlN_x sputtered on an AlGaAs/GaAs substrate system concluded that breakdown field is close to 10^6 V/cm and interface trap density between the AlN_x and AlGaAs layers to be 10^{12} /cm^2. Future work includes materials and device characterizations of annealing effects on the AlN/AlGaAs interface traps and electrical breakdown conditions. An extensive X-ray analysis will be carried out to analyze the material quality of AlN_x. To lower the interface trap density, we will perform the following: (1) reactive sputtering of AlN_x at elevated substrate temperatures; (2) modification of AlN_x via nitrogen ion implantation will be attempted to create a stoichiometric nitride material; (3) annealing of AlN_x samples at different temperatures and with various nitrogen-containing ambients.

A publication on the influence of substrate deep traps on GaAs FET leakage and backgating appeared in print.[4] This study reported a modification to a popular 2D device simulator PISCES -- the addition of deep traps with appropriate dynamic models so that the simulator can adequately simulate GaAs devices with significant density of deep traps in the substrate. Through simulations it was found that traps of the donor type in a certain energy range can significantly reduce the backgating effect.

References

[1] Wang, D.H, and Liang, Guo, "Optical Properties of Sputtered ALN Films and Coated GaAs," *Thin Solid Films*, pp. 39-43, 158(1988).

[2] Ho, Kwok-L., Jensen, Klavs F., et. al., "MOCVD of Wide Bandgap III-V Semiconductors by Using Novel Precursors" *M[illegible]es.Soc.Symp.Proc.*, pp. 605-610, Vol. 162.

Publications

[1] Chan, James, "Low Energy Ion Beam Modification of AlN_x Thin Film for Insulated Gate Field Effect Transistors", to be presented at the MRS Spring Meeting, San Francisco, April 27 - May 1, 1992.

[2] P. George, P.K. Ko, and C. Hu, "The Influence of Substrate Compensation on Inter-Electrode Leakage and Back-Gating in GaAs MESFETs," *Solid State Electronics*, Vol. 34, No. 3, March 1991, pp. 233-252.

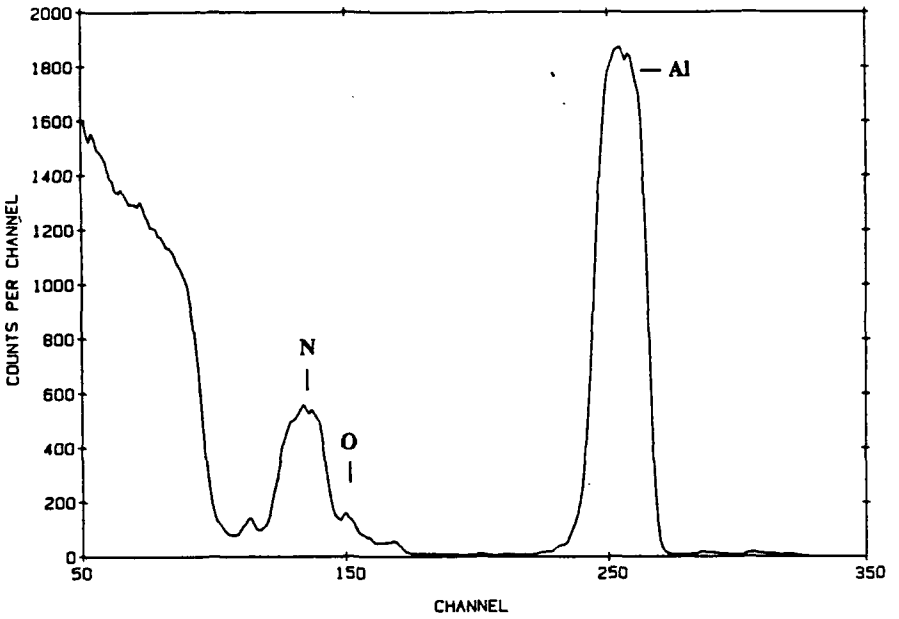


Figure 1: RBS analysis of AlN_x film sputtered on carbon substrate at 300W, Ar/N_2 ratio of 115 sccm:40 sccm.

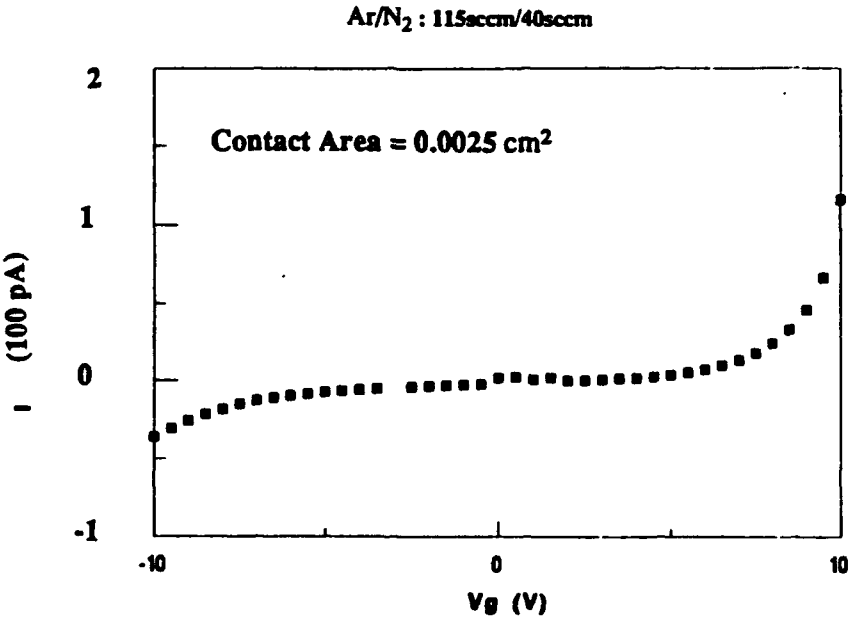


Figure 2: Current Voltage characteristics for AlN_x sputtered on MBE substrate; BV ~ 10V, E_{max} ~ 10^6 V/cm.

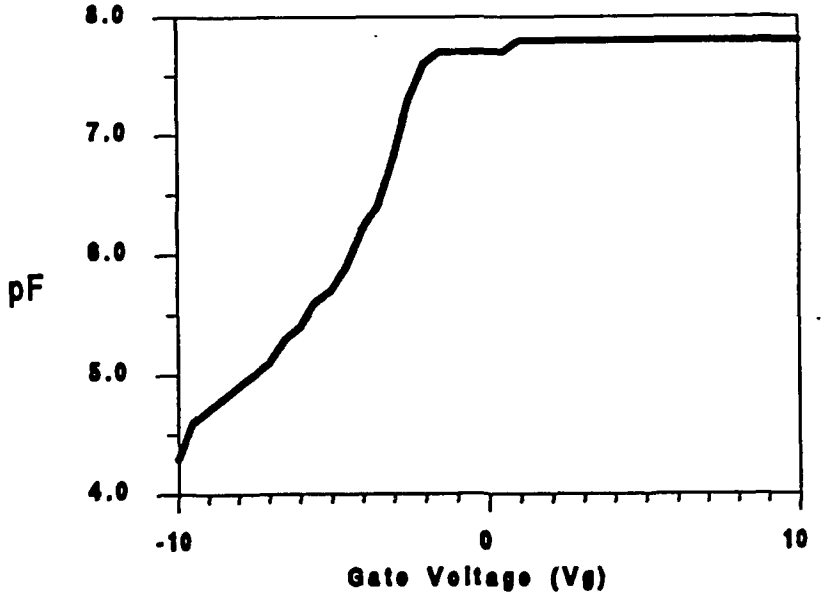


Figure 3: Capacitance-voltage characteristics for 300 angstroms AlN_x sputtered on MBE substrate.

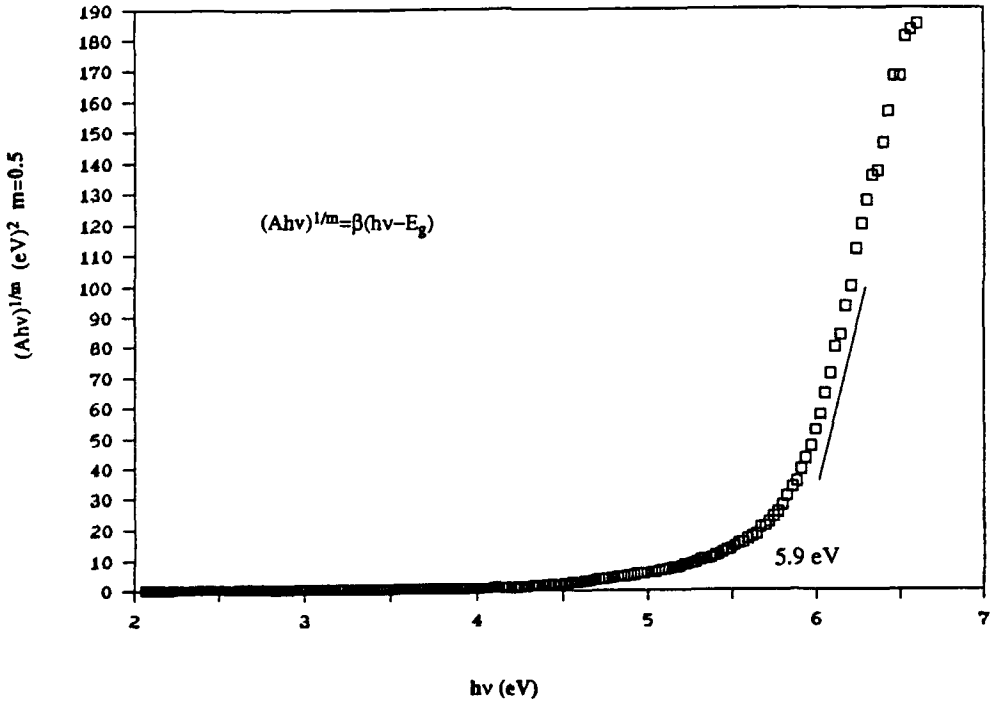


Figure 4: $(Ahv)^{1/m}$ vs. hν plot of AlN_x on sapphire, with extrapolated E_g = 5.9eV.

III-A. Stochastic Neural Networks and Applications to Signal Processing

Professor Avideh Zakhor and Francesco Lari

We address the classification of one dimensional signals using Time-Frequency(TF) techniques. TF techniques have the basic advantage of a physically meaningful interpretation that can not be claimed by classical detection methods such as matched filtering; thus human observers can interpret and classify the return signal from a target in the TF domain and not in the time domain. Moreover, the TF representation has the additional advantage that image processing techniques can be used to reject the noise while preserving the signal of interest.

In this project we compare two TF Transforms: (1) the Wigner-Ville Transform and (2) the Wavelet Transform. The Wigner-Ville Transform has been chosen because of its desirable mathematical properties that permit matching mathematical optimality with physical interpretation. We chose not to use the windowed versions of the Wigner-Ville, to avoid an adaptive choice of the window. This choice is justified because the presence of cross terms does not necessarily degrade a classification algorithm, as long as the cross termed transform remains representative of a given class. This is certainly accomplished if the cross terms remain limited to the ones due to interaction between terms in the positive frequency plane, i.e. if the analytic signal is used. Thus, we take the Hillbert Transform of the input real data to obtain the analytical signal. We then take the modulo of the resulting Wigner-Ville transform. The Wavelet Transform has been chosen because of its useful characteristics: in the wavelet transform the signal is transformed at different resolutions, providing a meaningful physical representation of the signal. Moreover, there is no need to window the signal, since the wavelet function itself provides a variable length window at the different resolutions. Another advantage of the wavelet transform is that it does not have the problem of the cross terms. We use the Gaussian Wavelet Function. In this way we obtain a real Wavelet Transform, of which we take the squared modulo(scalogram). We consider the scalogram for having a description equivalent to the one

given by the Wigner-Ville Transform, in which the integration over an area in the TF plane gives the power contained into that interval of time and frequency.

We then extract features from the two dimensional (2-D) TF Transform. If the dimension of the input vector to the classifier is too large, then a problem commonly referred to as "the curse of dimensionality" occurs: since the classifier has to be trained with a finite amount of data, a large dimensional feature vector makes the training data dispersed in the space, and undermines the generalization capability of the classifier. Thus it is not acceptable to simply feed the whole 2-D TF transform into the classifier input. As for the choice of the features to be used by the classifier, the first problem to be solved is the independence of the classification procedure from the absolute time. We solve this problem by identifying the specular return and taking all the times relative to the specular return time. We choose as features the results of integrations over squares in the TF domain. In this way we provide the classifier with information about the power distribution in the TF plane, with only a limited number of features. The locations and dimensions of the integration areas have been chosen to maximize the classifier performance while keeping the number of features low to avoid the "curse of dimensionality" discussed above.

For the classification of the features we use a decision tree. We choose this kind of classifier because of its performances and speed. Our experimental data consists of computer simulated returns from two solid elastic cylinders with hemispherical endcaps and a length to diameter ratio of 10 and 5 respectively, along with a solid sphere, an alluminum spherical shell and two steel shell of different thickness. This results in 6 different classes; moreover the data from the cylinders are obtained for various angles of incidence, resulting in a total of 7 different aspect ratios for each of the 2 cylinder classes. We add random white Gaussian noise of up to -4dB of SNR to the original data. In this way we generate 12,000 return data, of which 8,000 were used as learning sample for the decision tree, and 4,000 were used as test sample for optimal pruning of the same tree. The training sample size was chosen experimentally: we determined that further

increment in the sample size results in marginal improvements of the classification rate. Another set of data was then generated to verify the performance of the classifier.

The resulting performance of the classification algorithm is shown below; detection denotes the correct classification of the return in one of the 6 classes, while classification denotes the correct classification of the class and the aspect ratio.

SNr	DETECTION		CLASSIFICATION	
	Wign-Vill	Wav	Wign-Vill	Wav
12	100	100	100	100
10	100	100	99	99
8	100	97	99	95
6	100	97	97	94
4	99	95	95	92
2	97	75	94	65
0	95		90	
-2	86		73	
-4	71		47	

Publications

[1] F. Lari and A. Zakhor, "Automatic Classification of Active Sonar Returns using Time-Frequency Techniques," submitted to SPIE's 1992 International Symposium on Optical Applied Science and Engineering, San Diego, California, July 1992.

III-B. Learning and Generalization by Neural Networks

Professor Alberto Sangiovanni-Vincentelli and Arlindo L. Oliveira

There are two main reasons why neural networks are of interest for the scientific community. The first is their ability to generalize from incomplete information i.e., their ability to perform induction. The second is the promise of massive processing parallelism inherent to the architecture of neural nets.

So far, two major obstacles have hindered the use of neural nets in a wider variety of real life applications. The first is the inexistence of algorithms that can be used to train large networks for complex tasks. Current training algorithms usually scale badly and fail to achieve the desired objectives. The second obstacle is the fact that hardware implementations of threshold units with real-valued weights is expensive and commercial applications have been limited to relatively small numbers of neurons per chip.

The work in progress addresses these two problems by adopting a different view to the problem of generating networks that learn. In fact, the ability to generalize is not inherent to networks of threshold units with real-valued weights. Ability to generalize is closely related to data compaction. Simple descriptions of observed data have high predictive value while long and complex ones have low predictive value. We are currently developing a special purpose logic synthesizer that is able to generate minimal size networks that give the right output for the data available in the training set. Since very compact descriptions are usually generated, the resulting networks exhibit generalization abilities superior to the ones obtained with traditional learning algorithms like back-propagation. Furthermore, the types of the nodes used can be forced bo be easily implementable in hardware. Two major alternatives exist: either force the nodes to be logic gates (and/or/not) or threshold gates with small integer weights. In the first case implementation is straightforward, and networks with several tens of thousands nodes can be easily realized. For problems that map more easily into networks consisting of threshold gate nodes, the fact that the

weights are small integers also makes it much more easier to implement than general threshold gate networks.

In previous work, we developed techniques for the synthesis of two-level networks. The results have shown both that the approach is appropriate for problems that accept a compact two-level representation and that a more general approach is required for problems that require multi-level networks.

Ongoing work is directed at obtaining a logic synthesizer that efficiently generates compact multi-level networks using some of the properties of the problem. Multi-level network synthesis requires two steps: factorization and simplification. Special purpose algorithms for these two steps were developed and are currently being evaluated in several problems. Both algorithms use the fact that the input data consists only of minterms to avoid explicit representations of the don't-care set. Traditional algorithms, like the ones used in the MIS/SIS system need to explicitly derive a don't-care network and are inappropriate for the problems of interest, since the size of any explicit representation of the don't-care set is, in general, too large.

Publications

[1] Arlindo L. Oliveira and Alberto Sangiovanni-Vincentelli, "Learning Concepts by Synthesizing Minimal Threshold Gate Networks," *Proceedings of the Eighth International Workshop in Machine Learning*, Chicago, Illinois, June 1991.

[2] Arlindo L. Oliveira and Alberto Sangiovanni-Vincentelli, "LSAT: An Algorithm for the Synthesis of Threshold Gate Networks," *1991 International Conference in Computer Aided Design*, Santa Clara, California, November 1991.

[3] Arlindo L. Oliveira and Alberto Sangiovanni-Vincentelli, "Synthesis of Minimal Multi-Level Networks" to be presented at Neural Networks for Computing Workshop, Snowbird, Utah, April 1992.

III-C. Reconfigurable Analog Elements for Neural Nets

Professors Ping K. Ko and Chenming Hu

Investigation of the use of EEPROM devices as reconfigurable analog weights continued. The CMOS-EEPROM fabrication technology has been found to be fully functional and has been completely characterized based upon last December's set of 8 wafers. The EEPROM devices programmed and retained charge successfully.

Additionally, a novel CCD-EEPROM structure that will be a more efficient analog weight element is being prepared for the next set of wafers. This structure uses minority-carrier programming to control the amount of available programming charge. By taking advantage of the inherent physics of the MOS system to control the amount of charge available for tunneling, we hope to avoid the need for additional feedback circuitry that is required when using just an EEPROM. The first run of these devices was not entirely successful but experimentation showed that the major unknown of this composite device, minority-carrier tunneling, will work.

Further, since industry is progressing toward the use of Flash EEPROM (as opposed to the standard EEPROM), we are investigating the use of such a device as an analog weight element. Other research at this university has developed a Flash EEPROM process and will be used for experimentation toward such an end.

Publications

[1] Chi-Kai Sin, "EEPROM as an Analog Storage Element," Master's Report, University of California at Berkeley, September 1990. (Not reported previously.)

[2] C. K. Sin, A. Kramer, V. Hu, R. Chu and P. K. Ko, "EEPROM as an Analog Storage Device with Particular Applications in Neural Networks," *1992 IEEE Transactions on Electron Devices*, accepted for publication in June 1992 issue.

[3] A. Kramer, P. K. Ko and A. Sangiovanni-Vincentelli, "Massively Parallel Analog Geometric Computation Using EEPROMs," abstract for Neural Networks for Computing Conference, Snowbird, Utah, April 2-5, 1991.

III-D. Architectural Issues in Parallel Computation

Heterogeneous Architectures for Artificial Neural Networks
Chedsada Chinrungrueng (Professor C. H. Séquin)

Existing *monolithic* artificial neural network architectures, are not sufficient to cope with complex problems, such as processing of speech or vision. To solve a complex task, a large scale network is needed. Building a large scale monolithic network in parallel VLSI hardware will typically be too costly because the structure of the network would have to be too large and because its connectivity requirements would be excessive. Furthermore, training a large-scale network as a monolithic system is impractical because the dimension of the adjustable parameter space is very large and this would result in unacceptably slow convergence rates.

The existence of *heterogeneous* organization in mammalian visual systems suggests that large scale networks should be composed of a variety of *heterogeneous* modules. One possible heterogeneous architecture might be composed of two level: the lower level, consisting of self-organizing modules that partitions the input domain into manageable subdomains; and the higher level, consisting of trainable modules that works in a supervised learning mode.

We have developed an heterogeneous architecture for classification. In this architecture, the input domain is automatically partitioned into non-overlapping subdomains. Associated with each subdomain, are separate classification networks that are adjusted by supervised training; since they have to handle smaller problems, they can be simpler and can learn more quickly. The partitioning algorithm is characterized by two mechanisms. The first one balances the classification error equally among all the subdomains so that all the systems resources are optimally utilized. The second mechanism is to dynamically adjust the learning rate based on the degree of resource utilization, i.e., the uniformity of the distribution of the classification error among all regions. These two mechanism enable the partitioning network to automatically and dynamically adjust its characteristics to match that of an assigned problem without any interference from the user. It is suitable for both stationary and non-stationary situations. The

advantages of this automatic partitioning scheme over more classical approaches to input space partitioning has been demonstrated on various examples.

Work is under way to evaluate the performance of the described heterogeneous architecture on applications to equalization of digital channel and automatic control of second order system and to compare it against MLP and RBF.

Publications

[1] C. Chinrungrueng and C. H. Séquin "Optimal Adaptive K-Means Algorithm with Dynamic Adjustment of Learning Rate", International Computer Science Institute (ICSI), Berkeley, TR-91-017, 1991, March. Also presented at IEEE IJCNN-91, Seattle, July 1991.

[2] C. Chinrungrueng and C. H. Séquin, "K-Means Competitive Learning, for Non-Stationary Environments", *Proceedings of the International Joint Conference on Neural Networks in Singapore (IJCNN-91-Singapore)*, Vol. III, pp. 2703-2708, November, 1991.

Fault Tolerance in Layered Artificial Neural Networks.
Reed Clay (Professor C. H. Séquin)

We have been continuing our investigation into the fault tolerance of feed-forward artificial neural networks. In particular, during 1991, we have been investigating the additional benefits of our training technique for fault tolerance as it relates to generalization and to the problem of overfitting to training data. We have also refined our fault tolerance training technique and tested it on more difficult, "real-world" applications.

The basic idea behind our fault tolerance training technique is to randomly introduce - during training - the types of failures that one might expect to occur during the actual operation of the network. We have shown that this can make the network tolerant not only against the specific faults trained for, but also against faults that have never been presented explicitly and which may also be considerably more severe (i.e., double or triple faults) than any faults ever presented to the network during training.

Now, we are exploring what additional side effects may arise from this particular training for fault tolerance. Specifically, we have examined the power of generalization of networks trained for fault tolerance. For applications that show a tendency to overfit the network to a noisy set of training data, and thus to generalize to new test data more poorly than is warranted, our fault tolerance training method was able to reduce the amount of overfitting noticeably. The most convincing result was obtained on an artificial demonstration example which involved "learning" a classification boundary between two fuzzy, partially overlapping Gaussian distributions. Applying a standard backpropagation algorithm to a sparse sample of data points from this distribution typically results in a ragged decision boundary that reduces the classification error for the training data to zero, but which produces a larger overall error than a single straight decision boundary. Our fault tolerance training technique, on the other hand, because of the constantly varying combination of active hidden units, prevents an overly tight approximation of the training data, and produces a smoother, more general boundary, which lies more closely to the optimal separation line between the two Gaussians. Thus the effect of our training method seems to be comparable to other techniques that inject some noise via the data input or directly into the hidden units in order to "smooth out" the internal representation of the learned task. In either case, the effect is not very pronounced, and on actual data from other applications, such as for example sunspot prediction, the effect of overfitting was much less clearly visible, and thus any improvements due to our alternative training method were barely detectable.

We have also studied our fault tolerance training technique on the more difficult, "real-world" application of character recognition. Several sets of handwritten characters are digitized onto a 10x10 grid of analog values. These patterns are classified via a three level ANN into 26 classes representing the capital letters of the alphabet. We have been able to show that, given additional hidden units (beyond the 10 units required to solve this task without faults), our algorithm readily handles this more difficult task. Specifically, we have shown that the network can

become robust to all single, double, and triple faults if it is trained with samples of these types of faults - even for the cases where the hidden units are stuck at -1 or at +1. This is typically more difficult to handle than a simple loss of the signal from the hidden units, which is equivalent to a stuck-at-0 fault. Also, we have seen how prolonged training can make the network tolerant to faults that are worse than any that it has been trained on.

Another one of our fault tolerance techniques involves using a "supervisor" network that compares the outputs of several subnetworks which each individually approximate an analog function. Previously we used an ad hoc criterion for determining if one of the subnetworks differed significantly from the other subnetworks (thus eliminating it from the averaging of the subnets to determine the best estimate of the output). Now we are using a more appropriate criterion based on the "Chi-squared" test, from Statistics, which determines how likely it is that the variance of the outputs of the subnetworks is due to random noise rather than a more significant underlying source of disturbance - as in our case, a fault in one of the subnets. We have used this technique on the task of predicting the number of sunspots based on knowledge of their number in previous years. This test has shown that this "supervisor" technique generalizes nicely to such a more difficult analog function approximation task.

Network Design Issues for Fast Global Communication
Professor A. Ranade with M.T. Raghunath

A central problem in building large scale parallel machines is the design of the interconnection network. The choice of the network significantly affects the total cost, as well as the ease of programming such machines. Of the models proposed for programming large parallel computers, the most attractive are the ones based on a shared address space. In this model each processor is capable of accessing any part of the shared address space directly. The efficiency with which such models can be supported depends critically upon network performance.

The goal of our research is to develop high performance, cost effective networks for interconnecting processors. While network design depends upon many factors, it is largely constrained by packaging technology. Large networks need to be partitioned and packaged in a hierarchical manner. Wires at the higher levels of the hierarchy are likely to account for a greater fraction of the total cost than the wires at lower levels. We expect that for large-scale systems, the primary packaging constraint at the top level of the hierarchy will be the number of pins per module. We have evaluated different network organizations for a 1000 processor machine based on this constraint. Our results indicate the following general design principles: 1) Making the networks denser at the lower levels of the packaging hierarchy has a significant positive impact on global communication performance, 2) It is better to organize a fixed amount of communication bandwidth as a smaller number of high bandwidth channels 3) For shared memory based communication primitives it is better to make the number of memory modules smaller than the number of processors and 4) Providing the processors with the ability to tolerate latencies (by using *multithreading* is very useful in improving performance.

We are currently working on more detailed models of packaging constraints which take into account other cost metrics such as the number of cables, number of connectors etc. We are also looking at ways to explicitly factor into our cost model the difference in costs at multiple levels of the hierarchy.