# Thin-Body Silicon FET Devices and Technology

by

Varadarajan Vidya

B.Tech. (Indian Institute of Technology, Bombay) 2001
M.Tech. (Indian Institute of Technology, Bombay) 2001

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering – Electrical Engineering
and Computer Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Tsu-Jae King Liu, Chair
Professor Chenming Hu
Professor Eugene E. Haller

Spring 2007

The dissertation of Varadarajan Vidya is approved:

| | |
|---|---|
| **Chair** | **Date** |
| | Date |
| | Date |

University of California, Berkeley

Spring 2007

Thin-Body Silicon FET Devices and Technology

Copyright 2007

by

Varadarajan Vidya

**Abstract**

# Thin-Body Silicon FET Devices and Technology

by

Varadarajan Vidya

Doctor of Philosophy in Engineering –
Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Tsu-Jae King Liu, Chair

Continued miniaturization of bulk silicon CMOS transistors is being limited by degrading short channel effects. Traditionally, higher channel doping, shallower source/drain junctions, and thinner gate dielectrics have been employed to improve gate control and enhance performance as the gate length is scaled down. However, these techniques are rapidly approaching material and process limits. Alternate transistor architectures such as the planar ultra-thin body (UTB) FET and double-gate MOSFET may be necessary to continue gate length scaling down to the sub-10nm regime.

The non-planar FinFET is perhaps the most promising double-gate structure for integrated circuit manufacture. However, some key fabrication issues that are yet to be resolved include super-steep source/drain junction formation with high uniformity and dual metal gate technology with work function engineering for dense layouts. Furthermore, statistical variations due to factors such as dopant fluctuation effects and line-edge roughness demand the need for variation-tolerant device and circuit designs.

One way to tackle the challenges in the formation of uniform and steep source/drain junctions in FinFETs is to completely eliminate all p-n junctions in the device. By having the source/drain and channel regions of the same doping type, the device can be operated in accumulation mode instead of inversion mode. It is found that this architecture shows comparable intrinsic delay and lower gate leakage compared to the enhancement mode FinFET, but higher sensitivity to process variations.

Single gate work function CMOSFET design is studied to make metal gate technology more practical for FinFETs. It is shown that high performance, low standby power, and low operating power CMOS can be implemented by utilizing the electrical channel length ($L_{eff}$) as a $V_T$ tuning variable, where, $L_{eff}$ is optimized through an optimum choice of side-wall spacer thickness ($L_{SP}$) and source/drain gradient abruptness ($\sigma_{SD}$). In this methodology, $L_{eff}$ and silicon fin thickness ($T_{Si}$) are the only optimization variables and a tradeoff between performance and variability is inherent to this device design scheme. Through 3D atomistic simulations, it is identified that lean spacers and steep junctions, along with a relatively thick $T_{Si}$ will be necessary to minimize variations.

The concept of thin-body MOSFETs is extended to three-dimensional integration through a novel, low thermal budget, cost-effective integration methodology. Unlike conventional techniques, the proposed method focuses on building FETs directly within interconnect wires. The technique has been demonstrated on the aluminum-silicon system using the concept of aluminum induced crystallization of silicon.

Professor Tsu-Jae King Liu
Dissertation Committee Chair

***To my husband and my family, for their love and encouragement***

# Table of Contents

Chapter 1 : Introduction ........ 1
1.1 Challenges of Bulk Silicon CMOS Scaling ........ 1
1.2 Advanced MOSFET Structures ........ 3
1.3 Interconnect Scaling and 3-D Integration ........ 7
1.4 Organization ........ 9
1.5 References ........ 11
Chapter 2 : Accumulation Mode Thin-body MOSFETs ........ 16
2.1 Introduction ........ 16
2.2 Thin-body Accumulation Mode MOSFETs ........ 18
2.2.1 Device Structure and Operation ........ 18
2.2.2 Potential Benefits and Issues ........ 21
2.3 Device Design ........ 21
2.3.1 Design Optimization ........ 22
2.3.2 Short Channel Effects and Sensitivities ........ 28
2.4 Device Fabrication Issues ........ 29
2.4.1 Gate Last Process ........ 30
Shortcomings of Gate Last Process ........ 33
2.4.2 Gate First Process ........ 34
2.5 Summary ........ 35
2.6 References ........ 36
Chapter 3 : $V_T$ Adjustment via $L_{eff}$ Engineering ........ 40
3.1 Introduction ........ 40
3.2 $V_T$ Tuning by S/D Engineering ........ 42
3.2.1 Device Simulation Set-up ........ 43
SRAM Noise Margin Variation Simulation Setup ........ 45
3.2.2 Impact of $L_{eff}$ on $I_{ON}$-$I_{OFF}$ ........ 47
3.2.3 Choice of a Good S/D Design: Impact of Variations ........ 51

3.2.4 Evaluating HP and LOP Designs with Single Gate Work Function ........... 54
$V_T$ Adjustment and Sensitivity to Variations ........ 55
3.3 S/D Process Requirements for $L_{eff}$ Engineering ........ 60
3.3.1 Effect of Ion-Implantation Conditions ........ 61
Effect of Tilt and Rotation ........ 62
Effect of Implant Energy ........ 63
3.4 Summary ........ 64
3.5 References ........ 65
Chapter 4 : Statistical Dopant Fluctuation Effects in FinFETs ........ 70
4.1 Introduction ........ 70
4.1.1 Sources of Variations ........ 70
4.2 3D Atomistic Simulation Setup ........ 72
4.2.1 Random Dopant Generation and Placement ........ 72
4.2.2 Dopant Definition in a Drift-Diffusion Simulator ........ 76
4.2.3 Complete 3D Atomistic Simulation Algorithm ........ 81
4.3 Variation Tolerant FinFET Design ........ 82
4.3.1 Continuum domain device optimization ........ 82
4.3.2 Design Strategies for Variation Tolerance to SDF ........ 84
Requirement of S/D Abruptness $\sigma_{SD}$ ........ 84
Optimization of $T_{Si}$ ........ 85
Shortcomings of the Current Approach ........ 87
4.4 Summary ........ 87
4.5 References ........ 88
Chapter 5 : WireFET Technology ........ 94
5.1 Introduction ........ 94
5.1.1 Concept of WireFET ........ 96
5.1.2 Initial Simulations ........ 96
5.2 Aluminum Induced Crystallization (AIC) of Silicon ........ 98
5.3 AIC Process Optimization ........ 100
5.3.1 Proof-of-Concept ........ 100
5.3.2 Role of Native Oxide Layer ........ 101
5.3.3 Optimum Si:Al Thickness Ratio ........ 103

5.3.4 TEM Analysis of AIC Polysilicon ........ 105
5.4 Improved WireFET Process ........ 106
5.4.1 Improved WireFET Process Flow ........ 106
5.4.2 Electrical Measurements ........ 107
5.4.3 Further Material Characterization ........ 110
5.5 Summary ........ 114
5.6 References ........ 115
Chapter 6 : Conclusions ........ 119
6.1 Summary ........ 119
6.2 Suggestions for Future Research ........ 123
Appendix A: Accumulation mode MOSFET Process Steps ........ 126
Gate Last Process ........ 126
Gate First Process ........ 131
Appendix B: Useful Codes for Atomistic Simulations ........ 136
Appendix C: WireFET Process Steps ........ 144

# Acknowledgements

I would like to express most sincere and heartfelt gratitude to my advisor, Prof. Tsu-Jae King Liu, for her excellent guidance and encouragement through all years of my doctoral research. Her vast knowledge and technical expertise immensely helped me in getting a good understanding of device engineering and identifying the relevant research problems. At a personal level, she has been very kind and considerate to me and I am especially thankful to her for allowing me to work from home during my final semester.

I would like to thank Prof. Chenming Hu and Prof. Eugene Haller for serving on my dissertation committee. I am grateful to Prof. Jan Rabaey for serving on my Qualifying Examination committee. His valuable feedback led to a significant portion of my doctoral work and made my research more relevant. I would also like to thank Prof. J. W. Morris and Prof. Nathan Cheung for serving on my Qualifying Examination committee.

I would like to thank the Microfabrication Laboratory staff for providing the infrastructure as well as technical support for running my processes. I am especially thankful to Joe Donnelly for all his timely efforts and ensuring smooth running of my processes in the Novellus tool. Thanks also to Jimmy Chang for providing practical solutions and ideas for implementing my processes in the Microlab. I would also like to thank Dr Xiaofan Meng for helping me with evaporations and AFM measurements.

I am very grateful to Hideki Takeuchi, without whose guidance, it would have been very difficult for me start working in the microlab. He has been very patient with me and has taught me almost all the tools I have used and all processes I have run. I am

indebted to him for his selfless efforts over many months, even after office hours, in helping me debug my processes and in coming up with the most practical solutions.

I would like to thank Dr. Lee Smith from Synopsys Inc., for sharing his extensive knowledge and guiding me with the development of the tool for atomistic simulations. I would like to thank Dr. Hsing-Huang Tseng from Sematech for giving me the wonderful opportunity to work with him as a summer intern and gain the knowledge on high-k gate dielectrics and their characterization. I would also like to thank Dr. Wade Xiong and Dr. Cleavelin Rin from ATDF for insightful discussions on single gate FinFET technology.

I am thankful to Ruth Gjerde for very patiently answering all my administrative queries and I am very grateful for her timely guidance through all the administrative responsibilities at all stages of my PhD. I would also like to thank Linda Manly, Charlotte Jones and Misty Kiuchi for ensuring smooth monetary transactions.

It has been a great pleasure to work with and share the friendship of my batchmates, Sriram Balasubramanian, Marie-Ange Naida Eyoum, Mohan Dunga, Blake Lin, Kyoungsub Shin, and Hiu-Yung Wong. I would also like to thank my seniors, Leland Chang, Pushkar Ranade, Igor Polishchuk, Qiang Lu, Yang-Kyu Choi, Kevin Yang, Charles Kuo, Yu (Kevin) Cao, Yee-chia Yeo, Shiying Xiong, Min She, Patrick Xuan, Jane Xi, and Daewon Ha for helping me get started with my research and gaining the relevant background. Thanks to all the current students at 373 Cory, Joanna Lai, Pankaj Kalra, Andrew Carlson, Donovan Lee, Chung-Hsun Lin, Alvaro Padilla, Noel Arellano, Darson Lu, Xin Sun, Reinaldo Vega, Dan Good, Hei (Anderson) Kam, Hui Wan, Kinyip Phoa, Yu-Chih Tseng, Anupama Bowonder, Pratik Patel, Cheuk Chi Lo, David Carlton, and Rhesa Nathaniel, for making my stay in Berkeley, an enjoyable and

unforgettable experience. I would like to thank Joanna for her timely and significant help with my final processes in the microlab. I would like to thank the visiting scholars and post doctoral members, Koichi Fukuda, Woo Young Choi, Wesley Chang, Yuri Yasuda, Takuro Matsutoya, Akira Hokazono and Taro Osabe for sharing their technical expertise with us. I am particularly thankful to Koichisan, Yurisan, Woo Young, Wesley, and Noel for their friendship and for making it easy to work long hours in the lab!

I am deeply indebted to my parents and my family for their love, their sacrifices, and encouragements all through my life. I am very grateful to my father for building my foundations in physics and to my brother for my foundations in mathematics. My mother has been most encouraging with all my endeavors in life and she always inspired me to work all the way towards a PhD and beyond. I am very thankful to my brother Venkatesh, and my Manni Ramyaa for making my stay in the Bay Area so comfortable and for being by my side through all good and bad times, from the day I started. I am also very fortunate to have the immense love of my little niece Prerana (Bumble), who has always been very caring and relieved all my stresses in moments! Finally, I would like to thank my dear husband, Srinath, for his love, for being with me in good and bad times and for walking me through my doctoral work. I am grateful to him for encouraging me, inspiring me and valuing my work over all the years, and of course, for helping me in sustaining through all the all-nighters in the lab! Without his programming expertise, it would have been very difficult for me to develop the atomistic simulator. I would also like to thank my in-laws Amma and Nannaru for their love and encouragement.

Most of all, I sincerely thank God for His blessings, for showing me the way and giving me the capability to receive a good education.

# Chapter 1 : **Introduction**

The journey of the modern day Integrated Circuit (IC) began in the late 1950s with the first ICs independently demonstrated by Jack Kilby (1958) and Robert Noyce (1959) [1], and has led to the present semiconductor industry. Since its conception, there has been an exponential scaling down of transistor dimensions (following the trend predicted by G. E. Moore [2]) targeted towards making the circuits smaller in order to fit more and more functionalities in a given chip area. The modern day circuits have over 100 million transistors [3] per chip and this exponential growth since early 1960s is expected to continue for at least another decade [4].

## 1.1 Challenges of Bulk Silicon CMOS Scaling

The main goals behind scaling transistors in a circuit are achieving (i) improved performance ($I_{ON}$) (ii) more on-chip functionality and hence, (iii) reduced cost per functionality. The transistors in ICs have been traditionally implemented using Bulk Silicon CMOS technology. As the lateral and vertical dimensions of the transistor are scaled, several effects come into play, which make further scaling of the traditional device architecture increasingly challenging.

As the gate length ($L_G$) of a transistor is decreased (lateral scaling), the source and drain regions come closer and the drain electric field starts reducing the source-channel

potential barrier. The capacitively coupled gate tends to lose control on the channel, especially in the sub-surface region. This capacitive coupling can be improved by decreasing the gate dielectric thickness (vertical scaling). The chief outcome of this is an increase in the off-state leakage ($I_{OFF}$) of the transistors which contribute to the total stand-by power in a circuit. The key sources of leakage include (i) subthreshold leakage and parasitic components like (ii) sub-surface leakage due to poor short channel control and (iii) gate leakage due to direct tunneling through the gate dielectric, which are all exponential functions of the dielectric thickness or the source-channel potential barrier height. Therefore, there has been an exponential increase in the standby power of chips (**Figure 1.1**) and modern day ICs can dissipate almost 50% power in standby leakage [3, 5]!

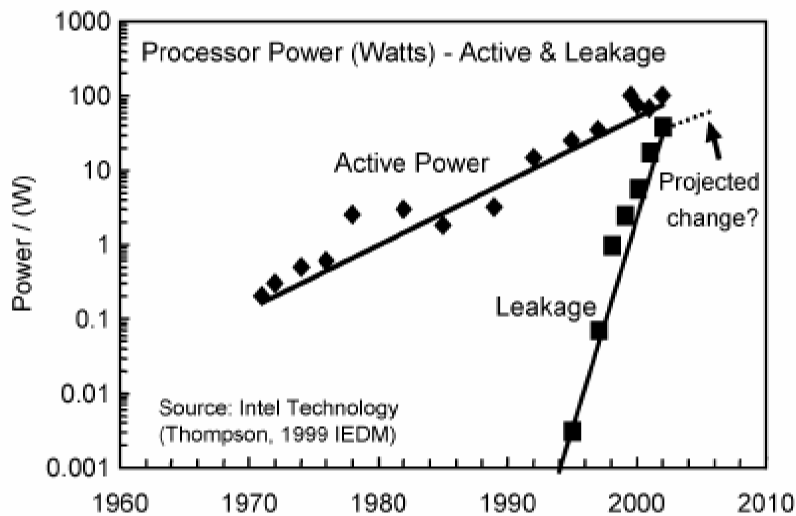



**Figure 1.1**: Trend in microprocessor active versus leakage power with scaling [5]. The leakage power is almost 50% of the total power on a modern-day chip.

Conventionally, the sub-surface leakage and subthreshold leakage components have been alleviated by increasing the channel doping (and hence the source-channel potential barrier). With scaling however, channel doping becomes a serious concern due

to statistical dopant fluctuation effects [6]. The gate leakage however, can be mitigated by the use of high-k gate dielectrics, though it faces multiple challenges in practical realization [7, 8].

Bulk CMOS technology faces other challenges as well. In order to keep short channel effects under control, ultra shallow junctions with very high doping abruptness and yet, high degrees of dopant activation, are required. Although methods such as laser annealing and flash lamp annealing are currently being investigated these may not work for future technology nodes. In addition, the poly-Si gate depletion effect contributes significantly towards the effective oxide thickness and hence the threshold voltage and performance. This effect can be completely eliminated by moving back to metal gate technology. NMOS and PMOS devices, however, need separate gate materials to achieve the required work functions, leading to process integration challenges.

Even though bulk CMOS technology with $SiO_2$ gate dielectric and poly-Si gates has been the most suitable and well-understood technology, the above mentioned challenges (and others not mentioned here) strongly push the need for alternate device structures and processing techniques.

## 1.2 Advanced MOSFET Structures

In order to mitigate some of the issues of the bulk CMOS MOSFET, advanced transistor structures such as ultra-thin body FET (UTB FET), Fully-depleted SOI (FD-SOI), Double-gate FET (DG-FET) like FinFET, Multi-Gate FET (MuGFET) and tri-gate FET (**Figure 1.2**) [9-13] have been studied as alternatives to the conventional device architecture. The key benefit of all these structures is that the conduction is confined to a thin silicon film, thereby physically eliminating the sub-surface leakage component. The

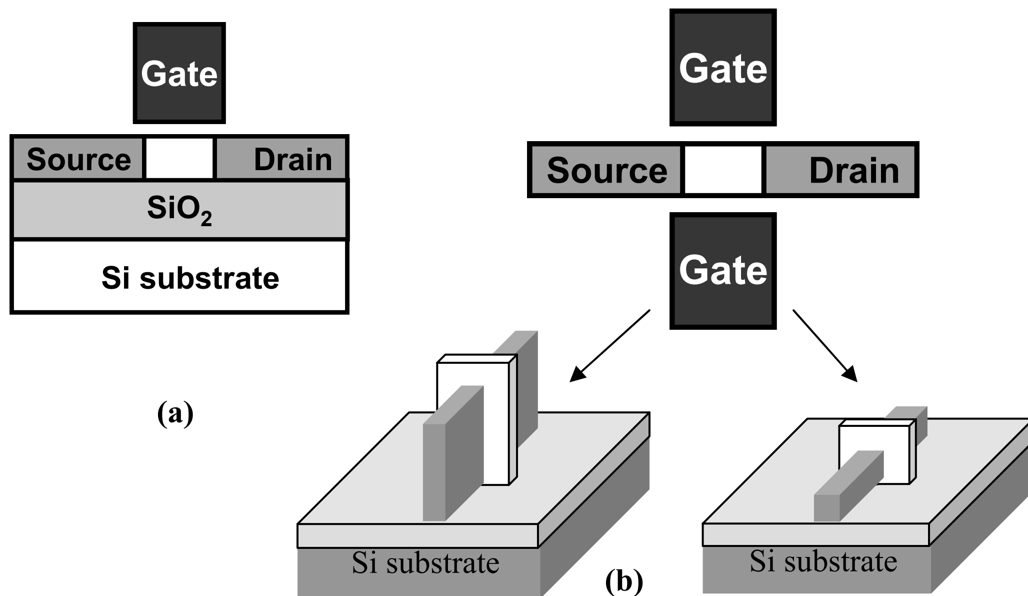



**Figure 1.2**: Schematic representation of advanced FET devices like the (a) Ultra-thin body MOSFET (UTB FET) and (b) Multiple-gate MOSFETs like FinFET and the Tri-gate MOSFET. Confining the channel to a thin silicon film provides better short channel control, performance and scalability compared to the bulk MOSFET.

layouts and process steps are very close to the conventional bulk CMOS flow. An undoped channel is used to reduce the effect of statistical dopant fluctuations. They also have additional benefits of better short channel control and reduced parasitic capacitance (no source-bulk or drain-bulk capacitances exist). Overall, these devices show superior performance ($I_{ON}/I_{OFF}$ as well as intrinsic delay) compared to their bulk counterpart [14, 15]. For reasonable performance, the body thickness requirement for a UTB-FET is about $L_G/3$, while that for a DG-FET is about $2L_G/3$ [16, 17]. This proves the better scalability of the DG-FET. The benefits of advanced FET structures are promising enough that the 2005 International Technology Roadmap for Semiconductors (ITRS [4]) predicts

possible introduction of these technologies in manufacturing in future technology nodes, summarized in **Figure 1.3**.

| $L_G$ (HP) | 2005 | 2007 | 2008 | 2011 | 2012 | 2015 | 2018 | 2020 |
|---|---|---|---|---|---|---|---|---|
| Bulk (nm) | 32 | 25 | 22 | 16 | 14 | - | - | - |
| UTB (nm) | - | - | 22 | 16 | 14 | 10 | - | - |
| DG (nm) | - | - | - | 16 | 14 | 10 | 7 | 5 |

**(a)**

| $L_G$ (LSTP) | 2005 | 2007 | 2008 | 2012 | 2013 | 2015 | 2018 | 2020 |
|---|---|---|---|---|---|---|---|---|
| Bulk (nm) | 65 | 45 | 37 | 22 | 20 | - | - | - |
| UTB (nm) | - | - | - | 22 | 20 | 17 | 14 | 12 |
| DG (nm) | - | - | - | 22 | 20 | 16 | 12 | 10 |

**(b)**

| $L_G$ (LOP) | 2005 | 2007 | 2008 | 2011 | 2012 | 2013 | 2016 | 2020 |
|---|---|---|---|---|---|---|---|---|
| Bulk (nm) | 45 | 32 | 28 | 20 | 18 | - | - | - |
| UTB (nm) | - | - | - | 20 | 18 | 16 | 11 | - |
| DG (nm) | - | - | - | 20 | 18 | 16 | 11 | 7 |

**(c)**

**Figure 1.3**: International Technology Roadmap predictions (of gate lengths used in high-volume IC manufacturing) [4] for the expected termination of bulk CMOS and introduction of advanced FETs for (a) High Performance, (b) Low Standby Power and (c) Low Operating Power applications.

Even though some of the issues of bulk CMOS technology get resolved in moving to advanced FETs, some challenges still question the practical implementation of these devices. Some of the issues are introduced here, though this is, by no means, a comprehensive list. For non-planar architectures, such as the FinFET, the topography poses an important integration challenge. In order to achieve uniform S/D doping along the height of the Si Fin, large angle tilted ion implantation or alternative doping strategies like plasma doping [18], will be needed. For dense memory layouts, like SRAM arrays it is impossible to have angular implants (due to shadowing effects), thereby, forcing the

need for introduction of alternate doping techniques. The p-n junction gradient requirements however, would be similar to the conventional MOSFET. Another issue that remains unresolved is the high-k dielectric with metal gate integration. With an undoped channel, NMOS and PMOS FinFETs require different gate work functions for optimum performance [15], which do not correspond to Si band edge values as in bulk MOSFETs, thus making it inevitable to introduce metal gate technology. Separate gate work functions for NMOS and PMOS can be obtained through gate work function engineering by selective ion implantation [19, 20]. However, with scaling and reduction in fin pitch, gate work function engineering would become impossible, and it would be necessary to implement single gate work function CMOS technology.

With scaling, we are quickly approaching material and process limits (no matter what device architecture is used) and this leads to a new inevitable problem arising out of statistical fluctuations in processes, which reveal themselves in the form of variations in the performance and power of a circuit. Some examples of the sources of variations include (i) statistical dopant fluctuations in the channel and/or source and drain [6, 21-23], (ii) Line-edge roughness of gate and fin coming from lithography and etching [24, 25], (iii) layout dependence on mechanical strain and performance [26] (iv) film thickness variations [6], (v) on-chip temperature gradients [27] and (vi) electrical stress related change in parameters [27]. Considering the complexity and cost of modern day chips, it is very important to quantify the overall variations and acquire design solutions prior to fabricating them. This requires significant and timely contribution from the Computer Aided Design (CAD) industry, and many of the above phenomena are yet to be modeled

accurately for complicated layouts. It is also important to identify variation-tolerant device designs, in conjunction with variation-tolerant circuit designs and architectures.

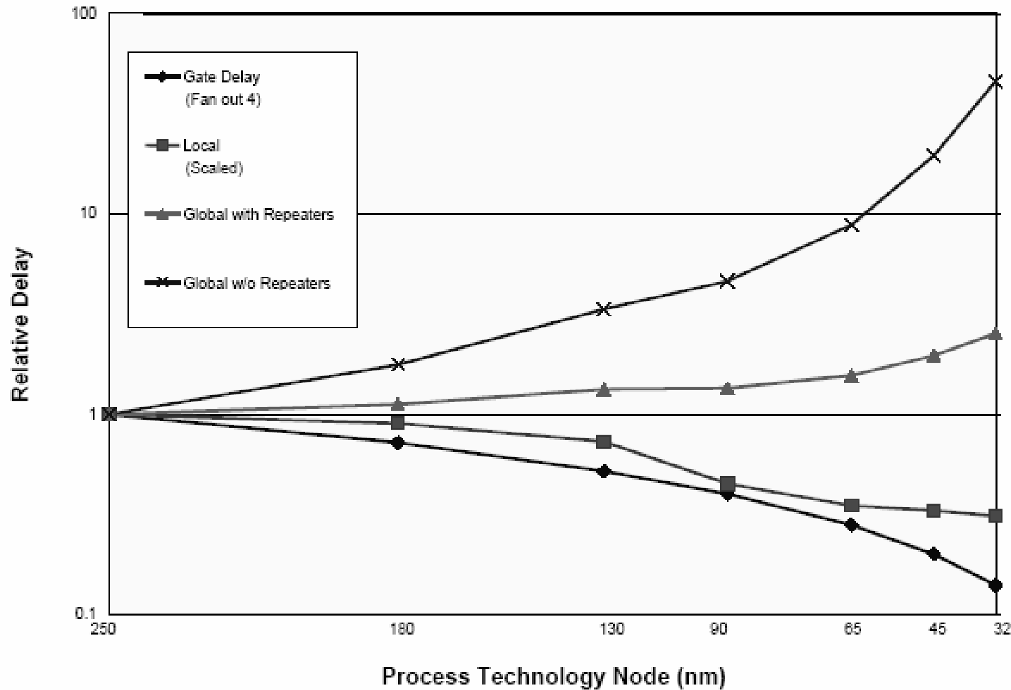



**Figure 1.4**: Non-scaling or reverse scaling of global interconnects makes their contribution to the total chip delay significant [29] with progressing technology nodes.

## 1.3 Interconnect Scaling and 3-D Integration

A lot of emphasis is given to extract the maximum possible performance from a MOSFET in a given technology. But interconnect scaling has been following an altogether different trend and the interconnect delay has become a significant fraction of the total delay in modern ICs [28]. Global interconnects are typically not scaled or reverse scaled to maintain the required resistivity and to handle high current densities. However, the global interconnect delay increases with technology node (as shown in **Figure 1.4** [29]) due to gradually increasing resistivities. The local interconnects lengths scale with technology [29] and do not contribute significantly towards the increasing

percentage of interconnect delay. Advanced solutions like 3-dimensional (3-D) integration and optical interconnects will be required in future technology nodes to keep the interconnect loading overhead under control.

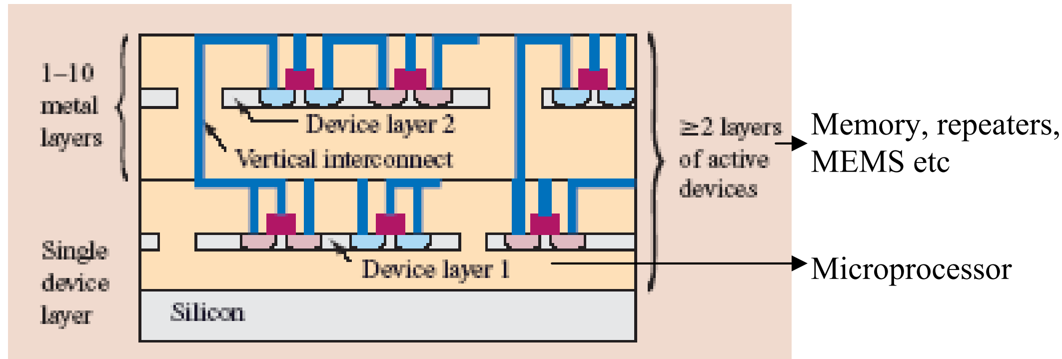



**Figure 1.5**: Schematic representation of a 3-dimensionally integrated circuit [31]. The first device layer could be the microprocessor function, followed by multiple active layers to execute memory, repeaters and other functions.

3-D Integration focuses on stacking different parts/functionalities of a chip in the vertical dimension rather than laterally (**Figure 1.5**). This way, interconnects routing together different blocks of a chip (which are typically several millimeters in length) can be significantly reduced in size (to several micrometers) and their contribution to the overall chip delay can be considerably reduced. Several fabrication schemes have been explored for this solution such as (i) epitaxial growth of silicon, (ii) low temperature crystallization of silicon [30] and (iii) processed wafer bonding [31], with wafer bonding being one of the more popular approaches. Approaches (i) and (ii) suffer from the big challenge of thermal budget. The crystallized silicon film can be a large grained poly crystalline film rather than a single crystalline film, which may affect performance. In

approach (iii) above, the different parts of the chip are fabricated on different silicon substrates and then individually transferred to a primary substrate by wafer bonding techniques. This is a favorable method in terms of the processing thermal budget and the ability to achieve good performance through the use of single crystalline silicon. But this is certainly a complex fabrication scheme and it has inherent issues like ultra high aspect ratio via etching and stringent alignment requirement between different device layers. The available approaches till date are therefore, not very cost-effective. The concept of 3-D integration is very promising but needs a lot more research and development before introduction into real products in a cost-effective manner.

## 1.4 Organization

In this dissertation, some of the above discussed problems have been approached with the aim of making advanced FET structures more manufacturable. Since FinFETs and MuGFETs have a more relaxed silicon body thickness requirement and are more scalable compared to FDSOI or UTB FETs, we will focus on FinFET type devices for the most part. Several designs of FinFETs have been optimized and analyzed through process and device simulations to judge their performance and manufacturability. A methodology to study statistical dopant fluctuations using commercial drift-diffusion simulators has been developed to study variations in highly scaled FinFET devices. A new approach toward 3-D integration has also been introduced and some initial results with ultra-thin body MOSFETs have been shown. Even though a FinFET like architecture may be developed using this 3-D integration scheme, it requires further process optimization and development beyond the work presented here.

This dissertation has been organized into six chapters. In Chapter 2, the double-gate MOSFET will be studied in the accumulation mode of operation rather than the conventional enhancement mode. The key benefits and disadvantages of opting for accumulation mode of operation, in terms of manufacturability and electrical operation will be presented. Two approaches for device fabrication will be illustrated (for planar back-gated architecture) and the issues of using a planar architecture will be highlighted.

Chapter 3 focuses on studying the conventional FinFET design, with the primary goal of implementing single gate CMOS technology. Using a single mid-gap work function gate, the feasibility to implement CMOS for High-Performance (HP), Low-Operating Power (LOP) and Low Standby Power (LSTP) devices will be evaluated. A device design optimization scheme will be developed through 2-D and 3-D device and process simulations. Sensitivities to gate length, body thickness variations and statistical dopant fluctuations for optimal designs will be presented and a brief discussion on SRAM noise-margin variability will be provided.

In Chapter 4, an attempt at modeling statistical dopant fluctuation effects has been presented. The developed methodology aims at using a conventional drift-diffusion simulator to run 3-D atomistic simulations. Various details of the methodology and algorithms will be discussed. This methodology will be applied to a highly scaled HP FinFET and simple device design rules for tolerance to variation from statistical dopant fluctuation effects will be put forward.

In Chapter 5, a novel, low-thermal budget, simple and cost-effective approach to 3-dimensional integration, called wireFET will be introduced. Initial process optimizations towards building these devices and some preliminary device fabrication

data will be shown. The issues with the current process will be demonstrated through electrical and material characterizations. Approaches to refine the process further will be discussed in chapter 6.

The key points of this dissertation are summarized in Chapter 6. Key research contributions and suggestions for future work will be highlighted. The appendices following the conclusive chapter provide the detailed process flows of various processes presented in this work and the programs to run atomistic device simulations.

## 1.5 References


[1] *http://nobelprize.org/educational_games/physics/integrated_circuit/history/*.

[2] G. E. Moore, "Cramming more components onto integrated circuits," *Electronics,* vol. 38, pp. 114 - 117, 1965.

[3] G. E. Moore, "No exponential is forever: but "Forever" can be delayed!," in *Proceedings of IEEE International Solid-State Circuits Conference*, 2003, pp. 20-23.

[4] *International Technology Roadmap for Semiconductors, PIDS Chapter,* http://public.itrs.net.

[5] S. E. Thompson, R. S. Chau, T. Ghani, K. Mistry, S. Tyagi, and M. T. Bohr, "In Search of "Forever" Continued Transistor Scaling One New Material at a Time," *IEEE Transactions on Semiconductor Manufacturing,* vol. 18, No. 1, pp. 26-36, Feb 2005.

[6] A. Asenov, A. R. Brown, J. H. Davies, S. Kaya, and G. Slavcheva, "Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale MOSFETs," *IEEE Transactions on Electron Devices,* vol. 50, pp. 1837-52, 2003.

[7] J. D. Plummer, "Material and Process Limits in Silicon VLSI Technology," *Proceedings of the IEEE,* vol. 89, No. 3, pp. 240-58, 2001.

[8] E. P. Gusev, V. Narayanan, and M. M. Frank, "Advanced high-K dielectric stacks with polySi and metal gates: recent progress and current challenges," *IBM Journal of Research and Development,* vol. 50, pp. 387-410, 2006.

[9] Y.-K. Choi, K. Asano, N. Lindert, V. Subramanian, T.-J. King, J. Bokor, and C. Hu, "Ultrathin-body SOI MOSFET for deep-sub-tenth micron era," *IEEE Electron Device Letters,* vol. 21, pp. 254-5, 2000.

[10] Y.-K. Choi, T.-J. King, and C. Hu, "Nanoscale CMOS spacer FinFET for the terabit era," *IEEE Electron Device Letters,* vol. 23, pp. 25-7, 2002.

[11] D. Hisamoto, W.-C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T.-J. King, J. Bokor, and C. Hu, "FinFET-a self-aligned double-gate MOSFET scalable to 20 nm," *IEEE Transactions on Electron Devices,* vol. 47, pp. 2320-5, 2000.

[12] B. Doyle, B. Boyanov, S. Datta, M. Doczy, S. Hareland, B. Jin, J. Kavalieros, T. Linton, R. Rios, and R. Chau, "Tri-Gate fully-depleted CMOS transistors: fabrication, design and layout," in *2003 Symposium on VLSI Technology*, 2003, pp. 133-4.

[13] P. Verheyen, N. Collaert, R. Rooyackers, R. Loo, D. Shamiryan, A. De Keersgieter, G. Eneman, F. Leys, A Dixit, M.Goodwin, Y.S.Yim, M. Caymax, K. De Meyer, P. Absil, M. Jurczak, and S. Biesemans, "25% Drive Current Improvement for p-type Multiple Gate FET (MuGFET) Devices by Introduction

of Recessed $Si_{0.8}Ge_{0.2}$ in the Source and Drain Regions," in *2005 Symposium on VLSI Technology*, 2005, pp. 194-5.

[14] L. Chang, "Nanoscale Thin-Body CMOS Devices," in *PhD Thesis*: University of California, Berkeley, 2003.

[15] L. Chang, S. Tang, T.-J. King, J. Bokor, and C. Hu, "Gate length scaling and threshold voltage control of double-gate MOSFETs," in *International Electron Devices Meeting*, 2000, pp. 719-22.

[16] B. Doyle, R. Arghavani, D. Barlage, S. Dutta, M. Doczy, J. Kavalieros, A. Murthy, and R. Chau, "Transistor Elements for 30nm Physical Gate Lengths and Beyond," *Intel Technology Journal,* vol. 6, No. 2, pp. 42-54, May 2002.

[17] D. J. Frank, Y. Taur, and H. S. Wong, "Generalized scale length for two-dimensional effects in MOSFETs," *IEEE Electron Device Letters,* vol. 19, pp. 385-7, 1998.

[18] D. Lenoble, K. G. Anil, A. De Keersgieter, P. Eybens, N. Collaert, R. Rooyackers, S. Brus, P. Zimmerman, M. Goodwin, D. Vanhaeren, W. Vandervorst, S. Radovanov, L. Godet, C. Cardinaud, S. Biesemans, T. Skotnicki, and M. Jurczak, "Enhanced performance of PMOS MUGFET via integration of conformal plasma-doped source/drain extensions," in *2006 Symposium on VLSI Technology*, 2006, pp. 168-9.

[19] P. Ranade, Y.-K. Choi, D. Ha, A. Agarwal, M.Ameen, and T.-J. King, "Tunable work function molybdenum gate technology for FDSOI-CMOS," in *IEEE International Electron Devices Meeting*, 2002, pp. 363-6.

[20] W. P. Maszara, Z. Krivokapic, P. King, J. S. Goo, and M. R. Lin, "Transistors with dual work function metal gates by single full silicidation (FUSI) of polysilicon gates," in *IEEE International Electron Devices Meeting*, 2002, pp. 367-70.

[21] T. Ezaki, T. Ikezawa, and M. Hane, "Investigation of realistic dopant fluctuation induced device characteristics variation for sub-100 nm CMOS by using atomistic 3D process/device simulator," in *IEEE International Electron Devices Meeting*, 2002, pp. 311-14.

[22] M. Hane, T. Ikezawa, and T. Ezaki, "Coupled atomistic 3D process/device simulation considering both line-edge roughness and random-discrete-dopant effects," in *IEEE International Conference on Simulation of Semiconductor Processes and Devices*, 2003, pp. 99-102.

[23] V.-Y. A. Thean, M. Sadd, and J. Bruce E. White, "Effects of Dopant Granularity on Superhalo-Channel MOSFET's According to Two- and Three-Dimensional Computer Simulations," *IEEE Transactions on Nanotechnology,* vol. 2, No. 2, pp. 97-101, June 2003.

[24] A. Asenov, S. Kaya, and A. R. Brown, "Intrinsic Parameter Fluctuations in Decananometer MOSFETs Introduced by Gate Line Edge Roughness," *IEEE Transactions on Electron Devices,* vol. 50, No. 5, pp. 1254-60, May 2003.

[25] S. Xiong, J. Bokor, Q. Xiang, P. Fisher, I. Dudley, P. Rao, H. Wang, and B. En, "Is Gate Line Edge Roughness A First Order Issue In Affecting The Performance Of Deep Sub-micron Bulk MOSFET Devices?," *IEEE Transactions on Semiconductor Manufacturing,* vol. 17, pp. 357-61, August 2004.

[26] V. Moroz, G. Eneman, P. Verheyen, F. Nouri, L. Washington, L. Smith, M. Jurczak, D. Pramanik, and X. Xu, "The Impact of Layout on Stress Enhanced Transistor Performance," *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD),* pp. 143-6, 2005.

[27] K. Bernstein, D. J. Frank, A. E. Gattiker, W. Haensch, B. L. Ji, S. R. Nassif, E. J. Nowak, D. J. Pearson, and N. J. Rohrer, "High-performance CMOS variability in the 65-nm regime and beyond," *IBM Journal of Research and Development,* vol. 50, pp. 433-49, 2006.

[28] J. A. Davis, R. Venkatesan, A. Kaloyeros, M. Beylansky, S. J. Souri, K. Banerjee, K. C. Saraswat, A. Rahman, R. Reif, and J. D. Meindl, "Interconnect limits on gigascale integration (GSI) in the 21st century," *Proceedings of the IEEE,* vol. 89, pp. 305-24, 2001.

[29] *International Technology Roadmap for Semiconductors, Interconnect Chapter,* 2005, (http://public.itrs.net).

[30] V. Subramanian and K. C. Saraswat, "A Novel Technique for 3-D Integration: Ge-seeded Laterally Crystalluzed TFTs," *1997 Symposium on VLSI Technology,* pp. 97-98, 1997.

[31] A. W. Topol, D. C. La Tulipe Jr., L. Shi, D. J. Frank, K. Bernstein, S. E. Steen, A. Kumar, G. U. Singco, A. M. Young, K. W. Guarini, and M. Ieong, "Three-dimensional Integrated Circuits," *IBM Journal of Research and Development,* vol. 50, No. 4/5, pp. 491-506, 2006.

# Chapter 2 : **Accumulation Mode Thin-body MOSFETs**

## 2.1 Introduction

As a consequence of scaling, state-of-the-art ICs today have transistors with gate length ($L_G$) below 35nm [1]. In order to suppress short channel effects, very thin gate dielectric (EOT<1nm using high-k gate dielectrics) and ultra-shallow source/drain junctions are employed. Further, to alleviate the issue of polysilicon gate depletion effect, metal gates will be employed in the near future [2]. These new materials and processes have been facing a lot of challenges [3] in getting introduced in the conventional CMOS process flow. In spite of these changes, short channel effects (e.g, DIBL and sub-surface leakage) are likely to be a major showstopper in scaling bulk CMOS FETs down to 10nm $L_G$. The ultra-thin body MOSFET (UTB FET) and double-gate MOSFET [4-6] (**Figure 2.1**) are alternative device structures which limit the channel conduction to a thin semiconductor layer, thus eliminating sub-surface leakage and improving short channel control. The multiple gate FET architectures like the Multi-gate FET (MuGFET) [7], Tri-gate FET [8] and Gate All Around FET (GAA FET) [9] have the additional benefit of controlling the channel from more than one direction, thus providing superior gate

control (hence performance) [10] compared to the UTB FET and the bulk MOSFET, as shown in **Figure 2.2**.

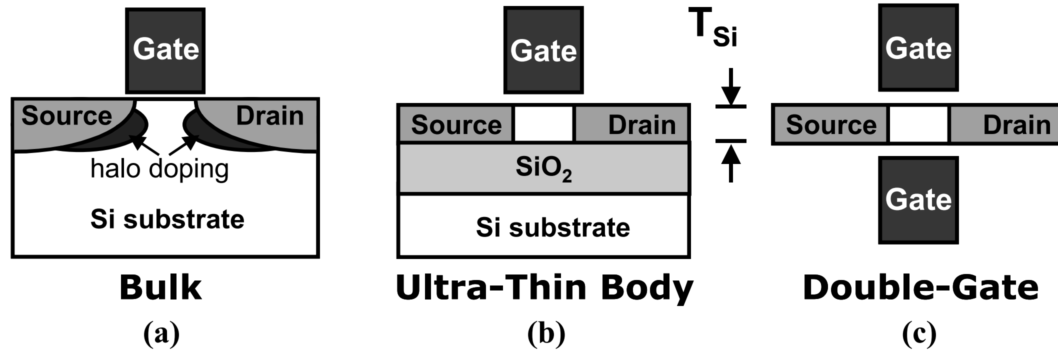



**Figure 2.1**: Schematic cross-sections showing the (a) conventional bulk MOSFET and advanced FET structures such as (b) planar Ultra-Thin Body (UTB) MOSFET and (c) double-gate (DG FET) MOSFET [10]. Confining conduction to a thin silicon film, significantly improves gate control of the channel.

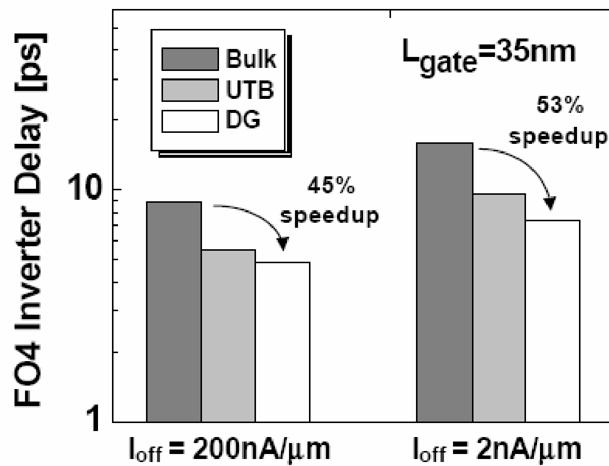



**Figure 2.2**: Owing to thin silicon channel and increased gate control, DG FET has better performance than UTB and bulk MOSFET [10].

In this chapter, double gate MOSFET design will be studied and evaluated for the accumulation mode design rather than the conventional enhancement mode [11]. The

working principle, device design and, potential advantages and disadvantages will be discussed in detail. Ideally, in enhancement mode of operation, the optimum gate work function for NMOS and PMOS are different [12], making it necessary to have two different gate metals for a CMOS process. Single gate work function CMOS design and technology will be dealt in greater detail in the next chapter. In this chapter, the feasibility and tradeoffs of implementing accumulation mode single gate work function CMOS technology will be briefly presented.

## 2.2 Thin-body Accumulation Mode MOSFETs

### 2.2.1 Device Structure and Operation

The structure of an accumulation mode DG FET (**Figure 2.3**) is very similar to an enhancement mode DG FET. The main difference lies in the channel doping for the two cases. For an accumulation mode (ACC) MOSFET, the S/D doping type is the same as the channel doping type, thus eliminating source-channel and drain-channel p-n junctions.

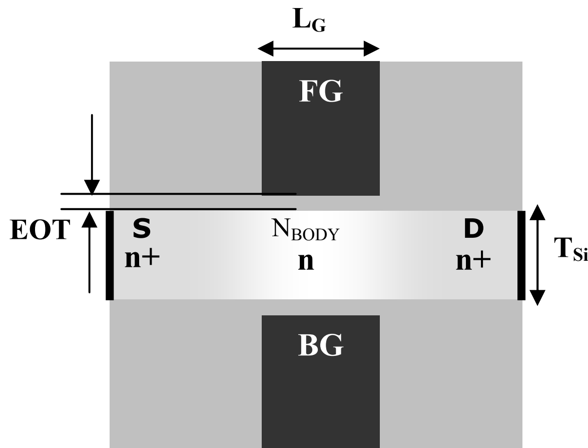



**Figure 2.3**: Schematic diagram of a 2D Accumulation mode DG nMOSFET.

| Parameter | Value |
|---|---|
| EOT | 6A of $SiO_2$ |
| $N_{BODY}$ (ACC) | 2E16-1E19 $cm^{-3}$ |
| $N_{BODY}$ (ENH) | 2E16 $cm^{-3}$ |
| $N_{S/D}$ | 1E20 $cm^{-3}$ |
| $L_G$ | 13nm, 9nm |
| $V_{DD}$ | 0.6V |
| $T_{Si}$ ($L_G$=13nm) | ~$2L_G/3$ |
| $I_{OFF}$ ($L_G$=13nm) | 1E-6 A/μm |

**Table 2.1**: Device parameters used in simulations

In bulk accumulation mode MOSFETs, a buried channel device is utilized [13]. The conduction path being physically far away from the gate-oxide interface, gives the benefit of lower noise in ACC mode devices, but leads to poor short channel control than its bulk enhancement mode (ENH) counterpart (i.e degrades performance). By utilizing a heterostructure thin-body device [14], it is possible to reap the benefits a buried channel device without sacrificing performance, but at the cost of additional process complexity.

The device operation of the device architecture studied in this work is summarized in **Figure 2.4**. In the OFF-state ($V_{DS}$=$V_{DD}$ and $V_{GS}$=0V), the gate completely depletes the channel region, while in the ON-state ($V_{DS}$=$V_{DD}$ and $V_{GS}$=$V_{DD}$), the depletion region recedes leading to resistive conduction. Further increase in gate bias leads to the formation of a channel accumulation region. It must be noted that in the OFF-state, the depletion region spreads in the lateral and longitudinal directions, thus leading to a larger effective channel length ($L_{eff}$) than in the ON-state. The electric potential profiles in the ON and OFF states, perpendicular and parallel to the direction of current flow, are shown in **Figure 2.5**. In the OFF state, the ACC mode design has more energy-band-bending in the vertical direction (perpendicular to current flow) than the ENH mode design, indicating that carriers are confined more to the center of the semiconductor film, resulting in higher leakage current $I_{OFF}$ for a given ON-state current specification. In the ON state, the shape of the vertical potential profile is similar for the ACC mode and ENH mode designs. The lateral (parallel to current flow) potential profile is similar for the ACC mode and ENH mode designs in the OFF state. The ON-state lateral potential profiles are similar, except that the voltage drop between the source and drain regions is

smaller for the ENH mode design due to non-zero voltage dropped across parasitic source/drain series resistances (absent in ACC mode design due to heavy channel doping).

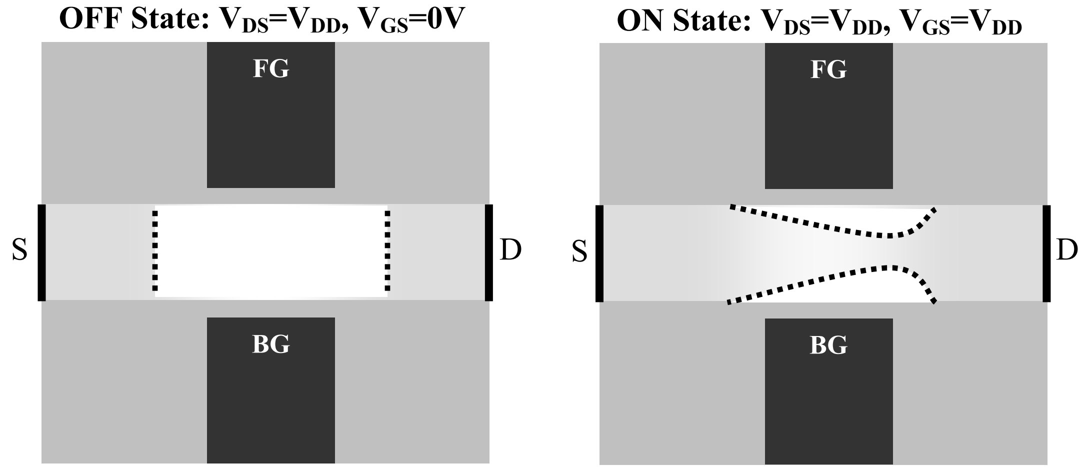



**Figure 2.4**: Pictorial representation of the ACC mode DG FET operation. In the OFF state, the $L_{eff}$ is larger than in the ON state owing to the extent of the depletion region. Current conduction is through the center of the fin (in ON and OFF state) for heavily doped channel.

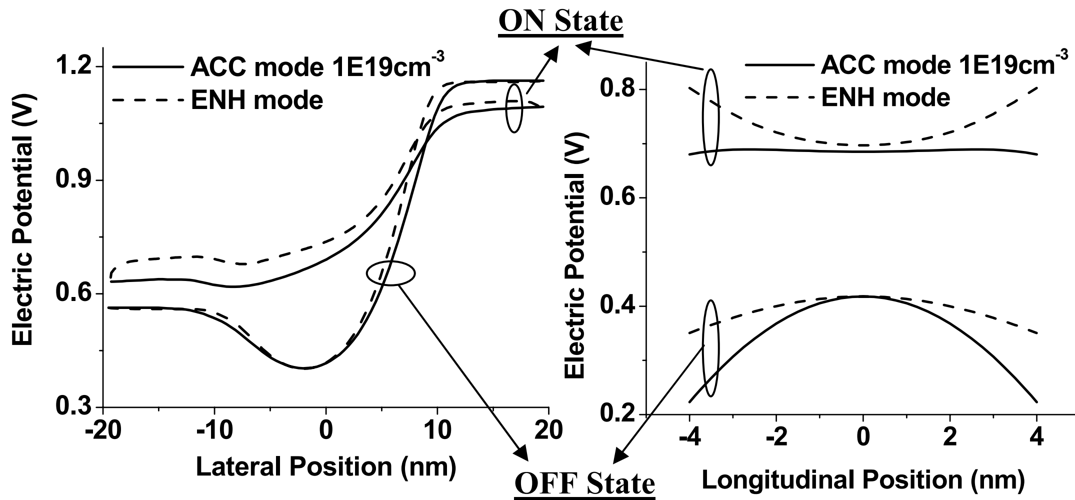



**Figure 2.5**: Electric potential in the lateral (direction of current flow) and longitudinal (perpendicular to current flow) direction in ON and OFF state for the two FET designs. For ACC mode FET (heavily doped), $V_{GS} > V_{DD}$ eventually leads to a longitudinal profile similar to ENH mode FET and, volume conduction.

### 2.2.2 Potential Benefits and Issues

From a first glance at the ACC mode structure, it is simpler to fabricate than its ENH mode counterpart due to absence of S/D p-n junctions. The ON and OFF state behavior also provides some benefits over the ENH mode architecture. The longitudinal direction (perpendicular to current flow) potential profile results in good quantum confinement in the OFF-state which is good for lower OFF-state current. In the ON state however, the depletion region band bending is eliminated thus reducing the quantum confinement effect, making it better for conduction. However, the shrinking of the $L_{eff}$ in the ON state is likely to degrade the short channel effects.

It is possible to adjust the body doping of the ACC mode device to achieve single gate work function for NMOS and PMOS (will be discussed in a later section), which makes it a potentially manufacturable solution. A high body doping yields lower parasitic series resistance (can yield better performance), but leads to (i) ionic scattering and (ii) high dopant fluctuation effects, which can lead to performance degradation as well as higher variability. Clearly, different phenomena contribute towards performance enhancement or degradation. The net effect on performance will be analyzed in the sections to follow and the accumulation mode MOSFET will be evaluated as a potential candidate for highly scaled CMOS technology.

## 2.3 Device Design

In this section, device design strategy will be discussed through 2D Device simulations of DG ACC mode nMOSFETs. The simulations have been done using the Taurus-Device simulator [15], with drift-diffusion transport and 1D Schrödinger solution

for quantum confinement effects. The device structure is as shown in **Figure 2.3** and the device parameters (taken from ITRS roadmap [16]) are summarized in **Table 2.1**.

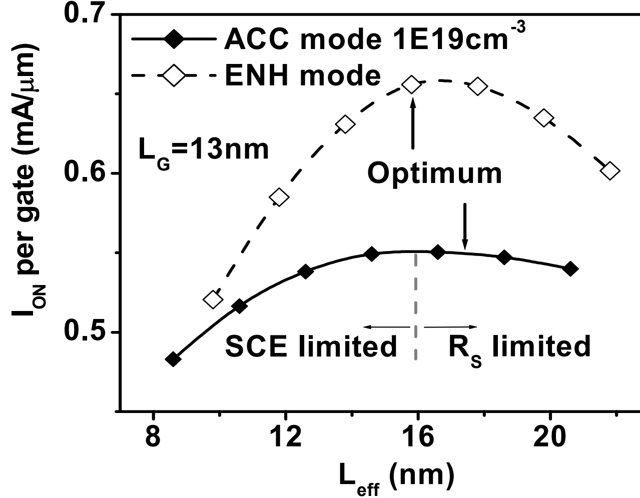



**Figure 2.6**: $I_{ON}$ vs $L_{eff}$ showing short channel effect limited and series resistance limited regions of operation. Optimum device is identified as one with maximum $I_{ON}$ for DIBL < 100mV/V. Ionic scattering limits the performance of heavily doped ACC mode FET.

### 2.3.1 Design Optimization

The ACC and ENH mode transistor designs are each optimized for maximum performance ($I_{ON}$) by adjusting the effective channel length, $L_{eff}$. For each value of $L_{eff}$, the gate work function is adjusted so as to meet the ITRS leakage current specification. If $L_{eff}$~$L_G$, short channel effects limit the ON-state current ($I_{ON}$). As $L_{eff}$ is increased, short channel control and hence $I_{ON}$ improves; but beyond a certain value of $L_{eff}$, increasing series resistance due to gate-to-source/drain underlap results in decreasing $I_{ON}$ [17]. The optimized device design is selected to be the one which provides the highest $I_{ON}$, with DIBL less than or equal to 100mV/V. The $I_{ON}$ *vs.* $L_{eff}$ curves for 13nm $L_G$ ACC-mode and ENH-mode FETs are shown in **Figure 2.6**. Due to heavy channel doping which

extends into the 'underlap' region as well, the accumulation mode MOSFET current does not fall off as fast as the ENH mode device in the series resistance limited regime. However, Coulombic scattering leads to a lower optimal $I_{ON}$ for the ACC mode FET compared to the ENH mode FET. It should be noted that in general the peak performance point in the $I_{ON}$-$L_{eff}$ plot may not be the optimum point from short channel effects point of view. (An example of this will be seen in Chapter 4).

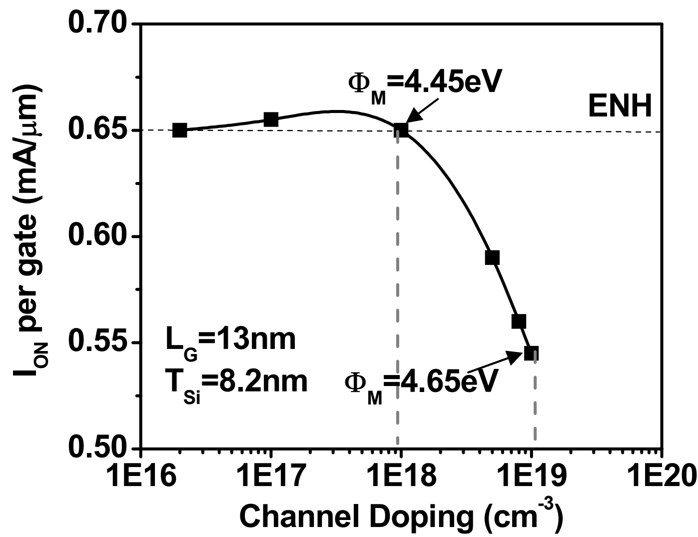



**Figure 2.7**: Optimum $I_{ON}$ (per gate) as function of channel doping for ACC mode FET. Single gate work function CMOS with midgap work function is feasible by appropriately tuning the channel doping ($I_{ON}$ is degraded for high channel doping).

Since the optimum $I_{ON}$ is limited by Coulombic scattering, it would strongly depend on the body doping. **Figure 2.7** shows $I_{ON}$ as a function of channel doping for the ACC mode FET, where each point has been independently optimized for best performance for the target $I_{OFF}$ and DIBL specifications. At low channel doping, the channel charge is almost insignificant and the device behaves exactly like an

enhancement mode device. The histogram shown in **Figure 2.8** summarizes the optimized $I_{ON}$ for differently doped ACC mode designs and the ENH mode design for two different gate lengths. To achieve high performance (comparable to the ENH mode FET), body doping less than 1E18cm$^{-3}$ is desirable. However, the low doped device behaves essentially like an undoped body device (in this case enhancement mode and accumulation mode are merely two different ways of referring to exactly the same operation).

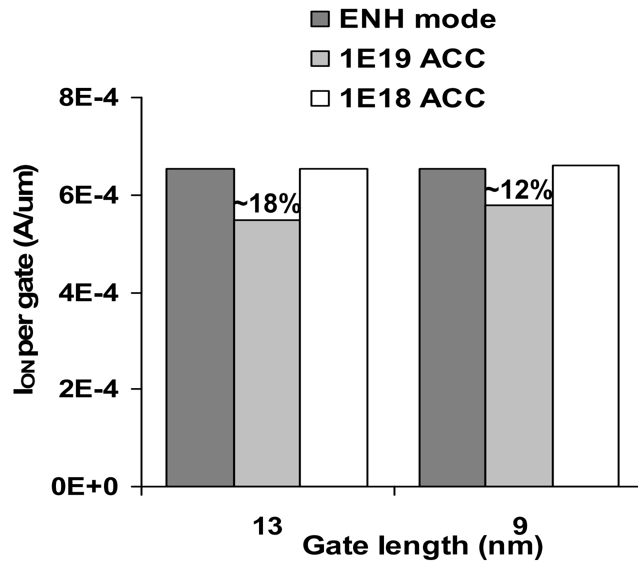



**Figure 2.8**: $I_{ON}$ comparison for ACC mode and ENH mode FET for different $L_G$(s). Coulombic scattering limits the performance of the heavily doped ACC mode FET.

An important consideration for transistor scaling is the gate work function requirement. The optimum gate work function for an undoped channel ENH mode double gate design requires two different gate work functions for NMOS and PMOS devices (4.45eV for NMOS and ~5eV for PMOS) to achieve symmetric threshold voltages ($|V_{TNMOS}| = |V_{TPMOS}|$) [12]. Referring back to **Figure 2.7**, the required gate work function for an ACC mode device depends on the channel doping level. For channel doping of

~1E19cm$^{-3}$, the required gate work function for a NMOS ACC mode FET is about 4.65eV (midgap). Therefore, a PMOS ACC mode FET with the same structure and channel doping (but p-type instead of n-type) will also require the same work function for optimal performance and a complementary threshold voltage. This indicates the viability of a single gate work function technology at the expense of a slightly lower performance. Along the same lines, an ENH mode NMOS and ACC mode PMOS (with appropriate body doping) can also be used with a single gate work function technology.

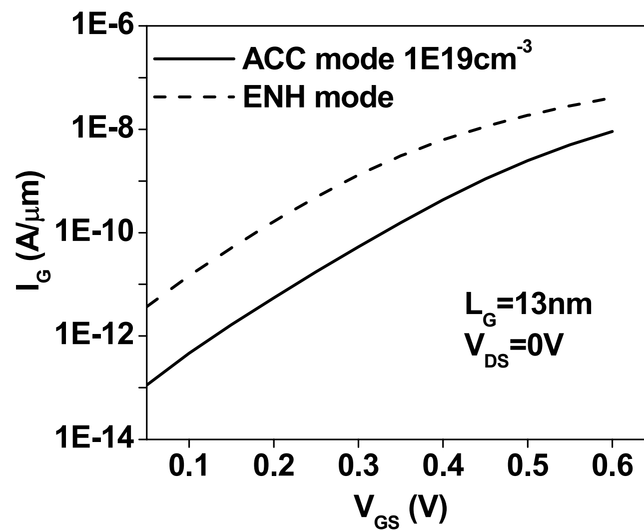



**Figure 2.9**: Comparison of gate leakage current densities for ACC vs. ENH mode designs. The ACC mode design has a significantly lower gate current due to lower vertical electric field.

An important benefit of the ACC mode device lies in its lower gate current (direct tunneling current), which is a very serious concern for future generations of MOSFETs. For a heavily doped ACC mode device, the carrier centroid lies at the center of the channel in both OFF and ON state, leading to a lower vertical electric field in the channel. On the other hand the undoped body device has its carrier centroids closer to the two

gates in the ON state, leading to a higher vertical electrical field in the channel and a higher gate current density (**Figure 2.9**).

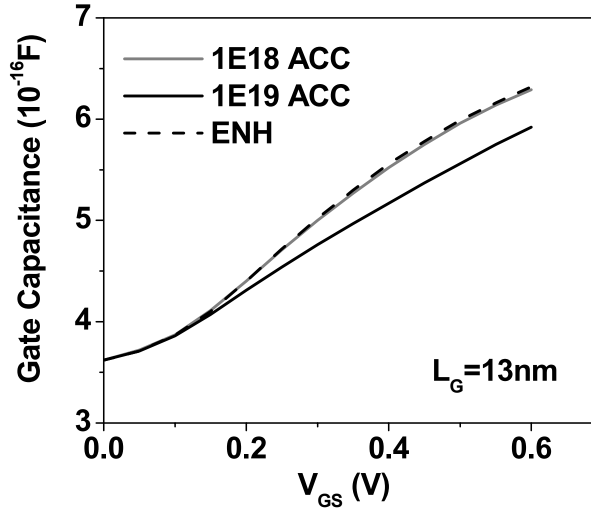



**Figure 2.10**: Gate capacitance comparison showing lower ON state capacitance for heavily doped ACC mode FET. This is another outcome of the lower vertical electric field due to the position of the charge centroid.

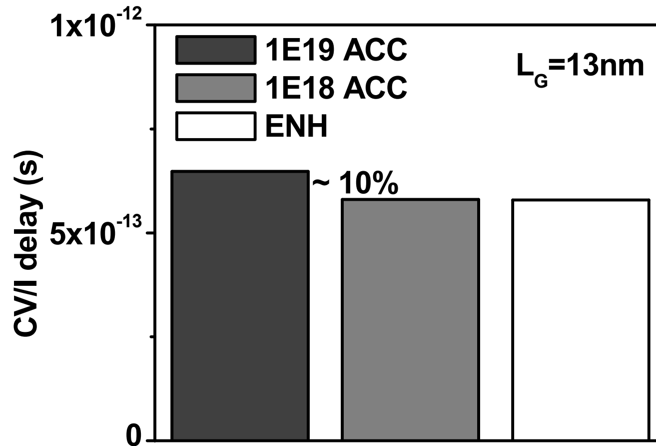



**Figure 2.11**: The intrinsic delay of heavily doped ACC mode design is within 10% of ENH mode design in spite of it much lower $I_{ON}$.

The charge centroid location in the ON state also provides a benefit in the capacitance for the heavily doped ACC mode design, while the lightly doped ACC mode

design behaves like the ENH mode design. The intrinsic charging delay of a transistor is determined not only by its $I_{ON}$ but also by its gate capacitance ($C_G$): $\tau = C_G*V_{DD}/I_{ON}$. **Figure 2.10** and **Figure 2.11** compare the gate capacitances and intrinsic delays of the two devices. Even though the $I_{ON}$ of a heavily doped ACC mode FET is significantly lower than that of the ENH mode device, the lower gate capacitance of the ACC mode device compensates for it and helps to improve its intrinsic delay. As shown in **Figure 2.11**, this delay, for an ACC mode design, is within 10% of the intrinsic delay of the ENH mode device. Thus, the accumulation mode DG FET performance is comparable to the enhancement mode DG FET, with benefits in (i) S/D design (ii) lower gate current (iii) lower parasitic S/D resistance and, (iv) feasibility of single gate work function CMOS implementation. In the next subsection, we shall focus on short channel control comparisons and sensitivities to gate length and body thickness variations, which are also very important for highly scaled MOSFETs.

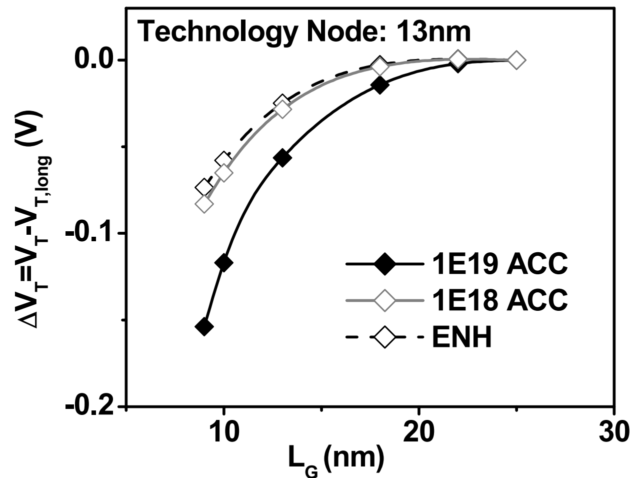



**Figure 2.12**: $V_T$ roll-off plot comparing the ACC and ENH mode devices. Inherent poor short channel control leads to higher gate-length sensitivity of the heavily doped ACC mode design.

## 2.3.2 Short Channel Effects and Sensitivities

**Figure 2.12** plots the $V_T$ roll-off curves (sensitivity to gate length variations) for the different ACC mode designs and the ENH mode design. The heavily doped ACC mode design is much more sensitive to $L_G$ variations. This is expected because the intrinsic source-to-channel potential barrier for this device is almost zero since the source and the channel are doped heavily of the same type. This leads to easier penetration of the drain electric field to the source-channel barrier, thus providing weaker short channel control. **Figure 2.13** compares the sensitivity of the devices to body thickness variations. Once again, the heavily doped ACC mode device has a much higher sensitivity than the ENH mode FET. Small changes in the body thickness of a heavily doped channel correlate to significant changes in the channel ionic charge contributing towards the $V_T$ and higher sensitivities.

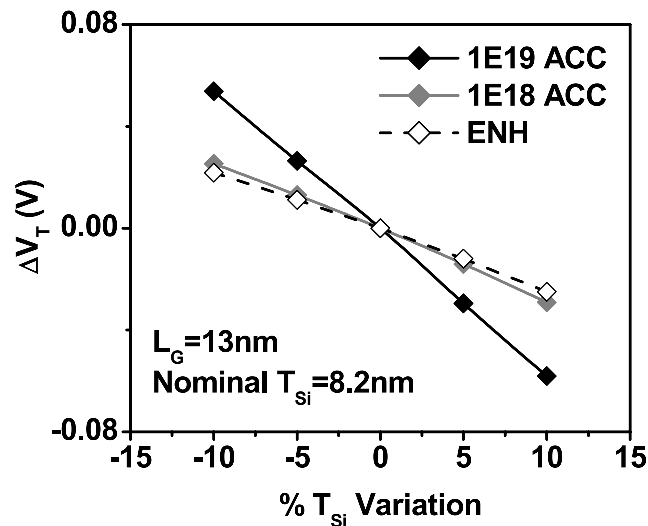



**Figure 2.13**: The heavily doped ACC mode design shows higher sensitivity to body thickness variations. A change in $T_{Si}$ leads to significant change in channel ionic charge and hence $V_T$.

An important aspect of variations not been explored here, is the effect of statistical dopant fluctuations. For a very small channel volume, the number of dopant atoms is very few and therefore, their location and number (presence or absence of a dopant atom) becomes important with respect to variations. For a heavily doped ACC mode MOSFET, variations due to SDF effects will further add on to the sensitivities in $L_G$ and $T_{Si}$ fluctuations caused by LER effects. Overall, the accumulation mode MOSFET is more prone to variations than the enhancement mode design.

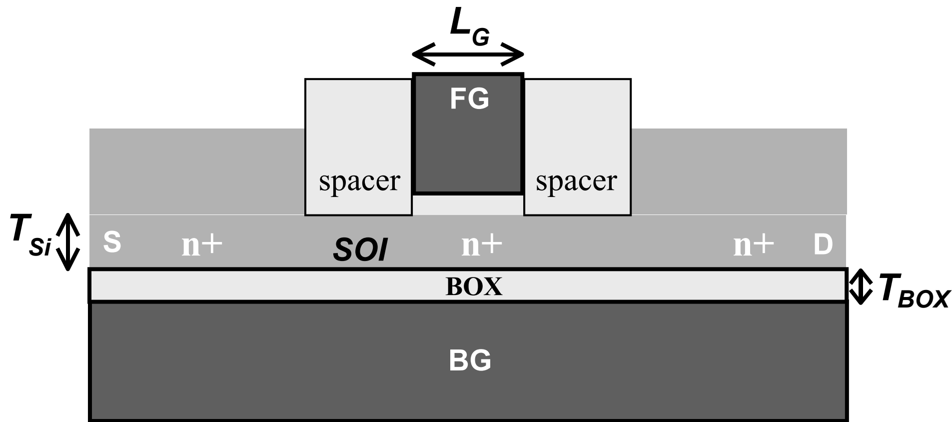



**Figure 2.14**: Schematic cross-section of the fabricated ACC mode FET. The channel doping was targeted ~ $5E19cm^{-3}$ requiring $T_{Si}$<11nm. A planar architecture is designed in order to minimize topography related issues.

## 2.4 Device Fabrication Issues

The fabrication of the accumulation mode MOSFET was attempted using (i) gate last as well as (ii) gate first approaches. Planar MOSFETs have an inherent topography advantage over non-planar architectures like the FinFET architecture. Therefore, in this work, a planar architecture was targeted and the device was designed in the back-gated mode (substrate is used as the back-gate and the silicon channel is formed on super thin buried oxide for reasonable back-gate control). A schematic cross-section of the target

device structure is shown in **Figure 2.14**. It can be clearly seen that in order to turn off the device, a very thin silicon channel will be needed. In order to avoid an impractically thin silicon channel thickness, back-gate control is required. Further, to maintain a reasonable gate control, a very thin buried oxide (BOX) layer is needed. Even though the S/D design becomes rather simple, the rest of the process turns out to be more complicated in the attempt to maintain a planar structure. The process optimization and the issues in fabrication of the back-gated accumulation mode MOSFET will be briefly discussed in this section.

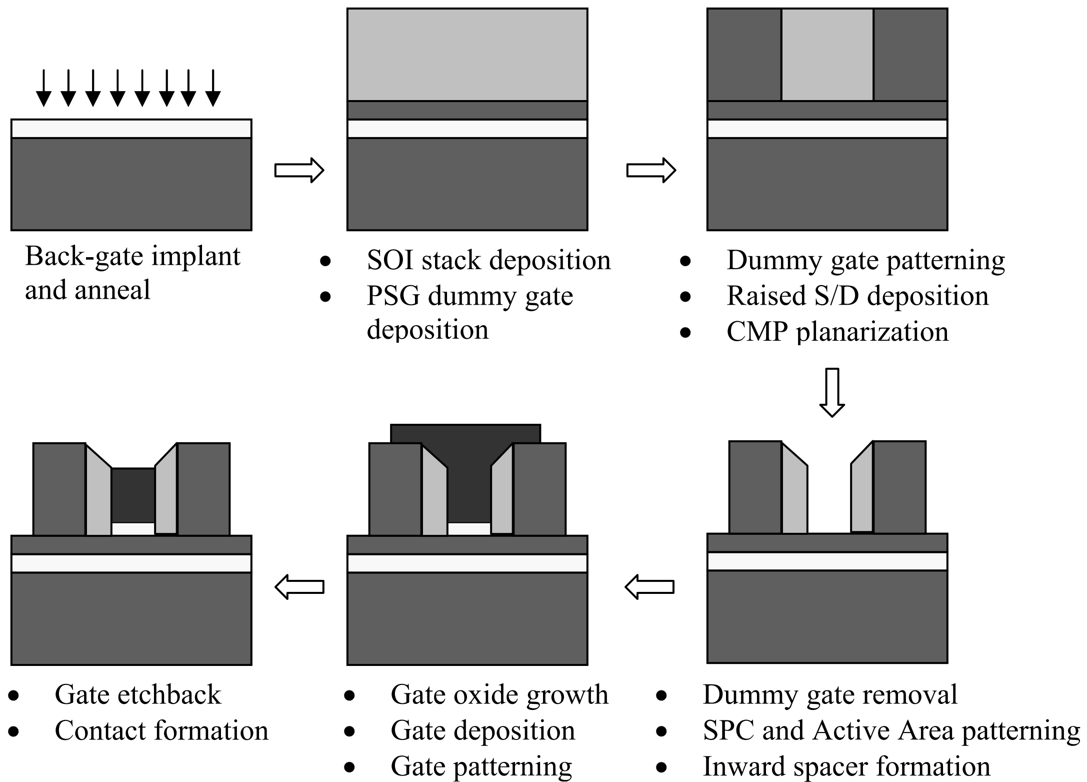



**Figure 2.15**: Basic process flow for the gate last process.

### 2.4.1 Gate Last Process

A *gate last* approach was first adopted to fabricate the ACC mode back-gated MOSFET. The process flow is summarized diagrammatically in **Figure 2.15** and the flow table is provided in Appendix A. The motivation for looking at this approach is twofold:

(i) the $L_G$ of the FET can be made very small using an inverse spacer process, rather than relying on lithography techniques like photoresist ashing and trimming and, (ii) the device can be subjected to high temperature processing almost until the last process steps, thus providing process flexibility.

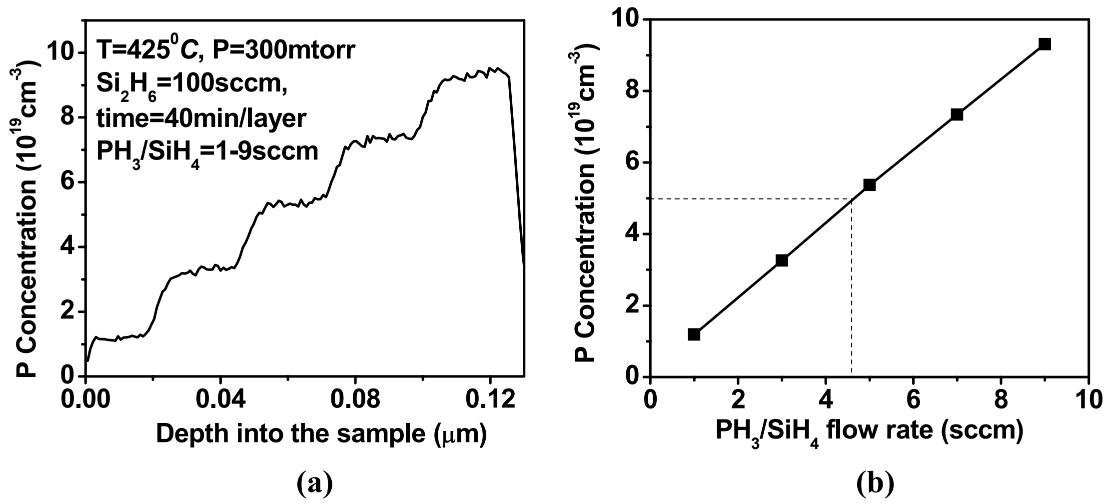



**Figure 2.16**: SIMS analysis to identify precise process conditions for 5E19cm$^{-3}$ channel doping. (a) SIMS raw data shows phosphorus concentration as a function of depth for a 5 step deposition process and (b) phosphorus concentration for each step (varying $PH_3$ flow rate) showing required $PH_3/SiH_4$ flow rate to be 4.7sccm.

In this process, since the BOX layer is very thin, standard available SOI wafers could not be used. An amorphous silicon (a-Si) film was deposited and solid-phase crystallized to achieve a polycrystalline silicon channel. Heavily doped polysilicon is expected to have similar transport properties as single crystalline silicon, therefore, justifying the use of a polycrystalline film. The heavy doping in the silicon channel was achieved by in situ doping during the a-Si deposition in an LPCVD furnace (tystar19). A short flow process run was conducted to identify the doping concentration of phosphorus

in silicon using the in situ doping process. The deposition was done at $425^0$C and 300mtorr, using $Si_2H_6$=10sccm, $PH_3/SiH_4$=1-9sccm (in steps), and time=40min/layer. The SIMS analysis for phosphorus is shown in **Figure 2.16**. In order to achieve a doping concentration of ~5E19cm$^{-3}$, $PH_3/SiH_4$ flow rate of 4.7 sccm was identified. The channel silicon deposition was targeted for thicknesses of 5nm and 10nm at $425^0C$. In order to achieve a uniform deposition with minimal pin-hole density, a two step deposition process (10A seed layer deposition + doped silicon film deposition) was implemented. The films were crystallized by SPC (furnace, tylan6) at $550^0C$ for 12 hours.

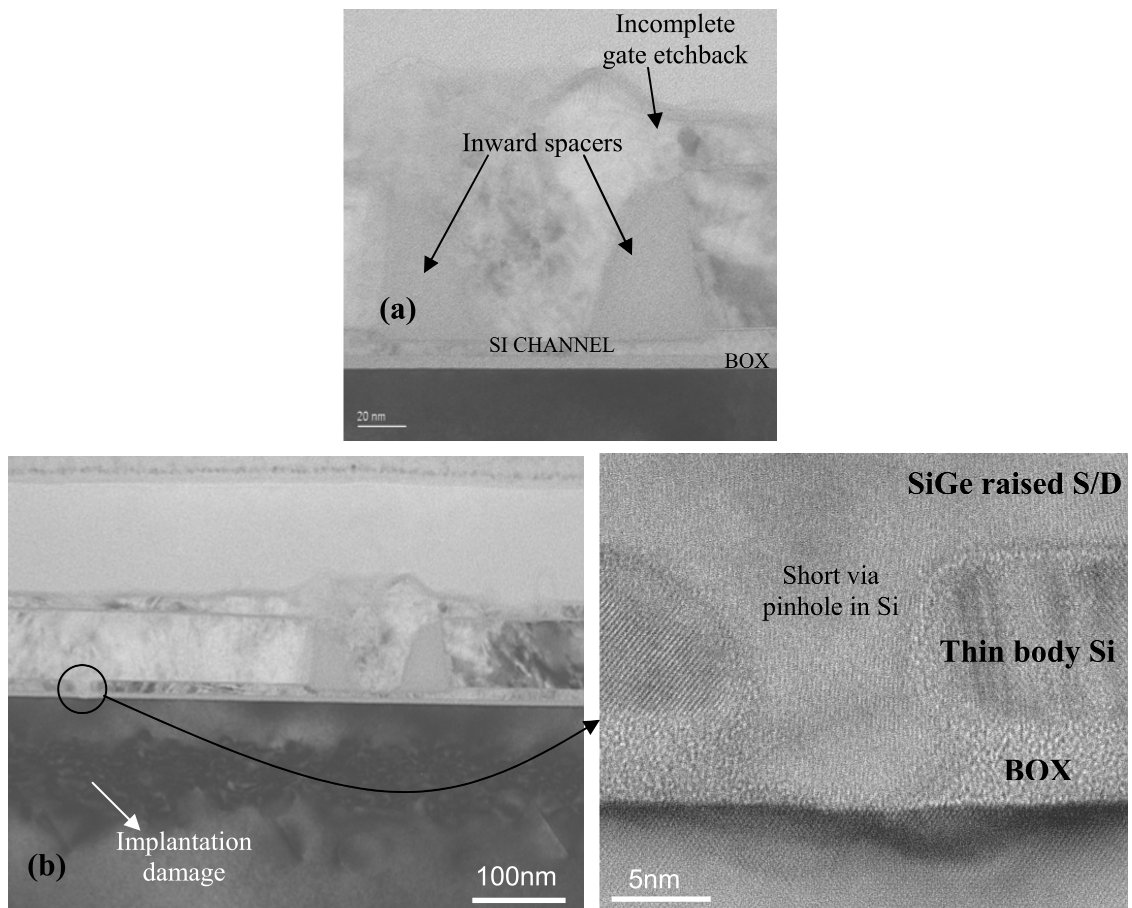



**Figure 2.17**: Issues of the gate last process: (a) Cross-sectional TEM micrograph of the completed gate last process for the ACC mode FET (b) Pinholes in Si layer cause HF etching of underlying BOX and shorts to substrate during SiGe S/D deposition.

With the above mentioned process for the channel, the complete process (as described in Appendix A) was implemented. This process required five lithography steps, two being critical. The inverse spacer process is promising from lithography point of view, but is very difficult to implement since identifying the etching end point (for inwards spacer) is challenging, especially for very short gate length devices. A two step etching (initial timed dry etch up to 90% + final wet etch in 100:1 HF) is required to ensure complete etching of the inverse spacer. The overetch time should also be carefully chosen to make sure that the inverse spacer is not completely lost. The TEM image of the completed device is shown in **Figure 2.17**(a) showing the complete device structure. One issue seems that the gate etchback process is not completed, as it was a timed etch process. Further, zooming into the S/D regions, as shown in **Figure 2.17**(b), 5nm shorts from S/D to substrate are observed. The most likely step at which these shorts were formed, are at the wafer precleaning step with a 100:1 HF dip prior to the SiGe raised S/D deposition. The HF could have percolated through tiny pinholes in the silicon film and etched through the ultra thin BOX film, to create the shorts to the substrate. A pure $SiO_2$ BOX layer is therefore, not recommended.

### Shortcomings of Gate Last Process

Summarizing, the gate last process faces several drawbacks, listed as follows:

- The inverse spacer etch process is difficult to implement due to inability to identify the end point.
- An ultra thin $SiO_2$ buried oxide film is undesirable to avoid accidentally etching it during wafer precleaning steps (which are unavoidable for a conventional process flow). A pure $Si_3N_4$ dielectric is also undesirable, as $Si:Si_3N_4$ dry etch selectivity is

difficult to achieve for the active area etching step. A double layer stack of $Si_3N_4$ and $SiO_2$ is therefore, the most appropriate buried oxide definition to evade all the problems.

- There is a lot of additional process complexity because of the replacement gate process. Moreover, it is not easy to time the end of gate etch back process.

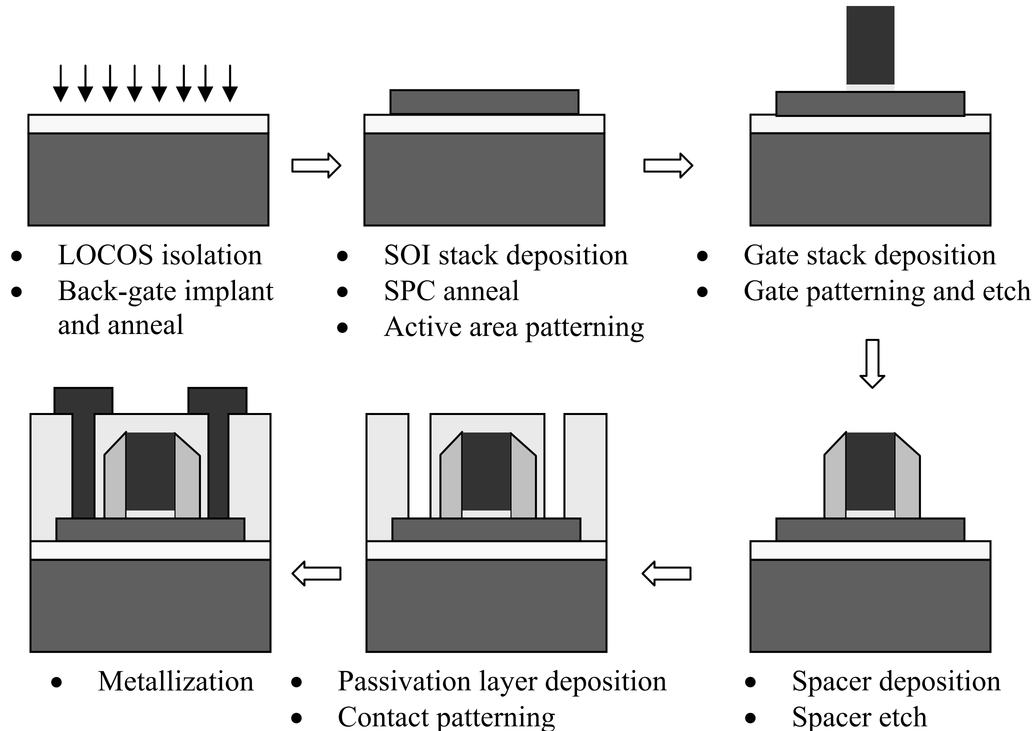



**Figure 2.18**: Process flow for gate first process. It is much simpler and closer to the standard bulk CMOS process than the gate last process. However, this process poses stringent requirements on the selectivities for active and gate etching steps.

## 2.4.2 Gate First Process

In order to circumvent some of the issues of the gate last process, the *gate first* approach was tested. Once again, the process flow is summarized pictorially in **Figure 2.18** and in tabular form in Appendix A at the end of this thesis. In this process, a double

layer stack was used for the back-gate dielectric and Molybdenum gate with polysilicon cap was used for the top/front gate. The Mo layer is kept thin (200A) in order to be able to successfully stop the Mo etch on the ultra-thin silicon channel without completely etching away the thin silicon. This process is simpler than the previous process flow, but still suffers from the stringent dry etching requirements and was therefore, not implemented successfully. The ultra-thin Mo layer was easily oxidized during the furnace loading process for poly-Si cap deposition right after the Mo sputtering. Additionally, the active area dry etch went too deep on some wafers, once again, yielding shorts to the substrate.

Even though a planar architecture is very promising from a topography standpoint, implementing a reliable ultra-thin buried oxide process still remains a challenge. For future studies, increasing the metal gate thickness and reducing the polysilicon thickness may help alleviate the Mo oxidation problem, though the Mo gate-etch (to stop on poly-Si channel) selectivity would become even more stringent. It may therefore, be more suitable to opt for a non-planar architecture if there is a reasonable performance benefit from using the ACC mode design!

## 2.5 Summary

In this chapter, Accumulation mode double-gate MOSFET has been studied as a possible alternative to the conventional enhancement mode double-gate MOSFET for highly scaled sub-15nm MOSFETs. The chief benefit of the accumulation mode FET lies in the more manufacturable channel and S/D design, and feasibility to implement single gate work function CMOS technology. The performance of a heavily doped channel FET is limited by Coulombic scattering as expected, but has the important benefits of (a)

lower capacitance and (b) lower gate leakage current than the undoped channel enhancement mode design. These benefits arise from a lower vertical electric field in the on-state owing to a favorable charge centroid position (center of the channel). Therefore, intrinsic delay of the doped channel accumulation mode MOSFET is within 10% of its undoped channel enhancement mode counterpart. However, due to absence of an intrinsic source-channel barrier, the accumulation mode MOSFET has poorer short channel control, thus making it more sensitive to variations. In addition, statistical dopant fluctuation effects will add another component to the variations.

It has been shown that a planar architecture is not very simple to fabricate due to the challenges caused by the need for ultra-thin BOX and ultra-thin channel layers. These films make the etching processes very stringent and require highly selective etch recipes. It may therefore, be more beneficial to look at the non-planar FinFET-like architecture using spacer lithography process to fabricate this device.

Putting together all the pros and cons of the accumulation mode design, its device behavior, from performance as well as processing standpoints, is not significantly better over the enhancement mode design.

## 2.6 References

[1] P. Bai, C. Auth, S. Balakrishnan, M. Bost, R. Brain, V. Chikarmane, R. Heussner, M. Hussein, J. Hwang, D. Ingerly, R. James, J. Jeong, C. Kenyon, E. Lee, S. H. Lee, N. Lindert, M. Liu, Z. Ma, T. Marieb, A. Murthy, R. Nagisetty, S. Natarajan, J. Neirynck, A. Ott, C. Parker, J. Sebastian, R. Shaheed, S. Sivakumar, J. Steigerwald, S. Tyagi, C. Weber, B. Woolery, A. Yeoh, K. Zhang, and M. Bohr, "A 65nm logic technology featuring 35nm gate lengths, enhanced channel strain,

8 Cu interconnect layers, low-k ILD and 0.57 μm$^2$ SRAM cell," in *2004 International Electron Devices Meeting. San Francisco, CA*, 2005, pp. 657-60.

[2] P. Ranade, T. Ghani, K. Kuhn, K. Mistry, S. Pae, L. Shifren, M. Stettler, K. Tone, S. Tyagi, and M. Bohr, "High performance 35nm $L_{GATE}$ CMOS transistors featuring NiSi metal gate (FUSI), uniaxial strained silicon channels and 1.2nm gate oxide," in *International Electron Devices Meeting*, 2005, pp. 217-20.

[3] E. P. Gusev, V. Narayanan, and M. M. Frank, "Advanced high-K dielectric stacks with polySi and metal gates: recent progress and current challenges," *IBM Journal of Research and Development,* vol. 50, pp. 387-410, 2006.

[4] L. Chang, Y.-K. Choi, D. Ha, P. Ranade, S. Xiong, J. Bokor, C. Hu, and T.-J. King, "Extremely scaled silicon nano-CMOS devices," *Proceedings of the IEEE,* vol. 91, pp. 1860-73, 2003.

[5] Y.-K. Choi, T.-J. King, and C. Hu, "Nanoscale CMOS spacer FinFET for the terabit era," *IEEE Electron Device Letters,* vol. 23, pp. 25-7, 2002.

[6] D. Hisamoto, W.-C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T.-J. King, J. Bokor, and C. Hu, "FinFET-a self-aligned double-gate MOSFET scalable to 20 nm," *IEEE Transactions on Electron Devices,* vol. 47, pp. 2320-5, 2000.

[7] P. Verheyen, N. Collaert, R. Rooyackers, R. Loo, D. Shamiryan, A. De Keersgieter, G. Eneman, F. Leys, A Dixit, M.Goodwin, Y.S.Yim, M. Caymax, K. De Meyer, P. Absil, M. Jurczak, and S. Biesemans, "25% Drive Current Improvement for p-type Multiple Gate FET (MuGFET) Devices by Introduction

of Recessed $Si_{0.8}Ge_{0.2}$ in the Source and Drain Regions," in *2005 Symposium on VLSI Technology*, 2005, pp. 194-5.

[8] B. Doyle, B. Boyanov, S. Datta, M. Doczy, S. Hareland, B. Jin, J. Kavalieros, T. Linton, R. Rios, and R. Chau, "Tri-Gate fully-depleted CMOS transistors: fabrication, design and layout," in *2003 Symposium on VLSI Technology*, 2003, pp. 133-4.

[9] H. Lee, L.-E. Yu, S.-W. Ryu, J.-W. Han, K. Jeon, D.-Y. Jang, K.-H. Kim, J. Lee, J.-H. Kim, S. C. Jeon, G. S. Lee, J. S. Oh, Y. C. Park, W. H. Bae, H. M. Lee, J. M. Yang, J. J. Yoo, S. I. Kim, and Y.-K. Choi, "Sub-5nm all-around gate FinFET for ultimate scaling," in *2006 Symposium on VLSI Technology.*, 2006, pp. 58-9.

[10] L. Chang, "Nanoscale Thin-Body CMOS Devices," in *PhD Thesis*: University of California, Berkeley, 2003.

[11] E. Rauly, B. Iniguez, and D. Flandre, "Investigation of deep submicron single and double gate SOI MOSFETs in accumulation mode for enhanced performance," *Electrochemical and Solid-State Letters,* vol. 4, pp. G28-30, 2001.

[12] L. Chang, S. Tang, T.-J. King, J. Bokor, and C. Hu, "Gate length scaling and threshold voltage control of double-gate MOSFETs," in *International Electron Devices Meeting*, 2000, pp. 719-22.

[13] S. Wolf, *Silicon Processing for the VLSI Era* vol. 3 - The Submicron MOSFEt Lattice Press, Sunset Beach, California.

[14] T. Krishnamohan, C. Jungemann, and K. C. Saraswat, "A novel, very high performance, sub-20nm depletion-mode double-gate (DMDG) $Si/Si_xGe_{(1-x)}/Si$

channel PMOSFET," in *IEEE International Electron Devices Meeting*, 2003, pp. 687-90.

[15] "Taurus-Device, v. 2003.12," Synopsys Inc., 2003.

[16] *International Technology Roadmap for Semiconductors, PIDS Chapter,* http://public.itrs.net.

[17] S. Balasubramanian, L. Chang, B. Nikolic, and T.-J. King, "Circuit-Performance Implications for Double-Gate MOSFET Scaling below 25nm," *Proceedings of the 2003 Silicon Nanoelectronics Workshop,* pp. 16-17, June 2003.

# Chapter 3 : $V_T$ Adjustment via $L_{eff}$ Engineering

## 3.1 Introduction

Thin-body enhancement-mode MOSFETs with enhanced carrier mobilities, will likely be required to meet power-performance targets with gate-length scaling. The FinFET [1-3] is a promising structure because it offers the superior scalability of the double-gate structure together with a fabrication process and layout similar to that of the conventional bulk-Si MOSFET [3]. An undoped channel is desired to eliminate $V_T$ variations due to statistical dopant fluctuation (SDF) effects in the channel, and to attain the highest possible carrier mobilities to achieve high $I_{ON}$ [1]. However, it necessitates an alternative means for adjusting $V_T$. Figure 3.1 [4] shows the threshold voltage dependence on gate work function for undoped channel NMOS and PMOS FinFETs. In order to achieve symmetric $V_T$ (s), two different gate work functions are needed and dual gate work function ($\Phi_M$) technology has been explored as one of the techniques for adjusting the threshold voltages. Different work functions for NMOS and PMOS can be achieved by engineering the work function by methods such as phase engineering [5, 6], and masked ion-implantation [7-10]. These are promising approaches, but face drawbacks when applied to dense circuits.

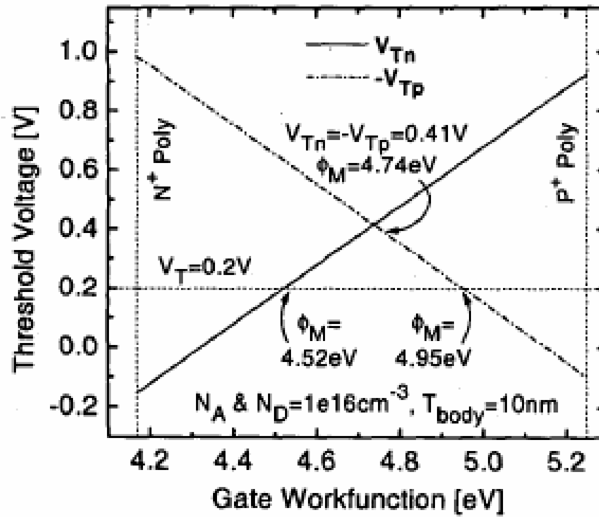



**Figure 3.1**: Threshold voltage requirement for undoped channel thin body NMOS and PMOS FinFETs [4]. Different gate work functions for NMOS and PMOS are needed to achieve symmetric high performance (HP) $V_T$ values, thus motivating dual gate technology.

For compact circuit layouts such as those used in static memory (SRAM) arrays [11], it is not possible to separately implant the gate electrodes in the n- and p-channel regions if the FinFET structure is employed. This is because the gate layer fills the entire region in-between the n-channel and p-channel fins (Figure 3.2) making it impossible to selectively and reliably change the work function of one of the FET gates. Thus, a single gate work function must be used for both devices, and another means for adjusting the threshold voltage should be investigated. In this chapter, $L_{eff}$ engineering (by S/D engineering) is explored as the alternative means to adjust $V_T$, and the advantages and limitations of this technique are presented.

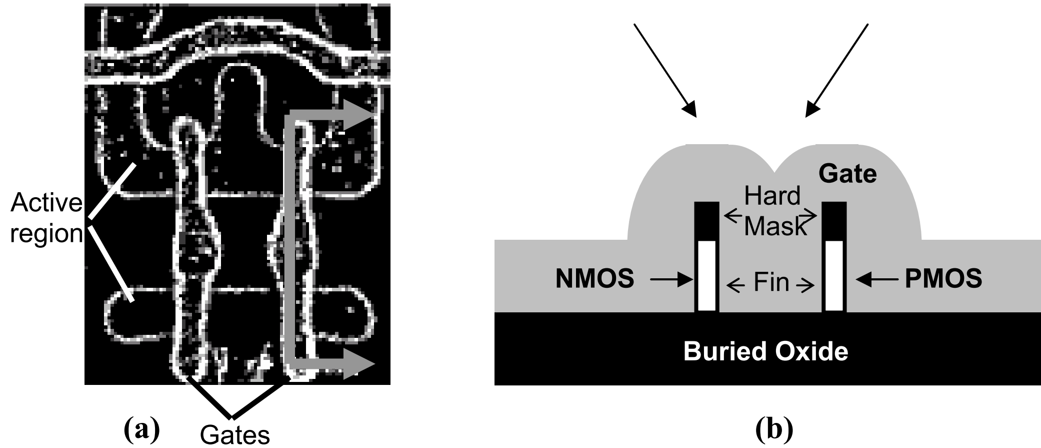



**Figure 3.2**: (a) SRAM layout and (b) schematic cross-section of a CMOS FinFET SRAM [11] inverter. If the separation between NMOS and PMOS active areas (Si fins) is less than twice the gate thickness, the gate completely fills the region in-between, so that it is difficult to achieve different gate work functions for the NMOS and PMOS gates.

## 3.2 $V_T$ Tuning by S/D Engineering

For a given technology node, the transistor physical parameters such as gate length ($L_G$) and gate oxide thickness ($T_{OX}$) are predetermined [12]. The fin thickness ($T_{Si}$) is somewhat flexible, based on the desired control of short channel effects [13, 14] and the available process window (to minimize $V_T$ variations due to fin line-edge roughness), but is likely to be identical for NMOS and PMOS devices for a given technology node. The threshold voltage ($V_T$) for a FinFET structure can therefore be adjusted using one or more of three device parameters: channel doping ($N_{Body}$), gate work function ($\Phi_M$), and electrical channel length ($L_{eff}$). Since channel doping and gate work function engineering are not good strategies for $V_T$ tuning for high density FinFET layouts, the methodology available for $V_T$ tuning is $L_{eff}$ engineering, keeping a fixed gate work function for all devices.

Previous studies [15] have shown that, in order to optimize the trade-off between parasitic series resistance and short-channel effects, a gate-underlapped structure will be required to achieve peak circuit performance for sub-20nm gate length. $V_T$ becomes a strong function of $L_{eff}$, in this case. The $L_{eff}$ can be adjusted by engineering the S/D profile, which in turn, is a function of the gate-sidewall spacer thickness ($L_{SP}$) and doping lateral abruptness ($\sigma_{SD}$). Since the S/D profile is the only control "knob", either the short channel effect or the $V_T$ (and not both at the same time) can be adjusted using this method. Our goal will be to achieve the desired $V_T$ and performance by the appropriate S/D design. To operate within a target $V_T$ roll-off constraint, the fin thickness ($T_{Si}$) is used as another variable (partially), but the same value must be maintained for NMOS and PMOS devices. In the following sections, the feasibility of implementing single-gate work function CMOS FinFETs is investigated by 2D and 3D device simulations.

### 3.2.1 Device Simulation Set-up

To investigate the methodology discussed above, 2-D and 3-D device simulations have been carried out using Taurus-Device [16] using self-consistent Drift-diffusion transport and 1D-Schrodinger (or MLDA quantum model) solutions. It should be noted that all simulations have (i) the quantum mechanical model for inversion charge confinement, (ii) no carrier temperature dependent current enhancement models and (iii) no mobility enhancement models (strain engineering) implemented. These simulations therefore, underestimate the performance ($I_{ON}$) owing to the models that have been used. 2D simulations have been implemented for LSTP FinFETs in this section to demonstrate

the $L_{eff}$ engineering concept. 3D simulations have been used to extend the concept to HP and LOP devices, and are presented in section 3.2.4.

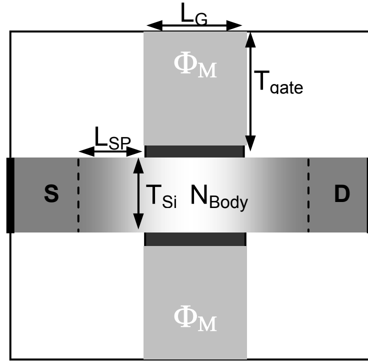



**Figure 3.3**: Schematic of the 2D simulation structure. $L_{SP}$ denotes the location at which the S/D gradient begins (referenced to the gate edge).

| Device Parameters | |
|---|---|
| Gate length $L_G$ | 18 nm |
| Gate-dielectric EOT | 9 Å |
| Body thickness, $T_{Si}$ | 9 or 12 nm |
| Gate thickness, $T_{Gate}$ | 27 nm |
| Body doping, $N_{body}$ | 1E16 $cm^{-3}$ |
| S/D doping, $N_{S/D}$ | 1E20 $cm^{-3}$ |
| Gate $\phi_M$ | 4.7 eV |
| $V_{DD}$ | 0.9 V |
| Target $I_{OFF}$ | 80 pA/µm |
| Target $I_{ON}$ | 880 µA/µm |

**Table 3.1**: Nominal device parameter values used for the simulations.

The 2D simulation structure and the device parameters (taken from ITRS specification for $L_G$=18nm LSTP device) are shown in Figure 3.3 and Table 3.1, respectively. Details of the 3D simulation structure and parameters are provided at the beginning of section 3.2.4. The gate work function is chosen to be 4.7eV (mid-gap) in order to be able to achieve symmetric $V_T$ values for both NMOS and PMOS devices with the same (or at least similar) S/D designs. The S/D implant conditions would be different for NMOS and PMOS devices, but the spacer thickness should be targeted to be identical, in order to have a reasonable fabrication process flow. In order to meet the required $I_{ON}/I_{OFF}$ specification, a thinner EOT is required, and so the corresponding high-performance EOT value has been chosen for this purpose. The $L_{eff}$ is defined to be the lateral separation between the locations at which the S/D doping falls to $1E10^{19}$ $cm^{-3}$. It is engineered using two parameters, shown as $L_{SP}$ (Figure 3.3) and $\sigma_{SD}$. $L_{SP}$ corresponds to

the location of the peak doping concentration of the S/D gradient region with respect to the gate-edge and $\sigma_{SD}$ corresponds to the lateral abruptness of the S/D doping (assuming it to be a Gaussian profile, as modeled in the device simulator). In a CMOS fabrication process, $L_{SP}$ and $\sigma_{SD}$ can be controlled by an optimal choice of the spacer thickness, and S/D implant and anneal conditions, respectively.

It will be shown that multiple S/D designs can achieve the target threshold voltage. The impact of variations on these designs is studied to identify the optimal design. The two main sources of $V_T$ variations are (i) line edge roughness in the fin and gate and (ii) statistical dopant fluctuation (SDF) in the S/D gradient region. The former effect is studied in this chapter by analyzing its effect on the SRAM noise margin, while SDF is modeled and studied using 3D simulations, in this and the next chapter.

### SRAM Noise Margin Variation Simulation Setup

For FinFETs, $T_{Si}$ should to be smaller than $L_G$ to suppress short channel effects. But, since $L_G$ and $T_{Si}$ are both defined by the same technology, the variation/roughness in the line-width can be assumed to have the same distribution. In this study, it is assumed that this distribution is Gaussian with mean equal to the nominal value and a standard deviation $\sigma$ given by $3\sigma$ = 10%*nominal for both $L_G$ and $T_{Si}$. $\sigma$ = 0.6nm for $L_G$ = 18nm. It should be noted that in a realistic situation, (i) the line-edge roughness and variations in $L_G$ and $T_{Si}$ are 3D effects and (ii) LER is likely to have a much retarded scaling rate [17, 18] than the nominal $L_G$ (strong function of the lithography technique). Therefore, the precise value of the variation in the SRAM Static Noise Margin (SNM) is likely to be inaccurate, but the trends for the different designs under consideration should be qualitatively valid.

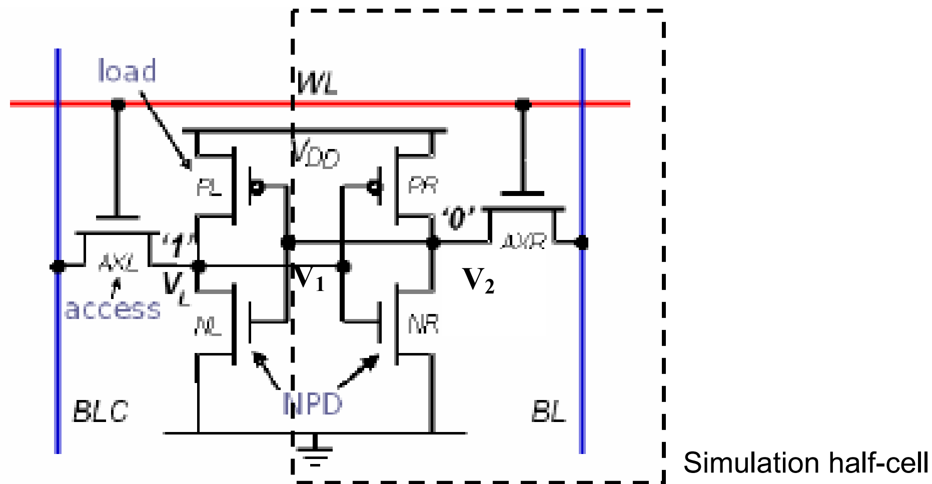



**Figure 3.4**: Schematic of a 6T SRAM cell. The dotted box encloses a half-cell, which is used in the simulation to generate the transfer characteristics for the butterfly curves. The butterfly curves do not depict the feedback and transient characteristics of the cell; therefore, simulating a half-cell is sufficient to characterize the noise margin.

The simulation and noise margin extraction methodologies used in this work have been adopted from the technique used by Guo et al in [19]. The transfer characteristic of each half cell (**Figure 3.4**) is simulated independently in order to keep the simulations less computationally intensive. This kind of setup is perfectly reasonable to simulate noise-margins, as we do not need to account for the circuit feedback and transients to determine the static noise margin. In terms of the simulation set-up, every transistor in the half-cell has its $L_G$ and $T_{Si}$ variations generated randomly from the above discussed Gaussian distribution. Once a pair of half-cells is simulated from randomly generated devices, their transfer characteristics give the butterfly curve, from which the noise margin can be extracted by the traditional method.

### 3.2.2 Impact of $L_{eff}$ on $I_{ON}$-$I_{OFF}$

For the simulation structure considered, the $L_{SP}$ is varied from 4nm to 14nm and the lateral abruptness factor $\sigma_{SD}$ is varied from 1.5nm to 5nm to span a reasonable range of $L_{eff}$ values. The fin thickness ($T_{Si}$) is varied from $L_G/2$ to $2L_G/3$, which is a suitable range for thin-body devices for good control of short channel effects [13]. **Figure 3.5** summarizes the $I_{ON}$-$I_{OFF}$ tradeoff for multiple combinations of $L_{SP}$ and $\sigma_{SD}$, for different values of $T_{Si}$. The target device is a LSTP device with a low $I_{OFF}$ specification.

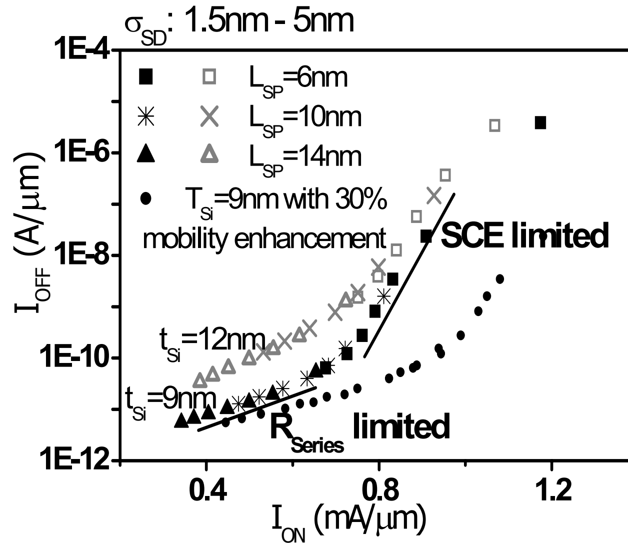



**Figure 3.5**: Trade-off between $I_{ON}$ and $I_{OFF}$ showing short channel effect limited and series resistance limited regions of operation. All combinations of $L_{SP}$ and $\sigma_{SD}$ fall on a single curve and multiple combinations with same $L_{eff}$ can yield the required value of $I_{OFF}$.

The scatter plot shows two regions of operation. In the low leakage regime, the $L_{eff}$ is much larger than $L_G$ and the $I_{ON}$ is limited by series resistance. In the high performance (hence, high leakage as well) regime, the $L_{eff}$ becomes comparable or shorter than $L_G$ and $I_{ON}$ is limited by the need to increase $V_T$ to achieve low leakage,

giving rise to the SCE (Short-Channel effect) limited operation regime. It is possible to cover a large range of $I_{OFF}$ and $I_{ON}$ depending on the extent of SCE immunity desired.

When $T_{Si} = 2L_G/3$, owing to poor short channel effect immunity, it is difficult to meet the LSTP $I_{OFF}$ specifications without a severe hit in $I_{ON}$. A larger $T_{Si}$ requires a larger $L_{eff}$ to reduce the off-state leakage. But that moves the device deeper into the series-resistance limited regime and can severely degrade the performance. A good choice of $T_{Si}$ is therefore critical especially for Low Standby Power operation. With $T_{Si} = L_G/2$, it is possible to achieve performance within 20% of the target specification, which can be further enhanced using mobility enhancement techniques. The structure simulated in this section has a uniformly thin fin even in the S/D regions, which results in large parasitic S/D series resistance. It has been experimentally demonstrated [20] that epitaxial SiGe can be used to flare out the S/D regions and significantly reduce the series resistance problem. In section 3.2.4, we will see that for High Performance and Low Operating Power applications, the tradeoff between $I_{OFF}$ and $I_{ON}$ is not as severe, since the target values for $I_{OFF}$ are higher.

It must be noted that all the data-points for a given $T_{Si}$ lie on the same curve even though each data-point corresponds to a different combination of $L_{SP}$ and $\sigma_{SD}$. Thus, multiple designs can provide the same $I_{ON}$ and $I_{OFF}$. This shows that a fundamental device parameter ($L_{eff}$) determined by $L_{SP}$ and $\sigma_{SD}$ is of key importance in controlling the $V_T$. **Figure 3.6** shows the S/D doping profiles corresponding the multiple solutions of interest to us, namely for $I_{OFF}$ = 80pA/μm. All the designs correspond to the same $L_{eff}$ (as expected). The $L_{eff}$ is defined as the distance between the points in the source and drain doping profiles, where the doping concentration is $1E10^{19}$ $cm^{-3}$ and this corresponds to

the point of carrier injection at the source-channel barrier. **Figure 3.7** shows that two of the extreme S/D designs yield the same SCE immunity. This allows for flexibility in process design, showing that we do not have to restrict ourselves to ultra steep S/D profiles and ultra thin spacers. In the next section, we shall briefly go over some analysis of variations to see if one design is actually more optimal than the other.

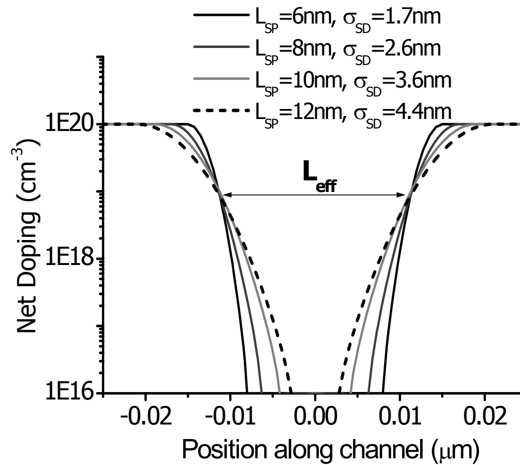



**Figure 3.6**: Multiple combinations of $L_{SP}$ and $\sigma_{SD}$ are optimal, all with the same $L_{eff}$ (distance between the points where the S/D doping concentrations = $1\times10^{19}$ $cm^{-3}$).

In order to use $L_{eff}$ engineering as a $V_T$-tuning methodology, it is important to quantify its $V_T$ tuning range and also note the position of the desired device design, on the $V_T$ roll-off curve. **Figure 3.8** shows a plot of $V_T$ vs $L_{eff}$ for $T_{Si}$ = 9nm. In this study, $V_T$ is defined as the gate voltage at which the drain current is equal to 100nA/μm. This plot looks very similar to a $V_T$ roll-off plot with the series resistance regime and short channel regime clearly identifiable. In order to avoid sensitivity to variations in gate length, it is preferable to tune the $V_T$ close to the series resistance regime. At $L_G$ = 18nm, about 0.2V of $V_T$ tuning is possible with a DIBL less than 100mV/V.

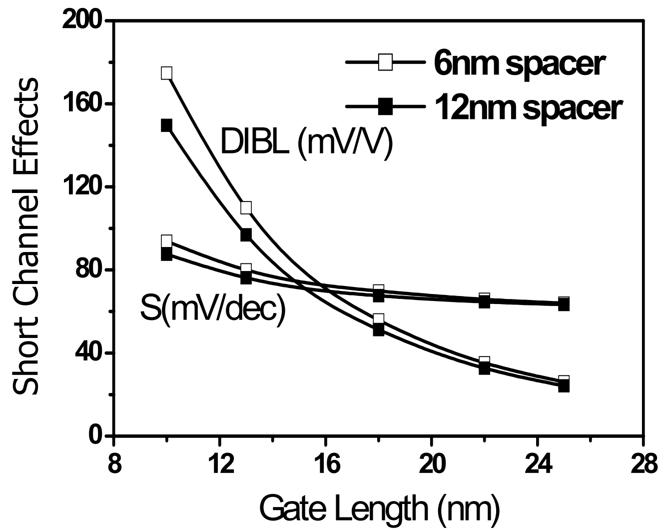



**Figure 3.7**: SCE comparison for two optimal S/D doping profile designs. A thin gate-sidewall spacer (small $L_{SP}$) with a steep doping gradient (small $\sigma_{SD}$) yields similar short channel behavior as a thick gate-sidewall spacer with relaxed $\sigma_{SD}$, thus providing flexibility in process design.

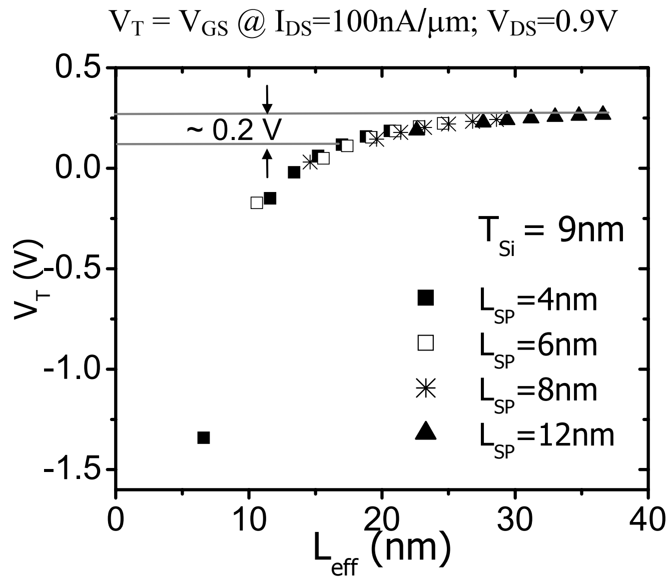



**Figure 3.8**: $V_T$ as a function of electrical channel length ($L_{eff}$) showing the range of $V_T$ tuning. It is possible to adjust $V_T$ by ~0.2V with reasonable short channel immunity.

In a CMOS technology, $L_{SP}$ should be the same for NMOS and PMOS devices. We therefore, repeat the methodology for PMOS FinFETs with the same $T_{Si}$ and the same range of S/D designs as done for the NMOS FinFETs. **Figure 3.9**(a) shows that PMOS $I_{OFF}$ specifications can also be met using the same $L_{SP}$ designs as for NMOS. This technique can therefore be used to implement CMOS technology for FinFETs. The performance specification is met within 20% of the target value (**Figure 3.9**(b)) without any mobility enhancement.

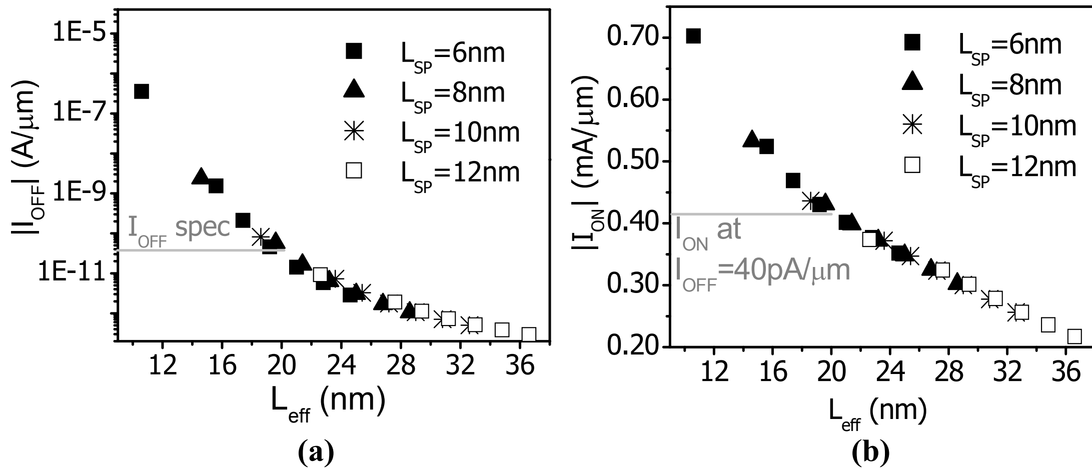



**Figure 3.9**: PMOS performance specifications, (a) $I_{OFF}$=40pA/µm and (b) $I_{ON}$=440µA/µm, can be met (within 20% of target) with the same values of $L_{SP}$ as for the NMOS device.

### 3.2.3 Choice of a Good S/D Design: Impact of Variations

In this section, the various S/D designs possible to achieve the desired $V_T$ have been further analyzed. As long as these designs share the same physical parameters such as $L_{eff}$, $T_{OX}$, and $T_{Si}$, they show very similar behavior in performance and short channel immunity. Another important aspect is the immunity to variations from sources like (i) statistical dopant fluctuation (SDF) effects in the S/D regions and, (ii) $L_G$ and $T_{Si}$ variations. A detailed description of the methodology developed for studying SDF effects

is given in the next chapter. Here, we shall simply use that methodology to extract the $V_T$ variations for different S/D designs.

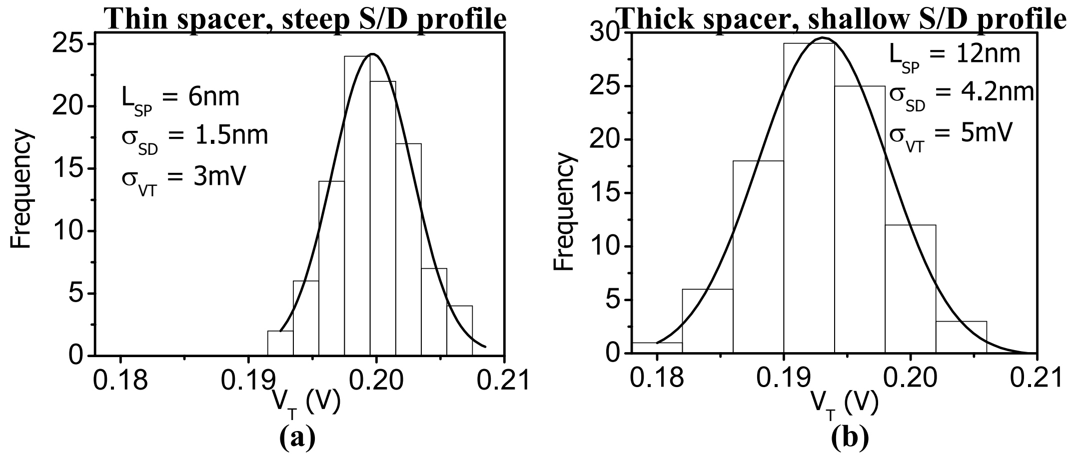



**Figure 3.10**: Statistical $V_T$ data obtained through 100 3-D atomistic device simulations indicate that SDF-induced variations will not be of concern at $L_G$ = 18nm. However, with $L_G$ scaling, lean spacers and steep S/D profiles will be favorable for minimizing $V_T$ fluctuations due to SDF. The Si fin height was assumed to be 50nm.

We first look at the effect of SDF on the device performance. 3D atomistic simulations have been carried out for the $L_G$ = 18nm LSTP devices for two of the extreme S/D designs with (i) $L_{SP}$ = 6nm, $\sigma_{SD}$ = 1.5nm and, (ii) $L_{SP}$ = 12nm, $\sigma_{SD}$ = 4.2nm, identified in the previous section. As shown in **Figure 3.10**, the $\sigma_{VT}$ is insignificant and SDF due to S/D dopants is probably not of concern at this gate length for a FinFET design. This provides a lot of process flexibility at this technology node. This is consistent with other data that has been reported on FinFET devices [21]. The point to note though is that there is a difference in the two designs in terms of variation tolerance, and as we scale down to future technology nodes, ultra thin spacers and steep S/D gradients will be essential to suppress the $V_T$ variations due to SDF in S/D regions.

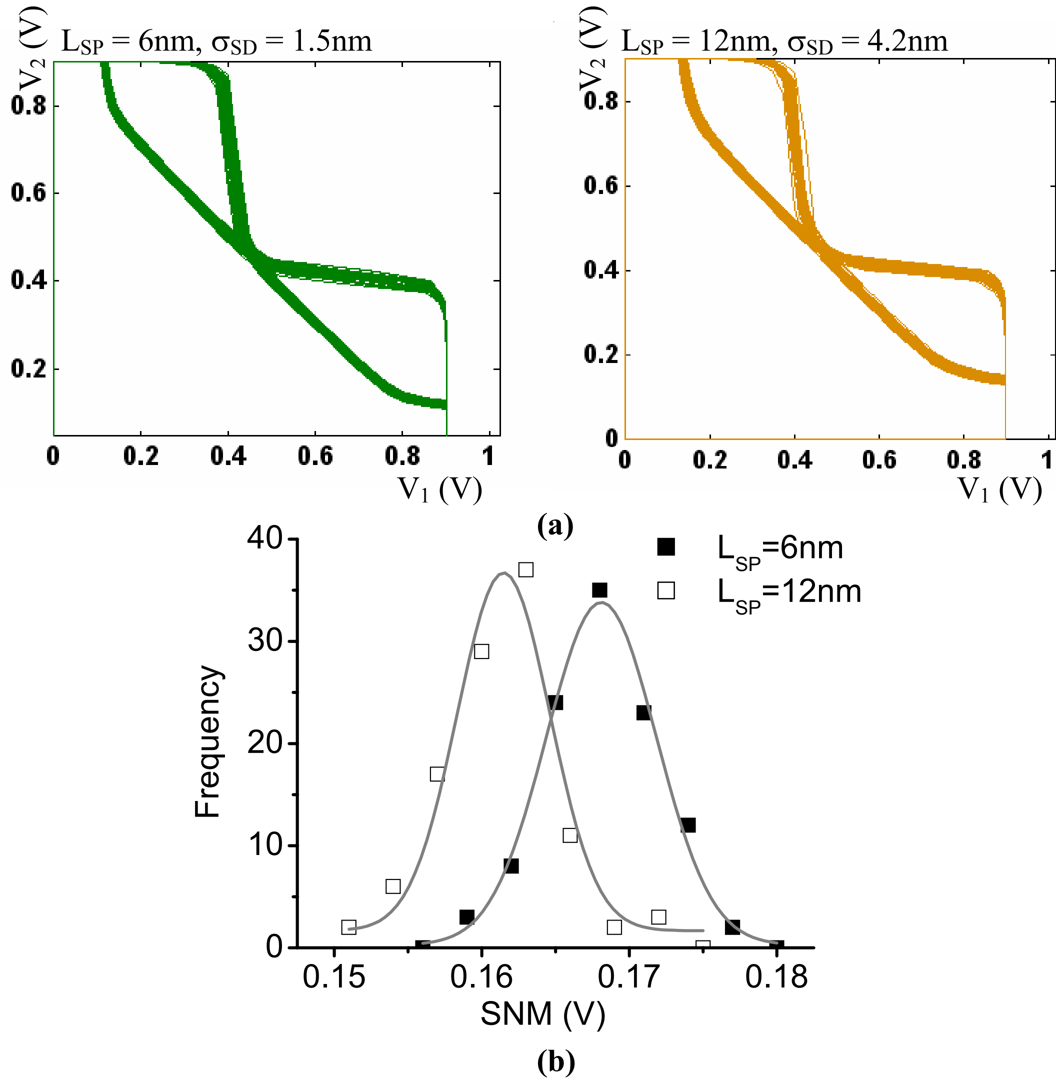



**Figure 3.11**: (a) SRAM noise margin simulation data for the two extreme S/D designs with $3\sigma = 0.1L_G$ variation in $L_G$ and $T_{Si}$. (b) The SNM variation predicts small $\sigma_{SNM}$ due to line width fluctuations at $L_G$=18nm. The two S/D designs have very marginal difference in their response to LER induced fluctuations, as expected.

Next, we look at the effect of line width variations on the SRAM noise margin (SNM). The methodology used for this analysis has been adopted from [19]. **Figure 3.11** summarizes the results obtained from extractions of SNM from an ensemble of 108 pairs

of SRAM half-cell simulations. **Figure 3.11**(a) shows the variations in the butterfly curves for the two extreme designs of $L_{SP}$=6nm and $L_{SP}$=12nm. The extracted statistical plot of SNM is shown in **Figure 3.11**(b). For $L_G$=18nm, LSTP technology, the simulated variation in SNM given by $\sigma_{SNM}$ is quite small for both the designs: $\sigma_{SNM,\ 6nm}$ = 4.1mV and $\sigma_{SNM,\ 12nm}$ = 4.3mV. These values are reasonably small, but due to inaccuracies in 2D Line-Edge Roughness (LER) simulations and the line width variation amount, the trend rather than the precise value of $\sigma_{VT}$ is important. The design with $L_{SP}$=6nm is marginally better than the $L_{SP}$=12nm design, but they show more or less similar variation in SNM due to LER. This is as expected because these devices have the same physical parameters ($L_{eff}$, $T_{OX}$ and $T_{Si}$), short channel immunity and occupy the same position on the $V_T$ roll-off curve (the $V_T$-$L_{eff}$ curve in this case), making them similar in their sensitivity to gate-length and body-thickness variations.

### 3.2.4 Evaluating HP and LOP Designs with Single Gate Work Function

In this subsection, the analysis of using S/D engineering as a methodology to tune $V_T$ will be extended to High Performance (HP) and Low Operating Power (LOP) devices. For HP devices, due to high $I_{ON}$ and $I_{OFF}$ (compared to LSTP), it is difficult to design in the series resistance regime. It would therefore, be particularly important in this case, to look at the position of the optimized design on the $V_T$ roll-off curve. For this study, 3D simulations of MuGFET devices [22] have been carried out, with hp2007 and lop2007 specifications taken from the ITRS roadmap. The MuGFET is similar to the FinFET, except that it utilizes the top surface of the fin, along with the fin sidewalls, for conduction. The simulated structures are shown in **Figure 3.12**. The *flared fin* MuGFET has been compared with the *flat fin* MuGFET to demonstrate the benefits arising from

lower S/D series resistance in the flared fin case. The device parameters used for this study are summarized in Table 3.2 [12].

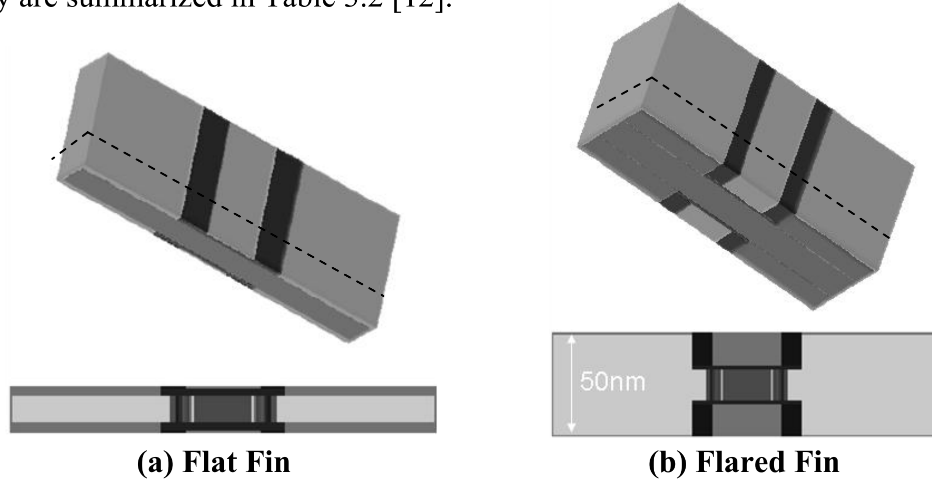



**Figure 3.12**: (a) Flat fin and (b) Flared fin 3D MuGFET structures used in device simulations. The flared fin structure has been simulated to demonstrate the advantage in its performance coming from reduced parasitic S/D series resistance.

| $L_G$ | 24nm | 35nm |
|---|---|---|
| $T_{OX}$ (EOT) | 11A | 12A |
| $T_{Si}$ | 10, 15, 20nm | 10, 15, 20nm |
| $L_{SP}$ | 3, 5, 8, 10nm | 3, 5, 8, 10nm |
| $\sigma_{SD}$ | 2.5 – 4nm | 2.5 – 4nm |
| $V_{DD}$ | 1.1V | 0.8V |
| Spec. Contact Res. | 9.5E-8 $\Omega$-cm$^2$ | 9.5E-8 $\Omega$-cm$^2$ |
| Target $I_{OFF}$ | 0.2µA/µm | 5nA/µm |
| Target $I_{ON}$ | 1.2mA/µm | 573µA/µm |
| Target $V_T$ | 0.15V | 0.25V |

**Table 3.2**: Device parameters used for 3D HP and LOP MuGFET simulations.

## $V_T$ Adjustment and Sensitivity to Variations

With the different combinations of $L_{SP}$ and $\sigma_{SD}$, a reasonable range of $L_{eff}$ values could be simulated. The $V_T$ vs. $L_{eff}$ for different values of $T_{Si}$ for HP and LOP MuGFETs

is shown in **Figure 3.13**. In this case, the definition of the threshold voltage is a modification of the simple constant current $V_T$ definition. The $g_{m,max}$ $V_T$ is extracted from the low $V_{DS}$ ($V_{DS}$=50mV) $I_D$-$V_G$ data and $I_D$=$I_{DLOW}$ at that $V_T$ is evaluated. The $V_T$ is then defined as the gate voltage at which the $I_D$=$I_{DLOW}$ for $V_{DS}$=$V_{DD}$. As can be seen from **Figure 3.13**, the flared fin and flat fin MuGFETs show similar $V_T$ vs $L_{eff}$ behavior. This is as expected because the threshold voltage is a biasing condition in weak inversion for which the series resistance does not play a significant role.

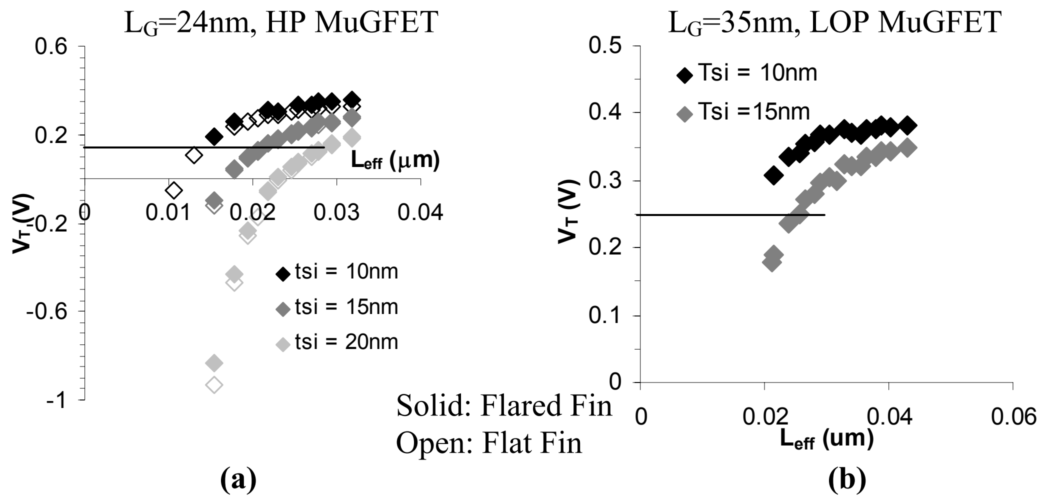



**Figure 3.13**: The target $V_T$ values for (a) high performance and (b) low operating power MuGFETs are achievable by $L_{eff}$ engineering. A thicker $T_{Si}$ requires a larger $L_{eff}$ for the same $V_T$ and gives better immunity to variations with some tradeoff in performance.

It can be clearly seen from **Figure 3.13** that the $V_T$ for HP and LOP devices are achievable using $L_{eff}$ engineering. An interesting point to be noted is as follows. For a given gate length, when $T_{Si}$ is increased, the short channel control is degraded leading to a decrease in $V_T$. Therefore, to achieve the target $V_T$, $L_{eff}$ has to be larger for the thicker $T_{Si}$ case. The increase in $L_{eff}$ counterbalances the effect of increase in $T_{Si}$ such that the

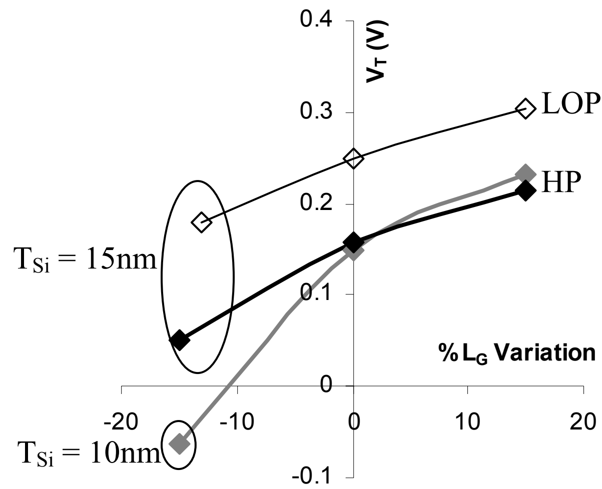



**Figure 3.14**: $V_T$ sensitivity to gate length variation for HP and LOP devices. A thicker $T_{si}$ device pushes the operation towards the series resistance limited regime, thus leading to a trade-off between performance and tolerance to LER induced fluctuations.

optimal design moves towards (or deeper into) the series resistance regime, making it a better device in terms of sensitivity to variations. This of course, leads to a trade-off between variation tolerance and performance. **Figure 3.14** shows the variation in $V_T$ for +/-15% variation in gate length for optimal designs at $T_{Si}$ = 10nm and 15nm. Consistent with the above explanation, the $V_T$ sensitivity for the thinner $T_{Si}$ is worse than that for the thicker body case at the target $V_T$. The reduction in $I_{ON}$ for the HP device, caused by the larger $L_{eff}$ is shown in **Figure 3.15**(a). With a thicker $T_{Si}$, the $I_{ON}$ can degrade by about 20-30%, which is a significant sacrifice to achieve a significant tolerance to variation. A careful choice of $T_{Si}$ and $L_{eff}$ is therefore very crucial to get the best performance within a target specification of $V_T$ sensitivity. Moreover, since $L_{eff}$ and $T_{Si}$ are the two variables, only two out of the three crucial parameters (namely $I_{ON}$, SCE and $\sigma_{VT}$) can be precisely controlled, which is a limitation of this technique. Another point to be noted is that the

flared fin device shows a significant $I_{ON}$ advantage over the flat fin device, as expected. On the other hand, there is no significant benefit in $I_{OFF}$ (**Figure 3.15**(b)), which is again, as expected as $I_{OFF}$ is not significantly affected by S/D series resistance.

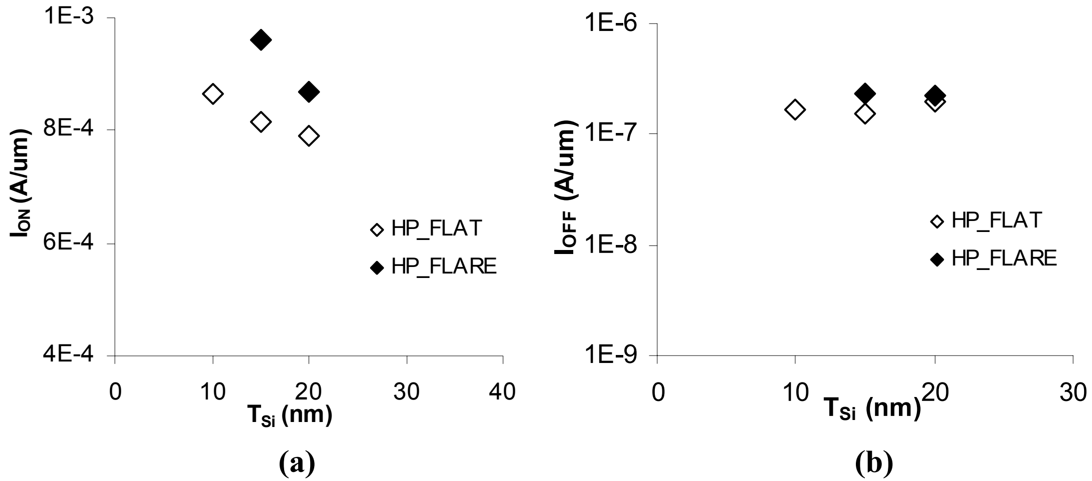



**Figure 3.15**: (a) Performance trade-off with increasing $T_{Si}$ (for better variation tolerance). Moreover, flared fin architecture shows significant benefit over flat fin MuGFET demonstrating the need for low resistance S/D architectures. (b) $I_{OFF}$ is not affected significantly by S/D parasitic series resistance, as expected.

It has been shown that it is feasible to adjust the $V_T$ values for HP and LOP devices (with some variation tolerance) by an intelligent choice of body thickness and S/D design. A further benefit for this method would exist if both HP and LOP MuGFETs can be designed with the similar physical parameters (same $T_{Si}$, $L_{SP}$ and $\sigma_{SD}$ but of course, different $L_G$ and $T_{OX}$). This would enable easy integration of HP and LOP devices on the same chip. In order to look into this, the data in **Figure 3.13** is replotted slightly differently and is shown in **Figure 3.16**. The $V_T$-$L_{eff}$ for HP and LOP devices are plotted together on the same plot for a fixed value of $T_{Si}$. The optimum gate underlap/overlap for

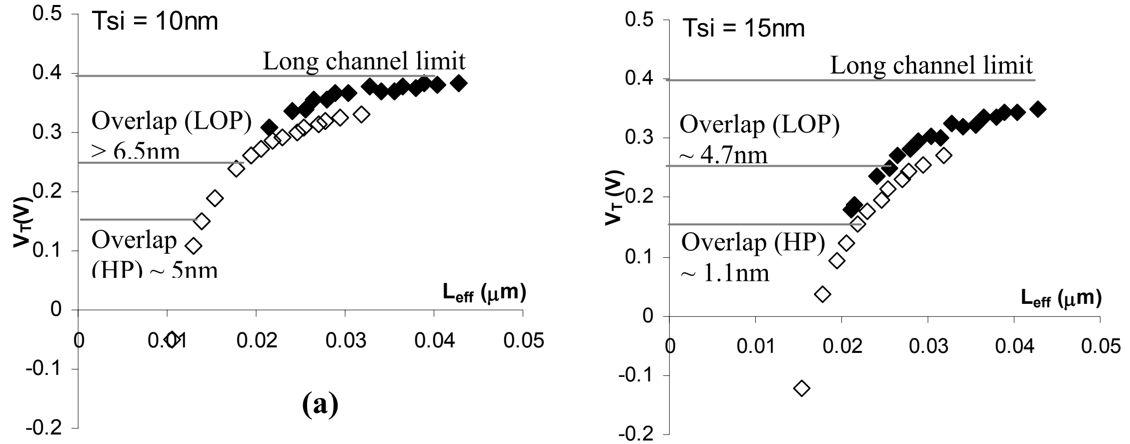



**Figure 3.16**: Optimum gate overlaps for LOP and HP devices for (a) $T_{Si}$=10nm and (b) $T_{Si}$ = 15nm respectively. Identical combination of $T_{Si}$, $L_{SP}$ and $\sigma_{SD}$ cannot be used for HP and LOP as well as LOP MuGFETs.

achieving the required $V_T$ values (0.15V for HP, 0.25V for LOP) are also indicated. It is evident that it is not possible to achieve HP and LOP requirements with identical $T_{Si}$, $L_{SP}$ and $\sigma_{SD}$, and at least one of these three parameters has to be different to achieve the target specifications. In a typical process, the $T_{Si}$ and $L_{SP}$ are likely to be common for all devices to have a reasonable process flow. It would therefore, be necessary to employ multiple S/D implantation conditions to achieve the NMOS and PMOS $V_T$ values for HP and LOP devices on the same chip. We shall choose $T_{Si}$ = 15nm for both types of devices for the remaining discussion. **Figure 3.17** summarizes the $V_T$ variation with $T_{Si}$ and $L_G$ variations for optimum HP and LOP designs at $T_{Si}$=15nm. The $T_{Si}$ is varied from 10-20nm and the $L_G$ is varied by +/- 15%, which is quite significant for any given technology and therefore, the variation in $V_T$ also seems high for these designs. This of course, demonstrates the need for a good control of line edge roughness induced fluctuations to keep $V_T$ variation under control.

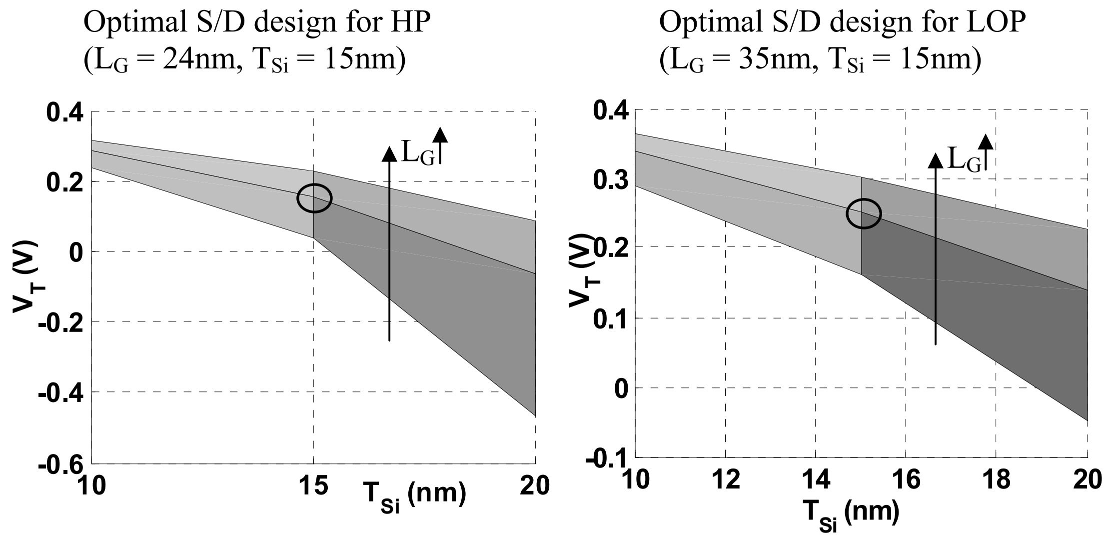



**Figure 3.17**: $V_T$ variation with large variations in $T_{Si}$ and $L_G$ (LER induced variations) for optimum HP and LOP MuGFETs at $T_{Si}$=15nm. Good control of LER is critical to ensure good control of $V_T$ fluctuations.

## 3.3 S/D Process Requirements for $L_{eff}$ Engineering

The process requirements for implementing this technology will now be discussed through 3D process simulations of S/D engineering. The main goal for this study is to examine $L_{eff}$ uniformity along the fin height and study the effect of implant conditions on the same. The practical issues related with an "optimum" process will also be discussed. For this analysis, the physical parameters considered are: $L_G$=55nm, $T_{OX}$=2nm, $T_{Si}$=20nm, $L_{SP}$<50nm and $H_{fin}$=60nm. These parameters are chosen based on a current generation, industry developed LSTP MuGFET process in order to be as close to a realistic design as possible. The simulations have been carried out using Taurus-Process [23] simulator.

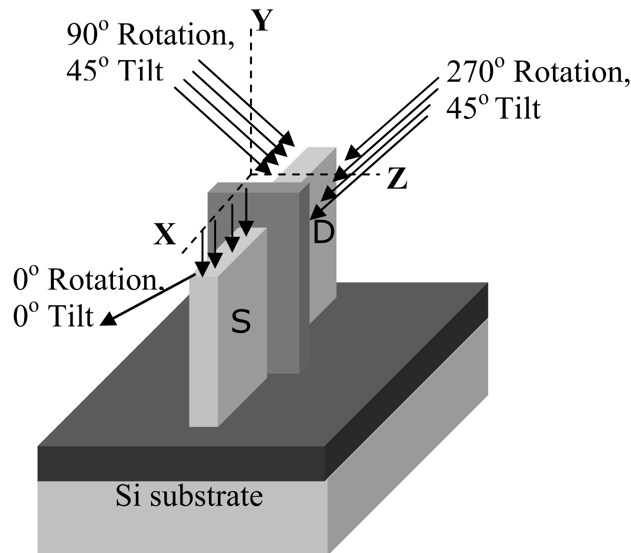



**Figure 3.18**: Ion implantation convention used in the 3D process simulations.

### 3.3.1 Effect of Ion-Implantation Conditions

The controllable parameters of an ion-implantation process are the energy, dose, implant tilt and wafer rotation. The dose should be high in our case, in order to minimize S/D series resistance by maximizing the active doping concentration. The ion implantation convention followed in the process simulations is as shown in **Figure 3.18**. To reduce computational time, the gate and spacer together are represented as one ideal oxide layer (typical spacer material) covering the whole fin. The thickness of this layer is the same the total gate stack thickness in the real process. The exact shape of the spacer is not simulated to minimize simulation time and complexity, but it was verified through some initial simulations, that the gate-stack is tall enough that the shape and the material of the spacer do not affect the final S/D profile.

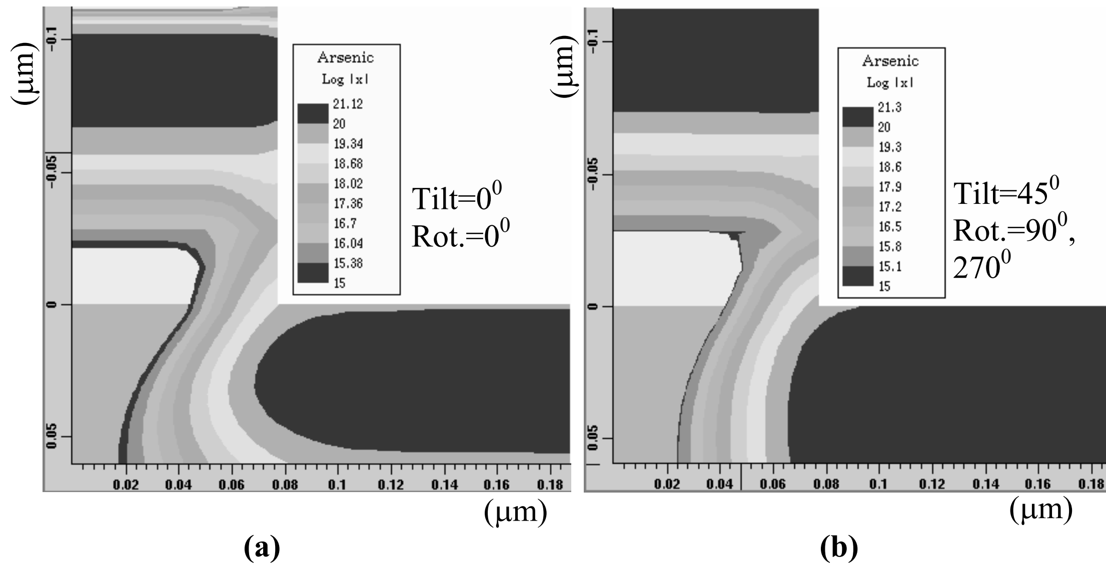



**Figure 3.19**: As implanted doping profiles through the centre of the fin for (a) tilt = $0^0$, rotation = $0^0$ and, (b) tilt = $45^0$, rotation = $90^0$, $270^0$. Implant conditions: Species: As, Energy: 35keV, Dose: 3E15cm$^{-2}$. A tilted implant is necessary to ensure better uniformity along fin height.

### Effect of Tilt and Rotation

The $L_{eff}$ of a MuGFET is a function of $L_{SP}$ and $\sigma_{SD}$. The $L_{SP}$ fixes the point where the S/D gradient begins and is equal to the spacer thickness of the process. The $\sigma_{SD}$ is a function of the (i) implant energy, tilt and rotation and (ii) activation anneal conditions. For rotations of $0^0$ (untilted), $90^0$ and $270^0$ (tilted) (**Figure 3.18**), $\sigma_{SD}$ is determined by the lateral straggle and lateral diffusion of the implanted species. Two implantation conditions were simulated: (i) single pass tilt=rotation=0 (which is a straight down implantation condition) and, (ii) double pass implant with tilt=$45^0$, rotation=$90^0$ and, tilt=$45^0$, rotation=$270^0$ (which is an angled implant on both sides of the fin). The other process conditions were: Species=Arsenic, Energy=35keV, Dose=3E15 cm$^{-2}$ and RTA @

$1000^0$C for 10s. The as implanted doping profiles along the z=0 plane passing through the center of the fin, are shown in **Figure 3.19**. The uniformity along the fin height is better for the two pass tilted implants, which is as expected. For a straight down implant (tilt=rotation=$0^0$), the implant energy should be high to ensure heavy doping along the entire height of the fin, but it leads to a large lateral straggle, lateral gradient and non-uniformity along the fin height. For tilted implants on the other hand, the energy can be much lower as the implant range only needs to be approximately half the fin thickness (which is much smaller than the fin height!), and can potentially yield steep S/D profiles.

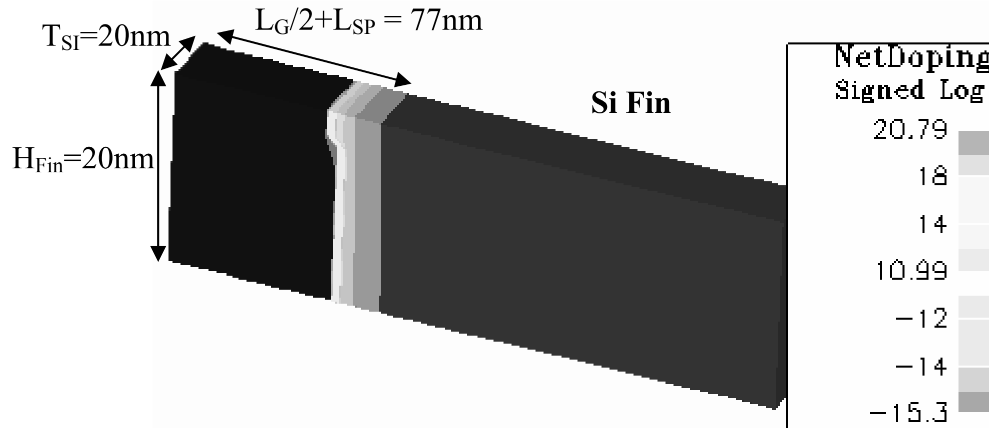



**Figure 3.20**: Net doping profile after implant and RTA, for a two pass tilted implant at 5keV shows good uniformity along the fin height. The actual doping concentrations and gradients are shown in Figure 3.21

## Effect of Implant Energy

To study the implant energy requirement, double-pass tilted implants were simulated, with the same dose and activation conditions as before, with implant energy varying from 5keV to 30keV. The lower limit of energy was set to 5keV in order to have the implant range > $T_{Si}/2$. **Figure 3.20** shows the doping profile for a 5keV implant after RTA. **Figure 3.21** shows the doping profile for 5keV and 10keV implants after the RTA

simulation. The lateral abruptness after the RTA simulation was observed to be 3nm/decade and 3.5nm/decade for the 5keV and 10keV implants respectively.

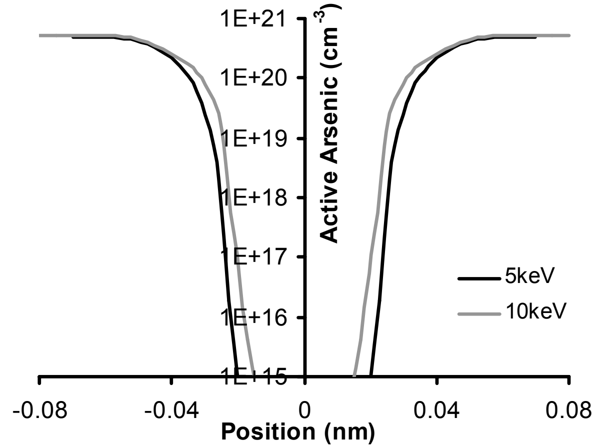



**Figure 3.21**: S/D profile along the z=0 plane, after implant and anneal for 5keV and 10keV energies, showing lateral abruptness of 3nm/decade and 3.5nm/decade respectively. This simulation has been done for a complete MuGFET structure.

For implementing $L_{eff}$ as a method to control $V_T$, good control on the S/D gradient and uniformity are required, which requires large tilt angle, low energy, high dose ion implantation. In dense memory array architectures, this is not a practical solution for doping the fin. Alternative doping strategies such as plasma doping [24] will be required to achieve good control of the S/D profiles.

## 3.4 Summary

In this chapter implementation of single gate work-function technology for FinFETs has been explored. It has been shown that $L_{eff}$ engineering is an effective methodology to tune the $V_T$ of NMOS and PMOS devices, for high performance, low operating power and low standby power categories of devices with a mid-gap work function gate. For given physical parameters of $L_G$, $T_{OX}$ and $T_{Si}$, target $V_T$ values can be

achieved by engineering the spacer thickness ($L_{SP}$) and the S/D gradient ($\sigma_{SD}$) to yield multiple designs, all of which share the same value of $L_{eff}$. For channel lengths larger than $L_G$=15nm, the variation in $V_T$ caused by statistical dopant fluctuation (SDF) effects in the S/D regions is quite small. Therefore, all optimum S/D designs are equally good in terms of variation tolerance, thus providing flexibility in process design. With $L_G$ scaling however, ultra thin spacers and steep S/D gradients may be necessary to suppress $V_T$ fluctuations due to SDF. Additionally, a thicker silicon body would be better to achieve better tolerance to line edge roughness (LER) induced $V_T$ fluctuations, at the cost of some performance ($I_{ON}$). It has been shown through process simulations that low energy, high dose, tilted implants will be required to achieve uniformly steep S/D junctions in the FinFET structure. Alternatively, methods such as plasma doping can practically achieve such profiles.

## 3.5 References


[1] D. A. Antoniadis, "MOSFET scalability limits and "new frontier" devices," in *2002 Symposium on VLSI Technology*, 2002, pp. 2-5.

[2] D. Hisamoto, W.-C. Lee, J. Kedzierski, E. Anderson, H. Takeuchi, K. Asano, T.-J. King, J. Bokor, and C. Hu, "A folded-channel MOSFET for deep-sub-tenth micron era," in *International Electron Devices Meeting* 1998, pp. 1032-34.

[3] N. Lindert, L. Chang, Y.-K. Choi, E. H. Anderson, W.-C. Lee, T.-J. King, J. Bokor, and C. Hu, "Sub-60-nm quasi-planar FinFETs fabricated using a simplified process," *IEEE Electron Device Letters,* vol. 22, pp. 487-9, 2001.

[4] L. Chang, S. Tang, T.-J. King, J. Bokor, and C. Hu, "Gate length scaling and threshold voltage control of double-gate MOSFETs," in *International Electron Devices Meeting*, 2000, pp. 719-22.

[5] J. A. Kittl, A. Veloso, A. Lauwers, K. G. Anil, C. Demeurisse, S. Kubicek, M. Niwa, M. J. H. van Dal, O. Richard, M. A. Pawlak, M. Jurczak, C. Vrancken, T. Chiarella, S. Brus, K. Maex, and S. Biesemans, "Scalability of Ni FUSI gate processes: phase and Vt control to 30 nm gate lengths," in *2005 Symposium on VLSI Technology*, 2005, pp. 72-73.

[6] A. Veloso, T. Hoffmann, A. Lauwers, S. Brus, J. F. de Marneffe, S. Locorotondo, C. Vrancken, T. Kauerauf, A. Shickova, B. Sijmus, H. Tigelaar, M. A. Pawlak, H. Y. Yu, C. Demeurisse, S. Kubicek, C. Kerner, T. Chiarella, O. Richard, H. Bender, M. Niwa, P. Absil, M. Jurczak, S. Biesemans, and J. A. Kittl, "Dual work function phase controlled Ni-FUSI CMOS (NiSi NMOS, $Ni_2Si$ or $Ni_{31}Si_{12}$ PMOS): manufacturability, reliability & process window improvement by sacrificial SiGe cap," in *2006 Symposium on VLSI Technology*, 2006, pp. 94-5.

[7] J. Kedzierski, D. Boyd, C. Cabral, Jr., P. Ronsheim, S. Zafar, P. M. Kozlowski, J. A. Ott, and I. Meikei, "Threshold voltage control in NiSi-gated MOSFETs through SIIS," *IEEE Transactions on Electron Devices,* vol. 52, pp. 39-46, 2005.

[8] W. P. Maszara, Z. Krivokapic, P. King, J. S. Goo, and M. R. Lin, "Transistors with dual work function metal gates by single full silicidation (FUSI) of polysilicon gates," in *IEEE International Electron Devices Meeting*, 2002, pp. 367-70.

[9] P. Ranade, Y.-K. Choi, D. Ha, A. Agarwal, M.Ameen, and T.-J. King, "Tunable work function molybdenum gate technology for FDSOI-CMOS," in *IEEE International Electron Devices Meeting*, 2002, pp. 363-6.

[10] K. Shin and T.-J. K. Liu, "Reduction in Molybdenum Gate Work Function via High Temperature Forming Gas Annealing," *IEEE Electron Device Letters,* submitted 2006.

[11] C. B. Oh, M. H. Oh, H. S. Kang, C. H. Park, B. J. Oh, Y. H. Kim, H. S. Rhee, Y. W. Kim, and K. P. Suh, "Double raised source/drain transistor with 50 nm gate length on 17 nm UTF-SOI for 1.1 $\mu m^2$ embedded SRAM technology," in *IEEE International Electron Devices Meeting*, 2003, pp. 31-34.

[12] *International Technology Roadmap for Semiconductors, PIDS Chapter,* http://public.itrs.net.

[13] B. Doyle, R. Arghavani, D. Barlage, S. Dutta, M. Doczy, J. Kavalieros, A. Murthy, and R. Chau, "Transistor Elements for 30nm Physical Gate Lengths and Beyond," *Intel Technology Journal,* vol. 6, No. 2, pp. 42-54, May 2002.

[14] D. J. Frank, Y. Taur, and H. S. Wong, "Generalized scale length for two-dimensional effects in MOSFETs," *IEEE Electron Device Letters,* vol. 19, pp. 385-7, 1998.

[15] S. Balasubramanian, L. Chang, B. Nikolic, and T.-J. King, "Circuit-Performance Implications for Double-Gate MOSFET Scaling below 25nm," *Proceedings of the 2003 Silicon Nanoelectronics Workshop,* pp. 16-17, June 2003.

[16] *Taurus-Device Users Manual, version 2005, Synopsys Inc.*

[17] A. Asenov, S. Kaya, and A. R. Brown, "Intrinsic Parameter Fluctuations in Decananometer MOSFETs Introduced by Gate Line Edge Roughness," *IEEE Transactions on Electron Devices,* vol. 50, No. 5, pp. 1254-60, May 2003.

[18] S. Xiong, J. Bokor, Q. Xiang, P. Fisher, I. Dudley, P. Rao, H. Wang, and B. En, "Is Gate Line Edge Roughness A First Order Issue In Affecting The Performance Of Deep Sub-micron Bulk MOSFET Devices?," *IEEE Transactions on Semiconductor Manufacturing,* vol. 17, pp. 357-61, August 2004.

[19] Z. Guo, S. Balasubramanian, R. Zlatanovici, T.-J. King, and B. Nikolic, "FinFET based SRAM Design," in *Proceedings of the 2005 International Symposium on Low Power Electronics and Design*, 2005, pp. 2-7.

[20] H. Shang, L. Chang, X. Wang, M. Rooks, Y. Zhang, B. To, K. Babich, G. Totir, Y. Sun, E. Kiewra, M. Ieong, and W. Haensch, "Investigation of FinFET devices for 32nm technologies and beyond," in *2006 Symposium on VLSI Technology*, 2006, pp. 54-55.

[21] K. Tanaka, K. Takeuchi, and M. Hane, "Practical FinFET Design Considering GIDL for LSTP (Low Standy Power) Devices," in *IEEE International Electron Devices Meeting*, 2005.

[22] P. Verheyen, N. Collaert, R. Rooyackers, R. Loo, D. Shamiryan, A. De Keersgieter, G. Eneman, F. Leys, A Dixit, M.Goodwin, Y.S.Yim, M. Caymax, K. De Meyer, P. Absil, M. Jurczak, and S. Biesemans, "25% Drive Current Improvement for p-type Multiple Gate FET (MuGFET) Devices by Introduction of Recessed $Si_{0.8}Ge_{0.2}$ in the Source and Drain Regions," in *2005 Symposium on VLSI Technology*, 2005, pp. 194-5.

[23] *Taurus-Process Users Manual, version 2005, Synopsys Inc.*

[24] D. Lenoble, K. G. Anil, A. De Keersgieter, P. Eybens, N. Collaert, R. Rooyackers, S. Brus, P. Zimmerman, M. Goodwin, D. Vanhaeren, W. Vandervorst, S. Radovanov, L. Godet, C. Cardinaud, S. Biesemans, T. Skotnicki, and M. Jurczak, "Enhanced performance of PMOS MUGFET via integration of conformal plasma-doped source/drain extensions," in *2006 Symposium on VLSI Technology*, 2006, pp. 168-9.

# Chapter 4 : **Statistical Dopant Fluctuation Effects in FinFETs**

## 4.1 Introduction

Statistical variation in transistor performance and leakage is becoming more and more important with transistor scaling [1-4]. This variation is a net outcome of variations in the (i) device parameters such as the threshold voltage, [5] (ii) parasitic resistances and capacitances [6] and (iii) interconnect geometries (leading to variations in interconnect loading) [7]. The sources of these variations are multifold and will be discussed in detail in this section. In this chapter, $V_T$ variations, particularly caused by statistical dopant fluctuation (SDF) effects in the S/D regions of a FinFET, will be discussed. Along with the simulation algorithm, design guidelines and trade-offs towards variation tolerant device design for highly scaled high performance FinFET devices will be provided.

### 4.1.1 Sources of Variations

The sources of variations on an integrated-circuit chip can be categorized into *spatial* and *temporal* components [8]. The spatial component relates to the variations in the polygon geometries (or positions) on the chip. These include variations in film thickness, dopant placements, polygons size and shape, and the overall geometry of the

devices. Some of the examples include (i) Line Edge Roughness (LER), [9-11] (ii) Layout dependent variations, [12, 13] (iii) Systematic and random variations in film thicknesses [14] (iv) Statistical dopant fluctuation effects [11, 14-16] etc. These are typically process dependent variations. We can further categorize spatial variation into *systematic* and *random* components [17, 18]. Systematic components can be modeled and can be reduced significantly by careful process and layout optimization. Random components on the other hand (as the name suggests) are completely random and do not have any trends associated with them. This component is unavoidable and is likely to become a significant component of the variations with $L_G$ scaling. The method to alleviate the trouble caused by this component is to come up with robust device designs, circuit strategies and architectural strategies that can be "tolerant" to these variations [4].

The temporal component of variations corresponds to the run-time variations on the chip. Some examples of such variations [8] include (i) Device wearout due to electrical stress (NBTI and PBTI) effects, (ii) body effect in SOI devices, (iii) electromigration related changes in interconnect loading [19] and, (iv) temperature dependent parameter variations [20, 21] (different points on the chip have different temperatures during operation). The temporal component can also be alleviated to some extent by well optimized device, circuit and architectural design.

With $L_G$ scaling, non-classical transistor structures such as thin-body MOSFETs may become necessary due to their superior immunity to short-channel effects (SCE) and higher carrier mobilities as compared to the classic bulk-Si MOSFET [22]. Although these structures yield improved off-state leakage ($I_{OFF}$) and on-state drive current ($I_{ON}$), they are also prone to random variations caused by SDF, gate LER and Fin LER, which

will be significant at very small gate lengths. As we saw in the previous chapter [23], SDF induced effects for LSTP devices for $L_G$>15nm are not serious. The advanced FET structures are most likely to first find their way into low power memory (SRAM) circuits, which are very sensitive to variations, thereby making it important to study variation tolerant FET design. In this chapter, we shall concentrate particularly on the modeling of SDF effects and looking at variation tolerant FinFET design to minimize $V_T$ variation due to dopant fluctuation effects.

## 4.2 3D Atomistic Simulation Setup

The SDF effects have been studied through 3D atomistic simulations using Taurus-Device, a conventional drift diffusion simulator [24]. The main challenge of using a commercial drift-diffusion simulator for atomistic simulations is that it does not understand "atoms". The presence of an atom has to be conveyed accurately in the form of a "doping profile" or "doping concentration" to the simulator. The problem therefore, can be broken into two parts: (i) discretize any given set of continuum doping profiles in a FET by random generation and placement of atoms (ii) for any given random placement of atoms, convey the message accurately to the drift diffusion simulator.

### 4.2.1 Random Dopant Generation and Placement

Several dopant placement algorithms have been explored by several researchers over the past decade [15, 25-27]. Wong et al [27] used a 3D atomistic simulation approach. The silicon volume of interest is expanded by 8000 times and dopant atoms (equal to the 8000 times the expected mean number of atoms) are randomly placed according to a uniform expected channel doping. Atom positions within a selected

"discrete doping region" ($1/8000^{th}$ the silicon region) are retained and the atom count in this region follows a Poisson distribution. This idea is simple and accurate, but its simplicity works only for uniform doping profiles and is not easy to implement on a modern MOSFET. Stolk et al [25, 28] on the other hand introduce randomness by replacing the continuum doping concentration at each simulation grid point by a stochastically varying doping concentration following a Poisson distribution. In recent years, more advanced Monte Carlo methods have been adopted for random generation and placement of dopant atoms, which can be used to represent any doping profile (or set of doping profiles as in a modern day MOSFET) with reasonable accuracy. Ezaki et al. [15] use Monte Carlo ion implantation and diffusion to accurately generate randomly placed atoms for any given process and then transform it to the device simulation domain to obtain the spread in the I-V characteristics. This method is the most accurate representation of a real scenario since it includes the implantation and diffusion physics, but it is computationally intensive (since Monte Carlo process simulations typically require a lot of computational resources) on a commercial process simulator. Another interesting approach is to use a Monte Carlo elimination scheme [5, 26, 29] to decide the placement of dopant atoms. In [26], the random dopant placement is done on a virtual silicon lattice to imitate the real scenario as well as possible. It gives a reasonable representation of dopant atom count as well as placement without intensive computation time and therefore, has been adopted as the placement algorithm for this work.

The dopant generation and placement algorithm is shown in **Figure 4.1**. The algorithm is implemented in MATLAB and the code is provided in Appendix B. The volume of silicon which needs to be atomized is converted to a cubic grid with the silicon

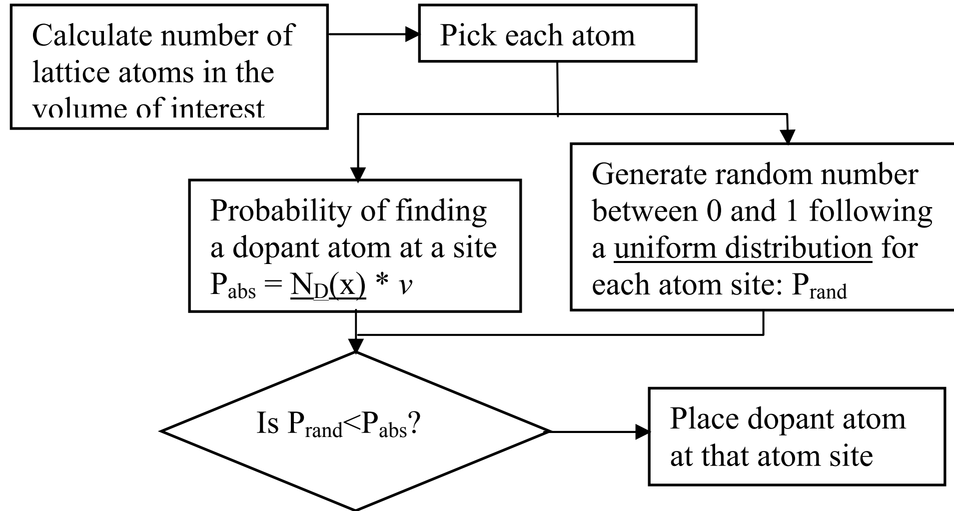
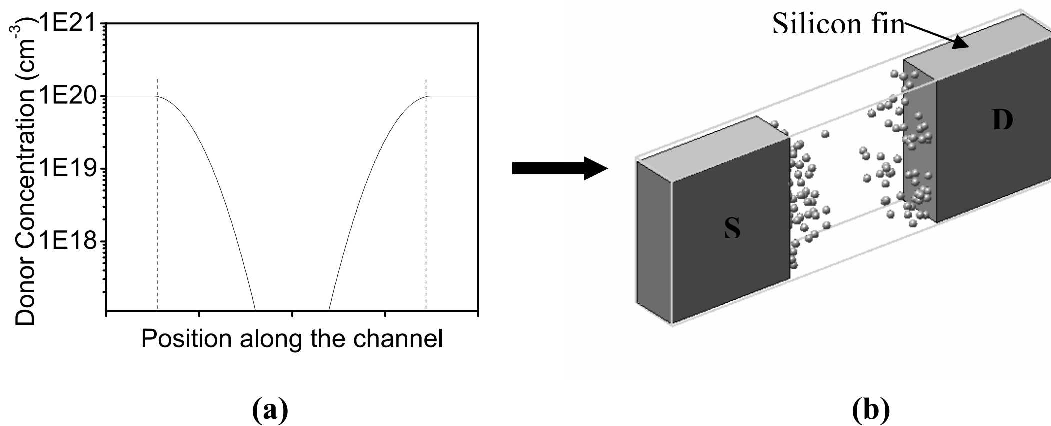



**Figure 4.1**: Dopant generation and placement algorithm adopted from [26]. The use of a pseudo silicon lattice grid to place atoms makes it an accurate Monte Carlo placement technique applicable to any doping profile.

lattice spacing in each direction, with each grid point representing a silicon atom in this pseudo silicon lattice. Depending on the expected doping concentration at any grid point (x,y,z) given by $N_D(x,y,z)$, the absolute probability of finding a dopant atom at the site is given by $P_{abs} = N_D(x,y,z)*v$, where $v$ is the atomic volume. In this study the doping varies in the x direction only and so $N_D(x,y,z)=N_D(x)$. Each atomic site is also assigned a random number ($P_{rand}$) between 0 and 1 following a uniform distribution. An elimination scheme compares the absolute probability $P_{abs}$ with $P_{rand}$. If $P_{abs} > P_{rand}$, a dopant atom is placed at the lattice site. By this scheme, more sites get picked in the heavily doped region and fewer sites get picked in the lightly doped region. An example of running this algorithm on a FinFET structure with a certain S/D doping gradient is shown in **Figure 4.2**. In this example, only the S/D gradient region has been discretized.




**(a)** **(b)**

**Figure 4.2**: An example of dopant generation and placement showing: (a) Continuum domain expected S/D profile and, (b) discretized profile for 3D atomistic simulations. Only the S/D gradient region has been discretized.

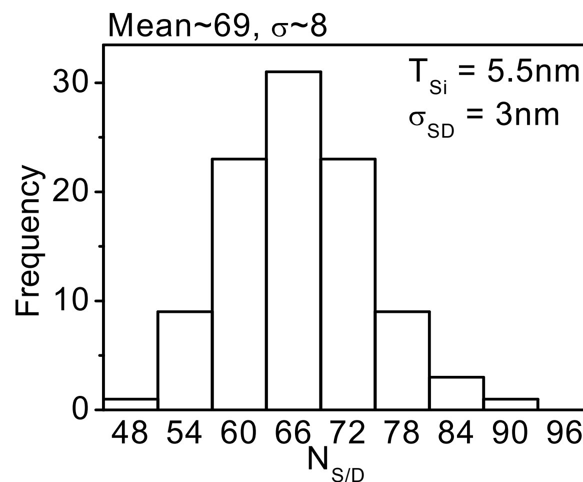



**Figure 4.3**: Typical distribution of number of dopant atoms as a result of running the dopant generation and placement algorithm on an undoped channel FinFET.

**Figure 4.3** shows a typical statistical distribution of the number of atoms obtained by running the above placement algorithm 100 times for a given continuum domain profile. This has been done for a FinFET device with a finite S/D doping gradient and undoped channel. In this example, the expected value of the total number of atoms in the

gradient region, computed from the continuum domain profile is about 70 atoms, which matches well with the mean value obtained from the atomized profile. The standard deviation is close to the square-root of the mean, suggesting that the data follows a Poisson distribution. Any statistic that looks at a count of number of events (number of dopant atoms) occurring in a certain duration of time (in our case there is no temporal process, but each of the 100 runs may be considered as a unit of time) is expected to be a Poisson process and therefore, the result obtained from this algorithm is as expected. The next part of the algorithm will deal with conveying the atomistic information accurately to a drift-diffusion simulator.

### 4.2.2 Dopant Definition in a Drift-Diffusion Simulator

This part of the algorithm for atomistic simulations communicates directly with the drift-diffusion simulator and therefore, the output format would depend on the simulator being used. For this study, Taurus-Device [24] simulator has been used and some modifications will be needed in the output format for any other simulator. Another methodology will also be briefly mentioned which does not require this part of the algorithm and can be used with any simulator.

Several methodologies have been adopted to represent a dopant atom in the form of a doping profile. The most intuitive strategy is to take the simulation grid (note, this is not the same as the pseudo-lattice grid) and for each dopant atom in the pseudo lattice, find the closest simulation grid point and snap the atom to this location on the simulation grid. The volume occupied by the grid point ($V_{grid}$) is computed based on the local grid spacing, and a doping concentration of $1/V_{grid}$ is assigned to that grid location. This method is pictorially demonstrated in **Figure 4.4**. **Figure 4.4**(a) shows the 2D projection

along a silicon fin (X-axis from source to drain and Y-axis along the fin height) and the discrete diamonds denote all the dopant atoms snapped to the simulation grid.

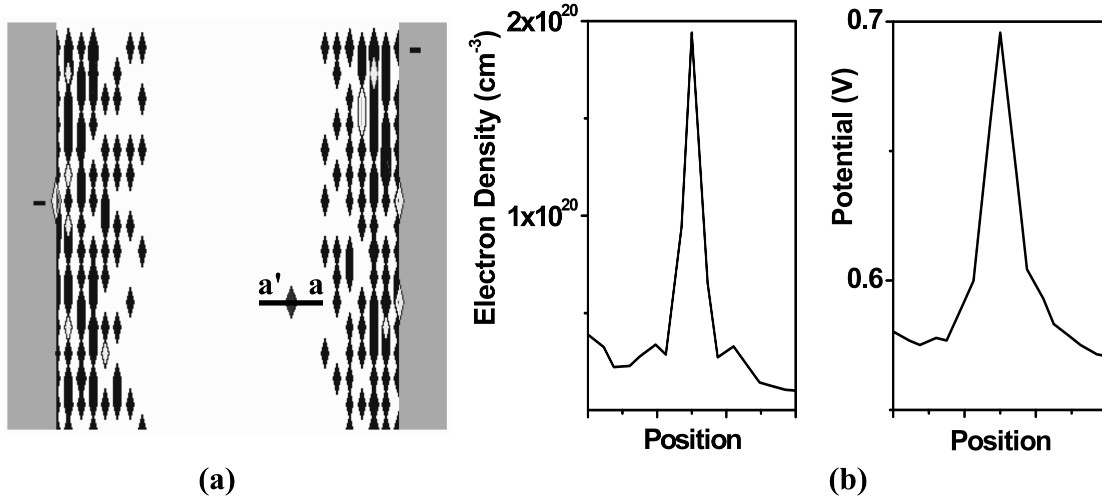



**Figure 4.4**: Demonstration of point charge definition of dopant atoms using a 3D FinFET device (a) Random dopant atoms are snapped to simulation grid and are defined as point doping concentrations assigned to the grid point. (b) This leads to singularities in potential and mobile electron density (cross section a'a through a dopant atom) which in turn introduces errors in the simulated $V_T$, $I_{ON}$ and $I_{OFF}$.

There are several drawbacks of this approach. One clear shortcoming of this method is the fact that $V_{grid}$ is a function of the grid spacing and therefore, the doping concentration (and hence $V_T$ and $I_{ON}$) assigned to a dopant atom will be very grid sensitive. In addition, snapping a dopant atom to the nearest grid point introduces an error in the simulated position of the dopant atom. **Figure 4.4**(b) shows the potential and mobile electron density along one of the dopants in the ON state. A point charge definition of a dopant atom gives rise to unrealistic singularities in potential and charge density, which in turn leads to further errors in $V_T$, $I_{OFF}$ and $I_{ON}$.

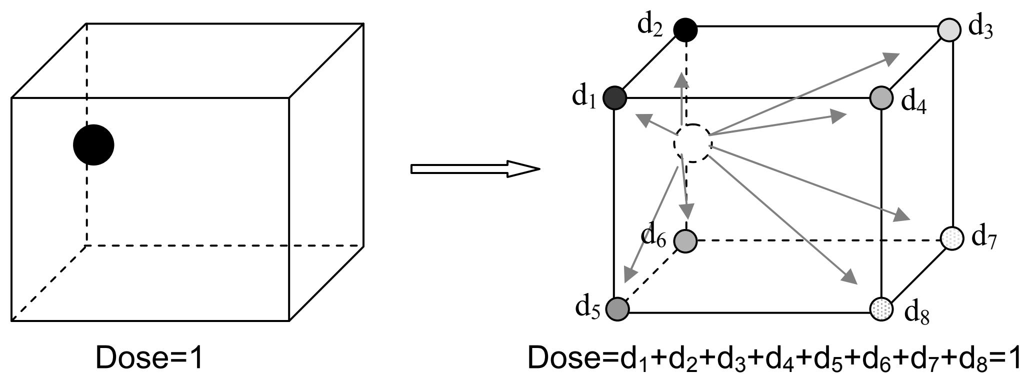



**Figure 4.5**: Pictorial representation of the *nearest neighbors* approach. A single dopant atom is defined by a weighted dose distribution over its eight neighbors in the simulation grid. This ensures maintaining the accuracy of the position of the generated dopant atom independent of the simulation grid.

A modification of the above approach alleviates some of its limitations. Instead of snapping the dopant atom to one grid point, it would be more accurate to share its effect over its *nearest neighbors* (formed by the eight vertices of the cuboid that encloses the atom) in a weighted manner. This is better represented in **Figure 4.5**. This way the "centroid" of the dopant atom still remains at the point where it was supposed to be defined. But the issues of grid sensitivity and potential singularities still remain. A much better representation of an atom was provided by Sano et al. [30, 31], which focuses on representing the "influence" of a dopant atom rather than a point charge. Each dopant atom has a Coulomb potential associated with it which follows a 1/r decay (**Figure 4.6**). In a silicon lattice, any conduction electron will almost always see the tail part of this potential, also known as the *long range Coulomb potential*. The long range and short range Coulomb potentials can be modeled separately and the long range potential does

not have a potential singularity at the location of the dopant. The long range charge density corresponding to this potential is given by:

$$\rho(r) = \frac{ek_c^3}{2\pi^2} \frac{\sin(k_c r) - (k_c r)\cos(k_c r)}{(k_c r)^3}$$

where, $k_c$ is the inverse screen length or inverse of *Debye length* and basically relates to the radius of influence of the dopant atom. In a heavily doped region, the radius of influence will be shorter than in the lightly doped region.

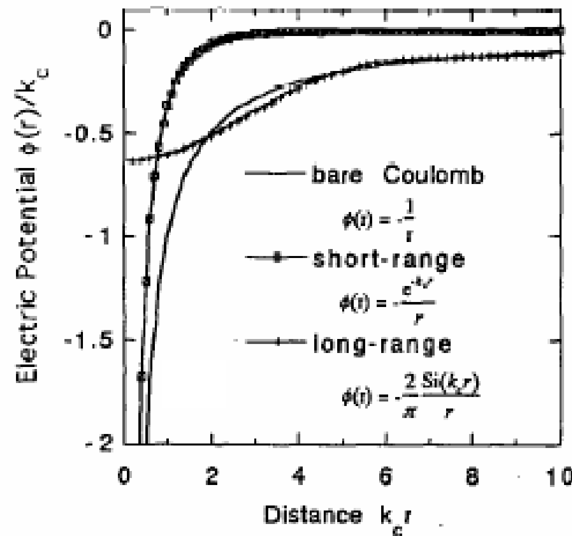



**Figure 4.6**: Coulomb potential of a single dopant atom in a silicon lattice showing the long range and short range components as well as the corresponding model [31]. An electron in the lattice is more likely to experience the long range Coulomb potential of a dopant atom.

This method can be easily incorporated in a commercial drift-diffusion simulator and has been adopted in this work. The distance (r) of each grid point from a dopant atom is computed. Using this, the doping concentration at each grid point due to the dopant

atom is computed using the long range charge density. This is repeated for all dopant atoms and the net doping concentration at any grid point is computed as the sum of the doping concentrations due to each dopant atom. In a small silicon volume (as in a silicon fin in a FinFET), the total dose of each atom represented by the long range charge density may not integrate to one and dose normalization should be done to conserve the dose. An example of a dopant atom represented by this method is shown in **Figure 4.7**.

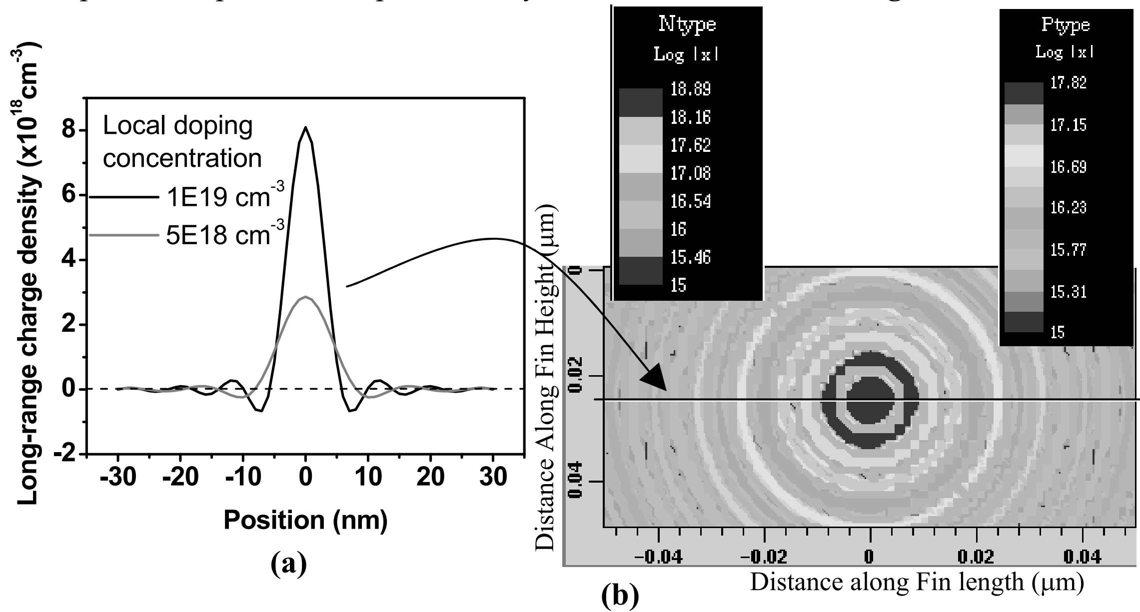



**Figure 4.7**: (a) Long range charge density for a single dopant atom in a simulation, shown for two different continuum local doping concentrations. (b) Representation of a dopant atom using long-range charge density approach in a drift-diffusion simulator.

The above method is quite accurate and alleviates the problems of grid sensitivity and charge singularities. Another approach that has been recently explored [32, 33], uses point charge representation of a dopant atom, but uses quantum mechanical models to smoothen the potential singularities. This is also a very accurate method (may be more accurate than our approach, particularly for thin body devices) which minimizes all the drawbacks of the point charge representation and is simple to implement.

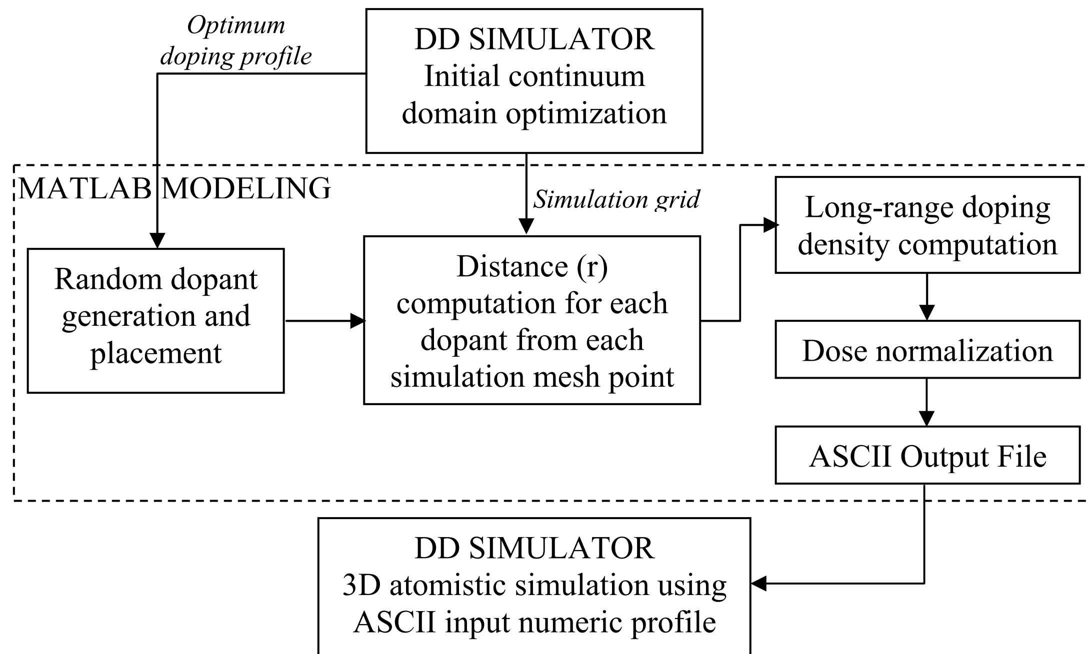



**Figure 4.8**: Complete algorithm for running 3D atomistic simulations in a Drift-Diffusion (DD) simulator.

### 4.2.3 Complete 3D Atomistic Simulation Algorithm

The overall 3D atomistic simulation methodology is shown in **Figure 4.8**. The continuum device is first optimized in the drift diffusion simulator to get the doping profiles to be discretized. The grid information and doping profiles of the optimized device are provided to MATLAB. Based on the profiles, the code generates random dopants, represents them as long range charge density profiles and generates an output file containing the grid information appended with the charge density at each grid point. This output file can now be used as an input Numeric Profile in the simulator. Different simulators may need different output formats. If a simulator cannot handle numeric profiles, it may be simpler to use the density gradient quantum model approach to simulate the dopant atoms.

## 4.3 Variation Tolerant FinFET Design

In this section statistical dopant fluctuation effects in the S/D gradient region of a FinFET will be studied through 3D device simulations [34]. The device simulators use drift-diffusion transport and 1D-Schrodinger quantum models. The target device is a highly scaled HP FinFET device (**Figure 4.9**), the continuum domain device parameters of which are summarized in **Table 4.1**. Through 3D atomistic simulations, simple device design strategies, along with their tradeoffs, are presented.

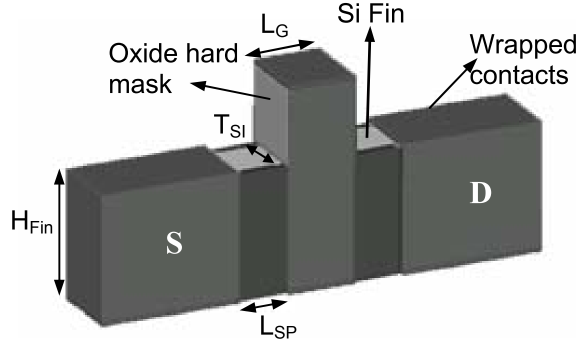



**Figure 4.9**: 3D simulation structure of a highly scaled HP FinFET used in the atomistic simulations.

| Parameter | Value |
|---|---|
| $L_G$ | 9nm |
| EOT | 7Å |
| $\rho_C$ | 8.69E-9Ω-cm$^2$ |
| $V_{DD}$ | 0.8V |
| $I_{OFF}$ | 210nA/µm |
| DIBL | 100mV/V |
| $N_{SD}$ | 1E20cm$^{-3}$ |
| **$\Phi_M$ $T_{Si}$, $L_{SP}$, $\sigma_{SD}$** | **Variables** |

**Table 4.1**: Continuum domain device parameters for optimization

### 4.3.1 Continuum domain device optimization

The first step is to identify optimum devices on which the statistical simulations can be implemented. The goal is to study the effect of different FinFET device parameters ($\sigma_{SD}$, and $T_{Si}$) on variation tolerance from SDF effects in the S/D gradient

region. For the device optimization, $L_{SP}$ (the lateral offset between the gate edge and the region of uniformly high S/D doping, $1\times10^{20}/cm^3$) and gate work function ($\Phi_M$) are used as the optimization parameters.

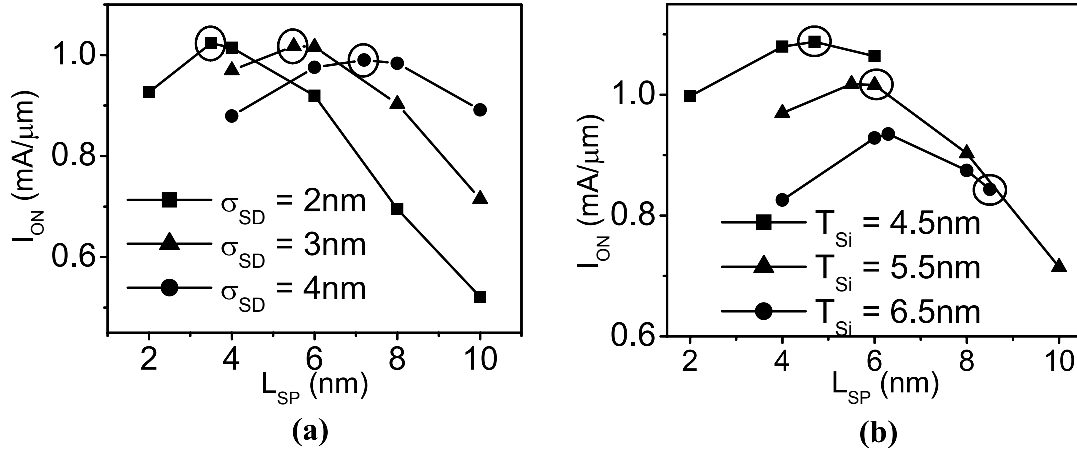



**Figure 4.10**: Continuum domain selection of optimal FinFET designs for atomistic simulations. FinFET designs for (a) $T_{Si}$=5.5nm, varying $\sigma_{SD}$ and, (b) $\sigma_{SD}$=3nm, varying $T_{Si}$. Larger $L_{SP}$ is needed to achieve reasonable SCE control for thicker $T_{Si}$.

In order to look at the effect of $\sigma_{SD}$, $T_{Si}$ = 5.5nm (~ $2L_G/3$) is chosen. The aspect ratio $H_{fin}$:$T_{Si}$ is fixed at 3:1 in consideration of etch process limits versus layout efficiency. The value of $\sigma_{SD}$ is varied from 2nm to 4nm. For each value of $\sigma_{SD}$, the $L_{SP}$ is varied (the gate work function is used to adjust the $I_{OFF}$) until $I_{ON}$ is maximized for a DIBL≤100mV/V at $I_{OFF}$=210nA/µm. When $T_{Si}$, $T_{OX}$ and $L_G$ are the same, the optimal devices for the different S/D doping gradients, have the same $L_{eff}$. The optimization for different $\sigma_{SD}$ values is shown in **Figure 4.10**(a), and the encircled datapoints correspond to the optimal designs.

In order to look at the effect of $T_{Si}$, the $\sigma_{SD}$ value is fixed at 3nm and $T_{Si}$ is varied from 4.5nm to 6.5nm. At each value of $T_{Si}$, $L_{SP}$ and $\Phi_M$ are optimized to identify design points with maximum $I_{ON}$ for DIBL≤100mV/V and $I_{OFF}$=210nA/μm. When $T_{Si}$ is large, the short channel control is worse and therefore the optimum design point does not coincide with the maximum achievable performance point. The encircled points denote the optimum design points in **Figure 4.10**(b).

### 4.3.2 Design Strategies for Variation Tolerance to SDF

With the optimized design points, we proceed with the 3D atomistic simulations to evaluate the variability of each of the designs. An ensemble of 100 atomistic simulations is done for each optimized device to analyze the statistical behavior.

#### Requirement of S/D Abruptness $\sigma_{SD}$

As $\sigma_{SD}$ increases, the mean number of dopant atoms in the S/D gradient regions ($N_{S/D}$) increases, and this helps in reducing $\sigma_{VT}$. At the same time, the variation in the position of the dopant atoms is higher, leading to increased variation of the "average" $L_{eff}$ (and hence $\sigma_{VT}$). These are two counteracting effects. **Figure 4.11** shows the resultant SDF-induced variation in threshold voltage $V_T$ as a function of $\sigma_{SD}$. **Figure 4.12** shows the statistical results for $I_{OFF}$ and $I_{ON}$ as a function of $\sigma_{SD}$. Steeper S/D gradient provides higher tolerance to SDF despite a decrease in the average number of dopant atoms. The increase in mean $I_{OFF}$ and decrease in mean $I_{ON}$ with increasing $\sigma_{SD}$ are due to increased SCE variation and increased parasitic S/D series resistance ($R_{series}$), respectively. As a result, the smallest value of $\sigma_{SD}$ yields the highest mean $I_{ON}/I_{OFF}$, in contrast to the

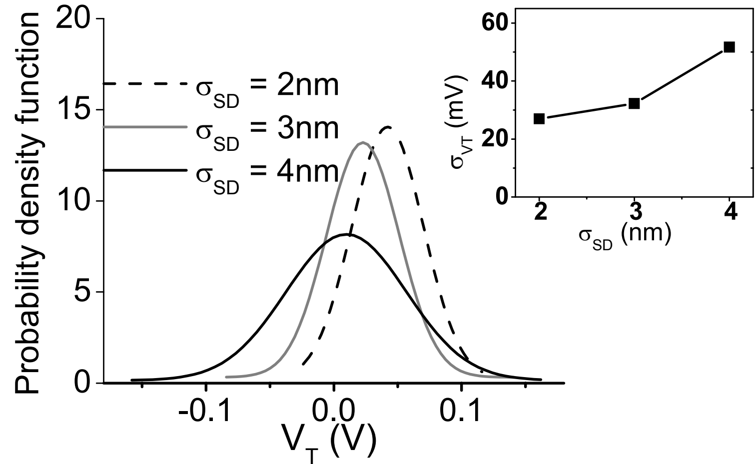



**Figure 4.11**: Statistical variation in saturation $V_T$ of the HP FinFET for varying $\sigma_{SD}$. $V_T$ is defined here, as $V_{GS}$ @ $I_{DS}$ = 800nA/μm and $V_{DS}$ = 0.8V.

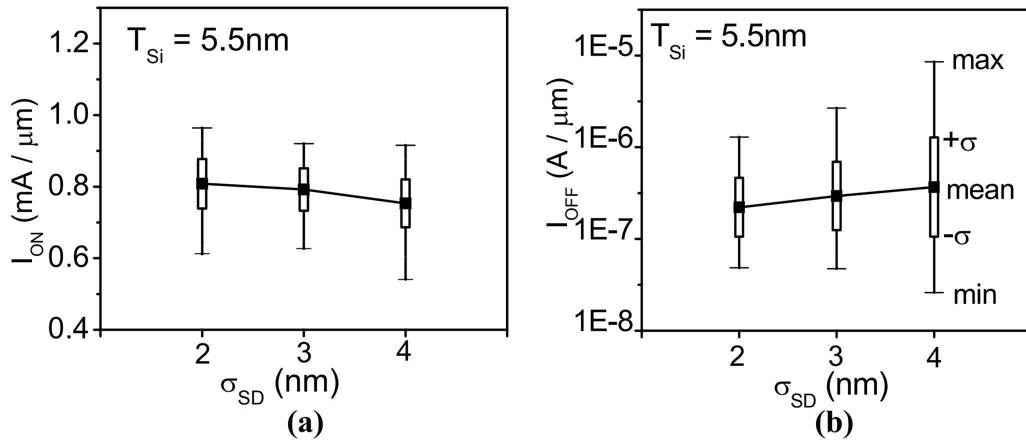



**Figure 4.12**: Statistical results of FinFET atomistic simulations, for varying $\sigma_{SD}$: **(a)** $I_{OFF}$ ($V_{GS}$=0V, $V_{DS}$=$V_{DD}$) and, **(b)** $I_{ON}$ ($V_{GS}$=$V_{DS}$=$V_{DD}$). $T_{Si}$ = 5.5nm.

continuum simulation results which show comparable $I_{ON}/I_{OFF}$ for different values of $\sigma_{SD}$. This indicates that the effect of $L_{eff}$ variation is stronger than the effect of $N_{S/D}$ variation.

## Optimization of $T_{Si}$

Device designs with the same $\sigma_{SD}$ and SCE but different $T_{Si}$ differ only in $N_{S/D}$. Therefore, thicker $T_{Si}$ is expected to yield better tolerance to SDF as observed in the

atomistic simulations (**Figure 4.13**). This of course comes at the cost of degraded mean $I_{ON}$ due to poorer short channel control for a thicker $T_{Si}$. Therefore, once again, there exists a tradeoff between performance and variation tolerance to SDF (as well as LER, from previous chapter) and an intelligent choice of $T_{Si}$ is important.

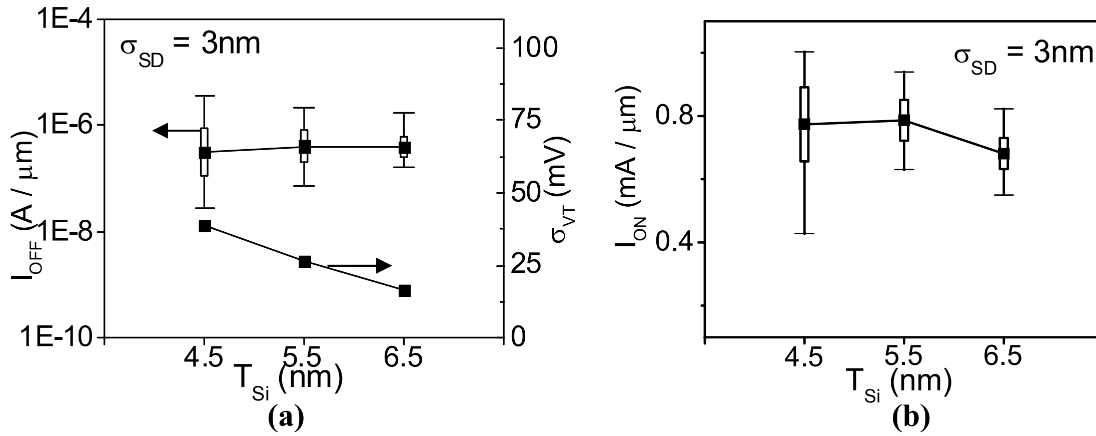



**Figure 4.13**: Statistical results of FinFET atomistic simulations, for varying $T_{Si}$: **(a)** $I_{OFF}$ ($V_{GS}$=0V, $V_{DS}$=$V_{DD}$) and $V_T$ ($V_T$=$V_{GS}$ @ $I_{DS}$=800nA/μm and $V_{DS}$=0.9V) **(b)** $I_{ON}$ ($V_{GS}$=$V_{DS}$=$V_{DD}$).

Even though the trends and conclusions are accurate in both the above analyses, there is one point to be carefully noted. It is observed that the mean $I_{ON}$ in the statistical simulations is lower than the continuum domain optimized $I_{ON}$. There are two possibilities for this: (i) the S/D resistance is higher when a doping profile is discretized into dopant atoms and/or, (ii) the long range charge density profile is still not efficient enough in removing potential singularities completely. It is difficult to avoid this issue in the current approach, but an alternative approach may be considered for simulations, as discussed in the next subsection.

### Shortcomings of the Current Approach

Even though the present approach for modeling variations due to SDF is one of most accurate strategies available, it is not perfect. For thin silicon films (as in FinFETs and UTB-FETs), the long-range charge density profile is not an accurate representation of a dopant atom. This is due to two reasons: (i) The long range Coulomb potential derivation assumes an infinite extent of silicon lattice, which breaks down when you have a thin silicon film and, (ii) In the high doping regions, the shape of the long range charge density profile is taller and thinner than in the moderately doped regions. In the worst case (say, in the 1E20cm$^{-3}$ doped region), this once again tends towards a point charge definition and can cause singularities in potential. But using the quantum mechanical model should alleviate this problem.

A simpler strategy for SDF modeling may be to go back to the *nearest neighbors* approach (**Figure 4.5**) and, use a 3D quantum mechanical model (like the Density Gradient model [32]) instead of the 1D Schrödinger solution.

## 4.4 Summary

In this chapter, a methodology for modeling statistical dopant fluctuation effects in a commercial drift-diffusion simulator has been demonstrated. The algorithm for random dopant generation and placement is realized in MATLAB by a Monte Carlo elimination scheme implemented on a pseudo silicon lattice. The presence of a dopant atom can be accurately communicated to the simulator by modeling the long-range charge density which is an outcome of the long-range Coulomb potential of a dopant atom. Modeling the long-range Coulomb potential (i) minimizes the singularities in

potential and electron density and (ii) removes grid sensitivity, otherwise caused by point charge definition of a dopant.

The above methodology has been used to study variation tolerant thin body device design using a highly scaled high-performance (HP) FinFET structure. Lean spacers and super steep S/D profiles (2nm/decade) will be needed to minimize the $V_T$ variation due to SDF effects. Additionally, thick silicon body, at the cost of some performance can further reduce SDF induced $\sigma_{VT}$.

Another approach for dopant definition, not discussed in this work, is to use a point charge definition but with a 3D quantum mechanical model which can completely eliminate singularities in potential and electron density.

## 4.5 References


[1] S. Borkar, T. Karnik, S. Narendra, J. Tschanz, A. Keshavarzi, and V. De, "Parameter variations and impact on circuits and microarchitecture," in *Proceedings 2003. Design Automation Conference. Anaheim, CA*, 2003, pp. 338-42.

[2] S. Nassif, "Delay variability: sources, impacts and trends," in *2000 IEEE International Solid-State Circuits Conference. Digest of Technical Papers.* , 2000, pp. 368-9.

[3] Y. Cao, H. Qin, R. Wang, P. Friedberg, A. Vladimirescu, and J. Rabaey, "Yield optimization with energy-delay constraints in low-power digital circuits," in *IEEE Conference on Electron Devices and Solid-State Circuits* 2003, pp. 285-8.

[4] S. Borkar, "Designing reliable systems from unreliable components: the challenges of transistor variability and degradation," *IEEE Micro,* vol. 25, pp. 10-16, 2005.

[5] A. R. Brown, A. Asenov, and J. R. Watling, "Intrinsic fluctuations in sub 10-nm double-gate MOSFETs introduced by discreteness of charge and matter," *IEEE Transactions on Nanotechnology,* vol. 1, pp. 195-200, 2002.

[6] J. O. Plouchart, J. Kim, B. J. Gross, K. Wu, R.Trzcinski, V. Karam, P. Hyde, R. Williams, M.-H. Na, J. M. Cullen, and W. Clark, "SOI 90-nm ring oscillator sub-ps model-hardware correlation and parasitic-aware optimization leading to 1.94-ps switching delay," in *International Electron Devices Meeting 2005*, 2005.

[7] Y. Jie, E. Cohen, C. Tabery, N. Rodriguez, and M. Craig, "An up-stream design auto-fix flow for manufacturability enhancement," in *2006 Design Automation Conference*, 2006, pp. 73-6.

[8] K. Bernstein, D. J. Frank, A. E. Gattiker, W. Haensch, B. L. Ji, S. R. Nassif, E. J. Nowak, D. J. Pearson, and N. J. Rohrer, "High-performance CMOS variability in the 65-nm regime and beyond," *IBM Journal of Research and Development,* vol. 50, pp. 433-49, 2006.

[9] A. Asenov, S. Kaya, and A. R. Brown, "Intrinsic Parameter Fluctuations in Decananometer MOSFETs Introduced by Gate Line Edge Roughness," *IEEE Transactions on Electron Devices,* vol. 50, No. 5, pp. 1254-60, May 2003.

[10] S. Xiong, J. Bokor, Q. Xiang, P. Fisher, I. Dudley, and P. Rao, "Study of Gate Line Edge Roughness Effects in 50nm Bulk MOSFET Devices," *Proceedings of SPIE,* vol. 4689, pp. 733-41, March 2002.

[11] M. Hane, T. Ikezawa, and T. Ezaki, "Coupled atomistic 3D process/device simulation considering both line-edge roughness and random-discrete-dopant effects," in *IEEE International Conference on Simulation of Semiconductor Processes and Devices*, 2003, pp. 99-102.

[12] L.-T. Pang and B. Nikolic, "Impact of layout on 90nm CMOS process parameter fluctuations," in *2006 Symposium on VLSI Circuits*, 2006, pp. 69-70.

[13] C. Tabery, M. Craig, G. Burbach, B. Wagner, S. McGowan, P. Etter, S. Roling, C. Haidinyak, and E. Ehrichs, "Process window and device variations evaluation using array-based characterization circuits," in *Proceedings of the 2006 7th International Symposium on Quality Electronic Design*, 2006, p. 6.

[14] A. Asenov, A. R. Brown, J. H. Davies, S. Kaya, and G. Slavcheva, "Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale MOSFETs," *IEEE Transactions on Electron Devices,* vol. 50, pp. 1837-52, 2003.

[15] T. Ezaki, T. Ikezawa, and M. Hane, "Investigation of realistic dopant fluctuation induced device characteristics variation for sub-100 nm CMOS by using atomistic 3D process/device simulator," in *IEEE International Electron Devices Meeting*, 2002, pp. 311-14.

[16] S. Xiong and J. Bokor, "Sensitivity of double-gate and FinFET Devices to process variations," *IEEE Transactions on Electron Devices,* vol. 50, pp. 2255-61, 2003.

[17] B. E. Stine, D. S. Boning, and J. E. Chung, "Analysis and decomposition of spatial variation in integrated circuit processes and devices," *IEEE Transactions on Semiconductor Manufacturing,* vol. 10, pp. 24-41, 1997.

[18] P. Friedberg, Y. Cao, J. Cain, R. Wang, J. Rabaey, and C. Spanos, "Modeling within-die Gate Length Spatial Correlation for Process Design Co-optimization," *Design and Process Integration for Microelectronic Manufacturing IV, SPIE,* vol. 5756, pp. 178-88, 2005.

[19] M. H. Lin, K. P. Chang, K. C. Su, and T. Wang, "Effects of width scaling, length scaling, and layout variation on electromigration in dual damascene copper interconnects," in *2006 IEEE International Reliability Physics Symposium proceedings*, 2006, pp. 671-2.

[20] Z. Lu, W. Huang, J. Lach, M. Stan, and K. Skadron, "Interconnect lifetime prediction under dynamic stress for reliability-aware design," in *ICCAD 2004. International Conference on Computer Aided Design*, 2004, pp. 327-34.

[21] H. Su, F. Liu, A. Devgan, E. Acar, and S. Nassif, "Full chip leakage-estimation considering power supply and temperature variations," in *International Symposium on Low-Power Electronics and Design* 2003, pp. 78-83.

[22] D. A. Antoniadis, "MOSFET scalability limits and "new frontier" devices," in *2002 Symposium on VLSI Technology*, 2002, pp. 2-5.

[23] V. Varadarajan and T.-J. K. Liu, "$V_T$ Adjustment by $L_{eff}$ Engineering for LSTP Single Gate Work-function CMOS FinFET Technology," *16th Biennial University/Government/Industry Microelectronics Symposium,* pp. 153-6, 2006.

[24] *Taurus-Device Users Manual, version 2003,* Synopsys Inc.

[25] P. A. Stolk and D. B. M. Klaassen, "The effect of statistical dopant fluctuations on MOS device performance," in *International Electron Devices Meeting. Technical Digest*, 1996, pp. 627-30.

[26] D. J. Frank, Y. Taur, M. Ieong, and H. S. Wong, "Monte Carlo modeling of threshold variation due to dopant fluctuations," in *Symposium on VLSI Circuits*, June 1999, pp. 169-70.

[27] H.-S. Wong and Y. Taur, "Three-dimensional "atomistic" simulation of discrete random dopant distribution effects in sub-0.1 μm MOSFET's," in *Proceedings of IEEE International Electron Devices Meeting*, 1993, pp. 705-8.

[28] P. A. Stolk, F. P. Widdershoven, and D. B. M. Klaassen, "Modeling statistical dopant fluctuations in MOS transistors," *IEEE Transactions on Electron Devices,* vol. 45, pp. 1960-71, 1998.

[29] V.-Y. A. Thean, M. Sadd, and J. Bruce E. White, "Effects of Dopant Granularity on Superhalo-Channel MOSFET's According to Two- and Three-Dimensional Computer Simulations," *IEEE Transactions on Nanotechnology,* vol. 2, No. 2, pp. 97-101, June 2003.

[30] T. Ezaki, T. Ikezawa, A. Notsu, K. Tanaka, and M. Hane, "3D MOSFET simulation considering long-range Coulomb potential effects for analyzing statistical dopant-induced fluctuations associated with atomistic process simulator," in *2002 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD 2002)*, 2002, pp. 91-4.

[31] N. Sano, K. Matsuzawa, M. Mukai, and N. Nakayama, "Role of long-range and short-range Coulomb potentials in threshold characteristics under discrete dopants in sub-0.1 μm Si-MOSFETs," in *International Electron Devices Meeting*, 2000, pp. 275-8.

[32] A. Asenov, G. Roy, C. Alexander, A. R. Brown, J. R. Watling, and S. Roy, "Quantum mechanical and transport aspects of resolving discrete charges in nano-CMOS device simulation," in *2004 4th IEEE Conference on Nanotechnology*, August 2004, pp. 334-6.

[33] G. Roy, A. R. Brown, A. Asenov, and S. Roy, "Bipolar quantum corrections in resolving dopants in 'atomistic' device simulation," *Superlattices and Microstructures,* vol. 34, No. 3-6, pp. 327-34, 1-5 Dec. 2003.

[34] V. Varadarajan, L. Smith, S. Balasubramanian, and T.-J. K. Liu, "Multi-Gate FET Design for Tolerance to Statistical Dopant Fluctuations," *Proceedings of the 2006 Silicon Nanoelectronics Workshop,* pp. 137-8, 2006.

# Chapter 5 : **WireFET Technology**

## 5.1 Introduction

As MOSFETs are scaled down to sub-20nm gate lengths ($L_G$), parasitic source/drain (S/D) series resistance and capacitance [1] and interconnect delays [2, 3] will become major limiters to the performance of integrated circuits (ICs). In order to overcome these issues, researchers have investigated solutions such as metallic-S/D transistors [4-6] and three-dimensional (3-D) integration [2, 7, 8]. 3-D integration focuses on stacking multiple device layers in the vertical dimension (which would otherwise be placed laterally on a conventional chip), thereby reducing chip delay and power by significantly reducing interconnect loading. Approaches explored to date include (**Figure 5.1**) (i) processed wafer bonding, [9] (ii) epitaxial silicon growth, and (iii) crystallization of amorphous silicon [10, 11]. In processed wafer bonding technique, multiple device layers are processed on separate single crystal silicon substrates and eventually bonded to a primary substrate. Chief processing challenges of this technique are stringent misalignment tolerances and very high aspect ratio via etching (to contact device layer with the interconnect layer). Crystallization of silicon approach is implemented by depositing amorphous silicon directly on the primary device wafer and then crystallizing through solid phase crystallization methods like metal induced crystallization. Achieving

a polycrystalline silicon channel and large crystallization thermal budget are the typical issues of this approach. But primarily, all the above techniques focus on building separate device layers and then dropping vias to make connections to them. This adds to process complexity as well as some area penalty (in dropping vias to the device layer), making them difficult to implement with conventional CMOS technology in a cost-effective manner for IC manufacturing. In this chapter, a novel, low thermal budget [12], cost-effective method to fabricate a transistor directly within a wire is described along with preliminary device fabrication data [13].

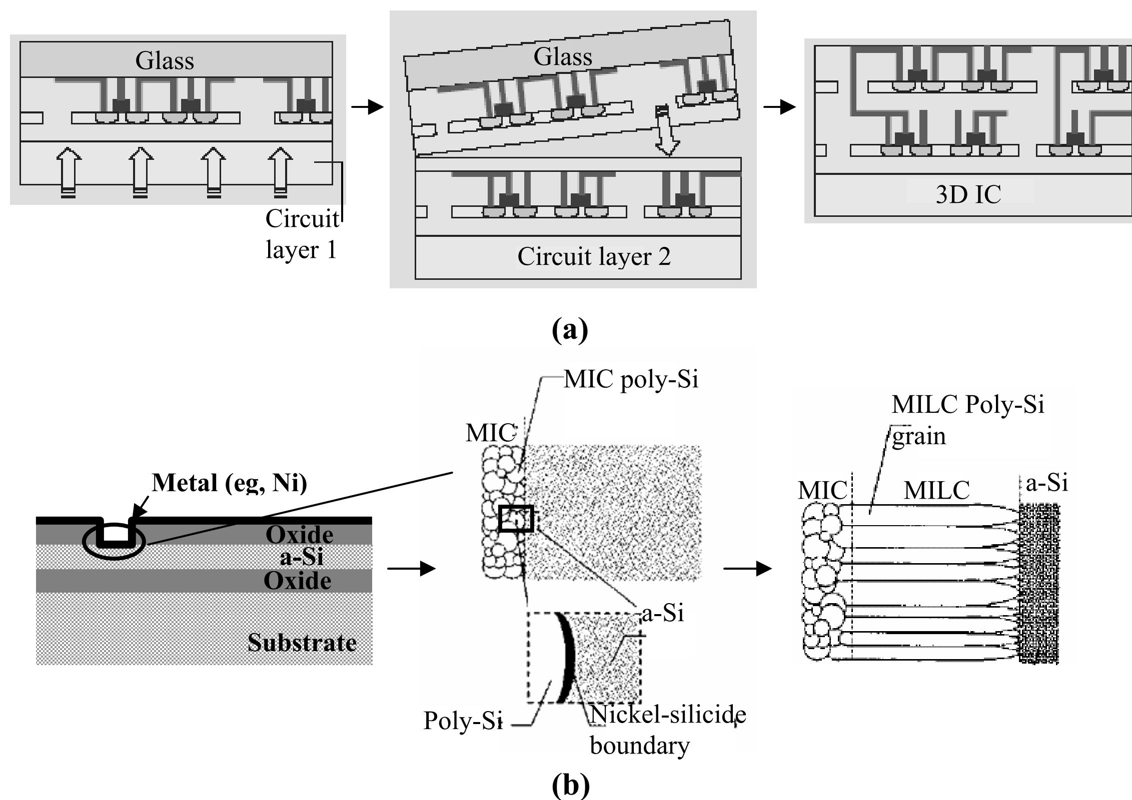



**Figure 5.1**: Current approaches to 3D-Integration include methods such as (a) processed wafer bonding (adopted from [9]) and (b) crystallization of a-Si (adopted from [11]). In these approaches the device layer is formed separately and connected to interconnect layers through vias.

### 5.1.1 Concept of WireFET

A schematic representation of the wireFET concept is shown in **Figure 5.2**. Consider two interconnect layers shown as *Metal1* and *Metal2*. If a silicon channel is formed selectively within a wire in *Metal1*, and *Metal2* is used to provide gate bias to this channel, then in principle, a metal S/D transistor is embedded within the interconnect layer. Three-dimensional integration for specific applications can be potentially targeted through this concept. It will be shown in the following sections that the phenomenon of aluminum-induced crystallization (AIC) of silicon can be employed to embed crystalline Si regions within an aluminum wire (potentially an aluminum interconnect wire). Other CMOS compatible metal-semiconductor systems are yet to be explored for fabricating wireFETs.

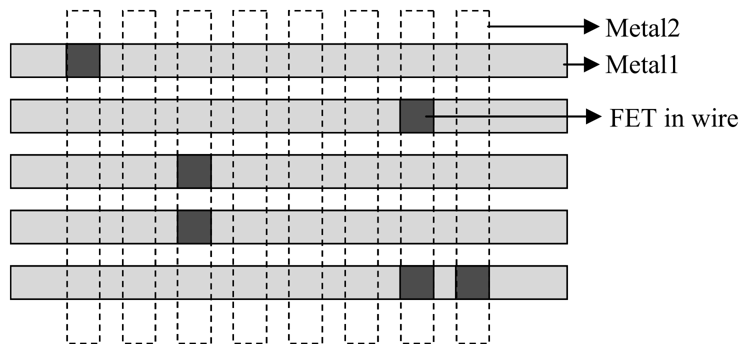



**Figure 5.2**: Schematic (plan view) representation of the proposed 3D-Integration scheme. The wireFET concept focuses on forming FETs directly within the interconnect wires rather than separate device layers.

### 5.1.2 Initial Simulations

2-D device simulations of the wireFET were performed using DESSIS-ISE [14], to assess the potential wireFET performance and guide device design. The simulated FET

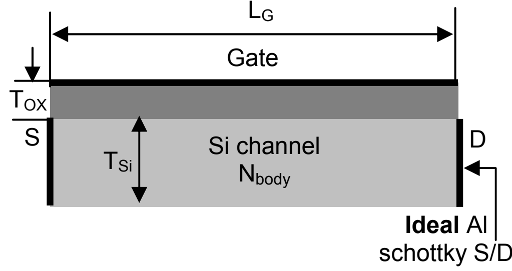



Figure 5.3: Simulation structure for wireFET simulations.

| Parameter | Value |
| --- | --- |
| $L_G$ | 0.25μm, 0.5μm, 1μm |
| $T_{ox}$ | 5nm |
| $T_{Si}$ | 20nm, 50nm, 100nm |
| $V_{DD}$ | 2.5V |
| $N_{body}$ | 5E18 $cm^{-3}$<br>(~Solid Solubility of Al in Si) |

**Table 5.1**: Device parameters used for wireFET simulations in DESSIS-ISE.

structure and device parameters are given in **Figure 5.3** and **Table 5.1**, respectively. Assuming an ideal Schottky aluminum-Si contact and a silicon body doping of 5E18$cm^{-3}$ (corresponding to bulk solid solubility of Al in Si [15]), an enhancement-mode NMOSFET operation is observed. **Figure 5.4**a shows that with the above assumptions, transistor action is feasible even with $T_{Si}$ as thick as 100nm. $I_{ON}/I_{OFF}$ ratio is a strong function of $V_{DS}$, varying between 250 and 1E12 which is as expected for a Schottky S/D FET. The $I_{ON}/I_{OFF}$ as a function of $T_{Si}$ and $L_G$ for large $V_{DS}$ is summarized in **Figure 5.4**b.

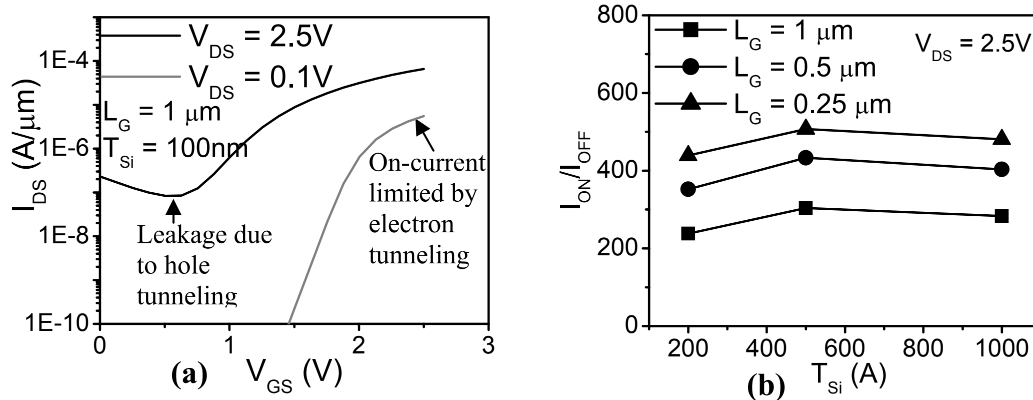



**Figure 5.4**: wireFET simulation results: (a) $I_D$-$V_{GS}$ characteristics show the feasibility of achieving FET action from the wireFET. $I_{ON}/I_{OFF}$ > 1E10 at low $V_{DS}$, and 250-400 at $V_{DS}$ = 2.5V, are achieved owing to different degrees of tunneling through the S/D Schottky barriers. (b) $I_{ON}$/ $I_{OFF}$ as a function of $T_{Si}$ and $L_G$ at $V_{DS}$ = 2.5V.

## 5.2 Aluminum Induced Crystallization (AIC) of Silicon

The wireFET concept can be realized by appropriately choosing a metal-silicon system such that (i) metal induced crystallization of silicon is possible in that system and (ii) the metal is CMOS compatible. The simplest such system is the Aluminum-Silicon system (phase-diagram shown in **Figure 5.5**), which is a simple eutectic system (no silicide phase), with eutectic temperature ($T_{eutectic}$) of $577^0C$. Recently, there has been a lot of interest in applying this system to devices like solar cells [16], and BJTs [17].

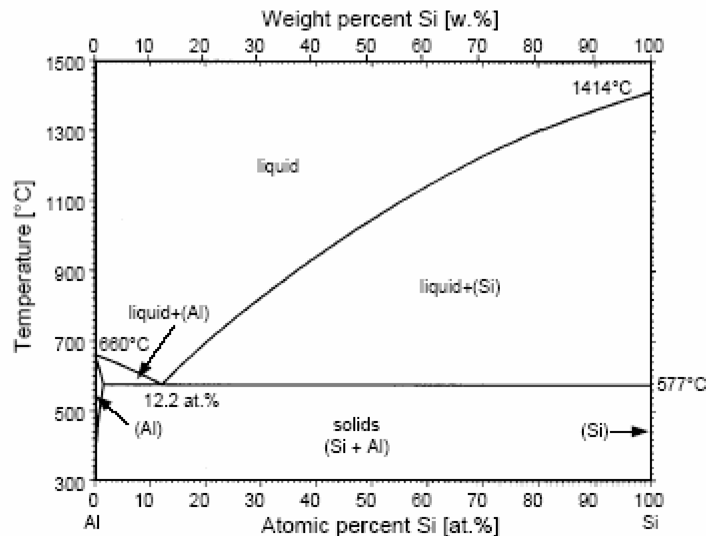



**Figure 5.5**: Binary phase diagram of the Al-Si system [19].

Annealing a stack of Al and amorphous silicon (a-Si) below $T_{eutectic}$ (as low as $150^0C$ reported), leads to aluminum induced crystallization (AIC) of silicon which also causes the aluminum and silicon layers to completely exchange positions, as shown in **Figure 5.6**a [18]. One of the proposed mechanisms for this inter-diffusion process is described in detail by Oliver Nast [18, 19]. In short (**Figure 5.6**b), the silicon (1) dissociates at the Al-Si interface (2) diffuses into the aluminum matrix, (3) locally supersaturates the aluminum, leading to nucleation of silicon grain followed by (4) grain growth within the

aluminum matrix and layer exchange. The polycrystalline silicon layer so formed is saturated with aluminum giving rise to a naturally p-type doped silicon film. To make n-type silicon channel, it will be required to start with n-type doped amorphous silicon film to compensate for the aluminum doping. In this work, AIC of undoped amorphous silicon has been employed to fabricate wireFETs.

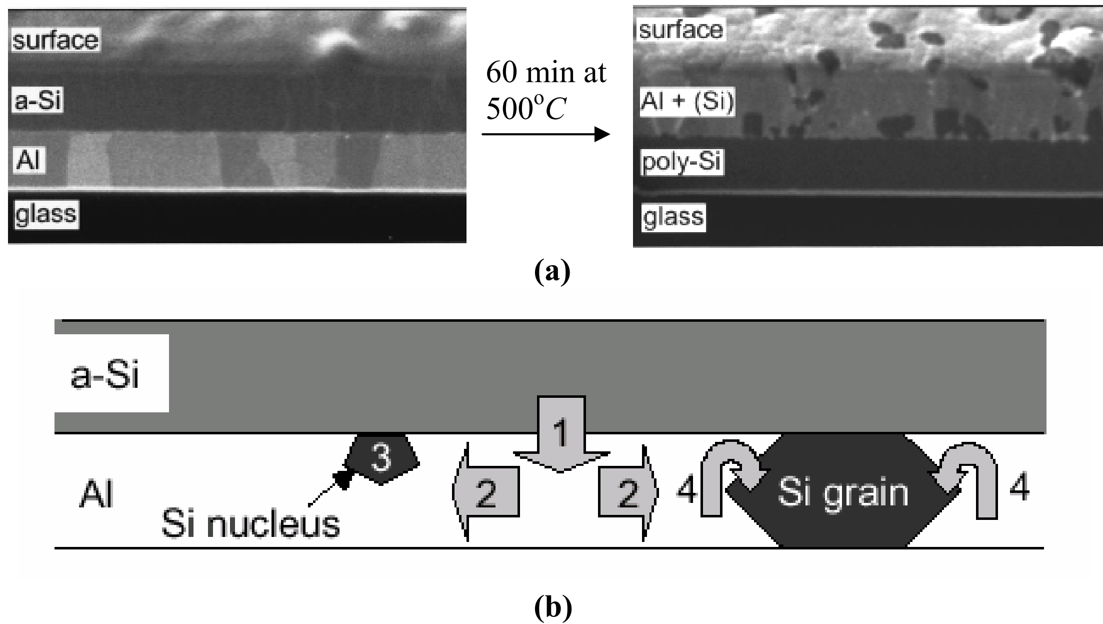



**Figure 5.6**: (a) Cross sectional FIB micrographs demonstrating the layer exchange between Al and Si. (b) A four step mechanism for AIC proposed by Nast [15].

Another interesting approach recently discussed in literature [20], follows a similar device concept using the nickel-silicon system. This system is not a simple eutectic system. In this approach, one starts with silicon nanowires in all the wire regions. All areas except the required channel regions, are converted into nickel silicide interconnects, thus forming wireFETs using Ni-Si system and nickel silicide interconnects. This requires moving to a completely different interconnect material, but

may be a very promising approach for 3D integration, given the properties of nickel silicide as interconnect.

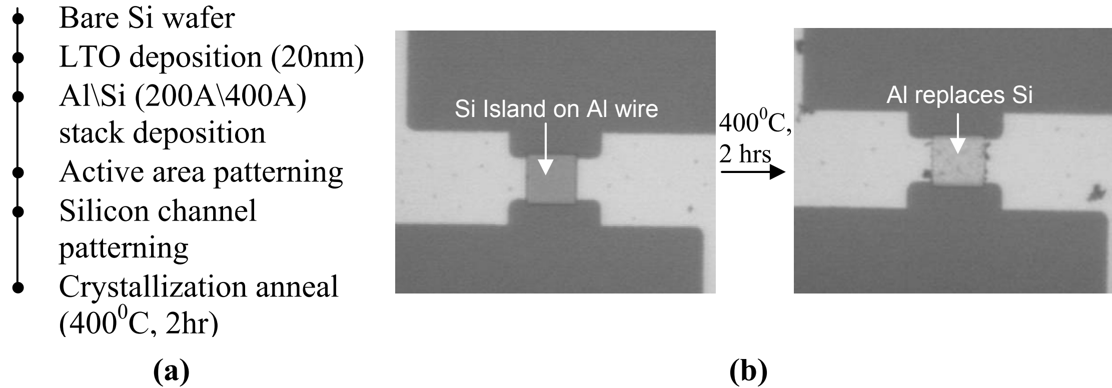



**Figure 5.7**: (a) Simple process flow for a gated-resistor structure. (b) Demonstration of layer exchange process implemented on an aluminum wire (L=10μm, W=10μm) at $400^0$C, 2hr. A distinct change in color is observed after the AIC anneal.

## 5.3 AIC Process Optimization

### 5.3.1 Proof-of-Concept

A simple proof-of-concept wireFET process flow (**Figure 5.7**a) involves formation of an aluminum wire, formation of a silicon island on top of it, and annealing the structure below the Al-Si eutectic temperature (2hr at $400^o$C in $N_2$ ambient). This results in a polycrystalline-Si (poly-Si) region embedded within the wire, which can subsequently serve as the channel of a transistor. **Figure 5.7**b shows micrographs of an Al wire test structure before and after AIC annealing showing a distinct change in the color of the channel region, indicative of the layer exchange process. Electrical measurements indicate an increase in the resistance of the wire after the layer exchange process (**Figure 5.8**a), and corresponds to Si doped to a concentration of ~5E18 $cm^{-3}$. By

applying a bias to the Si substrate (which is electrically isolated from the Al wire by a 20nm-thick $SiO_2$ layer), the resistance of the wire can be modulated (**Figure 5.8**b), proving that a semiconductor region exists within it. Hydrogen plasma treatment has been demonstrated [21] as a good passivation technique for TFTs, which typically employ polycrystalline silicon channels. Hydrogen plasma treatment (30min at 350°C) to passivate defects within the poly-Si further results in improved gate control, though FET like IV characteristics are not obtained. This could be because of the relatively thick and poor quality (Low Temperature LPCVD Oxide) gate oxide used in this run. A more refined process will be discussed in section 5.4.1.

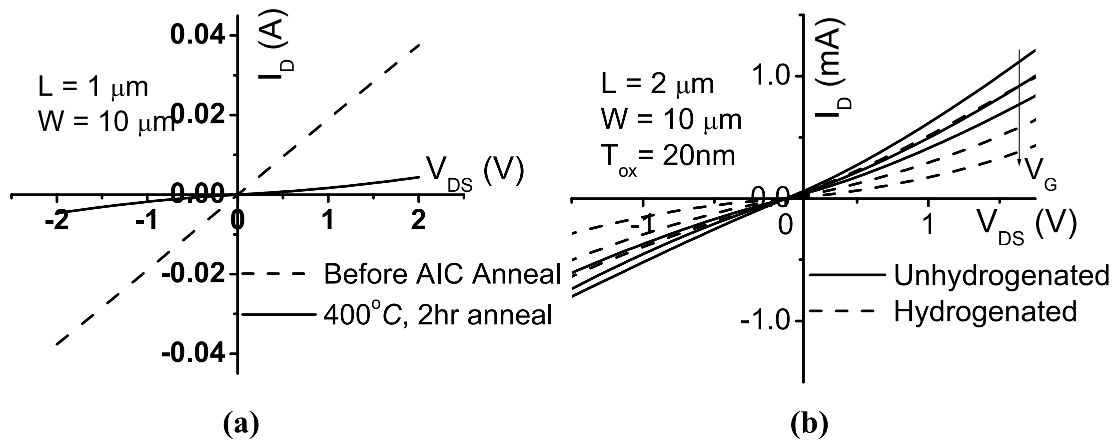



**Figure 5.8**: Change in electrical characteristics of an Al wire after AIC: (a) Increase in wire resistance and (b) field dependence with gate bias (substrate as gate) indicate the presence of silicon embedded within the Al wire. Hydrogenation further improves the effect of gate bias.

## 5.3.2 Role of Native Oxide Layer

The properties of the poly-Si film obtained from the AIC process can be influenced by various process parameters [18, 22, 23] such as Al grain size, Al:Si thickness ratio, Al-Si interface, and annealing temperature. Short-flow experiments were

conducted to study the impact of some of these process parameters on the quality of the AIC poly-Si, in order to optimize the process for wireFET fabrication.

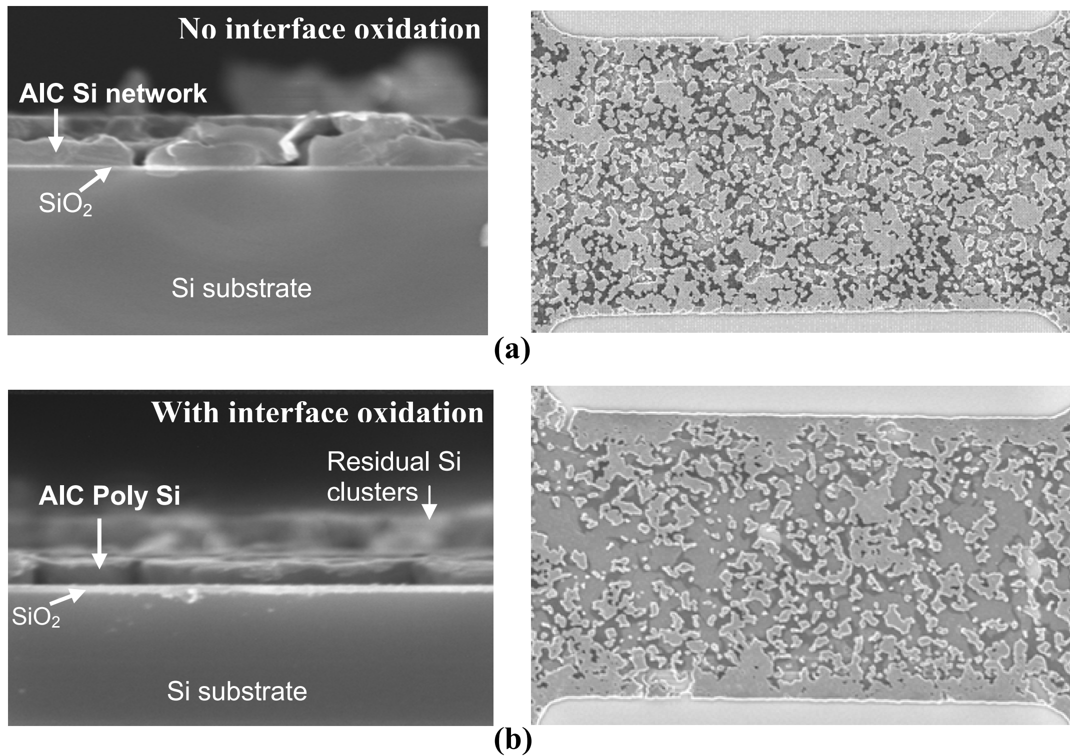


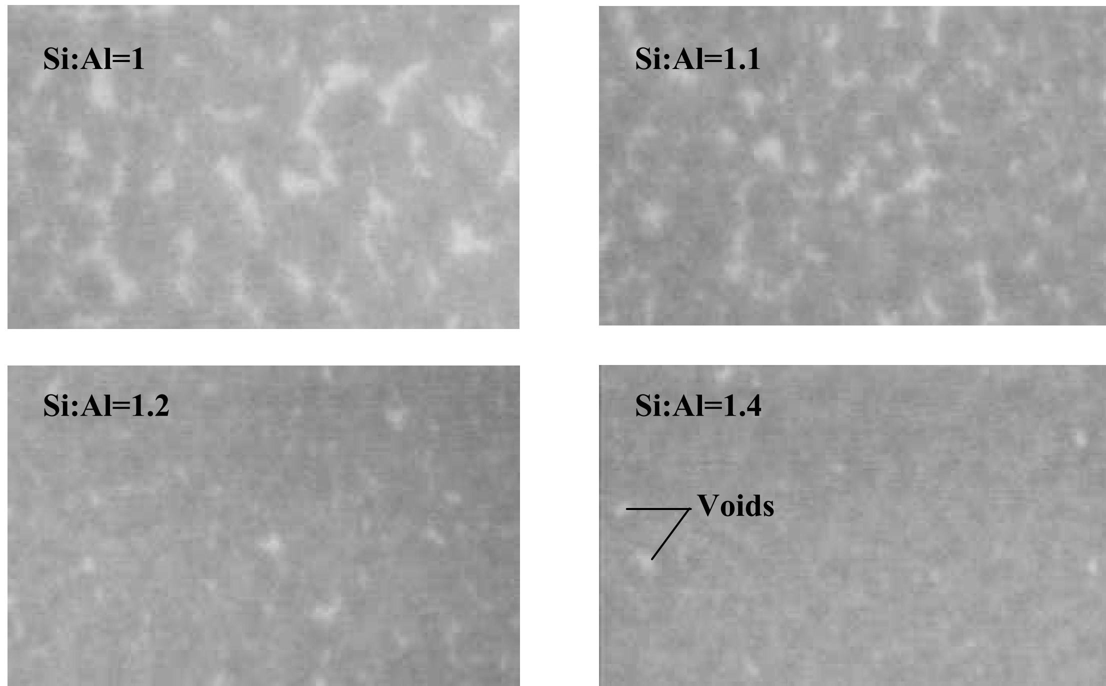

**Figure 5.9**: Effect of interfacial aluminum oxide layer between Al and Si during the AIC process: Cross sectional SEM micrographs (left-side) demonstrate the effect of (a) absence and (b) presence of interfacial aluminum oxide. A thin aluminum oxide layer is necessary to ensure a continuous silicon film. Top view SEM (right-side) micrographs demonstrate clustering by excess silicon on top of the AIC Si film.

Blanket films were used to study the impact of a native aluminum-oxide layer at the Al-Si interface. Samples were prepared on oxidized silicon substrates by sputtering Si over sputtered Al, with or without an intentional interfacial oxidation of Al. In order to avoid sources of contamination, the interfacial oxidation was done by flowing pure $O_2$ at ~2mtorr for 2 minutes on the aluminum film, immediately after sputter deposition

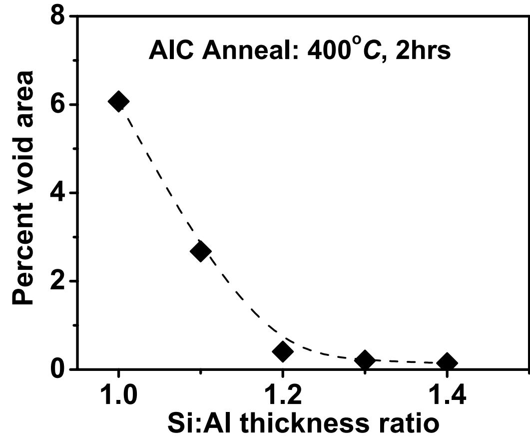
without breaking vacuum. This was followed by AIC anneal and removal of the displaced Al layer. Consistent with previous findings [22], an interfacial oxide layer causes the AIC process to form a very continuous Si film, whereas the absence of an interfacial oxide layer results in discontinuous Si islands (cross sectional SEM analysis shown in **Figure 5.9**) which cannot be used for FET fabrication. The presence of an aluminum oxide layer creates a membrane that tends to confine the nucleated poly-Si grains within the aluminum matrix to grow only within this matrix. However, some poly-Si clusters are also formed (by the excess silicon) on top of the continuous silicon film. This cluster formation can possibly be reduced by increasing the interfacial oxide thickness (thereby, increasing poly-Si grain size and improving its confinement) and reducing Si:Al ratio, with a trade off in the crystallization time. This has not been investigated in this work, but is strongly suggested for future work.

### 5.3.3 Optimum Si:Al Thickness Ratio

Reducing the Si:Al ratio can be used to minimize cluster formation by minimizing the excess silicon available for clustering. Blanket films with different Si:Al layer thickness ratios (in the range from 1 to 1.4) were formed on quartz substrates to allow for visual inspection of the AIC poly-Si layers from the wafer backsides. All samples were made with an intentional interfacial oxidation between the Al and Si films. **Figure 5.10**a shows micrographs from the wafer backsides for the different thickness ratios. Through simple image processing of these micrographs, the percentage area occupied by voids/discontinuities can be computed and is plotted against the thickness ratio in **Figure 5.10**b. Being an interdiffusion process, AIC inevitably tends to leave small agglomerates of Al within the crystallized silicon film and vice versa. Therefore, Si:Al layer thickness




**(a)**




**(b)**

**Figure 5.10**: Impact of starting Si:Al layer thickness ratio: (a) Backside optical views of the AIC silicon layer with different Si:Al starting thickness ratios. (b) Percentage of voids in the crystallized film, extracted from micrograph images. Si:Al thickness ratio of at least 1.4:1 is needed to ensure a continuous poly-Si film.

ratio should be greater than 1 to account for the loss of silicon in the displaced aluminum layer and achieve a continuous silicon film. It is observed that Si:Al ≥ 1.4:1 is desirable

to achieve a continuous poly-Si layer. Please note that increasing the interfacial oxide thickness is expected to decrease this ratio. A suitable balance between, interfacial oxide thickness (therefore, AIC anneal time) and Si:Al thickness ratio should be used.

### 5.3.4 TEM Analysis of AIC Polysilicon

**Figure 5.11** shows cross-sectional transmission electron micrographs before and after the layer exchange process. X-ray diffraction analysis confirms the crystallization of silicon by aluminum. However, the precise composition of the two layers after annealing is not confirmed. Further TEM and AFM analyses will be discussed in section 5.4.

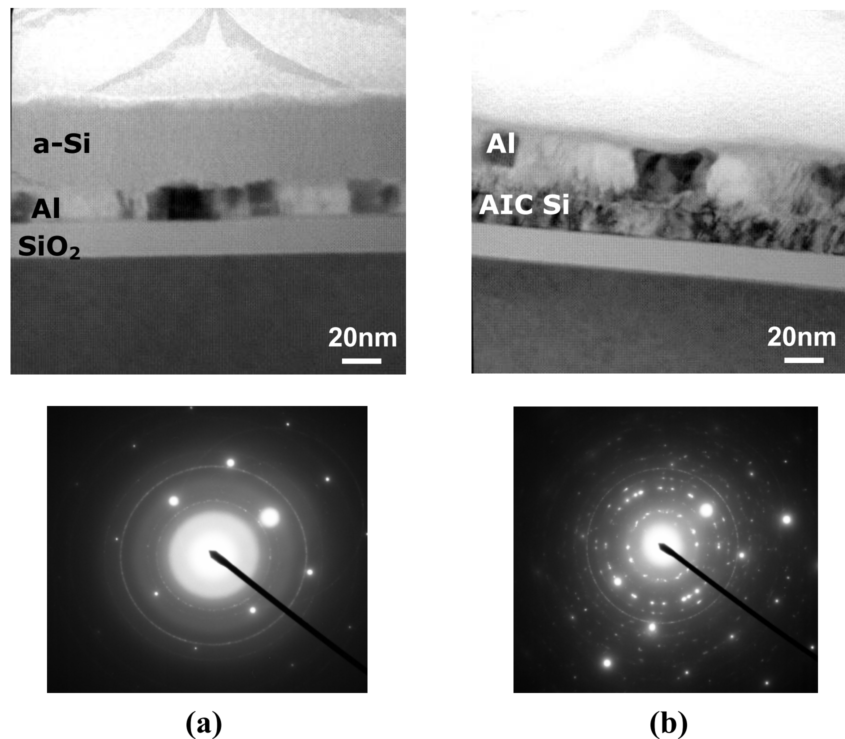



**Figure 5.11**: X-TEM micrographs and X-ray diffraction analyses (a) before anneal and (b) after $400^0C$, 2hours anneal, showing the layer exchange and crystallization of silicon. Presence of new silicon rings after annealing confirms the polycrystallinity of the silicon film.

## 5.4 Improved WireFET Process

The simple proof-of-concept wireFET process flow is not a very robust method and can fail frequently. The main drawback of this process is the incapability to selectively remove the aluminum layer formed on top of the AIC poly-Si layer without disturbing the rest of the aluminum wire. This leads to direct shorts from source to drain on multiple occasions and calls for a modified process flow. Moreover, the gate oxide thickness also needs to be thinned down for better gate control.

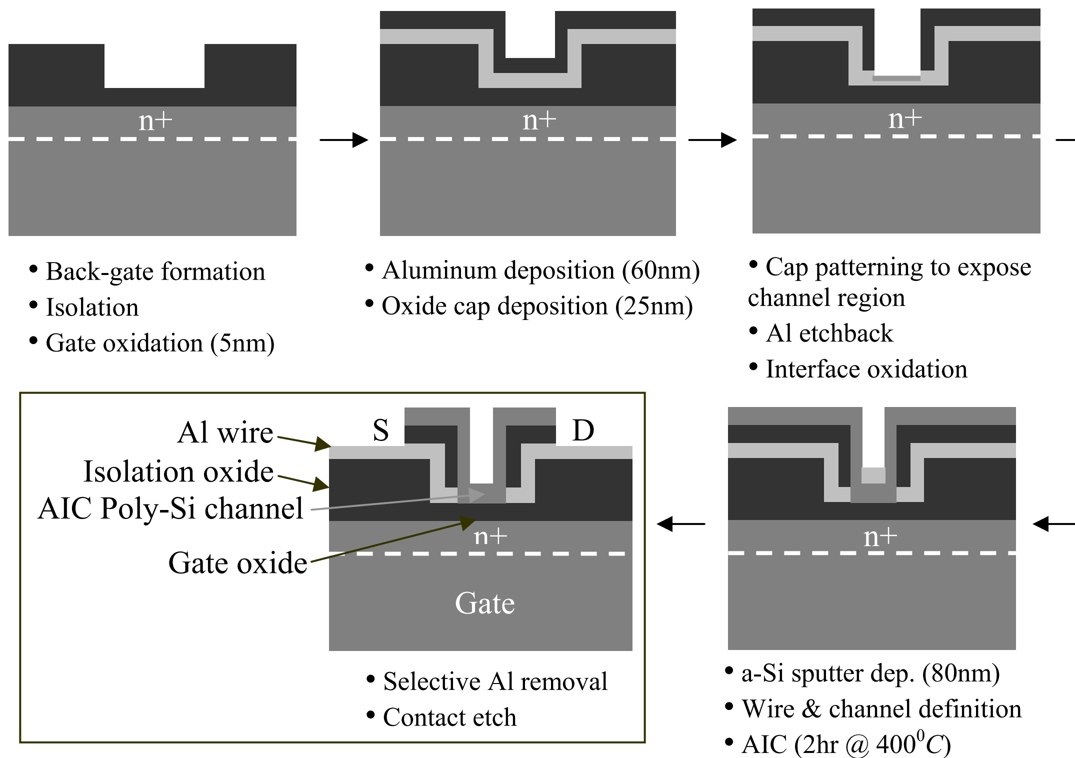



**Figure 5.12**: Schematic representation of improved wireFET process flow.

### 5.4.1 Improved WireFET Process Flow

The refined process flow is shown schematically in **Figure 5.12** and the process flow table is given in Appendix C. This process was used to fabricate wireFETs using the n+ doped Si substrate as the bottom gate. (Another interconnect layer can be used to form

the gate electrode, in a 3-D integrated process.) In this process, AIC occurs selectively in the channel region, and the layer-exchanged Al can be removed selectively without disturbing the rest of the aluminum wire. An example of this is demonstrated in the optical micrograph shown in **Figure 5.13**. Excluding the thermal anneal used to activate dopants in the Si substrate (which would not exist in a real wireFET process), the thermal budget required to fabricate the wireFETs is 90min at $400^0C$.

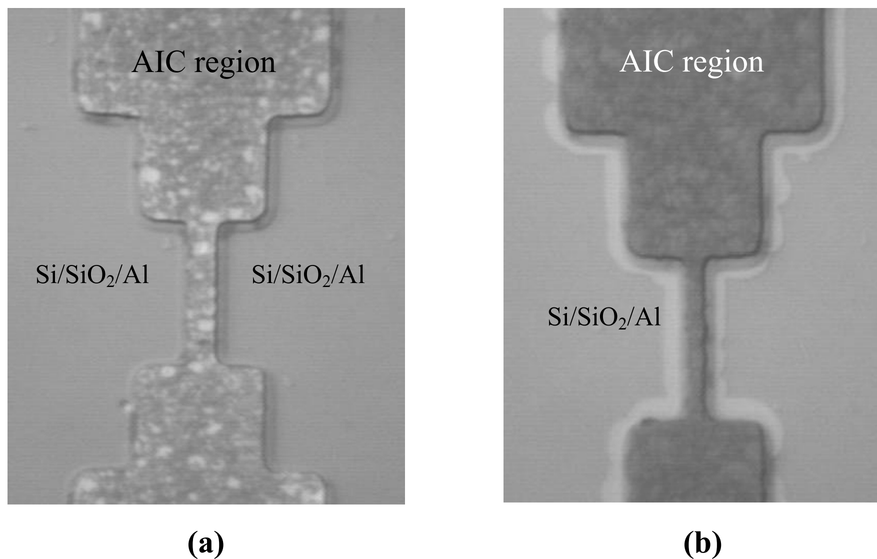



**Figure 5.13**: Optical micrographs (a) before and (b) after aluminum etch, demonstrating the ability to selectively etch Al on top of AIC silicon (without disturbing the remaining Al regions).

## 5.4.2 Electrical Measurements

**Figure 5.14**a shows the measured $I_D$-$V_D$ characteristics of a wireFET with $T_{Si}$ = 50nm. Gated resistor behavior is seen once again! This is inconsistent with the simulated results, suggesting that some of the ideal assumptions made in the simulations do not hold. A tunneling ohmic contact to Si is observed (inset **Figure 5.14**a) due to possible supersaturation of Al in poly-Si grain boundaries. This makes the device a depletion-

mode PMOSFET, which can be turned off with the application of gate bias only if $T_{Si}$ is less than the maximum depletion width. Simulation results with ohmic contacts and 5E18cm$^{-3}$ doping, for the fabricated device are shown in **Figure 5.14**b. The trend is consistent with the observed experimental data and confirms the need for a thinner $T_{Si}$. Following this, thin body devices with AIC poly-Si thickness between 10nm and 16nm were fabricated. But it is still found to be difficult to turn off the transistor. Furthermore, the gate leakage current in these FETs is very high, especially for large area devices.

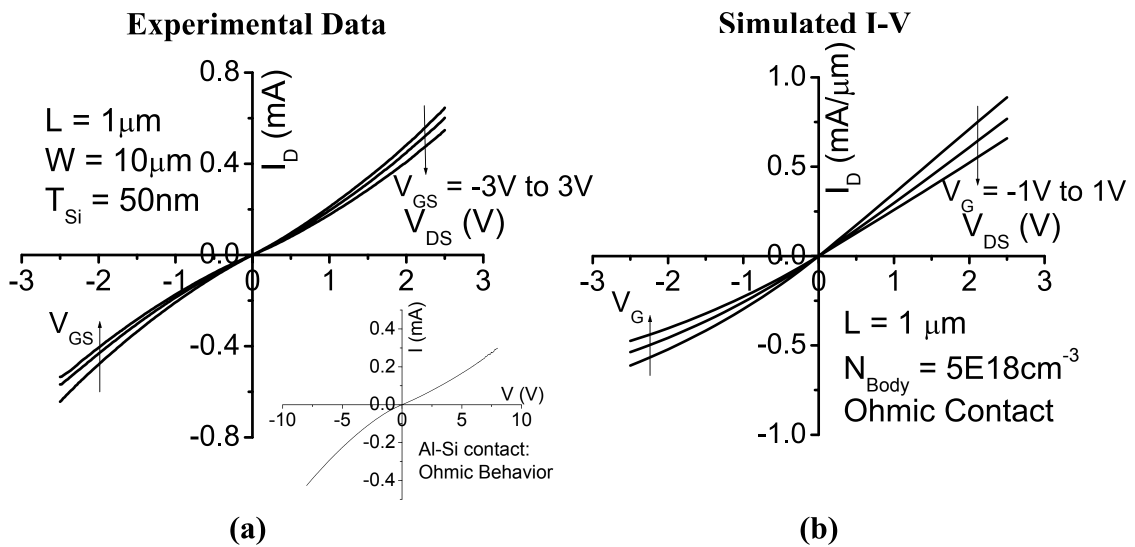



**Figure 5.14**: (a) $I_D$-$V_D$ obtained from fabricated FETs showing gated resistor behavior. Inset shows ohmic behavior of Al-Si contact giving rise to depletion-mode operation. (b) Simulations with ohmic contact between metal and Si, follows similar trend in device characteristics as experimental data.

The possible causes for the behavior described above can be multifold: (i) gate dielectric damage caused by aluminum sputtering (ii) aluminum concentration in thin poly-Si much higher than bulk solid solubility (iii) aluminum segregation near grain

boundaries creating conduction path from source to drain and (iv) poor gate-channel interface quality due to AIC process. Some of these factors have been analyzed here.

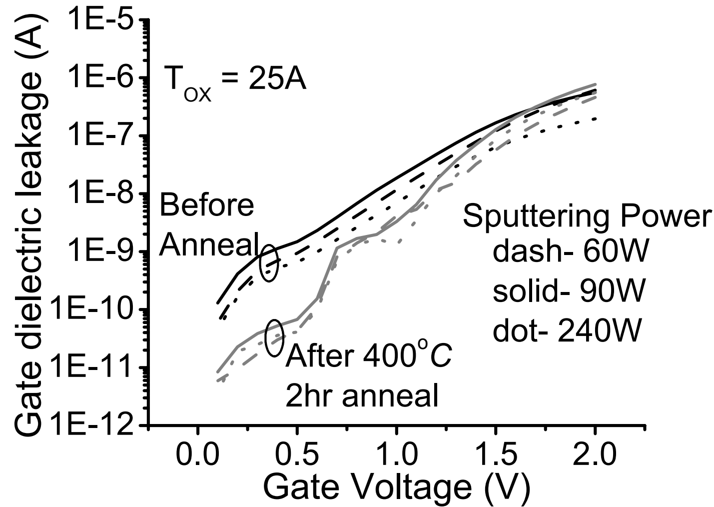



**Figure 5.15**: Gate leakage measurements on large area capacitor structures before and after anneal (AIC thermal budget) indicate no significant dielectric degradation caused by sputter damage.

Gate leakage measurements were conducted (aluminum sputtered on 25A thermal oxide) to study the effect of sputtering on gate dielectric quality. The gate leakage was measured before and after a test thermal anneal (2hrs at $400^0C$ corresponding to AIC thermal budget) for different sputtering powers. **Figure 5.15** shows the gate leakage as function of gate voltage for 100μmx10μm capacitor structures. There is no significant damage caused either by sputtering or by the thermal anneal of pure aluminum, eliminating hypothesis (i) described above.

**Figure 5.16** shows high frequency C-V measurements (100kHz) on capacitors formed by AIC poly-Si channel and the substrate as the gate. For this experiment a somewhat thick gate oxide (9nm) was used with AIC silicon thickness of 50nm to ensure

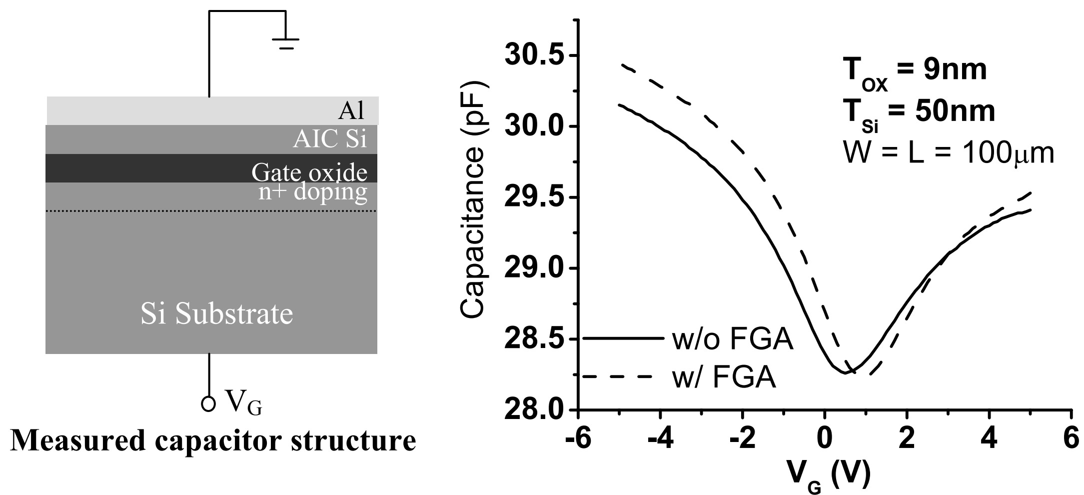



**Figure 5.16**: High frequency C-V measurement of channel formed by AIC shows accumulation depletion and inversion behavior. Onset of inversion and very high $C_{MIN}$ indicate poor interface quality and/or grain boundary defects.

full depletion of the channel. The C-V characteristic looks similar to a MOS capacitor (shows accumulation, depletion and inversion regions), proving the presence of a semiconductor channel. However, two things are peculiar about this data. Firstly, this was a high-frequency measurement and so, the inversion region should not come back up. Secondly, even though the accumulation $C_{OX}$ is reasonable, the $C_{MIN}$ value is too high (even 1E20cm$^{-3}$ aluminum doping in poly-Si does not explain it). Both these factors point towards a poor gate-oxide to channel interface and/or grain boundary defects (in favor of hypothesis (iv)) and needs further investigation.

### 5.4.3 Further Material Characterization

More material analyses have been conducted to further characterize the ultra-thin poly-Si channels formed by AIC. **Figure 5.17** shows high-resolution XTEM images of

the ultra-thin body wireFET channel region with the aluminum removed from the top after the AIC process.

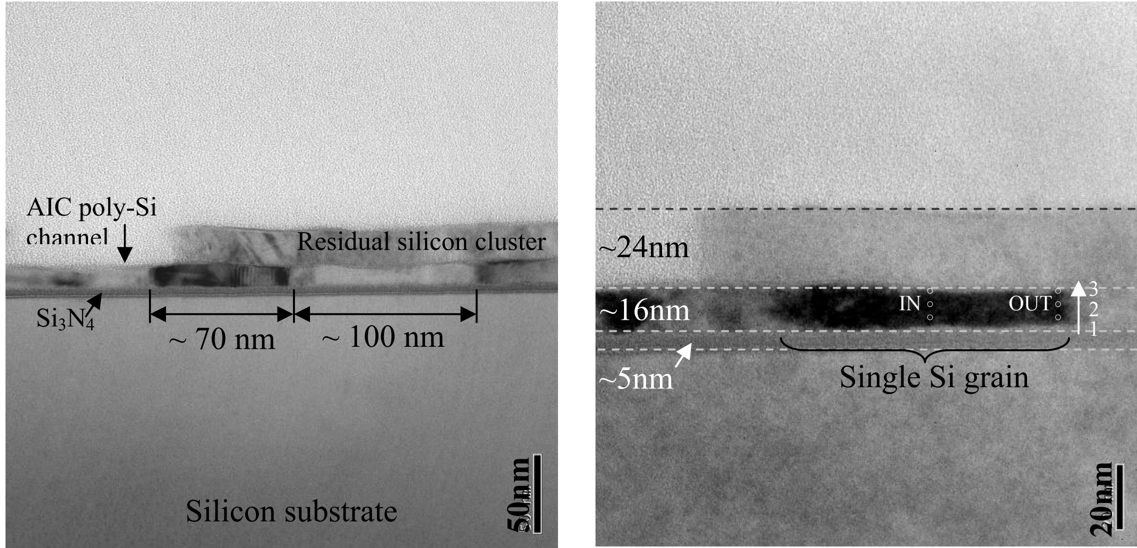



**Figure 5.17**: Cross sectional TEM analysis shows ultra thin AIC film with upto 100nm grains. The film thicknesses are as expected. EDS data has been collected at the 'IN' and 'OUT' points in the right figure.

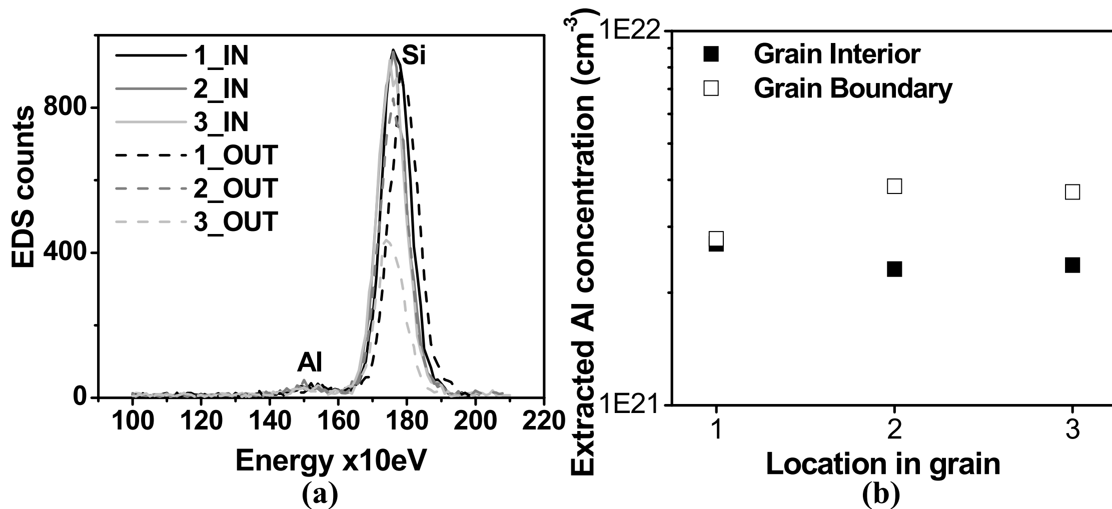



**Figure 5.18**: (a) EDS raw data at six different locations in a silicon grain (refer to Figure 5.17) shows finite Al and Si peaks. (b) Extracted aluminum concentration is in the $10^{21}cm^{-3}$ range (too high), with higher value near grain boundaries (Al segregation).

The film thicknesses and continuities/discontinuities are as expected and the largest AIC poly-Si grains are about 100nm in size for 16nm films. Energy dispersive spectroscopy (EDS) analysis (**Figure 5.18**a) was conducted at six different points on a grain, as shown (**Figure 5.17**), to study the aluminum content in silicon as a function of grain boundary location. A small but significant aluminum peak is obtained suggesting high aluminum doping concentration in the poly-Si. Based on the EDS signal intensities, the concentration of aluminum is related to that of silicon by the relation,

$$\frac{C_{Al}}{C_{Si}} = k_{Al-Si}\left(\frac{I_{Al}}{I_{Si}}\right)$$

Where, $k_{Al-Si}$ (the Cliff-Lorimer factor) is about 1.25 [24], $C_{Al}$, $C_{Si}$ are the concentrations of Al and Si and, $I_{Al}$ and $I_{Si}$ are the peak intensities of their EDS signals. From the peak intensities, aluminum doping greater than 1E21cm$^{-3}$ is extracted (**Figure 5.18**b), which is much higher than the expected solid solubility value suggesting (a) either significant super-saturation for ultra-thin films and/or (b) grain-boundary segregation of aluminum creating parasitic conductions paths from source to drain (hypotheses (ii) and (iii) from the previous subsection). The extracted doping concentration in the grain boundary region is typically higher than the grain interior (**Figure 5.18**b) indicating aluminum segregation. It should be noted that EDS is not the most accurate method for quantitative compositional analysis (especially where the concentration of an element is quite low) and therefore, it is important to look at other analyses to get a precise value of the doping concentration. AFM analysis data is shown in **Figure 5.19**. The data is as expected and consistent with the TEM pictures, further confirming the presence of a continuous AIC poly-Si film (with occasional pin holes) and silicon clusters on top of it.

The electrical measurements and microscopy analyses strongly indicate high aluminum concentration (possible aluminum segregation in grain boundaries) in conjunction with grain-boundary defects. Further process optimizations will therefore, be necessary to completely explore this device, and other metal-Si or metal-Ge systems should also be explored. Suggestions for further process improvements/optimizations are provided in the next chapter.

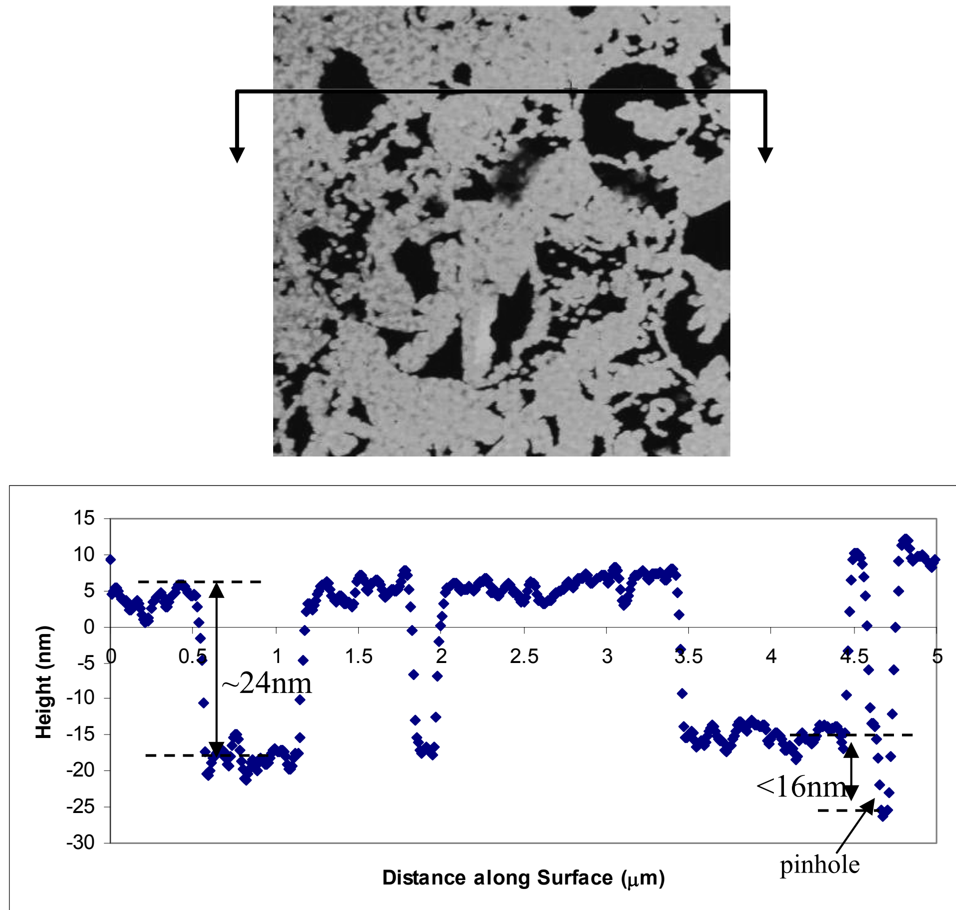



**Figure 5.19**: AFM scan in the channel region shows data consistent with TEM results. Occasional pinholes also detected in the AIC channel film.

## 5.5 Summary

A low thermal-budget, simple and cost-effective method to form FETs within interconnects (wireFET) has been introduced in this chapter. Aluminum-induced crystallization of silicon has been used to fabricate Si FETs directly within Al wires. The simplicity and low thermal budget of this technology makes it attractive for inexpensive implementation of configurable interconnects and tunable passive devices, toward 3-D integrated circuits.

Aluminum induced crystallization of silicon leads to a complete layer exchange between the aluminum and silicon layers. An interfacial aluminum oxide (between the a-Si and Al films), along with an optimum initial Si:Al thickness ratio ensure a uniform poly-Si film. Agglomeration of excess silicon on top of the AIC poly-Si is unavoidable. Although, not studied in this chapter, it is suggested that a good choice of interfacial oxide thickness and Si:Al should be identified to minimize this clustering effect.

Preliminary electrical measurements as well as electron microscopy analyses provide sufficient evidence for the layer exchange and the presence of a semiconductor channel. It is however, very difficult to turn off the FET device and it has a gated resistor behavior. Electrical measurements indicate an ohmic contact of Al to Si indicating the need for thin film FETs. Further, measurements as well as material characterization tend to indicate high concentrations of aluminum in the crystallized silicon film (possibly grain boundary segregation) as well as defect states (grain-boundary defects and/or interface states) making it challenging to achieve good FET characteristics with the Al-Si system.

## 5.6 References

[1] S. Balasubramanian, L. Chang, B. Nikolic, and T.-J. King, "Circuit-Performance Implications for Double-Gate MOSFET Scaling below 25nm," *Proceedings of the 2003 Silicon Nanoelectronics Workshop,* pp. 16-17, June 2003.

[2] J. A. Davis, R. Venkatesan, A. Kaloyeros, M. Beylansky, S. J. Souri, K. Banerjee, K. C. Saraswat, A. Rahman, R. Reif, and J. D. Meindl, "Interconnect limits on gigascale integration (GSI) in the 21st century," *Proceedings of the IEEE,* vol. 89, pp. 305-24, 2001.

[3] P. Kapur, G. Chandra, J. P. McVittie, and K. C. Saraswat, "Technology and reliability constrained future copper interconnects. II. Performance implications," *IEEE Transactions on Electron Devices,* vol. 49, pp. 598-604, 2002.

[4] D. Connelly, C. Faulkner, and D. E. Grupp, "Performance advantage of Schottky source/drain in ultrathin-body silicon-on-insulator and dual-gate CMOS," *IEEE Transactions on Electron Devices,* vol. 50, pp. 1340-5, 2003.

[5] J. Kedzierski, P. Xuan, E. H. Anderson, J. Bokor, T.-J. King, and C. Hu, "Complementary silicide source/drain thin-body MOSFETs for the 20 nm gate length regime," in *International Electron Devices Meeting.*, 2000, pp. 57-60.

[6] A. Kinoshita, C. Tanaka, K. Uchida, and J. Koga, "High-performance 50-nm-gate-length Schottky-source/drain MOSFETs with dopant-segregation junctions," in *2005 Symposium on VLSI Technology*, 2005, pp. 158-9.

[7] K. Banerjee, S. J. Souri, and K. C. Saraswat, "3-D ICs: A Novel Chip Design for Improving Deep Submicron Interconnect Performance and Systems-on-Chip Integration," *Proceedings of the IEEE,* vol. 89, No. 5, pp. 602-33, May 2001.

[8] A. Rahman, S. Das, A. P. Chandrakasan, and R. Reif, "Wiring requirement and three-dimensional integration technology for field programmable gate arrays," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 11, pp. 44-54, 2003.

[9] A. W. Topol, D. C. La Tulipe Jr., L. Shi, D. J. Frank, K. Bernstein, S. E. Steen, A. Kumar, G. U. Singco, A. M. Young, K. W. Guarini, and M. Ieong, "Three-dimensional Integrated Circuits," *IBM Journal of Research and Development,* vol. 50, No. 4/5, pp. 491-506, 2006.

[10] V. Subramanian and K. C. Saraswat, "A Novel Technique for 3-D Integration: Ge-seeded Laterally Crystalluzed TFTs," *1997 Symposium on VLSI Technology,* pp. 97-98, 1997.

[11] V. W. C. Chan, P. C. H. Chan, and M. Chan, "Three dimensional CMOS integrated circuits on large grain polysilicon films," in *IEEE International Electron Devices Meeting*, 2000, pp. 161-4.

[12] H. Takeuchi, A. Wung, X. Sun, R. T. Howe, and T.-J. King, "Thermal Budget Limits of Quarter-Micrometer Foundry CMOS for Post-Processing MEMS Devices," *IEEE Transactions on Electron Devices,* vol. 52, No. 9, pp. 2081-6, September 2005.

[13] V. Varadarajan, Y. Yasuda, S. Balasubramanian, and T.-J. K. Liu, "WireFET Technology for 3-D Integrated Circuits," *IEEE International Electron Devices Meeting,* pp. 149-52, 2006.

[14] *Dessis-ISE Users Manual, v. 8.5.5, Synopsys Inc.*

[15] Y. Taur and T. K. Ning, *Fundamentals of Modern VLSI Devices*: Cambridge University Press, 1998.

[16] S. Gall, J. Schneider, J. Klein, K. Hubener, M. Muske, B. Rau, E. Conrad, I. Sieber, K. Petter, K. Lips, M. Stoger-Pollach, P. Schattschneider, and W. Fuhs, "Large-Grained Polycrystalline Silicon on Glass for Thin-Film Solar Cells," *Thin Solid Films,* vol. 511-512, pp. 7-14, 2006.

[17] Y. Civale, L. K. Nanver, and P. Hadley, "Selective Solid-Phase Epitaxy of Ultra-Shallow p+ Aluminum-Doped Silicon Junctions for Integration in Nanodevices," *Proceedings of the 2006 Silicon Nanoelectronics Workshop,* pp. 55-6, 2006.

[18] O. Nast and S. R. Wenham, "Elucidation of the Layer Exchange Mechanism in the Formation of Polycrystalline Silicon by Aluminum-Induced Crystallization," *Journal of Applied Physics,* vol. 88, No. 1, pp. 124-32, July 2000.

[19] O. Nast, "The Aluminum-induced Layer Exchange Forming Polycrystalline Silicon on Glass for Thin-film Solar Cells," in *PhD Thesis*: Hahn-Meitner Institut, Berlin, 2000.

[20] W. Wang, "Interconnect-Driven Nanoelectronic Circuits," *Proceedings of the 2006 Silicon Nanoelectronics Workshop,* pp. 83-4, 2006.

[21] I.-W. Wu, T.-Y. Huang, W. B. Jackson, A. G. Lewis, and A. Chiang, "Passivation Kinetics of Two Types of Defects in Polysilicon TFT by Plasma Hydrogenation," *IEEE Electron Device Letters,* vol. 12, No. 4, pp. 181-3, April 1991.

[22] O. Nast and A. J. Hartmann, "Influence of Interface and Al Structure on Layer Exchange During Aluminum-Induced Crystallization of Amorphous Silicon," *Journal of Applied Physics,* vol. 88, No. 2, pp. 716-24, July 2000.

[23] J. Schneider, J. Klein, M. Muske, A. Schopke, S. Gall, and W. Fuhs, "Aluminum-Induced Crystallization of Amorphous Silicon: Influence of Oxidation Conditions," *Proceedings of the 3rd World Conference on Photovoltaic Energy Conversion,* vol. 1, pp. 106-9, 2003.

[24] P. J. Sheridan, "Determination of Experimental and Theoretical $k_{ASi}$ Factors for a 200-kV Analytical Electron Microscope," *Journal of Electron Microscopy Technique,* vol. 11, pp. 41-61, 1989.

# Chapter 6 : **Conclusions**

## 6.1 Summary

CMOS scaling has been successfully implemented for almost four decades following the trend predicted by Moore's Law. The transistors have primarily been implemented in bulk CMOS technology with $SiO_2$ based gate dielectric and poly-Si gate. This technology, however, is approaching physical limits due to fundamental material and process limits as well as degraded electrical performance due to short channel effects. Non-classical transistor structures (e.g, thin-body MOSFETs) have superior immunity to short-channel effects and are attractive candidates for extending CMOS technology scaling beyond the limits of the bulk MOSFET. They also have higher carrier mobilities due to an undoped channel, which further boosts their performance. Even though electrically they are superior to the conventional bulk MOSFET, they still face many process integration challenges which have raised serious concerns about their manufacturability.

This dissertation has addressed some of the processing issues by looking at device design variations of thin body double-gate MOSFETs. In order to eliminate the need for S/D p-n junctions (which need to be very abrupt and uniform), the accumulation mode double-gate MOSFET has been evaluated as a possible alternative to the conventional

enhancement mode design. In order to resolve the process integration issues involved in dual metal gate technology, especially for very dense layouts, feasibility and design of single gate work function CMOS has been assessed for high-performance, low operating power and low standby power applications. At such small scales, many factors (both process induced and temporal on-chip factors) contribute towards statistical variations in the electrical performance. The issue of process induced variations has been discussed and a detailed analysis of statistical dopant fluctuation effects and device design rules for variation tolerant transistor design has been provided. The issue of a manufacturable 3-D integration scheme using thin body MOSFETs has also been addressed and a novel technology towards cost-effective three-dimensional integration has been introduced.

Through a simulation study, the accumulation mode double-gate MOSFET has been evaluated for its performance and sensitivities to process variations. The drive current ($I_{ON}$) of a heavily doped channel accumulation mode FET is slightly lower than an enhancement mode device, limited by ionic scattering in the channel. Due to lower vertical electric field which arises from a sub surface conduction path, the accumulation mode design shows lower gate leakage and gate capacitance. The capacitance compensates for the lower $I_{ON}$ and the overall intrinsic delay given by $CV_{DD}/I_{ON}$ is within 10% of the enhancement mode MOSFET. The sensitivity to variations is worse for the accumulation mode design due to significant channel charge and poorer short channel effects due to sub-surface conduction and absence of an intrinsic source-channel potential barrier. An attempt at fabricating a planar back-gated architecture revealed many new challenges arising from the need for ultra-thin buried oxide and an ultra-thin silicon body.

The ultra-thin films pose stringent etch-selectivity requirements, both for the active-area and the gate etches.

Considering that the enhancement mode design has both performance and sensitivity advantages over the accumulation mode design, we get back to enhancement mode FETs. In order to meet the work-function requirements for FinFET devices and to eliminate the poly-Si gate depletion effect, the use of metal gates seems inevitable. However, gate work function engineering for a dual metal gate process becomes impractical for dense FinFET layouts due to close proximity of NMOS and PMOS fins. The feasibility of implementing single gate work function CMOS has been demonstrated for double gate FETs through a proper choice of the S/D design. The threshold voltage depends on the fundamental parameter, the effective channel length ($L_{eff}$), rather than the particular S/D design. Same performance and short channel effects can be obtained with lean spacers and steep S/D profiles or relaxed spacers with gentle gradients, thus providing process flexibility. However, for highly scaled FETs, variations caused by statistical dopant fluctuation effects, require implementation of lean spacers and super steep gradients.

It must be noted that this technology does not leave much room for optimization. The only parameters under control are the S/D design and the silicon body thickness. With a thicker silicon body, a larger $L_{eff}$ is needed to meet the required $I_{OFF}$ constraints. This takes the device deeper into the series resistance limited regime, which reduces $I_{ON}$, but improves sensitivity to process variations, consequently leading to a tradeoff between performance and variability.

In order to get a reasonable estimate of the variations due to statistical dopant fluctuation effects, an accurate methodology to implement 3-D atomistic simulation using a conventional drift-diffusion (DD) simulator has been demonstrated. Given a continuum domain S/D design, the silicon volume of interest is discretized into a pseudo silicon lattice and dopant atoms are randomly generated and placed. Each dopant atom is communicated to the DD simulator in the form of a long range doping profile, which is obtained from its long range Coulomb potential. This method minimizes point dopant induced singularities and simulation grid sensitivities, making it a very robust method. However, long range Coulomb potential assumes an infinite silicon lattice, which breaks down for very short channel advanced transistor structures. It is therefore, recommended that point dopant definition in conjunction with a quantum model (effective in all three directions), like the density gradient model, be used. It was shown through statistical atomitistic simulations that for highly scaled MOSFETs, lean spacers with steep S/D profiles and thick silicon body are desirable to achieve tolerance to variations.

Finally the concept of thin body devices has been extended to three-dimensional integration. Unlike conventional 3-D integration techniques, the idea is to form the FETs within each interconnect layer (instead of a separate device layer), thus reducing process complexity and cost. The aluminum-silicon material system is an ideal and simplest candidate for this process. On annealing an aluminum wire with a silicon island on top, inter-diffusion between Al and Si causes a layer exchange between the two, thus forming a poly-Si channel within an Al interconnect wire. In order to get uniform a poly-Si film, the presence of an interfacial aluminum oxide is very critical. Additionally the Si:Al thickness ratio should be more than 1.4 to ensure the merging of independently growing

grains. The FET formed by this technique responds to gate control and the C-V characteristic further shows success of the layer exchange process. However, turning off the device is found to be very difficult, limiting its application to passive devices at this time. Electrical and material characterization points towards presence of large amounts of defects and a very high concentration of Al within the layer exchanged silicon film. Further process optimization is suggested for successfully turning off the transistor.

## 6.2 Suggestions for Future Research

Considering that advanced FETs like FinFETs are likely to become the future devices in ICs, accumulation mode design is a good candidate to be considered to avoid S/D doping related issues. It was found in this work that planar back-gated design is difficult to implement and it might be a good idea to try a conventional FinFET design since the process is already known to some extent.

To mitigate the process integration issues of dual metal gate technology, both the enhancement mode and accumulation mode designs can be implemented through single gate work function technology. However, the enhancement mode design sees a clear benefit in performance and variation-tolerance, making it more attractive. It would therefore, be important to develop alternative doping techniques to dope the source and drain regions of a silicon fin with high uniformity and steepness. Plasma doping is one such candidate and is still in the evaluation phase. It would be very useful to explore this option in greater detail and enhance its ability towards utilizing it for FinFET applications.

Process induced variations are a serious concern for advanced FETs. In order to study variations in devices and to come up with variation tolerant designs, it is important to incorporate as many sources of variations as possible, not just statistical dopant

fluctuations (SDF). In case of SDF simulations, it is suggested that density gradient model instead of Sano's long-range Coulomb potential model be used to implement dopant atoms. The present approach for dopant generation and placement follows a rather non-physical and completely mathematical model. In order to improve the accuracy of SDF simulations, especially in very small scale devices, it may be important to simulate the physical process of implantation and diffusion phenomena (say Monte Carlo ion implantation and diffusion in a conventional process simulator) and accurately communicate the atomic positions to a DD simulator. It should be noted however, that this is likely to be a computationally expensive technique, with a poor convergence rate.

Finally, looking into the process development of the new 3-D integration technique using wireFETs, a lot more study is required before the process is completely evaluated. It is suggested that a more detailed study of the aluminum distribution in the silicon film be done. This would answer the fundamental question of whether Al-Si system is in fact a good candidate for making wireFETs. Looking into other metal-semiconductor systems would also be very attractive (for example, using Ge instead of Si or other CMOS metals etc) to understand the feasibility of this scheme. With the Al-Si system, it is suggested that the effect of interfacial oxidation time and its relationship to Si:Al ratio be evaluated in detail. This may lead to a convenient process condition, with minimum poly-Si clustering within the desired thermal budget. In order to minimize the amount of grain boundary defects, development of a single grain device with grain boundary passivation techniques should be explored. The current work indicates high concentration of doped aluminum in the poly-Si film. It may be worthwhile looking at aluminum-induced crystallization of silicon using trace amounts of aluminum (say with a

low dose implantation of aluminum in silicon or by using a very large Si:Al ratio). The concept of the wireFET, if implemented on an appropriate metal-semiconductor system, would be a very promising technique for low cost, low thermal budget 3-D integration!

# Appendix A: Accumulation mode MOSFET Process Steps

## Gate Last Process

| Step | Process Name | Equipment | Process Specification | Comments |
|---|---|---|---|---|
| **Back-gate Formation** | | | | |
| 1.1 | Wafer preclean | sink6 | Pirahna, $120^0C$, 10 min + 25:1 HF | |
| 1.2 | Sacrificial oxidation (200A) | tystar2 | $800^0C$, 25 min | Recipe:2WETOXA |
| 1.3 | Inspection | nanoduv | thickness measurement | Recipe: $SiO_2$/Si |
| 1.4 | $BF_2^+$ Implantation | Implanter | 5e15 $cm^{-2}$, 20keV | Core Systems Inc. |
| **Channel Stack deposition** | | | | |
| 2.1 | Wafer preclean | sink6 | Pirahna, $120^0C$, 10 min | |
| 2.2 | Dopant Activation | heatpulse3 | $1000^0C$, 5sec, $N_2$ ambient | |
| 2.3 | Sac. oxide etch | sink6 | 10:1 HF | check for dewetting |
| 2.4 | HTO deposition (50A) | tystar9 | 300mtorr, $800^0C$, $SiH_2Cl_2$ = 18sccm, $N_2O$ = 180sccm, 13min | Recipe:9VHTOA |
| 2.5 | Inspection | sopra | thickness measurement | |
| 2.6 | Channel a-Si deposition (50A/100A) | tystar19 | 300mtorr, $425^0C$, $PH_3/SiH_4$ = 4.7sccm, $Si_2H_6$=100sccm, 7/20 min | Recipe: SELDEPC.019 |
| 2.7 | Inspection | sopra | thickness measurement | |
| 2.8 | PSG dep. (200nm) | tystar11 | $450^0C$, 11 min | Recipe:11SDLTOA |
| 2.9 | Inspection | nanoduv | thickness measurement | Recipe: $SiO_2$/Si |
| **PSG Patterning** | | | | |
| 3.1 | HMDS | primeoven | $90^0$C, 1min | |
| 3.2 | PR coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
| 3.3 | PR Expose | gcaws | Dummy Gate, F: 257, E: 3.6s | |
| 3.4 | PR Develop | svgdev | Bake $120^0C$, 60s; Develop 60 sec, OPD4262 developer | Oven #1,Develop#1 |

| | | | | |
|---|---|---|---|---|
| 3.5 | Inspection | optiphot | | Use yellow filter |
| 3.6 | Hard bake | uvbake | $200^0C$, 60s | |
| 3.7 | Descum | technics-c | 50W, 1 min, $O_2$ plasma | |
| 3.8 | PR ashing | lam5 | 20mtorr, $RF_{top}$=100W, $RF_{bot}$=0, $O_2$=100sccm, ~160s | Recipe: 5962 |
| 3.9 | Inspection | leo | CD measurement | Repeat 3.8 if reqd. |
| 3.10 | PSG etch | lam5 | 13 mtorr, $RF_{top}$=200W, $RF_{bot}$=40W, $CF_4$=100sccm, 99s | Timed etch |
| 3.11 | Inspection | leo | CD measurement | |
| 3.12 | PR strip | technics-c | 300W, 6 min, $O_2$ plasma | |
| 3.13 | Residual PSG etch | sink7 | 100:1 HF, 70s | check for dewetting |
| 3.14 | Inspection | leo | Final PSG CD measurement | |
| 3.15 | Wafer postclean | sink8 | Pirahna, $120^0C$, 10 min | |
| **S/D Deposition and Planarization** | | | | |
| 4.1 | Wafer preclean | sink6 | Pirahna, $120^0C$, 10 min | |
| 4.2 | Native oxide etch | sink7 | 100:1 HF, 10s | |
| 4.3 | n+ S/D Deposition (2000A) | tystar19 | Step1: $350^0C$, 300mtorr, $GeH_4$ =200 sccm, 1min ; #2: $480^0C$, 300mtorr, $Si_2H_6$ =70sccm, $PH_3/SiH_4$=6sccm, 35 min | Recipe: GENUC.019 |
| 4.4 | Inspection | nanoduv | thickness measurement | Recipe: poly/Ox |
| 4.5 | CMP | cmp | poly recipe (~ 40s) | calibrate first |
| 4.6 | etchback | lam5 | B.T: $RF_{top}$=200W, $RF_{bot}$=40W, $CF_4$=100sccm, 13mtorr, 10s; M.E: $RF_{top}$=150W, $RF_{bot}$=75W, HBr=150sccm, 12mtorr | M.E time depends on CMP remains |
| 4.7 | Inspection | optiphot | | |
| 4.8 | Dummy gate etch | sink7 | 100:1 HF, 18min | Timed etch |
| 4.9 | Inspection | leo | | |
| **Field patterning and Channel Crystallization** | | | | |
| 5.1 | HMDS | primeoven | $90^0C$, 1min | |

| 5.2 | PR coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
|---|---|---|---|---|
| 5.3 | PR Expose | gcaws | Active area, F: 245, E: 4.5s | |
| 5.4 | PR Develop | svgdev | Bake $120^0C$, 60s; Develop 60s, OPD4262 developer | Oven #1,Develop#1 |
| 5.5 | Inspection | optiphot | | Use yellow filter |
| 5.6 | Hard bake | uvbake | $200^0$C, 60s | |
| 5.7 | Active area etch | lam5 | B.T: $RF_{top}$=200W, $RF_{bot}$=40W, $CF_4$=100sccm, 13mtorr, 35s O.E: $RF_{top}$=250W, $RF_{bot}$=150W, HBr=200sccm, $O_2$=5sccm, 35mtorr, 4s | Recipe: 5963, No main etch step Timed etch |
| 5.8 | PR strip | matrix | 400W, T=$200^0C$, 90s | |
| 5.9 | Wafer postclean | sink8 | Pirahna, $120^0C$, 10 min | |
| 5.10 | Wafer preclean | sink6 | Pirahna, $120^0C$, 10 min | |
| 5.11 | Solid phase crystallization | tylan6 | $600^0C$, 12 hrs | Recipe: SPC |
| **Inverse Spacer Formation** | | | | |
| 6.1 | HTO deposition (400A) | tystar9 | $700^0C$, $SiH_2Cl_2$=18sccm, $N_20$=180 sccm, 300mtorr, 10.5hrs | Recipe: 9VHTOA |
| 6.2 | Inspection | nanoduv | thickness measurement | Recipe: $SiO_2$/Si |
| 6.3 | blanket HTO etch | lam5 | 13 mtorr, $RF_{top}$=200W, $RF_{bot}$=40W, $CF_4$=100sccm, 10s | 200A timed etch, Recipe: 5963 |
| 6.4 | Inspection | leo | | |
| 6.5 | HMDS | primeoven | $90^0$C, 1min | |
| 6.6 | PR coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
| 6.7 | PR Expose | gcaws | Spacer window, F: 247, E: 6s | |
| 6.8 | PR Develop | svgdev | Bake $120^0C$, 60s; Develop 60 sec, OPD4262 developer | Oven #1,Develop#1 |
| 6.9 | Inspection | optiphot | | Use yellow filter |
| 6.10 | Hard bake | uvbake | $200^0$C, 60s | |
| 6.11 | Spacer etch | lam5 | 13 mtorr, $RF_{top}$=200W, $RF_{bot}$=40W, $CF_4$=100sccm, 15s | Recipe 5003 (B.T), Timed etch |

| 6.12 | PR strip | matrix | 400W, T=200$^0C$, 90s | |
|---|---|---|---|---|
| 6.13 | Wafer postclean | sink8 | Pirahna, 120$^0C$, 10 min | |
| **Gate Stack Deposition and Patterning** | | | | |
| 7.1 | Wafer preclean | sink6 | Pirahna, 120$^0C$, 10 min | |
| 7.2 | Residual spacer etch | sink8 | 100:1 HF, 20-30s | |
| 7.3 | Gate dielectric | heatpulse3 | 600C, 150s, $NH_3$ ambient | Recipe:600150C |
| 7.4 | Dielectric densification | heatpulse3 | 1050$^0C$, spike anneal, Ar ambient | |
| 7.5 | Inspection | sopra | thickness measurement | |
| 7.6 | Gate deposition (400A) | tystar19 | 450$^0C$, 300mtorr, $Si_2H_6$=100sccm, $B_2H_6$=90sccm, 20 min | Recipe: SELDEPC.019 |
| 7.7 | Inspection | nanoduv | thickness measurement | |
| 7.8 | Gate anneal | heatpulse3 | 800C, 30 sec, $N_2$ anneal | |
| 7.9 | HMDS | primeoven | 90$^0$C, 1min | |
| 7.10 | PR coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
| 7.11 | PR Expose | gcaws | Gate Mask | |
| 7.12 | PR Develop | svgdev | Bake 120$^0C$, 60s; Develop 60s, OPD4262 developer | Oven #1,Develop#1 |
| 7.13 | Inspection | optiphot | | Use yellow filter |
| 7.14 | Hard bake | uvbake | 200$^0$C, 60s | |
| 7.15 | Gate etch | lam5 | B.T: $RF_{top}$=200W, $RF_{bot}$=40W, $CF_4$=100sccm, 13mtorr, 10s M.E: $RF_{top}$=300W, $RF_{bot}$=150W, HBr=150sccm, 12mtorr | Recipe: 5003 |
| 7.16 | PR strip | matrix | 400W, T=200$^0C$, 90s | |
| 7.17 | Wafer postclean | sink8, sink6 | Pirahna, 120$^0C$, 10 min | |
| **Etchback Process** | | | | |
| 8.1 | Planarization coating (PC3-700) | spinner1 | 3000rpm, 40s, 8:1 dilution of pc3-700: diluent | Strong stirring required |
| 8.2 | Bake coating | hotplate | 200$^0C$, 2 min | |

| 8.3 | Coating thinning | lam5 | $RF_{top}$=100W, $RF_{bot}$=0W, $O_2$=100sccm, 20mtorr, 3.5min | Recipe: 5962 Timed etch |
|---|---|---|---|---|
| 8.4 | etch rate calibration | lam5 | $RF_{top}$=300W, $RF_{bot}$=150W, $O_2$=2sccm, $Cl_2$=50sccm, HBr=150sccm | Poly : Coating etch rate ~1.5:1; Recipe: 5001 |
| 8.5 | etchback | lam5 | same as 8.3, t=32-40s | |
| 8.6 | Coating strip | matrix | 400W, $200^0C$, 90s | |
| 8.7 | Wafer postclean | sink8, sink6 | Pirahna, $120^0C$, 10 min | |
| | | **Contact etching** | | |
| 9.1 | HMDS | primeoven | $90^0$C, 1min | |
| 9.2 | PR coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
| 9.3 | PR Expose | gcaws | Spacer window, F: 250, E: 6.5s | |
| 9.4 | PR Develop | svgdev | Bake $120^0C$, 60s; Develop 60 sec, OPD4262 developer | Oven #1,Develop#1 |
| 9.5 | Inspection | optiphot | | Use yellow filter |
| 9.6 | Hard bake | uvbake | $200^0$C, 60s | |
| 9.7 | Descum | technics-s | 50W, 60sec | |
| 9.8 | Contact etch | sink7 | 25:1 HF, 1min | |

## Gate First Process

| Step | Process Name | Equipment | Process Specification | Comments |
|---|---|---|---|---|
| | | **LOCOS Isolation** | | |
| 1.1 | Wafer preclean | sink6 | Pirahna, $120^0C$, 10min, 25:1 HF | |
| 1.2 | Pad oxidation (200A) | tystar2 | $800^0C$, 25min | Recipe:2WETOXA |
| 1.3 | Inspection | nanoduv | Thickness measurment | Recipe: Ox/Si |
| 1.4 | HTN deposition (500A) | tystar9 | $800^0C$, 300mtorr, $N_2$=80sccm, $SiH_2Cl_2$=5sccm, $NH_3$=15sccm, 44min | Recipe: 9VHTOA |
| 1.5 | HMDS | primeoven | $90^0$C, 1min | |
| 1.6 | Photoresist (PR) coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
| 1.7 | PR Expose | gcaws | Field Area | Active+Gate mask |
| 1.8 | PR Develop | svgdev | Bake $120^0C$, 60s; Develop 60 s, OPD4262 developer | Oven #1, Develop#1 |
| 1.9 | Inspection | optiphot | | Use yellow filter |
| 1.10 | Hard bake | uvbake | $200^0C$, 60s | |
| 1.11 | Resist ashing | technics-c | 200W, $O_2$ plasma, 3min 45s | Increase field area |
| 1.12 | Nitride etch | lam5 | M.E: $RF_{top}$=300W, $RF_{bot}$=150W, HBr=150sccm, 12mtorr, 42s | Recipe: 5003 Calib. timed etch |
| 1.13 | PR strip | matrix | 400W, $200^0C$, 1.5min | |
| 1.14 | Inspection | asiq | Verify step height | |
| 1.15 | Wafer postclean | sink8 | Pirahna, $120^0C$, 10min | |
| 1.16 | Wafer preclean | sink6 | Pirahna, $120^0C$, 10min | |
| 1.17 | Residual oxide etch | sink6 | 100:1 HF, 100s | Calib. timed etch |
| 1.18 | LOCOS oxidation | tystar2 | $850^0C$, 1hr | Recipe:2WETOXA |
| 1.19 | Inspection | nanoduv | Thickness measurment | Recipe: Ox/Si |
| 1.20 | Pad nitride removal | sink7 | $H_3PO_4$, $160^0C$, 35 min | Calib. timed etch |
| 1.21 | Inspection | asiq | Verify step height | |
| | | **Back-Gate Formation** | | |
| 2.1 | $BF_2^+$ Implantation | Implanter | 20keV, 2E15$cm^{-2}$ | Core Systems Inc. |

| | | | | |
|---|---|---|---|---|
| 2.2 | Wafer preclean | sink6 | Pirahna, $120^0C$, 10min | |
| 2.3 | Damage anneal | tylan6 | $550^0C$, 6hrs | Recipe: SPC |
| 2.4 | Dopant activation | heatpulse3 | $1000^0C$, 5s, $N_2$ ambient | |
| 2.5 | Pad oxide etch | sink6 | 25:1 HF, 125s | Calib. timed etch |
| **Channel Stack Deposition** | | | | |
| 3.1 | Interface nitridation | heatpulse3 | $600^0C$, 150s, $NH_3$ ambient | Recipe:600150C |
| 3.2 | Nitride densification | heatpulse3 | $1050^0C$, spike anneal, $N_2$ amb. | Recipe: spike |
| 3.3 | Inspection | sopra | Thickness measurment | |
| 3.4 | HTN deposition (30A) | tystar9 | $800^0C$, 300mtorr, $N_2$=80sccm, $SiH_2Cl_2$=5sccm, $NH_3$=15sccm, 2min | Recipe:9VNITA |
| 3.5 | Inspection | sopra | Thickness measurment | |
| 3.6 | HTO deposition (25A) | tystar9 | $700^0C$, 300mtorr, $SiH_2Cl_2$ =18sccm, $N_2O$=180sccm, 15min | Recipe:9VHTOA |
| 3.7 | Inspection | sopra | Thickness measurment | |
| 3.8 | a-Si deposition (50/100A) | tystar19 | step 1: 300mtorr, $425^0C$, $Si_2H_6$=100sccm, 1min, step2: 300mtorr, $425^0C$, $PH_3/SiH_4$=4.7sccm, $Si_2H_6$=100sccm, 6/16 min | Recipe: SIGEVARC.019 |
| 3.9 | Inspection | sopra | Thickness measurment | |
| 3.10 | Channel crystallization | tylan6 | $550^0$C, 12hrs | Recipe: SPC |
| **Active Area Patterning** | | | | |
| 4.1 | HMDS | primeoven | $90^0$C, 1min | |
| 4.2 | Photoresist (PR) coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
| 4.3 | PR Expose | gcaws | Active Area | |
| 4.4 | PR Develop | svgdev | Bake $120^0C$, 60s; Develop 60 s, OPD4262 developer | Oven #1, Develop#1 |
| 4.5 | Inspection | optiphot | | Use yellow filter |
| 4.6 | Hard bake | uvbake | $200^0C$, 60s | |

| | | | | |
|---|---|---|---|---|
| 4.7 | Active area etch | lam5 | B.T: $RF_{top}$=200W, $RF_{bot}$=40W, $CF_4$=100sccm, 13mtorr, 5s O.E: $RF_{top}$=250W, $RF_{bot}$=150W, HBr=200sccm, $O_2$=5sccm, 35mtorr, 15s | Recipe: 5963 (short B.T and no M.E) |
| 4.8 | Inspection | optiphot, leo | | |
| 4.9 | PR strip | matrix | 400W, $200^0C$, 1.5min | |
| 4.10 | Wafer postclean | sink8 | Pirahna, $120^0C$, 10min | |
| **Gate Stack Formation** | | | | |
| 5.1 | Wafer preclean | sink6 | Pirahna, $120^0C$, 10min | |
| 5.2 | Gate nitridation | heatpulse3 | $600^0C$, 150s, $NH_3$ ambient | Recipe:600150C |
| 5.3 | Nitride densification | heatpulse3 | $1050^0C$, spike anneal, Ar amb. | Recipe: spike |
| 5.4 | Inspection | sopra | Thickness measurment | |
| 5.5 | Mo deposition | novellus | 300W, 10mtorr, 200A | |
| 5.6 | Inspection | asiq | Thickness measurment | |
| 5.7 | PolySi cap deposition (1000A) | tystar19 | $600^0C$, 300mtorr, $SiH_4$ =100sccm, $PH_3$=5sccm, 54min | Recipe: SIGEVARC.019 |
| 5.8 | Inspection | nanoduv | Thickness measurment | Recipe: polysi/ox |
| 5.9 | Hard mask deposition (800A) | tystar11 | $450^0C$, 3.5min | Recipe: 11SULTOA |
| 5.10 | Inspection | nanoduv | Thickness measurment | Recipe: Ox/Si |
| **Gate Stack Patterning** | | | | |
| 6.1 | HMDS | primeoven | $90^0$C, 1min | |
| 6.2 | Photoresist (PR) coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
| 6.3 | PR Expose | gcaws | Gate, F: 254, E: 4.4s | |
| 6.4 | PR Develop | svgdev | Bake $120^0C$, 60s; Develop 60 s, OPD4262 developer | Oven #1, Develop#1 |
| 6.5 | Inspection | optiphot | | Use yellow filter |
| 6.6 | Hard bake | uvbake | $200^0C$, 60s | |
| 6.7 | PR Ashing | lam5 | 20mtorr, $RF_{top}$=100W, $RF_{bot}$=0, $O_2$=100sccm, ~160s | Recipe: 5962 |

| 6.8 | Inspection | leo | CD measurement | Repeat 6.7 if reqd. |
|---|---|---|---|---|
| 6.9 | Hard mask etch (600A) | lam5 | B.T: $RF_{top}$=200W, $RF_{bot}$=40W, $CF_4$=100sccm, 13mtorr, 35s | Recipe: 5003 Timed etch |
| 6.10 | Hard mask trimming | sink7 | 100:1 HF, ~80s | Calib. timed etch |
| 6.11 | PR strip | matrix | 400W, $200^0C$, 1.5min | |
| 6.12 | Wafer postclean | sink7 | 100:1 HF dip | |
| 6.13 | Inspection | asiq | Step height measurement | |
| 6.14 | Poly-Mo stack etch | lam5 | B.T: $RF_{top}$=200W, $RF_{bot}$=40W, $CF_4$=100sccm, 13mtorr, 7s<br>M.E: $RF_{top}$=300W, $RF_{bot}$=150W, HBr=150sccm, 12mtorr, E.P +20%O.E<br>Mo etch: $RF_{top}$=150W, $RF_{bot}$=100W, Cl2=70sccm, $O_2$=56sccm, 13mtorr, 8s | Recipe: 5003 |
| 6.15 | Inspection | asiq | Step height measurement | |
| 6.16 | Inspection | leo | | |
| | | **Side-wall Spacer Formation** | | |
| 7.1 | Wafer postclean | sink7 | 100:1 HF dip | |
| 7.2 | LTO spacer deposition (400A) | tystar11 | $450^0C$, 1min45s | Recipe: 11SULTOA |
| 7.3 | Spacer etch (300A) | lam5 | B.T: $RF_{top}$=200W, $RF_{bot}$=40W, $CF_4$=100sccm, 13mtorr, 20s | Recipe: 5003 Timed etch |
| 7.4 | Residual spacer etch | sink7 | 100:1 HF, ~80s | Calib. timed etch |
| | | **S/D Formation** | | |
| 8.1 | DI water rinse | sink8, sink6 | | No pirahna clean |
| 8.2 | Selective SiGe raised S/D dep. (600A) | tystar19 | $340^0C$, 600mtorr, $GeH_4$Hi=100sccm, 10min | Recipe: SELDEPC.019 |
| 8.3 | Inspection | optiphot | Look for white SiGe deposition | |
| 8.4 | LTO cap dep. (200A) | tystar11 | $450^0C$, 1s | Recipe:11SULTOA |
| 8.5 | Inspection | nanoduv | Thickness measurment | |
| 8.6 | S/D doping | Implanter | 30keV, 5E15$cm^{-2}$ | Core Systems Inc. |

| 8.7 | Activation anneal | heatpulse3 (silicides) | $650^0C$, 30s, $N_2$ ambient | |
|---|---|---|---|---|
| **Contact Formation** | | | | |
| 9.1 | Photoresist (PR) coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
| 9.2 | PR Expose | gcaws | Gate, F: 254, E: 4.4s | |
| 9.3 | PR Develop | svgdev | Bake $120^0C$, 60s; Develop 60 s, OPD4262 developer | Oven #1, Develop#1 |
| 9.4 | Inspection | optiphot | | Use yellow filter |
| 9.5 | Contact hole etch | sink7 | 100:1 HF, ~40s | Timed etch |
| 9.6 | Inspection | asiq | Step height measurement | |

# Appendix B: Useful Codes for Atomistic Simulations

Code 1: runarray.m: Main code to be executed on MATLAB command prompt

Note: Region of interest = Region of silicon where atoms need to be simulated

```
clear all; close all;
tic;
global lg finh finw xmin xmax ymin ymax zmin zmax
finw = 5.5e-7; % device width of interest
finh = 3*finw; % device height of interest
lg = 9e-7+(2*6e-7); % device length of interest (not necessarily LG)
xmin=-lg/2; xmax=lg/2; ymin=0; ymax=finh; zmin=-finw/2; zmax=finw/2;
% read in the grid and extract unique positions
coordinates = load('tsi5p5nodes.txt');
xc = unique(coordinates(:,1));
yc = unique(coordinates(:,2));
zc = unique(coordinates(:,3));
tol = 1e-12;
ix = find(xc>=(xmin-tol) & xc<=(xmax+tol));
iy = find(yc>=(ymin-tol) & yc<=(ymax+tol));
iz = find(zc>=(zmin-tol) & zc<=(zmax+tol));
% simulation grid coordinates within the region of interest
xcin = [xc(ix(1)-1); xc(ix); xc(ix(end)+1)];
ycin = yc(iy);
zcin = zc(iz);
% Compute "space" occupied by each grid point
dxcin = getd(xcin);
dycin = getd(ycin);
dzcin = getd(zcin);
% Pick lines from the grid file, that correspond to the region of interest
iregion = find( coordinates(:,1) >= xmin-tol & coordinates(:,1) <= xmax+tol & ...
                coordinates(:,2) >= ymin-tol & coordinates(:,2) <= ymax+tol & ...
                coordinates(:,3) >= zmin-tol & coordinates(:,3) <= zmax+tol);
% This will form [x, y, z, 0]
```

USER INPUT (lines finw, finh, lg)

USER INPUT (line coordinates = load(...))

```
regioncoords = [coordinates(iregion, :), zeros(length(iregion), 1)];
% This forms an Nx4 array which is of the form: [x, y, z, dv]
% dv is the volume occupied by the corresponding grid point in the simulation grid
for i=1:size(regioncoords,1)
    xi = find(xcin == regioncoords(i,1));
    yi = find(ycin == regioncoords(i,2));
    zi = find(zcin == regioncoords(i,3));
    dv = dxcin(xi)*dycin(yi)*dzcin(zi);
    regioncoords(i,4) = dv;
end
% Run the atomization file as many times as needed (typically 100 times).
N = 100;                                          <-- USER INPUT
for filecount = 1:N
    Silattice1;
end
toc;
```

Code 2: Sililattice.m: (i) Generate and places random dopant atoms and (ii) Generates output file for Taurus-device numeric profile based on Sano Profile definition of atom.

```
% Program to randomly place dopant atoms in a silicon pseudo-lattice
% The pseudo lattice has a uniform 'grid' structure and not a 'diamond cubic' structure.
% This program can be used independently or with 'runarray.m'.
global lg finh finw xmin xmax ymin ymax zmin zmax
% profile information: peak concentration, drain doping, channel doping
pdope = 1e20; % S/D peak doping
sigma = 3e-7; % std dev of S/D gradient profile          <-- USER INPUT
mean1 = -lg/2; % peak location of source profile
mean2 = lg/2; % peak location of drain profile
sirho = 5e22;
chvol = finw*finh*lg;
fprintf('\n\nRunning atomistic placement of dopant atoms... \n\n');
fprintf('Creates ASCII dopant files for the Taurus-Device Numerics statement\n\n');
% Calculation of number of lattice atoms, atomic spacing and atomic volume
nsi = sirho*chvol; atvol = 1/sirho; atsp = atvol^(1/3);
x = (xmin+atsp/2:atsp:xmax); nx = size(x,2);
y = (ymin+atsp/2:atsp:ymax); ny = size(y,2);
```

```
z = (zmin+atsp/2:atsp:zmax); nz = size(z,2);
n = nx*ny*nz;
fprintf('Total number of lattice atoms considered = %d\n', n);
fprintf('Error in calculating the number of atoms = %f%% \n', 100*(nsi-n)/nsi);
% Create the nprofile and pprofile initial files
nnodes = size(coordinates, 1);
initdope = zeros(nnodes,1);
initregiondope = zeros(size(regioncoords,1),1);
finalndope = [coordinates initdope];
finalpdope = finalndope;
xmin=-lg/2; xmax=lg/2; ymin=0; ymax=finh; zmin=-finw/2; zmax=finw/2;
boundary = [xmin xmax ymin ymax zmin zmax];
outnodes =
find((coordinates(:,1)<xmin)|(coordinates(:,1)>xmax)|(coordinates(:,2)<ymin)|(coordinates(:,2)>ymax)
|(coordinates(:,3)<zmin)|(coordinates(:,3)>zmax));

% RANDOM GENERATION OF DOPANT LOCATIONS ON SILICON LATTICE
[X,Y,Z] = ndgrid(x,y,z);
PROBS = atvol*pdope*(exp((-(X-mean1).^2)/(2*sigma^2)));
PROBD = atvol*(pdope*(exp((-(X-mean2).^2)/(2*sigma^2))));
APROB = abs(PROBS + PROBD);
RPROB = rand([nx, ny, nz]);
INDEX = findn(APROB > RPROB);
natoms = size(INDEX,1)
countn=0;
countp=0;

% EXPRESSING EACH ATOM AS A SANO PROFILE AND NORMALIZING THE DOSE
for i=1:natoms
   dop =
(PROBD(INDEX(i,1),INDEX(i,2),INDEX(i,3))+PROBS(INDEX(i,1),INDEX(i,2),INDEX(i,3)))/atvol
   if dop>=0
      countn = countn+1;
      posnx(countn) = X(INDEX(i,1),INDEX(i,2),INDEX(i,3));
      posny(countn) = Y(INDEX(i,1),INDEX(i,2),INDEX(i,3));
      posnz(countn) = Z(INDEX(i,1),INDEX(i,2),INDEX(i,3));
      r=[posnx(countn) posny(countn) posnz(countn)];
```

USER INPUT: Change depending on reqd. doping profile

USER INPUT: Currently assumes only ntype dopants. Include else statement if p-type also exists.

```
        r2 = (regioncoords(:,1)-r(1)).^2 + (regioncoords(:,2)-r(2)).^2 + (regioncoords(:,3)-r(3)).^2;
                % Sano profile generation routine is called here.
        D = doping(r2, dop);
        ipnode = find(D < 0);
                ndopants = D;
                ndopants(ipnode) = 0;
                pdopants = zeros(length(D),1);
                pdopants(ipnode) = -D(ipnode);
        % Dose normalization
                dose = sum(ndopants.*regioncoords(:,4)) - sum(pdopants.*regioncoords(:,4));
        fprintf('Total dose = %f doping = %e\n', dose,dop);
        finalndope(iregion,4) = finalndope(iregion,4) + ndopants/dose;
        finalpdope(iregion,4) = finalpdope(iregion,4) + pdopants/dose;
    end
end

% Final matrices of interest. They are of the form [x,y,z,ntype] or
% [x,y,z,ptype]. Two different output files stored for two different types
% of dopants.
finalndope=[10000*finalndope(:,1:3) finalndope(:,4)];
save(sprintf('nprofile%d.txt', filecount),'finalndope','-ASCII');     <- OUTPUT FILES
finalpdope=[10000*finalpdope(:,1:3) finalpdope(:,4)];
save(sprintf('pprofile%d.txt', filecount),'finalpdope','-ASCII');     <- OUTPUT FILES
fprintf('countn : %d\n', countn);
if countn ~=0
    posn = 10000*[posnx' posny' posnz'];
    save(sprintf('posn30nm%d',filecount),'posn','-ASCII');
end
clear nnodes; clear initdope; clear finalndope; clear finalpdope; clear posnx; clear posny; clear posnz;
clear posn; clear r; clear r2; clear outnodes; clear pnode; clear ndopes; clear pdopes;
if countp~=0
    clear pnode; clear pospx; clear pospy; clear pospz; clear posp;
end
```

Code 3: doping.m: Sano profile definition

```
function B=doping(r2,dopeval)
q=1.6e-19;
```

```
eps=11.8; eps0=8.85e-14; kT=0.026;
r=sqrt(r2);
kc = sqrt((dopeval*(q))/(eps*eps0*kT));
% PUTTING THE SANO PROFILE
% n-type dopant
if dopeval>0
   % Calculating Sano profile
   B=(kc^3/(2*pi^2))*((sin(kc*r)-((kc*r).*cos(kc*r)))./(kc*r).^3);
   i=find(r==0);
   B(i)=kc^3/(6*pi^2);
end
```

Codes 4: Other MATLAB functions, not requiring user intervention

getd.m: Compute distance occupied by a grid point in a given direction

```
function [dx] = getd(x)
% function to compute the 'distance/region' occupied by each grid point in a
% particular direction in the simulation grid
dx(1) = (x(2)-x(1))/2;
for i=2:(length(x)-1)
   dx(i)=(x(i+1)-x(i-1))/2;
end
dx(end+1) = (x(end)-x(end-1))/2;
```

findn.m: Additional MATLAB routine not included in usual packages

```
function ind=findn(arr);
%FINDN   Find indices of nonzero elements.
%   I = FINDN(X) returns the indices of the vector X that are
%   non-zero. For example, I = FINDN(A>100), returns the indices
%   of A where A is greater than 100. See RELOP.
%   This is the same as find but works for N-D matrices using
%   ind2sub function
%   It does not return the vectors as the third output arguement
%   as in FIND
%   The returned I has the indices (in actual dimensions)
%   x(:,:,1)           x(:,:,2)           x(:,:,3)
%      = [ 1 2 3          =[11 12 13         =[21 22 23
%          4 5 6             14 15 16           24 25 26
```

```
%           7 8 9]            17 18 19]         27 28 29]
%   I=find(x==25) will return 23
%   but findn(x==25) will return 2,2,3
%   Also see find, ind2sub
%   Loren Shure, Mathworks Inc. improved speed on previous version of findn
%   by Suresh Joel Mar 3, 2003

in=find(arr);
sz=size(arr);
if isempty(in), ind=[]; return; end;
[out{1:ndims(arr)}] = ind2sub(sz,in);
ind = cell2mat(out);
```

cell2mat.m: Additional MATLAB routine not included in usual packages

```
function m = cell2mat(c)
% Error out if there is no input argument
if nargin==0
   error('MATLAB:cell2mat:NoInputs',['No input argument specified. ' ...
      'There should be exactly one input argument.'])
end
% short circuit for simplest case
elements = numel(c);
if elements == 0
   m = [];
   return
end
if elements == 1
   if isnumeric(c{1}) | ischar(c{1}) | islogical(c{1})
      m = c{1};
      return
   end
end
% Error out if cell array contains any cell arrays or objects
ciscell = cellfun('isclass',c,'cell');
if isempty(c)
   cisobj = 0;
else
```

```
    cisobj = isobject(c{1});
end
if cisobj | any(ciscell(:))
    error('MATLAB:cell2mat:UnsupportedCellContent',['Cannot support cell ' ...
        'arrays containing cell arrays or objects.']);
end
% If the cell array has one element we can finish
%   up the conversion quickly
if elements == 1
    m = c{1};
    return
end
csize = size(c);
% Error out if cell array contains mixed data types
cellclass = class(c{1});
ciscellclass = cellfun('isclass',c,cellclass);
if ~all(ciscellclass(:))
    error('MATLAB:cell2mat:MixedDataTypes', ...
        'All contents of the input cell array must be of the same data type.');
end
% If cell array of structures, make sure the field names are all the same
if isstruct(c{1})
    cfields = cell(elements,1);
    for n=1:elements
        cfields{n} = fieldnames(c{n});
    end
    % Perform the actual field name equality test
    if ~isequal(cfields{:})
        error('MATLAB:cell2mat:InconsistentFieldNames', ...
            ['The field names of each cell array element must be consistent ' ...
            'and in consistent order.'])
    end
end
% If cell array is 2-D, execute 2-D code for speed efficiency
if ndims(c) == 2
    rows = size(c,1);
    m = cell(rows,1);
```

```
    % Concatenate each row first
    for n=1:rows
        m{n} = cat(2,c{n,:});
    end
    % Now concatenate the single column of cells into a matrix
    m = cat(1,m{:});
    return
end
% Treat 3+ dimension arrays
% Construct the matrix by concatenating each dimension of the cell array into
%   a temporary cell array, CT
% The exterior loop iterates one time less than the number of dimensions,
%   and the final dimension (dimension 1) concatenation occurs after the loops
% Loop through the cell array dimensions in reverse order to perform the
%   sequential concatenations
for cdim=(length(csize)-1):-1:1
    % Pre-calculated outside the next loop for efficiency
    ct = cell([csize(1:cdim) 1]);
    cts = size(ct);
    ctsl = length(cts);
    mref = {};
    % Concatenate the dimension, (CDIM+1), at each element in the temporary cell array, CT
    for mind=1:prod(cts)
        [mref{1:ctsl}] = ind2sub(cts,mind);
        % Treat a size [N 1] array as size [N], since this is how the indices
        %   are found to calculate CT
        if ctsl==2 & cts(2)==1
            mref = {mref{1}};
        end
        % Perform the concatenation along the (CDIM+1) dimension
        ct{mref{:}} = cat(cdim+1,c{mref{:},:});
    end
    % Replace M with the new temporarily concatenated cell array, CT
    c = ct;
end
% Finally, concatenate the final rows of cells into a matrix
m = cat(1,c{:});
```

# Appendix C: WireFET Process Steps

| Step | Process Step | Equipment | Process Specification | Comments |
|---|---|---|---|---|
| | | | **Backgate formation** | |
| 1.1 | Wafer preclean | sink6 | Pirahna $120^0C$, 10min + HF dip | |
| 1.2 | $POCl_3$ process | tystar 13 | Deposition:$1000^0C$ 60min, Drive in: $1000^0C$ 60min | Recipe:13POCL3A |
| 1.3 | Oxide strip | sink 6 | 25:1 HF or 10:1 HF | check for dewetting |
| 1.4 | $R_S$ measurement | 4ptprb | Expect ~ 2ohms/sq | |
| | | | **Isolation** | |
| 2.1 | HTO deposition (50A) | tystar9 | $800^0C$, 300mtorr, $SiH_2Cl_2$ =18sccm, $N_2O$=180sccm,16min | Recipe:9VHTOA |
| 2.2 | Inspection | sopra | Thickness measurement | |
| 2.3 | PSG dep.(600A) | tystar11 | $450^0C$, 2min | Recipe:11SDLTOA |
| 2.4 | Inspection | nanoduv | Thickness measurement | Recipe: Ox/Si |
| 2.5 | HMDS | primeoven | $90^0C$, 1min | |
| 2.6 | PR coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
| 2.7 | PR Expose | gcaws2 | Inverse gate mask | |
| 2.8 | PR Develop | svgdev | Bake $120^0C$, 60s; Develop 60 s, OPD4262 developer | Oven #1, Develop#1 |
| 2.9 | Inspection | optiphot | | Use yellow filter |
| 2.10 | Hard bake | uvbake | $200^0C$, 60s | |
| 2.11 | Inspection | leo | CD measurement | Leo SEM |
| 2.12 | PR ashing | technics-c | 30W, $O_2$ plasma, ~7min | Expand channel area |
| 2.13 | Inspection | leo | CD measurement | Repeat 2.7 if reqd |
| 2.14 | Oxide etching | sink7 | 100:1 BHF, 40s timed etch | RCA clean bath |
| 2.15 | Strip photoresist | matrix | 400W, $200^0C$, 1.5min, $O_2$ plasma | |
| 2.16 | Inspection | leo | Final gap measurment | |

| 2.17 | Wafer postclean | sink8 | Pirahna $120^0C$, 10min | |
|---|---|---|---|---|
| **PSG flow and Gate dielectric formation** | | | | |
| 3.1 | Wafer preclean | sink6 | Pirahna $120^0C$, 10min | |
| 3.2 | PSG reflow | tystar2 | $950^0C$, 30min | Recipe:2N2ANNLA |
| 3.3 | Gate oxidation | tystar1 | Oxidation: $750^0C$, 1min<br>Densification: $950^0$C, 30min | Recipe: 1THIN_OX |
| 3.4 | Inspection | sopra | Thickness measurement | |
| **Al deposition and Channel definition** | | | | |
| 4.1 | Al deposition (200A) | novellus | 3kW suppy, 5% power, 4mtorr, $25^0C$ | Recipe:ALRT300W |
| 4.2 | $R_S$ measurement | 4ptprb | | |
| 4.3 | TEOS cap dep. (500A) | p5000 | $300^0C$, 8s | Recipe:A-PE-US6300C |
| 4.4 | Inspection | nanoduv | | Recipe: Ox/Si |
| 4.5 | HMDS | primeoven | $90^0C$, 1min | |
| 4.6 | PR coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
| 4.7 | PR Expose | gcaws2 | Inverse gate mask | |
| 4.8 | PR Develop | svgdev | Bake $120^0C$, 60s; Develop 60 s, OPD4262 developer | Oven #1, Develop#1 |
| 4.9 | Inspection | optiphot | | Use yellow filter |
| 4.10 | Hard bake | uvbake | $200^0C$, 60s | |
| 4.11 | Inspection | leo | CD measurement | |
| 4.12 | TEOS etching | lam2 | $RF_{top}$=850W, Gap=0.38cm, $CHF_3$=30sccm, $CF_4$=90sccm 15s | Recipe:SIO2MON |
| 4.13 | resist strip | matrix | 400W, $200^0C$, 1.5min, $O_2$ plasma | |
| 4.14 | Inspection | leo | | |
| 4.15 | Post strip clean | sink5 | SVC-14, $80^0C$, 15min | metal cleaning |
| **Si deposition and implantation** | | | | |
| 5.1 | Al surface etchback (40A) | novellus | $75^0C$, 1kW supply, 25% power, 1min, self bias -200V | Precise control difficult |
| 5.2 | Interface | novellus | Degas station, pure $O_2$ gas, | |

|  |  |  |  |  |
|---|---|---|---|---|
|  | oxidation |  | 2mtorr, 2min, $25^0C$ |  |
| 5.3 | a-Si deposition (1.5*240A ) | novellus | $25^0C$, 3kW supply, 30% power, with PCT, 80s | No break between 5.2 and 5.3 |
| 5.4 | Inspection | nanospec | Thickness measurement | Recipe aSi/Si |
| 5.5 | PR coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
| 5.6 | Half wafer PR Expose | gcaws2 | clear glass mask, high exposure |  |
| 5.7 | PR Develop | svgdev | Bake $120^0C$, 60s; Develop 60 s, OPD4262 developer | Oven #1, Develop#1 |
| 5.8 | Inspection | optiphot |  | Use yellow filter |
| 5.9 | Passivation species implant | Implanter | Se implant (Rp=100A, dose=4E13/$cm^2$) | Innovion Inc. |
| 5.10 | PR strip | matrix | 400W, $200^0C$, 1.5min, $O_2$ plasma |  |
| 5.11 | Post strip clean | sink5 | SVC-14, $80^0C$, 15min | metal cleaning |
| **Channel Aluminum Induced Crystallization** |  |  |  |  |
| 6.1 | AIC anneal | tystar18 | $400^0C$, 90min, $N_2$ amobient | Just boat-in/out |
| 6.2 | Inspection | reichert |  |  |
| 6.3 | Al wet etching | sink8 | Premixed Al etchant, $50^0C$ | timed etch |
| 6.4 | Inspection | reichert |  |  |
| **Active area patterning** |  |  |  |  |
| 7.1 | PR coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
| 7.2 | PR Expose | gcaws2 | Active area mask |  |
| 7.3 | PR Develop | svgdev | Bake $120^0C$, 60s; Develop 60 s, OPD4262 developer | Oven #1, Develop#1 |
| 7.4 | Inspection | optiphot |  | Use yellow filter |
| 7.5 | Hard bake | uvbake | $200^0C$, 60s |  |
| 7.6 | etch silicon | lam3 | Airlock chamber, 1torr, $RF_{top}$=400W, $CF_4$=90sccm, $O_2$=10sccm | Calib. timed etch |
| 7.7 | Inspection | reichert |  |  |
| 7.8 | etch oxide | lam2 | $RF_{top}$=850W, Gap=0.38cm, $CHF_3$=30sccm, $CF_4$=90sccm | calib. timed etch |

| | | | | |
|---|---|---|---|---|
| 7.9 | Inspection | reichert | | |
| 7.10 | PR strip | matrix | 400W, $200^0C$, 1.5min, $O_2$ plasma | matrix |
| 7.11 | Inspection | reichert | | |
| **Contact formation** | | | | |
| 8.1 | TEOS cap dep. (1800A) | p5000 | $300^0C$, 30s | Recipe:A-PE-US6300C |
| 8.2 | PR coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
| 8.3 | PR Expose | gcaws2 | Contact area mask | |
| 8.4 | PR Develop | svgdev | Bake $120^0C$, 60s; Develop 60 s, OPD4262 developer | Oven #1, Develop#1 |
| 8.5 | Inspection | optiphot | | Use yellow filter |
| 8.6 | Hard bake | uvbake | $200^0C$, 60s | |
| 8.7 | TEOS + a-Si etching | lam2 | $RF_{top}$=850W, Gap=0.38cm, $CHF_3$=30sccm, $CF_4$=90sccm, 1min | all oxide and thin layer of s-Si can be etched |
| 8.8 | PR Strip | matrix | 400W, $200^0C$, 1.5min, $O_2$ plasma | |
| 8.9 | Inspection | reichert | | |
| **Metallization** | | | | |
| 9.1 | Al deposition, 1000A | novellus | 12kW supply, 10% power, 4mtorr, $25^0C$, 1min | novellus |
| 9.2 | PR coat | svgcoat1 | 1.1micron I-line resist | Coater #1, Oven #1 |
| 9.3 | PR Expose | gcaws2 | Metallization mask | |
| 9.4 | PR Develop | svgdev | Bake $120^0C$, 60s; Develop 60 s, OPD4262 developer | Oven #1, Develop#1 |
| 9.5 | Inspection | optiphot | | Use yellow filter |
| 9.6 | Hard bake | uvbake | $200^0C$, 60s | |
| 9.7 | Metal etching | sink8 | Premixed Al etchant, $50^0C$ | timed etch |
| 9.8 | PR Strip | matrix | 400W, $200^0C$, 1.5min, $O_2$ plasma | |
| 9.9 | Post strip clean | sink5 | SVC-14, $80^0C$, 15min | |
| 9.10 | FGA | tystar18 | $350^oC$, 30minutes, $H_2/N_2$ ambient | |