Study of Sub-0.5 μm SOI-with-Active-Substrate (SOIAS) Technology for Ultra-Low Power Applications

by

Isabel Y. Yang

B.S., Massachusetts Institute of Technology (1990)
M.S., Massachusetts Institute of Technology (1993)

Submitted to the Department of Electrical Engineering and Computer Science

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1996

© Massachusetts Institute of Technology 1996. All rights reserved.

Eng.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

MAR 06 1997

LIBRARIES

Author ....
Department of Electrical Engineering and Computer Science
September, 1996

Certified by ....
Dimitri A. Antoniadis
Professor of Electrical Engineering
Thesis Supervisor

Certified by ....
Henry I. Smith
Professor of Eletrical Engineering
Thesis Supervisor

Accepted by ....
Frederic R. Morgenthaler
Chairman, Departmental Committee on Graduate Students

Study of Sub-0.5 μm SOI-with-Active-Substrate (SOIAS) Technology for Ultra-Low Power Applications

by

Isabel Y. Yang

Submitted to the Department of Electrical Engineering and Computer Science
on September, 1996, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electrical Engineering

Abstract

In this thesis, a novel fully depleted (FD) Silicon-On-Insulator (SOI) based technology was developed for ultra-low power applications. This technology is called Silicon-On-Insulator-with-Active-Substrate (SOIAS). The SOIAS technology, in general, enables the integration in the third dimension of gates and interconnects by using buried under-layers. In this work, the buried under-layer served as a second gate, a back-gate, to the conventional single-gate FD SOI device. The function of the back-gate is to modulate the threshold voltage of the normal device. Since supply voltage scaling to 1V and below has been identified as a key approach for energy efficient high performance computing, the threshold voltage needs to be scaled concomitantly in order to maintain high performance levels. However, lowering of the threshold voltage results in high static leakage energy. In order to meet the opposing requirements of high performance and low power, a dynamic threshold voltage control scheme was proposed utilizing the SOIAS technology. This work covers the development of the SOIAS substrate fabrication, material characterization, CMOS device and circuit fabrication which also includes the integration of an x-ray lithography module for scaling into the deep submicrometer regime, device and circuit characterization, and energy evaluation of the SOIAS technology for low power applications.

The technology to fabricate the SOIAS wafers was developed using two different approaches; both approaches take advantages of the already existing technologies developed for SOI and bulk. The first method uses the bond-etch-back (BESOI) process, and the second method is a novel bonded SIMOX process. The process complexity is reduced in the bonded SIMOX approach with very good silicon film thickness uniformity control as compared to the BESOI method. The material quality of the former, however, is slightly inferior to that of the latter. Compared to bulk and conventional SIMOX in terms of electrical properties (e.g. mobility, intrinsic oxide breakdown, and interface state density), the BESOI SOIAS substrates are equal or even superior, and the bonded SIMOX SOIAS wafers are slightly inferior.

The implementation of a CMOS process on SOIAS substrates was demonstrated, and the integration of the x-ray lithography module was successful. A mix and match

scheme was developed and implemented to match the various lithography tools used in this process, and both positive and negative chemically amplified resist processes were optimized for 0.1 μm features sizes. Device and simple circuit fabrication was successful. Approximately 3 decades switch in off current and 1.5X increase in drive current was achieved for both NMOS and PMOS at V_{DS} of 1 V. Independent control of NMOS and PMOS back-gates were verified with 36% change in speed with 250 mV switch in the threshold voltage at 1 V supply voltage. Large dynamic threshold voltage control range was demonstrated, and the threshold voltage design band of 250 mV fits well within this range. Threshold voltage control is not deteriorated for effective channel length scaling down to approximately 0.2μm. Quasi-static control of threshold voltage at 20 MHz was verified.

Having demonstrated the technology, a theoretical energy consumption model was developed. This model was used to compare the SOIAS technology against a benchmark convention SOI technology. System and architecture parameters as well as technology design parameters were examined. The dynamic threshold voltage switching scheme is proposed to be implemented at the functional module level (e.g. adder, shifter and multiplier). The SOIAS technology was found to be most suitable for systems which operate in burst-mode such as the microprocessor of an x-server. Significant energy savings is expected for the SOIAS over the SOI technology for event-driven systems. The energy consumption model was also used in conjunction with measured static energy, dynamic energy, and delay data for optimization of threshold and power supply voltages at various system activities. From this analysis, the conclusion was that adaptively adjusting the threshold and supply voltages as the system activity changes can potentially allow energy efficient computation at all times.

Thesis Supervisor: Dimitri A. Antoniadis
Title: Professor of Electrical Engineering

Thesis Supervisor: Henry I. Smith
Title: Professor of Eletrical Engineering

Acknowledgments

Initially, the SOIAS project was not funded. With the help of various supporters, I was able to carry through most if not all of the research work. I am very grateful to Lisa Allen of IBIS Corp. for providing us with SIMOX wafers, Charles Goodwin of AT&T for helping with the wafer thinning process, D.P. Mathur of Hughs for providing the localized plasma thinning, and Craig Keast of MIT Lincoln Laboratory for helping with the CMP. I would also like to acknowledge AT&T for providing me with a graduate fellowship, without it I would not have had the freedom to pursue this work.

I was very fortunate to come under the guiding influence of my advisors. From them, I have learned how to be a good investigator, to think scientifically, and that no detail is too small to ignore. They gave me the freedom to pursue my work and establish my individuality, and at the same time offered advice and guidance whenever I needed help. I would like to thank Professor Dimitri Antoniadis for his support, and for imparting some of his infinite wisdom and knowledge to me. I would like to thank Professor Henry Smith for his technical and spiritual support, for his boundless enthusiasm, and for sharing his great love of science. I would also like to thank Professor Anantha Chandrakasan for his tutelage and insights on circuit and system design issues.

I would like to thank Lalitha Parameswaran for her valuable advice on silicon wafer bonding, Paul Tierney, Tim Tyson, Brian Foley, Dan Adams, Dick Westberg, Jim Bishop, Joe DiMaria, Joe Walsh, Bernard Alamariu, Pat Burkhart, Octavio Hurtado, Linus Cordes, and Vicki Diadiuk for their technical support and patience. My special thanks to Jimmy Carter and Mark Mondol for giving me the freedom to do crazy experiments in the NSL.

Through the years I've spent at MIT, I have met and befriended some of the brightest people who came to this fine institution. My thanks goes to: Dr. Lisa Su for her friendship and for those endless philosophical discussions that made life so interesting, Dr. Vincent Wong for his friendship and for being my NSL lab buddy, Dr.

Julie Tsai for being my apartment mate and for remaining my friend afterwards, Dr. Jarvis Jacobs for all the computer support he provided, Drs. Martin Burkhardt and Hang Hu for their help in the NSL, Melanie Sherony for her friendship and for lending a sympathetic ear whenever I needed to let out some steam, Mike Lim and Tim Savas for keeping me amused and up to date on the latest gossip, Euclid Moon and Juan Ferrera for their help in the NSL, and for just being a bunch of nice guys, Andy Wei and Mark Armstrong for their help on various occasions with Naxos, Siva Naren and Anthony Lochtefeld for their participation and interest in the SOIAS project.

I am also grateful to my friends who shared the first half of my journey through MIT and who continued to support me even after they have moved on: Winnie Chan, May Yip, Alice Yee, Frances Lam, Kay Hsu, and Rebecca Kao.

To my parents, I owe everything. A mere "thank you" to acknowledge their sacrifices can hardly do justice to the immense gratitude I feel. I hope that my accomplishments have made them proud. Last, but not least, a big thanks goes to my husband, Bill Chu. I would not have made it this far without his constant support, encouragement and infinite patience. His belief in me never wavered even during the most trying times when I had serious doubts about myself. I hope I will have the opportunity to be his "pillar of strength" as he has been to me. I look forward to the life we will share together.

Contents

1 Introduction **17**

1.1 Thesis Organization 18

2 Background **20**

2.1 The Need for Low Power 20

2.2 Dynamic Threshold Voltage Control Concept 22

2.3 SOI Technology 23

2.3.1 SOI Dual-gated Devices 24

2.4 Comparison of V_T Controllability in Bulk and SOI MOSFETs 27

2.4.1 Long Channel Device 27

2.4.2 Short Channel Device 29

2.5 SOIAS Device Structure 33

3 Silicon-On-Insulator-with-Active-Substrate: SOIAS **36**

3.1 Silicon Direct Bonding 36

3.1.1 The Bonding Mechanism 37

3.1.2 Voids 40

3.1.3 Bond Strength 41

3.2 SOIAS Technology 43

3.2.1 SOIAS Wafer Fabrication 44

3.2.2 Bonded Interface Evaluation 47

3.2.3 SOIAS Materials Characterization 51

4 X-ray Lithography Integration in CMOS Process **56**

4.1 Mix-and-Match Scheme 57

4.1.1 Experimental Procedure 58

4.1.2 Coarse Features and E-beam Field Marks Transfer 58

4.1.3 E-beam Patterning of Fine Features 61

4.2 X-ray Lithography Exposure and Alignment Procedure 63

4.3 Gate Pattern Transfer Process 70

4.3.1 Negative Resist Process of SAL-601 70

4.3.2 Positive Resist Process of ESCAP-X 73

4.4 Device Results 75

4.4.1 Bulk Devices 75

4.4.2 SOIAS Devices 79

5 SOIAS CMOS Devices **81**

5.1 Fabrication 81

5.2 Device Measurement Results 86

5.3 Dynamic Operation 87

5.4 SOIAS Circuits 88

6 SOIAS for Ultra-Low Power Applications **97**

6.1 Energy Consumption Model for SOIAS 97

6.1.1 Activity Factor Definitions 100

6.1.2 Simulation Criteria 102

6.2 Energy Trade-offs in Technology Design Space 103

6.3 Energy Trade-offs in A_{bg} and A_{fg} Space 105

7 Ultra-Low Voltage Design Space **113**

7.1 Measurement of Energies 114

7.2 Variable V_{DD} and V_T: 117

7.3 Dynamic Threshold Voltage Control 120

8 Summary **122**

List of Figures

2-1 Electron potential of dual-gated device in the subthreshold region and strong inversion [11]. 26

2-2 Measured threshold voltage control by either body bias in bulk or back-gate bias in SOIAS. The open triangles represent a bulk NMOS device with super-steep retrograde (SSR) channel doping; L_{eff} is approximately 0.75 μm. The filled triangles represent a single step channel doping device, L_{eff} is approximately 0.7 μm. Although the latter has a lower quiescent V_T, the body factor is much smaller than the SSR design. The SOIAS device dynamic V_T range is larger than both of the bulk devices. L_{eff} of the SOIAS device is approximately 0.4 μm. . 30

2-3 Charge sharing in bulk and FD SOI devices [23]. 31

2-4 SOIAS back-gated CMOS device schematic. 33

2-5 Simulated subthreshold I-V at different back-gate biases. 34

3-1 Hydrogen bond formation between two hydrophilic surfaces [26]. . . . 37

3-2 The bonding mechanism proposed by Stengl *et al* showing the transition from weak hydrogen bonding to covalent Si-O-Si bonding at two different temperatures [39]. 38

3-3 (a)Fracture strength for SiO_2 to SiO_2 bonding [47]. (b)Fracture strength for Si to Si bonding [27]. 42

3-4 Surface energy of SiO_2 to SiO_2 bonded interface (open symbols) [36], and that of Si to Si bonded interface (filled symbols) [50]. 43

3-5 The bond and etch back (BESOI) process for fabricating SOIAS substrates. 45
3-6 bonding chamber 46
3-7 The bonded SIMOX process for fabricating SOIAS substrates. 47
3-8 Infrared image of bonded wafers. 48
3-9 Acoustic microscope image of bonded wafers. 49
3-10 (a)SOIAS wafer prepared by the bonded SIMOX process. (b)SOIAS wafer prepared by the BESOI method. The absence of fringes indicates the better film thickness uniformity of the bonded SIMOX SOIAS as compared to the BESOI SOIAS. 50
3-11 Comparison of effective electron mobility for various substrate materials. 52
3-12 Comparison of cumulative percentage failure of 10x10 μm transistors for different substrates. 53
3-13 Charge pumping current as a function of back-gate bias on a SOIAS device. 54
3-14 Recombination charge per cycle as a function of frequency for different substrate materials. 55

4-1 Fizeau interferogram showing an ultra-flat SiN_X membrane, mesa-style x-ray mask. 59
4-2 Process of combining e-beam and optical lithography to produce an x-ray mask. (a)Patterning of coarse features and e-beam field alignment marks using a stepper.(b)E-beam writing of fine features by aligning to the optically defined field marks. (c)The final pattern on the x-ray mask as a result of optical and e-beam lithographies. 60

4-3 Quartz wafer patterning and DUV coarse feature transfer onto x-ray mask. (a)A 10X emulsion mask with coarse features and e-beam field marks was stepped down onto a chrome coated quartz wafer. The pattern was then wet etched into chrome. This quartz wafer then serves as the 1X mask for transferring the pattern onto the x-ray mask. (b)The DUV transfer process. The x-ray mask is placed on a fixture with holes drilled in the underside of the fixture. A jet of nitrogen is "blown" through the holes in the fixture forcing the x-ray mask membrane to conform to the quartz wafer. 62

4-4 E-beam writing process for fine gate features. 64

4-5 The completed x-ray mask with coarse features patterned by G-line stepper and fine featured patterned by e-beam writing. 65

4-6 An enlarged view of the completed four die mask (20x20 mm), and e-beam field view of a NAND gate in a NAND ring oscillator. The light scatter (white dots) outside the membrane area are of no concern. 66

4-7 (a)Schematic of alignment and exposure system. The alignment was done outside of the helium enclosure using the CCD camera and the microscope. Once alignment was completed, the stage was rolled on ball bearings into the helium enclosure. During exposure, the Cu_L x-rays pass through a silicon nitride vacuum window and irradiate the mask and wafer. (b)Aligned polysilicon gate on a narrow-width device active area. 67

4-8 One e-beam field view of a NAND gate in a NAND ring oscillator after the amorphous silicon gate etch. 69

4-9 Process for x-ray exposure using the chemically amplified resist, SAL601, and PVA topcoat to prevent the formation of a "skin" layer. The definition of the polysilicon gate was accomplished in two steps:(1)the LTO hard mask was etched in CHF_3 RIE, (2)then the polysilicon was etched in pure Cl_2 plasma. 71

4-10 (a)Sample with crosslinked "skin" layer with the gate pattern. The gate pattern not visible in the SEM, is visible under the optical microscope. (b)Sample without the "skin" layer. 72

4-11 Dry etching results:(a)SAL601 resist on top of LTO after CHF_3 RIE for a 80 nm gate, (b)120 nm gate pattern transferred into the LTO with the SAL601 stripped, (c)polysilicon gate after Cl_2 plasma etching using LTO as hardmask. 74

4-12 The gate patterning processing using ESCAP-X resist with nitride hard mask. 76

4-13 Schematic cross section of 0.1μm channel length nMOSFET. The gate oxide thickness was approximately 4.5 nm and the polysilicon thickness was 300 nm. The shading in the channel region depicts the retrograde doping. The source/drain extension and deep junction depths are ~40-50 nm and ~ 120 nm, respectively. The 180 nm thick nitride spacers prevented the encroachment of the deep source/drain junctions into the channel region, and ~50 nm of cobalt silicide was formed on the gate and source/drain. 77

4-14 Device output characteristics:(a)I-V characteristics for L_{eff}=85 nm gate length device with threshold voltage of 0.36 V, and gate voltage stepped from 0 to 2 V in 0.2 V increments. (b)Subthreshold characteristics of the same device. 78

4-15 (a) Subthreshold characteristics of a 0.2μm top-gate device showing three decade change in the off current with 4 V back-gate bias in idling mode with low quiescent V_T. With high quiescent V_T, 1.5x times drive current change at VDS of 1.0 V can be observed. (b) I-V characteristics of the same device. 80

5-1 Device schematic. 82

5-2 Two step masking and implant process for the formation of n and p back-gates. 83

5-3 Suprem3 simulated implant profiles for NMOS and PMOS devices and their back-gates. The final active dopant concentration is after all the thermals cycles in SOIAS CMOS process. 84
5-4 Schematic of back-gate contact cuts. 85
5-5 SEM micrograph of the SOIAS device. 85
5-6 Measured NMOS device characteristics showing the effects of switching the V_T on the off current and the drive current at V_{DS}=1 V. 89
5-7 Measured PMOS device characteristics. 90
5-8 Measured threshold voltage tuning range. The x-axis is the quiescent V_T determined by the silicon film thickness and the doping level. The y-axis is the tunable V_T set by the back-gate bias; the design range of 200 mV switch in V_T fits well within the tuning range. 91
5-9 Measured threshold voltage control as a function of gate length. . . . 92
5-10 Threshold voltage roll-off as a function of gate length at different back-gate biases for NMOS and PMOS. 92
5-11 Schematic of measurement setup for the 101 stage ring oscillators and 50 stage inverter chains. 93
5-12 Measured ring oscillator frequency as a function of independently controlled V_T through back-gate biasing. 93
5-13 Measured 101 stage ring oscillator output frequency as varied by changing V_T. A 36% change in the speed is observed for 200 mV change in V_T at V_{DD} of 1 V. 94
5-14 Schematic showing the measurement setup and device under test for dynamic switching of the back-gate experiment. The V_{out} at the drain for three different input frequencies are shown with one example of the input pulse on V_{gb} at 5 MHz. 95
5-15 Overlay of dynamic current due to switching of back-gate on DC characteristics at various static back-gate biases. The points marked X represent dynamic measurements at 10 MHz. 96

6-1 Capacitances of SOIAS and SOI devices. 99
6-2 Schematic illustration of the definitions for the various activities factors. 101
6-3 Comparison of the front-gate dynamic switching energy for SOIAS and SOI. The dark line indicates the boundary where the energies are the same. When the SOIAS front-gate switching energy becomes greater than that of the SOI, the overlap capacitance contribution becomes significant. 104
6-4 The front-gate and back-gate dynamic switching energies for the adder. 106
6-5 The front-gate and back-gate dynamic switching energies for the multiplier and shifter. 108
6-6 Comparison of dynamic switching energies (back-gate + front-gate) and the static leakage energy components for the SOIAS and SOI technologies. The static leakage for the SOIAS is not shown on the plot because it is off the scale. 109
6-7 Comparison of dynamic switching energies (back-gate + front-gate) and the static leakage energy components for the multiplier and shifter. 110
6-8 The adder total energy ratio for SOIAS and SOI. The dark line outlines the break-even plane. 111
6-9 The multiplier and shifter total energy ratio for SOIAS and SOI. Both fall below the break-even plane. 111
6-10 The total energy ratio for SOIAS and SOI with χ=100% (open symbols) and χ=2% (filled symbols). Squares represent the multiplier; circles represent the shifter and triangles represent the adder. Shown are two plots for different technology design parameters: (a)t_{box}=50 nm, s/d overlap=2L_g=0.5 μm. (b)t_{box}= 50 nm, s/d overlap=0.5L_g=0.125 μm. 112

7-1 Inverter transfer characteristics. 114
7-2 The measured effective switching capacitance. 115

7-3 The measured total energy for 101-stage ring oscillator running at three different frequencies. 115

7-4 The measured (a) static (b) dynamic energies, and (c) delay per stage as a function of V_T and V_{DD}. As guides, the arrows indicate the direction of V_{DD} and V_T scaling. 116

7-5 Energy and performance contours in V_T and V_{DD} space for an adder with 0.25 system activity, and the module activity as obtained by profiling a SPEC benchmark. 118

7-6 The optimal V_T and V_{DD} for the adder, shifter and multiplier modules with different system activity factors. The module activity factor, A_{fg}, for each functional unit was obtained from program profiling of the SPEC benchmark Espresso. 119

7-7 Comparison of normalized energy versus V_{DD} for different system activities. 119

List of Tables

3.1 Surface roughness comparison for different substrate materials. 51

3.2 Average interface state density for different substrate materials. . . . 55

6.1 Activity values from profiling the data encryption program (IDEA). . 102

7.1 Module activity factor, A_{fg}, and back-gate activity factor, A_{bg} obtained from profiling the SPEC benchmark program. 121

7.2 The optimal V_T and V_{DD} at minimum total energy for some modules as obtained from the energy consumption model and measured static and dynamic energies for burst-mode, χ=0.01, 100 MHz operation. . 121

A.1 Recipe 110 141

A.2 Recipe 114 141

A.3 Recipe 226 142

A.4 Recipe 230 142

A.5 Recipe 10 142

A.6 Recipe 15 142

A.7 Recipe 20 143

A.8 Recipe 32 143

A.9 Modified recipe 32 for gate etch. 143

Chapter 1

Introduction

The scope of this thesis spans a wide range. The main focus is placed on technology development with emphasis on application to low power electronics. With the tremendous infrastructure developed for silicon technology, it is unlikely that replacement by a completely different material would occur in the next 10 years. The main technology elements in this thesis, in my opinion, extends the longevity of silicon-based technology by leveraging off the established infrastructure for silicon bulk CMOS. The two main technology elements of this work are:

- Silicon-On-Insulator-with-Active-Substrate (SOIAS) is a variant of the fully depleted (FD) Silicon-On-Insulator (SOI) technology. In general, this technology can be extended to fabricate 3-dimensional structures as opposed to the conventional 2-dimensional planar structures that are prevalent in today's integrated circuits. Specifically, in this work, an additional back-gate is added to the conventional FD SOI MOSFET to electronically modulating the threshold voltage.

- X-ray lithography technology allows scaling of the MOSFET gate dimension into the sub-0.1 μm regime. This technology is applicable to both bulk and SOIAS CMOS. Scaling the MOSFET gate is another means of extending the life of silicon technology by squeezing more performance out of the device. In this work, the integration of x-ray lithography into a CMOS process is developed.

From a pragmatic standpoint, a new technology would be of no value unless it

can be implemented in applications that can greatly benefit from such a technology. In this work, the SOIAS technology with electronically variable threshold voltage is evaluated for ultra-low voltage, low power, and high performance systems. This is done using a combination of theoretical, simulation-based, as well as with actual measurement results.

1.1 Thesis Organization

In the ensuing chapters, the evolution of the SOIAS technology development, x-ray lithography integration in our CMOS process, and the theoretical evaluations of the applicability of SOIAS to low power systems will be presented in the following order,

- chapter 2 gives a brief background on the requirements for low power and the history of double-gated devices. The basis of the SOIAS device design is also described.

- Chapter 3 focuses on the development of the bonding technology for SOIAS substrates. First a brief history of the theory of silicon fusion bonding is given, followed by the SOIAS substrate preparation process. Finally, the material characterization through electrical measurements are presented.

- Chapter 4 concentrates on the other major technology component of this thesis, which is x-ray lithography for the fabrication of deep submicrometer MOSFETs. The integration of this technology into our CMOS process using a mix-and-match scheme with optical lithography is described, followed by details of the critical resist and dry etching processes. Some device results are presented at the end.

- Chapter 5 presents the measurements on SOIAS devices. First DC static measurement are presented, followed by dynamic measurement results.

- Chapter 6 presents the theoretical energy dissipation model for SOIAS, and benchmarking against a conventional, constant low V_T SOI technology.

- Chapter 7 explores the V_T and V_{DD} design space for optimal operating points for systems with various activities using measured data in conjunction with the energy model.

- Chapter 8 is the summary and conclusion of this work.

Chapter 2

Background

2.1 The Need for Low Power

The rapid emergence of portable wireless electronic systems has spurred researchers to explore various low power optimization techniques at all levels of the hierarchy: architecture, circuit, device and technology. The improvements in battery lifetime, weight, and size, however, does not parallel the explosive commercial growth of portable electronic systems. Therefore, more aggressive power reduction methodologies need to be investigated. Until recently, power consumption considerations in design has always been secondary to performance enhancement. Examples of several strategies adapted for low power design include: (1) switched capacitance reduction, (2) power supply voltage scaling, (3) use of gated clocks, and (4) powering down during long idling periods. The first two methods are at the technology level whereas the latter two are at the circuit and architecture design level. For total reduction in power, all of the strategies should be implemented. Since this thesis is based on technology, focus will be placed on issues related to such (e.g. method 1 and 2). The rationale behind the first two methodologies can be understood by examining the sources of power consumption. For a CMOS technology, there are three main components in the equation for total power dissipation, P_{tot}:

$$
\begin{aligned}
P_{tot} &= P_{dyn} + P_{stat} + P_{sc} \\
P_{dyn} &= \alpha C_{load} V_{dd}^{2} f_{clock} \\
P_{stat} &= I_{Leak} V_{dd} \\
P_{sc} &= I_{ave} V_{dd}
\end{aligned}
$$

where P_{dyn} is the dynamic power dissipation. The dynamic power dissipation is due to the switching of the load capacitance, C_{load}. Therefore, for heavily loaded circuits, the dynamic power dissipation dominates, and it is intuitively obvious that C_{load} and the power supply voltage, V_{dd}, should be reduced in order to decrease the total power dissipation without slowing down the clock. α can best be described as the node switching activity during one clock period, which is strongly dependent on the input signal statistics. P_{stat} is the power dissipation due to leakage currents in the subthreshold region and reverse bias diode leakage currents for bulk CMOS devices. Finally, the short-circuit power, P_{sc}, results from both NMOS and PMOS stacked devices being "ON" simultaneously providing a current path from power supply to ground. I_{ave} is the mean short-circuit current which is strongly dependent on the input and output rise times. The short circuit power can be kept at a low percentage of total power dissipation by balancing the NMOS and the PMOS through transistor sizing assuming the threshold voltages are of the same magnitude.

Capacitance reduction can be accomplished at many levels of design and is fairly straightforward. For example, a Silicon-On-Insulator technology can be used to lower the parasitic node capacitances, and the selection of appropriate logic style and proper transistor sizing can be done at the circuit design level to reduce switched capacitances. Supply voltage scaling is more complex because of the performance/throughput *vs* power consumption trade-off and is strongly dependent on the application. Supply and threshold voltages must be reduced concomitantly if performance losses cannot be tolerated especially at low V_{dd} (e.g 1 V and below). However, the price of re-

ducing the threshold voltage is the increased leakage current since the subthreshold slope cannot be scaled. Therefore, an optimum set of supply and threshold voltages exists which balances the dynamic and leakage power dissipations that will result in minimization of total energy. The switching activity, α, is also a determining factor in finding the optimal threshold and supply voltages. For systems which operate in burst mode or are "event-driven", in which intermittent computations are performed as a result of I/O activity with long periods of inactivity while awaiting I/O, the clock may be stopped completely during inactivity in which case the power consumption should be nearly zero if the leakage power is low, i.e. threshold voltage is high. However, for ultra-low-power applications, the targeted power supply voltages are less than 1 V, and hence the performance will surely suffer. A clever way to address the performance *vs* power trade-off is to dynamically control the threshold voltage. The granularity of the threshold voltage control depends on the system usage.

2.2 Dynamic Threshold Voltage Control Concept

Many system computations are either temporally or spatially localized. Systems that are frequently idle, i.e. doing computation only for a small fraction of the time, operate in burst-mode, and hence exhibit temporal locality. On the other extreme, systems that are active all of the time operate in "continuous" mode, and hence do not exhibit temporal locality. At the same time, a system may only have a fraction of its functional modules active all of the time, such systems exhibit spatial locality. This idea can be applied to lower levels of the hierarchy such as at the logic gate level or the transistor level. One of the widely adapted global strategies for achieving high performance and low power in continuously computing systems (e.g., modules of a video compression system) has been the simultaneous reduction of supply voltage, V_{DD}, and threshold voltage, V_T, where the optimal V_{DD} and V_T are found for minimum total system energy by trading off dynamic energy for static leakage energy[1, 2, 3]. CMOS-based high performance burst-mode computation systems (e.g. a microprocessor running an X-server or cellular phone which is idling more

than 90% of the time) will suffer high static leakage energy dissipation operating at low V_{DD} with constant low V_T even with clocks stopped. For example, even when a user is continuously entering data at the keyboard, the X-server is active, (i.e. doing computation), only 2-3% of the time [4]. In order to simultaneously achieve high performance during active period and low leakage power during idle period for burst-mode computational systems, several schemes of reducing the leakage current have been proposed. The multiple V_T CMOS design involves using high V_T transistors to gate the low V_T functional blocks[5, 6]. Both NMOS and PMOS transistors are needed in order to preserve state. These devices must be made large due to the finite resistance of these transistors. This incurs additional switching energy to switch these devices. Therefore, appropriate sizing of the high V_T transistors is crucial. Another approach is the dynamic control of V_T by biasing the bulk-CMOS wells[7]. A triple well technology is required for this scheme. Furthermore, well biasing is complicated by the N-well to P-well junction leakage current as well as source/drain junction to well leakage currents. Both of the above schemes are implemented at the functional module level; for example, in the well biasing scheme, all the transistors in the functional module have the same variable V_T which is dependent on the well bias.

2.3 SOI Technology

The aforementioned technologies have been mainly proposed for implementation in bulk silicon CMOS. However, the increasing maturity of the SOI technology in the past few years cannot be ignored, especially with the dramatic improvements in material quality. There are two modes of operation for SOI MOSFETs, fully depleted (FD) and partially depleted (PD) channel region (body). In the conventional strongly FD SOI device, the silicon film thickness is usually less than or equal to half of the depletion width of the bulk device. The surface potentials at the front and back interfaces are strongly coupled to each other and capacitively coupled to the front-gate and the substrate through the front-gate oxide and buried oxide, respectively. Therefore,

the potential throughout the silicon film, and hence the charge, is determined by the bias conditions on both the front-gate and the substrate. By replacing the substrate with a back-gate, the device becomes a dual-gated device. The FD design is unique to SOI because the front-gate and back-gate both have control of the charge in the silicon film. In the strongly PD SOI, the back-gate or substrate has no influence on the front surface potential. In the middle regime, the device is nominally PD and can become FD by applying a back-gate bias, thus, coupling of the front and back surface potentials still occurs. In order to implement the dynamic threshold voltage control scheme in SOI, devices must operate in the FD mode.

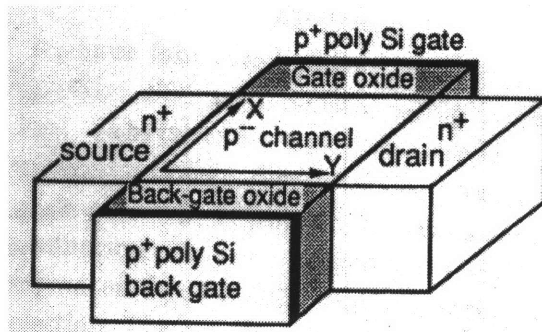
2.3.1 SOI Dual-gated Devices

Various researchers have exploited the use of FD SOI in dual-gated devices in which the top and bottom gates are tied and switched together, resulting in enhanced transconductance and reduced short channel effects [8, 9, 10, 11, 12, 13]. Sekigawa *et al* proposed the earliest dual-gated device with improved V_T roll-off from analytical calculations [13]. Colinge *et al* fabricated the "gate-all-around" MOSFET which they have claimed to show more than 2X enhanced transconductance near threshold compared to that of a single-gate device. They attributed this enhancement to "volume inversion" of the silicon film. However, in strong inversion, the enhancement is minimal, i.e. this device acts like two transistors in parallel. This was also confirmed by Venkatesan *et al* [14]. In the subthreshold region, the potential in the silicon film is more or less flat, i.e. the carriers are more or less evenly distributed in the film. The potential barrier in the film center with respect to the source is lower than in the single gate device, and thus facilitates charge injection from the source, hence more subthreshold current. Furthermore, since the vertical field is small, the carrier in the center of the film does not suffer from surface scattering, and hence may have a higher mobility. Once in strong inversion, the surface potential becomes pinned, and the carriers are predominantly near the surface, and the contribution to total conduction by the carriers in the center of the silicon film is insignificant. This is very nicely shown in the two dimensional plot of the electron potential by Tanaka *et al*, Figure 2-1,

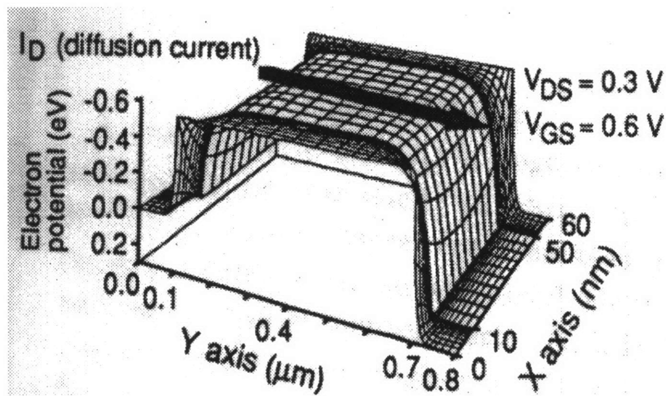
[11]. Balestra *et al* also claimed "volume inversion" in their dual-gated MOSFETs through simulation. However, their confirmation through measurements is dubious because their device is not a true dual-gated device. They used the substrate as the second gate and the bias voltage ratio of back to front-gate is 10X to compensate for the oxide thickness difference. In terms of current drive, the dual-gated device does not show any advantage over two single-gated devices unless it operates in the subthreshold region all the time. The main advantage of the dual-gated device is the scalability into the sub-100nm regime [15, 16, 17, 18, 19]. All these devices have the same work function for both top and bottom gates, i.e same polarity gates. Suzuki *et al* has shown that devices with opposite polarity (n^+-p^+) gates scale better than same polarity (p^+-p^+ or n^+-n^+) gates; and the V_T of the opposite polarity gates devices are in the right range for today's VLSI design requirements due to the work function difference of the top and bottom gates [17, 19]. However, for low voltage operations of the opposite polarity gates devices (⟨ 1 V), the bottom gate does not fully "turn on" due to the difference in V_T of the top and bottom gates. The function of the bottom gate, in this case, is to lower the V_T of the top-gate device during switching.

An ideal dual-gated device should have its top and bottom gates perfectly aligned. With the same thickness top and bottom gate oxides, misalignment of the bottom gate with the source and drain will greatly increase the parasitic capacitance and degrade the short channel effect immunity [18]. Recently, another group introduced a Dynamic Threshold Voltage MOSFET (DTMOS) in SOI [20]. In this structure, the gate of a PD-SOI MOSFET is tied to the body. Therefore, during switching of the gate, a bias is also applied to the body. As a result, the subthreshold slope reaches the theoretical limit of 60 mV/decade at room temperature, and the threshold voltage appears to be lowered. In order not to incur excessive body current, the gate voltage cannot be higher than 0.6 V. The observed improved subthreshold slope and V_T lowering are due to the forward biasing of the source-body p-n junction. By the conventional definition of dynamic threshold control, which is the parallel shift of the subthreshold I-V, the DTMOS does not comply.

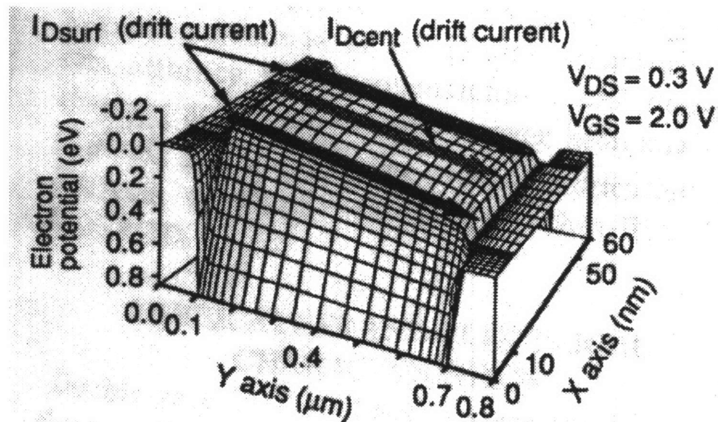
The fabrication of dual-gated devices is a complex process, and switching the



Device Schematic



Subthreshold Conduction



Strong Inversion Conduction

Figure 2-1: Electron potential of dual-gated device in the subthreshold region and strong inversion [11].

back-gate every clock cycle also incurs additional energy. In this thesis work, the development of a practical (i.e. easily integrated) technology to implement the dynamic threshold voltage control concept is investigated. This is a FD-SOI-based technology called SOI-with-Active-Substrate (SOIAS). The particular implementation of the SOIAS back-gated devices does not call for switching of the back-gate every clock cycle. V_T of blocks of devices, i.e. in a module, would be switched similar to the concept of well-biasing in the triple well CMOS.

2.4 Comparison of V_T Controllability in Bulk and SOI MOSFETs

2.4.1 Long Channel Device

In bulk CMOS, the threshold voltage is a square root function of the well bias, and the body factor determines the degree of control[21],

$$\begin{aligned} V_T &= V_{FB} + \phi_B + \gamma(\sqrt{V_{sb} + \phi_B}) \\ \gamma &= \frac{\sqrt{2q\epsilon_{Si}N_A}}{C_{ox}} \\ \phi_B &= 2\frac{kT}{q} ln\frac{N_A}{n_i} \end{aligned}$$

C_{ox} is the gate oxide capacitance, N_A is the channel dopant concentration, q is the electron charge, ϵ_{Si} is the silicon dielectric constant, and γ, the body factor, is strongly dependent on the channel doping concentration,

In FD SOI, the threshold voltage depends linearly on the back-gate bias, and the body factor, δ, does not depend on channel doping concentration[22],

$$V_{Tf} \simeq V_{Tf}^A - \delta(V_{Gb} - V_{Gb}^A)$$

$$
\begin{aligned}
&\simeq V_{Tf}^{I} - \delta(V_{Gb} - V_{Gb}^{I}) \\
V_{Tf}^{A} &= V_{FB}^{f} + (1 + \frac{C_b}{C_{ox}})2\phi_B - \frac{Q_b}{2C_{ox}} \\
V_{Tf}^{I} &= V_{FB}^{f} + 2\phi_B - \frac{Q_b}{2C_{ox}} \\
V_{Gb}^{A} &\simeq V_{FB}^{b} - \frac{C_b}{C_{box}}2\phi_B - \frac{Q_b}{2C_{box}} \\
V_{Gb}^{I} &\simeq V_{FB}^{b} + 2\phi_B - \frac{Q_b}{2C_{box}} \\
\delta &= \frac{C_b C_{box}}{C_{ox}(C_b + C_{box})}
\end{aligned}
$$

where $C_b = \frac{\epsilon_{Si}}{t_{Si}}$, $C_{box} = \frac{\epsilon_{ox}}{t_{box}}$. $Q_b = qN_A t_{Si}$ is the bulk charge for fully depleted SOI, and V_{FB} is the flatband voltage. V_{Gb}^{A} and V_{Gb}^{I} denote the back biases at the onset of back interface accumulation or inversion, respectively; V_{Tf}^{A} and V_{Tf}^{I} are the corresponding front-gate threshold voltages, which are pinned after the back interface has reached either accumulation or inversion.

Although the concept of threshold voltage control in bulk and SOIAS are the same, the design trade-offs are different. The reasons for choosing a SOI based technology rather than a bulk technology are manifold. The SOIAS technology has all the advantages of conventional fully depleted SOI, and offers more flexibility. In bulk CMOS, in order to increase the body factor, the channel doping must be increased. However, in doing so, the threshold voltage at zero bias is also increased. Therefore the body factor in bulk is coupled to the threshold voltage. Proper channel and well doping designs may resolve this problem. In the SOIAS technology, since the body factor is not dependent on the channel doping, the threshold voltage at zero bias can be set independently of the body factor. In threshold voltage switching, the SOIAS device back-gate can be switched in either direction to raise or lower the threshold voltage without suffering from junction leakage currents. In bulk, switching from low to high threshold voltages is more desirable than vice versa because the p-n junctions would be forward biased in the latter case resulting in unwanted leakage currents. The physical decoupling of the SOIAS device back-gate allows ample flexibility in its

design independent of the top-gate device design. In bulk, the well conductance (i.e. doping) cannot be arbitrarily changed without affecting design parameters of the device. Figure 2-2 illustrates the channel design trade-off of body factor and V_T in bulk NMOS devices and the large dynamic V_T range of SOIAS device. In order to achieve a large body factor and low quiescent V_T in bulk devices, the super-steep retrograde (SSR) channel doping design should be utilized. These bulk devices were measured with a chuck contacting the entire back surface of a wafer which clearly cannot be done in CMOS. In triple-well CMOS, in order to increase the well conductance, a retrograde well design may be necessary. In the SOIAS design, the vertical structure design specifications (e.g. t_{ox}, t_{si}, and t_{box} would determine the dynamic V_T range as long as the device is FD.

2.4.2 Short Channel Device

In both the bulk device and SOIAS device, as the gate is scaled into the deep submicrometer regime, unless proper device engineering is employed to reduce the lateral encroachment of the source/drain fields, i.e. short channel effects (SCE), the control of the threshold voltage through the body or the back-gate would decrease. The complex 2-dimensional potential distribution is difficult to treat analytically. For purpose of accuracy, numerical solution of the 2-D Poisson equation should be sought. The charge-sharing concept which most of the analytical models are based provides intuitive understanding of the SCE. Figure 2-3 schematically depicts the charge sharing concept for bulk and SOI devices [23]. The charge, Qd1, in the channel region balances the charge in the top-gate, the source and drain, and the body/back-gate. For a long channel device, the top of the trapezoid is L, the length of the gate, and the bottom of the trapezoid is approximately L. The top and bottom gates are assumed to control the charge in the trapezoidal region L. As the device is scaled, the bottom of the trapezoid becomes less than L and control by the body/back-gate also diminishes. The charge Qd1 in the trapezoidal region, under the control of the top-gate and the body/back-gate is also less. As the source and drain are brought closer to the channel region, in order for the top-gate and body/back-gate to maintain

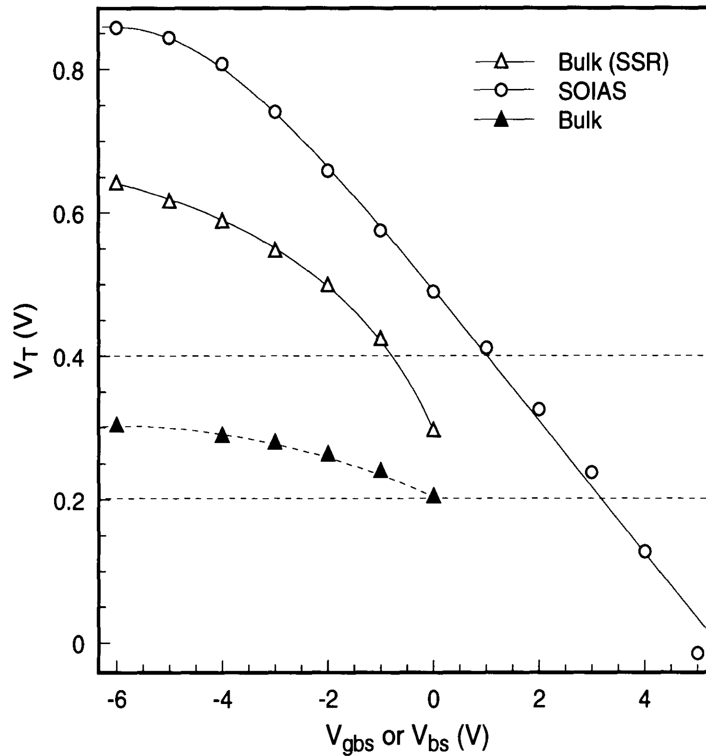


Figure 2-2: Measured threshold voltage control by either body bias in bulk or back-gate bias in SOIAS. The open triangles represent a bulk NMOS device with super-steep retrograde (SSR) channel doping; L_{eff} is approximately 0.75 μm. The filled triangles represent a single step channel doping device, L_{eff} is approximately 0.7 μm. Although the latter has a lower quiescent V_T, the body factor is much smaller than the SSR design. The SOIAS device dynamic V_T range is larger than both of the bulk devices. L_{eff} of the SOIAS device is approximately 0.4 μm.

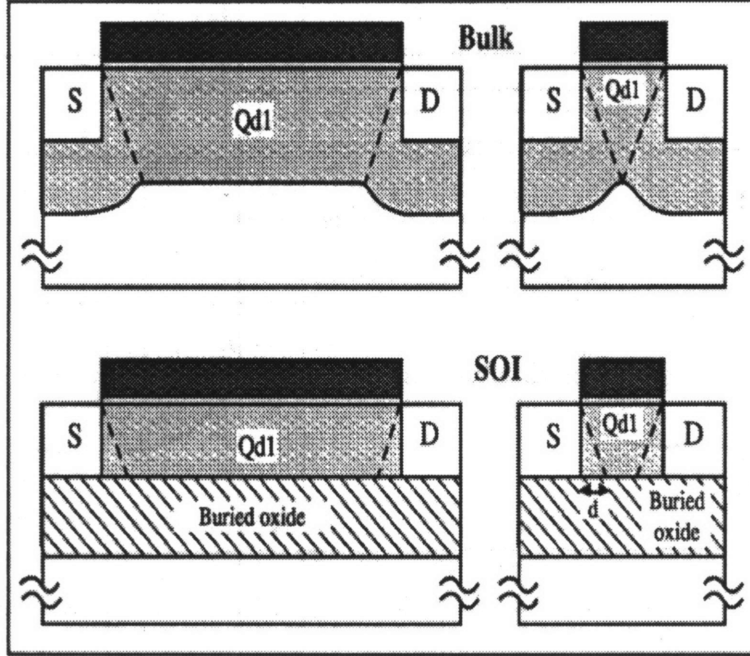


Figure 2-3: Charge sharing in bulk and FD SOI devices [23].

proportionately the same control, similar scaling in the vertical dimension must also be implemented. This requires the gate oxide and the source/drain junction be scaled, and the channel doping be increased. Scaling the gate oxide, physically brings the gate closer to the channel region. Increasing the channel doping decreases the depletion into the channel from the source/drain junctions, and decreasing the junction depth reduces the lateral fringing fields from the source/drain. From the charge-sharing concept and using geometrical constructs, the V_T equation becomes [21]

$$\begin{aligned} V_T &= V_{FB} + \phi_B + \gamma(\sqrt{V_{sb} + \phi_B})(1 - \frac{\alpha\zeta}{L}\sqrt{\phi_B + V_{sb}}) \\ \zeta &= \frac{\sqrt{2\epsilon_{Si}}}{qN_A} \end{aligned}$$

where α is a fitting parameter. As can be seen from the above equation, for short channel devices, there is a linear dependence of V_T on V_{sb}. In FD SOI, again from charge-sharing arguments, the V_T for a FD SOI device, with the back interface de-

pleted, involves replacing Q_b with Q_b(1-d/L) [24], where

$$d \simeq \frac{(V_{bi} - \Psi_{sb})}{E_{b(eff)}} \tag{2.1}$$

V_{bi} is the source/drain to channel junction built-in potential, and Ψ_{sb} is the back interface potential at half way between the source and the drain. $E_{b(eff)}$ is the effective lateral component of the electric field at the back interface which is given approximately by

$$\begin{aligned} E_{b(eff)} &= [qN_A(V_{bi} - \Psi sb)/2\epsilon_{Si}]^{1/2} \\ &\quad + f_a \frac{\epsilon_{ox}}{\epsilon_{Si}} \frac{V_{Gb} - {V_{FB}}^b - \Psi_{sb}}{t_{box}} \\ &\quad + f_b \frac{\epsilon_{ox}}{\epsilon_{Si}} \frac{V_{bi} - {V_{Gb}}^b + {V_{FB}}^b}{t_{box}} \end{aligned}$$

Ψsb and $E_{b(eff)}$ must both be solved iteratively in order to obtain d [24]. f_a and f_b are fitting parameters. In addition to reducing the front-gate oxide, the back-gate oxide also needs to be reduced in order to maintain the back-gate's control on the charges in the channel, and to reduce the coupling from the source/drain fringing fields, hence larger $E_{b(eff)}$, and smaller d. Reducing the silicon film thickness effectively decreases the junction depth. Increasing the silicon film doping also increases the $E_{b(eff)}$, i.e. reduces depletion width of the source/drain and channel junction. The downside of having a large body factor in both technologies is that the parasitic source/drain capacitances will increase, the saturation current will decrease, and stacked gates will suffer larger delays. The effects of the drain bias can be included in the above equations using the same charge-sharing arguments for bulk; in FD SOI, however, there is no simple way to incorporate this effect due to the coupling to the back interface, therefore, 2-D numerical simulations is the best way to study these device structures.

2.5 SOIAS Device Structure

The general structure of interest for this work is the back-gated device depicted in Figure 2-4. The top and bottom gates are designed with opposite polarities with the

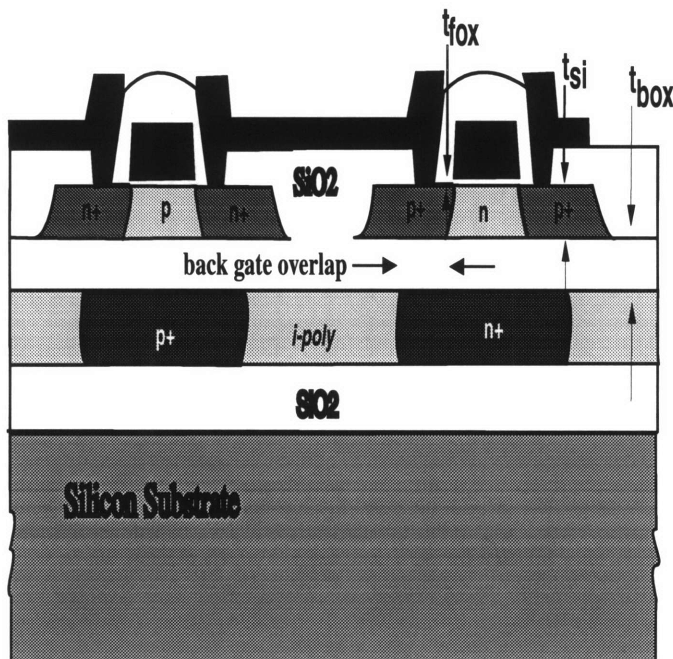


Figure 2-4: SOIAS back-gated CMOS device schematic.

back-gate having the same polarity as the silicon film. The top and bottom gates are separate and can be switched independently. Therefore, to mimic a dual-gated device, the top and the bottom gates can be tied and swept together. At 1 V or less on V_{DD}, the back-gate would still function as a threshold voltage modulator of the front-gate. The decision to either implement a dual-gated or back-gated device structure depends on the system usage (burst-mode *vs* continuously computing), and the bottom gate switching energy *vs* technology complexity/difficulty trade-offs since a perfectly self-aligned top and bottom gate device is impossible to achieve with technologies that exist to date. In choosing the device design parameters for the SOIAS devices, the emphasis was placed on the maximum process latitude which will yield devices that meet the V_T and off current requirements at 1 V supply voltage. Figure 2-5 shows an example of such a device from MINIMOS simulation. The specifications for this device is 3-4 decades change in off current at V_{DD} of 1 V, corresponding to 200 mV of switch

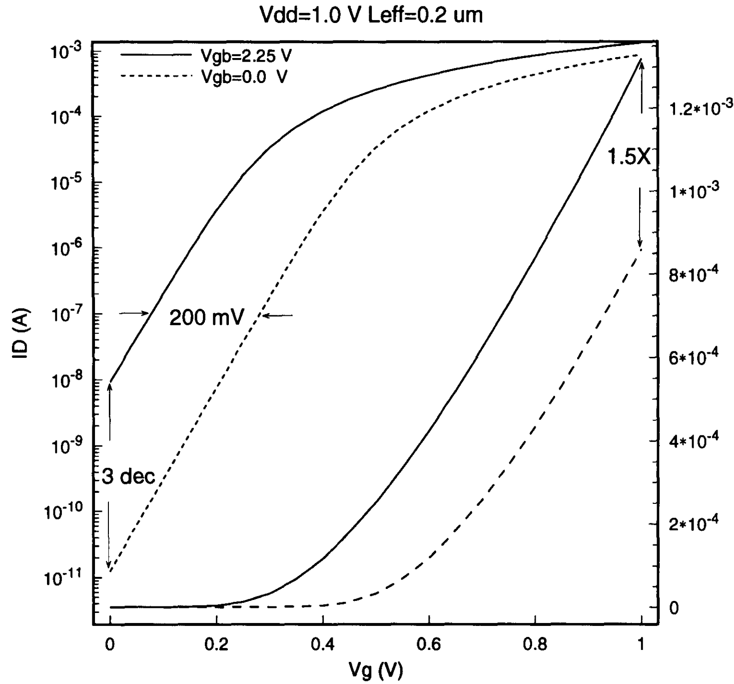


Figure 2-5: Simulated subthreshold I-V at different back-gate biases.

in V_T. The design parameters are 7 nm front-gate oxide, 40 nm silicon film thickness, 80 nm back-gate oxide thickness. The thick back-gate oxide allows maximum latitude in the formation of the back-gate since source/drain overlap capacitance is small in comparison to the switching capacitance of the front-gate. The silicon film is as thin as practically achievable to maximize the control of the back-gate on the front-gate V_T. The devices fabricated in this work are designed around these specifications and device parameters. For scaling into the sub-0.2 μm regime, the vertical structure (t_{ox}, t_{Si}, and t_{box}) must be redesigned to achieve the same threshold voltage control and minimize SCE.

Since the SOIAS device operates in the FD mode, some of the issues associated with the design of FD SOI devices are still present while others are ameliorated by the ability to control the threshold voltage, and to operate at lower supply voltages. One of the key challenges in properly designing the conventional SOI FD devices is obtaining a reasonable threshold voltage in very thin FD films. Especially when designing in the deep submicrometer regime, the tradeoffs of the silicon film thickness, channel doping, and buried oxide thickness needs to be carefully evaluated. Conven-

tionally, the FD SOI has 200-400 nm of buried oxide, and very thin silicon film (40-60 nm). The SOIAS device requires the buried oxide thickness to be reduced to less than 100 nm, and the silicon to approximately 40 nm, and thus reduces the short channel effects, as well as increases the the back-gate control on the front-gate threshold voltage. Since the threshold voltage of the SOIAS device can be precisely controlled electronically, the requirements on the channel doping can be relaxed as long as the film is fully depleted. The large series resistance problem associated with thin silicon films without a good silicide still exist. The reduction of the supply voltage to 1 V or below also avoids the problem of large field-induced drain breakdown.

Chapter 3

Silicon-On-Insulator-with-Active-Substrate: SOIAS

3.1 Silicon Direct Bonding

Silicon direct bonding or fusion bonding, as the name implies, is the joining together of two silicon wafers (with or without oxide) at room temperature without the aid of an high electric field or gluing agent, and subsequently annealing at high temperatures to seal the bond. Although this concept has been around for quite some time, with the first patent filed in 1966 [25], the demand for SOI material was insignificant at that time. Then in 1985, Lasky *et al* [26] demonstrated the fabrication of NMOS SOI devices using the bonding and etch-back methodology which spurred a flurry of research and development. Since then, much work has been published which contributed to the understanding of the bonding mechanism and interface quality characterization. Any two very flat and smooth surfaces when brought into contact form a weak bond. The materials of the two surfaces can be metal, semiconductor, insulator, or any combination of these [26, 27, 28, 29, 30]. Once the two surfaces are in contact, a "contact wave" spreads across the entire area from the point of

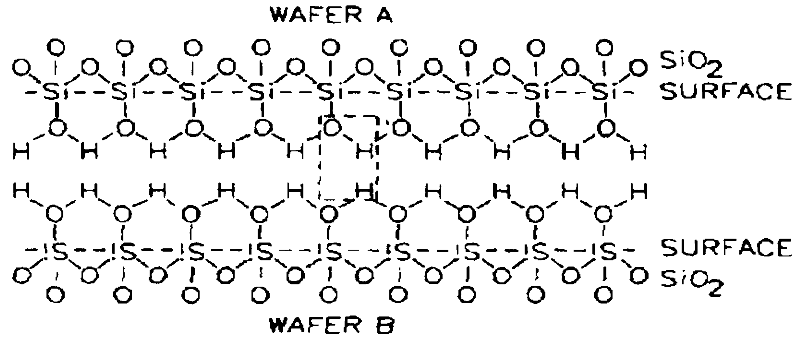


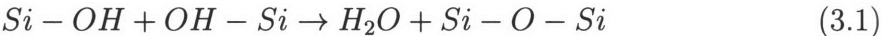
Figure 3-1: Hydrogen bond formation between two hydrophilic surfaces [26].

contact. The speed of the "contact wave" depends on the flatness of the wafers, the smoothness of the surfaces, and the chemical treatment of the surfaces prior to bonding. The surfaces of these wafers are usually chemically treated to render them either hydrophilic or hydrophobic. The type of bond formed and the bond strength at room temperature depend on the chemical surface treatment. The bond strength after annealing depends on the annealing temperature. The bonding mechanism and bond strength evaluation has been studied extensively. An excellent review paper on this topic can be found in reference [31].

3.1.1 The Bonding Mechanism

The wafer surface can be rendered hydrophilic through various treatments. Typically cleaning with heated hydrogen peroxide based solution (e.g. RCA [32] clean or pirahna clean (H_2SO_4:H_2O_2)) would yield hydrophilic surfaces. Some groups had tried treatment with hot nitric acid (HNO_3) after the above mentioned cleaning procedures to enhance the hydrophilicity of the surface [33, 34]. For hydrophilic surfaces, the bonding at room temperature is predominately caused by hydrogen bonds between hydroxyl groups and water molecules adsorbed on the surfaces when exposed to air[26], see Figure 3-1. Van der Waals forces between the atoms and the adsorbed molecules also exist [35]. These bonds are electrostatic and relatively weak. The bond strength improvement with annealing temperature is believed to occur in three different regimes with different mechanisms. The temperature boundary is by no means

clearly defined. When the room temperature bonded wafers are annealed at low temperatures (100-200^oC), enhanced mobility of the adsorbed molecules is believed to cause more bonds to bridge the gap between the two surfaces, and hence strengthen the bond [36]. For annealing in the 300 ^oC range, a chemical bonding reaction occurs which replaces the weak electrostatic bonds with covalent Si-O-Si bonds according to the following reaction [26, 36, 37]:



$$Si - OH + OH - Si \rightarrow H_2O + Si - O - Si \qquad (3.1)$$

Abe *et al* believed that the temperature at which the above reaction occurs is around 600 ^oC [38]. Stengl *et al* proposed a two step mechanism for the formation of the Si-O-Si bonds as depicted in Figure 3-2[39].

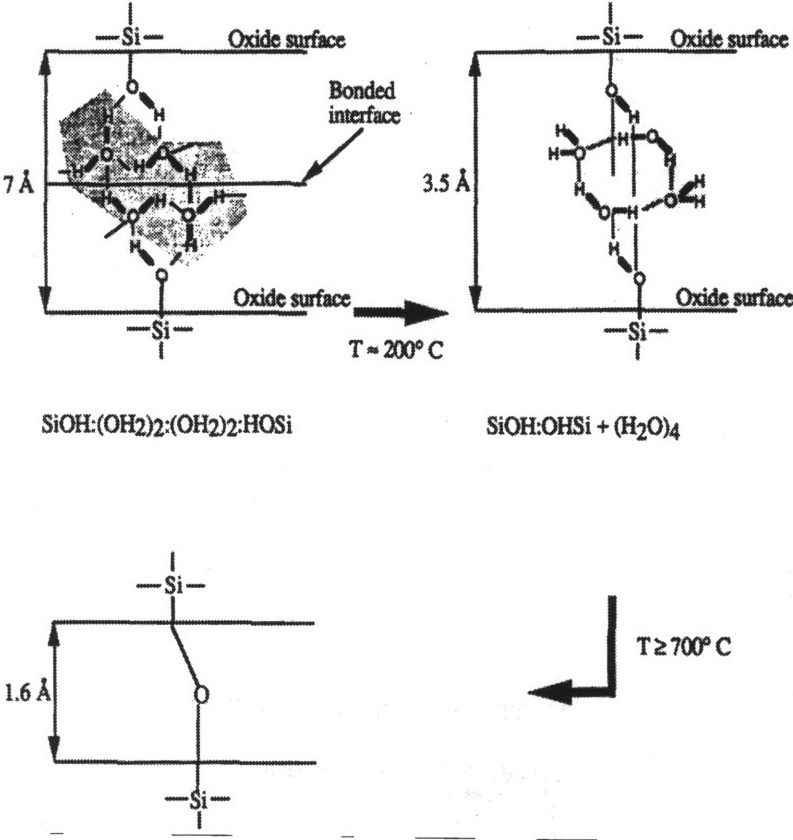


Figure 3-2: The bonding mechanism proposed by Stengl *et al* showing the transition from weak hydrogen bonding to covalent Si-O-Si bonding at two different temperatures [39].

These authors proposed that at temperatures above 200 ^oC, the adsorbed water molecules aggregate into clusters of four molecules, thus reducing the gap between the wafers. At temperatures equal to or above 700 ^oC, the water molecules diffuse away leaving the Si-O-Si bonds [39]. After the formation of the Si-O-Si bonds, annealing in the temperature range from approximately 600-1100 ^oC strengthens the bond more because the bonded area increases due to elastic deformation of the wafer around the peripheries of unbonded areas (e.g. a particle) [36]. At even higher temperatures (i.e. $\rangle$ 1100 ^oC), the viscous flow of the oxide fills in the microvoids, and the bond strength depends on the mass transport phenomenon, i.e. length of the anneal [36, 40].

Non-hydrophilic surfaces can also be bonded [41, 34, 42, 43, 44]. Spontaneous bonding usually does not occur on these surfaces, but with slight force the wafers can be bonded at room temperature. The hydrophobic silicon wafers are usually dipped in dilute HF prior to bonding. The bonding mechanism for hydrophobic surfaces still remains, for the most part, a mystery. Bengtsson and Engstrom proposed that after treatment in dilute HF($\sim$ 2%), the surface is predominately terminated by hydrogen atoms; however, the subsequent rinse in deionized water and exposure to air increases the density of hydroxyl groups on the wafer surface, although the surface remains hydrophobic [42]. This may be a possible explanation of adhesion between two hydrophobic surfaces. Himi *et al* used a similar explanation with highly concentrated HF ($\sim$ 49%). For treatment in highly concentrated HF, the surface is predominately terminated with fluorine atoms ($\sim$ 37%). Dipping such a wafer in deionized water will result in replacement of the fluorine atoms with OH groups. They believed that although the amount of OH groups for such a surface is less than the hydrophilic surface, it is enough to form the initial bond at room temperature [43]. Of course, the Van Der Waals forces always exist in both cases, and can also aid in the initial room temperature bonding. Therefore, a hydrophilic surface is not necessary for successful bonding. The bond strength after annealing at high temperatures for the hydrophobic surfaces is stronger than the hydrophilic case. This is due to the epitaxial bonding at the interfaces that is free of oxide [43].

3.1.2 Voids

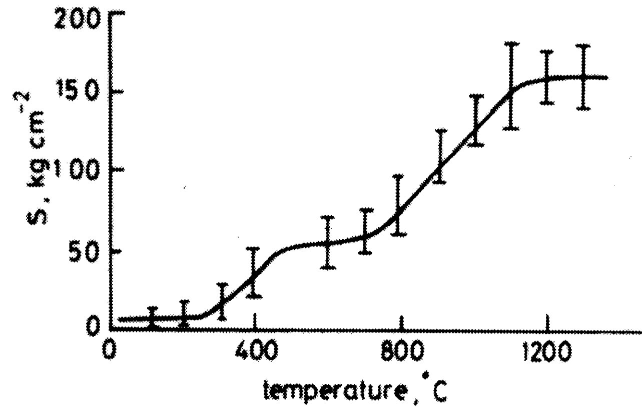
Achieving wafer bonding that is free of voids is an extremely challenging task. There are two category of voids; extrinsic voids and intrinsic voids. Extrinsic voids are those resulting from particulates on the wafer surfaces, local contamination, surface roughness, non-flat surfaces, or gas trapped in pockets when two or more "contact waves" are generated simultaneously at the mating of the two surfaces at room temperature. These type of voids are present at the time of room temperature bonding and would remain after the high temperature anneals. In order to reduce the extrinsic voids, wafer cleaning prior to bonding is critical, and bonding in a cleanroom environment is required. Wafer bowing tend to introduce voids due to multiple points of contact when the wafers are pressed together. Extreme wafer bowing prevents bonding altogether. This problem is more apparent in cases where stressed film layers are deposited on the wafer. Intrinsic voids are only present in silicon to silicon bonding with both hydrophilic and hydrophobic surfaces at intermediate annealing temperatures ($\sim 200 - 800^oC$). These voids can be observed after annealing at these temperatures following room temperature void-free bonding. Above 1000 ^{o}C, the voids disappear. There is no consensus on the origin of these voids. One explanation proposed was that the voids are formed from the release of water molecules during the Si-O-Si bond formation reaction [27, 39]. The belief is that, at elevated temperatures, the water in the voids would oxidize the silicon. Another group claimed that hydrocarbon contamination of the wafer surface is the cause of these temperature dependent voids [45, 46]. This theory is in part supported by the observation that HF treated wafers has more intrinsic voids compared to those without the HF treatment [45, 34]. At high enough annealing temperatures, either the hydrocarbons decompose into smaller groups and the gas in the voids migrates to the edge of the wafer, or diffuse into the wafer. The total effect may be a combination of the two theories. The biggest mystery, however, is the absence of intrinsic voids at these intermediate temperatures for bonding silicon to an oxidized wafer or quartz. Abe *et al* has shown that at least 20 nm of silicon dioxide is needed on one of the wafers to

produce void-free bonded pairs with annealing in the intermediate temperature range [28]. The same observation was made for bonding silicon to sapphire [28]. One explanation may be that whatever is trapped at the interface, be it water or hydrocarbon groups, can readily diffuse into the oxide at lower annealing temperatures. Annealing at lower temperatures is attractive for bonding a wide variety of materials, especially for metals used in interconnects. Recently, successful bonding of cobalt silicide and tungsten has been demonstrated at annealing temperatures below 900 °C [29, 30]. This author believes this will be the trend of wafer bonding in the future.

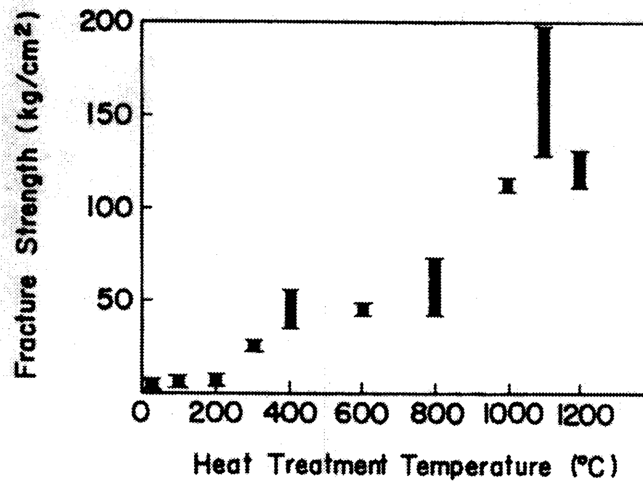
3.1.3 Bond Strength

The methods derived for evaluation of bonding strength are destructive and semi-quantitative. Two widely used methods are the measurements of fracture strength and surface energy [27, 36, 47, 48]. The fracture strength technique is a direct measurement of the bond strength and involves physical separation either at the bond interface, which yields the true bond quality, or fracture in one of the bonded wafers which implies that the bond strength is equivalent to or greater than the fracture strength of the substrate. Usually a strong adhesive is used to glue "handles" to the bonded wafers, and a force perpendicular to the bonded interface is applied to pull the wafers apart or until fracture. Figure 3-3 shows the measured fracture strength for silicon to silicon bonding and oxide to oxide bonding as a function of annealing temperature [27, 47].

The surface energy measurement technique was developed by Maszara *et al* [36] based on the theory of crack propagation in a cleaved bulk sample developed by Gillis and Gilman [49]. The surface energy measurement technique involves the insertion of a blade of a certain thickness at the bonding interface, and the length of the crack is the distance from the edge of the wafer where the blade was inserted to the point where the wafers still remained bonded. The surface energy is inversely proportional to this length [36]. Figure 3-4 shows the measured surface energy as a function of annealing temperature for silicon to silicon and oxide to oxide bonding [36, 50]. All these figures show a similar trend with three fairly distinct regions which is consistent



(a)



(b)

Figure 3-3: (a)Fracture strength for SiO_2 to SiO_2 bonding [47]. (b)Fracture strength for Si to Si bonding [27].

with the explanations given above for the bonding mechanisms at different annealing temperatures. The surface energy for silicon to silicon bonding seems to be a lot higher than oxide to oxide bonding which is contradictory to the direct fracture strength measurement. Since the loading force for these two measurements are different, perpendicular to the bonded interface for the fracture strength measurement and parallel for the surface energy measurement, a direct translation of surface energy to fracture strength may not be straightforward.

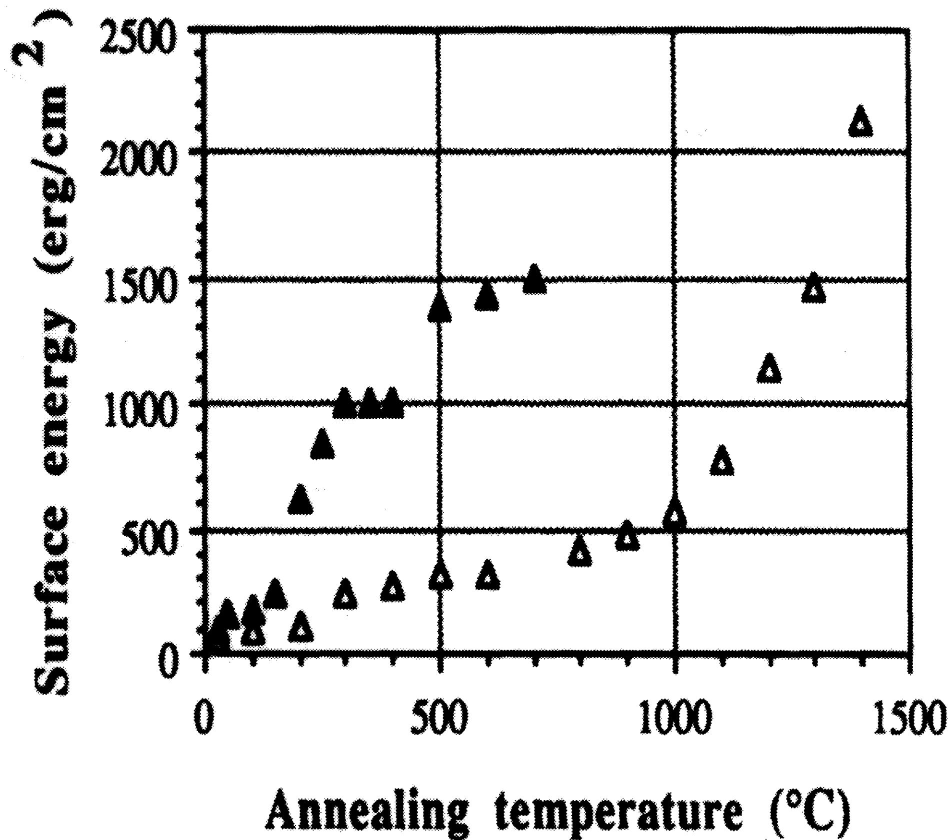


Figure 3-4: Surface energy of SiO_2 to SiO_2 bonded interface (open symbols) [36], and that of Si to Si bonded interface (filled symbols) [50].

3.2 SOIAS Technology

The development of the Silicon-On-Insulator-with-Active-Substrate (SOIAS) technology was guided by the understanding of the silicon bonding mechanism as presented in the previous sections. In conventional silicon fusion bonding, the bonded wafers usually do not contain more than two layers of thin films. The concept of SOIAS technology can be taken to many levels of complexity. The fundamental idea behind this technology is to add one or more conductive under-layers beneath the buried oxide of a SOI structure. Such layers can serve as buried interconnects or gates or both, i.e. enabling integration in the third dimension. The fabrication of SOIAS structures leverages off many of the technologies developed for bulk and SOI CMOS processes (e.g. CMP and wafer bonding). There are several options and var-

ious degrees in which the buried layer or layers can be rendered conductive. In one extreme, the buried layer can be a refractory metal such as tungsten, or silicides of such metals which can withstand subsequent high temperature processing. In this case, the buried conductive layer must be pre-patterned prior to bonding, and a dielectric must be deposited and planarized. This bonding process is more complicated and challenging. In the other extreme, a blanket insulating/semi-insulating layer (e.g.. intrinsic amorphous/polycrystalline silicon) can be used, and selective areas of the buried layer can be made conductive by ion implantation with dopants. This work focuses on the development of the latter approach.

3.2.1 SOIAS Wafer Fabrication

Two different approaches have been taken for preparing the SOIAS wafers. The first is the more traditional route of the bond-and-etch-back (BESOI) process. In this case, the device wafer includes the back-gate oxide (to be) which is obtained by dry thermal oxidation, and the back-gate material to be which is amorphous silicon (as deposited). Amorphous silicon was used as the back-gate material because as-deposited amorphous silicon is very smooth which facilitates direct bonding to the oxidized handle wafer. The handle wafer serves only as mechanical support for the various layers of film. The device wafer was then bonded to the handle wafer which was also oxidized to form approximately 1 μm of silicon dioxide. Therefore, the bonding interface is between the amorphous silicon and the thick insulating oxide. Figure 3-5 depicts the SOIAS preparation for the BESOI process. Immediately prior to bonding, the wafer surfaces were prepared by a RCA clean without the HF treatment. Thus, the surfaces are hydrophilic, facilitating bonding at room temperature. The bonding can be done either in air or in oxygen ambient. Bonding in air involves placing two wafers, with polished sides facing each other, together in a Teflon wafer carrier, and then applying slight pressure to the center at a point to initiate the "contact wave". Figure 3-6 schematically depicts the bonding apparatus used for bonding in an oxygen ambient. The wafers were placed, polished side facing each other, with the device wafer usually on top of the handle wafer, initially separated by

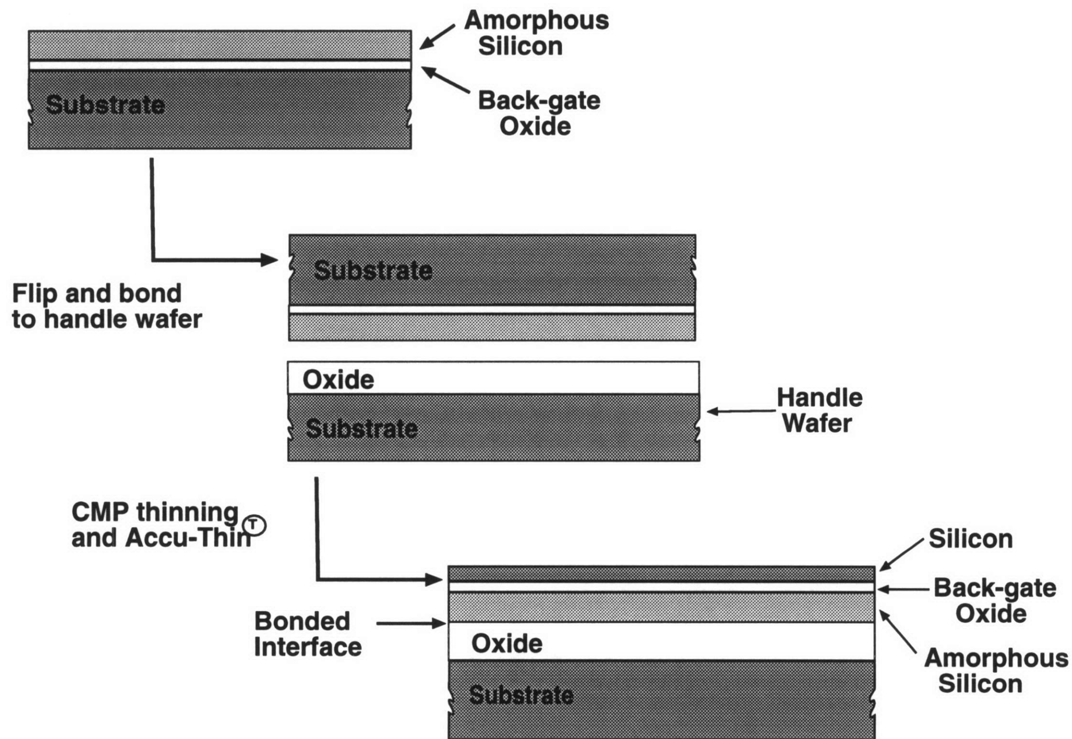


Figure 3-5: The bond and etch back (BESOI) process for fabricating SOIAS substrates.

Teflon spacers. These wafers were spun at ~2000 rpm with oxygen flowing. Once the chamber is completely purged with oxygen. The Teflon spacers were knocked out, and the top wafer drops onto the bottom wafer. Again, a slight pressure is applied to the center of the wafer to initiate the "contact wave". The device wafer thinning was accomplished by chemical and mechanical polishing. Localized plasma thinning (Accu-Thin^{TM}) [51] was used to improve silicon film thickness uniformity. A touch polish of the surface was performed after the localized plasma thinning and prior to thermal oxidation thinning to reach the final desired silicon film thickness.

The second approach involves the bonding of a SIMOX wafer and relying on the buried oxide for etch-stop using wet chemistry wafer etching. Bonding of SIMOX to quartz and silicon substrates has been shown to be feasible [52, 53]. The same layers as described above were grown on the SIMOX and the handle wafers, and the bonding interface was still between the amorphous silicon and the insulating oxide with the same surface treatment prior to bonding. Figure 3-7 depicts the SOIAS preparation

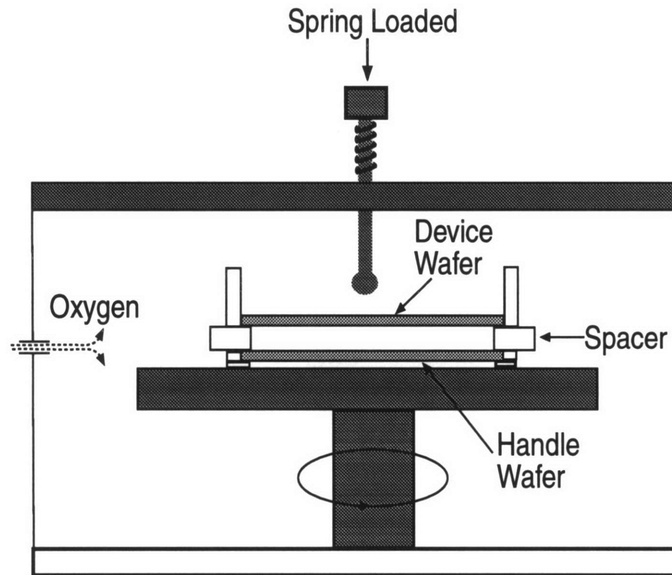


Figure 3-6: bonding chamber

for the bonded SIMOX process. The bulk of the SIMOX substrate was removed by coarse grinding and polishing leaving approximately 60 μm of silicon depending on the the bow of the wafer. The final removal was carried out by wet chemical etching in 25 wt% Tetramethyl Ammonium Hydroxide (TMAH), commonly used as photoresist developer, stopping on the buried oxide. The TMAH etch was done in a reflux condenser which allows the evaporated water to condense back into the solution thus maintaining the same concentration throughout the etch. The selectivity of silicon to oxide in TMAH is about 5000 to 1 at 80°C [54]. The buried oxide serves as a perfect etch-stop. Therefore, the resulting silicon film thickness is as uniform as that of the original SIMOX wafer. The buried oxide was then removed in three wet etching steps (50:1 H_2O:HF,TMAH,7:1 H_2O:HF). The first HF etch is to remove the few hundred angstroms of buried oxide and thus exposing the silicon inclusions typically found in the SIMOX buried oxide. A short TMAH etch follows the HF dip to remove the silicon inclusions in order to avoid particulate contamination. The final HF etch is to remove the remaining buried oxide. Final thinning of the silicon film was accomplished with thermal oxidation and wet oxide strip. In both the BESOI and bonded SIMOX approach, the thermal oxide and LPCVD amorphous silicon did

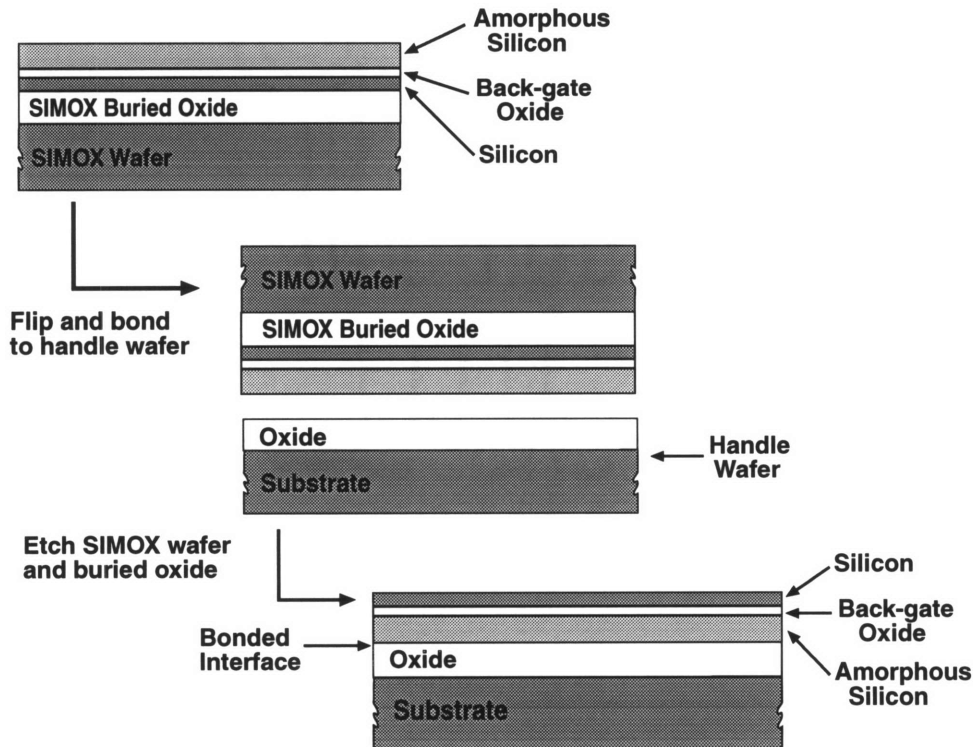


Figure 3-7: The bonded SIMOX process for fabricating SOIAS substrates.

not produce any significant wafer bowing because the films were uniformly deposited on both sides of the wafer.

3.2.2 Bonded Interface Evaluation

Infrared imaging provides immediate feedback of the bonding quality by making relatively large voids (with at least 0.25 μm separation between the wafers) visible to the eye through an infrared camera. Voids appear either as bright spots or interference fringes. The infrared imaging apparatus consists of a video camera, a broad band light source, and a TV monitor. Since undoped silicon is transparent to the infrared, the wafer itself acts as a filter. Figure 3-8 is a photograph of the infrared image of a bonded wafer prior to high temperature anneal. If the room temperature bonding is not satisfactory, the wafers can be separated at this point by inserting a blade between them to split them apart. After bonding, the wafers were annealed either in an inert N_2 ambient or an oxidizing ambient at 1000 ^oC for one hour to seal the bond. We have found that small voids formed during bonding in oxygen ambient can

be annihilated by annealing in an oxidizing ambient. Acoustic microscope imaging can also be used to evaluate the bonding quality. This method allows the detection of very small voids. Figure 3-9 is an acoustic microscope image of a bonded SIMOX wafer. Since this technique is not readily available in our research lab, only a few sample wafers were evaluated, and they do not represent the overall bonding quality of our wafers. We did not perform any destructive bond strength measurements such as the fracture test or the surface energy measurement since all our bonded wafers were used in fabricating CMOS devices and circuits. The bond was strong enough to survive grinding and polishing of the device wafer in the thinning process. Figure 3-10a-b show the final SOIAS substrates, after thinning of the device wafer and before thermal oxidation thinning, prepared by the bonded SIMOX and BESOI processes, respectively.

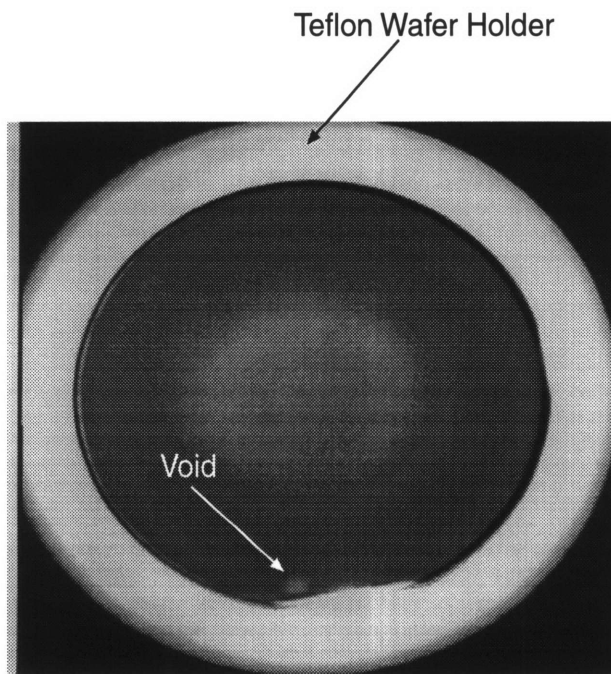


Figure 3-8: Infrared image of bonded wafers.

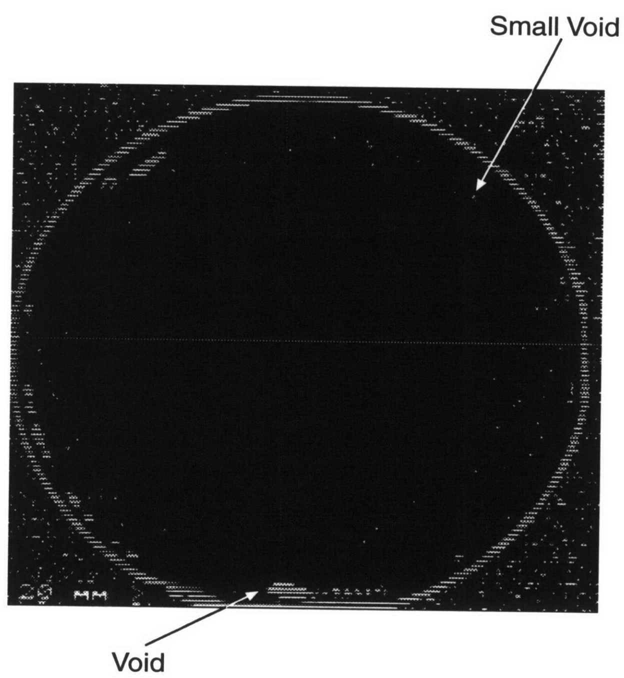
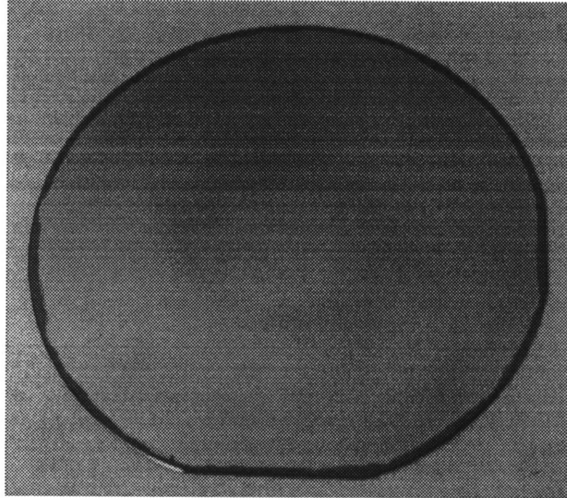


Figure 3-9: Acoustic microscope image of bonded wafers.



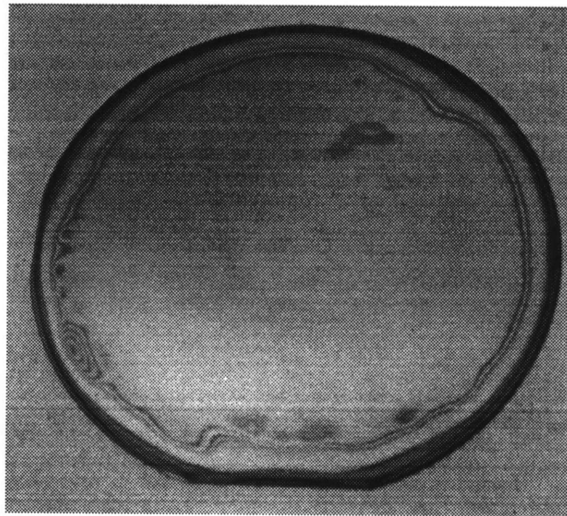


Figure 3-10: (a)SOIAS wafer prepared by the bonded SIMOX process. (b)SOIAS wafer prepared by the BESOI method. The absence of fringes indicates the better film thickness uniformity of the bonded SIMOX SOIAS as compared to the BESOI SOIAS.

3.2.3 SOIAS Materials Characterization

Since the ultimate goal of this work is to demonstrate functional CMOS devices and circuits on the SOIAS substrates, the integrity of the SOIAS substrates was evaluated mainly by electrical characterization through measurements of the effective electron mobility, Time Zero Dielectric Breakdown (TZDB) tests for intrinsic oxide quality, and average interface state density measurement using the charge pumping technique. The SOIAS substrate prepared by the bonded SIMOX technique is of particular interest since the front surface, where the devices would be built, was the back interface of the original SIMOX which is known to be rough. The only physical characterization used was surface profilometry with an Atomic Force Microscope (AFM) to determine the roughness. Table 3.1 shows the root mean square of surface height variation (RMS), and average surface roughness (RA) for different substrates with different sample areas. Figure 3-11 shows the effective electron mobility vs effective

	5x5 μm^2 Sample Area		205x205 nm^2 Sample Area	
	RMS(nm)	RA(nm)	RMS(nm)	RA(nm)
Bulk	0.15	0.12	0.14	0.10
SIMOX	0.77	0.62	0.40	0.30
SOIAS(BESOI)	0.26	0.20	0.22	0.15
SOIAS(BSIMOX)	2.35	1.86	0.64	0.40

Table 3.1: Surface roughness comparison for different substrate materials.

transverse electric field of the front-gate device for conventional SIMOX and SOIAS. All measurements were taken on 50x50 μm NMOS devices. The effective channel mobility was extracted from drain current and gate capacitance measurements using the following two equations:

$$\mu_{eff} = \frac{I_d}{\left(\frac{W}{L}\right) Q_i V_{ds}} \tag{3.2}$$

$$Q_i\left(V_{gs}\right) = \int_{-\infty}^{V_{gs}} C_g\left(V_{gs}\right) dV_{gs} \tag{3.3}$$

Since the carriers are confined to travel across the channel near the surface, surface

roughness would result in carrier scattering, reducing the effective mobility. The universality of the curves in Figure 3-11 indicates no apparent difference between the SOIAS and SIMOX substrates from a device operation point of view.

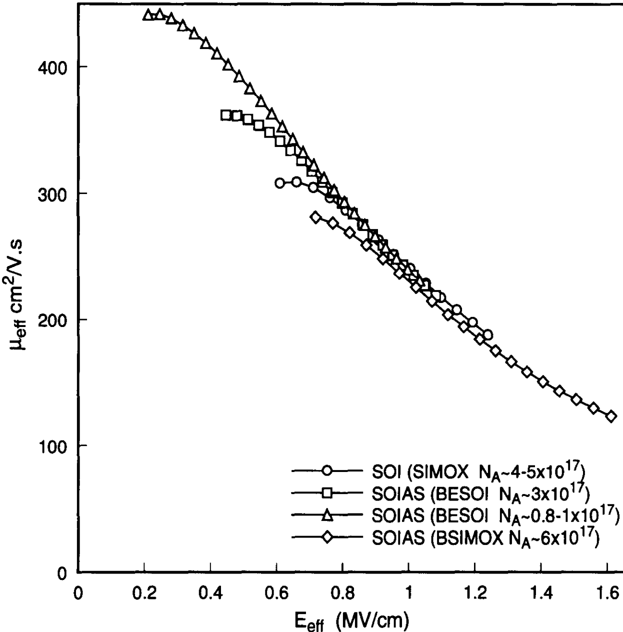


Figure 3-11: Comparison of effective electron mobility for various substrate materials.

Figure 3-12 shows the cumulative percentage failure comparison of SOIAS with bulk and SIMOX substrates in the TZDB test. All devices tested were 10x10 μm. Surface asperity would result in enhanced localized electric fields, thus causing oxide breakdown to occur at lower applied electric fields as compared to a smooth surface. The SOIAS substrate prepared by the bonded SIMOX method is slightly worse than the bond and etch-back SOIAS as well as the bulk and SIMOX. Overall, the intrinsic oxide breakdown of the SOIAS is comparable to those of bulk and SIMOX. The charge pumping technique for evaluating the Si/SiO_2 interface state density has been well established [55, 56]. Reference [55] provides a thorough and detailed analysis of the charge pumping technique, and therefore will not be elaborated here. The charge pumping current, measured at the body of the devices, is given by the classic equation 3.4. This equation relates the charge pumping current to the average interface states

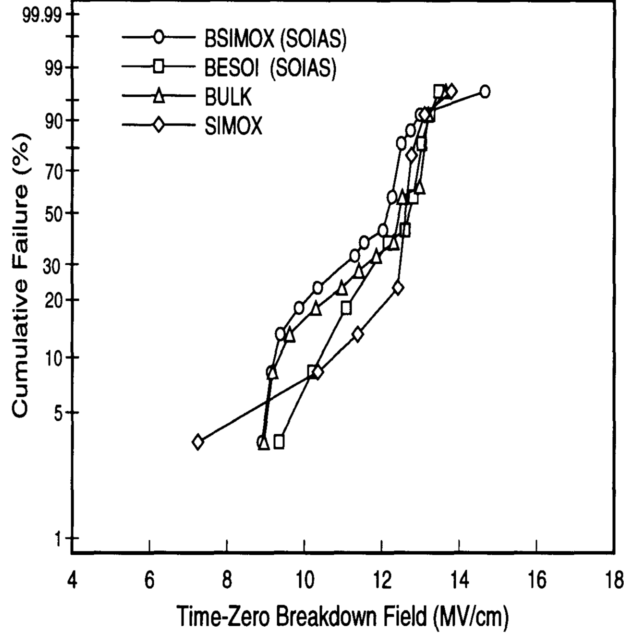


Figure 3-12: Comparison of cumulative percentage failure of 10x10 μm transistors for different substrates.

density, $\bar{D}_{it}$, the thermal carrier velocity, v_{th}, the capture cross section for electrons and holes, (σ_n,σ_p),the frequency of the gate pulse, and the pulse shape as indicated by the rise and fall times (t_r, t_f).

$$I_{cp} = 2qfWL\bar{D}_{it}kTln\left(v_{th}n_i\sqrt{\sigma_n\sigma_p}\frac{|V_{FB}-V_T|}{|\Delta V_G|}\sqrt{t_rt_f}\right) \tag{3.4}$$

Figure 3-13 shows charge pumping characteristics for a NMOS SOIAS device. Equation 3.4 is valid when the channel is pulsed from accumulation to inversion,i.e. the peak of the charge pumping current curves. For a trapezoidal pulse, in order to extract $\bar{D}_{it}$, V_T, V_{FB}, v_{th}, σ_n, and σ_p must be known. However, if the pulse is triangular with equal rise and fall times,i.e. $t_r + t_f = 1/f$ and $\sqrt{t_rt_f} = 1/2f$, then equation 3.4 reduces to the following

$$Q_{it} = \frac{I_{cp}}{fWL} = 2qkT\bar{D}_{it}ln\left(\frac{v_{th}n_i\sqrt{\sigma_n\sigma_p}|V_{FB}-V_T|}{|\Delta V_G|}\frac{1}{2f}\right) \tag{3.5}$$

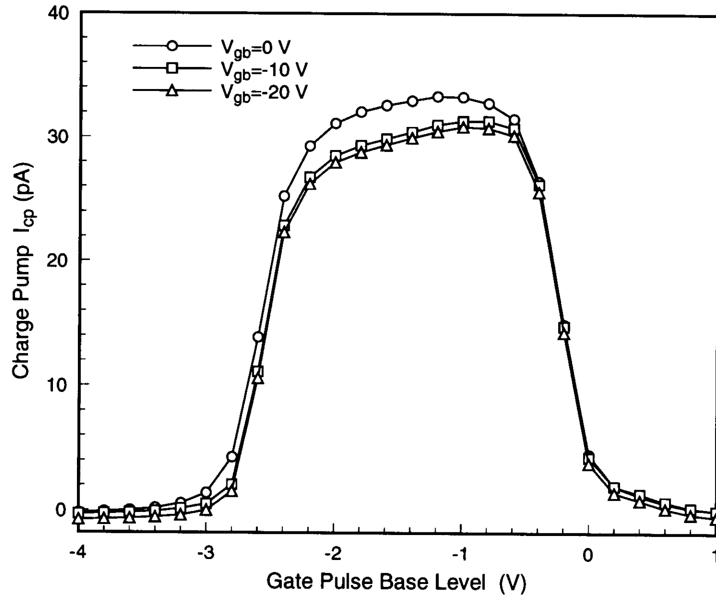


Figure 3-13: Charge pumping current as a function of back-gate bias on a SOIAS device.

where Q_{it} is the recombination charge per cycle. From the semilogarithmic plot of Q_{it} and ln f, the slope of the straight line yields $2qkT\bar{D}_{it}$. From this, the average interface states can be extracted without knowing any of the other parameters [57]. Figure 3-14 shows the semilogarithmic plots of Q_{it} and ln f. This technique works very well for bulk devices, however, for thin film SOI/SOIAS devices, the coupling between the two interfaces will lead to charge pumping currents which result from trapping and de-trapping of carriers at both interfaces. In order to isolate the top interface, the back interface must either be in strong accumulation or inversion so that all the interface state traps at the back are filled during charge pumping of the front interface. Figure 3-13 shows that by applying a back-gate bias to accumulate the back interface, the peak charge pumping current decreases and eventually saturates when the back is in strong accumulation, thus decoupling the two interfaces. Figure 3-14 shows the semilogarithmic plots from which the slope and consequently $\bar{D}_{it}$ was extracted for the various substrates. Table 3.2 summarizes the results. All of the material analysis results indicate that the BESOI SOIAS substrates are very similar to the bulk and even better than the conventional SIMOX wafers. The BESOI SOIAS preparation is a more complicated process due to the film thickness uniformity requirements for

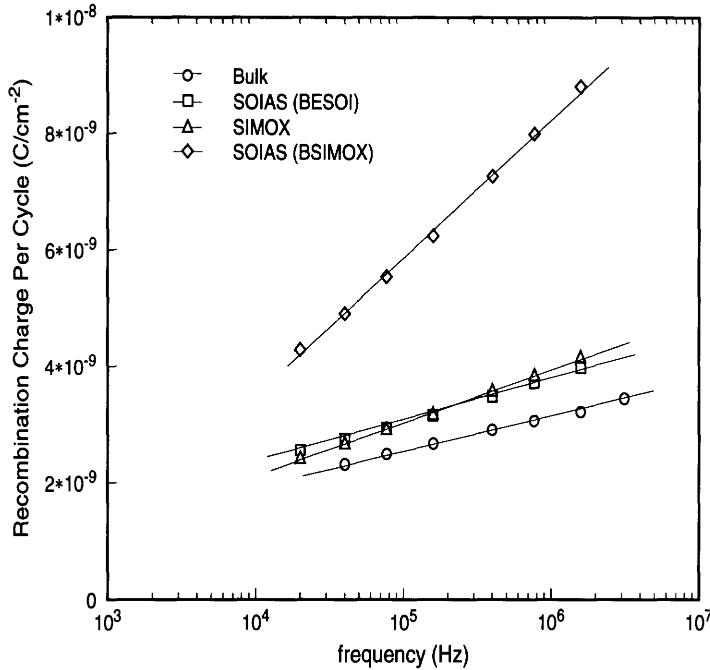


Figure 3-14: Recombination charge per cycle as a function of frequency for different substrate materials.

fully depleted SOI. The data for the bonded SIMOX SOIAS are very encouraging, although the overall material quality seem a bit worse than its counterparts, with touch polishing of the surface, the material can be improved. The bonded SIMOX approach is a simpler process with very good film thickness uniformity control, thus it is a viable technique enabling integration in the third dimension.

Substrate Type	Average Interface States (cm^{-2}-eV^{-1})
Bulk	$3.1X10^{10}$
SOIAS (BESOI)	$3.9X10^{10}$
SIMOX	$4.8X10^{10}$
SOIAS (BSIMOX)	$1.2X10^{11}$

Table 3.2: Average interface state density for different substrate materials.

Chapter 4

X-ray Lithography Integration in CMOS Process

Proximity x-ray lithography was chosen as the technology for gate patterning in the deep submicrometer region. X-ray lithography offers process latitude and robustness due to absorption without spurious scattering. The impact of photoelectron range and diffraction effects on the resolution limit of x-ray lithography were originally thought to be much more severe than is the case. It was assumed that the maximum photoelectron range defined the resolution limit of x-ray lithography [58]. Theoretical calculations of image degradation downstream of the mask based on a Kirchoff approximation and perfect spatial coherence led to the conclusion of very limited depth-of-focus and exposure latitude [59]. Both of these predictions were disproved by experimental results and subsequently by theory. Early *et al* has shown 30 nm lines in Au absorber can be faithfully replicated in PMMA resist at wavelengths of 0.83, 1.32, and 4.5 nm [60]. The explanation is that the Auger electrons released when an x-ray photon is absorbed put a strong peak in the point-spread function about 5nm wide, and that the effective range of photoelectrons is the relevant parameter, not the maximum range. Chu *et al* has shown that 50 nm lines and spaces can be printed with 2.3X variation in exposure dose at a gap three times larger than predicted by the earlier theoretical calculations [61]. Since proximity x-ray lithography shows no inherent limitations for patterning in the 0.1 μm and near sub-100 nm region, the

questions to be addressed are the applications of x-ray lithography and its feasibility for VLSI. The issues to be dealt with in this thesis specifically involve the integration of the existing x-ray lithography technology to our baseline CMOS. One of the main concerns for manufacturing is throughput, and even with a synchrotron, the use of PMMA is impractical. In recent years, resist engineers have developed a plethora of chemically amplified resists (CAR) for DUV which have also found their use in x-ray lithography. Paramount to fine lithography is faithful transfer of the pattern via dry etching. The fabrication of deep submicrometer bulk and SOIAS devices was a test vehicle for this technology.

4.1 Mix-and-Match Scheme

Since the resolution of x-ray lithography is not needed at all levels of the process, a mix-and-match scheme must be employed, in research as well as in manufacturing, which involves the matching of the optical, e-beam, and x-ray tools. To validate the matching of the tools, alignment and registration between the optical and x-ray levels must be demonstrated. The motivation for this work stems from the need to integrate x-ray lithography with a baseline CMOS process to fabricate 0.1 and sub-0.1 μm devices and circuits on bulk silicon and SOIAS. Currently, only the gate level is defined by x-ray lithography while the rest of the levels are patterned by optical projection. The x-ray mask is patterned using electron-beam lithography. One of the key challenges in this mixing of lithography tools is to reconcile or deal with the mismatch in the scaling, stepping, and distortion of these tools. In addition, for large chips and dense patterns at 0.1 μm, the e-beam time required to write even the gate level patterns on an x-ray mask can be prohibitively long (e.g., tens of hours on a research type of e-beam tool). In this work, a strategy of combining the three different lithography tools which best utilizes the strength of each, while resolving the problems of mismatch and long e-beam writing time was explored. Specifically, a G-line stepper, our VS-PL vector-scan e-beam, and our in-house x-ray lithography system were used. The key idea of this strategy is to pattern the coarse features and

e-beam field marks onto x-ray masks optically. Then, the e-beam tool writes only the fine features after aligning to the field marks on the x-ray mask, thus ensuring registration to the optical stepper generated features. The completed x-ray mask with coarse and fine features for the gate level is then used to pattern the device wafers. This mix-and-match scheme was demonstrated for 10x10 mm chips with four chips patterned onto one x-ray mask. The total pattern area is then 20x20 mm on a 31 mm diameter SiN_X membrane mesa-style x-ray mask. The subsequent x-ray exposure of device wafers was done with simultaneous alignment of all four dies.

4.1.1 Experimental Procedure

Figure 4-1 is a Fizeau interferogram of a SiN_X membrane, mesa-style x-ray mask. The membrane is flat to about 100 nm, less than one optical fringe. The mesa rim is flat to about one fringe. The non-flatness of the region outside the mesa, the Pyrex ring, is inconsequential as far as mask-sample gap is concerned. Figure 4-2 illustrates the overall process of mixing and matching the optical and e-beam tools. Since not all of the features at the gate level need the resolution of e-beam lithography, these features, along with the e-beam field marks are patterned optically onto the x-ray mask. By patterning the e-beam field marks optically, many distortion, scale, or stepping error in the optical stepper will be captured and duplicated by the e-beam registering to each field. The smaller the e-beam writing field, the better this technique will work. In our experiment, we chose to use 100x100 μm e-beam fields. The device and circuit patterns were placed such that they don't overlap with e-beam field marks . Figure 4-2a shows, as an example, the coarse features of the gate level of a NAND ring oscillator circuit, together with e-beam field marks. Only fine gates were then written by the e-beam, Figure 4-2b. The final combination of the coarse and fine features on the x-ray mask is shown in Figure 4-2c.

4.1.2 Coarse Features and E-beam Field Marks Transfer

Currently, our optical stepper is not configured to do exposures on x-ray masks.

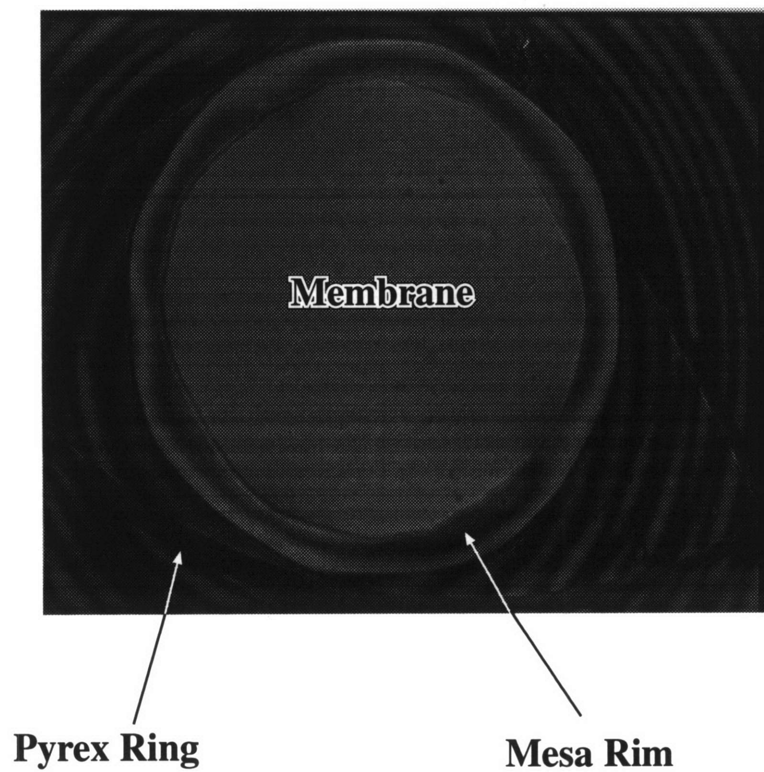
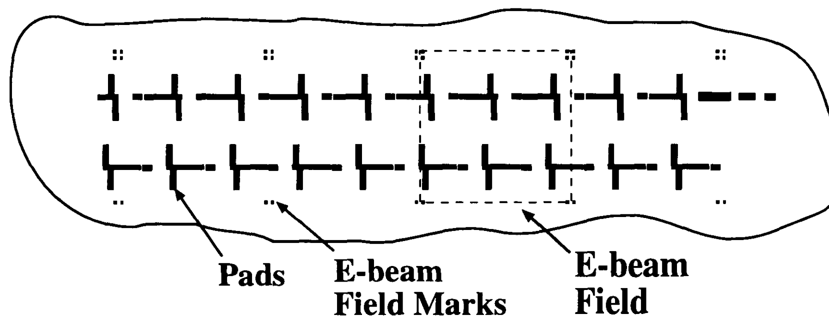
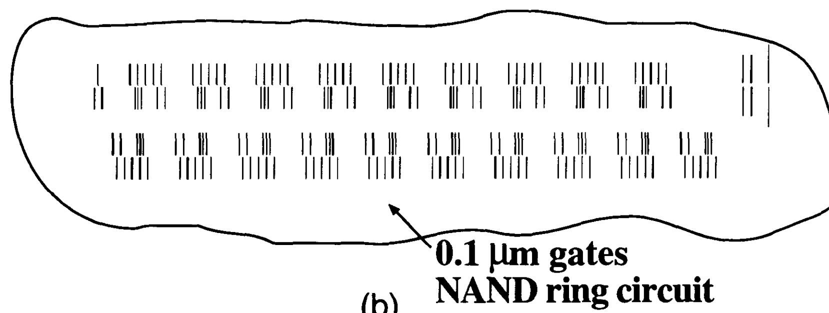


Figure 4-1: Fizeau interferogram showing an ultra-flat SiN_X membrane, mesa-style x-ray mask.





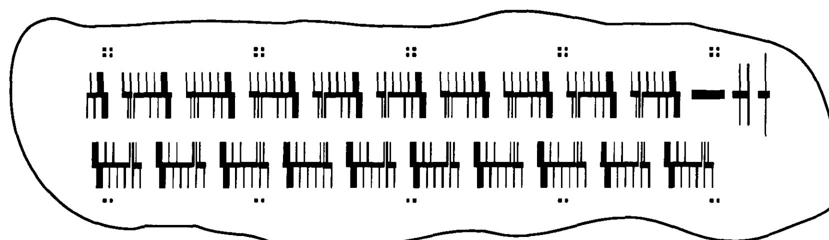
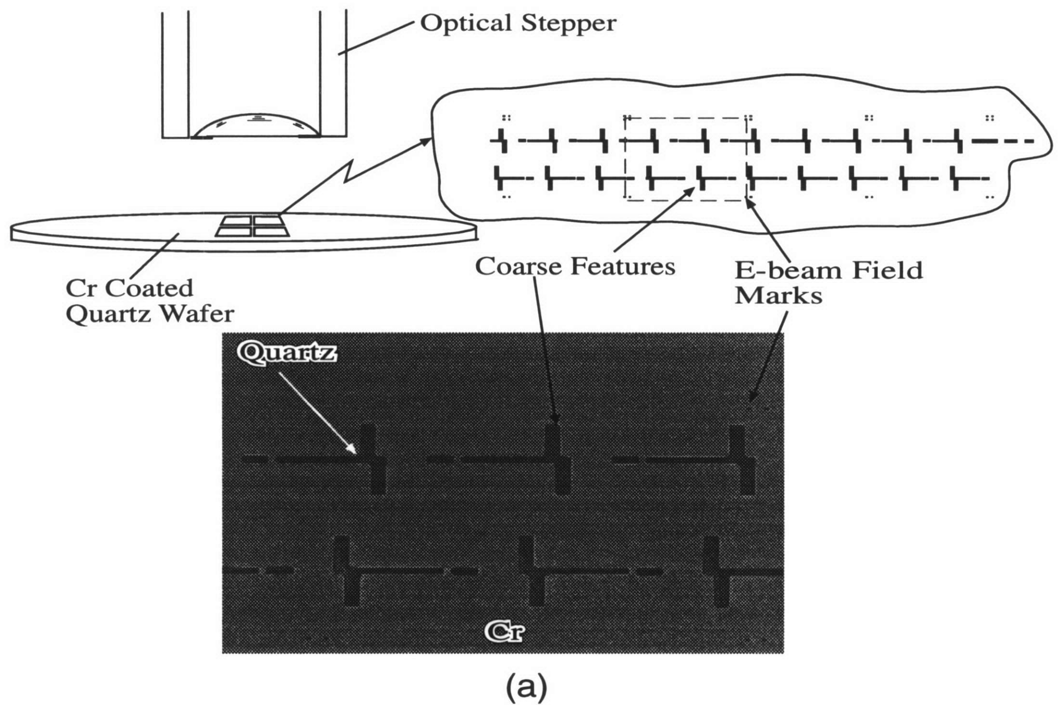


Figure 4-2: Process of combining e-beam and optical lithography to produce an x-ray mask. (a)Patterning of coarse features and e-beam field alignment marks using a stepper.(b)E-beam writing of fine features by aligning to the optically defined field marks. (c)The final pattern on the x-ray mask as a result of optical and e-beam lithographies.

Instead, we exposed the gate-level coarse patterns and e-beam field marks on 10 cm-diameter, chrome-coated quartz wafers, and then transferred these patterns onto the x-ray masks using DUV contact printing in 250 nm thick PMMA, followed by development and gold electroplating to a thickness of 80-100 nm. The process is depicted in Figure 4-3. An 10X dark-field emulsion mask of the coarse features and e-beam field marks was made for the g-line stepper. Four dies were then stepped down onto the chrome-coated quartz wafer and the chrome was wet etched. Figure 4-3b depicts the DUV pattern transfer onto the x-ray mask. In order to ensure faithful transfer of the pattern without introducing scaling errors, the gap between the quartz wafer and the membrane must be kept small, i.e. 1μm or less. Since the quartz wafer is not perfectly flat, and it rests on the x-ray mask mesa rim relying on gravity, the gap must be set by making the flexible membrane conform to the curvature of the quartz. This is accomplished by placing the x-ray mask on a fixture with holes drilled from underneath to allow a nitrogen jet to "blow" against the backside of the membrane. This system is not tightly sealed so that the pressure underneath the membrane does not build up. If the membrane goes into contact with the quartz wafer, a slight vacuum is applied to the backside to pull the membrane out of contact.

4.1.3 E-beam Patterning of Fine Features

The x-ray mask thus prepared can be tested for flatness, stress induced distortion, and alignment to previous optically defined levels prior to e-beam writing of the fine features. The PMMA can be stripped after the DUV exposure and gold electroplating, although we chose to keep the PMMA used in the DUV exposure for the e-beam exposure of the fine gates as well. Each of the four dies has a mark patterned at the left bottom corner to be used as an origin reference when initially setting up the e-beam coordinate system, see Figure 4-4. There are two sets of alignment marks designed for coarse alignment to the optical level; one set is for a clear-field x-ray mask, the other for a dark-field mask, Figure 4-4. These marks are located on the top and the right side of each die. The dark-field alignment marks were used for rotation



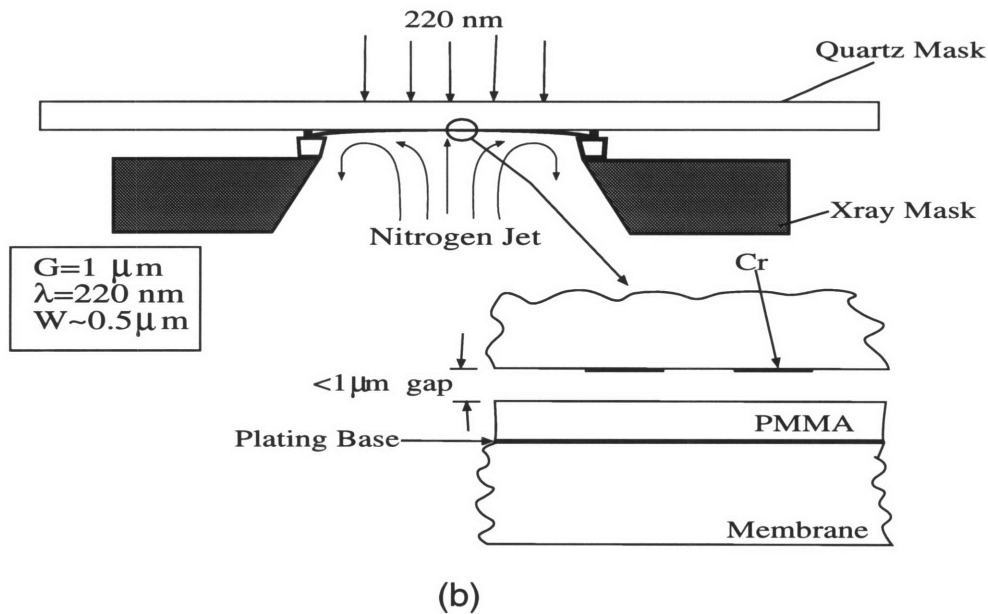


Figure 4-3: Quartz wafer patterning and DUV coarse feature transfer onto x-ray mask. (a)A 10X emulsion mask with coarse features and e-beam field marks was stepped down onto a chrome coated quartz wafer. The pattern was then wet etched into chrome. This quartz wafer then serves as the 1X mask for transferring the pattern onto the x-ray mask. (b)The DUV transfer process. The x-ray mask is placed on a fixture with holes drilled in the underside of the fixture. A jet of nitrogen is "blown" through the holes in the fixture forcing the x-ray mask membrane to conform to the quartz wafer.

alignment when e-beam writing a clear field mask; conversely, the clear field marks would be used when a dark field mask is desired. The CAD layout file is fractured into many subfiles, the left bottom e-beam field mark coordinates in these subfiles are referenced to the origin mark at the left bottom corner of each die. Once the e-beam registers to these field marks, two or more alignment iterations are made and the pattern in this particular e-beam field is written. The stage then steps to the subsequent fields performing the same alignment and writing in each field. Fields without any patterns are bypassed. This procedure was then repeated to write the other three dies. After e-beam exposure, the x-ray mask is then developed in 3:2 IPA:MIBK for 120 seconds at 21 °C. Gold is then plated up to a thickness of 200 nm in the fine features while the coarse features have a total thickness of 280-300 nm. A 200 nm thick gold absorber provides 10 dB attenuation at 1.3 nm wavelength [62]. Aluminum studs ($\sim 3 - 5$ μm in height), which served to set the mask-wafer gap, were then evaporated onto the mesa rim. Figure 4-5 shows the completed x-ray mask with both coarse and fine features plated up in Au. Figure 4-6 shows a close-up view of the patterned four dies (20x20 mm) on an x-ray mask with a 750X magnified view of an e-beam field in a NAND circuit. The e-beam registered very well to the optically patterned field marks as indicated by the continuity of the gate finger to the contact pads.

4.2 X-ray Lithography Exposure and Alignment Procedure

The gate level lithography was carried out in a custom built alignment and exposure system[63]. Figure 4-7(a) shows a schematic of the entire system. Since the patterns on the x-ray mask are already matched to the optical levels using the mix-and-match scheme described in [64], the alignment of the front-gate to the active regions was done in two steps. Since four dies must be aligned simultaneously, large crosses in each of the four dies were used to adjust for the gross rotation and x-y

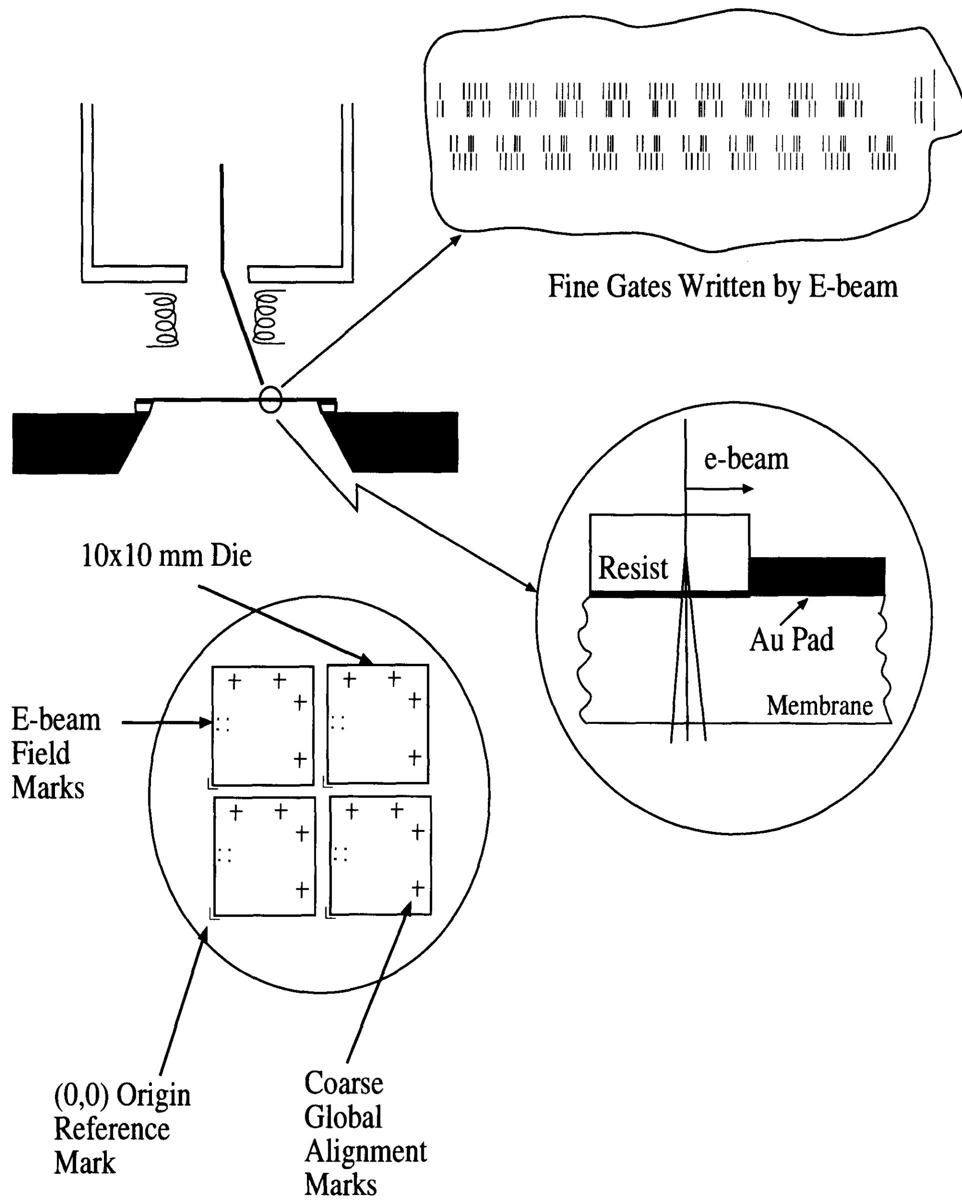


Figure 4-4: E-beam writing process for fine gate features.

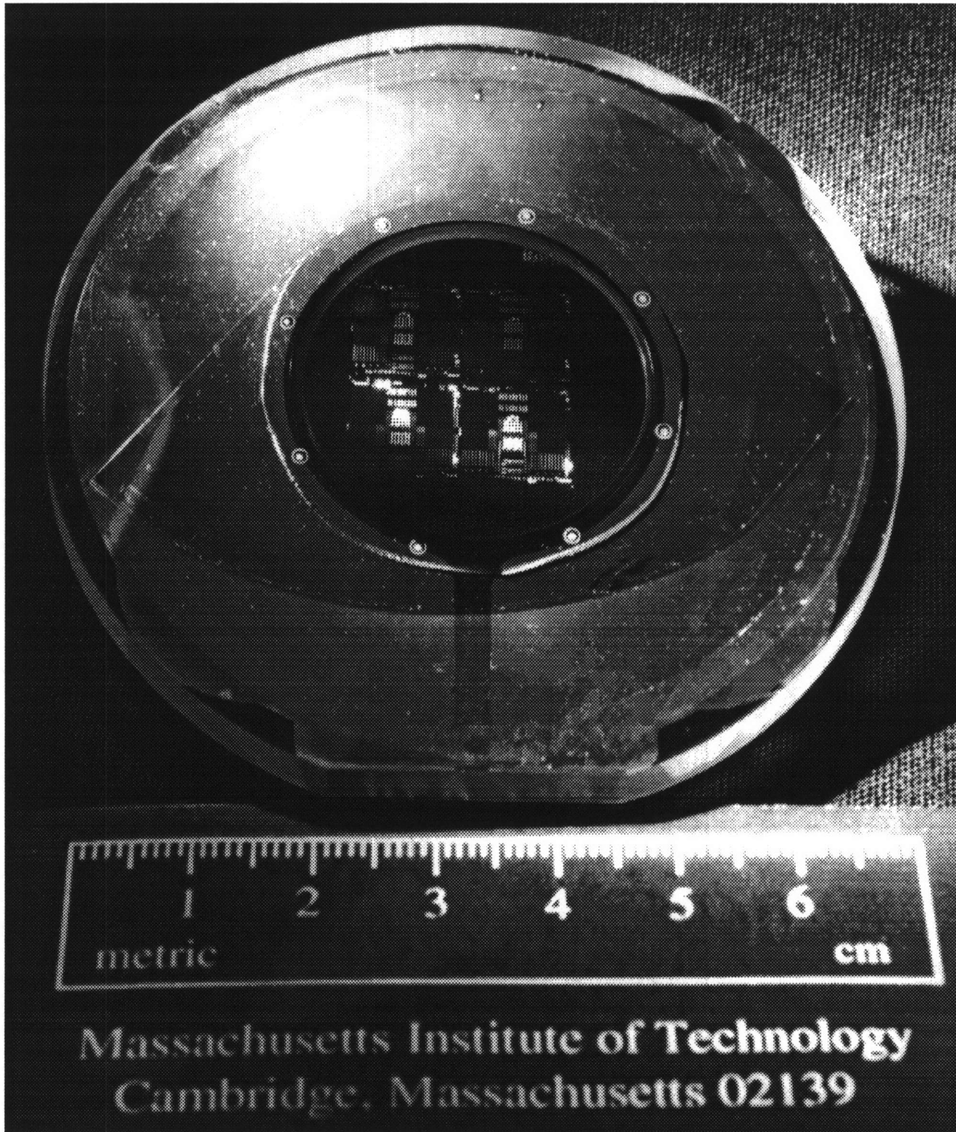


Figure 4-5: The completed x-ray mask with coarse features patterned by G-line stepper and fine featured patterned by e-beam writing.

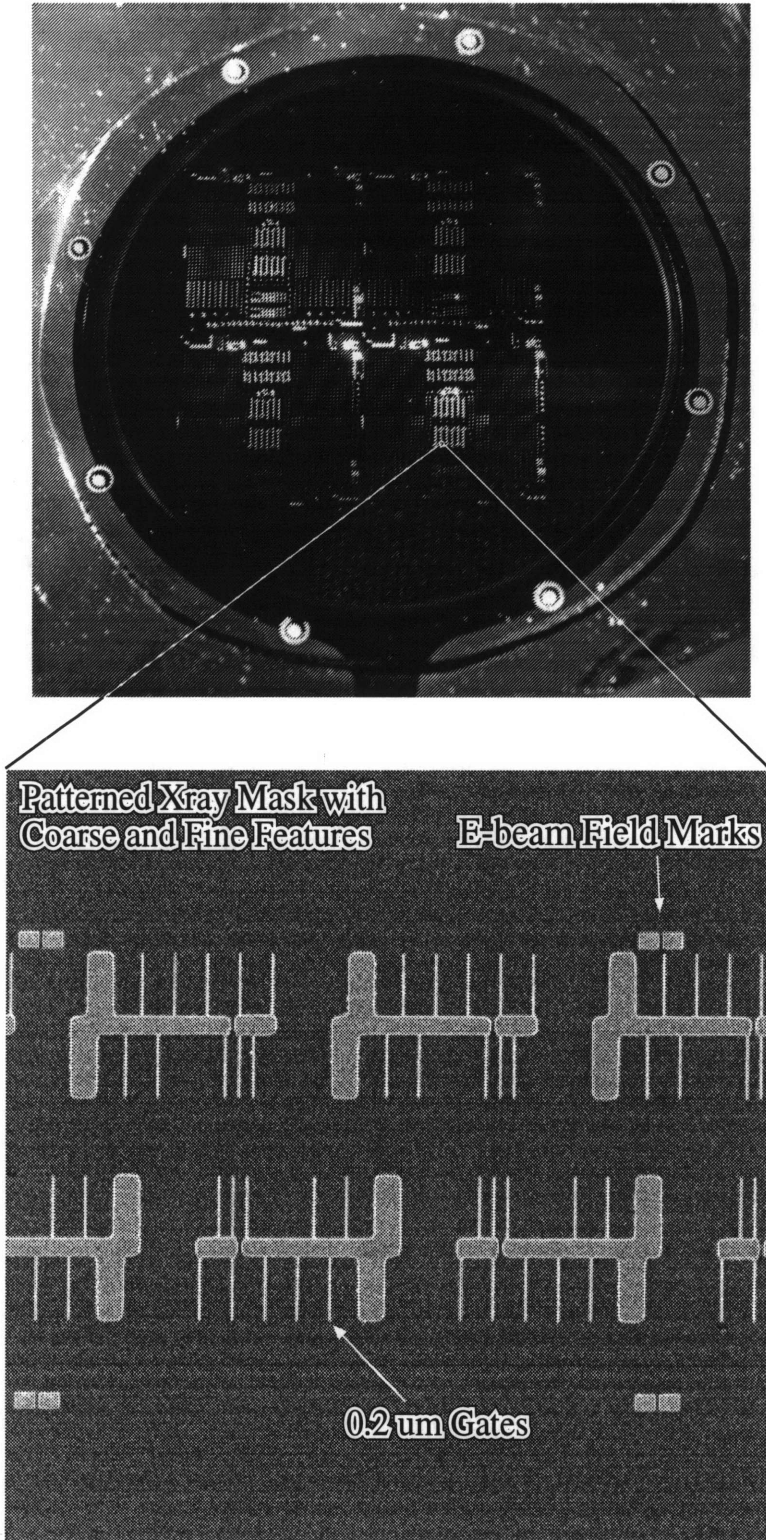
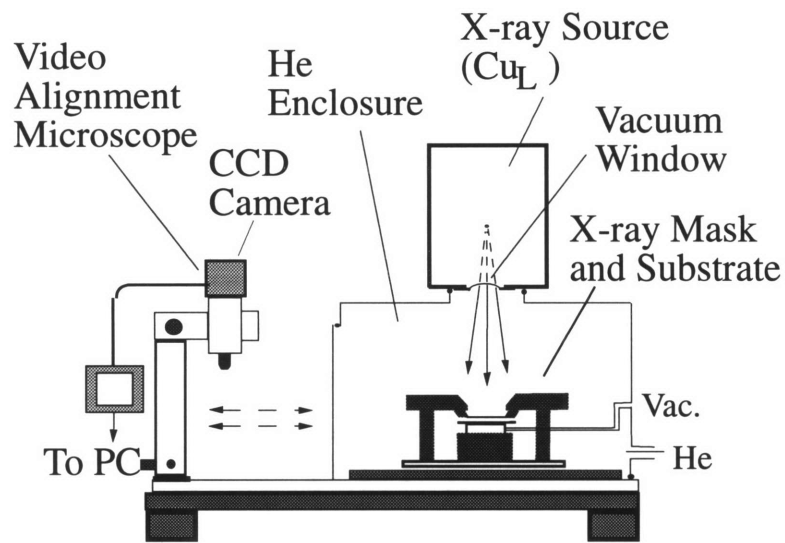
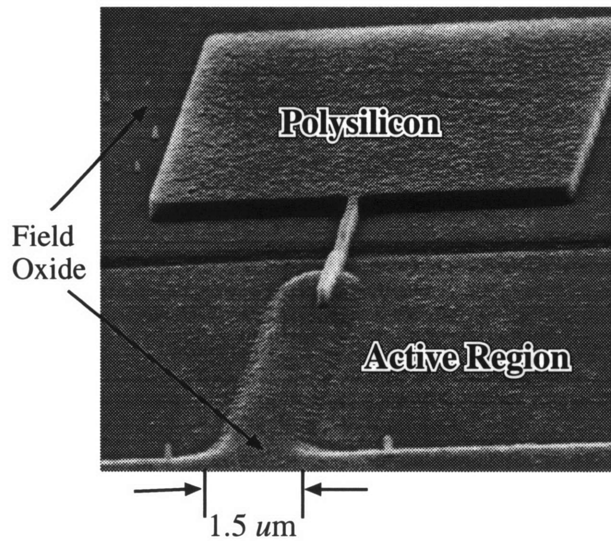


Figure 4-6: An enlarged view of the completed four die mask (20x20 mm), and e-beam field view of a NAND gate in a NAND ring oscillator. The light scatter (white dots) outside the membrane area are of no concern.



(a)



(b)

Figure 4-7: (a)Schematic of alignment and exposure system. The alignment was done outside of the helium enclosure using the CCD camera and the microscope. Once alignment was completed, the stage was rolled on ball bearings into the helium enclosure. During exposure, the Cu_L x-rays pass through a silicon nitride vacuum window and irradiate the mask and wafer. (b)Aligned polysilicon gate on a narrow-width device active area.

misalignments. Then using verniers with resolution of 100 nm, the fine alignment was done. The verniers were convenient because the measurements of the alignment before and after exposure were done with the same marks. Of course, a much more sensitive scheme can be employed, for example the grating-based interferometric technique called interferometric-broad-band imaging (IBBI) [65]. The 10 cm silicon wafer was held flat on a vacuum pin chuck which sits on a translation stage with x,y, and θ movements controlled by manual or piezo-actuated micrometers. This entire assembly was raised in the z direction towards an x-ray mask held in a tripod. The mask-leveling micrometers on the tripod control the leveling of the mask surface relative to the wafer. Alignment of the mask to the wafer was initially done at a 10-15 μm gap, then the gap was decreased to 5 μm for the final alignment before resting the mask on the wafer. The alignment latitude allowed between the optically-defined diffusion level and the x-ray-defined gate was 0.70 μm. Once alignment was completed, the substrate was raised towards the mask until the mask was lifted from the mask holder and rested on the wafer via the aluminum studs. No shift of the mask relative to the substrate was observed during this step. The alignment stage was then moved on ball bearings into a helium filled enclosure. To prevent motion of the mask relative to the substrate, a copper ring was placed on top of the x-ray mask to increase the frictional force of the studs on the wafer. The exposures were then carried out in helium atmosphere with oxygen concentration in the 200 ppm range, using a Cu_L (λ=1.3 nm) x-ray source. Figure 4-7(b) shows an aligned narrow-width device. In a manufacturing implementation, our simple means of setting the gap would be replaced by an automated scheme. Figure 4-8 shows the aligned NAND gates to the active regions after x-ray exposure. The worst case alignment of all four dies to the optically patterned diffusion level in the y-direction is 0.2 μm, and in the x-direction is 0.3 μm for 8 wafers. The best alignment was within 0.1μm which is the resolution limit of the verniers.

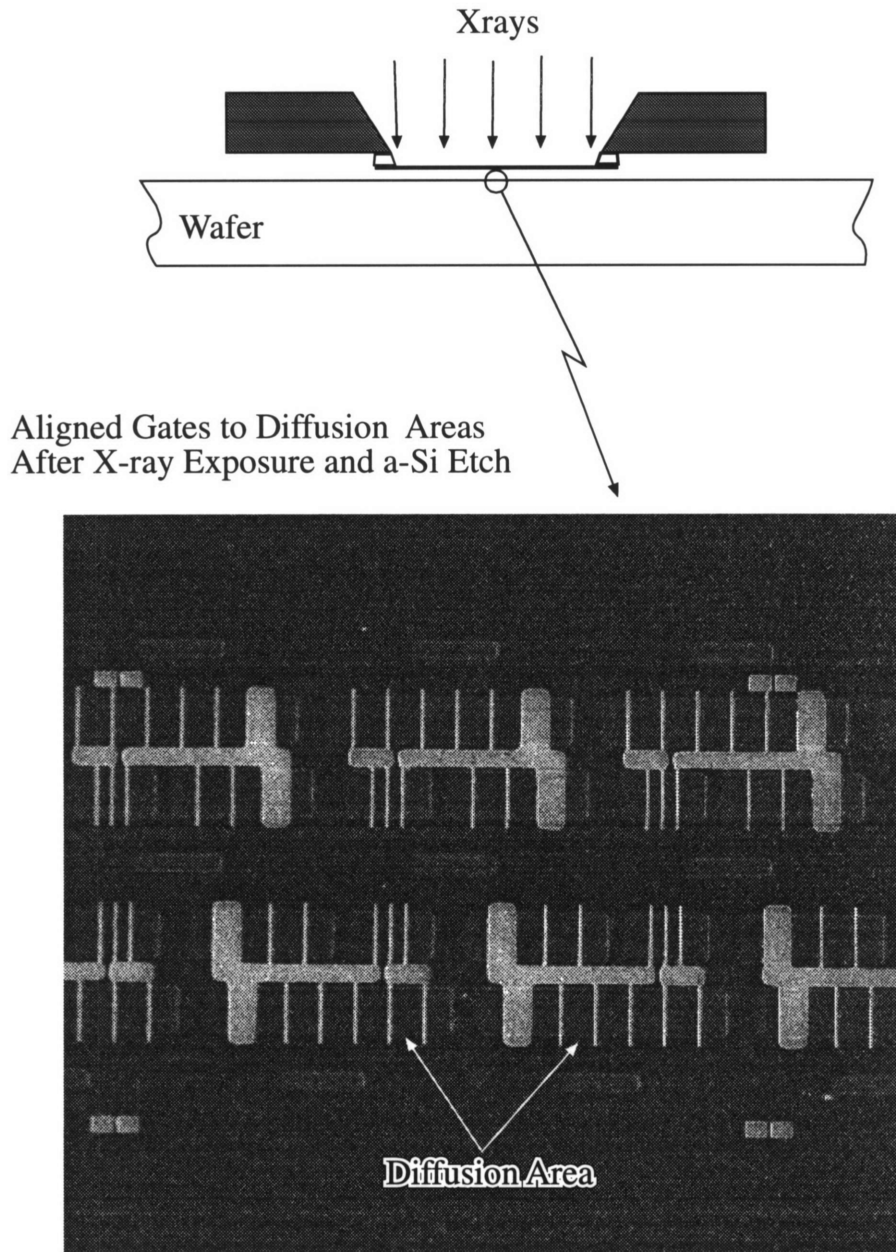


Figure 4-8: One e-beam field view of a NAND gate in a NAND ring oscillator after the amorphous silicon gate etch.

4.3 Gate Pattern Transfer Process

Chemically Amplified Resist (CAR) consist of three components, a base resin, a photo-acid generator, and a catalytically acid-activated cross-linking or dissolution agent. Upon exposure to photons or electrons, acids are generated. These acids act as catalysts to initiate the process of cross-linking or dissolution of the resist. Heat is usually added to speed up the process. Cross-linking occurs in negative-tone resist and dissolution occurs in positive-tone resist. Because the cross-linking (dissolution) process depends on the activation by photosensitive acids, control of the acid diffusion and generation is crucial for critical dimension control. In addition, depletion of acids may occur due to wafer surface or airborne contaminants which neutralize the acids. Consequently, resist residue and/or non-vertical profiles is the result. A typical CAR process includes, post-apply (PA) bake, resist exposure and post-exposure (PE) bake. The purpose of the PA bake is to dry out the solvents in the resist; the process latitude of this step is fairly large. The purpose of the PE bake is to induce the cross-linking or dissolution process with the photo-acid as a catalyst and to determine the extent of the 3-D random walk of the acid catalyst. The control of the time and temperature of this step is absolutely crucial. There have been many CARs developed in the past 10-15 years; two of which, SAL-601(Shipley) and ESCAP-X(IBM), we have found to give excellent results both in terms of resolution and process latitude for x-ray lithography. Results from several industry research groups are in agreement with our findings [66, 67]. SAL-601 was originally developed as an e-beam negative resist, and ESCAP was originally formulated as positive resist for DUV. Compared to Polymethyl Methacrylate (PMMA), the reduction in exposure time for SAL-601 is 5X and ESCAP-X is 7X, in our custom built x-ray system.

4.3.1 Negative Resist Process of SAL-601

SAL-601 is a novalak based resist similar to photoresist. Because of the stability of the aromatic rings that make up novalak, such resists have good etch resistance. Figure 4-9 shows the gate pattern transfer process flow for SAL-601. The purpose of

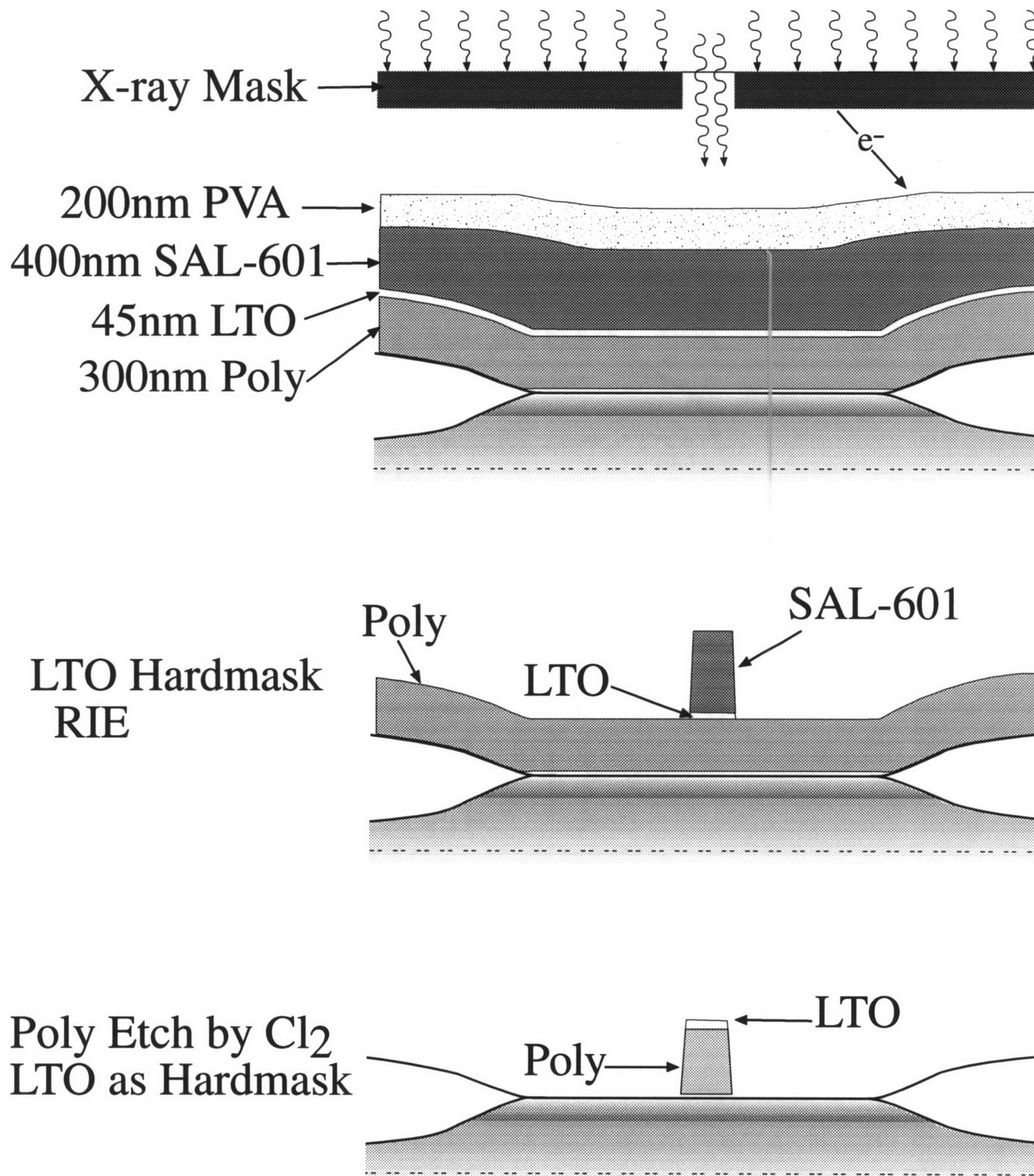


Figure 4-9: Process for x-ray exposure using the chemically amplified resist, SAL601, and PVA topcoat to prevent the formation of a "skin" layer. The definition of the polysilicon gate was accomplished in two steps:(1)the LTO hard mask was etched in CHF_3 RIE, (2)then the polysilicon was etched in pure Cl_2 plasma.

the low temperature oxide (LTO) was to serve as a hard mask for the dry etching of the polysilicon. Hexamethyl disilazane (HMDS) was used as an adhesion promoter. The HMDS was spun on at 4.5 krpm for 20 seconds and then baked at 110 ^{o}C for 2 minutes on a vacuum hot plate. The chemically amplified resist SAL601(Shipley) was then spun to 400 nm thickness and baked at 95 ^{o}C for 60 seconds on a vacuum hot plate. We observed the formation of a cross-linked "skin" layer on top of the SAL601, Figure 4-10(a). This may have been due to exposure by photoelectrons ejected from

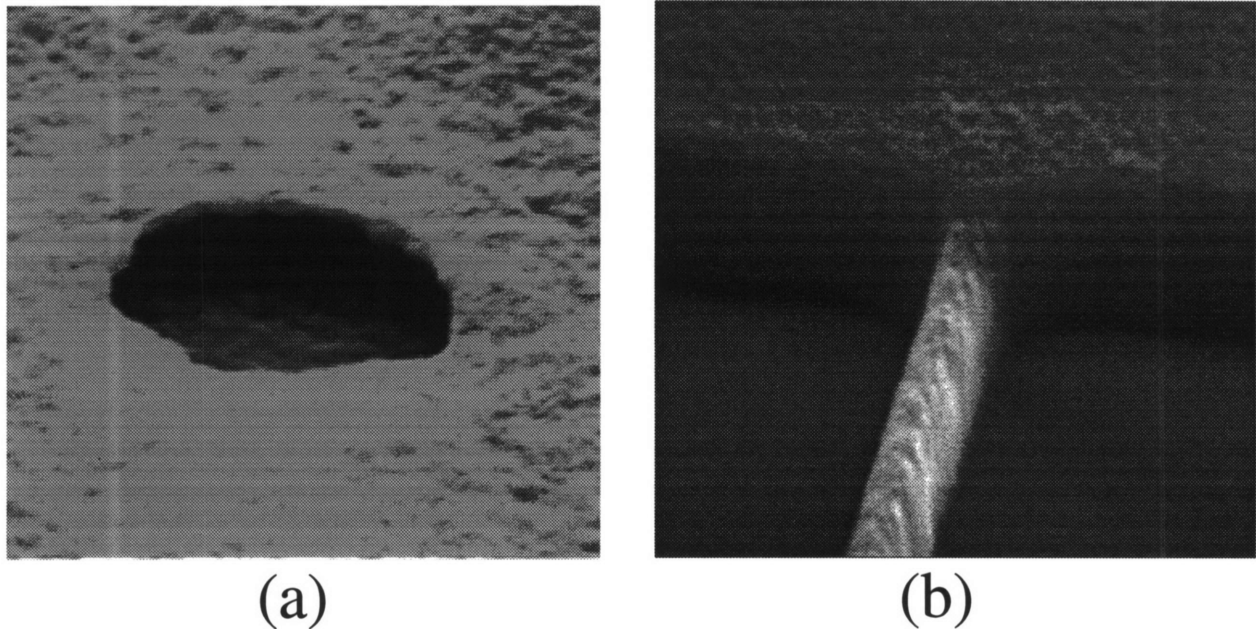


Figure 4-10: (a)Sample with crosslinked "skin" layer with the gate pattern. The gate pattern not visible in the SEM, is visible under the optical microscope. (b)Sample without the "skin" layer.

the gold absorber by the hard Bremsstrahlung and K line (0.15 nm) x-rays from our source. At small gaps (~3-5 μm), photoelectrons emitted from the gold absorber are not stopped by the helium ambient. To prevent the crosslinking of the resist top surface, we spun on 200-250 nm of water soluble polyvinyl alcohol(PVA). The PVA was a 6 wt% solution of Evanol 51-05 crystals (obtained from Dupont Chemicals) in high purity water (DI). PVA was chosen because it is immiscible with the SAL601. Figure 4-10(b) shows the result of the gate pattern exposed with a PVA top coat. After exposure, the PVA was rinsed off with DI water. The post-exposure bake (PEB)

of the SAL601 was at 110 °C for 40 seconds on a vacuum hot-plate. The PEB time and temperature were optimized for linewidths of approximately 0.1 μm. After the PEB, the wafer was quenched on an aluminum plate to room temperature. Resist development was done in a 1:1 solution of MF312:DI (Shipley developer). After development, a 3 minutes bake on a vacuum hot-plate at 110 °C was sufficient to harden the resist, which acted as the mask for the reactive ion etching (RIE) of LTO. The LTO was etched in 10 mTorr CHF_3 gas at 120 watts and 600 V DC bias in the NSL RIE. Since metal contamination, especially gold (NSL is a potential source of gold contamination), was a concern, "piranha" and RCA cleans were performed prior to the next etching step, which was done using a plasma etcher in our CMOS facility. The resist was stripped during the cleaning process and the LTO was used as a hard mask for the polysilicon gate etch which was done in 20 mTorr of pure Cl_2 at 100 watts [68]. Figure 4-11(a) shows a gate pattern stack of SAL601 on top of LTO after etching in CHF_3. No degradation of the resist is apparent. Figure 4-11(b) shows a pattern transferred into the LTO hardmask. Figure 4-11(c) shows a final gate pattern in polysilicon. It was faithfully transferred with approximately 100:1 selectivity of polysilicon to thermal oxide. The step coverage over 180 nm of field oxide was also excellent, as no detectable breaks in the gate fingers occurred.

4.3.2 Positive Resist Process of ESCAP-X

ESCAP-X, a positive resist developed at IBM [69], is not a novalak based resist. Figure 4-12 depicts the process for this resist. Post-apply bake was at 150 °C for 90 seconds, and post-exposure bake was at 140 °C for 45 seconds. Development was done in MF321(Shipley), which is 0.21N TMAH. Since this resist was very much immune to contaminants, no overcoat layer was needed. Exposure by photoelectrons from the x-ray mask was not a concern for positive resist. Since the photoelectrons only penetrates $\sim$ 20-30 nm of the resist, the resist would lose 20-30 nm in thickness during development. With these process conditions, the ESCAP resist yielded 0.15 μm gates with very smooth sidewalls as seen in Figure 4-12. The ESCAP resist was used as the etch mask for the 60 nm thick Si_3N_4 hard-mask which was etched in CHF_3 by the

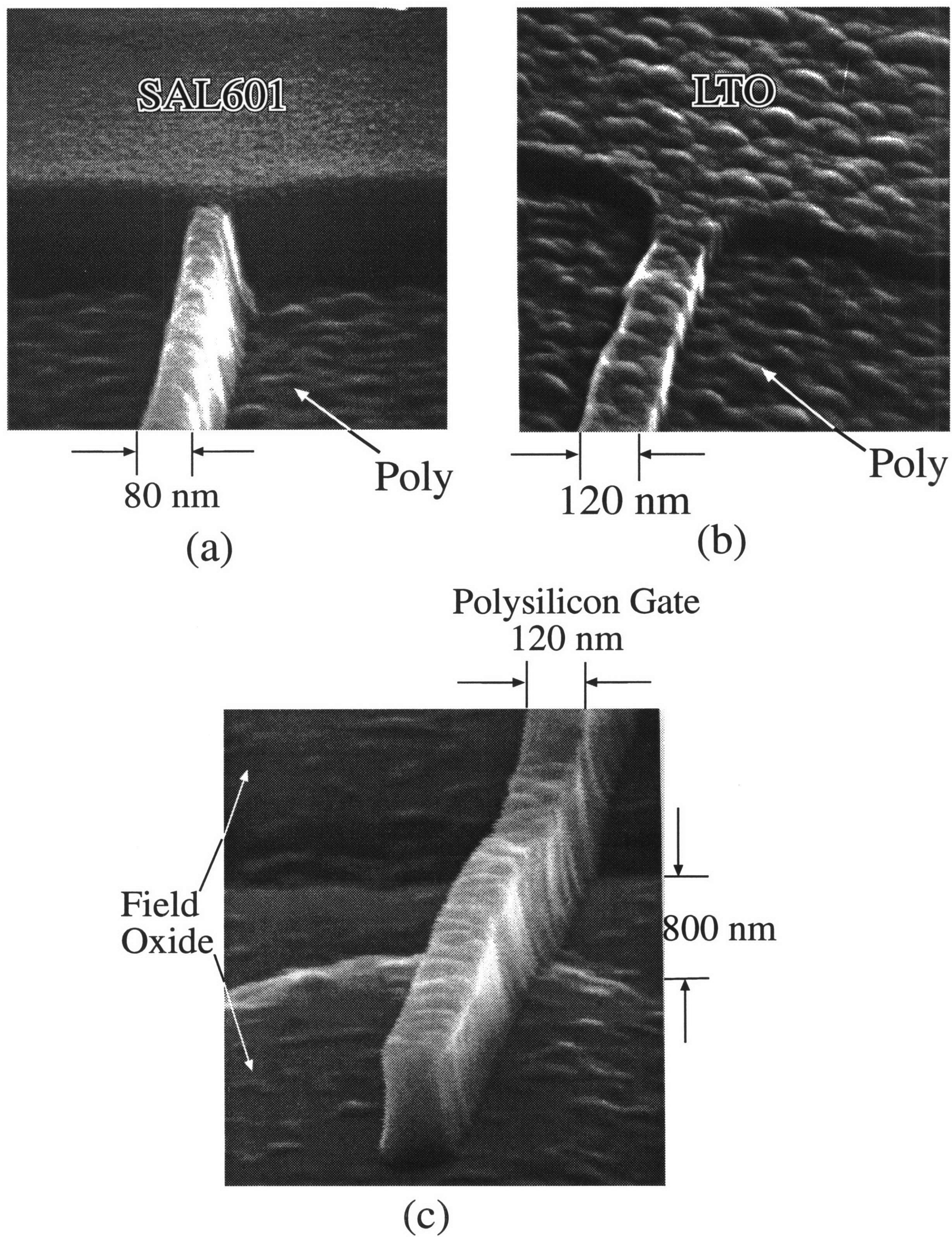


Figure 4-11: Dry etching results:(a)SAL601 resist on top of LTO after CHF_3 RIE for a 80 nm gate, (b)120 nm gate pattern transferred into the LTO with the SAL601 stripped, (c)polysilicon gate after Cl_2 plasma etching using LTO as hardmask.

NSL RIE at 10 mTorr pressure, 550 V DC bias, and 75 watts. The resist was stripped in "piranha" and the amorphous silicon was etched in pure Cl_2 gas at 100 watts in a plasma etcher using the nitride as the etch mask. Since the 400 nm of ESCAP resist was not able to withstand the pure Cl_2 etch, the hard mask was needed [70]. The ESCAP-X process is simpler due to its immunity to contaminants.

4.4 Device Results

4.4.1 Bulk Devices

Figure 4-13 shows the device cross section. The gate oxide thickness was approximately 4.5 nm and the polysilicon gate thickness 300 nm. The shallow source/drain extension junctions were 40-50 nm in depth, and the deeper source/drain junctions (~120 nm) provide low sheet resistance and easier cobalt silicide formation. The nitride spacers were formed to prevent the encroachment of the deep source/drain junction into the channel region, and to allow the formation of a self-aligned silicide on the gate, source, and drain. The active areas were defined by LOCOS isolation. The Super Steep Retrograde channel doping profiles were obtained by using shallow indium (100 KeV, 5x10^{12} cm^{-2}) and deep boron (50 KeV, 5x10^{12} cm^{-2}) implants. The polysilicon gate patterns were defined by x-ray lithography and anisotropic dry etching, as discussed in the previous section. The source/drain shallow junction extensions were formed using a low energy arsenic implant (9x10^{14} cm^{-2}, 10 KeV) with indium preamorphization and counter-doping (10^{14} cm^{-2}, 40 KeV). The deep source/drain junctions were formed using arsenic implantation at 5x$10^{15}$$cm^{-2}$ and 20 KeV. Cobalt silicide was then formed on the source/drain and the gate using a novel $CoSi_2$ technology [71]. Figure 4-14(a) shows the I-V characteristics of a sub-0.1 μm nMOSFET with L_{eff} of 85 nm and gate width of 50 μm. In this set of devices, the gate width ranged from 0.8 μm to 50 μm. The L_{eff} was extracted electrically using the method of Terada and Muta [72]. The device in Figure 4-14 has a current drive of 0.74 mA/μm and a saturated transconductance of 500 mS/mm at 2.0 V power supply

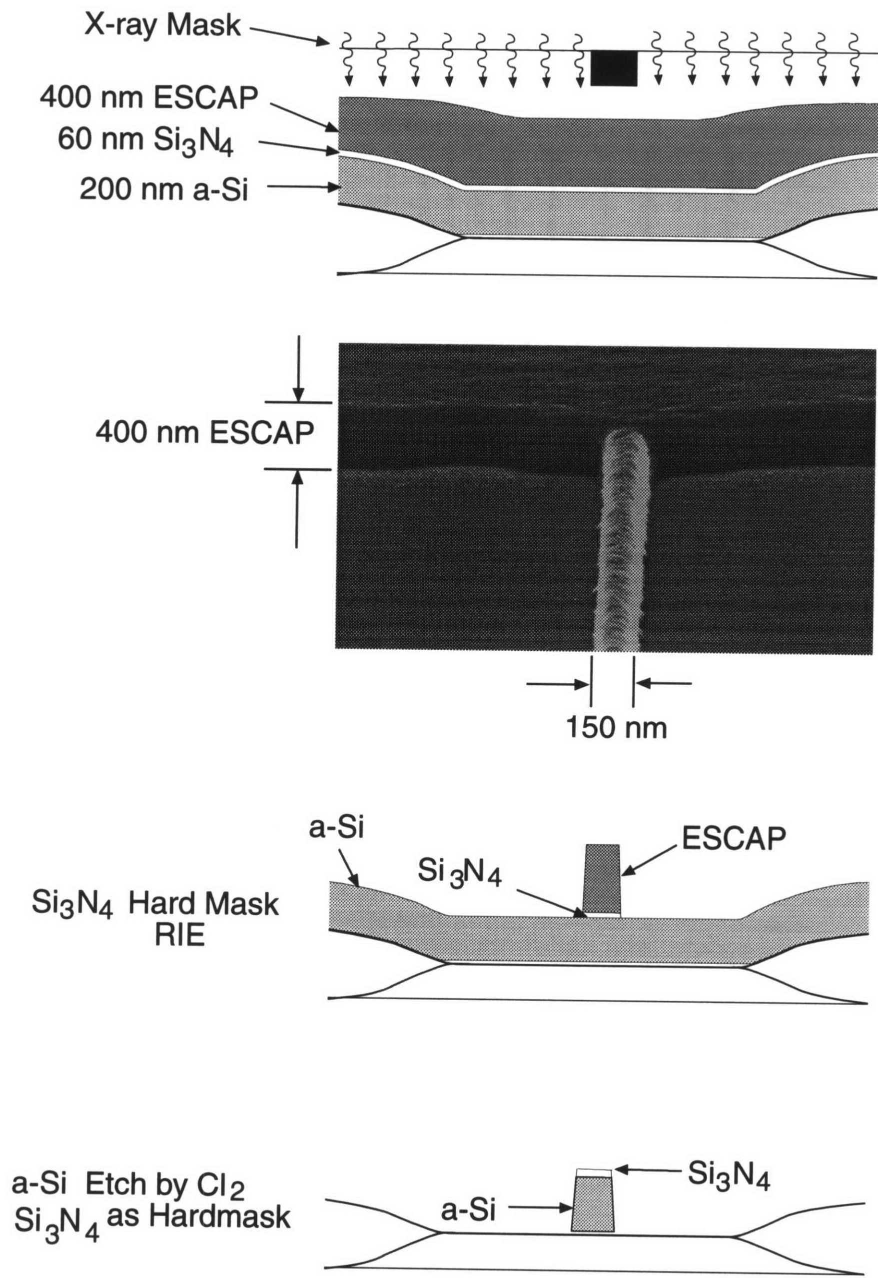


Figure 4-12: The gate patterning processing using ESCAP-X resist with nitride hard mask.

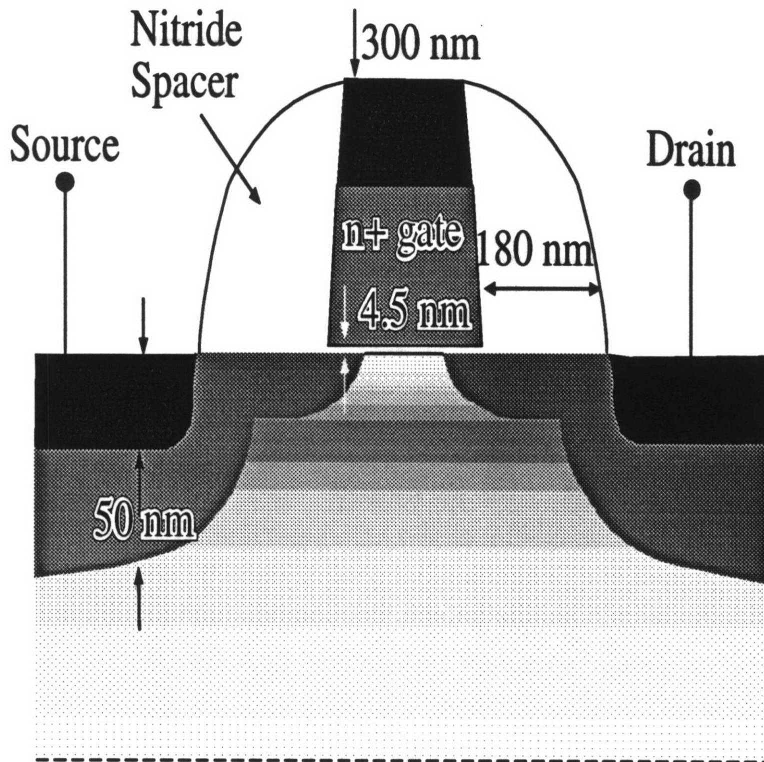
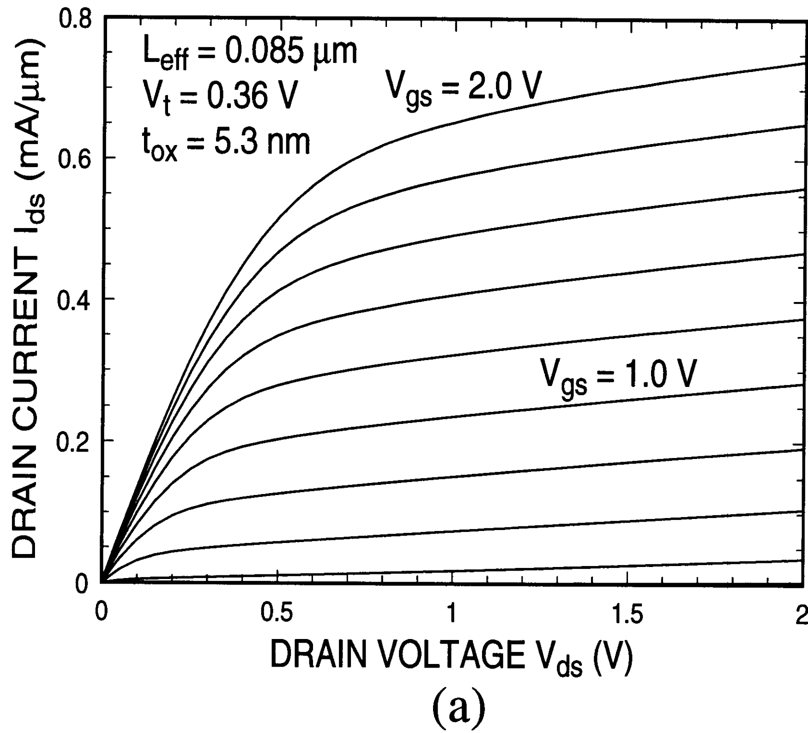
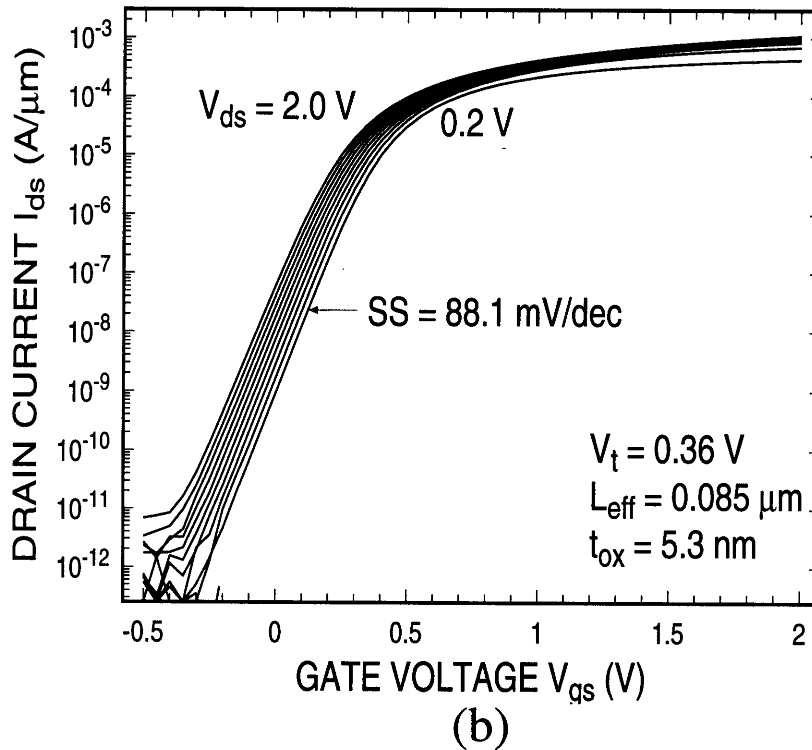


Figure 4-13: Schematic cross section of 0.1μm channel length nMOSFET. The gate oxide thickness was approximately 4.5 nm and the polysilicon thickness was 300 nm. The shading in the channel region depicts the retrograde doping. The source/drain extension and deep junction depths are ~40-50 nm and ~ 120 nm, respectively. The 180 nm thick nitride spacers prevented the encroachment of the deep source/drain junctions into the channel region, and ~50 nm of cobalt silicide was formed on the gate and source/drain.

voltage, with a threshold voltage of 0.36 V and a sub-threshold slope of 88.1 mV/dec. The sharp turn-on in the linear region indicates that the parasitic resistance is very low through the choice of source/drain implant profile and the cobalt silicide process. Since soft x-rays were used and a subsequent 1000 °C RTA activation of the source/drain dopants followed the gate patterning step, residual damage of the device was unlikely. Also, very low off-current is achieved, as indicated by the subthreshold characteristics in Figure 4-14(b). The shifting of the curves in Figure 4-14(b) as a function of Vds indicates that this device has some amount of Drain Induced Barrier Lowering (DIBL). A more complete analysis of the device results is given in ref. [73].



(a)

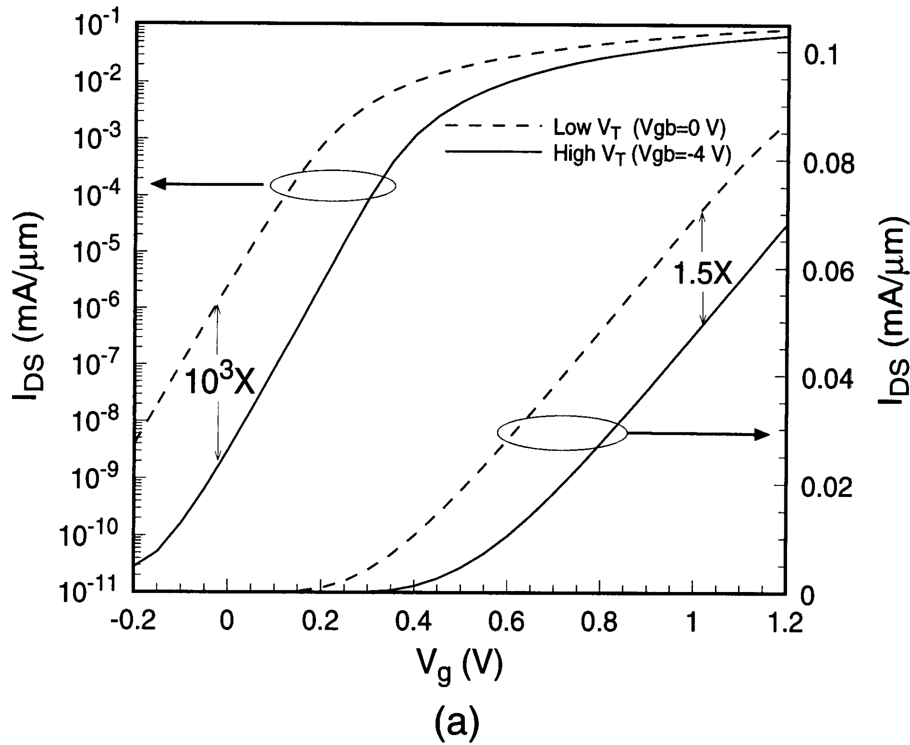


(b)

Figure 4-14: Device output characteristics:(a)I-V characteristics for L_{eff}=85 nm gate length device with threshold voltage of 0.36 V, and gate voltage stepped from 0 to 2 V in 0.2 V increments. (b)Subthreshold characteristics of the same device.

4.4.2 SOIAS Devices

The detailed description of the SOIAS device fabrication will be given in the next chapter. Only representative short channel SOIAS device characteristics is shown here. Figure 4-15 shows the subthreshold characteristics of a back-gated device. This device has 6.5 nm effective gate oxide, 38 nm silicon thickness, back-gate oxide of 100 nm, and an effective channel length of 0.2 μm. This is an example of how the dynamic threshold scheme can work. The quiescent V_T for this device was designed to be low ($\sim$200 mV). Therefore, with a -4 V bias on the back-gate which raises the V_T to approximately 450 mV, the off current can be decreased three decades when the device (circuit) is idling. On the other hand, the device can be designed with a high quiescent V_T. Therefore, when the circuit is active, with a +4 V bias on the back-gate, the drive current at VDS=1 V can be increased 1.5X. Figure 4-15(b) show the I-V characteristics of the device in Figure 4-15(a) with back-gate bias at -4 and 0 V. The PMOS results are similar. Therefore, these devices can be "turned off" hard in order reduce static leakage power, and can be "turn on" hard when high performance is needed, thus, meeting the opposing requirements of high performance and low power in the same device.



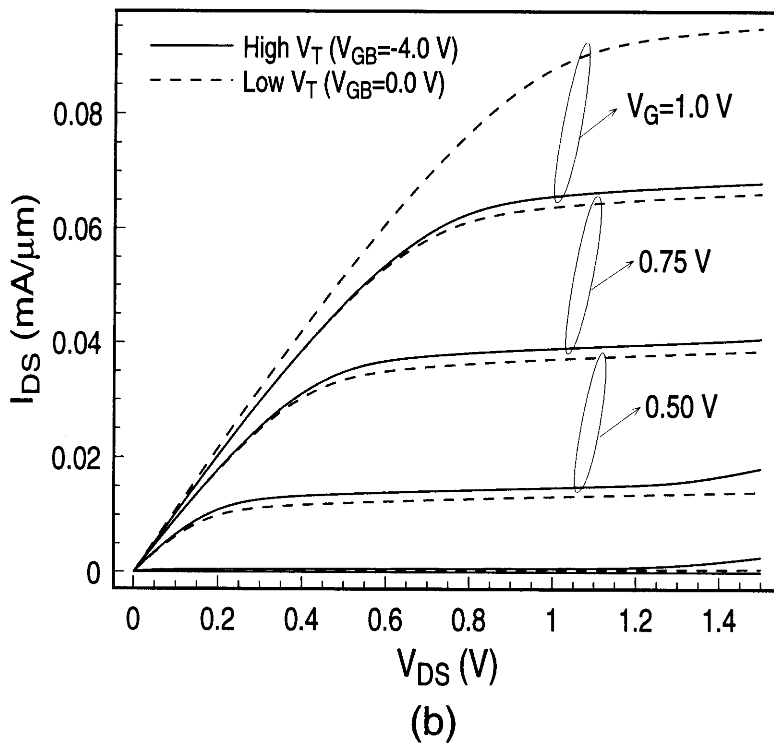


Figure 4-15: (a) Subthreshold characteristics of a 0.2μm top-gate device showing three decade change in the off current with 4 V back-gate bias in idling mode with low quiescent V_T. With high quiescent V_T, 1.5x times drive current change at VDS of 1.0 V can be observed. (b) I-V characteristics of the same device.

Chapter 5

SOIAS CMOS Devices

5.1 Fabrication

The starting material for CMOS device fabrication is the SOIAS substrates prepared by the method described in Chapter 2. The SOIAS substrate is a multi-layered blanket film stack consisting of the silicon wafer, oxide, intrinsic polysilicon, back-gate oxide, and silicon film. Figure 5-1 illustrates the final device schematic. The device fabrication on SOIAS substrate follows the conventional CMOS SOI self-aligned process with two additional steps. The first step of the process is the patterning of global alignment marks in the upper silicon film. These global marks allow the registration of the active-region patterns to the implanted back-gates since the masking and implantation of the back-gates do not leave any visible marks on the wafer. The back-gates are formed first by ion implantation through the silicon film in two masking steps using a 1 μm thick photo resist as an implant mask. This results in islands of p and n polysilicon, isolated by intrinsic poly after thermal anneal, Figure 5-2. Using the same type of doping in the back-gate poly and silicon film resulted in near-zero flatband voltage at the back-gate. By properly tailoring the energy and dose of the implant, the back-gate and the V_T-adjust implants for setting quiescent V_T value (i.e V_T at zero back-gate bias) can be done in one step. The peak of the back-gate implant is placed deep in the back-gate poly, and the leading edge of the implant is used to dope the silicon film. Figure 5-3a and b shows examples of Suprem3

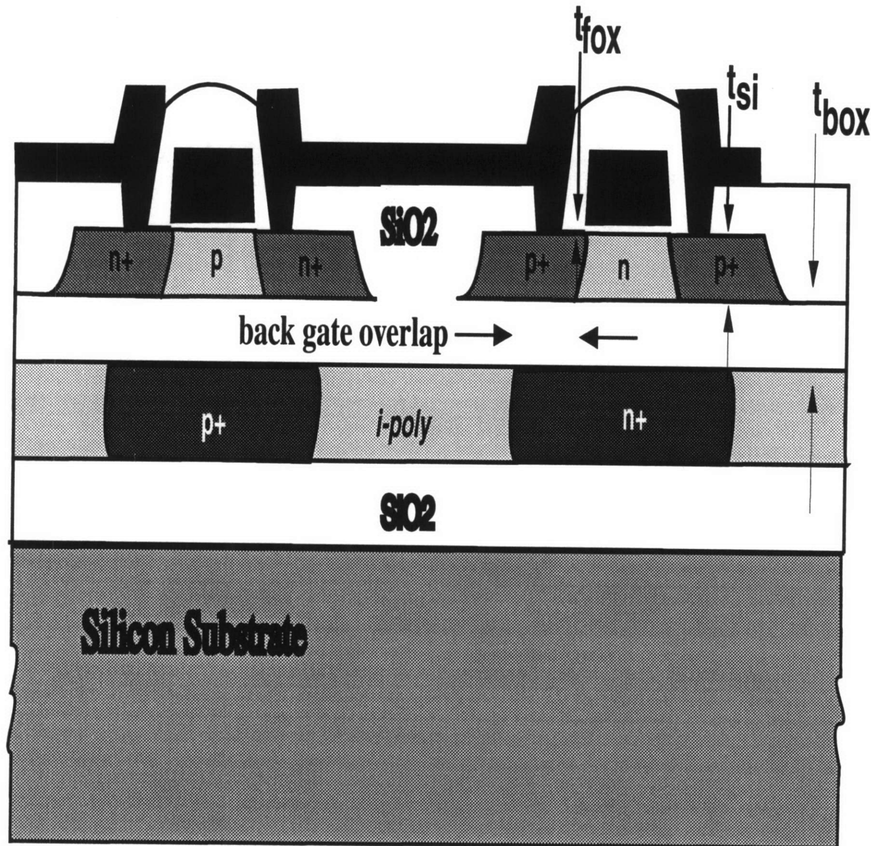


Figure 5-1: Device schematic.

simulations of the as-implanted and final boron and phosphorus concentrations in the silicon and back-gate. Typical sheet resistance of the back-gate poly is in the range of 1-5 KΩ/square range for the dopant concentrations shown. The back-gate can never be the same size as the front-gate due to the lateral diffusion which occurs during the subsequent high temperature steps after the back-gate formation. The source/drain to back-gate overlap is not negligible in this case. Therefore in order to reduce excessive parasitic overlap capacitance, the back-gate oxide has to be made relatively thick (80-100nm). The front-gate device is then built as in a conventional SOI CMOS process. After the implants, LOCOS oxidation was carried out at 950 ^{o}C for the isolation. The gate stack was then deposited and gate definition was done both optically and by x-ray lithography, as described in the previous chapter. The back-gates were contacted from the top, with the contacts cut through the field oxide and the back-gate oxide. Figure 5-4 shows the top and cross sectional views for the

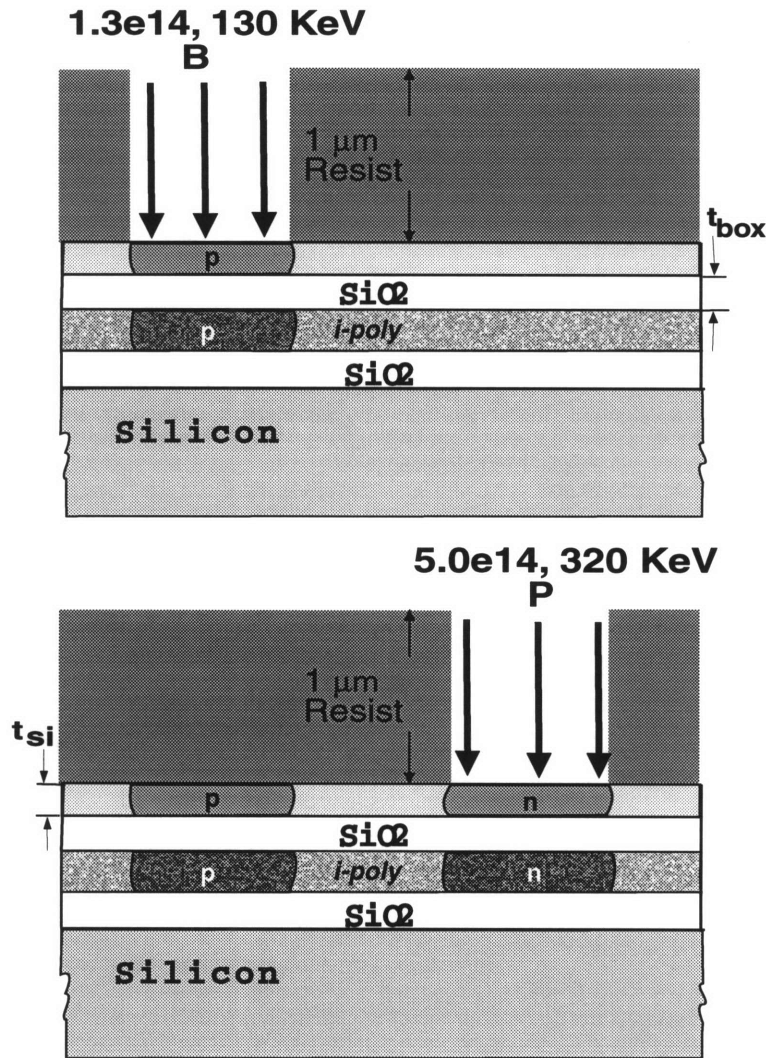
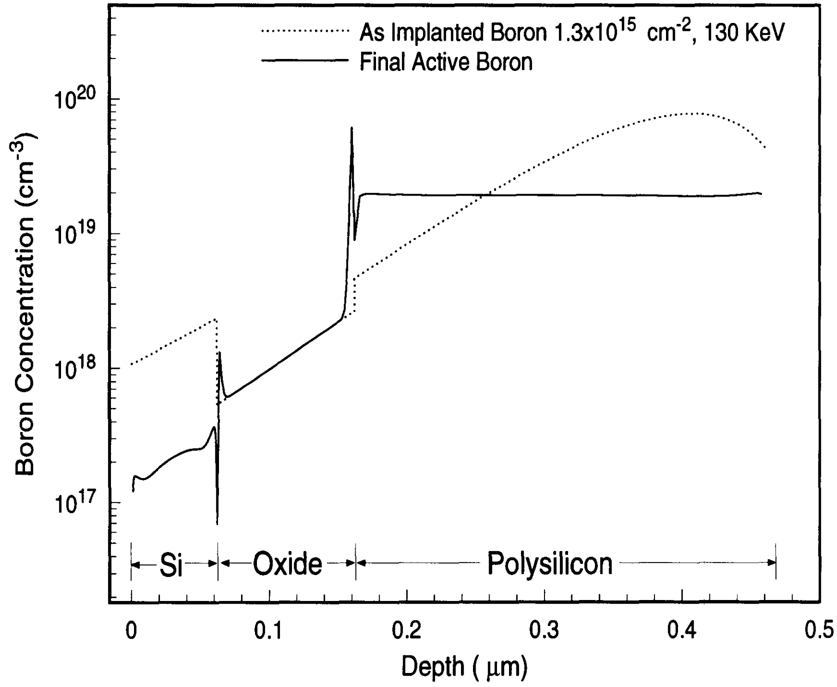


Figure 5-2: Two step masking and implant process for the formation of n and p back-gates.

back-gate contact cuts. Currently, our CMOS process has only one level of metal. Figure 5-5 is a SEM micrograph of the SOIAS device. The coupling between the front and back-gates depends on the ratio of the critical film thicknesses: front-gate oxide thickness (t_{fox}), silicon film thickness (t_{si}), and back-gate oxide thickness (t_{box}). We have chosen to demonstrated SOIAS with effective 9 nm t_{fox}, 40 nm t_{si}, and 100 nm t_{box} nominal design parameters which should yield 200 mV change in V_T per 3 V on the back-gate. A complete process flow detailing all of the steps can be found in Appendix A.



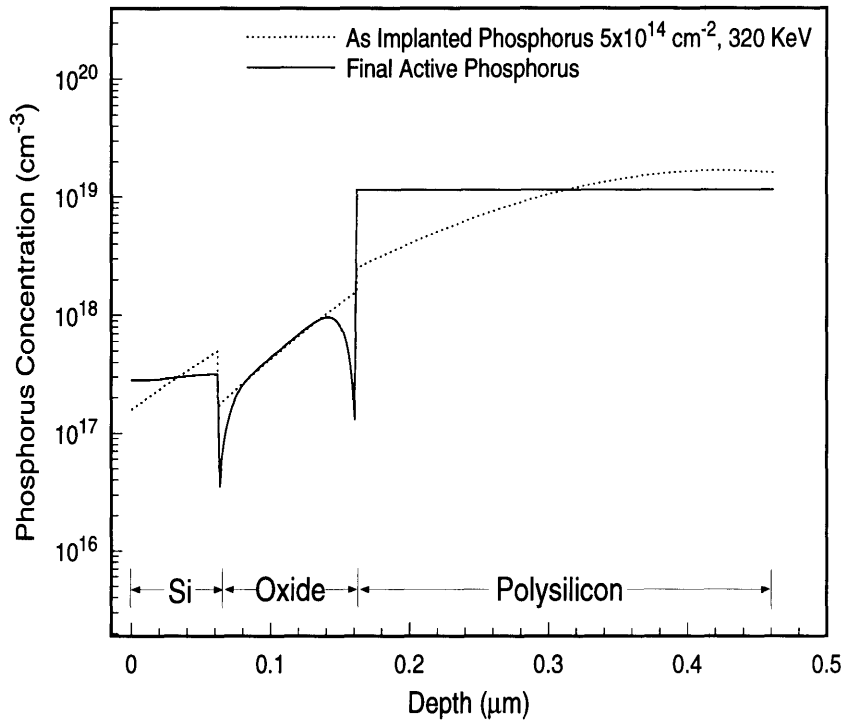
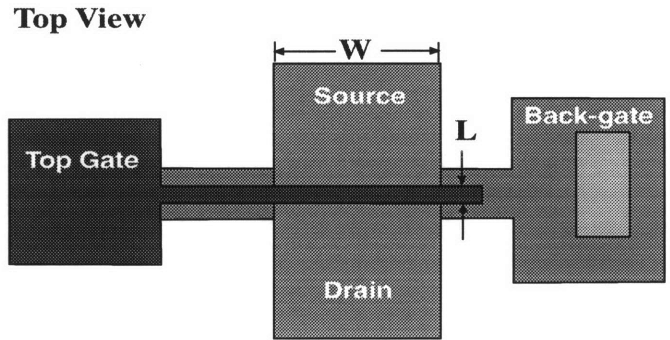


Figure 5-3: Suprem3 simulated implant profiles for NMOS and PMOS devices and their back-gates. The final active dopant concentration is after all the thermals cycles in SOIAS CMOS process.

Top View



Cross Setional View Along the Width of Device

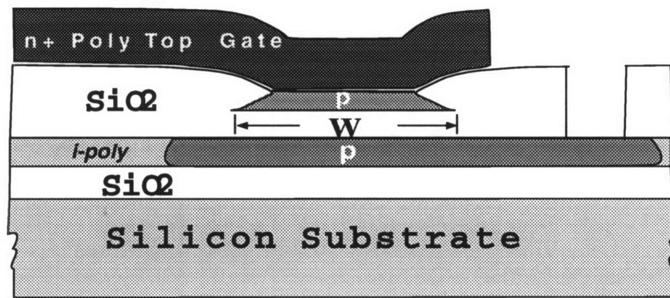


Figure 5-4: Schematic of back-gate contact cuts.

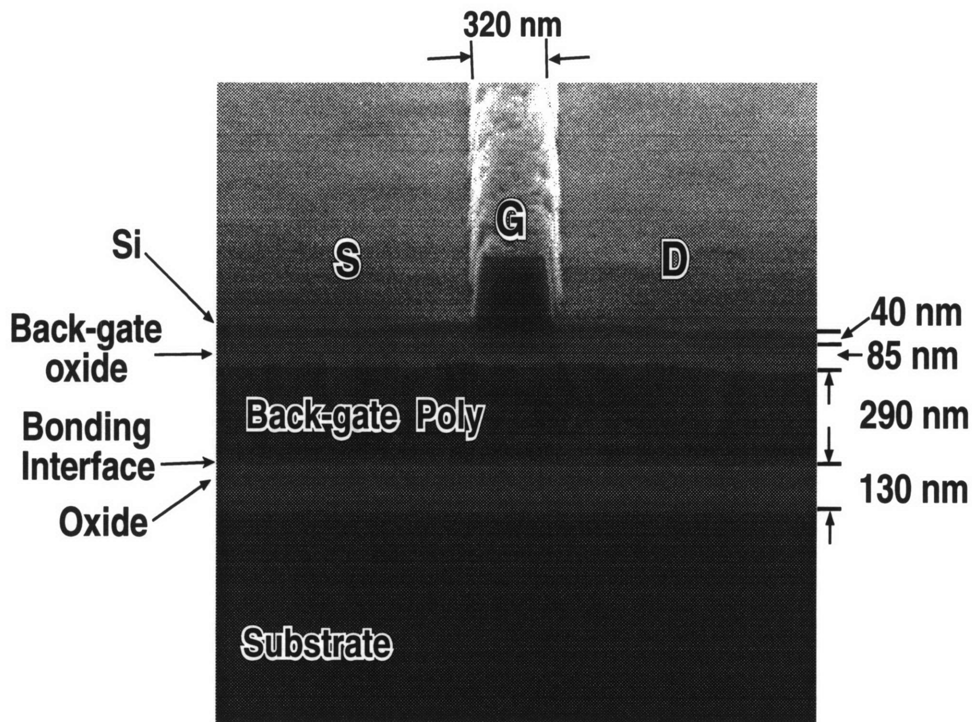


Figure 5-5: SEM micrograph of the SOIAS device.

5.2 Device Measurement Results

Figures 5-6 and 5-7 show the measured I-V and subthreshold device characteristics for NMOS and PMOS at L_{eff} is 0.44 μm and 0.35 μm, respectively, at two different threshold voltages tuned by biasing the back-gate. A 250 mV change in threshold voltage results in a 3.5 and 4 decade reduction in off current and a 50 and 80% current increase at 1 V operation for PMOS and NMOS, respectively. Figure 5-8 shows the measured maximum and minimum tunable V_T limits for the above nominal design parameters. The x-axis is the designed quiescent V_T. The quiescent V_T was obtained either by varying the doping or the silicon film thickness. Therefore, for low V_Ts at zero back-gate bias, the film is strongly fully depleted, i.e. either the doping level is low or the silicon film is thin. Similarly, for the high V_Ts at zero back-gate bias, either the doping level is high or the silicon film is thick. The y-axis, tunable V_T, was obtained by applying various back-gate biases. The tunable V_T range is quite large (approximately 1V) for fully depleted back interface as can be seen for the lowest V_T case (V_T=-0.2 V) at zero back-gate bias. The limits of the upper and lower tunable V_T range are determined by the back interface becoming either accumulated or inverted, in which case the back-gate becomes decoupled from the front-gate. Even for the partially depleted, highest V_T case (V_T=1.0 V), there is still a reasonable tuning range (approximately 0.5 V). This has implications for making FD SOI a viable technology since the threshold voltage and the device operating mode can be controlled precisely by the back-gate. Figure 5-8 demonstrates that V_T can be fine tuned over a wide range despite variations in t_{si}(average thickness=48.4 nm, maximum thickness= 69.9 nm and minimum thickness=37.6 nm) and L_{eff}. For example, a nominal V_T of 500 mV can be reached even for a ± 400 mV deviation by using a ±5-6 V back-gate bias. Typically, only a 200 mV switch in the V_T is sufficient to achieve approximately three decades reduction in the subthreshold leakage current. This design range fits well-within the limits of the tunable V_T band for the given film thicknesses and doping levels. For a lithographic shrink on the front-gate without changing any of the other design parameters, the control of the back-gate on the

front-gate V_T is not deteriorated for NMOS and PMOS effective channel length down to 0.25 μm and 0.17 μm as shown respectively in Figure 5-9. Figure 5-10 shows the measured threshold voltage roll-off for NMOS and PMOS devices. The threshold voltage, in this case is defined as the gate voltage at I_{DS}=1μA/μm. The two curves for zero and 3 V back-gate bias remain approximately parallel, indicating that the back-gate bias does not affect the roll-off. Furthermore, the small separation between V_{DS}=0.05 and 1.0V curves at a back-gate bias of 3 V shows that the drain-induced barrier lowering (DIBL), i.e. parallel shift of a given subthreshold current as a function of V_{DD}, is small.

5.3 Dynamic Operation

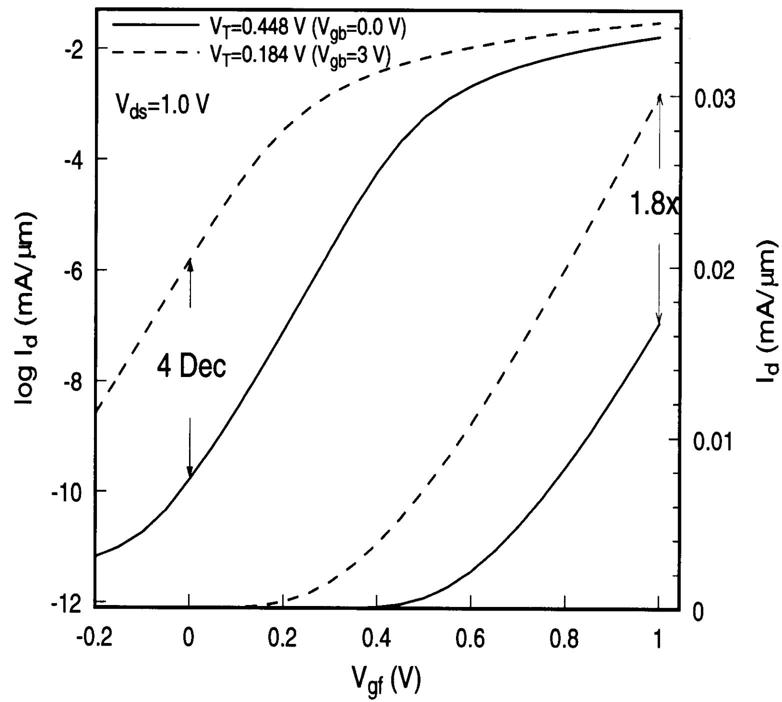
Figure 5-11 depicts schematically the measurement setup for testing ring oscillator and inverter chains. The power supply, ground, back-gate power supply for NMOS and PMOS are all separate for the buffer and the ring/chain. Figure 5-12 shows the frequency of a 101-stage ring oscillator as a function of varying the back-gate-controlled V_T for either the NMOS or PMOS only; hence, there is complete independent control of the NMOS and PMOS device threshold voltages. Figure 5-13 shows the actual output of the ring oscillator. For a 200 mV change in V_T for both the NMOS and PMOS, the result is a 36% change in the speed at V_{DD} of 1V.

In order for this scheme of dynamic threshold control to work properly, the V_T must change quasi-statically with back-gate switching. This is of concern because of the relatively high back-gate sheet resistance. The verification of this quasi-static control of the V_T was carried out with an experiment depicted in Figure 5-14. An NMOS transistor's gate was tied to V_{DD} of 0.5 V which is close to the threshold voltage of this device at zero back-gate bias, and a 50 ohm resistor was placed between V_{DD} and the drain. The back-gate was pulsed at various frequencies and pulse heights while the output at the drain, V_{out} was monitored. As shown in the inset to Figure 5-14, the device under test has an annular gate with a large 23x25 μm back-gate, with one contact off to the side. When the input pulse applied to V_{gb} is high, V_{out} is low

because more current is pulled through the resistor due to a lowered V_T. Figure 5-14 shows the V_{out} at 5, 10, and 20 MHz pulses on the back-gate. Even at 20 MHz, the V_{out} is still following the input pulse for this fairly large back-gate. Knowing the V_{out} value, the dynamic current due to lowering of the V_T (switching of the back-gate) can be overlayed onto the DC measured currents for various back-gate biases, i.e. pulse heights. The dynamic current was simply calculated as $(V_{DD}-V_{out})/R$, where R is the resistor value. Figure 5-15 shows the composite of these two measurements. The x-points lying precisely on the DC measured current curves is indicative of the quasi-static control of the device V_T through dynamic back-gate biasing.

5.4 SOIAS Circuits

Fabrication of an 8 bit ripple adder and an 8 bit multiplier in our SOIAS CMOS process with one of level metal has been successful. Functionality of these circuits was verified. The speed of these circuits was, however, limited by the use of excessive polysilicon routing in the layout due to availability of only one level of metal. The same or similar circuits will be built in a two-level metal process at MIT Lincoln Laboratory, in which case, the true speed of these circuits can be tested.



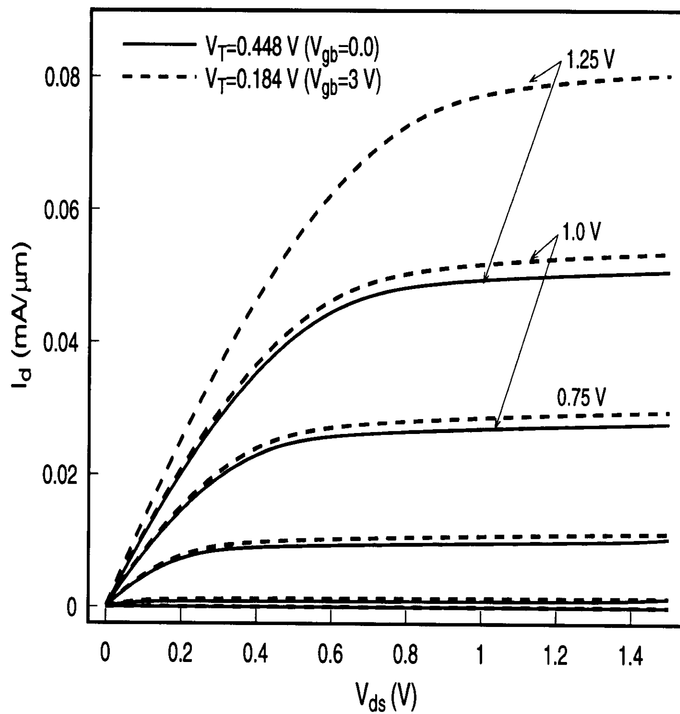
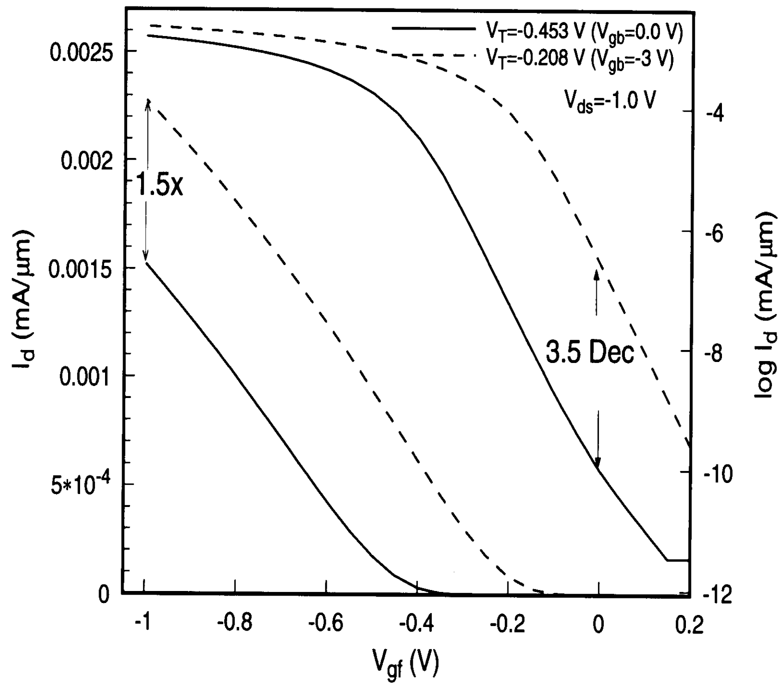


Figure 5-6: Measured NMOS device characteristics showing the effects of switching the V_T on the off current and the drive current at V_{DS}=1 V.



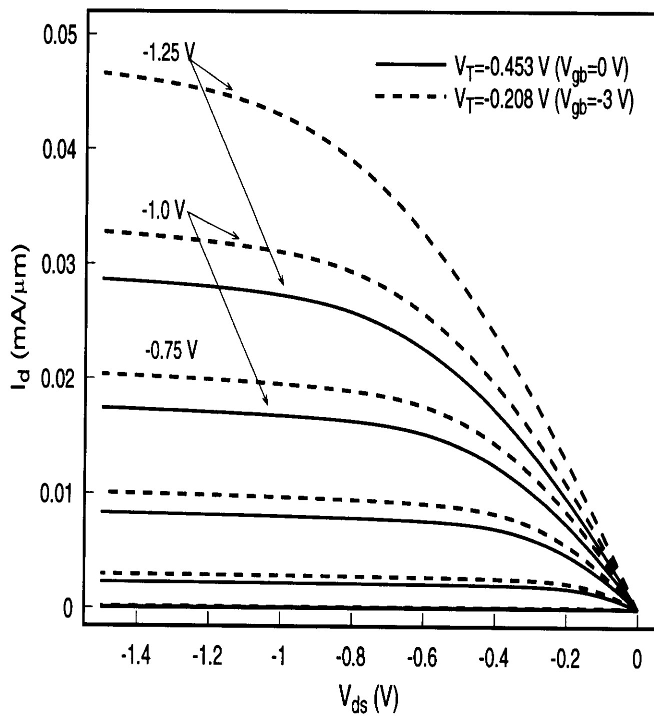


Figure 5-7: Measured PMOS device characteristics.

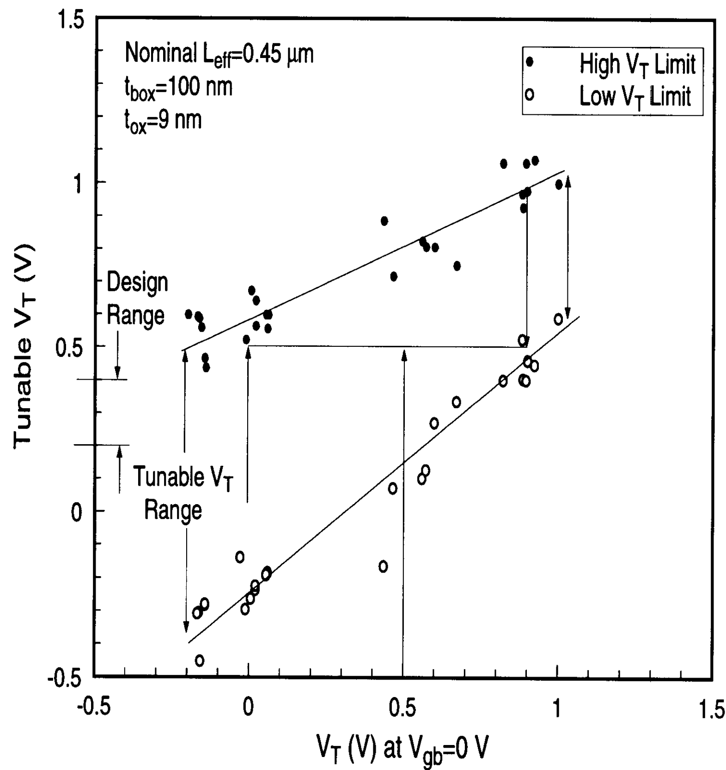


Figure 5-8: Measured threshold voltage tuning range. The x-axis is the quiescent V_T determined by the silicon film thickness and the doping level. The y-axis is the tunable V_T set by the back-gate bias; the design range of 200 mV switch in V_T fits well within the tuning range.

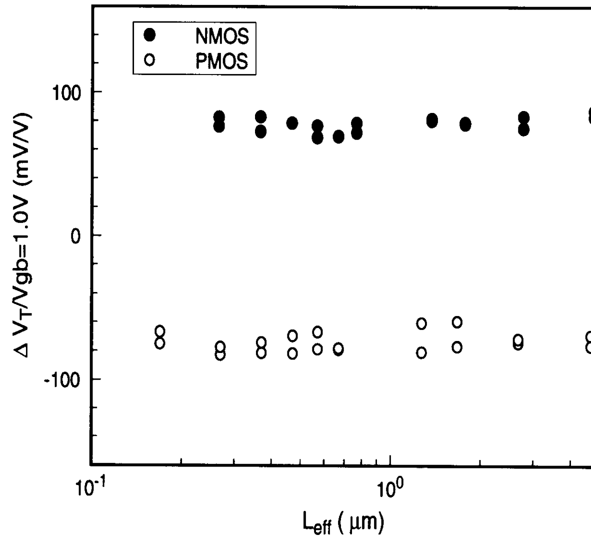


Figure 5-9: Measured threshold voltage control as a function of gate length.

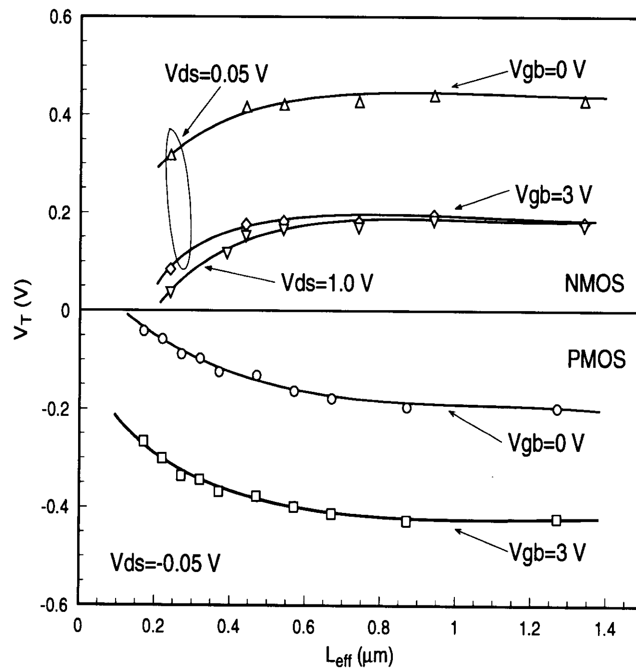


Figure 5-10: Threshold voltage roll-off as a function of gate length at different back-gate biases for NMOS and PMOS.

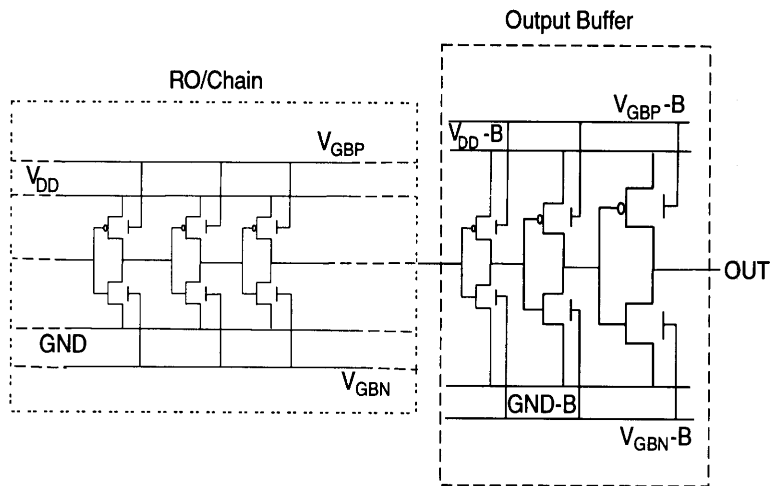


Figure 5-11: Schematic of measurement setup for the 101 stage ring oscillators and 50 stage inverter chains.

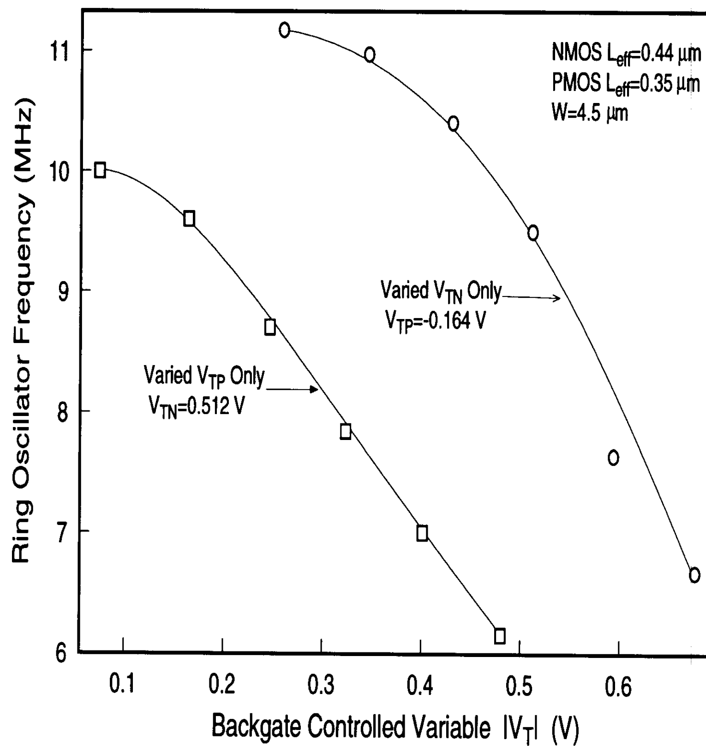


Figure 5-12: Measured ring oscillator frequency as a function of independently controlled V_T through back-gate biasing.

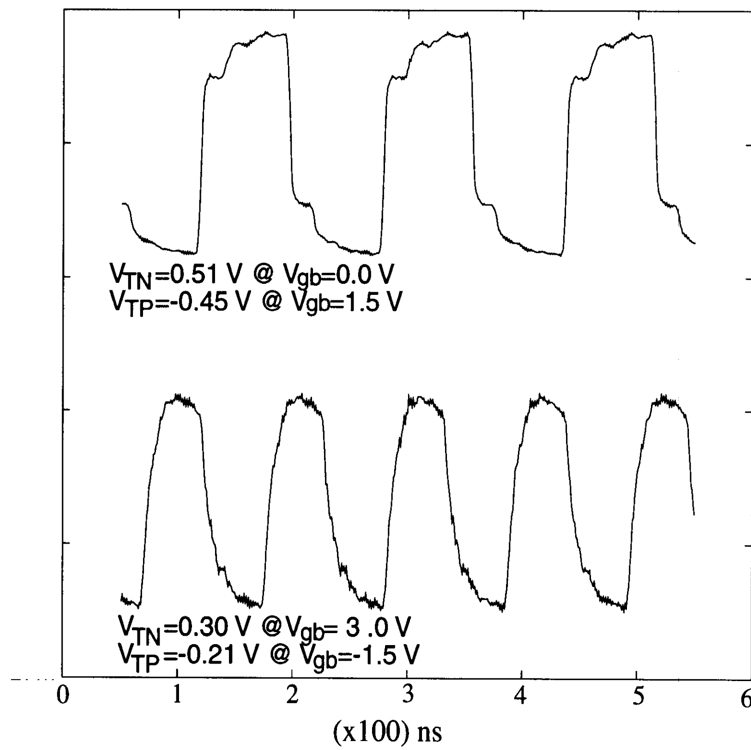


Figure 5-13: Measured 101 stage ring oscillator output frequency as varied by changing V_T. A 36% change in the speed is observed for 200 mV change in V_T at V_{DD} of 1 V.

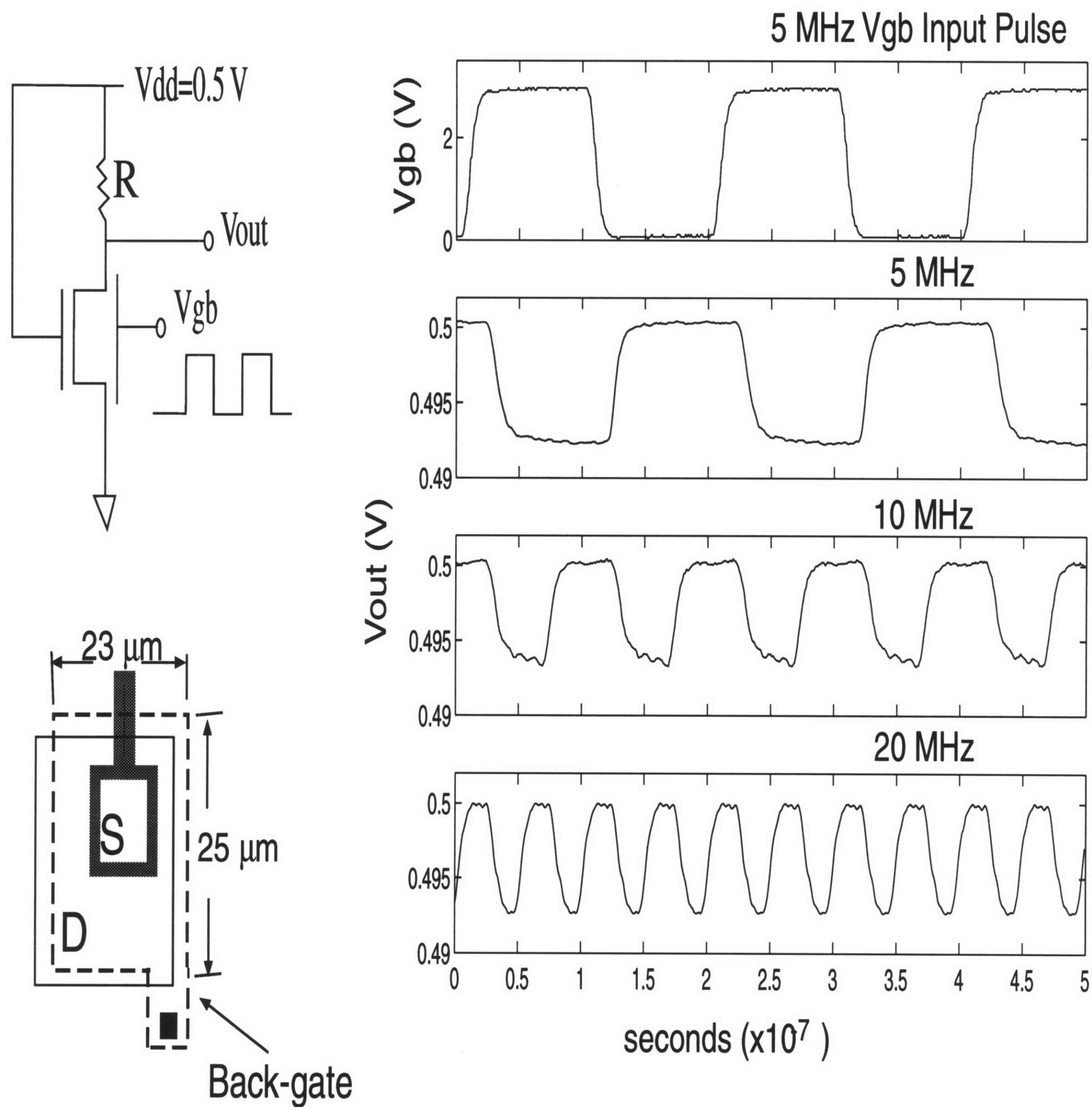


Figure 5-14: Schematic showing the measurement setup and device under test for dynamic switching of the back-gate experiment. The V_{out} at the drain for three different input frequencies are shown with one example of the input pulse on V_{gb} at 5 MHz.

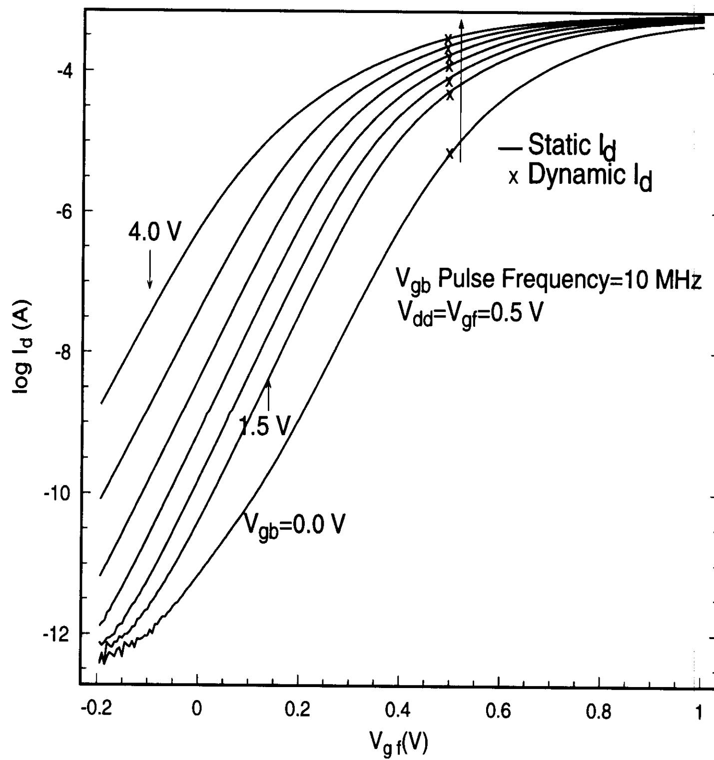


Figure 5-15: Overlay of dynamic current due to switching of back-gate on DC characteristics at various static back-gate biases. The points marked X represent dynamic measurements at 10 MHz.

Chapter 6

SOIAS for Ultra-Low Power Applications

The SOIAS technology for dynamic threshold voltage control is proposed to be implemented at the functional module level, similar to the well biasing scheme in bulk. The back-gate controls the (V_T) of the front-gate device, and the NMOS and PMOS back-gates are switched independently from each other and from the front-gates. For burst-mode high performance and low power applications, the threshold voltage would be raised during idle periods to reduce the static leakage current, and lowered during active periods to achieve high performance. For continuous mode operation, the back-gate would not be switched.

6.1 Energy Consumption Model for SOIAS

A theoretical model was developed to evaluate and compare the total energy dissipation for the SOIAS variable V_T technology *vs* an optimized low power SOI CMOS technology. We have chosen to assume a model of operation in which functional modules share a common V_T, i.e. all same polarity transistors in such modules have the same V_T. This implies that in addition to the module's conventional gated clock in the SOI implementation, another gated clock would be needed for the back-gate control. Under this model, an active module's idle devices are left in a low-leakage

state. In the modeling of a microprocessor's energy dissipation, various modules were considered such as the ALU adder unit, the shifter, and the integer multiplier. In order to analyze the applicability of the SOIAS technology to low-power static CMOS logic, we have developed total energy dissipation equations including switching and static energies for a SOIAS and the benchmark SOI technology:

Variable V_T technology (SOIAS):

$$\begin{aligned} E_{SOIAS} &= \chi A_{fg} \gamma C_1 V_{dd}^2 && \textit{Dynamic Energy} \\ &+ \chi A_{fg} I_{off(low)} V_{dd} t_c + (1 - \chi A_{fg}) I_{off(high)} V_{dd} t_c && \textit{Static Leakage Energy} \\ &+ \chi A_{bg} C_{box} V_{gb}^2 && \textit{Back-gate Switching Energy} \end{aligned}$$

Benchmark constant V_T technology (SOI):

$$\begin{aligned} E_{SOI} &= \chi A_{fg} (\gamma C_2 V_{dd}^2) && \textit{Dynamic Energy} \\ &+ I_{off(low)} V_{dd} t_c && \textit{Static Leakage Energy} \end{aligned}$$

These equations include:

(a) *Algorithm and architecture parameters:*

A_{fg} = module activity factor

A_{bg} = back-gate activity factor

γ = node switching probability during active period

χ = system activity of processor during interactive computation

(b) *Technology and circuit parameter:*

$t_c = 1/f_{clock}$

(c) *Technology parameters:*

C_1,C_2=total physical capacitance (C_g+C_{gdo}+C_f+C_{bgdo})

$I_{off(low)}$=low V_T off current

$I_{off(high)}$=high V_T off current

C_{box}=back-gate oxide capacitance

V_{gb}=back-gate bias

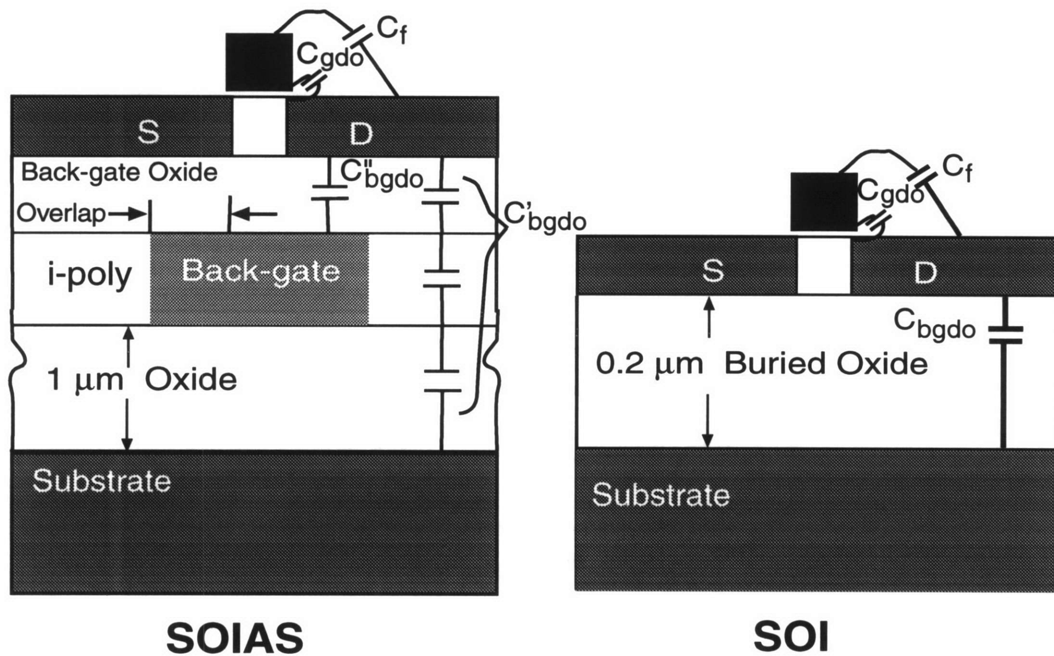


Figure 6-1: Capacitances of SOIAS and SOI devices.

C_g is the front-gate capacitance, C_{gdo} is the gate to source/drain overlap capacitance, C_f is the gate to source/drain fringing capacitance, and C_{bgdo} is the source/drain to back-gate overlap capacitance. Figure 6-1 shows the various capacitances. For the SOIAS device, C_{bgdo} is composed of two components, as seen in Figure 6-1; the overlap with the heavily doped back-gate (C_{bgdo}") and the overlap with the intrinsic region of the back-gate poly (C_{bgdo}'). The total energy equation for the SOIAS is composed of three components, the dynamic switching energy, the static leakage energy, and the overhead energy required to switch the back-gate. For the SOI technology, the total energy is composed of the dynamic switching energy and the static leakage energy. The applicability of SOIAS technology is a strong function of transistor usage, functional block usage, and interactive system activity; i.e. the parameters A_{fg}, A_{bg}, γ, and χ.

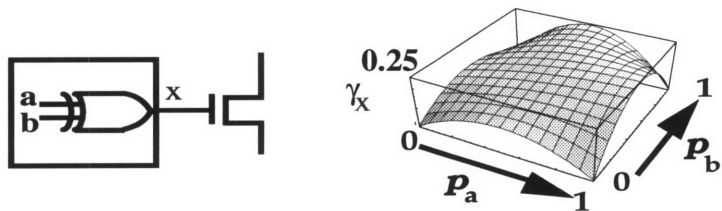
The dynamic switching energies for the SOIAS and SOI technologies are approximately the same for burst-mode as well as for continuously computing systems provided that the source/drain to back-gate overlap capacitance, C_{bgdo}, of the SOIAS device does not contribute significantly to the total capacitance driven. In order for

the SOIAS technology to provide energy saving over the SOI, the sum of the static leakage energy and the back-gate switching energy for the SOIAS must be less than the static leakage energy of the constant low V_T SOI for low power applications. For burst-mode computational systems, the A_{fg} and A_{bg}, are numbers much smaller than one. Therefore, the back-gate switching overhead energy is small due to A_{bg} being a small number. The static energy for the SOIAS technology would also be much less than that of the SOI technology because the low V_T leakage energy is weighted by a very small number, A_{fg}, and the high V_T leakage energy is low due to low sub-threshold leakage. The contribution of each of the components to the total energy are presented and analyzed for the adder, multiplier and shifter modules.

6.1.1 Activity Factor Definitions

Figure 6-2 illustrates how the activity factors are defined. A_{fg} is the module activity factor which is the fraction of time a module (e.g. an adder) is "on", i.e. doing computation. From Figure 6-2, the A_{fg} can be defined as the number of CLK cycles that the adder is "on", i.e. the number of periods in CLK-ADD, over the total number of CLK cycles observed. A_{bg} is the back-gate activity factor which is dependent on how frequently the module is "on". The gated-clock for the back-gate, CLK-Backgate, makes a 0 to 1 transition whenever the CLK-ADD is "on". Therefore, A_{bg} can be defined as the number of 0 to 1 transitions of the back-gate over the total number of CLK cycles observed. One important point to note here is that the back-gate activity factor (A_{bg}) is always less than or equal to the module activity factor (A_{fg}). For example, an adder can be active several clock cycles in a row, and hence the V_T for this module is then left in the low V_T state during those cycles. Therefore the back-gate only needs to be switched once for all the consecutive cycles that the module is active. The individual node transition activity, γ, is strongly dependent on the signal statistics. For example, in Figure 6-2, the node transition activity of node X is a function of the probability of inputs A and B making 0 to 1 transitions. Since the maximum probability of such a transition for each input is 0.5, the maximum γ_X is 0.25.

Circuit node transition activity:



Functional module activities:

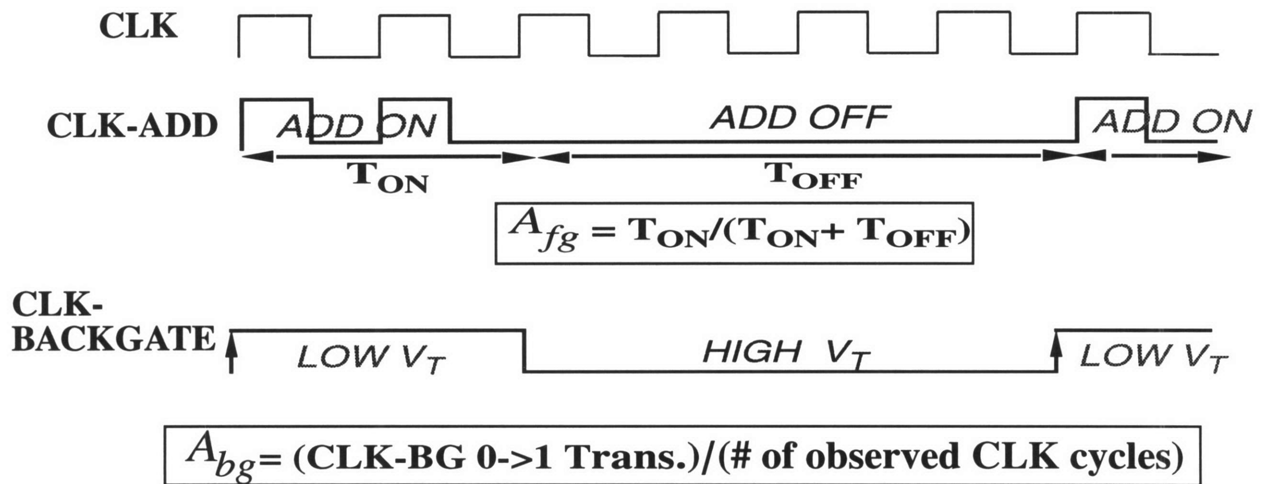


Figure 6-2: Schematic illustration of the definitions for the various activities factors.

	A_{fg}	A_{bg}
Adder	0.69	0.21
Shifter	0.11	0.09
Multiplier	0.008	0.008

Table 6.1: Activity values from profiling the data encryption program (IDEA).

In order to determine functional block usage patterns (A_{fg} and A_{bg}), a series of program profiling experiments were performed using the ATOM code instrumentation interface [74] for a particular microprocessor implementation, compiler technology, and various algorithms. In order to map the assembly language instructions to the functional blocks use, certain assumptions about the implementation were made. The ALU adder, for example, is used for all add, compare, load, and store instructions. The ratio of the compiled relevant instructions for a specific module to the total number of instruction calls in a program gives the A_{fg} value. The ratio of the number of blocks of functional module usage to the total number of instruction calls give the A_{bg} value. Since the back-gate activity depends on the module activity as well as the module usage pattern, A_{bg} can be correlated to A_{fg} by the following relationship [75]

$$A_{bg} = A_{fg} - {A_{fg}}^2 \tag{6.1}$$

Table 6.1 shows the profiling results for the data encryption program, IDEA.

6.1.2 Simulation Criteria

The simulations were carried out with the following design specifications: t_{si}=40 nm for the practical limit of thinning the silicon film for both SOIAS and SOI, t_{fox}=7 nm for a 0.25 μm L_{eff} technology, t_{box} for the SOI technology is 200 nm, and for the SOIAS technology is varied from 70 to 100 nm, C_f and C_{gdo} were obtained from measurements of our current SOIAS CMOS technology, C_g, C_{box}, and C_{bgdo} are calculated assuming a parallel-plate capacitor configuration. The fringing capacitance

for the back-gate can be calculated by solving, numerically, the equations developed by Greeneich [76] for all combination of t_{box} and back-gate overlap. In this analysis, the fringing capacitance of the back-gate is assumed to be a small fraction of the total back-gate capacitance, and therefore is neglected. The node transition probability, γ, in one active cycle is assumed to be 40% obtained by estimating a ripple carry adder under random input pattern, and χ, the interactive computation activity, is assumed to be 2% as in an X server. The clock frequency is 100 MHz, $V_{T,SOIAS}$(low)=$V_{T,SOI}$=200 mV, $V_{T,SOIAS}$(high)=400 mV, and V_{DD} = 1.0 V. The drive current per unit width at V_{DD}=1 V and V_T of 200 mV is the same for both SOIAS and SOI, as obtained from MINIMOS simulation. The 200 mV switch in V_T is maintained as the t_{box} is scaled, thus the back-gate bias is scaled concomitantly. The leakage current per unit width and the V_{gb} necessary to maintain 200 mV of V_T switch were also obtained using MINIMOS simulation. The total energy dissipation is per unit width.

6.2 Energy Trade-offs in Technology Design Space

For the given algorithmic parameters obtained from profiling the data encryption algorithm, IDEA, and assuming χ=2%, the latitude in the technology design can be shown in technology parameter space, in this case, t_{box} and source/drain overlap with back-gate are the parameters of choice. These two parameters were chosen because their contribution to the total energy is considered to be overhead. The scaling of t_{box} results in significant reduction of the back-gate switching energy, i.e. the overhead energy. However, scaling the back-gate also requires that the source/drain to back-gate overlap be reduced simultaneously in order to minimize the parasitic capacitance. The overlap is normalized with respect to the gate length, L_g. Figure 6-3 compares the front-gate dynamic energy component for the SOIAS and SOI technology. The source/drain overlap capacitance dominates at thin back-gate oxide and large overlap. The dark line demarcates the design space range where the SOIAS front-gate switching energy is equal to or less than that of the SOI.

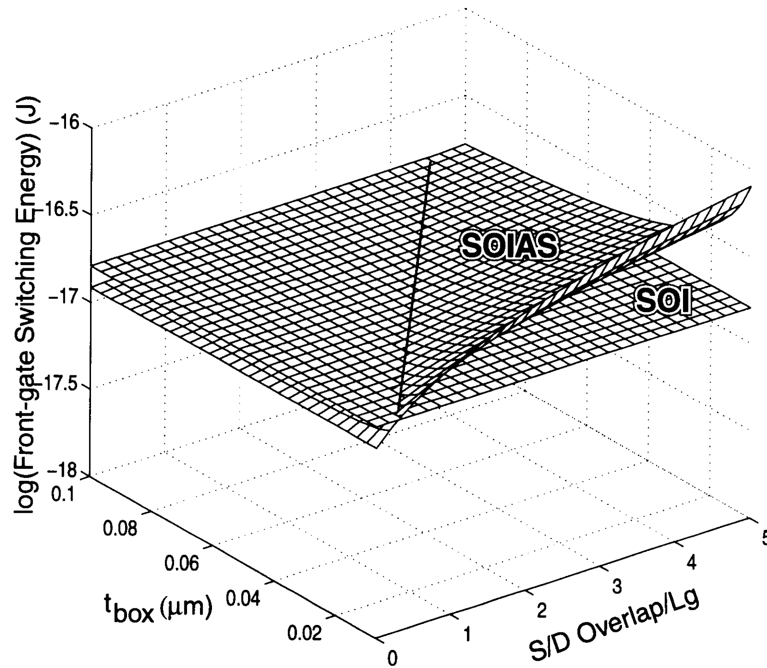


Figure 6-3: Comparison of the front-gate dynamic switching energy for SOIAS and SOI. The dark line indicates the boundary where the energies are the same. When the SOIAS front-gate switching energy becomes greater than that of the SOI, the overlap capacitance contribution becomes significant.

Figures 6-4 and 6-5 compare the front-gate and back-gate switching energies for the adder, multiplier and shifter. From these figures, it is clearly evident that the magnitude of these energies depends strongly on the relative magnitude of the activity factors. In order to minimize the overhead back-gate switching energy, the t_{box} and overlap values should fall in the range where the back-gate switching energy is less than or equal to the front-gate energy. For the adder, this is achievable for a fairly large range of t_{box} and source/drain overlap because the front-gate activity is relatively high, and the back-gate activity is approximately 1/3 of that of the front-gate; this is not the case for the multiplier and shifter since their front-gate and back-gate activities are very low and approximately equal. For the adder, for example, if 0.1 μm thick back-gate oxide is used, then for equal front-gate and back-gate switching energies, the maximum overlap allowed is 0.67% of the front-gate length. For the 0.25 μm technology, this requires the overlap to be no more than 0.17 μm. With the current DUV lithographic tools, this overlay requirement is easily attainable.

For the multiplier and shifter, the back-gate switching energy swamps out the front-gate switching energy, which seem to imply that the design space for the shifter and multiplier is very small. However, in benchmarking against the SOI technology, the total energy must be considered, and since the activities of these two modules are small, the total dynamic and static leakage energies is expected to be small for the SOIAS as compared to the SOI technology. Figure 6-6 and 6-7 show the relative contribution of the dynamic and static energies for the two technologies for all three modules. The static leakage dominates the total energy for the SOI technology while the dynamic component dominates for the SOIAS. This implies that as long as the energy cost of switching the back-gate plus the front-gate switching energy for the SOIAS is less than the constant low V_T static leakage energy, it is still worthwhile to dynamically switch the V_T.

The ratio of the total energies for the SOIAS and SOI are plotted in Figures 6-8 and 6-9. The plane outlined in dark is the break-even plane for SOIAS and conventional SOI technology. For low activity modules (the multiplier and shifter), the design space in favor of the SOIAS technology spans the entire parameter range under study for both t_{box} and source/drain overlap; i.e. even the maximum back-gate switching energy is still less than the constant low V_T static leakage energy. Therefore, even for the most energy costly back-gate design (thick t_{box} with maximum source/drain overlap), the SOIAS dynamically controllable V_T scheme still results in a total energy savings over the constant low V_T SOI technology for the multiplier and shifter. Of course, the better the technology (e.g. alignment overlay), the greater the energy savings. For relatively high activity modules (the adder), the design space in favor of SOIAS technology is smaller. There is, however, an optimal range of t_{box} which allows the most S/D overlap, i.e. the most process latitude, where the total energy is minimized for the adder module.

6.3 Energy Trade-offs in A_{bg} and A_{fg} Space

The ratio of the total energy dissipation for SOIAS and SOI was analyzed as

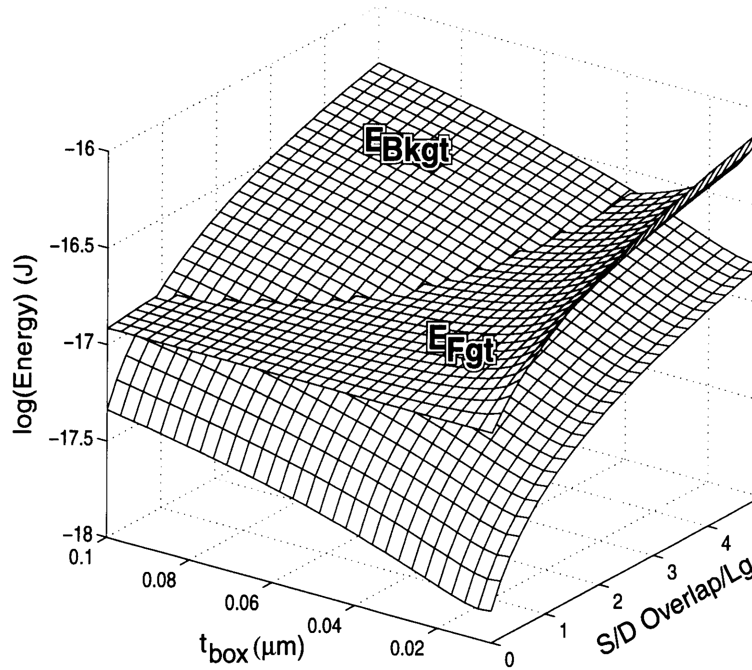
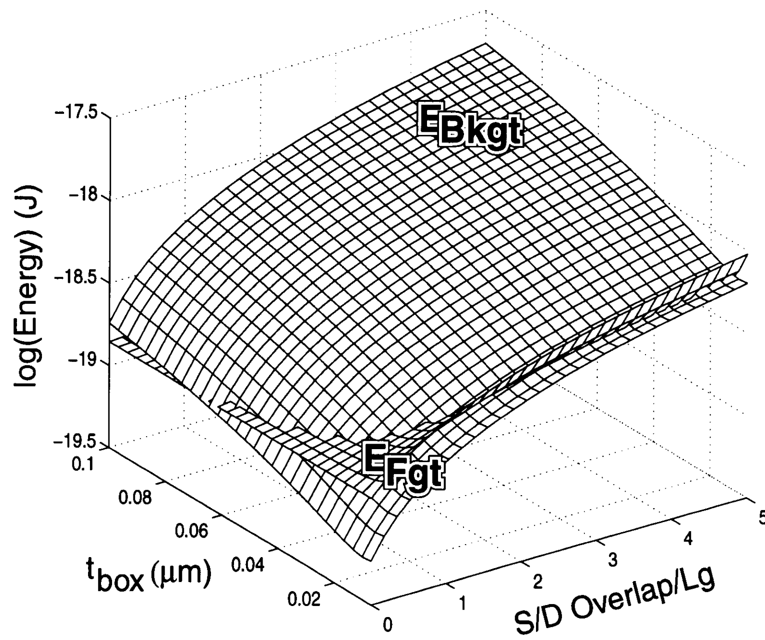


Figure 6-4: The front-gate and back-gate dynamic switching energies for the adder.

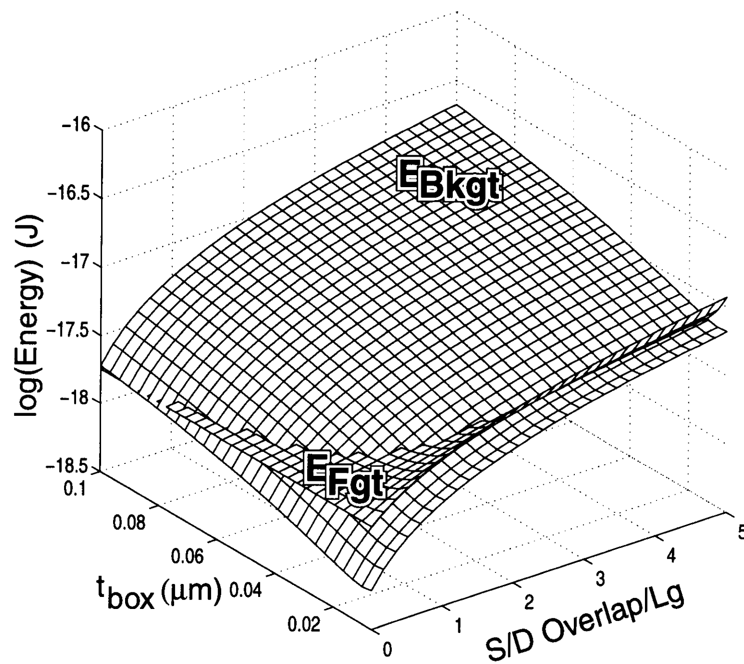
a function of algorithm and architecture dependent parameters(A_{fg} and A_{bg}), as shown in Figure 6-10). The dark line demarcates the break-even contour. For near continuous functional block usage which does not exhibit strong temporal locality, χ=100%, (e.g adder and shifter in a continuously computing system), the back-gates would not be switched, and hence E_{SOIAS} would be equal to E_{SOI} at constant low V_T. However, if the back-gate were to be switched in such systems, then the cost in energy for switching the back-gate is high, and, therefore ,the E_{SOIAS} would be greater than E_{SOI}. This is indicated in Figure 6-10 by the open symbols. In a system which is frequently idle while awaiting I/O, such as an X-server which is active approximately 2% of the time, the SOIAS technology dissipates much less energy than conventional SOI; this is depicted by the filled symbols. Two sets of technology design parameters were used differing only in the back-gate to source/drain overlap, with all others being equal. In Figure 6-10a, the source/drain to back-gate overlap was assumed to be $2L_g$ for a 0.25 μm technology, i.e. 0.5 μm overlap. These points are indicated in the technology design space of Figure 6-8 and 6-9. Total overlap for both source and drain would be 1.0 μm. This design rule is extremely lax with respect to the

overlay capabilities of current steppers. Even so, the energy savings in burst mode (χ=2%) is 43% for the adder (A_{fg}=69%, A_{bg}=21%), 80% for the shifter(A_{fg}=11%, A_{bg}=9%) , and 97% for the multiplier(A_{fg}=0.8%, A_{bg}=0.8%). In Figure 6-10b, the source/drain to back-gate overlap is reduced to 0.5L_g, i.e. 0.125μm (total of 0.25 μm for both source and drain). This overlay is certainly achievable in the current DUV stepper tools. In this case, the multiplier, even in continuous mode (χ=100%), shows energy savings with dynamic switching of the V_T. Of course, in burst mode, the energy savings is even more with this design rule than that of the 2L_g design rule for all three modules.

From the analysis done with the energy model, it is clear that being able to predict the system activity is important in determining the suitability of dynamic V_t control for such a system intended to minimize total energy dissipation. Furthermore, in evaluating the total energy in the technology design space enables the designer to narrow down the range of device design parameters that will minimize total system energy and meet the capabilities of current semiconductor tools.



(a) multiplier



(b) shifter

Figure 6-5: The front-gate and back-gate dynamic switching energies for the multiplier and shifter.

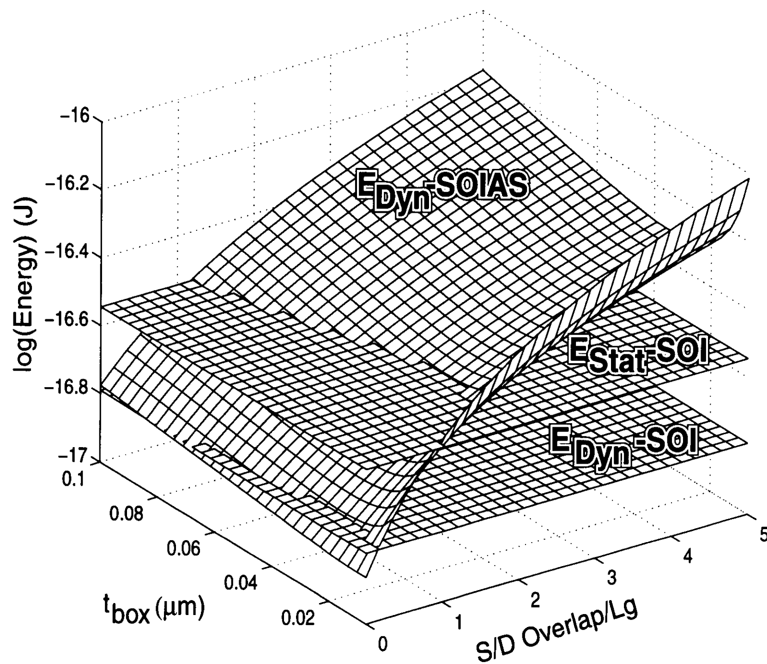
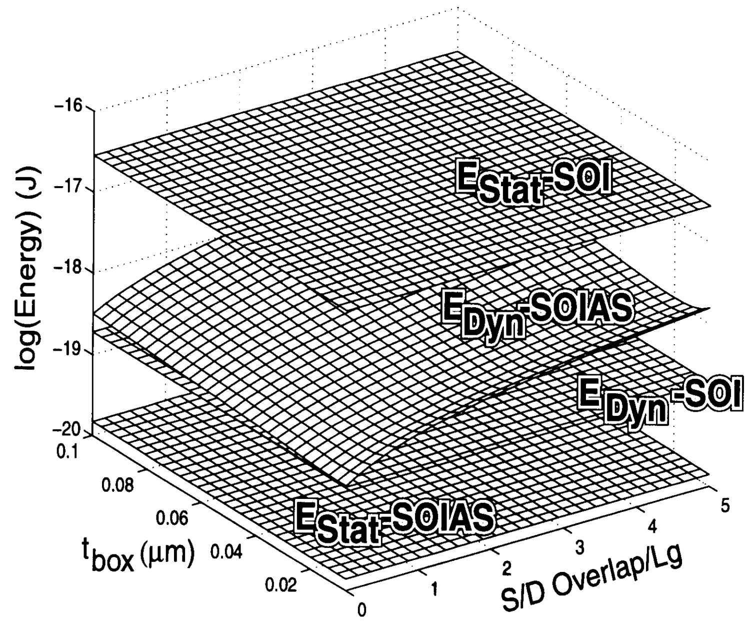
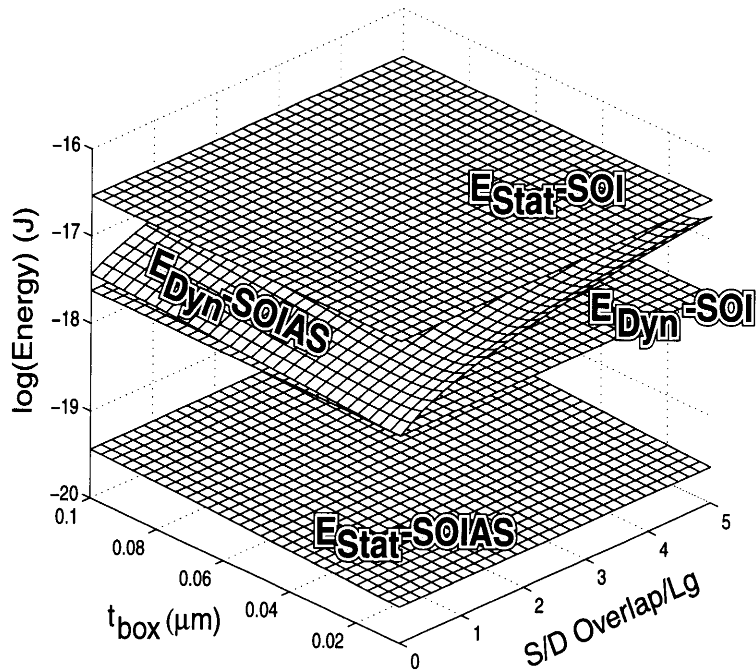


Figure 6-6: Comparison of dynamic switching energies (back-gate + front-gate) and the static leakage energy components for the SOIAS and SOI technologies. The static leakage for the SOIAS is not shown on the plot because it is off the scale.



(a) multiplier



(b) shifter

Figure 6-7: Comparison of dynamic switching energies (back-gate + front-gate) and the static leakage energy components for the multiplier and shifter.

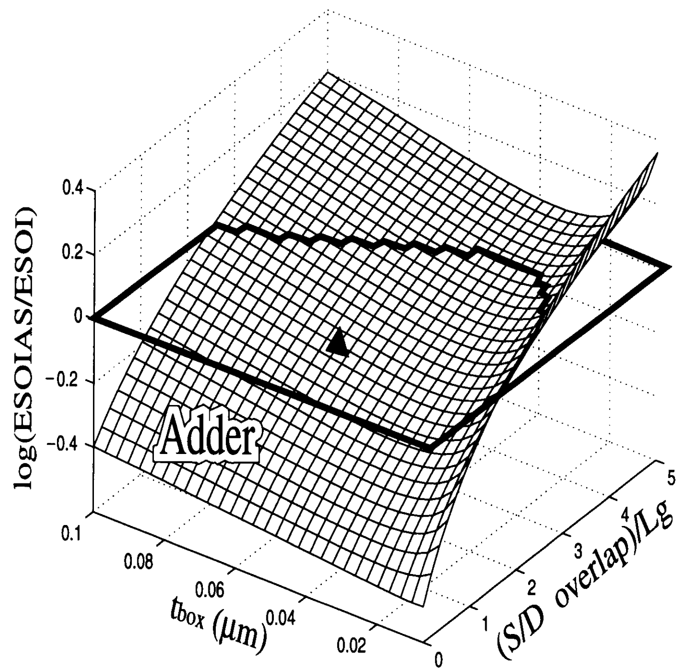


Figure 6-8: The adder total energy ratio for SOIAS and SOI. The dark line outlines the break-even plane.

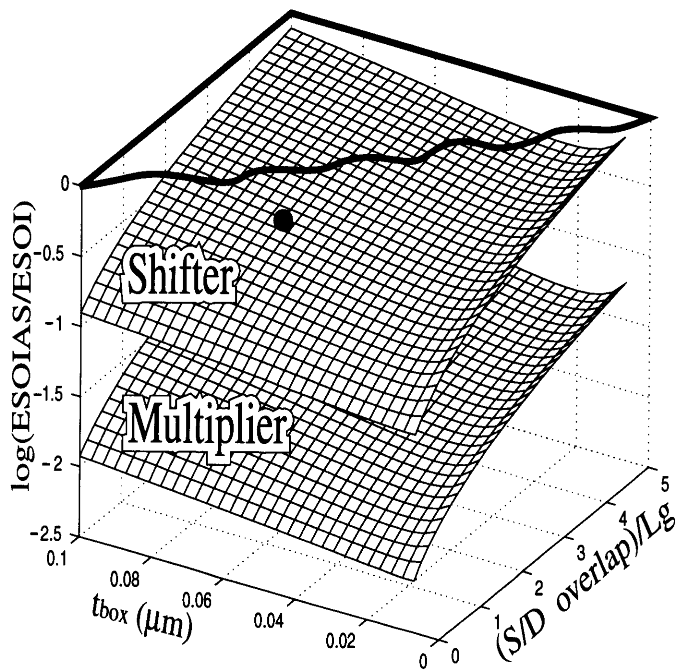
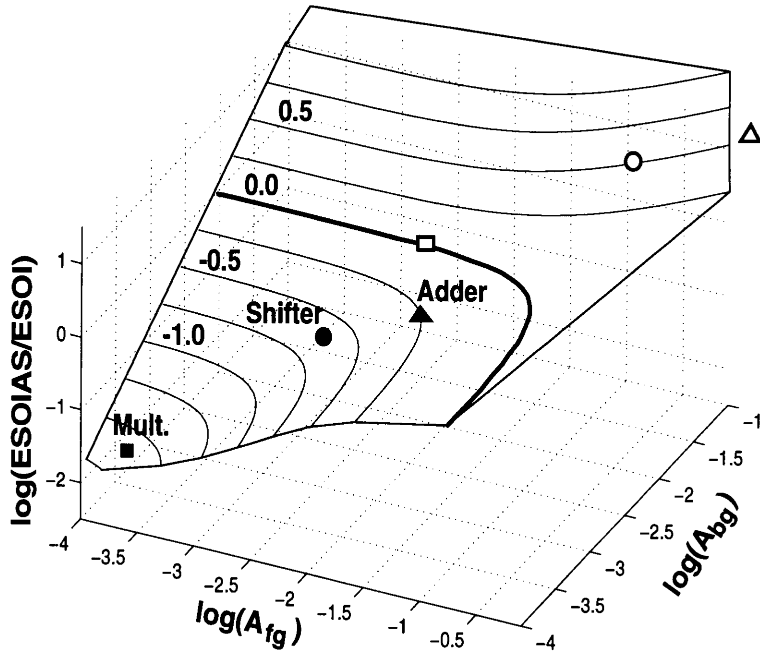


Figure 6-9: The multiplier and shifter total energy ratio for SOIAS and SOI. Both fall below the break-even plane.



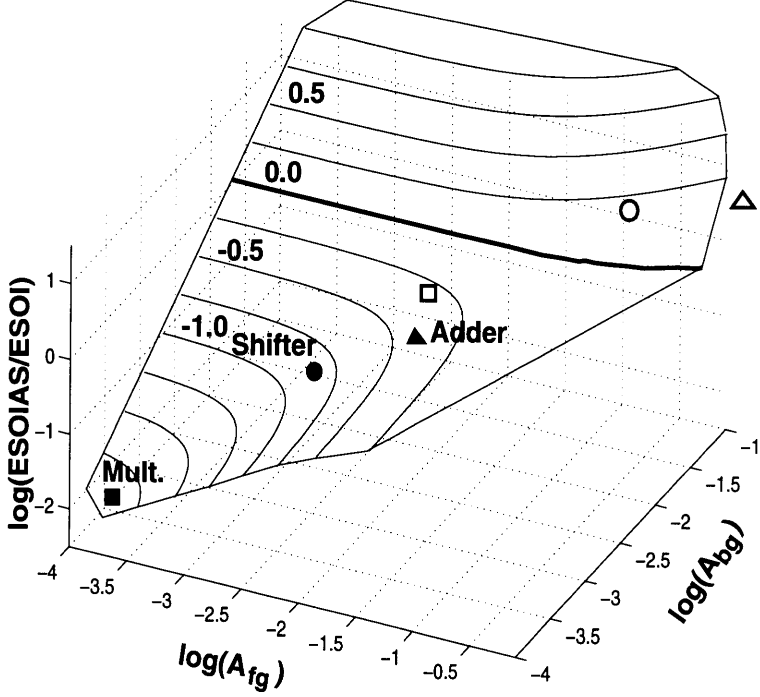


Figure 6-10: The total energy ratio for SOIAS and SOI with χ=100% (open symbols) and χ=2% (filled symbols). Squares represent the multiplier; circles represent the shifter and triangles represent the adder. Shown are two plots for different technology design parameters: (a)t_{box}=50 nm, s/d overlap=$2L_g$=0.5 μm. (b)t_{box}= 50 nm, s/d overlap=$0.5L_g$=0.125 μm.

Chapter 7

Ultra-Low Voltage Design Space

In the previous chapter, only burst mode operation with system activity, χ=0.02, was considered for a fixed V_{DD}=1 V with dynamic control of the V_T. The flexibility of electronically controllable V_T allows the evaluation of the entire V_{DD} and V_T design space for optimal operating points at minimum energy for various system activities. In this study, the V_{DD} range was 0.17 V to 2.6 V, and the V_T range was 50 mV to 650 mV. Typically in a microprocessor, the activity can vary over a large range depending on the application run. If both V_{DD} and V_T can be adaptively changed as the activity changes, the system can always operate at minimum total energy. A DC-DC converter has been demonstrated which generates arbitrary (and variable) supply voltages with high efficiency [77], and the SOIAS technology provides the electronically variable V_T. In this analysis, we examine two levels of optimization:

- optimal V_{DD} and V_T at minimum energy for various system activities (χ ranging from 1 to 0.01) for adaptively changing V_{DD} and V_T.
- optimal V_{DD} and V_T at minimum energy for burst mode operation (χ=0.01) with dynamic V_T control during application run.

Our approach is to apply the energy consumption model to measured dynamic energy and static leakage energy data for the entire V_{DD} and V_T design space under study.

7.1 Measurement of Energies

The design parameters for the devices under study are: silicon thickness 35 nm, top gate oxide 7 nm, back-gate oxide 100 nm. The different threshold voltages on the same device/circuit are obtained by biasing the back-gates. The quiescent V_T for these devices were $\pm$340 mV, both polarity back-gate biases were applied to achieve a full range of threshold voltages from $\pm$50 mV to $\pm$650 mV. Figure 7-1 shows the inverter transfer characteristics at various V_{DD}s. The delay per stage was obtained from 101 stage inverter ring oscillators, and the dynamic and static leakage currents were measured on identical inverter chains. Figure 7-2 shows the extracted effective capacitance from the energy measurements. The non-flat surface is due to non-linearity of the gate capacitance and the short-circuit component. Figure 7-3 shows the measured total energy for a 101-stage ring oscillator at three different frequencies. The optimal V_T for minimum total energy is between 200 to 250 mV; the tradeoff between dynamic switching energy and static leakage energy is clearly evident. Figure 7-4a-c shows the measured static leakage energy, dynamic switching energy and delay per gate as a function of varying V_{DD} and V_T. The actual leakage can be obtained by scaling the numbers in Figure 7-4a to the appropriate clock cycle.

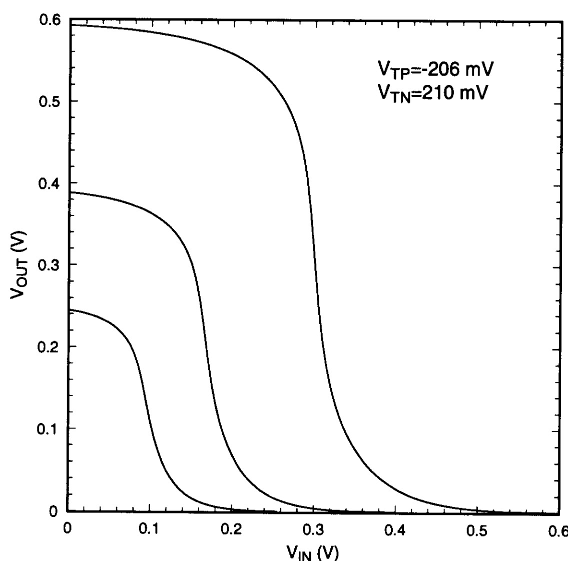


Figure 7-1: Inverter transfer characteristics.

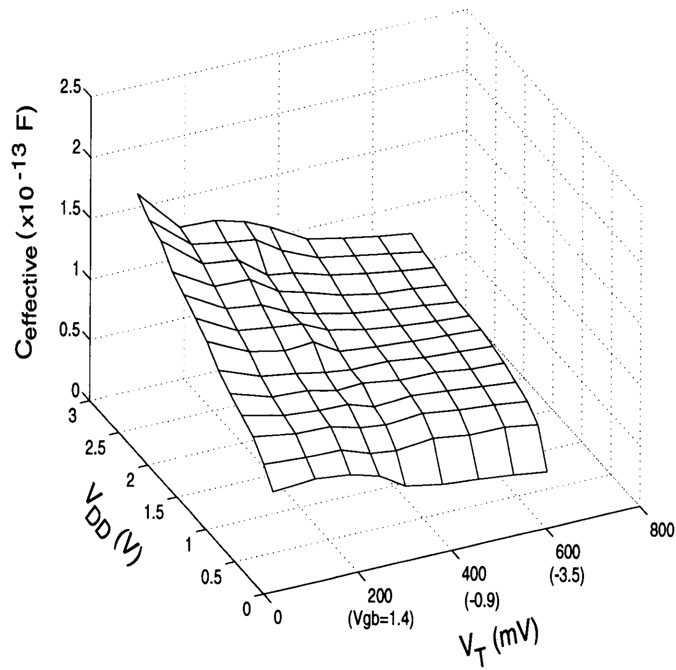


Figure 7-2: The measured effective switching capacitance.

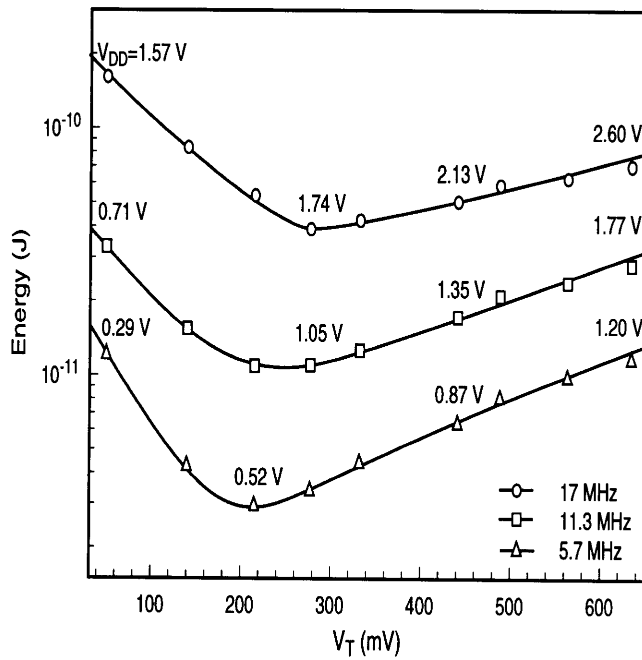
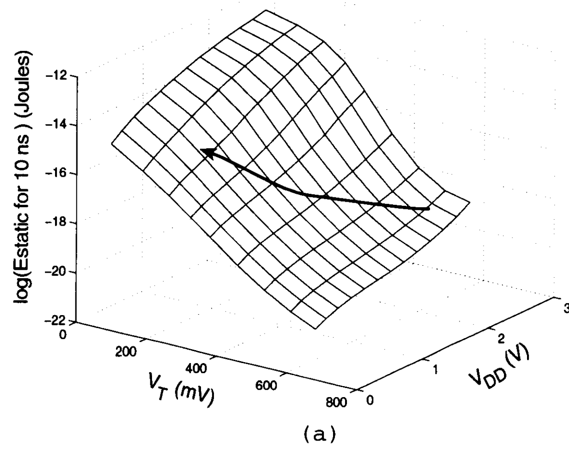
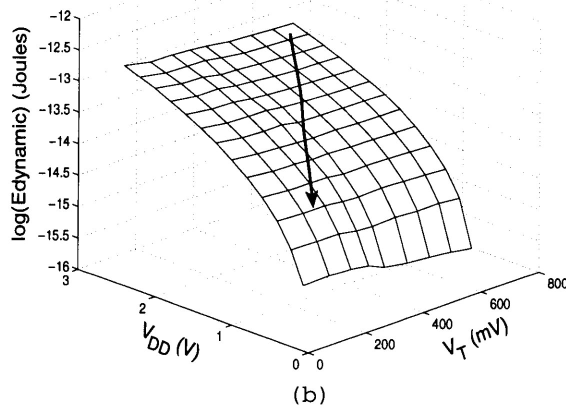


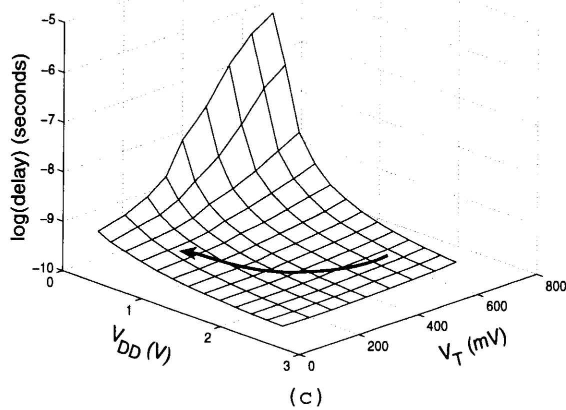
Figure 7-3: The measured total energy for 101-stage ring oscillator running at three different frequencies.



(a)



(b)



(c)

Figure 7-4: The measured (a) static (b) dynamic energies, and (c) delay per stage as a function of V_T and V_{DD}. As guides, the arrows indicate the direction of V_{DD} and V_T scaling.

7.2 Variable V_{DD} and V_T:

In this analysis, the optimal V_{DD} and V_T was found based on the metric of minimum energy dissipation at maximum possible clock rate. Typically, there are 10 gates in the critical path of a microprocessor. The minimum possible clock period is, therefore, 10 times the delay per stage. The relative contribution of the static leakage energy and dynamic leakage energy depends on the mode of operation. The optimal V_{DD} and V_T points are found from contour plots like Figure 7-5. This figure shows contours of energy and performance for 25% system activity for an adder. For a given performance, optimal V_{DD} and V_T can be found for minimum energy as indicated by the arrows in Figure 7-5. Motion perpendicular to the delay contours shows the trade-off in energy as performance improves. The energy contours change shape and gradient as the activity changes because the relative contributions of the dynamic and static components change, thus shifting the optimal V_{DD} and V_T point of minimum energy. Figure 7-6 shows the optimal V_{DD}s and V_Ts, at minimum energy, for three different modules, running at 100 MHz clock frequency, with various system activities. Each point was extracted from graphs similar to Figure 7-5 at the approximate minimum total energy. As can be seen from Figure 7-6, even with the same system activity, the optimal operating point for the three modules are different due different module activities. The module activities, shown in Table 7.1, were obtained from program profiling the SPEC benchmark Espresso. Figure 7-6 suggests that as the system activity changes, the V_{DD} and V_T for each module should be adaptively changed to achieve energy efficient computation. To demonstrate this point, Figure 7-7 shows the normalized energy as a function of V_{DD} (V_T is changed as V_{DD} is varied to maintain the 100 MHz clock frequency) of an adder for different system activities. If the V_{DD} and V_T are fixed, based on optimizing for χ=1, then as the system activity decreases (e.g. χ=0.01), the adder would be expending 23 times more energy than if it were operating at the optimized V_{DD} and V_T for χ=0.01. If the system activity is known for running a particular application, then the V_{DD} and V_T can be electronically set for that application run. In this scheme, the V_{DD} and

V_T are held constant through the computation. This is different from the dynamic V_T control scheme.

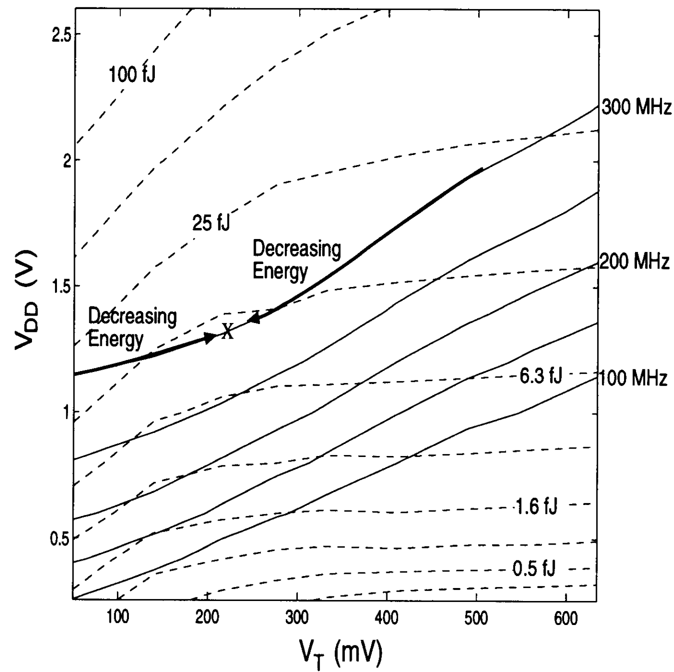


Figure 7-5: Energy and performance contours in V_T and V_{DD} space for an adder with 0.25 system activity, and the module activity as obtained by profiling a SPEC benchmark.

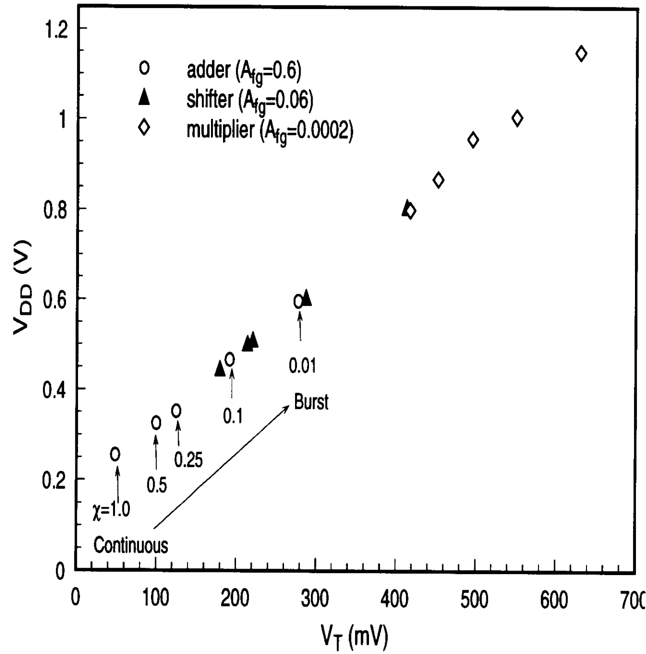


Figure 7-6: The optimal V_T and V_{DD} for the adder, shifter and multiplier modules with different system activity factors. The module activity factor, A_{fg}, for each functional unit was obtained from program profiling of the SPEC benchmark Espresso.

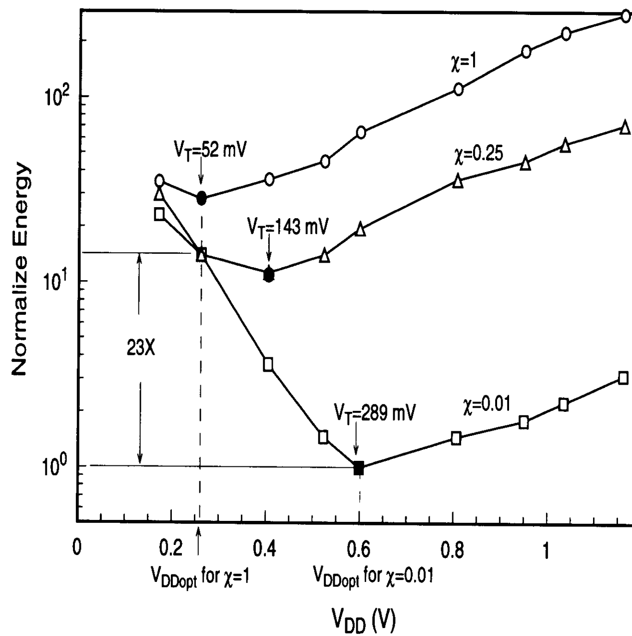


Figure 7-7: Comparison of normalized energy versus V_{DD} for different system activities.

7.3 Dynamic Threshold Voltage Control

Systems which operate in the burst-mode, i.e. those with very low activity, suffer from high static leakage energy dissipation at constant low V_T. The SOIAS technology with dynamic threshold voltage control can minimize this energy component by raising the V_T when the system is idle. From Figure 7-6, the optimal V_{DD} and V_T increases as the system activity increases. The multiplier, for example, has very low module activity, therefore, its optimum operating points are higher than both the adder and shifter, for the same clock frequency. By using the measured switching and static leakage energies in the SOIAS energy consumption model, the optimal V_{DD} and V_T can be found for dynamic switching of the V_T. The switch in the V_T was 250 mV. Module and back-gate activity factors were obtained from profiling of a SPEC benchmark program. Table 7.1 lists the module and back-gate activity factors for the adder, shifter and multiplier. Table 7.2 compares the optimal V_T and V_{DD} and energy for constant and dynamic switching of the V_T. Dynamic switching of the V_T allows reduction of optimal V_T and V_{DD} without suffering from high static leakage energy, and thus provides significant energy savings while maintaining the same performance. The optimal V_Ts shown in Table 7.2 are the low V_Ts during the active cycles of the module. The high V_Ts are 250 mV above those listed in the table. The dynamic switching of the V_T is the second level of optimization in addition to that of the adaptively set V_{DD} and V_T for burst-mode applications.

Table 7.1: Module activity factor, A_{fg}, and back-gate activity factor, A_{bg} obtained from profiling the SPEC benchmark program.

	A_{fg}	A_{bg}
Adder	0.6039	0.1954
Shifter	0.0633	0.0541
Mult.	0.0002	0.0002

Table 7.2: The optimal V_T and V_{DD} at minimum total energy for some modules as obtained from the energy consumption model and measured static and dynamic energies for burst-mode, χ=0.01, 100 MHz operation.

	Optimal	Optimal	Normalized
	V_T (mV)	V_{DD} (V)	Total Energy
Const. V_T Adder	277	0.59	1
Var. V_T Adder	51	0.26	0.21
Const. V_T Shifter	413	0.81	1
Var. V_T Shifter	139	0.37	0.23
Const. V_T Mult.	629	1.15	1
Var. V_T Mult.	333	0.69	0.4

Chapter 8

Summary

In this thesis, the development of the SOIAS technology for ultra-low voltage and low power applications was presented. This is a novel technology that adds a third dimension to the conventional planar MOSFET device structure. The development of this technology involved four stages: (1) substrate fabrication and material characterization; (2) device fabrication and characterization; (3) circuit fabrication; and (4) energy evaluation for low power applications.

The SOIAS wafer preparation involved wafer bonding of various materials. Based on the current understanding of silicon wafer bonding, we successfully bonded amorphous silicon to thermal oxide. Two different approaches were taken for the fabrication of SOIAS substrates. The BESOI approach yielded device quality wafers with excellent material qualities. The mobility, intrinsic oxide breakdown, and interface state density are all comparable if not better than bulk. One drawback of this technique is the poor control of silicon film thickness uniformity. An alternative approach taken to improve the silicon film thickness uniformity is the bonding of SIMOX wafers. The bonded SIMOX method yielded device quality wafers with slightly worse material qualities compared to its bulk counterpart. This was due to the roughness of the surface upon which the devices were built. With touch polishing of the bonded SIMOX SOIAS wafer surface prior to device fabrication, the material quality should improve.

CMOS device fabrication on both BESOI and bonded SIMOX wafers were suc-

cessful. Proper back-gate and channel implant engineering yielded back-gates with sheet resistance in the 1-5 KΩ range, and low quiescent V_T. Independent V_T control of NMOS and PMOS was verified through device and ring oscillator measurements. Quasi-static V_T control was verified with dynamic switching of device back-gate up to 20 MHz. The design range of 200 mV switch in V_T fits well within the measured V_T control range. For scaling the effective gate length down to 0.17 μm, the back-gate's control on the V_T does not show any degradation for the design parameters used in this work. Further scaling of the gate length will require the scaling of the back-gate oxide in order to maintain the same amount of control on the V_T. For sub-100 nm gate length fabrication, x-ray lithography was used. The integration of our x-ray lithography technology into our CMOS process was successful utilizing a mix-and-match scheme. The mix-and-match scheme was developed to match the optical, e-beam and x-ray lithography tools. The characterization of both positive and negative chemically-amplified resist processes for fabrication of 100 nm lines was completed. Devices with effective gate length of 83 nm was demonstrated in bulk.

Fabrication of an 8 bit adder and 8 bit multiplier was also successful. Functionality of these modules has been verified. The chip containing these arithmetic modules has approximately two thousand transistors, and the yield per wafer of this chip was approximately 50-60%. These circuits did not run at speed due to excessive poly routing. A silicide technology or two-level metal process should be implemented to address this problem.

Having developed the SOIAS technology with successful demonstration of working devices and circuits, a theoretical energy consumption model was developed. This energy consumption model was used to benchmark the SOIAS technology against the conventional SOI technology with constant low V_T operating at V_{DD} of 1 V. For burst-mode applications, significant energy savings are expected for dynamically controlled V_T as opposed to constant low V_T. Evaluation of the technology design space for burst-mode applications indicates that there is ample latitude in the back-gate design in terms of oxide thickness and source/drain overlap. When the back-gate is designed at the limit of the current fabrication tools, the energy savings is even

greater. The energy consumption model was also used in conjunction with measured dynamic and static leakage energies for exploring the V_{DD} and V_T design space. From this study, we have concluded that for energy efficient computing, V_{DD} and V_T should be adaptively changed as the system activity changes, and dynamic control of V_T during an application run lowers the optimal V_{DD} and V_T point for burst-mode computation.

Bibliography

[1] Y. Mii, Y. Taur, Y. Lii, D. Klaus, and J. Bucchignano, "An Ultra-Low Power 0.1 μm CMOS," in *Symposium on VLSI Technology*, (Honolulu, Hawaii), pp. 9–10, IEEE, 1994.

[2] D. Liu and C. Svensson, "Trading Power for Low Power by Choice of Supply and Threshold Voltages," *IEEE Solid-State Circuits*, vol. 28, no. 1, pp. 10–17, January 1993.

[3] A. Chandrakasan and R. Brodersen, "Minimizing Power Consumption in Digital CMOS Circuits," *IEEE Proceedings*, pp. 498–523, April 1995.

[4] M. Srivastava, A. Chandrakasan, and R. Brodersen, "Predictive System Shutdown and Other Architectural Techniques for Energy Efficient Programmable Computation," *IEEE Transactions on VLSI Systems*, Dec 1995.

[5] D. Takashima, S. Watanabe, K. Sakui, H. Nakano, and K. Ohuchi, "Standby/Active Mode Logic for sub-1 V 1G/4Gb DRAMS," in *ISSCC Digest of Technical Papers*, pp. 83–84, IEEE, 1994.

[6] T. Yabe, F. Matsuoka, K. Sato, S. Hyakawa, K. Sato, M. Matsui, A. Aono, H. Yoshimura, K. Ishimaru, H. Gojohbori, S. Morita, Y. Unno, M. Kakumu, and K. Ochii, "High-Speed and Low-Standby-Power Circuit Design of 1 to 5V Operating 1Mb Full CMOS SRAM," in *ISSCC*, pp. 107–108, IEEE, 1994.

[7] K. Seta, H. Hara, T. Kuroda, M. Kakumu, and T. Sakurai, "50% Active-power Saving without Speed Degradation using Standby Power Reduction (SPR) Circuit," in *ISSCC*, p. 318, IEEE, 1995.

[8] J. Colinge, M. Gao, A. Romano-Rodriguez, H. Maes, and C. Claeys, "Silicon-On-Insulator "Gate-All-Around-Device"," in *International Electron Device Meeting*, pp. 595–598, IEEE, 1990.

[9] F. Balestra, S. Cristoloveanu, M. Benachir, J. Brini, and T. Elewa, "Double-Gate Silicon-on-Insulator Transistor with Volume Inversion: A New Device with Greatly Enhanced Performance," *IEEE Electron Device Letters*, vol. EDL-8, no. 9, pp. 410–412, September 1987.

[10] T. Tanaka, K. Suzuki, H. Horie, and T. Sugii, "Ultrafast Low-Power Operation of p+-n+ Double-Gate SOI MOSFET," in *Symposium on VLSI Technology*, (Honolulu, Hawaii), pp. 11–12, IEEE, 1994.

[11] T. Tanaka, H. Horie, S. Ando, and S. Hijiya, "Analysis of P+ Poly Si Double-Gate Thin-Film SOI MOSFETs," in *International Electron Device Meeting*, (Washington D.C.), pp. 683–686, IEEE, 1991.

[12] J. Denton and G. Neudeck, "Fully Depleted Dual-Gated Thin-Film SOI P-MOSFET With an Isolated Buried Polysilicon Backgate," in *International SOI Conference*, (Tuscon, AZ), pp. 135–136, IEEE, 1995.

[13] T. Sekigawa and Y. Hayashi, "Calculated Threshold-Voltage Characteristics of an XMOS Transistor Having an Additional Bottom Gate," *Solid-State Electronics*, vol. 27, no. 8, pp. 827–828, 1984.

[14] S. Venkatesan, G. Neudeck, and R. Pierret, "Dual-Gate Operation and Volume Inversion in n-Channel SOI MOSFET's," *IEEE Electron Device Letters*, vol. 13, no. 1, p. 44, January 1992.

[15] R.-H. Yan, A. Ourmazd, and K. Lee, "Scaling the Si MOSFET: From Bulk to SOI to Bulk," *IEEE Transactions on Electron Devices*, vol. 39, no. 7, p. 1704, July 1992.

[16] C. Fiegna, H. Iwai, T. Wada, E. Sangiorgi, and B. Ricco, "A New Methodology For The 0.1-0.025 μm MOSFET," in *Symposium on VLSI Technology*, p. 33, IEEE, 1992.

[17] K. Suzuki, T. Tanak, Y. Tosaka, H. Horie, and Y. Arimoto, "Scaling Theory for Double-Gate SOI MOSFET's," *IEEE Transactions on Electron Devices*, vol. 40, no. 12, p. 2326, December 1993.

[18] H.-S. Wong, D. Frank, Y. Taur, and J. Stork, "Design and Performance Considerations for Sub-0.1 μm Double-Gate SOI MOSFET's," in *International Electron Device Meeting*, (San Francisco), p. 747, IEEE, 1994.

[19] K. Suzuki, Y. Tosaka, and T. Sugii, "Analytical Threshold Voltage Model for Short Channel n^+-p^+ SOI MOSFET's," *IEEE Transactions on Electron Devices*, vol. 43, no. 5, p. 732, May 1996.

[20] F. Assaderaghi, D. Sinistsky, S. Parke, J. Bokor, P. Ko, and C. Hu, "A Dynamic Threshold Voltage MOSFET (DTMOS) for Ultra-Low Voltage Operation," in *International Electron Device Meeting*, pp. 809–812, IEEE, 1994.

[21] Y. Tsividis, *Operation and Modeling of the MOS Transistor*. McGraw-Hill, 1987.

[22] H. Lim and J. Fossum, "Threshold Voltage of Thin-Film Silicon-onInsulator (SOI) MOSFET's," *IEEE Transactions on Electron Devices*, vol. ED-30, no. 10, pp. 1244–1251, October 1983.

[23] J. Colinge, *Silicon-On-Insulator Technology: Materials to VLSI*. Kluwer Academic Publishers, 1991.

[24] S. Veeraraghavan and J. G. Fossum, "A Physical Short-Channel Model for the Thin-Film SOI MOSFET Applicable to Device and Circuit CAD," *IEEE, Transactions on Electron Devices*, vol. 35, no. 11, p. 1866, November 1988.

[25] T. Nakamura, 1966. U.S. Patent 3.228.656.

[26] J. Lasky, S. Stiffler, F. White, and J. Abernathey, "Silicon-On-Insulator (SOI) by Bonding and Etch-Back," in *International Electron Device Meeting*, (San Francisco), p. 684, IEEE, 1985.

[27] M. Shimbo, K. Furukawa, K. Fukuda, and K. Tanzawa, "Silicon-to-Silicon Direct Bonding Method," *Journal of Applied Physics*, vol. 60, no. 8, p. 2987, October 1986.

[28] T. Abe and Y. Nakazato, "Thinned Silicon Layers on Oxide Film, Quartz and Sapphire by Wwafer Bonding," *IEICE Trans. Electron.*, vol. E77-C, no. 3, p. 342, March 1994.

[29] K. Ljungber, A. Soderbarg, A. Tiensuu, S. Johansson, G. Thugstrom, and C. S. Petersson, "Buried Cobalt Silicide Layers in Silicon Created by Wafer Bonding," *Journal of Electrochemical Society*, vol. 141, no. 10, p. 2829, October 1994.

[30] A. Plettner, K. Haberger, and K. Neumier, "Improved BESOI Substrates for High Speed ICs," in *International SOI Conference*, p. 70, IEEE, 1995.

[31] S. Bengtsson, "Semiconductor Wafer Bonding: A Review of Interfacial Properties and Applications," *Journal of Electronic Materials*, vol. 21, no. 8, p. 841, 1992.

[32] W. Kern and D. Puotinen, "Cleaning Solutions Based on Hydrogen Peroxide for use in Silicon Semiconductor Technology," *RCA Review,*, p. 187, 1970.

[33] Y. Backlund, K. Hermansson, and L. Smith, "Bond Strength Measurements Related to Silicon Surface Hydrophilicity," in *Proceedings of the First International Symposium on Semiconductor Wafer Bonding: Science, Technology and Applications*, (Pennington, NJ), p. 82, Electrochem. Soc., 1992.

[34] S. Bengtsson and O. Engstrom, "Low-Temperture Preparation of Silicon/Silicon Interfaces by the Silicon-to-Silicon Direct Bonding Method," *Journal of Electrochem. Soc.*, vol. 137, no. 7, p. 2297, July 1990.

[35] J. Haisma, G. Spierings, U. Biermann, and J. Pals, "Silicon-On-Insulator Wafer Bonding-Wafer Thinning Technological Evaluations," *Japanese Journal of Applied Physics*, vol. 28, p. 1426, 1989.

[36] W. Maszara, G. Goetz, A. Caviglia, and J. McKitterick, "Bonding of Silicon Wafers for Silicon-On-Insulator," *Japanese Journal of Applied Physics*, vol. 64, no. 10, p. 4943, November 1988.

[37] P. Barth, "Silicon Fusion Bonding for Fabrication of Sensors, Actuators and Microstructures," *Sensors and Actuators*, vol. A21-A23, p. 919, 1990.

[38] T. Abe, T. Takei, and A. Uchiyama, "Silicon Wafer Bonding Mechanism for Silicon-On-Insulator Structures," *Japanese Journal of Applied Physics*, vol. 29, no. 12, p. L2311, December 1990.

[39] R. Stengl, K.-Y. Ahn, and U. Gösele, "A Model for the Silicon Wafer Bonding Process," *Japanese Journal of Applied Physics*, vol. 28, p. 1735, 1989.

[40] M. Horiuchi and S. Aoki, "Characteristic of Silicon Wafer-Bond Strengthening by Anneling," *Journal of Electrochem. Soc.*, vol. 139, no. 9, p. 2589, September 1992.

[41] V. Watt and R. Bower, "Low Temperature Direct Bonding of Non-hydrophilic Surfaces," *Electronics Letters*, vol. 30, no. 9, p. 693, April 1994.

[42] S. Bengtsson and O. Engstrom, "Interface Charge Control of Directly Bonded Silicon Structures," *Japanese Journal of Applied Physics*, vol. 66, no. 3, p. 1231, August 1989.

[43] H. Himi, M. Matsui, S. Fujino, and T. Hattori, "Silicon Wafer Direct Bonding without Hydrophilic Native Oxides," *Japanese Journal of Applied Physics*, vol. 33, Part 1, no. 1A, p. 6, January 1994.

[44] T. Abe, A. Uchiyama, and K. Yoshizawa, "Surface Impurity Encapsulated by Silicon Wafer Bonding," *Japanese Journal of Applied Physics*, vol. 29, no. 12, p. L2315, December 1990.

[45] K. Mitani, V. Lehmann, R. Stengl, D. Feijoo, U. Gösele, and H. Z. Massoud, "Causes and Prevention of Temperature-Dependent Bubbles in Silicon Wafer Bonding," *Japanese Journal of Applied Physics*, vol. 30, no. 4, p. 615, April 1990.

[46] V. Lehmann, U. Gösele, and K. Mitani, "Contamination Protection of Semiconductor Surfaces by Wafer Bonding," *Solid State Technology*, p. 91, April 1990.

[47] Q.-Y. Tong, X.-L. Xu, and H. Shen, "Diffusion and Oxide Viscous Flow Mechanism in SDB Process and Silicon Wafer Rapid Tthermal Bonding," *Electronics Letters*, vol. 26, p. 697, 1990.

[48] X.-L. Xu, J. Zhan, and Q.-Y. Tong, "Silicon on Quartz by Solid-State Diffusion Bonding (SSDB) Technology," *Electronics Letters*, vol. 24, p. 691, 1988.

[49] P. Gillis and J. Gilman, "Double-Cantilever Cleavage Mode of Crack Propagation," *Journal of Applied Physics*, vol. 35, p. 647, 1964.

[50] G. Cha, W.-S. Yang, D. Feijoo, W. Taylor, R. Stengl, and U. Gösele, "Silicon Wafers with Cavities Bonded in Different Atmospheres," in *First Int. Symp. on Semiconductor Wafer Bonding* (E. Society, ed.), (Phoenix, AZ), Electrochemical Society, 1992.

[51] "Accu-Thin^{TM} is a trademark technology of Hughes."

[52] Q.-Y. Tong and U. Gösele, "Fabrication of Ultrathin SOI by SIMOX Wafer Bonding (SWB)," *Journal of Electronic Materials*, vol. 22, no. 7, p. 763, 1993.

[53] Q.-Y. Tong, U. Gösele, T. Martini, and M. Reiche, "Ultrathin Single-crystalline Silicon on Quartz (SOQ) by 150 ^{o}c Wafer Bonding," *Sensors and Actuators*, vol. A 48, p. 117, 1995.

[54] U. Schnakenberg, W. Benecke, and P. Lange, "TMAHW Etchants For Silicon Micromachining," in *TRANSDUCERS, International Conference on Solid-State Sensors and Actuators*, (San Francisco, CA), p. 815, IEEE, 1991.

[55] P. Heremans, J. Witters, G. Groeseneken, and H. Maes, "Analysis of the Charge Pumping Technique and Its Application for the Evaluation of MOSFET Degradation," *IEEE Transactions on Electron Devices*, vol. 36, no. 7, p. 1318, 1989.

[56] P. Heremans, H. Maes, and N. Saks, "Evaluation of Hot Carrier Degradation of N-Channel MOSFET's with the Charge Pumping Technique," *IEEE Electron Device Letters*, vol. EDL-7, no. 7, p. 428, 1986.

[57] T. Ouisse, S. Cristoloveanu, T. Elewa, H. Haddara, G. Borel, and D. Iannou, "Adaptation of the Charge Pumping Technique to Gated p-i-n Diodes Fabricated on Silicon on Insulator," *IEEE Transactions on Electron Devices*, vol. 38, no. 6, p. 1432, 1991.

[58] R. Feder, E. Spiller, and J. Topalian *Journal of Vacuum Science and Technology*, vol. 12, p. 1332, 1975.

[59] B. Lin, "A Comparison of Projection and Proximity Printings–From UV to X-ray," *Microelectronics Engineering*, vol. 11, p. 137, 1990.

[60] K. Early, M. Schattenburg, and H. Smith, "Absence of Resolution Degradation in X-ray Lithography for λ from 4.5 nm to 0.83 nm," *Microelectronics Engineering*, vol. 11, pp. 317–321, 1990.

[61] W. Chu, H. Smith, and M. Schattenburg, "Replication of 50 nm Linewidth Device Patterns using Proximity X-ray Lithography at Large Gaps," *Applied Physics Letters*, vol. 59, pp. 1641–1643, 1991.

[62] W. Chu, H. Smith, S. Rishton, D. Kern, and M. Schattenburg, "Fabrication of 50 nm Line-and-space X-ray Masks in Thick Au Using a 50 kev Electron Beam System," *J. Vac. Sci. Technol. B*, vol. 10, no. 6, p. 118, 1992.

[63] A. Moel, M. Schattenburg, J. Carter, and H. I. Smith, "A Compact, Low Cost System for Sub-100 nm X-ray Lithography," *J. Vac. Sci. Technol. B*, vol. 8, p. 1648, 1990.

[64] I. Y. Yang, S. Silverman, J. Ferrera, K. Jackson, J. Carter, D. A. Antoniadis, and H. Smith, "Combining and Matching Optical, Electron-beam, and X-ray Lithographies in the Fabricaton of Si Complementary Metal-Oxide-Semiconductor Circuits with 0.1 and sub-0.1 μm Features," *Journal of Vac. Sci. Technol.*, vol. B 13, no. 6, p. 2741, Nov 1995.

[65] E. Moon, P. Everett, and H. I. Smith, "Immunity to Signal Degradation by Overlayers Using a Novel Spatial-Phase-Matching Alignment System," *Journal of Vac. Sci. Technol.*, vol. B 13, no. 6, p. 2648, Nov 1995.

[66] K. Deguchi, K. Miyoshi, H. Ban, H. Kyuragi, and S. Konaka, "Application of X-ray Lithography with a Single-Layer Resist Process to Subquartermicron Large Scale Integrated Circuit Fabrication," *Journal of Vacuum Science and Technology*, vol. B 10, no. 6, p. 3145, November 1992.

[67] A. Pomerene, K. Petrillo, and D. Seeger, "Optimization of ESCAP Photoresist for X-ray Lithography," *SPIE*, vol. 2194, p. 162, 1994.

[68] I. Y. Yang, H. Hu, L. T. Su, V. V. Wong, M. B. Burkhardt, E. E. Moon, J. Carter, D. Antoniadis, H. I. Smith, K. W. Rhee, and W. Chu, "High Performance Self-aligned Sub-100 nm Metal-Oxide Semiconductor Field-Effect Transistors using X-ray Lithography," *J. Vac. Sci. Technol. B*, vol. 12, no. 6, p. 4051, 1994.

[69] Courtesy of Don Hofer IBM Almaden.

[70] I. Y. Yang, C. Vieri, A. Chandrakasan, and D. A. Antoniadis, "Back-gated CMOS on SOIAS For Dynamic Threshold Voltage Control," in *IEDM*, (Washington D.C.), p. 877, IEEE, 1995.

[71] L. Su, M. Sherony, H. Hu, J. Chung, and D. Antoniadis, "Optimization of Series Resistance in Sub-0.2 μm SOI MOSFETs," in *IEDM*, (Washington D.C.), p. 723, IEEE, 1993.

[72] K. Terada and H. Muta, "A New Method to Determine Effective MOSFET Channel Length," *Jap. J. Appl. Phys.*, vol. 18, no. 5, p. 593, 1979.

[73] H. Hu, L. Su, I. Y. Yang, D. Antoniadis, and H. I. Smith, "Channel and Source/Drain Engineering in Hight-Performance Sub-0.1μm nMOSFETs Using X-ray Lithography," in *Symposium on VLSI Technology*, (Honolulu, Hawaii), p. 17, IEEE, 1994.

[74] A. Srivastava and A. Eustace, "ATOM: A System for Building Customized Program Analysis Tools," tech. rep., 1994.

[75] C. Vieri, I. Y. Yang, C. Chandrakasan, and D. A. Antoniadis, "SOIAS: Dynamically Variable Threshold SOI with Active Substrate," in *Symposium on Low Power Electronics*, p. 86, IEEE, 1995.

[76] E. W. Greeneich, "An Analytical Model for the Gate Capacitance of Small-Geometry MOS Structures," *IEEE, Transactions on Electron Devices*, vol. ED-30, no. 12, p. 1838, December 1983.

[77] V. Gutnik and A. Chandrakasan, "An Efficient Controller for Variable Supply Voltage Low Power Processing," in *VLSI Circuits Symposium*, IEEE, June 1996.

Appendix A

MTL Low Power SOIAS Back-gated CMOS Process

This process flow only represents the core flow for our SOIAS CMOS lots for thin silicon films, i.e. the implant conditions were simulated via Suprem3 for 40 nm silicon films. For every run, splits were done for different silicon film thicknesses, and different V_Ts on the same wafer. Different silicon film thicknesses are achieved with different oxidation thinning conditions. Film thicknesses ranged from 35-100 nm. Usually 3-4 V_T implant splits are done on each wafer with target values ranging from 0.0-0.8 V. In addition, regular SIMOX wafers with the same splits, and bulk wafers are always carried along as controls. For our bulk CMOS, we use a single well process which is different from the MTL baseline. The key steps that differ are the N-well implant, well drive-in, field implant, field oxidation, and punch-through implant conditions. The V_T splits on bulk wafers are the same as those of the thick film SOIAS/SOI wafers ($\sim$ 100 nm).

Process Step	Description
Diffusion (1)	**Silicon Film Oxidation Thinning**
	Recipe 114
Lithography (1)	**Alignment Marks Pattern**
	First level alignment marks
Dry Etch (1)	**Plasma Silicon Etch to Define Alignment Marks**
	Recipe 10, timed etch for 20 seconds
Ashing (1)	**O_2 Plasma Resist Strip**
Lithography (2)	**NMOS Back-gate Pattern**
	Align to first level alignment marks
Implant (1)	**Boron Back-gate Implant (for NMOS)**
	Boron: 1.3e15 cm^{-2}, 130 KeV
Ashing (2)	**O_2 Plasma Resist Strip**
Lithography (3)	**PMOS Back-gate Pattern**
	Align to first level alignment marks
Implant (2)	**Phosphorus Back-gate Implant (for PMOS)**
	Phosphorus: 6.5e14 cm^{-2}, 320 KeV
Ashing (3)	**O_2 Plasma Resist Strip**
Clean (1)	**Piranha Clean**
	3:1 H_2SO_4:H_2O_2 for 10 minutes
Diffusion (2)	**Stress-Relief-Oxide-22 nm**
	Recipe 240

Diffusion (3) **LPCVD-Silicon-Nitride**

Recipe 410; 800 °C deposition temperature

Lithography (4) **Active-Area-Pattern**

Align to first level alignment marks.

Dry Etch (2) **Nitride-Plasma-Etch**

Recipe 15

Ashing (4) **O_2 Plasma Resist Strip**

Lithography (5) **NMOS Field Implant Pattern**

Align to diffusion level

Implant (3) **NMOS Field Implant**

BF_2: 5e13 cm^{-2}, 60 KeV

Ashing (5) **O_2 Plasma Resist Strip**

Lithography (6) **PMOS Field Implant Pattern**

Align to diffusion level.

Implant (4) **PMOS Field Implant**

Phosphorus: 5e12 cm^{-2}, 35 KeV

Ashing (6) **O_2 Plasma Resist Strip**

Clean (2) **Piranha**

Diffusion (4) **Field Oxidation**

Recipe 114

Wet Etch (1) **Nitride Strip**

180 °C phosphoric acid for 40 minutes

Wet Etch (2)	**Strip 22 nm SRO**
	Done in the 50:1 HF in the RCA prior to dummy oxidation
Diffusion (5)	**Dummy Oxide Growth**
	Recipe 110
Lithography (7)	**NMOS V_T Implant Pattern**
	Align to diffusion level
Implant (5)	**NMOS V_T Implants**
	BF_2: 2.8e11 cm^{-2}, 25 KeV
Ashing (7)	**O_2 Plasma Resist Strip**
Lithography (8)	**PMOS V_T Implant Pattern**
	Align to diffusion level
Implant (6)	**PMOS V_T Implant**
	Phosphorus: 1.7e12 cm^{-2}, 25 KeV
Ashing (8)	**O_2 Plasma Resist Strip**
Wet Etch (3)	**Dummy Gate Removal**
	Done with 50:1 HF dip in RCA prior to gate oxidation
Diffusion (6)	**7 nm Gate Oxidation**
	Recipe 226
Diffusion (7)	**LPCVD Amorphous Silicon (230 nm) Deposition**
	Recipe 705; 560 ^oC deposition temperature
Diffusion (8)	**LPCVD Nitride (60 nm) Deposition**
	Nitride hardmask for gate etching

Lithography (9) **Gate Pattern**

Align to diffusion level

Done either optically or by x-ray lithography

Dry Etch (3) **RIE in NSL**

CHF_3 nitride hard mask definition

Clean (3) **Piranha in TRL**

Dry Etch (4) **Polysilicon Plasma Etch (230 nm)**

Modified recipe 32

Diffusion (9) **Reoxidation**

Recipe 226

Lithography (10) **Back-gate Contact Pattern**

Align to diffusion level

Dry Etch (5) **Plasma Si Etch To Contact The Back-gate**

Modified Recipe 10; timed etch for 30 seconds

Wet Etch (4) **Back-gate Oxide Etch**

Use a dummy 100 nm to monitor etch rate

Lithography (11) **N+ Poly-Source/Drain-Pattern**

Align to gate level

Implant (7) **N+ Poly-Source/Drain-Implant**

Arsenic: 4e15 cm^{-2}, 25 KeV

Ashing (9) **O_2 Plasma Resist Strip**

Clean (4) **Piranha**

Diffusion (10) **RTA Anneal**

1000 ^{o}C for 10 seconds

Lithography (12)	**P+ Source/Drain-Pattern**
	Align to gate level
Implant (8)	**P+ Source/Drain-Implant**
	BF_2, 3e15 cm^{-2}, 20 KeV
Ashing (10)	**O_2 Plasma Resist Strip**
Clean (5)	**Piranha**
Diffusion (11)	**RTA Anneal**
	950 ^oC 5 seconds
Diffusion (12)	**LTO Deposition (20 nm)**
	Nitride spacer etch stop; Modified recipe 436
	425 ^oC deposition temperature
Diffusion (13)	**LPCVD Nitride Deposition (150 nm)**
	Nitride spacer
Dry Etch (6)	**Nitride Plasma Etch**
	Recipe 15; nitride spacer etch
Clean (6)	**SPECIAL-RCA-CLEAN**
	strip LTO in RCA 50:1 HF until dewet
Deposition (1)	**E-Beam Evaporation**
	Cobalt: 6.5 nm Ti: 3.5 nm
Diffusion (14)	**RTA Anneal**
	550 ^oC, 30 seconds
Wet Etch (5)	**Ti-Cobalt-Strip**
	Piranha strip in dedicated quartzeware

Diffusion (15)	**RTA Anneal**
	700 °C, 60 seconds
Diffusion (16)	**LTO Deposition (400 nm)**
	Recipe 439; passivation oxide
Lithography (13)	**Top Resist Coat**
	Do a hardbake. Recipe 82
Wet Etch (6)	**400 nm Backside LTO Removal**
	7:1 BOE until dewet
Dry Etch (7)	**Backside Nitride Removal**
	Recipe 15
Wet Etch (7)	**20 nm Backside LTO Removal**
Dry Etch (8)	**Backside Poly Removal**
	Recipe 12
Wet Etch (8)	**Backside Oxide Removal**
Ashing (11)	**O_2 Plasma Resist Strip**
Lithography (14)	**Contact-Pattern**
	Align to gate level
Dry Etch (9)	**300 nm LTO Etch**
	Recipe 20; contact cuts
Wet Etch (9)	**100 nm LTO Removal**
	7:1 BOE dip to clear the contact cuts
Ashing (12)	**O_2 Plasma Resist Strip**

Deposition (2) **Metal-Deposition**

Al with 1% Si

Lithography (15) **Metal-Pattern**

Align to contact level

Dry Etch (10) **Metal Etch**

Recipe 32

Ashing (13) **O_2 Plasma Resist Strip**

Diffusion (17) **Sinter**

Recipe 710; 400 ^oC

Table A.1: Recipe 110

800 ^oC	30 min	DRY O_2
800	25	N_2

Table A.2: Recipe 114

800 ^oC	10 min	N_2
950	15	N_2
950	30	DRY O_2
950	FF	WET O_2
950	30	DRY O_2
800	70	N_2

FF is the variable time interval that can be set to yield the right oxide thickness.

Table A.3: Recipe 226

800 °C	10 min	N_2
900	20	N_2
900	FF	DRY O_2
900	15	N_2
800	40	N_2

Table A.4: Recipe 230

800 °C	25 min	N_2
950	10	N_2
950	37	DRY O_2
950	30	N_2
800	60	N_2

Table A.5: Recipe 10

	Step 1	Step 2	Step 3	Step 4	step 5
Press.	500 mT	500 mT	200 mT	200 mT	200 mT
RF	0 W	300 W	0 W	300 W	200 W
Gap	1.5 cm	1.5 cm	1.5 cm	1.5 cm	1.5 cm
CCl_4	0 sccm	0 sccm	130 sccm	130 sccm	130 sccm
O_2	200 sccm	200 sccm	20 sccm	20 sccm	20 sccm
He	100 sccm	100 sccm	70 sccm	70 sccm	130 sccm
	Stability	6 seconds	Stability	1 min 30 sec. Endpt	Overetch 25 %

Table A.6: Recipe 15

	Step 1	Step 2	Step 3	Step 4	step 5
Press.	500 mT	500 mT	375 mT	375 mT	375 mT
RF	0 W	300 W	0 W	250 W	100 W
Gap	1.5 cm	1.5 cm	1.0 cm	1.0 cm	1.5 cm
O_2	200 sccm	200 sccm	0 sccm	0 sccm	0 sccm
He	100 sccm	100 sccm	20 sccm	20 sccm	20 sccm
SF_6	0 sccm	0 sccm	50 sccm	50 sccm	60 sccm
	Stability	6 seconds	Stability	1 minute Endpt	10 seconds

Table A.7: Recipe 20

	Step 1	Step 2	Step 3	Step 4	step 5	step 6
Press.	3 T	3 T	3 T	3 T	3 T	3 T
RF	0 W	50 W	0 W	900 W	0 W	800 W
Gap	0.5 cm	0.5 cm	0.36 cm	0.36 cm	0.38 cm	0.38 cm
O_2	100 sccm	100 sccm	0 sccm	0 sccm	0 sccm	0 sccm
He	200 sccm	200 sccm	200 sccm	200 sccm	125 sccm	125 sccm
CHF_3	0 sccm	0 sccm	30 sccm	30 sccm	28 sccm	28 sccm
CF_4	20 sccm	20 sccm	90 sccm	90 sccm	25 sccm	25 sccm
	Stability	10 seconds	Stability	1 minute Endpt	Stability	10 seconds

Table A.8: Recipe 32

	Step 1	Step 2	Step 3	Step 4	step 5
Press.	220 mT	220 mT	220 mT	220 mT	220 mT
RF	0 W	250 W	0 W	175 W	175 W
$CHCl_3$	0 sccm	0 sccm	15 sccm	15 sccm	15 sccm
N_2	0 sccm	0 sccm	10 sccm	10 sccm	10 sccm
Cl_2	40 sccm	40 sccm	27 sccm	27 sccm	27 sccm
BCl_3	90 sccm	90 sccm	90 sccm	90 sccm	90 sccm
	Stability	10 seconds	Stability	1 min. 30 sec. Endpt	Overetch 50%

Table A.9: Modified recipe 32 for gate etch.

	Step 1	Step 2	Step 3	Step 4
Press.	50 mT	50 mT	50 mT	50 mT
RF	0 W	0 W	100 W	80 W
Cl_2	20 sccm	20 sccm	20 sccm	20 sccm
	Stability	10 seconds	1 min. 30 sec. Endpt	5 seconds