Title
Nanoscale Bulk MOSFET Design and Process Technology for Reduced Variability

Author
Sun, Xin

Publication Date
2010

Peer reviewed|Thesis/dissertation

Nanoscale Bulk MOSFET Design and Process Technology for Reduced Variability

By

Xin Sun

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 David Brillinger

Spring 2010

Abstract

Nanoscale Bulk MOSFET Design and Process Technology for Reduced Variability

by

Xin Sun

Doctor of Philosophy in Engineering – Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Tsu-Jae King Liu, Chair

Historically, the steady miniaturization of the conventional (planar bulk) MOSFET by simply scaling the device dimensions with minimal changes to the conventional transistor design and CMOS process flow has been effective to provide for continual improvements in integrated circuit performance and cost per function with every technology node. However, transistor scaling has become increasingly difficult in the sub-100 nm regime. Increased leakage current and variability in transistor performance are the major challenges for continued scaling of bulk-Si CMOS technology.

The benefit of using a spacer gate lithography process to mitigate the effect of gate line edge roughness (LER) is assessed using statistical 3-D device simulations. The simulation results indicate that spacer gate lithography is a scalable technology which can dramatically reduce LER-induced variation in transistor performance.

A tri-gate bulk MOSFET design combining retrograde channel doping with a multi-gate structure is proposed to provide an evolutionary pathway for bulk CMOS scaling. The scalability, design optimization, and the effect of systematic and random variations on transistor performance are investigated. As compared with the classic planar MOSFET design, the tri-gate bulk MOSFET provides for superior electrostatic integrity and reduced variability. As compared with SOI FinFET design, the tri-gate bulk MOSFET design is more scalable and less sensitive to device design parameters. As compared with the bulk FinFET design, the tri-gate bulk MOSFET offers comparable performance and variability. Its low-aspect-ratio channel structure is favorable for ease of manufacturing. Thus, the tri-gate bulk MOSFET is a promising structure for CMOS scaling to the end of the technology roadmap.

The fabrication process flow and the most critical processes for tri-gate bulk MOSFET fabrication are discussed. Initial device results show that tri-gate bulk MOSFET design is beneficial for reduced variability.

To my family,
for their unbounded love and support

Table of Contents

Acknowledgements

Pursuing a Ph.D. is a long-term journey, with both the joys and frustrations. It is fun, it may get hard, and then it gets fun again. I am grateful for all the encouragement and help I get during my time at Berkeley.

First and foremost, I would like to express my sincere gratitude to my advisor, Professor Tsu-Jae King Liu, for her extraordinary support since the first day I came to Berkeley and for her encouragement and guidance throughout the research.

I would also like to thank Professor Chenming Hu for his enthusiasm and help, and for chairing my qualifying examination committee. I also wish to thank Professor David Brillinger and Professor Ming Wu for their valuable feedback as members of my qualifying examination.

I am grateful to Dr. Qiang Lu for his help and advice during my internship at Synopsys. I would also like to thank Dr. Ken Rim and Dr. Noah Zamdmer for their mentoring during my summer internship at IBM.

There are several people whom I would like to thank with regards to this work specifically. Dr. Andrew Carlson helped me develop the processes for the corrugated substrate. Dr. Harry Levinson and Tom Wallow at AMD provided the photoresist data for the study of gate line edge roughness. Dr. Qiang Lu and Victor Moroz helped answer my questions about TCAD simulations. Mr. Hideki Takeuchi at ATDF helped device fabrication. The UC Berkeley Microfabrication lab staff helped me debug different processes.

I am deeply grateful to the former and current members of Berkeley device research group for their encouragement and help, for being my colleague and years of friends. Thanks everyone.

This work has been sponsored by Tokyo Electron Limited and the Semiconductor Research Corporation, FLCC, IMPACT and the UC Discovery Grant program.

Chapter 1

Introduction

1.1 The Need for CMOS Scaling

Since the integrated circuit (IC) was invented [1, 2], the evolution of IC technology has enabled the rapid growth of the electronics industry for more than forty years. Technological advancements have yielded dramatic improvements in product performance as well as cost per function, making electronics products more capable and affordable for consumers. Take the personal computer (PC), for example: the first IBM PC equipped with a 4.77 MHz Intel 8088 microprocessor and 16 kilobytes of memory was sold for around $1,600 in 1981 (which would be around $4,000 today); nowadays a typical laptop PC with a 2.66 GHz Intel Core 2 Duo microprocessor and 4 gigabytes of memory costs only about one-third of $4,000. These phenomenal improvements are ascribed to the sustained scaling of the metal-oxide-semiconductor field-effect transistor (MOSFET) [3] – the most important and fundamental building block for constructing integrated circuits.

The simplicity of the MOSFET structure and the complementary symmetry of n-type and p-type MOSFETs have contributed to the success of complementary-MOSFET (CMOS) technology [4, 5]. Over the past few decades, CMOS scaling generally followed Moore's Law [6] (Fig. 1.1 [7]). Scaling theory [8, 9] provided simple rules for transistor design to increase circuit speed and density (Table 1.1). The improved circuit performance and density enable more complicated functionality, since more transistors and functionalities can be integrated on one single chip. The circuit density improvement also provides for reduction in manufacturing cost per function: with transistor scaling, the silicon die area is smaller for a given functionality, thus more dies can be fabricated from a single silicon wafer at very little marginal cost. Smaller die size is also beneficial for

yield improvement [10]. As shown in Fig. 1.2 [11], cost per function has been exponentially decreased with every technology node.

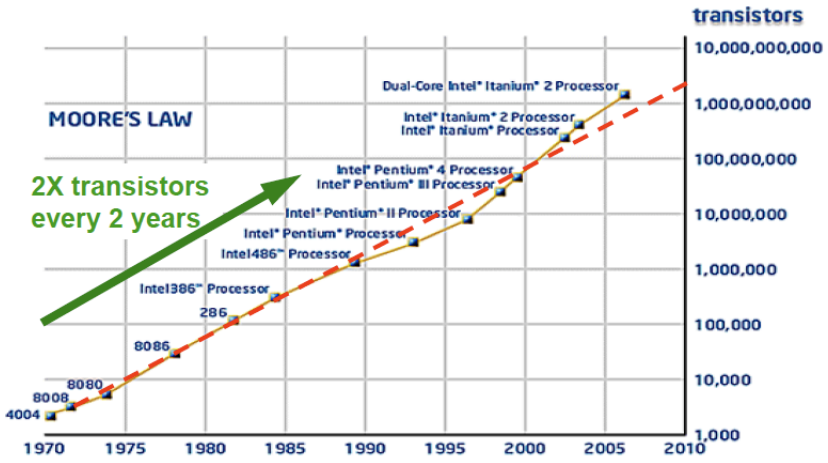


Figure 1.1. Moore's Law: from few to billions of transistors (adapted from [7]).

Device and circuit parameters	Constant-field scaling	Generalized scaling
Device dimensions (L_G, T_{OX}, X_J, W)	$1/\kappa$	$1/\kappa$
Body doping concentration (N_B)	κ	$\alpha\kappa$
Supply voltage (V_{DD})	$1/\kappa$	α/κ
Electric field (E)	1	α
Transistor current (I)	$1/\kappa$	α/κ
Area (A)	$1/\kappa^2$	$1/\kappa^2$
Capacitance ($C = \varepsilon_{OX}A/T_{OX}$)	$1/\kappa$	$1/\kappa$
Intrinsic delay ($\tau \sim CV_{DD}/I$)	$1/\kappa$	$1/\kappa$
Power dissipation ($P \sim IV_{DD}$)	$1/\kappa^2$	α^2/κ^2
Power density (P/A)	1	α^2

Table 1.1. Guidelines for MOSFET scaling (adapted from [8, 9, 12]). κ is the scaling factor for dimensions (> 1), and α is a different scaling constant, ~ 1.15 [13].

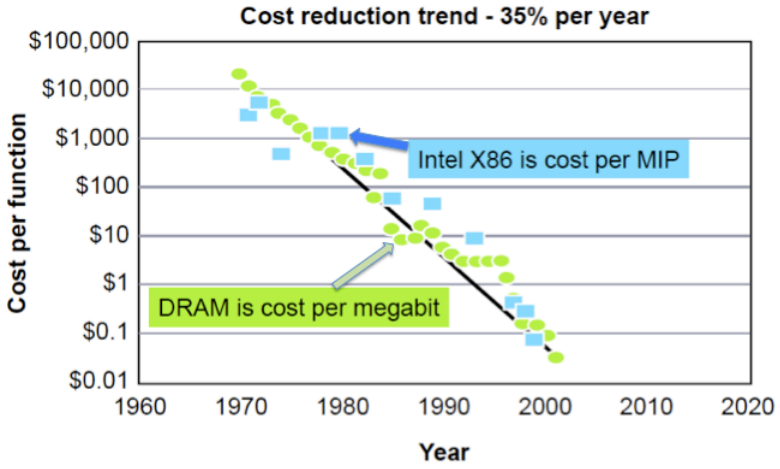


Figure 1.2. Cost per function has been exponentially decreased (adapted from [11]).

1.2 MOSFET Scaling Challenges

Historically, the steady miniaturization of the conventional (planar bulk) MOSFET by simply scaling the device dimensions with minimal change to the conventional CMOS process flow has been effective to provide for continual improvements in integrated circuit performance and cost per function with every technology node. However, transistor scaling has become increasingly difficult in the sub-100 nm regime.

Material limits impose a natural barrier to continued scaling. Carrier mobilities are degraded due to higher vertical electric fields in the MOSFET channel [14, 15]. Mobility enhancement techniques such as strained-Si [16, 17, 18] and high-mobility channel materials [19, 20] have been actively pursued. Another fundamental material limit is faced by the gate oxide. Gate tunneling leakage becomes significant below 1 nm gate oxide thickness [21]. Further scaling can be realized by replacing the oxide/oxynitride gate dielectric with a high-permittivity (high-κ) gate dielectric [22].

As transistor dimensions are reduced, parasitic resistances and capacitances both scale unfavorably with reduced pitch. These parasitic elements will diminish the performance gain by transistor scaling [23].

Furthermore, power dissipation must be carefully considered in MOSFET scaling. Nowadays, passive power (due to transistor off-state leakage) constitutes a significant portion of the total power dissipation in high-performance CMOS products. The source-to-drain OFF-state leakage current (I_{OFF}) needs to be suppressed to mitigate passive power dissipation. I_{OFF} increases exponentially with decreasing V_T/S, where V_T is the threshold voltage and S is the sub-threshold swing. Since S is fundamentally limited to be no less than 60mV/decade at room temperature due to the thermal distribution of carriers within energy bands, V_T scaling has slowed down to avoid dramatic increases in I_{OFF}. To achieve large drive current, the gate overdrive (V_{DD} - V_T) needs to be significant and therefore V_{DD} scaling also has to slow down, which results in increased active power density.

Finally, increased variability in transistor performance with each technology node poses critical challenges to future technology development, manufacturing, and design [24]. Improvements in transistor architecture and process control are needed to reduce variability. In addition, variability is detrimental to manufacturing yield, which will ultimately increase the cost per function, reducing the benefits of scaling. Fig. 1.3 [25] shows the distributions of chip frequency and leakage current of a large number of microprocessors. The variability in transistor performance causes about 20X variation in chip leakage and 30% variation in chip frequency, which substantially affects the yield [25]. Variability also affects power dissipation, since the nominal value of V_T needs to be kept high to avoid unacceptably large I_{OFF} and hence the pace of V_{DD} scaling needs to be

reduced to guarantee high performance. The impact of variability on design includes mismatch in analog/RF circuits and timing issues in digital circuits. To cope with the increased variability of scaled transistors, analog/RF circuit designers must increase the size of the critical transistors to reduce the impact of mismatch, and digital circuit designers must increase design margins [24]. These methods limit die area scaling and thus the economical benefit of scaling.

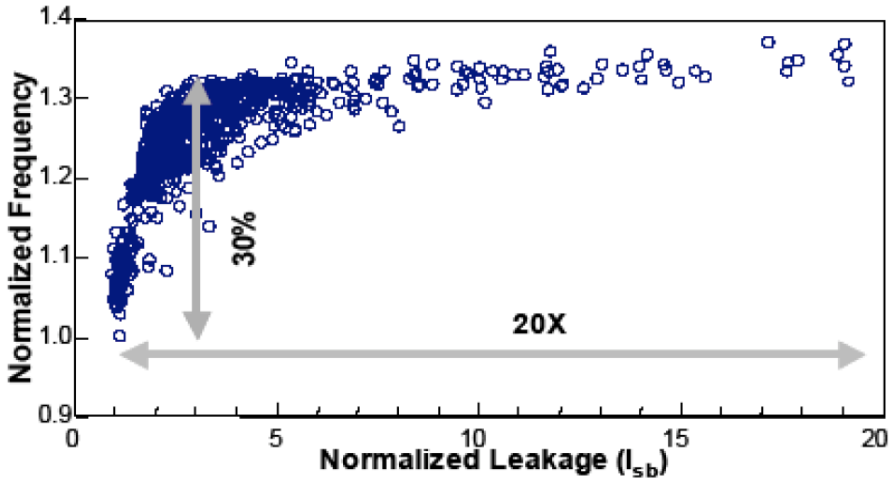


Figure 1.3. Leakage and frequency variations (adapted from [25]).

1.3 Sources of Variability

Variation in transistor characteristics can be characterized by different root causes of these types of variation. There are two main categories of variation: systematic and random variations.

Systematic variation can further be divided into inter-die systematic variation and intra-die systematic variation [26]. Inter-die systematic variation is due to normal manufacturing tolerances. One example is gate length variation due to the length of lithographic exposure. Intra-die systematic variation comes from differences in either the device layout or the device neighborhood. For example, optical proximity effects cause the gate length to vary as a function of local layout. Resolution enhancement techniques such as optical proximity correction (OPC) and phase-shift masking (PSM) have been used in manufacturing to reduce systematic intra-die variation. However, these techniques are costly and increase process sensitivity. The use of stress to boost transistor performance is another source of systematic intra-die variation, since process-induced stress is dependent on layout.

Examples of random variation are photoresist line edge roughness (LER) (Fig. 1.4) and random dopant fluctuations. These variations are caused by fundamental phenomena associated with today's CMOS manufacturing processes. LER- and RDF-induced variations increase significantly with continued device scaling.

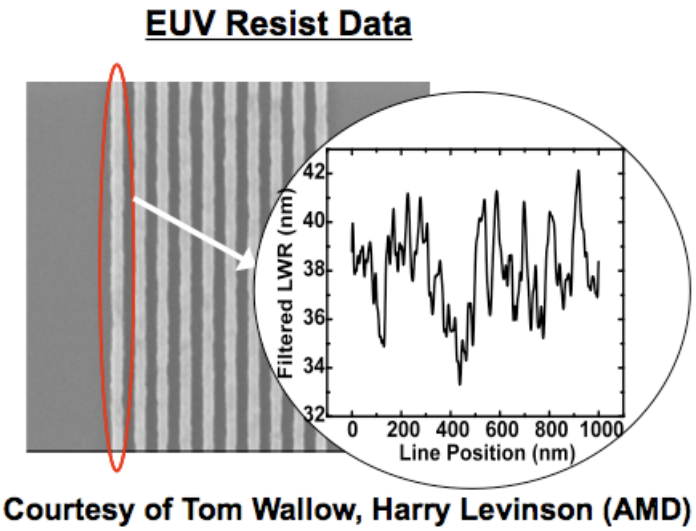


Figure 1.4. Plan-view scanning electron micrograph (SEM) of photoresist resist lines.

1.4 Advanced MOSFET Structures

To address the scaling challenges of planar bulk CMOS technology, alternative MOSFET structures based on silicon-on-insulator (SOI) technology have been developed. Fig. 1.5 [27] illustrates three SOI MOSFET structures.

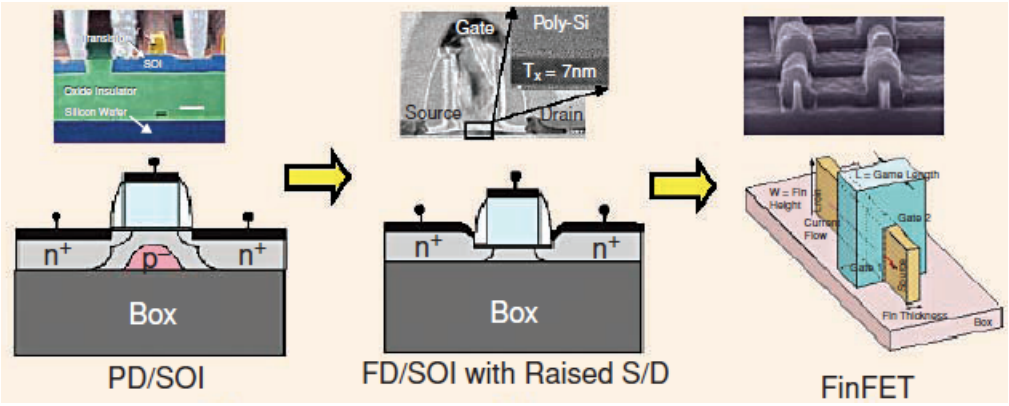


Figure 1.5. Advanced MOSFET structures: from PD SOI [28] to FD SOI with raised source/drain [29] to FinFET [30] (adapted from [27]).

Partially depleted (PD) SOI was the first SOI technology introduced for high-performance microprocessor applications [28, 31, 32]. The buried oxide (BOX) layer serves to provide for reduced source/drain junction capacitance. PD/SOI MOSFETs offer a good trade-off between power and performance.

Fully depleted (FD) SOI MOSFETs achieve superior control of short channel effects (SCE) and therefore are beneficial for low-power technologies which have stringent

leakage specifications. FD/SOI MOSFETs rely on a thin body to suppress leakage, which presents news challenge for transistor fabrication, mainly the tight control of the physical body thickness and doping profiles. Series resistance in the source and drain extension regions is another issue. An elevated source/drain structure is needed to mitigate this parasitic resistance, but this structure results in increased gate-to-drain capacitance (the Miller capacitance) and thus degrades circuit performance.

A thin-body structure with front- and back-gate electrodes allows the body thinness requirement of the FD/SOI MOSFET to be relaxed. This makes the back-gated thin-body MOSFET (e.g. an FD/SOI MOSFET with ultra-thin BOX) a good candidate for CMOS scaling.

The double-gate FinFET structure [30] can provide for improved electrostatic integrity over planar thin-body MOSFET structures. However, it requires the formation of high-aspect-ratio "fin" structures (of narrow width to suppress SCE and tall height to achieve high current per unit area) which pose a significant challenge for process control and design flexibility. By making the fin wider and much shorter, a tri-gate SOI MOSFET structure is formed, which is advantageous for manufacturability but which has poorer layout area efficiency (*i.e.* less current per unit area).

In summary, advanced MOFET structures have been developed to enable continued transistor scaling beyond the limit of the planar bulk MOSFET, but they also pose new challenges for integrated circuit manufacturing.

1.5 This Work

This dissertation discusses MOSFET design and process technology for reduced variability. A quasi-planar tri-gate bulk MOSFET structure is proposed to provide an evolutionary pathway for CMOS scaling to the end of the roadmap. Through detailed modeling and simulation, scaling issues, device performance and variability of nanoscale MOSFETs are evaluated. The fabrication process for tri-gate bulk MOSFETs is also discussed.

In Chapter 2, the effect of gate line edge roughness (LER) on bulk-Si MOSFET performance is studied using statistical 3-D device simulations. The benefit of using a spacer gate lithography process to mitigate the effect of LER is assessed, with consideration of source/drain placement and spacer width variation.

In Chapter 3, a quasi-planar tri-gate bulk MOSFET design utilizing a low-aspect-ratio channel region is proposed to provide an evolutionary pathway for continued bulk CMOS scaling. The scalability and design optimization of tri-gate bulk MOSFET are investigated.

In Chapter 4, a variability study of the tri-gate bulk MOSFET is performed. The simulation results show that an optimized tri-gate bulk MOFET design offers less systematic variation as compared to planar bulk MOSFET and SOI FinFET designs, and comparable random variation as the SOI FinFET.

In Chapter 5, the fabrication process flow and the most critical processes for tri-gate bulk MOSFET fabrication are discussed. Initial device results show that tri-gate bulk MOSFET design is beneficial for reduced variability.

Chapter 6 summarizes the contributions of this work and offers directions for further research.

1.6 References

[1] R. N. Noyce, "Semiconductor device and lead structure," U.S. Patent #2,981,877, issued in April 1961.

[2] J. S. Kilby, "Miniaturized electronic circuits," U.S. Patent #3,138,743, issued in June 1964.

[3] D. Kahng, "Electric field controlled semiconductor device," U.S. Patent #3,102,230, issued in August 1963.

[4] F. M. Wanlass and C. T. Sah, "Nanowatt logic using field-effect metal-oxide semiconductor triodes," *International Solid State Circuits Conference*, *Digest of Technical Papers*, pp. 32-33, 1963.

[5] F. M. Wanlass, "Low stand-by power complementary field effect circuitry," U.S. Patent #3,356,858, issued in December 1967.

[6] G. E. Moore, "Progress in digital integrated electronics," *IEDM Tech. Dig.*, pp. 11-13, 1975.

[7] T. Ghani, "Challenges and innovations in nano-CMOS transistor scaling," *Nikkei Presentation*, October 2009.

[8] R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideout, E. Bassous, and A. LeBlanc, "Design of ion-implanted MOSFETs with very small physical dimensions," *IEEE J. Solid-State Circuits*, vol. 9, pp. 256-268, October 1974.

[9] G. Baccarani, M. R. Wordeman, and R. H. Dennard, "Generalized scaling theory and its application to a 1/4 micrometer MOSFET design," *IEEE Trans. Electron Devices*, vol. 31, pp. 452-462, April 1984.

[10] C. N. Berglund, "A unified yield model incorporating both defect and parametric effects," *IEEE Trans. Semiconductor Manufacturing*, vol. 9, No. 3, pp. 447-454, August 1996.

[11] S. W. Jones, "Exponential trends in the integrated circuits industry", *ICKnowledge*, 2004.

[12] T.-J. King Liu, and L. Chang, "Transistor scaling to the limit," *Moore's Law: Beyond Planar Silicon CMOS and into the Nano Era*, Springer-Verlag, 2008.

[13] E. J. Nowak, "Maintaining the benefits of CMOS scaling when scaling bogs down," *IBM J. Research and Development*, vol. 46, pp. 169-180, 2002.

[14] A. Lochtefeld and D. A. Antoniadis, "On experimental determination of carrier velocity in deeply scaled NMOS: How close to the thermal limit?," *IEEE Electron Device Lett.*, vol. 22, pp. 95-97, February 2001.

[15] M. V. Fischetti and S. E. Laux, "Long-range Coulomb interactions in small Si devices. Part I: Performance and reliability," *J. Appl. Phys.*, vol. 89, pp. 1205-1231, January 2001.

[16] J. Welser, J. L. Hoyt, and J. F. Gibbons, "Electron mobility enhancement in strained-Si n-type metal-oxide-semiconductor field-effect transistors," *IEEE Electron Device Lett.*, vol. 15, pp. 100-102, March 1994.

[17] T. Mizuno, S. Takagi, N. Sugiyama, H. Satake, A. Kurobe, and A. Toriumi, "Electron and hole mobility enhancement in strained-Si MOSFET's on SiGe-on-insulator substrates fabricated by SIMOX technology," *IEEE Electron Device Lett.*, vol. 21, pp, 230-232, May 2000.

[18] M. Saitoh, N. Yasutake, Y. Nakabayashi, K. Uchida, and T. Numata, "Understanding of strain effects on high-field carrier velocity in (100) and (110) CMOSFETs under quasi-ballistic transport", *IEDM Tech. Dig.*, pp. 469-472, 2009.

[19] D. Kuzum, T. Krishnamohan, A. Nainani, Y. Sun, P. A. Pianetta, H. S-. P. Wong, and K. C. Saraswat, "Experimental demonstration of high mobility Ge NMOS," *IEDM Tech. Dig.*, pp. 453-456, 2009.

[20] T. Ashley, M. T. Emeny, D. G. Hayes, K. P. Hilton, R. Jefferies, J. O. Maclean, S. J. Smith, A. W-H. Tang, D. J. Wallis, and P. J. Webber, "High-performance InSb based quantum well field effect transistors for low-power dissipation applications," *IEDM Tech. Dig.*, pp. 849-852, 2009.

[21] T. Ghani, K. Mistry, P. Packan, S. Thompson, M. Stettler, S. Tyagi, and M. Bohr, "Scaling challenges and device design requirements for high performance sub-50nm gate length planar CMOS transistors," *Symp. VLSI Tech. Dig.*, pp. 174-175, 2000.

[22] G. D. Wilk, R. M. Wallace, and J. M. Anthony, "High-κ gate dielectrics: Current status and materials properties considerations," *J. Appl. Phys.*, vol. 89, pp. 5243-5275, May 2001.

[23] J.W. Sleight, I. Lauer, O. Dokumaci, D. M. Fried, D. Guo, B. Haran, S. Narasimha, C. Sheraw, D. Singh, M. Steigerwalt, X. Wang, P. Oldiges, D. Sadana, C.Y. Sung, W. Haensch, and M. Khare, "Challenges and opportunities for high performance 32 nm CMOS technology," *IEDM Tech. Dig.*, pp. 1-4, 2006.

[24] S. Saxena, C. Hess, H. Karbasi, A. Rossoni, S. Tonello, P. McNamara, S. Lucherini,

S. Minehane, C. Dolainsky, and M. Quarantelli, "Variation in transistor performance and leakage in nanometer-scale technologies," *IEEE Trans. Electron Devices*, vol. 55, pp. 131-144, January 2008.

[25] S. Borkar, T. Karnik, S. Narendra, J. Tschanz, A. Keshavarzi, and V. De, "Parameter variations and impact on circuits and microarchitecture", *Proc. Conference on Design Automation*, pp. 338-342, 2003.

[26] P. S. Zuchowski, P. A. Habitz, J. D. Hayes, and J. H. Oppold, "Process and environmental variation impacts on ASIC Timing," *Proc. International Conference on Computer-Aided Design*, pp. 336-342, 2004.

[27] C.T. Chuang, K. Bernstein, R. V. Joshi, R. Puri, K. Kim, E. J. Nowak, T. Ludwig, and I. Aller, "Scaling planar silicon devices," *IEEE Circuits & Devices Magazine*, pp. 6-19, February 2004.

[28] C.T. Chuang, P.F. Lu, and C.J. Anderson, "SOI for digital CMOS VLSI: Design considerations and advances," *Proc. IEEE*, vol. 86, pp. 689-720, April 1998.

[29] B. Doris, M. Ieong, T. Kanarsky, Y. Zhang, R.A. Roy, O. Dokumaci, Z. Ren, F.F. Jamin, L. Shi, W. Natzle, H.J. Huang, J. Mezzapelle, A. Mocuta, S. Womack, M. Gribelyuk, E.C. Jones, R.J. Miller, H-S.P. Wong, and W. Haensch, "Extreme scaling with ultra-thin Si channel MOSFETs," *IEDM Tech. Dig.*, pp. 267-270, 2002.

[30] J. Kedzierski, D. M. Fried, E. J. Nowak, T. Kanarsky, J. Rankin, H. Hanafi, W. Natzle, D. Boyd, Y. Zhang, R. Roy, J. Newbury, C. Yu, Q. Yang, P. Saunders, C. Willets, A. Johnson, S. Cole, H. Young, N. Carpenter, D. Rakowski, B. A. Rainey, P. Cottrell, M. leong, and P. Wong, "High-performance symmetric-gate and CMOS-compatible V_T asymmetric-gate FinFET devices," *IEDM Tech. Dig.*, pp. 437-440, 2001.

[31] M. Canada, C. Akrout, D. Cawthron, J. Corr, S. Geissler, R. Houle, P. Kartschoke, D. Kramer, P. McCormick, N. Rohrer, G. Salem, and L. Warriner, "A 580 MHz RISC microprocessor in SOI," *ISSCC Tech. Dig.*, pp. 430-431, 1999.

[32] D.H. Allen, A.G. Aipperspach, D.T. Cox, N.V. Phan, and S.N. Storino, "A 0.2 μm 1.8 V SOI 550 MHz 64 b PowerPC microprocessor with copper interconnects," *ISSCC Tech. Dig.*, pp. 438-439, 1999.

Chapter 2

Spacer Gate Lithography for Reducing Variability

2.1 Introduction

Variability in transistor performance is one of the major challenges for continued scaling of CMOS technology. As transistor gate lengths are scaled down, gate line edge roughness (LER) is not reduced commensurately [1]. At the same time, thermal process budgets are reduced to achieve shallower and more abrupt source/drain junctions, so that the effects of gate LER become increasingly significant (Fig. 2.1). Thus, gate LER will be the dominant source of variability for channel lengths below 25nm [1]. The use of spacer lithography (also known as sidewall transfer lithography) to define fins has been experimentally demonstrated to be effective in reducing FinFET variability [2,3]. In this chapter, the effect of gate LER on bulk-Si MOSFET performance is studied using three-dimensional (3D) device simulations [4]. The use of spacer lithography to define the gate electrode is shown to be effective for reducing gate-LER-induced variability in MOSFET performance.

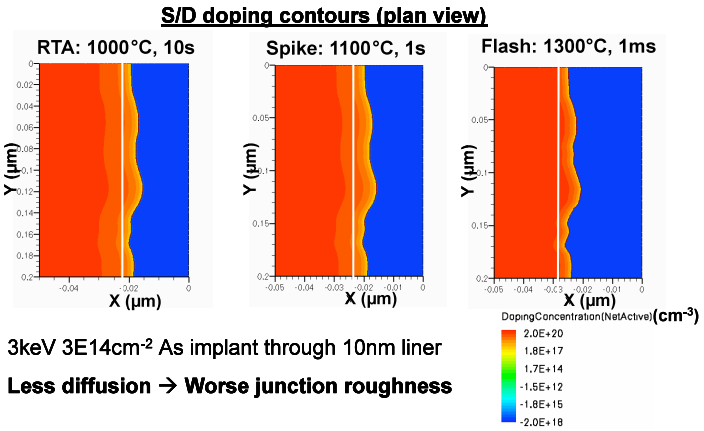


Figure 2.1. Plan view of source/drain (S/D) doping contours generated by the Sentaurus process simulator [4]. As the thermal budget for the dopant activation anneal is reduced, the effect of gate LER becomes increasingly significant.

2.2 Spacer Lithography Process

Spacer lithography is advantageous for defining fine-line features (lines and spaces) with tight line-width distribution [5]. An exemplary spacer lithography process flow is shown in Fig. 2.2. First, a sacrificial layer is deposited and patterned (into a geometrically regular array of lines and spaces) by conventional lithography. Then, a hard-mask layer is conformally deposited and anisotropically etched, so that hard-mask spacers remain along the sidewalls of the sacrificial material. The spacer width is determined by the thickness of the deposited hard-mask layer, which can be much smaller than the minimum feature size defined by conventional lithography, and very uniform (Fig. 2.3). Note that LER in the patterned sacrificial layer will result in LER for the "inner edges" of the spacers; the LER for the "outer edges" of the spacers will be correlated with that of the "inner edges," except that it will be smoothened due to the conformal nature of the hard-mask deposition process. As a result, the spacer width will always be equal to or greater than the thickness of the deposited hard-mask layer.

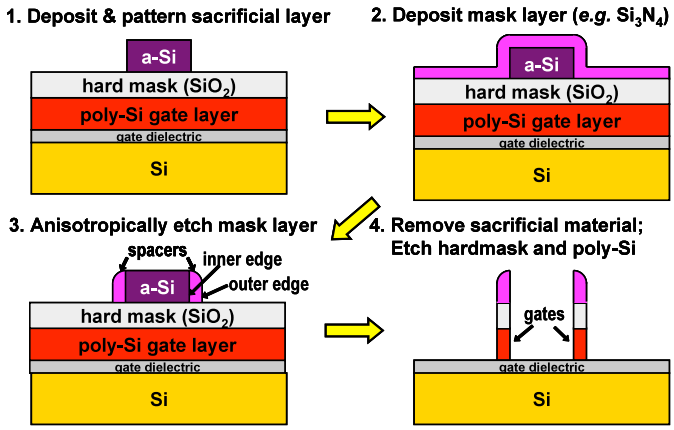


Figure 2.2. Illustrative spacer lithography process flow.

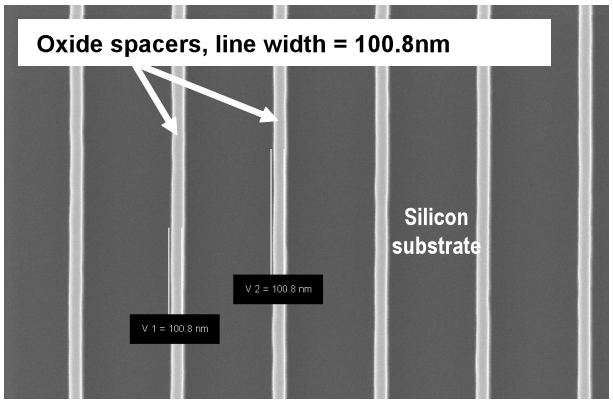


Figure 2.3. Plan-view scanning electron micrograph of oxide spacers formed using 248nm lithography to pattern the sacrificial layer. Very uniform line widths of sub-lithographic dimension (100.8 nm) are achieved.

After selective removal of the sacrificial material, the remaining hard-mask spacers can then be patterned (using conventional lithography and etch) to remove the spacers where gate lines are not desired. Larger feature sizes (if desired) can then be defined in another resist layer using conventional lithography. The resulting composite spacer + resist mask pattern can then be transferred to the underlying gate layer by a selective etch. Note that the minimum feature size and alignment tolerance for the lithography processes used to pattern the spacers and to define the larger feature sizes are much larger than those for a lithography process utilizing a single mask to pattern the gate layer, hence their associated costs should be much lower. Also, the geometrically regular mask used to pattern the sacrificial layer can be used for multiple chip designs, to save cost.

2.3 Gate Line Edge Roughness (LER) Study

2.3.1 Simulation Approach

Gate line edge profiles were generated by MATLAB [6] based on measured LER data, and then input to Sentaurus [4] to generate 3D device structures for 3D device simulation (Fig. 2.4). This procedure was iterated to simulate many devices; then the simulated device *I-V* curves were used to assess variability in device performance. The LER and the line width roughness (LWR) values used for a conventional gate lithography process are 4nm and 6.4nm, respectively. The LWR for the electrical channel length (L_{eff}) was assumed to be the same as that for the gate, because this is the worst case (relevant for future ultra-shallow junction technology). To simulate gate line edge profiles for a spacer lithography process, the profile for one edge (the "inner edge") was duplicated, shifted by the nominal gate length L_G, and then smoothened by filtering out the LER components with spatial frequency higher than $1/(1.5*L_G)$ -- approximating the planarizing effect of a conformal deposition process -- to generate the profile for the other edge (the "outer edge"), as shown in Fig. 2.5. Note that, as a result of the smoothening, the width of the gate electrode at each location along the width of the transistor channel is

equal to or greater than L_G, for a spacer gate lithography process. This is in contrast to a conventional gate lithography process, for which the width of the gate electrode can also be less than L_G.

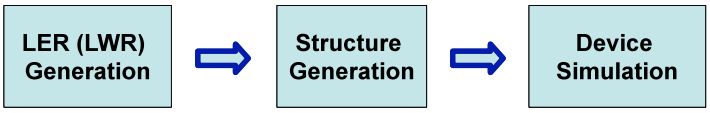


Figure 2.4. Procedure used to study the impact of LER on MOSFET performance.

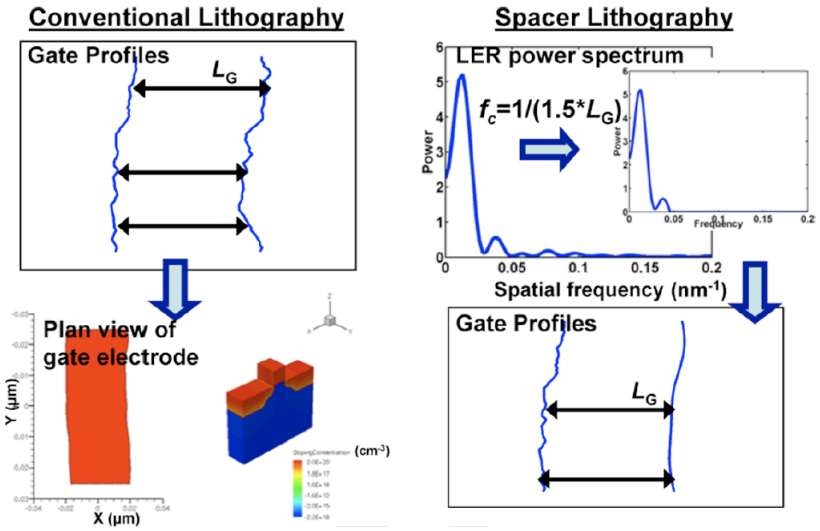


Figure. 2.5. Gate electrode structure generation for conventional lithography vs. spacer gate lithography.

Simulations were performed for the following bulk-Si n-channel MOSFET designs, in order to assess the benefit of spacer gate lithography with gate-length scaling: 1) L_G = 32nm, channel width W_G = 64nm, equivalent oxide thickness EOT = 1.2nm, source/drain extension junction depth X_J = 11nm, supply voltage V_{DD} = 1.1V; 2) L_G = 14nm, W_G = 28nm, EOT = 0.5nm, X_J = 5nm, V_{DD} = 0.9V. The channel region is uniformly doped (p-type) in each case, to a concentration of 3×10^{18} cm^{-3} and 6×10^{18} cm^{-3} for L_G = 32nm and L_G = 14nm, respectively, to achieve drain induced barrier lowering less than 100mV/V. These design parameters are based on the International Technology Roadmap for Semiconductors (ITRS) [7]. It is assumed that gate LER is not reduced with L_G scaling.

2.3.2 Results and Discussion

A. Spacer gate lithography vs. conventional gate lithography

The simulated drain current vs. gate voltage characteristics (I_D-V_G) and I_{ON}-I_{OFF} scatter plots (where I_{ON} is the on-state drive current for V_{GS} = V_{DS} = V_{DD}, and I_{OFF} is the off-state leakage current for V_{DS} = V_{DD} and V_{GS} = 0) for conventional gate lithography and spacer gate lithography for L_G = 32nm and L_G = 14nm are shown in Fig. 2.6 and Fig.

2.7, respectively. Spacer gate lithography dramatically reduces LER-induced variation. Fig. 2.8 shows that variability is well suppressed with gate-length scaling, even if LER does not scale. The average I_{ON} and I_{OFF} values for conventional gate lithography and spacer gate lithography are also indicated in Fig. 2.8. The average value and variance of I_{OFF} are significantly reduced for spacer lithography; the variance of I_{ON} is also reduced, while the average value of I_{ON} is slightly degraded. This is because the average gate length along the width of the channel is larger for spacer lithography than for conventional gate lithography, as explained above. Thus, spacer gate lithography is preferable for smaller L_G: variability is well suppressed, and I_{OFF} is significantly reduced without significantly reducing I_{ON}. (Note: as the thickness of the deposited hard-mask layer is reduced to scale down L_G, the amount of LER smoothening for the "outer edge" is reduced and hence the increase in average gate length will be reduced.).

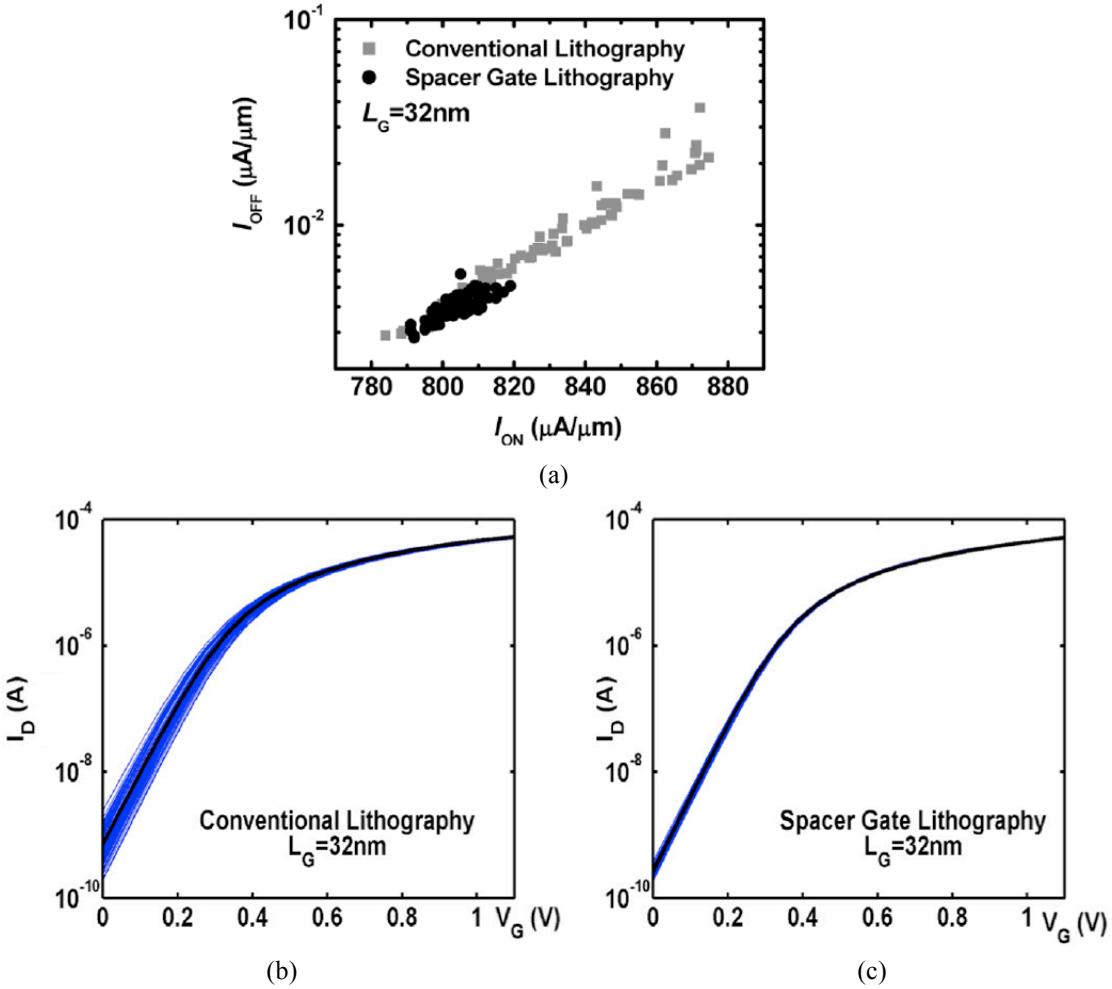


Figure 2.6. (a) Comparison of I_{ON}-I_{OFF} for spacer *vs.* conventional gate lithography. L_G=32nm. (b) I_D-V_G curves for conventional gate lithography -- the heavy black line indicates the average I_D-V_G curve. (c) I_D-V_G curves for spacer gate lithography -- the heavy black line indicates the average I_D-V_G curve.

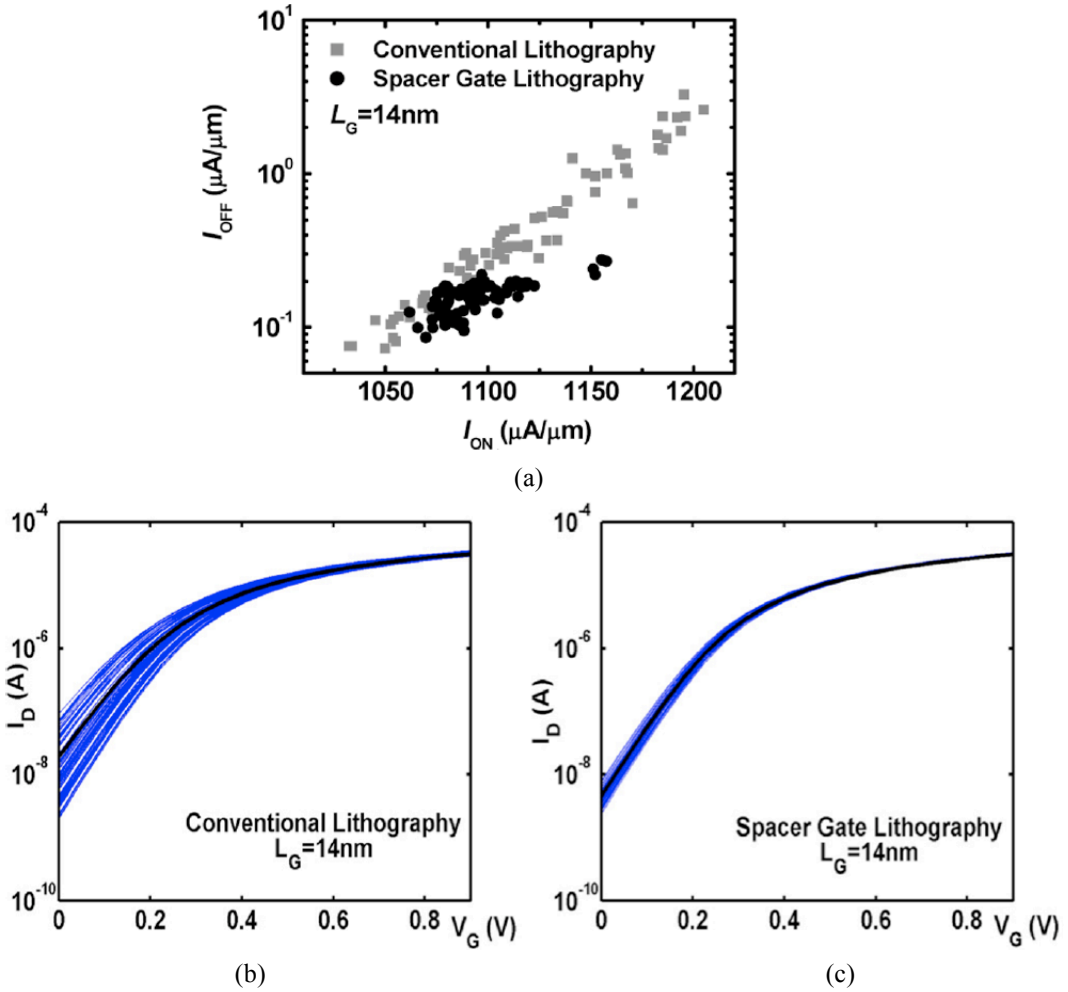


Figure. 2.7. (a) Comparison of I_{ON}-I_{OFF} for spacer *vs.* conventional gate lithography. L_G=14nm. (b) I_D-V_G curves for conventional gate lithography -- the heavy black line indicates the average I_D-V_G curve. (c) I_D-V_G curves for spacer gate lithography -- the heavy black line indicates the average I_D-V_G curve.

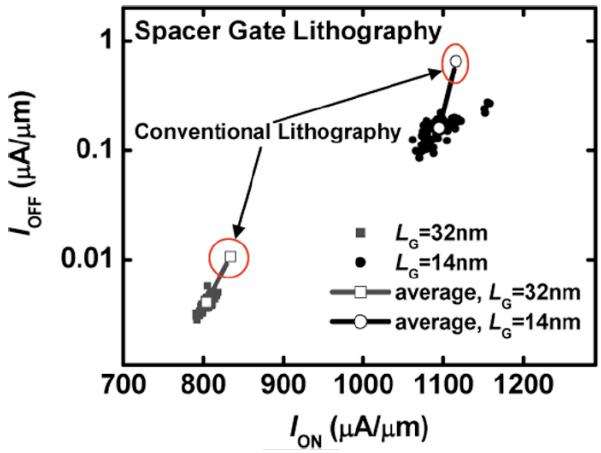


Figure 2.8. Comparison of I_{ON}-I_{OFF} for L_G=14nm *vs.* L_G=32nm, for spacer gate lithography. The average values for conventional gate lithography are also shown for reference.

B. Effects of source/drain placement for a spacer-defined gate electrode

Since the "outer edge" of a spacer-defined gate is smoothened, the source and drain edge profiles are not perfectly correlated. The effect of source/drain placement is shown in Fig. 2.9. Since current is limited by thermionic emission over the source-to-channel potential barrier in the off state, variations in barrier height caused by electric-field crowding due to source-side gate edge roughness have more impact on I_{OFF}; therefore, more reduction in I_{OFF} spread is obtained when the source edge profile is smoothened. The devices with smooth source profile have higher average I_{ON} and lower average I_{OFF}, because the electric-field crowding due to the drain-side gate edge roughness effectively reduces channel length, while the source-to-channel potential barrier lowering due to the electric-field crowding is eliminated. Since random dopant fluctuations (RDF) also have significant impact on deeply scaled MOSFETs [1,8], the effect of LER together with RDF in the source/drain and channel regions was studied using Kinetic Monte Carlo simulations, for the 14nm device design ($L_{G,\ nominal}$ = 14nm, W_G = 28nm, EOT = 0.5nm, X_J = 5nm, V_{DD} = 0.9V, channel doping concentration = 6×10^{18} cm^{-3}). Fig. 2.10 shows that the performance advantages of devices with smooth source profile are retained in the presence of RDF.

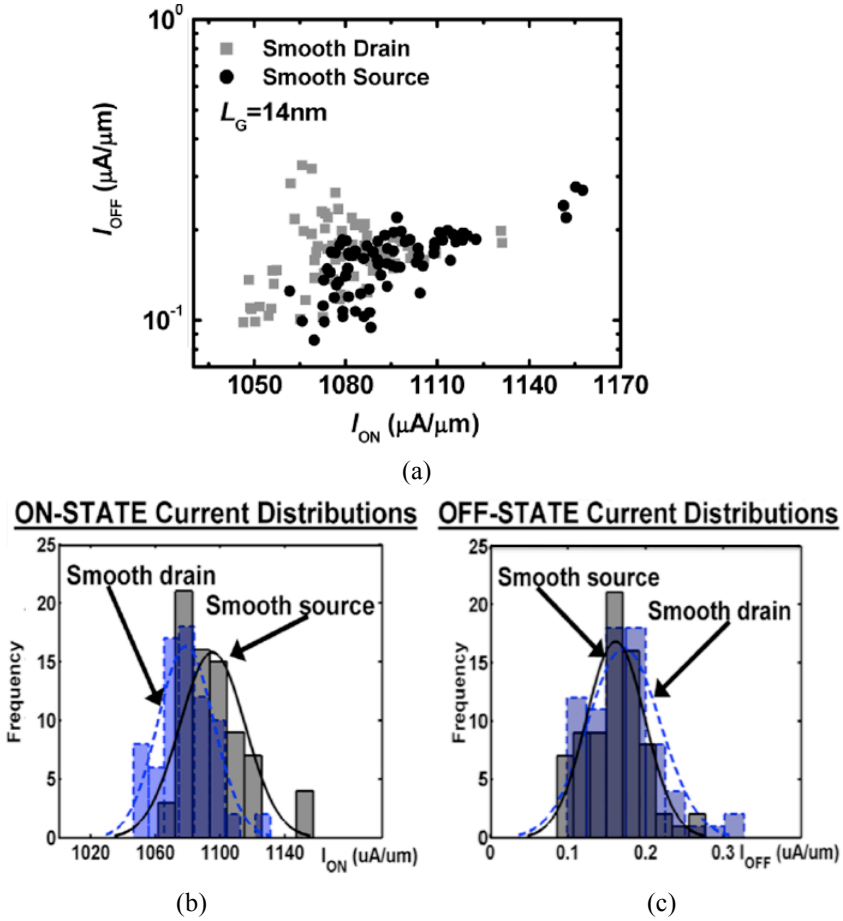
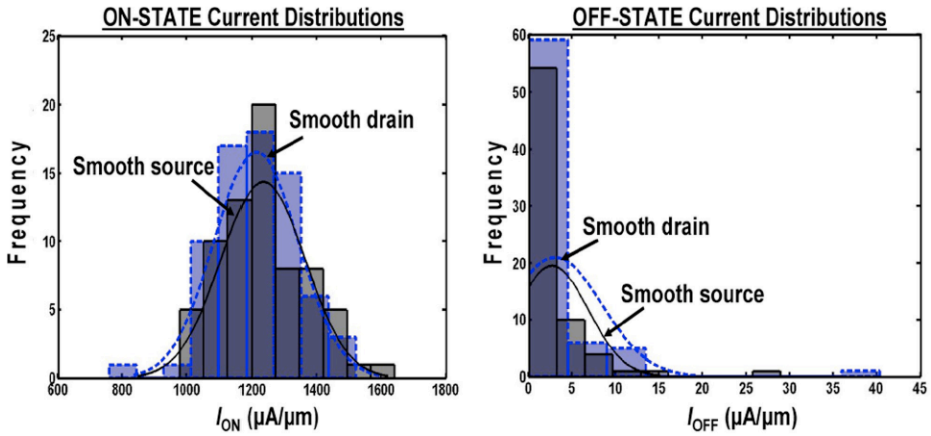


Figure. 2.9. Comparison of I_{ON}-I_{OFF} for smooth source *vs.* smooth drain. L_G=14nm. (a) I_{OFF}-I_{ON}. (b) I_{ON} distributions. (c) I_{OFF} distributions.

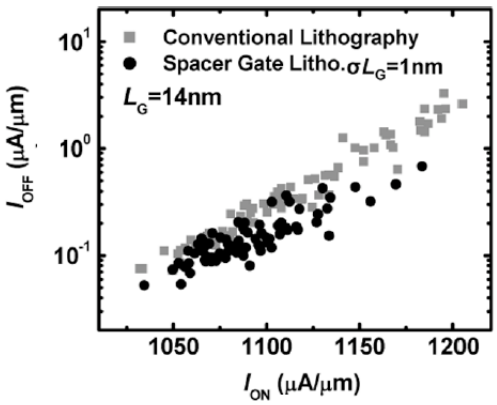


(a) (b)

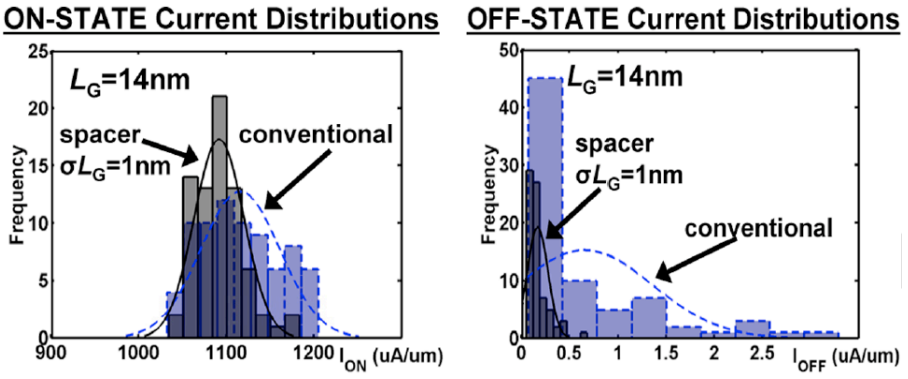
Figure. 2.10. Comparison of I_{ON}-I_{OFF} for smooth source *vs.* smooth drain, including the effects of RDF. L_G=14nm. (a) I_{ON} distributions. (b) I_{OFF} distributions.

C. Effects of spacer width variation

Spacer gate lithography reduces LER-induced variation because the gate-edge profiles are correlated, resulting in lower LWR. However, other sources of variation are introduced in the spacer lithography process, such as film thickness non-uniformity, etch non-uniformity, etc. Device performance distributions for conventional gate lithography vs. spacer gate lithography are compared in Fig. 2.11 and Fig. 2.12 for spacer-width standard deviation values of 1nm and 2nm, respectively. The results show that spacer gate lithography reduces variability if the standard deviation in spacer width is less than 2nm. This is consistent with the LWR specified in the simulation inputs. (LWR is 6.4nm, which is very close to the 3σ value.) For a 14nm spacer process, ≤2nm standard deviation in spacer width is reasonable. Therefore, despite the additional sources of variation (other than LER), spacer gate lithography can significantly reduce variability in MOSFET performance.

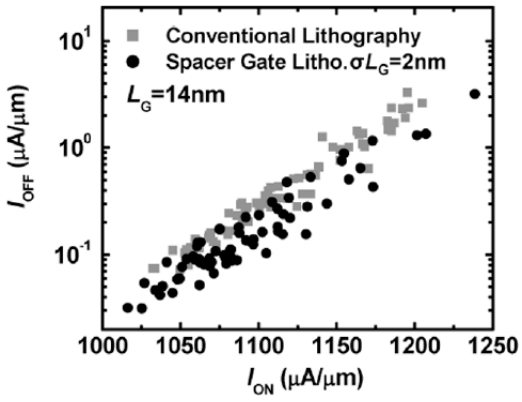


(a)

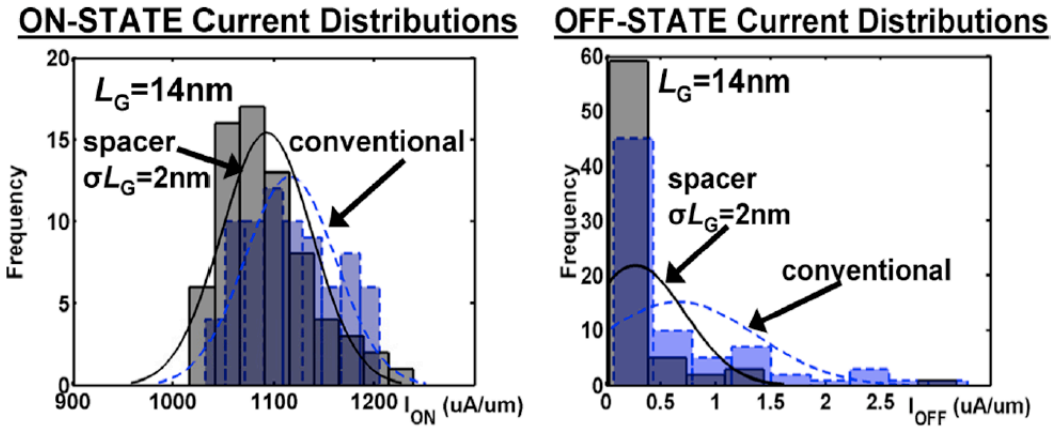


(b) (c)

Figure 2.11. Comparison of I_{ON} and I_{OFF} for spacer gate lithography (σL_G=1nm) *vs.* conventional gate lithography. L_G=14nm, smooth source. (a) I_{OFF} *vs.* I_{ON}. (b) I_{ON} distributions. (c) I_{OFF} distributions.



(a)



(b) (c)

Figure. 2.12. Comparison of I_{ON} and I_{OFF} for spacer gate lithography (σL_G=2nm) *vs.* conventional gate lithography. L_G=14nm, smooth source. (a) I_{OFF} *vs.* I_{ON}. (b) I_{ON} distributions. (c) I_{OFF} distributions.

2.4 Summary

As gate lengths and thermal process budgets are reduced, the effect of gate line edge roughness becomes increasingly significant. Statistical 3D device simulations show that spacer gate lithography is a scalable technology which can significantly reduce LER-induced variability in MOSFET performance.

2.5 References

[1] A. Asenov, "Simulation of statistical variability in nano MOSFETs," *Symp. VLSI Tech. Dig.*, pp. 86-87, 2007.

[2] A. Kaneko, A. Yagishita, K. Yahashi, T. Kubota, M. Omura, K. Matsuo, I. Mizushima, K. Okano, H. Kawasaki, S. Inaba, T. Izumida, T. Kanemura, N. Aoki, K. Ishimaru, H. Ishiuchi, K. Suguro, K. Eguchi, and Y. Tsunashima, "Sidewall transfer process and selective gate sidewall spacer formation technology for sub-15nm FinFET with elevated source/drain extension," *IEDM Tech. Dig*., pp. 863-866, 2005.

[3] R. Rooyackers, E. Augendre, B. Degroote, N. Collaert, A. Nackaerts, A. Dixit, T. Vandeweyer, B. Pawlak, M. Ercken, E. Kunnen, G. Dilliway, F. Leys, R. Loo, M. Jurczak, and S. Biesemans, "Doubling or quadrupling MuGFET fin integration scheme with higher pattern fidelity, lower CD variation and higher layout efficiency," *IEDM Tech. Dig*., pp. 993-996, 2006.

[4] Sentaurus User's Manual, v. 2006.06 (Synopsys, Inc.)

[5] Y.-K. Choi, T.-J. King, and C. Hu, "A spacer patterning technology for nanoscale CMOS," *IEEE Trans. Electron Devices*, Vol. 49, pp. 436-441, March 2002.

[6] MATLAB User's Manual (The MathWorks, Inc.)

[7] *Int'l Technology Roadmap for Semiconductors*, 2006 Update.

[8] M. Miyamura, T. Fukai, T. Ikezawa, R. Ueno, K. Takeuchi, and M. Hane, "SRAM critical yield evaluation based on comprehensive physical / statistical modeling, considering anomalous non-Gaussian intrinsic transistor fluctuations," *Symp. VLSI Tech. Dig.*, pp. 22-23, 2007.

Chapter 3

Tri-Gate Bulk MOSFET Design

3.1 Introduction

Challenges for continued scaling of bulk-Si CMOS technology include increased leakage current (I_{OFF}) [1] and variability in transistor performance [2]. Non-classical transistor structures such as thin-body MOSFETs [3] provide better control of short-channel effects (SCE) without the need for heavy channel surface doping and thus can reduce variability in threshold voltage (V_T) due to random dopant fluctuations (RDF). Although it is generally agreed that such structures (e.g. the FinFET [4] or the tri-gate silicon-on-insulator (SOI) FET [5]) will be necessary in the sub-20 nm gate length (L_G) regime, they require either expensive SOI substrates or more complex fabrication processes (e.g. [6]) and advanced compact models for circuit design, which present barriers to their adoption. For example, to circumvent the need for an SOI substrate, body-tied FinFETs have been investigated recently [7, 8, 9]. These devices have large channel aspect ratio, because they rely on a narrow and tall (fin-like) channel to suppress SCE and achieve good layout area efficiency, respectively. In this chapter we propose a tri-gate bulk MOSFET design which utilizes a low-aspect-ratio channel, for improved manufacturability.

3.2 Tri-Gate Bulk MOSFET Structure

3.2.1 Device Structure

The channel region of a tri-gate bulk MOSFET consists of one or more parallel segments ("stripes") of equal width, W_{STRIPE}, which can be substantially larger (by as much as ~ 4×) than L_G. Within each stripe, the channel- and source/drain-doping profiles

are similar to those in a planar bulk MOSFET, as illustrated in Fig. 3.1. The thickness of the lightly doped channel, T_{Si}, corresponds to the depth of the retrograde channel doping, which can be controlled precisely by ion implantation. Note that the stripes within a single multi-stripe transistor are isolated by very shallow trench isolation (VSTI) oxide, which extends to a depth below the source/drain extension regions, but which can be much shallower than the STI oxide that is used to isolate individual transistors, so that there is no need to form high-aspect-ratio fins. The deep source/drain regions within a single multi-stripe transistor may be contiguous underneath the VSTI oxide. Such a segmented MOSFET structure may seem to be very complicated to manufacture. However, it can be fabricated in a relatively straightforward manner using a conventional process flow, starting with a substrate having a corrugated semiconductor surface planarized by VSTI oxide [10].

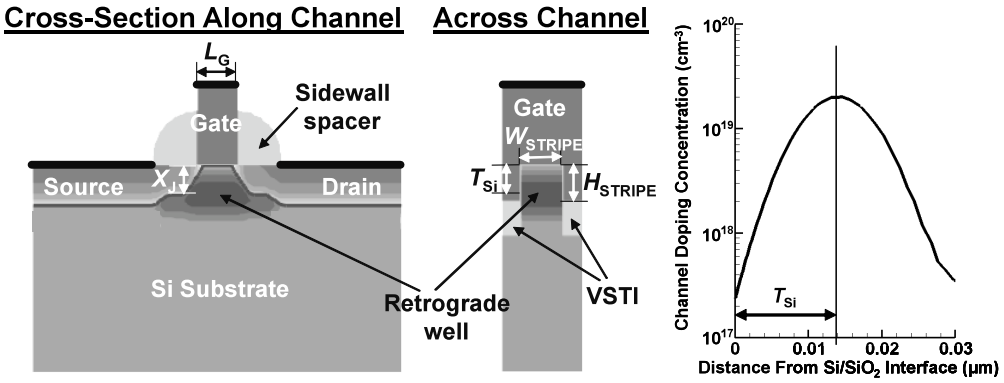


Figure 3.1. Cross-sectional views of the simulated 3-D tri-gate bulk MOSFET structure.

To provide for improved electrostatic integrity and layout area efficiency (effective channel width per unit layout width), the VSTI/STI oxide surrounding the active-area stripe(s) can be recessed by a small amount, H_{STRIPE}, just prior to gate stack formation, so that the gate electrode wraps around the top portion of each stripe, resulting in a tri-gate structure. It should be noted that narrow-width MOSFETs (such as those used in SRAM cells) can be readily made into (single-stripe) tri-gate MOSFETs in this manner, *i.e.* simply by recessing the STI oxide prior to gate stack formation in a standard CMOS fabrication process. The combination of retrograde channel doping [11] with a multi-gate structure provides for superior electrostatic integrity as well as significantly reduced V_T variation due to RDF [12, 13].

3.2.2 Simulation Results and Discussion

Sentaurus 3-dimensional process and device simulations were performed using advanced physical models [14] to study the performance of tri-gate bulk MOSFETs *vs.* planar bulk MOSFETs. The retrograde well doping profiles were identical for these two device designs. Device parameters were selected based on ITRS specifications [15]. The well doping profile steepness is 4 nm/decade, which is less abrupt than required for the

source/drain extensions [15]. No mobility enhancement was assumed, although well-established techniques for inducing strain in the channel region can be readily applied.

A. Tri-Gate Bulk MOSFET Benefit for Improved Scalability

The use of retrograde well doping (or halo doping) allows the tri-gate bulk MOSFET to achieve electrostatic integrity superior to that of the tri-gate SOI MOSFET design [5] because drain electric field penetration through the buried oxide (BOX) is eliminated. Therefore W_{STRIPE} and/or T_{Si} ($\sim H_{STRIPE}$) can be larger for the same degree of SCE control, as shown in Fig. 3.2. Note that the use of heavy doping beneath the BOX is less effective for the tri-gate SOI MOSFET, even if an ultra-thin (10nm-thick) BOX is used. For example, W_{STRIPE} can be comparable to L_G without requiring a channel thickness less than $\frac{2}{3}L_G$, which is advantageous for improving manufacturability (since W_{STRIPE} does not need to be less than L_G) and achieving good layout area efficiency (because the gated stripe sidewalls contribute to current conduction). The retrograde well doping profile also provides for light doping at the upper channel surfaces to mitigate corner conduction effects [16].

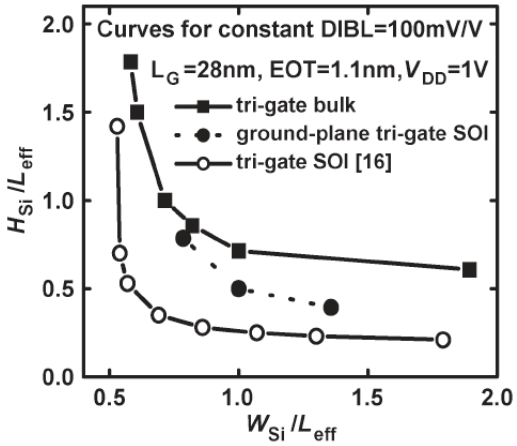


Figure 3.2. Stripe dimensions required to adequately suppress short-channel effects, from 3-D device simulations. For a specified amount of drain-induced barrier lowering (100 mV/V), the tri-gate bulk MOSFET offers higher layout efficiency than the tri-gate SOI transistor because taller stripes can be used (providing for more sidewall current conduction). This is the case even if the SOI device has a "ground plane" (doped to a concentration of 10^{20} cm^{-3}) beneath an ultra-thin (10nm-thick) BOX.

B. Tri-Gate Bulk MOSFET Benefit for Reduced Variability

Fig. 3.3 shows drain current *vs.* gate voltage (I_{DS}-V_{GS}) characteristics (for 100 simulated cases) of tri-gate bulk MOSFETs and planar bulk MOSFETs with atomistic doping effects included, following the methodology described in [17, 18]. (Dopant atoms within the channel and the source/drain gradient regions were randomly placed (*e.g.* Fig.

3.3a inset) using a Monte Carlo algorithm [19].) Because the nominal (continuum) body and source/drain doping profiles are identical for the tri-gate and planar MOSFETs, the reduction in V_T variation is due solely to their structural difference. The depletion charge density per unit channel width is smaller for the tri-gate structure, so that RDF effects are mitigated.

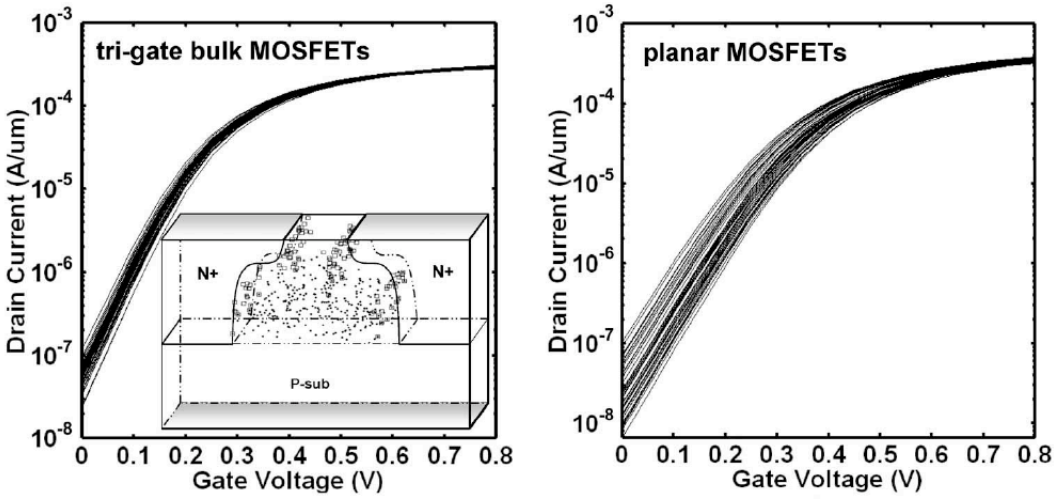
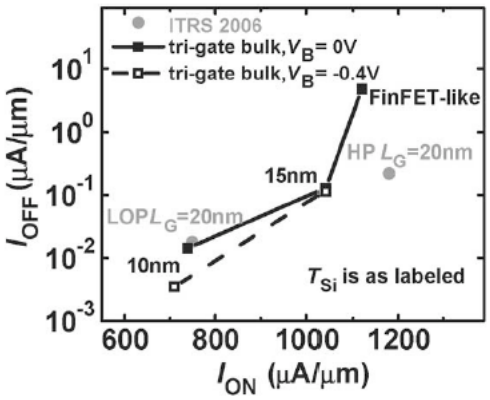


Figure 3.3. Simulated I_{DS}-V_{GS} curves for transistors (L_G = 20 nm) with discretely placed dopant atoms. Each curve corresponds to a different random placement of the source/drain-gradient and retrograde-channel dopant atoms, *e.g.* as shown in the inset in (a) where each square corresponds to an n-type dopant atom and each dot corresponds to a p-type dopant atom. Currents are normalized to effective channel width. (a) tri-gate bulk MOSFETs with W_{STRIPE} = 20 nm, H_{STRIPE} = 14 nm, $W_{SPACING}$ = 20 nm; (b) planar MOSFETs. Device layout width is 40 nm. EOT = 0.9 nm, X_J = 14nm, depth of retrograde channel doping T_{Si} = 14nm, gate work function = 4.1eV.

C. V_T Adjustment via Doping and Body Biasing

The dependences of I_{ON} and I_{OFF} (normalized to layout width) on T_{Si} (~H_{STRIPE}) and W_{STRIPE} are shown in Fig. 3.4. (No strain-induced mobility enhancement is assumed.) The depth of the retrograde doping can be used to tune I_{ON} *vs.* I_{OFF}, so that V_T can be adjusted without heavy channel surface doping or gate work function tuning. If T_{Si} is less than one-half of W_{STRIPE}, body biasing is effective for dynamic V_T control to optimize the power *vs.* delay tradeoff, or for compensation of process-induced V_T variations for improved parametric yield [1].



	Stripe width (nm)	Stripe spacing (nm)	Stripe height (nm)	VSTI depth (nm)
Moderately recessed	20	20	10	30
Very recessed	20	20	15	35
FinFET-like	10	20	20	40

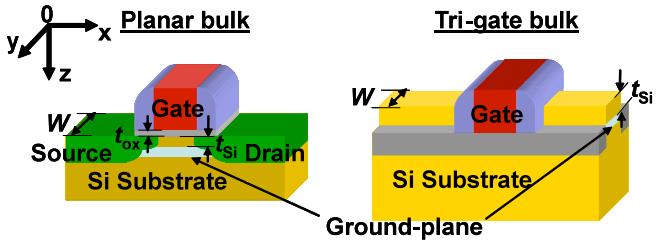
Figure 3.4. Simulated I_{ON} & I_{OFF} (normalized to layout width) for n+ poly-Si-gated tri-gate bulk MOSFETs with steep retrograde channel doping (peak concentration = $2x10^{19}$ cm^{-3}, steepness = 4nm/dec). L_G = 20nm, W_{STRIPE} = 20nm, EOT = 0.9nm, V_{DD} = 1V. Values of the depth of retrograde channel doping T_{Si} (~H_{STRIPE}) are indicated. X_J and H_{STRIPE} each increase with T_{Si}. No mobility enhancement is assumed for the tri-gate bulk MOSFETs, in contrast to the ITRS specifications.

The tri-gate bulk MOSFET design provides an evolutionary pathway for bulk CMOS scaling to the end of the roadmap. The combination of retrograde channel doping with a multi-gate structure provides for superior electrostatic integrity, so that an ultra-thin/ultra-narrow channel is not required in order to suppress short-channel effects. As compared with the classic planar MOSFET design, the tri-gate bulk MOSFET also provides for reduced variability due to random dopant fluctuations. It also allows for V_T tuning via channel doping depth (*vs.* dose), as well as the possibility of dynamic V_T adjustment via body biasing for further yield improvement and/or versatility in circuit design.

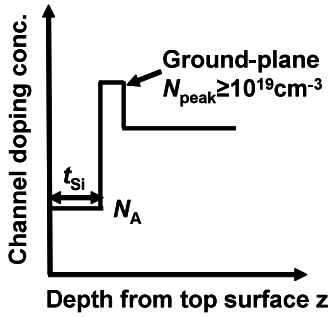
3.3 Scale Length Assessment

3.3.1 Introduction

The planar ground-plane bulk MOSFET and double-gate SOI MOSFET structures have been studied as candidates to extend transistor scaling to its ultimate limit [11][20][21]. In section 3.2, the tri-gate bulk MOSFET design was proposed to provide the advantages of a multi-gate SOI MOSFET together with the advantages of a planar bulk MOSFET [22]. To assess its scaling limit, 3-dimensional (3-D) effects need to be considered. This section analytically derives the scale length for tri-gate and planar ground-plane bulk MOSFETs, and compares the scaling limit of different MOSFET designs (Fig. 3.5).



(a)



(b)

Figure 3.5. (a) 3-D view of MOSFET designs studied in this work. $x = 0$, $y = 0$, and $z = 0$ correspond to the location of the source-side gate edge, the center of the channel along its width, and the gate oxide-channel interface, respectively. (b) Channel doping profile for tri-gate and planar ground-plane bulk MOSFETs.

3.3.2 Scale Length Derivation

The scale length of the tri-gate bulk MOSFET is derived herein, for the coordinate system shown in Fig. 3.5. The potential distribution $\Phi(x,y,z)$ in the channel region is determined by Poisson's equation

$$\frac{d^2\Phi}{dx^2}+\frac{d^2\Phi}{dy^2}+\frac{d^2\Phi}{dz^2}=\frac{qN_A}{\varepsilon_{Si}} \qquad (1)$$

for $0 \le x \le L_{eff}$, $-\frac{W}{2} \le y \le \frac{W}{2}$, and $0 \le z \le t_{Si}$. N_A is the channel dopant concentration. Parabolic potential variation between the two lateral gates (in the y-direction) is assumed [23] to describe the potential distribution:

$$\Phi(x,y,z) \approx a_0(x,z)+a_1(x,z)y+a_2(x,z)y^2 \qquad (2).$$

At $y = 0$, we assume a simple parabolic function in the z-direction [24]:

$$\Phi(x,0,z) = a_0(x,z) \approx c_0(x)+c_1(x)z+c_2(x)z^2 \qquad (3).$$

The boundary conditions are: $\Phi(x,0,0) = \Phi_f(x) = a_0(x,0) = c_0(x)$ (4)

$$\left.\frac{d\Phi}{dz}\right|_{y=0,z=0} = \frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{\Phi_f(x)-\Phi_{gs}}{t_{ox}} = c_1(x) \quad (5)$$

$$\Phi(x,0,t_{Si}) = \Phi_{sb} = a_0(x,t_{Si}) = c_0(x) + t_{Si}c_1(x) + t_{Si}^2 c_2(x) \quad (6)$$

$$\Phi(x,-\frac{W}{2},z) = \Phi(x,\frac{W}{2},z) \quad (7)$$

$$\Phi(x,\pm\frac{W}{2},z) = \Phi_f(x) = a_0(x,z) + a_2(x,z)\frac{W^2}{4} \quad (8)$$

where $\Phi_f(x)$ and Φ_{sb} are the potentials at $z=0$ and $z=t_{Si}$. $a_0(x,z)$ can be determined by the boundary conditions in Eqns. (4-6). From Eqn. (7), we obtain $a_1(x,z)=0$. $a_2(x,z) = \frac{4}{W^2}\left[\Phi_f(x) - a_0(x,z)\right]$ is obtained by solving Eqn. (8). Using these boundary conditions, we obtain the following expression for the potential distribution within the channel:

$$\Phi(x,y,z) = \Phi_f(x) + \frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{\Phi_f(x)-\Phi_{gs}}{t_{ox}}z - \left[\frac{\Phi_f(x)-\Phi_{sb}}{t_{Si}^2} + \frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{\Phi_f(x)-\Phi_{gs}}{t_{ox}t_{Si}}\right]z^2$$
$$-\frac{4}{W^2}\left\{\frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{\Phi_f(x)-\Phi_{gs}}{t_{ox}}z - \left[\frac{\Phi_f(x)-\Phi_{sb}}{t_{Si}^2} + \frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{\Phi_f(x)-\Phi_{gs}}{t_{ox}t_{Si}}\right]z^2\right\}y^2 \quad (9).$$

The following equation is obtained by substituting (9) into (1) and setting y to 0:

$$\left(1 + \frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{z}{t_{ox}} - \frac{z^2}{t_{Si}^2} - \frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{z^2}{t_{ox}t_{Si}}\right)\frac{d^2\Phi_f(x)}{dx^2} - \left(\frac{8}{W^2}\frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{z}{t_{ox}} - \frac{8}{W^2}\frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{z^2}{t_{ox}t_{Si}} + \frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{2}{t_{ox}t_{Si}}\right)\cdot$$
$$\left[\Phi_f(x)-\Phi_{gs}\right] - \left(\frac{2}{t_{Si}^2} - \frac{8}{W^2}\frac{z^2}{t_{Si}^2}\right)\left[\Phi_f(x)-\Phi_{sb}\right] = \frac{qN_A}{\varepsilon_{Si}} \quad (10).$$

By making the following transformations

$$\lambda = \sqrt{\frac{1 + \frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{z}{t_{ox}} - \frac{z^2}{t_{Si}^2} - \frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{z^2}{t_{ox}t_{Si}}}{\frac{8}{W^2}\left(\frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{z}{t_{ox}} - \frac{z^2}{t_{Si}^2} - \frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{z^2}{t_{ox}t_{Si}}\right) + \frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{2}{t_{ox}t_{Si}} + \frac{2}{t_{Si}^2}}} \quad (11)$$

and

$$\phi(x) = \left(1 + \frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{z}{t_{ox}} - \frac{z^2}{t_{Si}^2} - \frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{z^2}{t_{ox}t_{Si}}\right)\Phi_f(x) + \frac{qN_A}{\varepsilon_{Si}}\lambda^2$$
$$-\left(\frac{8}{W^2}\frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{z}{t_{ox}} - \frac{8}{W^2}\frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{z^2}{t_{ox}t_{Si}} + \frac{\varepsilon_{ox}}{\varepsilon_{Si}}\frac{2}{t_{ox}t_{Si}}\right)\Phi_{gs}\lambda^2 - \left(\frac{2}{t_{Si}^2} - \frac{8}{W^2}\frac{z^2}{t_{Si}^2}\right)\Phi_{sb}\lambda^2 \quad (12),$$

Eqn. (10) can be simplified to

$$\frac{d^2\phi(x)}{dx^2} - \frac{\phi(x)}{\lambda^2} = 0 \quad (13),$$

where λ is the scale length. In calculating λ, the value of z corresponding to the effective location of the conduction path of off-state leakage, *i.e.* the "weak spot" in the channel, should be used. Given the symmetry of the structure, $y = 0$ at the "weak spot":

$$\Phi(x,y,z) = \Phi_f(x) + \frac{\varepsilon_{ox}}{\varepsilon_{Si}} \frac{\Phi_f(x) - \Phi_{gs}}{t_{ox}} z - \left[\frac{\Phi_f(x) - \Phi_{sb}}{t_{Si}^2} + \frac{\varepsilon_{ox}}{\varepsilon_{Si}} \frac{\Phi_f(x) - \Phi_{gs}}{t_{ox} t_{Si}} \right] z^2 \quad (14).$$

Eqn. (14) is a quadratic function of z, which reaches a maximum at

$$z_{peak} = \frac{\frac{\varepsilon_{ox}}{\varepsilon_{Si}} \frac{\Phi_f(x) - \Phi_{gs}}{t_{ox}}}{2 \cdot \left[\frac{\Phi_f(x) - \Phi_{sb}}{t_{Si}^2} + \frac{\varepsilon_{ox}}{\varepsilon_{Si}} \frac{\Phi_f(x) - \Phi_{gs}}{t_{ox} t_{Si}} \right]} = \frac{1}{\frac{2\varepsilon_{Si}}{\varepsilon_{ox}} \frac{t_{ox}}{t_{Si}^2} \frac{\Phi_f(x) - \Phi_{sb}}{\Phi_f(x) - \Phi_{gs}} + \frac{2}{t_{Si}}} \quad (15).$$

In the off-state, $\Phi_{sb} = \Phi_{gs} = 0$, and Eqn. (15) is simplified to $z_{peak} = \frac{1}{\frac{2\varepsilon_{Si}}{\varepsilon_{ox}} \frac{t_{ox}}{t_{Si}^2} + \frac{2}{t_{Si}}}$, which corresponds to the location of the highest leakage current density.

3.3.3 Scale Length Comparison

Fig. 3.6 shows scale length contour lines for different values of channel width (W) and lightly doped channel thickness (t_{Si}). For reference, the scale length for the double-gate SOI MOSFET [20] design is also shown. As W increases, the scale length for the tri-gate bulk MOSFET approaches the value of the scale length for the planar ground-plane bulk MOSFET ($W >> t_{Si}$): $\lambda = \sqrt{\frac{1 + \frac{\varepsilon_{ox}}{\varepsilon_{Si}} \frac{z_{peak}}{t_{ox}} - \frac{z_{peak}^2}{t_{Si}^2} - \frac{\varepsilon_{ox}}{\varepsilon_{Si}} \frac{z_{peak}^2}{t_{ox} t_{Si}}}{\frac{\varepsilon_{ox}}{\varepsilon_{Si}} \frac{2}{t_{ox} t_{Si}} + \frac{2}{t_{Si}^2}}}$. It should be noted that this formula yields a larger scale length for the planar ground-plane MOSFET than the formula given in [11], because the scale length formula derived in [11] assumed that the off-state leakage path is at the gate oxide-channel interface so that it is overly optimistic. For given values of W and t_{Si}, the scale length is smallest for the tri-gate bulk MOSFET design because it *combines* the benefit of a multi-gate structure (improved gate control) with the benefit of a ground-plane structure (leakage suppression). Fig. 3.7 shows the off-state electron potential distribution along the leakage path, for each of the transistor structures with identical channel dimensions (W and t_{Si}). The tri-gate bulk MOSFET has the least amount of drain-induced barrier lowering and therefore is most scalable.

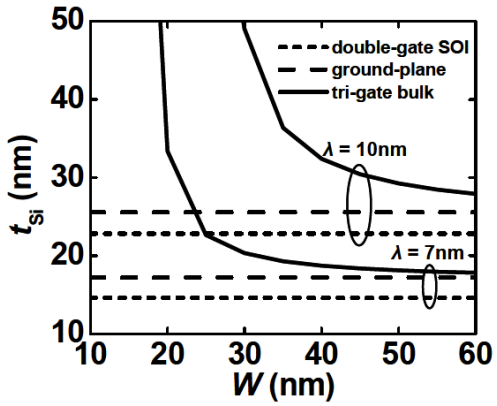


Figure 3.6. Channel dimensions required to achieve 7nm and 10nm scale length. t_{ox} = 1nm.

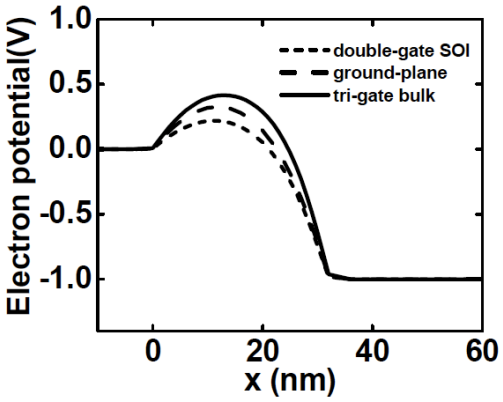


Figure 3.7. Off-state electron potential distribution along the leakage path. L_{eff} = 32nm, W = 20nm, t_{Si} = 20nm, t_{ox} = 1nm, V_{dd} = 1V.

Fig. 3.8 shows the scale length sensitivity to t_{ox} for different designs. Tri-gate and planar ground-plane bulk MOSFETs are less sensitive to t_{ox} variation as compared with the double-gate SOI MOSFET, because they each employ a ground-plane to help suppress short channel effects (SCE), whereas the double-gate SOI structure relies on good gate control (*i.e.* a combination of ultra-thin t_{ox} and thin t_{Si}) to suppress SCE.

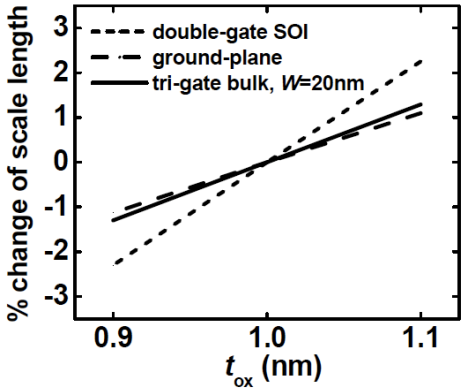


Figure 3.8. Scale length sensitivity to t_{ox}. The t_{si} values are selected to obtain λ = 7nm at t_{ox} = 1nm.

3.4 Design Optimization of Tri-Gate Bulk MOSFET

As CMOS technology scaling continues, short-channel effects (SCE) and variability in transistor performance become increasingly difficult problems for the planar bulk MOSFET design [1, 2], so that alternative MOSFET designs eventually will be needed to extend transistor scaling into the sub-20nm gate length regime [15]. The FinFET [25] is a leading candidate: it utilizes the combination of a thin channel (which eliminates sub-surface leakage paths) with a double-gate structure (which increases capacitive coupling between the gate and the channel) to suppress SCE and variability. However, the FinFET presents significant challenges for manufacturing because it requires the formation of narrow (sub-gate-length) fins [4] with uniform width [26] and large (>1) aspect ratio [27] particularly if a bulk silicon wafer is to be used [28].

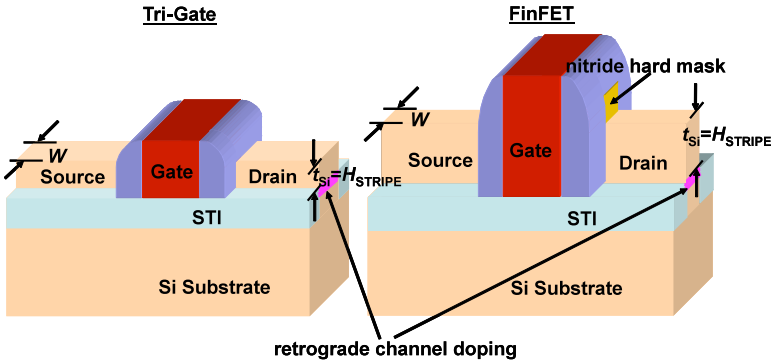
The tri-gate bulk MOSFET design was proposed in section 3.2 to provide a more evolutionary pathway for continued transistor scaling [22]. It utilizes a combination of retrograde channel doping (which suppresses drain-induced barrier lowering, DIBL) with a triple-gate structure (which increases capacitive coupling between the gate and the channel) to suppress SCE and variability. Thus, it offers superior electrostatic integrity (hence scalability) as compared to the double-gate MOSFET (*i.e.* the FinFET) (as discussed in section 3.3, [29]), without requiring the formation of sub-gate-length or high-aspect-ratio features. In this section, design optimization and performance of tri-gate bulk MOSFETs [22] *vs*. bulk FinFETs [28] are compared.

Sentaurus 3-dimensional (3-D) device simulations were performed using advanced physical models [14] to study transistor performance as a function of the retrograde channel doping profile and the effective channel length (L_{eff}), physical channel width (W) and height (H_{STRIPE}). 3-D quantization effects were included using the density gradient quantization model. Hydrodynamic model (physical parameters tuned according to Monte Carlo simulation results) was used to model the transport of the carriers [14]. The values used for the supply voltage V_{DD} and other device design parameters were based on ITRS low-operating-power (LOP) specifications at gate length (L_G) values of 18nm and 13nm [15]. For simplicity, no mobility enhancement was assumed.

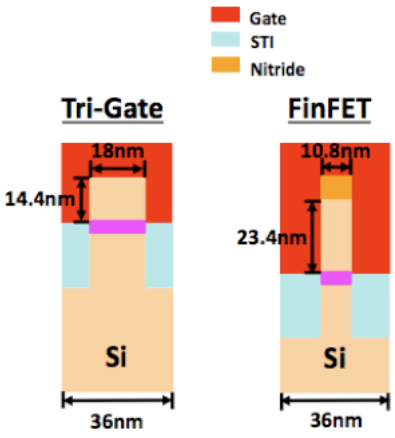
3.4.1 Tri-Gate vs. FinFET Design Optimization

In principle, the height (H_{STRIPE}) of a FinFET can be made to be very tall to achieve high layout efficiency, *i.e.* large effective channel width (W_{eff}) per unit layout area. In practice, however, shorter fins are preferred for ease of manufacture and for design flexibility (*i.e.* to allow for finer increments in designed W_{eff}). Therefore, optimized bulk tri-gate and FinFET designs of the same W_{eff} are compared (for each H_{STRIPE} of the tri-gate design). Fig. 3.9 shows 3-D and cross-sectional views (across the channel) of the tri-gate and FinFET structures, for L_G = 18nm. Note that they are assumed to have the same channel stripe pitch ($2*L_G$, which is an aggressive estimation), as this is set by

lithography limitations. For the tri-gate design, H_{STRIPE} values are set to 0.6*L_G (lower bound for high I_{ON} and good manufacturability) and 0.8*L_G (upper bound for less than 100mV/V DIBL when the silicon stripe width W and effective channel length L_{eff} are each equal to L_G, from Fig. 3.10). t_{Si} (the peak depth of the retrograde channel doping profile) is set to be equal to H_{STRIPE}. L_{eff} is then optimized (in practice by adjusting the width of the gate-sidewall spacers) to maximize the on-state drive current (I_{ON}) for a fixed 7nA/μm off-state leakage current (I_{OFF}) as specified for double-gate FETs [15], while maintaining DIBL to be less than 100mV/V. The optimal design is then obtained to minimize intrinsic delay C_{total}*V_{DD}/I_{eff}, where C_{total} is the total gate capacitance, and I_{eff} is the average of the drain current I_D for V_{GS} = V_{DD} and V_{DS} = $V_{DD}/2$ and I_D for V_{GS} = $V_{DD}/2$ and V_{DS} = V_{DD}.



(a)



(b)

Figure 3.9. (a) 3-D and (b) cross-sectional views (across the channel) of the simulated tri-gate and FinFET bulk MOSFET structures, for L_G = 18nm.

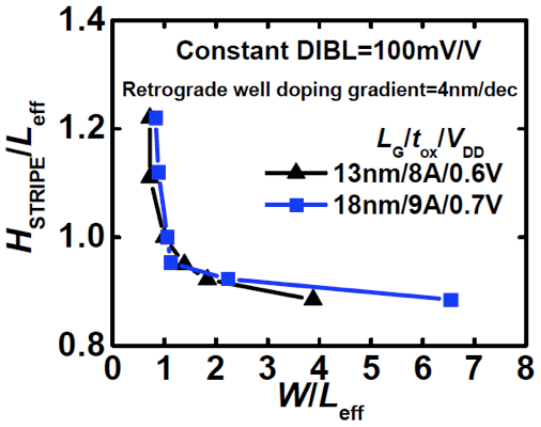


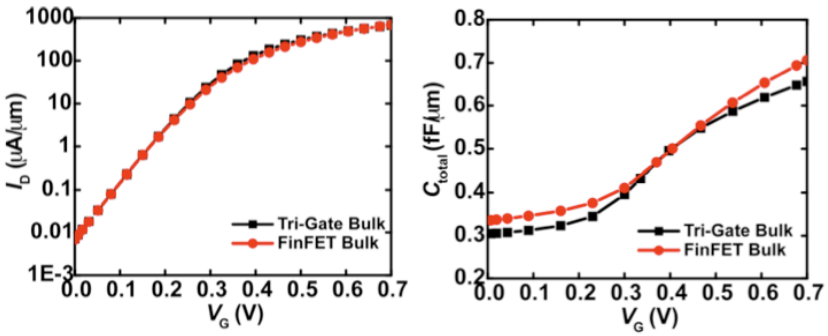
Figure 3.10. Stripe dimensions required to adequately suppress short-channel effects, from 3-D device simulations.

For the FinFET design, the value of $H_{STRIPE} = t_{Si}$ is selected to provide the same W_{eff} as the tri-gate design, *i.e.* $H_{STRIPE,FinFET} = 0.5*[W_{Tri\text{-}Gate}+2*H_{STRIPE,Tri\text{-}Gate}]$, and the maximum value of W that meets the 100mV/V DIBL specification is selected ($W = 0.6*L_G$). L_{eff} is then adjusted to maximize I_{ON} and to minimize delay. The thickness of the nitride hard-mask layer is assumed to be equal to L_G.

A lower-aspect-ratio tri-gate bulk MOSFET design, with $W = 2*L_G$ and channel stripe pitch = $4*L_G$ (which is the worst case for layout area efficiency and fringing capacitance), is also investigated to find the optimal physical channel width for tri-gate design.

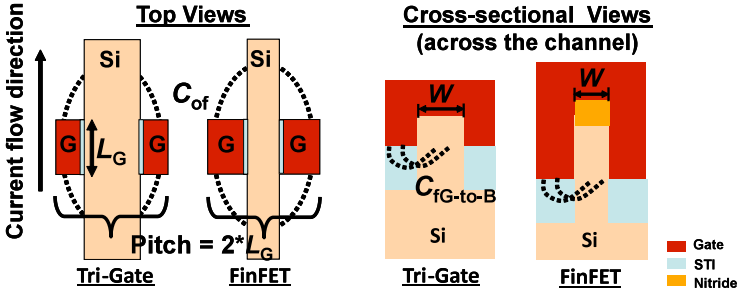
For all designs, source/drain contacts are assumed to be made along the top and sidewall surfaces of the source/drain regions. The gate work function is assumed to be around 4.2 to 4.5eV, which is experimentally achievable [30].

The DC and AC characteristics of the tri-gate and FinFET bulk MOSFETs with same W_{eff} are shown in Fig. 3.11. I_{ON} normalized to W_{eff} is larger for the FinFET design, due to lower average channel doping. Note that the source/drain lateral doping gradient (2 nm/dec) [15] and retrograde well doping gradient (4 nm/dec) are assumed to be the same for the tri-gate and FinFET designs. Since the doping profiles in a FinFET likely would be less abrupt in practice (because it is more difficult to form a steep retrograde profile and abrupt source/drain profiles at the base of a tall stripe), I_{ON} is overestimated for the FinFET. 3-D simulated total gate capacitance is also larger for the FinFET design, due to larger fringing capacitances, as illustrated in Fig. 3.12. Since the FinFET has a narrower and taller stripe, the thickness and height of the gate electrode along the channel-stripe sidewalls are larger for the FinFET design, resulting in larger outer fringing capacitance (C_{of}). Similarly, the gate-to-substrate coupling capacitance is also larger for the FinFET design.



(a) (b)

Figure 3.11. Comparison of simulated characteristics for bulk MOSFETs: (a) DC characteristics, (b) AC characteristics. L_G = 18nm, t_{ox} = 9A, V_{DD} = 0.7V, L_{eff} = 22nm.

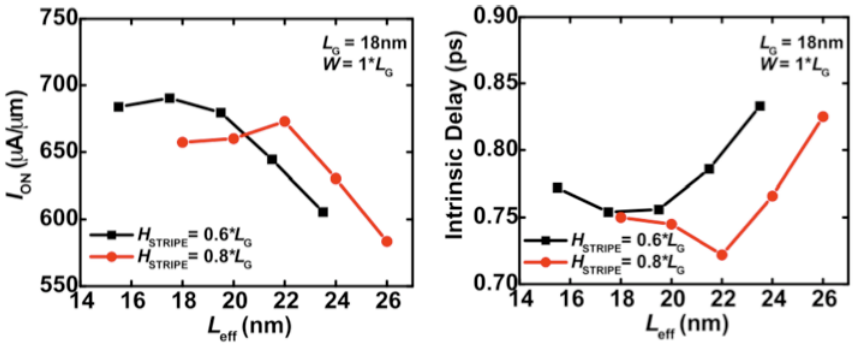


(a) (b)

Figure 3.12. Illustration of parasitic capacitances in multi-gate MOSFET structures. (a) Outer fringing capacitance C_{of}, (b) Gate-to-substrate capacitance $C_{fG\text{-}to\text{-}B}$.

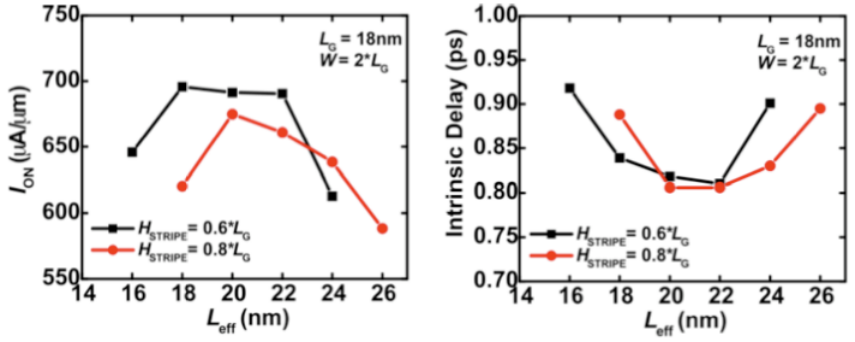
Figs. 3.13a and 3.13b show the L_{eff} dependence of I_{ON} and intrinsic delay, respectively, for different values of H_{STRIPE}, with $W = 1*L_G$, channel stripe pitch = $2*L_G$. It can be seen from Fig. 3.13b that the optimal design with larger H_{STRIPE} has larger L_{eff} (wider gate-sidewall spacers, smaller C_{of}) and smaller minimum intrinsic delay. The tri-gate bulk MOSFET design with $W = 2*L_G$ and channel stripe pitch = $4*L_G$ shows poorer intrinsic delay (Fig. 3.14b) due to degraded I_{ON}, consistent with a larger scale length [29]. For the FinFET design ($W = 0.6*L_G$, channel stripe pitch = $2*L_G$), a larger value of H_{STRIPE} is advantageous to achieve smaller delay, since the ratio of C_{of} over total gate capacitance is smaller (Fig. 3.15b). Based on the results in Figs. 3.13, 3.14, and 3.15, the tri-gate bulk MOSFET design with $W = 1*L_G$ is optimal to minimize intrinsic delay. Fig. 3.16 show the effect of retrograde channel doping gradient on MOSFET performance. Note that the retrograde doping gradients that are used in the simulations range from 5 nm/dec and aggressively down to 1 nm/dec, for the purpose of finding the optimal doping gradient value. In practice, steep retrograde doping profiles (3 nm/dec ~ 4 nm/dec) can be achieved by utilizing diffusion-barrier layers [31]. For the tri-gate design, there exists an optimal value of retrograde doping gradient that optimizes the trade-off between mobility and SCE suppression; for the FinFET design, which does not rely on retrograde channel doping to suppress SCE, the retrograde doping gradient generally should be as steep as possible in order to minimize mobility degradation. Therefore, the optimal

retrograde channel doping gradient for the tri-gate design is less steep than that for the FinFET design, and is easier to attain in practice.



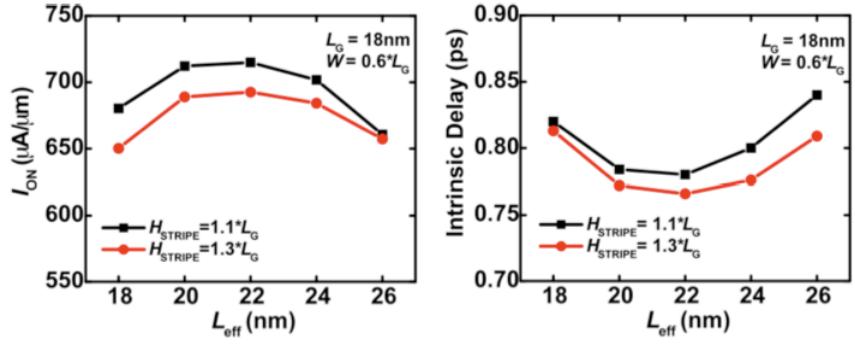
(a) (b)

Figure 3.13. Tri-gate bulk MOSFET design optimization for channel stripe width W = L_G. (a) I_{ON} *vs*. L_{eff}, for I_{OFF} = 7 nA/μm, (b) Delay *vs*. L_{eff}. L_G = 18nm, t_{ox} = 9A, $X_{J,SD}$ = 1.2*H_{STRIPE}, t_{Si} = H_{STRIPE}, V_{DD} = 0.7V.



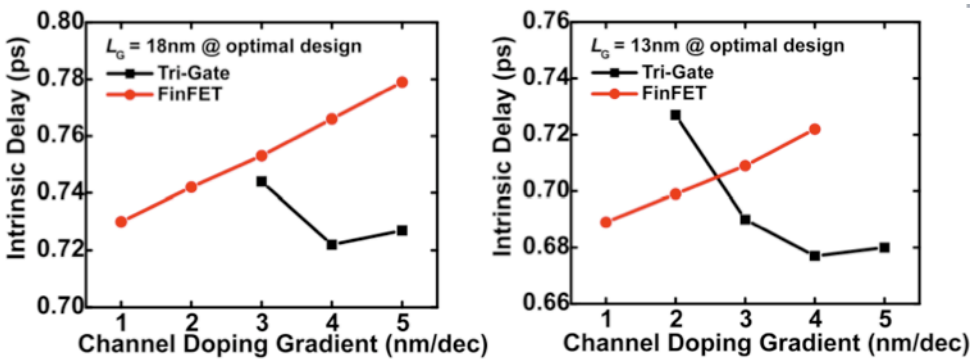
(a) (b)

Figure 3.14. Tri-gate bulk MOSFET design optimization for channel stripe width W = 2*L_G. (a) I_{ON} *vs*. L_{eff}, for I_{OFF} = 7 nA/μm, (b) Delay *vs*. L_{eff}. L_G = 18nm, t_{ox} = 9A, $X_{J,SD}$ = 1.2*H_{STRIPE}, t_{Si} = H_{STRIPE}, V_{DD} = 0.7V.



(a) (b)

Figure 3.15. FinFET bulk MOSFET design optimization. (a) I_{ON} *vs*. L_{eff}, for I_{OFF} = 7 nA/μm, (b) Delay *vs*. L_{eff}. L_G = 18nm, W = 0.6*L_G, t_{ox} = 9A, $X_{J,SD}$ = 1.2*H_{STRIPE}, t_{Si} = H_{STRIPE}, V_{DD} = 0.7V.

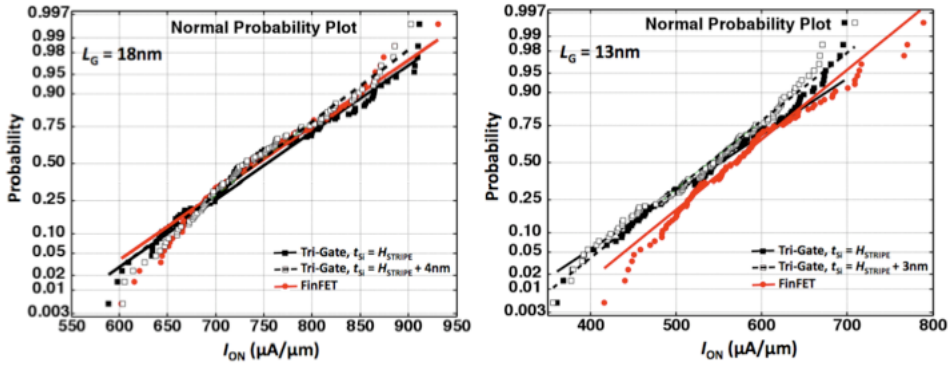


(a) (b)

Figure 3.16. Impact of retrograde channel doping gradient on the performance of tri-gate bulk and FinFET bulk MOSFETs. For the tri-gate design, $W = 1*L_G$, $H_{STRIPE} = 0.8*L_G$; for the FinFET design, $W = 0.6*L_G$, $H_{STRIPE} = 1.3*L_G$. (a) L_G = 18nm, t_{ox} = 9A, $t_{Si} = H_{STRIPE}$, V_{DD} = 0.7V, (b) L_G = 13nm, t_{ox} = 8A, $t_{Si} = H_{STRIPE}$, V_{DD} = 0.6V.

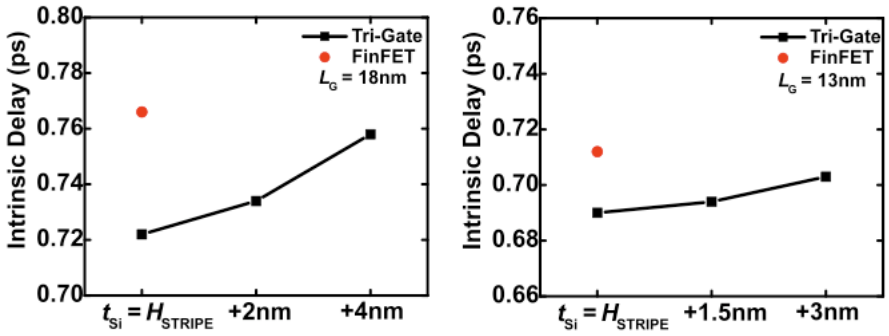
3.4.2 Variability Study

The impact of gate line edge roughness (LER) together with random dopant fluctuations (RDF) on bulk tri-gate and FinFET designs is studied using Kinetic Monte Carlos simulations [14]. The LER value used in the simulations is 4nm. For a fair comparison, the nominal structures for tri-gate and FinFET designs are the optimal structures (with minimum intrinsic delay) with the same source/drain lateral doping gradient (2nm for L_G = 18nm, 1.5nm for L_G = 13nm) and retrograde well doping gradient (4nm for L_G = 18nm, 3nm for L_G = 13nm). Figs. 3.17a and 3.17b show I_{ON} distributions for L_G = 18nm and L_G = 13nm, respectively. As compared to FinFET designs, the I_{ON} variations are larger for tri-gate designs with $t_{Si} = H_{STRIPE}$ due to higher average channel doping. Comparable I_{ON} variations can be achieved by making t_{Si} larger than H_{STRIPE} in tri-gate designs while still maintaining smaller intrinsic delay than FinFET designs (Fig. 3.18).



(a) (b)

Figure 3.17. I_{ON} distributions due to LER and RDF. (a) L_G = 18nm, t_{ox} = 9A, V_{DD} = 0.7V, (b) L_G = 13nm, t_{ox} = 8A, V_{DD} = 0.6V.



(a) (b)

Figure 3.18. Tri-gate design intrinsic delay dependence on t_{Si}. (a) L_G = 18nm, t_{ox} = 9A, V_{DD} = 0.7V, (b) L_G = 13nm, t_{ox} = 8A, V_{DD} = 0.6V.

Table 3.1 summarizes the parameters for the optimized tri-gate and FinFET bulk MOSFET designs. The tri-gate bulk MOSFET design offers comparable (or even less) delay, with less aggressive features (stripe width, stripe aspect ratio, retrograde doping gradient) for improved manufacturability.

	Parameter	FinFET		Tri-Gate	
Design	L_G (nm)	18	13	18	13
	T_{ox} (nm)	0.9	0.8	0.9	0.8
	Gate work function (eV)	4.51	4.45	4.35	4.34
	L_{eff} (nm)	22	17	22	17
	W/L_G	0.6	0.6	1	1
	H_{STRIPE}/L_G	1.3	1.3	0.8	0.8
	Retrograde gradient (nm/dec)	1	1	4	4
Performance (V_{DD} = 0.7V for L_G = 18nm, V_{DD} = 0.6V for L_G = 13nm)	I_{ON} (μA/μm) (for I_{OFF} = 7 nA/μm)	754	592	673	531
	CV/I_{eff} (ps)	0.73	0.69	0.72	0.68

Table 3.1. Optimal designs for the tri-gate and FinFET bulk MOSFETs.

3.5 Summary

The tri-gate bulk MOSFET structure is proposed to provide an evolutionary pathway for continued CMOS scaling. An analytical equation for the scale length of a tri-gate bulk MOSFET is derived. As compared with the double-gate SOI MOSFET, the tri-gate bulk MOSFET design is more scalable and less sensitive to device design parameters. As

compared with the bulk FinFET design, the tri-gate bulk MOSFET design offers comparable performance and variability. Its low-aspect-ratio channel structure is favorable for ease of manufacturing. Thus, the tri-gate bulk MOSFET is a promising structure for CMOS transistor scaling to the end of the technology roadmap.

3.6 References

[1] S. Borkar, "Circuit techniques for subthreshold leakage avoidance, control, and tolerance," *IEDM Tech. Dig*., pp. 421-424, 2004.

[2] H. Masuda, S. Ohkawa, A. Kurokawa, and M. Aoki, "Challenge: variability characterization and modeling for 65- to 90-nm process," *Custom Integrated Circuits Conf*., pp. 593-599, 2005.

[3] Y.-K. Choi, K. Asano, N. Lindert, V. Subramanian, T.-J. King, J. Bokor, and C. Hu, "Ultra-thin-body SOI MOSFET for deep-sub-tenth micron era," *IEEE Electron Device Lett.*, vol. 21, pp. 254–255, May 2000.

[4] 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 Lett.*, vol. 22, pp. 487–489, October 2001.

[5] B.S. Doyle, S. Datta, M. Doczy, S. Hareland, B. Jin, J. Kavalieros, T. Linton, A. Murthy, R. Rios, and R. Chau, "High performance fully-depleted tri-gate CMOS transistors," *IEEE Electron Device Lett.*, Vol. 24, pp. 263-265, April 2003.

[6] M. Jurczak, T. Skotnicki, M. Paoli, B. Tormen, J. Martins, J. L. Regolini, D. Dutartre, P. Ribot, D. Lenoble, R. Pantel, and S. Monfray, "Silicon-on-Nothing (SON)-an innovative process for advanced CMOS," *IEEE Trans. Electron Devices*, Vol. 47, pp. 2179-2187, November 2000.

[7] T. Park, S. Choi, D.H. Lee, J.R. Yoo, B.C. Lee, J.Y. Kim, C.G. Lee, K.K. Chi, S.H. Hong, S.J. Hynn, Y.G. Shin, J.N. Han, I.S. Park, U.I. Chung, J.T. Moon, E. Yoon, and J.H. Lee, "Fabrication of body-tied FinFETs (Omega MOSFETs) using bulk Si wafers," *Symp. VLSI Tech. Dig*., pp. 135-136, 2003.

[8] T.-S. Park, H. J. Cho, J. D. Choe, I. H. Cho, D. Park, E. Yoon, and J.H. Lee, "Characteristics of body-tied triple-gate pMOSFETs," *IEEE Electron Device Lett.*, vol. 25, pp. 798-800, December 2004.

[9] T.-S. Park, H. J. Cho, J. D. Choe, S. Y. Han, D. Park, K. Kim, E. Yoon, and J.H. Lee, "Characteristics of the full CMOS SRAM cell using body-tied TG MOSFETs (bulk FinFETs)," *IEEE Trans. Electron Devices*, Vol. 53, pp. 481-487, March 2006.

[10] U.S. Patent #7,190,050.

[11] R.-H. Yan, A. Ourmazd, and K.F. Lee, "Scaling the Si MOSFET: from bulk to SOI to bulk," *IEEE Trans. Electron Devices*, Vol. 39, pp. 1704-1710, July 1992.

[12] K. Takeuchi, T. Tatsumi, and A. Furukawa, "Channel engineering for the reduction of random-dopant-placement-induced threshold voltage fluctuation," *IEDM Tech. Dig*., pp. 841-844, 1997.

[13] A. Asenov, and S. Saini, "Suppression of random dopant-induced threshold voltage fluctuations in sub-0.1μm MOSFET's with epitaxial and δ-doped channels," *IEEE Trans. Electron Devices*, Vol. 46, pp. 1718-1724, August 1999.

[14] *Sentaurus User's Manual*, v. 2006.06 (Synopsys, Inc.)

[15] *Int'l Technology Roadmap for Semiconductors*, 2006 Ed.

[16] J.G. Fossum, J.-W. Yang, and V.P. Trivedi, "Suppression of corner effects in triple-gate MOSFETs," *IEEE Electron Device Lett.*, vol. 24, pp. 745-747, December 2003.

[17] V. Vidya, L. Smith, S. Balasubramanian, and T.-J. King Liu, "Multi-gate FET design for tolerance to statistical dopant fluctuations," *Proc. IEEE Silicon Nanoelectronics Workshop*, pp. 137-138, 2006.

[18] 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," *IEDM Tech. Dig*., pp. 275-278, 2000.

[19] D. J. Frank, Y. Taur, M. Leong, and H.-S.P. Wong, "Monte Carlo modeling of threshold variation due to dopant fluctuations," *Symp. VLSI Tech. Dig*., pp. 169-170, 1999.

[20] K. Suzuki, T. Tanaka, Y. Tosaka, H. Horie, and Y. Arimoto, "Scaling theory for double-gate SOI MOSFET's," *IEEE Trans. Electron Devices,* Vol. 40, pp. 2326-2329, December 1993.

[21] D. J. Frank, Y. Taur, and H.-S. P. Wong, "Generalized scale length for two-dimensional effects in MOSFET's," *IEEE Electron Device Letters*, pp. 385-387, October 1998.

[22] X. Sun, Q. Lu, V. Moroz, H. Takeuchi, G. Gebara, J. Wetzel, S. Ikeda, C. Shin, and T.-J. King Liu, "Tri-gate bulk MOSFET design for CMOS scaling to the end of the roadmap," *IEEE Electron Device Letters*, pp. 491-493, May 2008.

[23] K. Akarvardar, A. Mercha, S. Cristoloveanu, P. Gentil, E. Simoen, V. Subramanian, and C. Claeys, "A two-dimensional model for interface coupling in triple-gate transistors," *IEEE Trans. Electron Devices*, Vol. 54, pp. 767-775, April 2007.

[24] K. K. Young, "Short-channel effect in fully-depleted SOI MOSFETs," *IEEE Trans. Electron Devices*, vol. 36. pp. 399-402, February 1989.

[25] 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 Trans. Electron Devices*, Vol. 47, pp. 2320-2325, December 2000.

[26] X. Wu, P. C. H. Chan, and M. Chan, "Impacts of nonrectangular fin cross section on the electrical characteristics of FinFET," *IEEE Trans. Electron Devices*, Vol. 52, pp. 63-68, January 2005.

[27] T.-J. King, "FinFET promise and challenges," *Extended Abstracts of the 2003 International Conference on Solid-State Devices and Materials* (Tokyo, Japan), pp. 280-281.

[28] C.-H. Lee, J. M. Yoon, C. Lee, H. M. Yang, K. N. Kim, T. Y. Kim, H. S. Kang, Y. J. Ahn, D. Park, and K. Kim, "Novel body tied FinFET cell array transistor DRAM with

negative word line operation for sub 60nm technology and beyond," *Symp. VLSI Tech. Dig.,* pp. 130-131, 2004.

[29] X. Sun, and T.-J. King Liu, "Scale length assessment of the tri-gate bulk MOSFET design," *IEEE Trans. Electron Devices*, vol. 56, no. 11, pp. 2840-2842, 2009.

[30] Y.-S. Suh, G. P. Heuss, and V. Misra, "Electrical characteristics of $TaSi_xN_y/SiO_2/Si$ structures by Fowler–Nordheim current analysis," *Applied Physics Letters*, vol. 80, pp. 1403-1405, February 2002.

[31] A. Hokazono, H. Itokawa, N. Kusunoki, I. Mizushima, S. Inaba, S. Kawanaka, and Y. Toyoshima, "Steep channel & halo profiles utilizing boron-diffusion-barrier layers (Si:C) for 32 nm node and beyond," *Symp. VLSI Tech. Dig.,* pp. 112-113, 2008.

Chapter 4

Variation Study of Tri-Gate Bulk MOSFET

4.1 Introduction

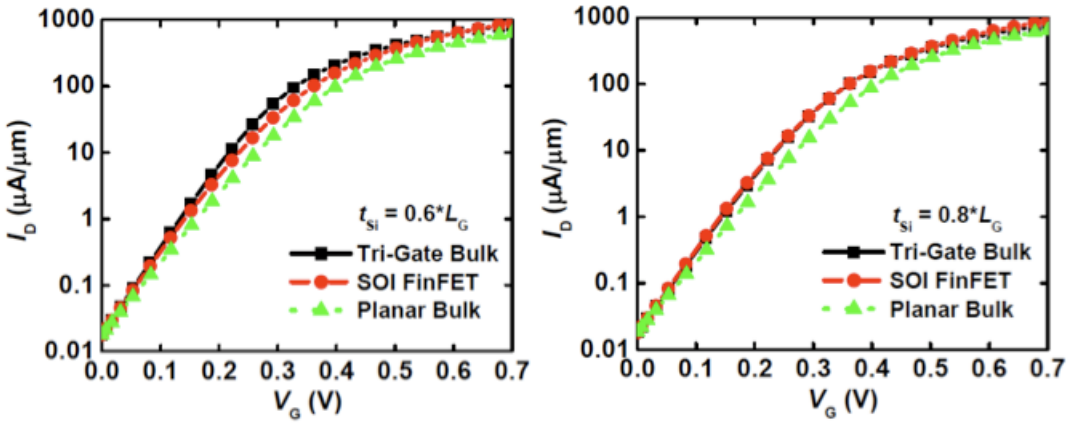
Variability in transistor performance increases dramatically as the critical dimension (the gate length) is reduced well below 50nm. This poses critical challenges for continued CMOS scaling and cost-effective utilization of scaled technologies [1]. Variations in transistor performance can be divided into two main categories: systematic variation and random variation. Sources of systematic variation include process-induced variations in gate length (L_G), channel width (W_{STRIPE}), gate-oxide thickness (t_{ox}) variation and layout-dependent channel stress. Sources of random variation are gate line edge roughness (LER) and random dopant fluctuations (RDF), which are also intrinsic sources of variation [2] (*i.e.* fundamental to the transistor architecture and manufacturing processes).

As candidate structures to extend transistor scaling to its ultimate limit, the planar ground-plane bulk MOSFET, SOI FinFET and tri-gate bulk MOSFET structures were discussed in Section 3.3 [3-6]. A comparison of the scale lengths for these structures shows that the tri-gate bulk MOSFET is more scalable than the other two designs [6]. Besides the comparative performance analysis of transistor nominal designs, a comparison of variability in transistor performance must also be considered to determine which MOSFET structure is the most scalable.

4.2 Nominal MOSFET Designs

The device structures are identical to those described in Fig. 3.5a, except that a realistic retrograde channel doping profile (Fig. 3.1, peaked at a depth t_{Si} with 4nm/dec gradient [7, 8]) is used in the Sentaurus 3-dimensional (3-D) device simulations [9].

Simulations were performed for the following n-channel MOSFET nominal designs: physical gate length L_G = 20nm, equivalent oxide thickness t_{ox} = 9Å, supply voltage V_{DD} = 0.7V. These design parameters are based on the International Technology Roadmap for Semiconductors (ITRS) [10]. For the tri-gate bulk MOSFET design, silicon stripe width $W_{STRIPE} = L_G$, two sets of t_{Si} (equal to oxide recess depth H_{STRIPE}) are simulated (t_{Si} = 0.6*L_G and 0.8*L_G), and source/drain extension junction depth X_J = 1.2*H_{STRIPE}; for the planar bulk MOSFET design, same values of W_{STRIPE} and t_{Si} are assumed, X_J = 7nm [10]; for the SOI FinFET design, W_{STRIPE} (which is labeled as t_{Si} in Fig. 3.5a) is set to 0.6*L_G to suppress short channel effects (SCE), fin height H_{STRIPE} (which is labeled as W in Fig. 3.5a) is chosen to achieve the same effective channel width (W_{eff}) as the tri-gate bulk MOSFET design. For each design, the effective channel length (L_{eff}) is optimized to minimize intrinsic delay C_{total}*V_{DD}/I_{eff}, where C_{total} is the total gate capacitance, and I_{eff} is the average of the drain current I_D for $V_{GS} = V_{DD}$ and $V_{DS} = V_{DD}/2$ and I_D for $V_{GS} = V_{DD}/2$ and $V_{DS} = V_{DD}$, at a fixed 18nA/μm off-state leakage current (I_{OFF}) as specified for planar bulk MOSFETs [10]. The I_D-V_G characteristics of the optimal designs at t_{Si} = 0.6*L_G and 0.8*L_G are shown in Fig. 4.1. The tri-gate bulk MOSFET design with t_{Si} = 0.6*L_G shows the best subthreshold swing (S) due to superior gate control. The parameters for the optimized tri-gate bulk MOSFET, planar bulk MOSFET, and SOI FinFET are summarized in Table 4.1. Only one set of parameters are shown for SOI FinFET, as there is no difference in the optimal SOI FinFET design parameters. The optimal tri-gate bulk MOSFET design shows about 7% less delay than the optimal SOI FinFET and 35% less delay than optimal planar bulk MOSFET design.



(a) (b)

Figure 4.1. Simulated MOSFET DC characteristics. (a) t_{Si} = 0.6*L_G, (b) t_{Si} = 0.8*L_G.

	Parameter	Tri-Gate Bulk		Planar Bulk		FinFET SOI
Design	t_{Si}/L_G	0.6	0.8	0.6	0.8	1.1 or 1.3
	L_G (nm)	20	20	20	20	20
	T_{ox} (Å)	9	9	9	9	9
	Gate work function (eV)	4.265	4.367	4.064	4.171	4.519
	L_{eff} (nm)	21	21	21	21	25
	W/L_G	1	1	1	1	0.6
	Retrograde gradient (nm/dec)	4	4	4	4	No channel doping
Performance (V_{DD} = 0.7V)	I_{ON} (μA/μm) (for I_{OFF} = 18 nA/μm)	820	778	629	647	889
	CV/I_{eff} (ps)	0.65	0.65	1	1	0.7

Table 4.1. Optimal designs for the tri-gate bulk MOSFET, planar bulk MOSFET, and SOI FinFET.

4.3 Impact of Systematic Variations

In this section, the effects of systematic variations in L_G, W_{STRIPE} (for tri-gate and FinFET structures), t_{ox} and channel stress distribution (for tri-gate and planar MOSFETs) are discussed.

Fig. 4.2 shows threshold voltage (V_T) sensitivity to L_G variation. At $t_{Si} = 0.6*L_G$, V_T variation in tri-gate bulk MOSFET design is the smallest among the three structures. However, tri-gate bulk MOSFET V_T variation increases to the same level as that of the planar bulk MOSFET design, due to the weaker gate control and deeper X_J. (X_J for planar bulk MOSFET is fixed at 7nm.)

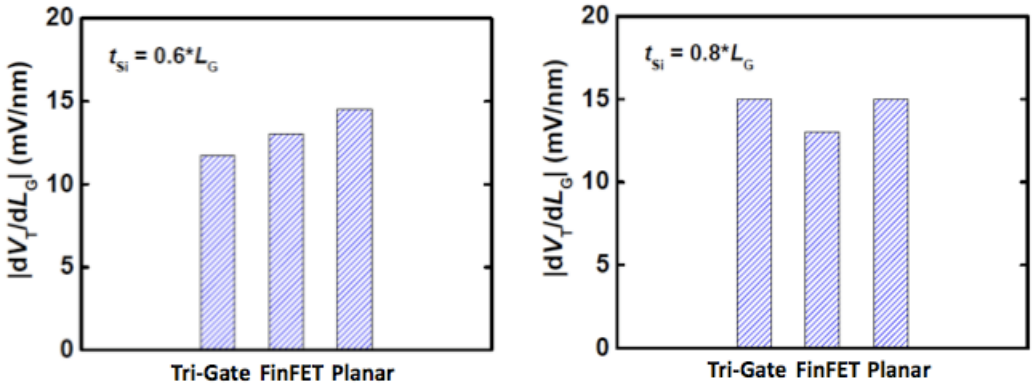


Figure 4.2. Simulated V_T sensitivity to L_G variation. (a) $t_{Si} = 0.6*L_G$, (b) $t_{Si} = 0.8*L_G$.

For the multi-gate structures (tri-gate and FinFET), V_T is dependent on W_{STRIPE} since the side gates influence the channel potential. Since the FinFET relies on thin W_{STRIPE} to suppress SCE, it is much more sensitive to W_{STRIPE} variation than the tri-gate bulk MOSFET design which employs retrograde channel doping to help suppress SCE, as shown in Fig. 4.3. Note that the V_T sensitivity to W_{STRIPE} is calculated at nominal W_{STRIPE} with ±10% variation. The tri-gate bulk MOSFET design with t_{Si} = 0.8*L_G exhibits smallest (close to zero) V_T sensitivity to W_{STRIPE} variations, since the top gate constitutes a smaller portion of total gate area as compared to the design with t_{Si} = 0.6*L_G.

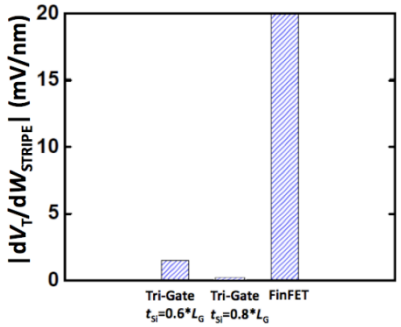
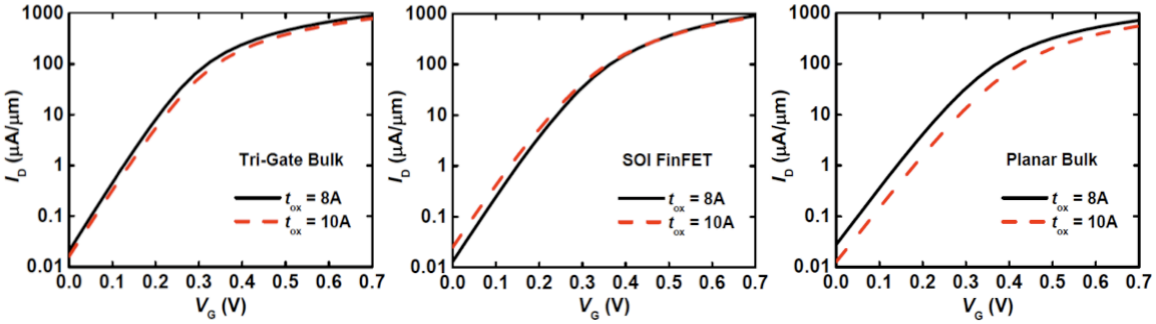


Figure 4.3. Simulated V_T sensitivity to W_{STRIPE} variation.

The effects of t_{ox} variation for t_{Si} = 0.6*L_G are shown in Fig. 4.4. The simulation results for t_{Si} = 0.8*L_G show a similar trend. For tri-gate and planar bulk MOSFET designs, V_T increases (thus I_{OFF} and I_{ON} decrease) with increasing t_{ox}. Because the gate electrode wraps around the channel region in the tri-gate bulk design, the amount of depletion charge (due to ionized channel dopants) per unit channel width is smaller than for a planar bulk device; thus the V_T shift induced by t_{ox} variation is smaller for the tri-gate design. The SOI FinFET design shows a different performance change: as t_{ox} increases, I_{OFF} increases due to worse electrostatic integrity and I_{ON} decreases due to worse substhreshold slope. This is because the SOI FinFET structure relies on good gate control (*i.e.* a combination of ultra-thin t_{ox} and thin t_{Si}) to suppress SCE, whereas the planar and tri-gate bulk designs both employ retrograde channel doping to help improve electrostatic integrity.



(a) (b) (c)

Figure 4.4. Simulated I_D-V_G characteristics with t_{ox} variation, t_{Si} = 0.6*L_G. (a) Tri-gate bulk MOSFET design, (b) SOI FinFET design, (c) planar bulk MOSFET design.

Additional process-induced systematic variations, such as that due to layout-dependent channel stress induced by shallow-trench isolation (STI) oxide, have become significant for scaled CMOS technologies [11]. Fig. 4.5 compares transverse stress profiles and STI-induced hole mobility variations for a planar bulk MOSFET *vs*. a tri-gate bulk MOSFET. Since the surface of each channel stripe above the STI is elevated, mobility dependence on layout is dramatically reduced.

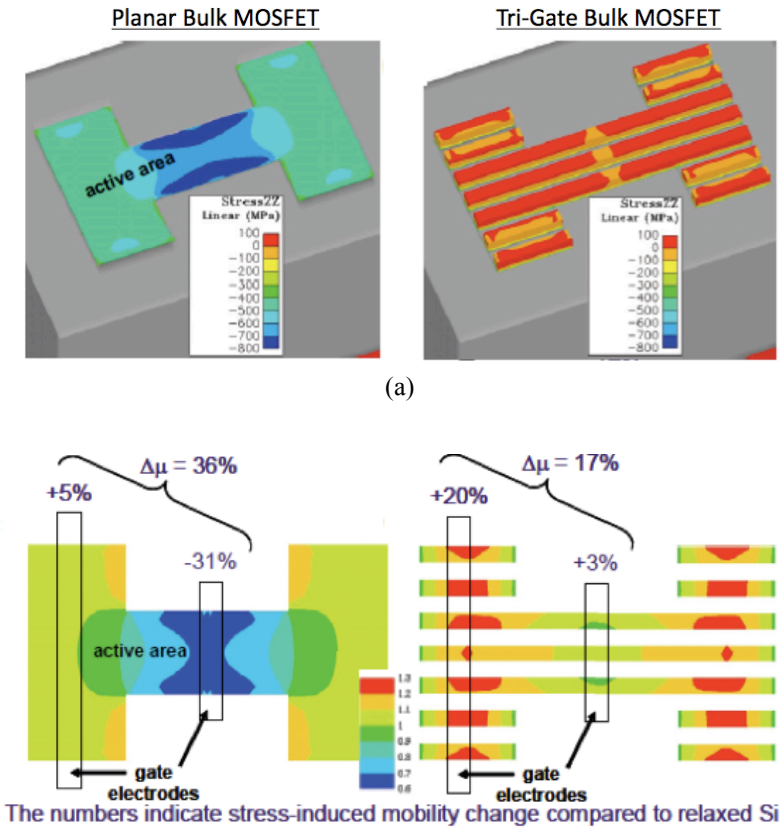


Figure 4.5. Comparison of (a) transverse stress profiles and (b) STI stress-induced hole mobility variations, for planar *vs.* tri-gate (W_{STRIPE} = 20nm) bulk MOSFETs, using Taurus-3D [12]. The silicon stripe height is 10nm for the tri-gate bulk MOSFETs, which provides for more uniform channel mobility because of reduced STI-induced channel stress. Courtesy of Victor Moroz (Synopsys, Inc.).

Stressors used to boost transistor performance are another source of systematic variation; an example is the use of contact etch stop liner (CESL) for mobility enhancement. Fig. 4.6 shows the CESL-induced stress distributions along the channel (current flow) direction within the planar and tri-gate bulk MOSFET structures. The narrow and geometrically regular channel structure in the tri-gate bulk MOSFET results in more channel stress (and thus more mobility enhancement) and less variation in channel stress with changes in effective channel width.

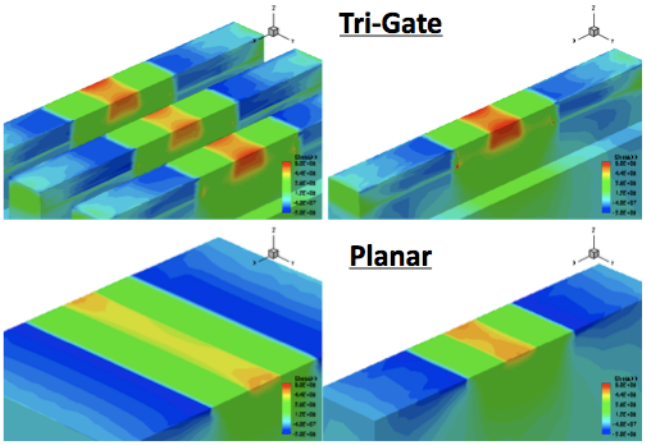


Figure 4.6. CESL-induced stress distribution in planar and tri-gate bulk MOSFETs. The CESL is assumed to be a 30nm-thick silicon nitride with 2GPa tensile stress. L_G = 20nm, t_{ox} = 9Å, gate electrode thickness (T_{GATE}) = 40nm, spacer width (L_{SPACER}) = 20nm, W_{STRIPE} = 20nm, $W_{SPACING}$ = 20, H_{STRIPE} = 10nm.

4.4 Impact of Random Variations

In this section, LER- and RDF-induced variations are compared for the three MOSFET structures. The LER profiles are generated in the same manner as described in Chapter 2. Fig. 4.7 shows an example of tri-gate bulk MOSFET structure with gate LER. The RDF profiles are generated using Kinetic Monte Carlo simulations [9]. The LER and the line width roughness (LWR) values used for the simulations are 4nm and 6.4nm, respectively. 3-D device simulations were performed for the optimal nominal designs (for $t_{Si} = 0.6*L_G$ and $t_{Si} = 0.8*L_G$) to investigate the effects of LER only, and also to investigate the effects of LER together with RDF in the source/drain and channel regions.

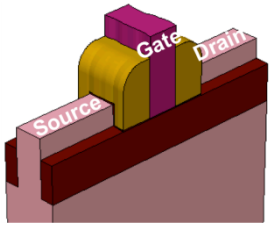
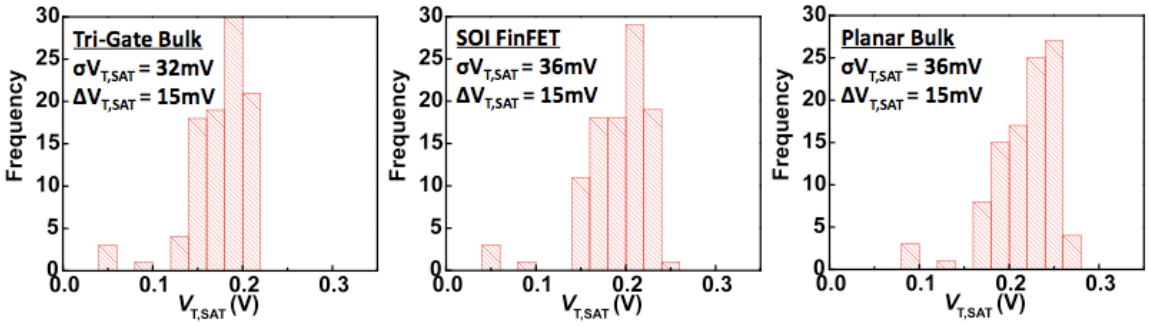


Figure 4.7. Example of a simulated tri-gate bulk MOSFET structure with gate LER.

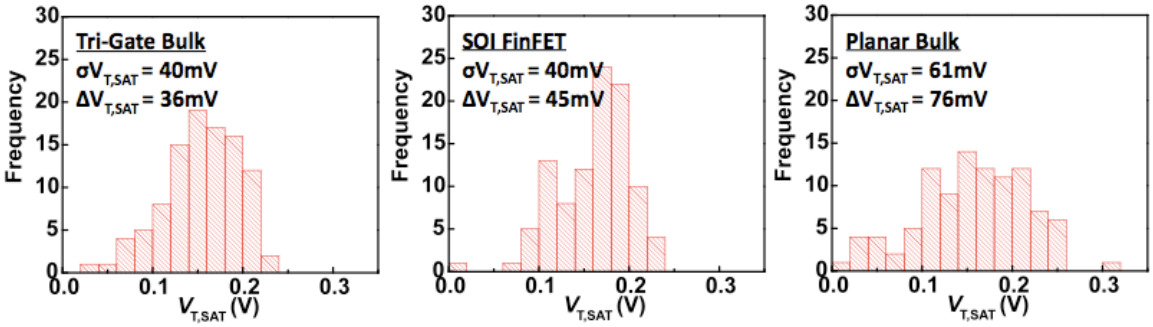
Fig. 4.8 shows the gate LER-induced V_T variations for the SOI FinFET, planar bulk MOSFET and tri-gate bulk MOSFET with $t_{Si} = 0.6*L_G$. The variation is smallest for the tri-gate bulk MOSFET design, due to its good suppression of SCE. Although it has better electrostatic integrity, the SOI FinFET has V_T variation comparable to that of the planar bulk MOSFET. This is because the two sidewall gates of the FinFET have discrete, different gate lengths due to gate LWR, while the effects of gate LWR on planar bulk MOSFETs are somewhat averaged across the channel width. V_T variations due to the presence of both LER and RDF are shown in Fig. 4.9. V_T lowering is smallest for the tri-gate bulk MOSFET due to its superior electrostatic integrity. The tri-gate bulk MOSFET also shows comparable V_T variation as an SOI FinFET. Because V_T variation induced by

the RDF in the source/drain gradient regions increases with smaller W_{STRIPE} [13], and also because of larger LER-induced variation, the SOI FinFET does not provide for reduced random variation as compared to the tri-gate bulk MOSFET design even though there are no channel dopants in the SOI FinFET. Fig. 4.10 shows the effect of t_{Si} on the random variation. For the tri-gate bulk MOSFET, t_{Si} = 0.6*L_G yields the smallest random variation, while for the planar bulk MOSFET, t_{Si} = 0.8*L_G is beneficial for reduced RDF-induced variability since the average number of channel dopants is smaller number.



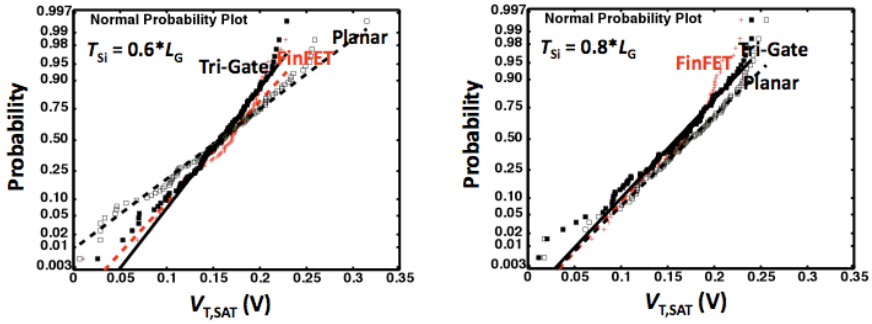
(a) (b) (c)

Figure 4.8. LER-induced variation in (a) tri-gate bulk MOSFET, (b) SOI FinFET, and (c) planar bulk MOSFET. t_{Si} = 0.6*L_G.



(a) (b) (c)

Figure 4.9. LER- and RDF- induced variation in (a) tri-gate bulk MOSFET, (b) SOI FinFET, and (c) planar bulk MOSFET. t_{Si} = 0.6*L_G.



(a) (b)

Figure 4.10. Effect of t_{Si} on LER- and RDF- induced variation. (a) t_{Si} = 0.6*L_G, (b) t_{Si} = 0.8*L_G.

4.5 Summary

The impacts of process-induced systematic and random variations on transistor performance are investigated for three different transistor structures. As compared to the planar bulk MOSFET and SOI FinFET, the tri-gate bulk MOSFET design shows the least variability as well as the best nominal performance. Thus it is a promising device architecture for transistor scaling to the end of the technology roadmap.

4.6 References

[1] H. Masuda, S. Ohkawa, A. Kurokawa, and M. Aoki, "Challenge: variability characterization and modeling for 65- to 90-nm process," *Custom Integrated Circuits Conf.*, pp. 593-599, 2005.

[2] 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 Trans. Electron Devices*, vol. 50, no. 9, pp. 1837–1852, September 2003.

[3] R.-H. Yan, A. Ourmazd, and K.F. Lee, "Scaling the Si MOSFET: from bulk to SOI to bulk," *IEEE Trans. Electron Devices*, Vol. 39, pp. 1704-1710, July 1992.

[4] K. Suzuki, T. Tanaka, Y. Tosaka, H. Horie, and Y. Arimoto, "Scaling theory for double-gate SOI MOSFET's," *IEEE Trans. Electron Devices,* Vol. 40, pp. 2326-2329, December 1993.

[5] D. J. Frank, Y. Taur, and H.-S. P. Wong, "Generalized scale length for two-dimensional effects in MOSFET's," *IEEE Electron Device Letters*, pp. 385-387, October 1998.

[6] X. Sun, and T.-J. King Liu, "Scale length assessment of the tri-gate bulk MOSFET design," *IEEE Trans. Electron Devices*, vol. 56, no. 11, pp. 2840-2842, 2009.

[7] A. Hokazono, H. Itokawa, N. Kusunoki, I. Mizushima, S. Inaba, S. Kawanaka, and Y. Toyoshima, "Steep channel & halo profiles utilizing boron-diffusion-barrier layers (Si:C) for 32 nm node and beyond," *Symp. VLSI Tech. Dig.,* pp. 112-113, 2008.

[8] X. Sun, Q. Lu, V. Moroz, H. Takeuchi, G. Gebara, J. Wetzel, S. Ikeda, C. Shin, and T.-J. King Liu, "Tri-gate bulk MOSFET design for CMOS scaling to the end of the roadmap," *IEEE Electron Device Letters*, pp. 491-493, May 2008.

[9] *Sentaurus User's Manual*, v. 2006.06 (Synopsys, Inc.)

[10] *Int'l Technology Roadmap for Semiconductors*, 2006 Update.

[11] G. Scott, J. Lutze, M. Rubin, F. Nouri, and M. Manley, "NMOS drive current reduction caused by transistor layout and trench isolation induced stress," *IEDM Tech. Dig.*, pp. 827–830, 1999.

[12] *Taurus User's Manual* (Synopsys, Inc.)

[13] V. Varadarajan, L. Smith, S. Balasubramanian, and T.-J. King Liu, “Multi-gate FET design for tolerance to statistical dopant fluctuations,” *Proc. IEEE Silicon Nanoelectronics Workshop*, pp. 137-138, 2006.

Chapter 5

Segmented Bulk MOSFET Technology

5.1 Introduction

As discussed in Chapter 3, the channel region of a tri-gate bulk MOSFET consists of one (*e.g.* SRAM) or more (*e.g.* general purpose logic) parallel segments ("stripes") of equal width (W_{STRIPE}). Fig. 5.1 shows the structure of a multi-stripe tri-gate bulk MOSFET. The stripes in a multi-stripe transistor are isolated by very shallow trench isolation (VSTI) oxide. The VSTI oxide extends to a depth (X_{VSTI}) below the source/drain extension regions, but can be much shallower than the STI oxide that is used to isolate individual transistors, so that there is no need to form high-aspect-ratio stripes. This relatively low-aspect-ratio ($X_{VSTI}/W_{STRIPE} < 2$) is beneficial for ease of fabrication. The gate electrode wraps top portion of each stripe to form a tri-gate structure. Retrograde channel doping (located at the base of the gated stripe region) is used to suppress sub-surface leakage. Within each stripe, the channel- and source/drain-doping profiles are similar to those in a planar bulk MOSFET, as illustrated in Fig. 5.1.

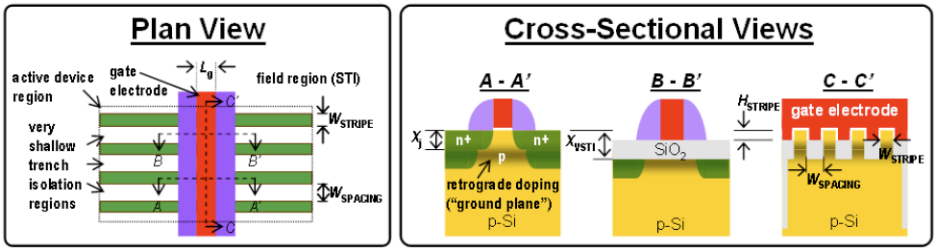


Figure 5.1. Schematic diagrams illustrating an n-channel multi-stripe tri-gate bulk MOSFET structure (adapted from [1]).

The tri-gate bulk MOSFETs can be fabricated using a process that is nearly identical to that for planar bulk MOSFETs, except that a corrugated substrate (comprised of stripes

isolated by VSTI) is used as the starting material. The fabrication process flow is discussed in section 5.2. Sections 5.3 and 5.4 will discuss the most critical processes for tri-gate bulk MOSFET fabrication - formation of corrugated substrate and control of VSTI recess.

5.2 Device Process Flow

The tri-gate bulk MOSFET fabrication process flow is outlined in Fig. 5.2. Starting with a substrate having a corrugated semiconductor surface planarized by VSTI oxide (step 1), the tri-gate bulk MOSFET can be fabricated in a relatively straightforward manner using a conventional process flow. Since the features on the corrugated substrate are geometrically very regular, small-pitch and high-resolution patterning techniques such as multiple patterning or spacer lithography [2] can be readily used to achieve long stripes of uniform width with very fine pitch (details will be discussed section 5.3).

Active areas are defined by photolithography and oxide/silicon etch (step 2). The trenches are filled by oxide to form STI (step 3). After well implantations, the stripe surfaces are slightly elevated above the STI to form a tri-gate structure (step 4). The following steps are straightforward: implant channel and form gate stacks (step 5), form source/drain extension and sidewall spacers (step 6), grow epitaxial material in source/drain regions (step 7, which is optional), and finally dope source/drain regions and form silicide (step 8).

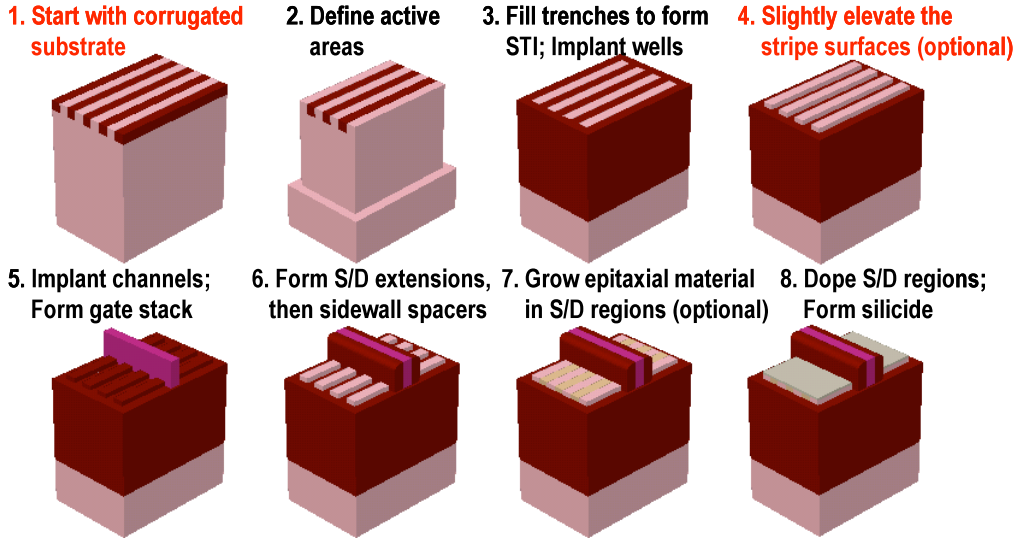


Figure 5.2. SegFET fabrication process (front-end-of-line steps). By defining the channel segments first (prior to active-area definition and gate-stack formation), precise control of segment width (hence channel width) can be achieved.

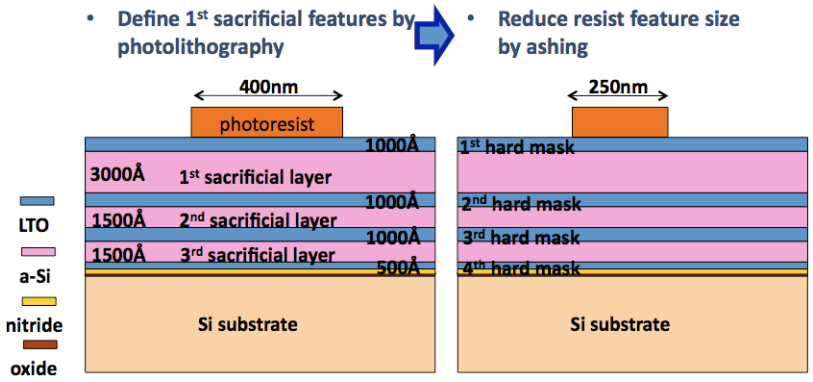
5.3 Corrugated Substrate Process

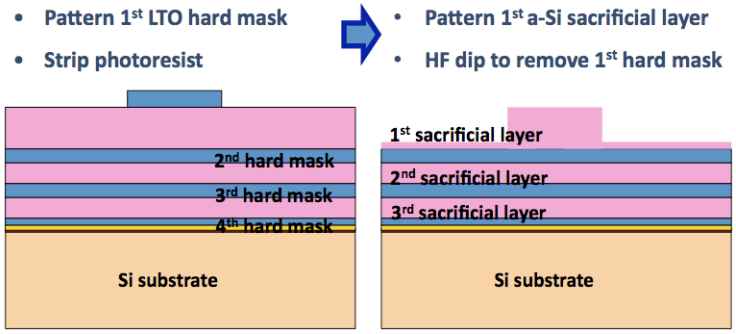
In order to achieve highly uniform stripe width (for tight V_T control), spacer lithography is used to form the corrugated substrate. An exemplary spacer lithography process flow is shown in Chapter 2 (Fig. 2.2). Since the spacer width is determined by the thickness of the deposited spacer layer (very uniform and controllable), it can be much smaller than the minimum feature size defined by conventional lithography, and very uniform.

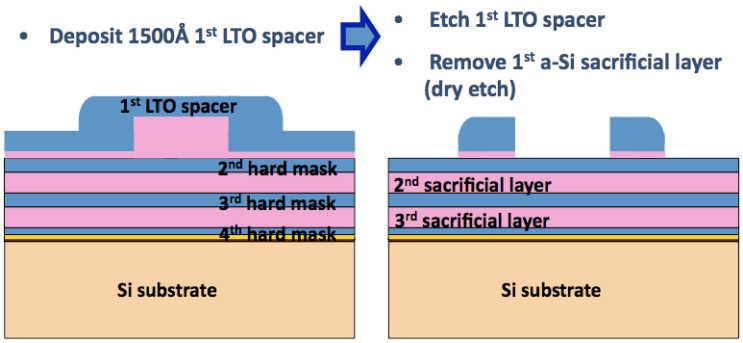
Spacer lithography can be iterated to achieve ultra high densities. With *n* iterations, the pitch reduction is 2^n. Iterated spacer lithography has been demonstrated by forming spacers on the sidewalls of existing spacers [3, 4]. This may introduce additional line width variation. If the etch process is not perfectly anisotropic, the resultant sloped sidewalls result in sloped spacers, and these sloped spacers result in next set of sloped spacers. To mitigate this variation, a multi-tiered hard mask was proposed to ensure every set of spacers is formed along sidewalls with identical slope [5]. Section 5.3.1 will discuss the iterative spacer lithography using this multi-tiered hard mask process to enable 8x pitch reduction with no increase in critical dimension (CD) variation.

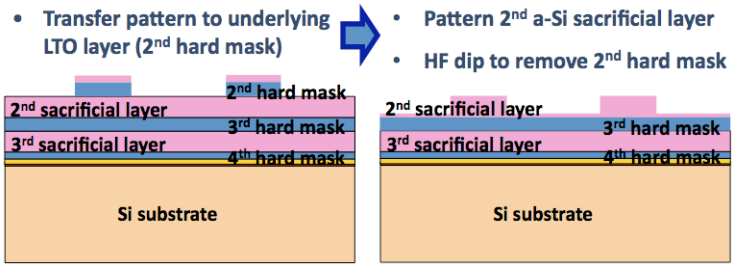
5.3.1 Iterative Spacer Lithography Process

A three-iteration spacer lithography process is illustrated in Fig. 5.3. Starting with the silicon substrate covered by a multi-tiered hard mask (each tier of the hard mask consists of an amorphous silicon (a-Si) sacrificial layer and a thin underlying low temperature oxide (LTO) etch-stop layer), pattern the LTO hard mask layer (1^{st} hard mask) by photolithography (400nm as-printed linewidth, 250nm linewidth after photoresist ashing, 800nm pitch). Then perform a timed anisotropic etch on the a-Si sacrificial layer (1^{st} sacrificial layer) using a chlorine (Cl_2) and hydrogen bromide (HBr) plasma. Note that at least 50nm a-Si is needed to protect the underlying LTO hard mask (2^{nd} hard mask) during the following hydrofluoric acid (HF) etch to remove the remaining portions of the 1^{st} hard mask. A 150nm LTO layer is then deposited and anisotropically etched (in trifluoromethane (CHF_3) plasma) to form the spacers (1^{st} spacer). After removing the a-Si sacrificial layer using a Cl_2 and HBr plasma, transfer the spacer patterns to the underlying LTO hard mask (using a CHF_3 plasma etch). Repeat the sacrificial layer timed etch, HF dip, spacer formation and sacrificial layer removal steps. Thusly, uniform LTO patterns with 100nm pitch are achieved after 3 iterations of spacer lithography, as shown in Fig. 5.4. These LTO lines are then used as a mask to pattern the underlying hard mask layers and silicon to form uniform silicon stripes.


• Define 1st sacrificial features by photolithography
• Reduce resist feature size by ashing
400nm
250nm
photoresist
1000Å
3000Å
1st sacrificial layer
1000Å
1500Å
2nd sacrificial layer
1000Å
1500Å
3rd sacrificial layer
500Å
Si substrate
1st hard mask
2nd hard mask
3rd hard mask
4th hard mask
Si substrate
LTO
a-Si
nitride
oxide


• Pattern 1st LTO hard mask
• Strip photoresist
• Pattern 1st a-Si sacrificial layer
• HF dip to remove 1st hard mask
2nd hard mask
3rd hard mask
4th hard mask
Si substrate
1st sacrificial layer
2nd sacrificial layer
3rd sacrificial layer
Si substrate


• Deposit 1500Å 1st LTO spacer
• Etch 1st LTO spacer
• Remove 1st a-Si sacrificial layer (dry etch)
1st LTO spacer
2nd hard mask
3rd hard mask
4th hard mask
Si substrate
2nd sacrificial layer
3rd sacrificial layer
Si substrate


• Transfer pattern to underlying LTO layer (2nd hard mask)
• Pattern 2nd a-Si sacrificial layer
• HF dip to remove 2nd hard mask
2nd hard mask
2nd sacrificial layer
3rd hard mask
3rd sacrificial layer
4th hard mask
Si substrate
2nd sacrificial layer
3rd hard mask
3rd sacrificial layer
4th hard mask
Si substrate

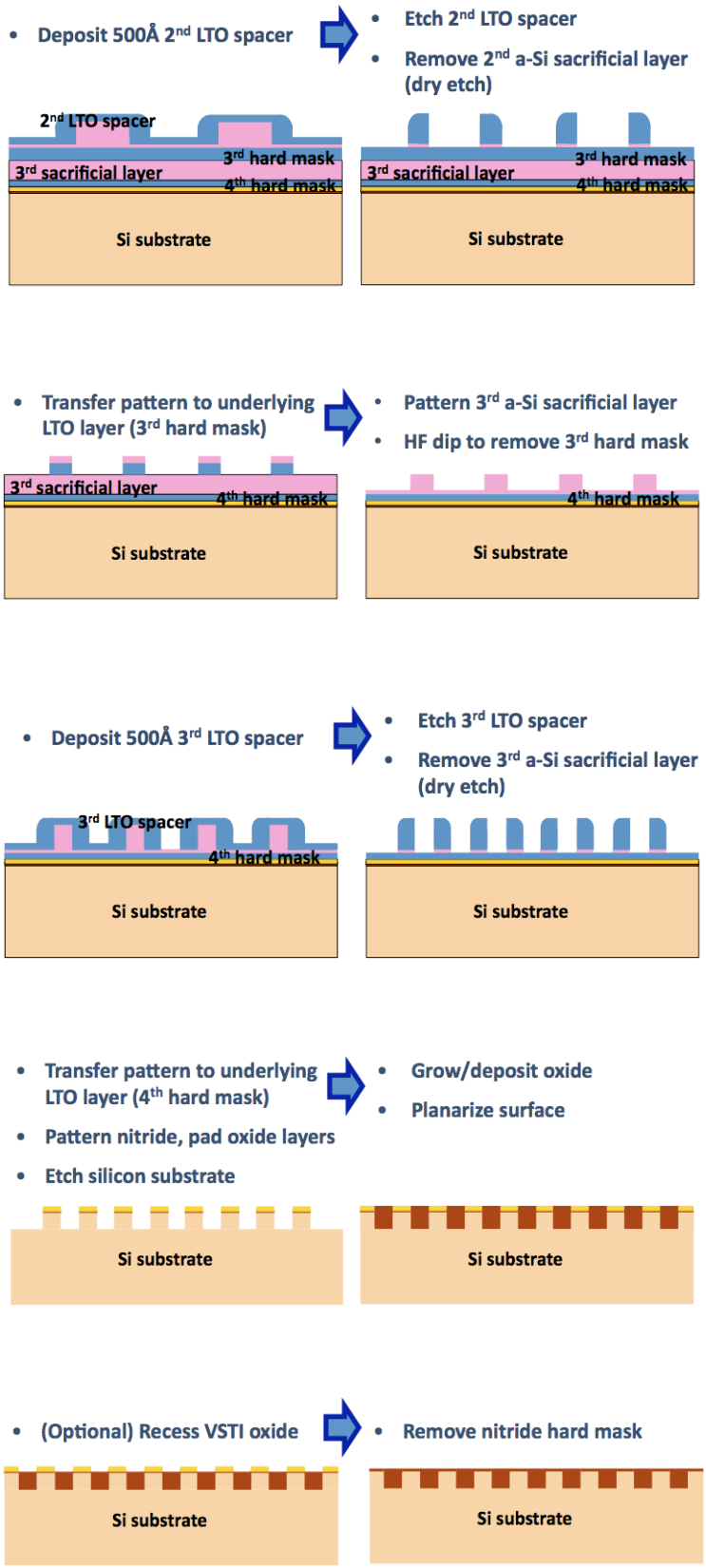


Figure 5.3. Illustration of iterative spacer lithography process (3 iterations).

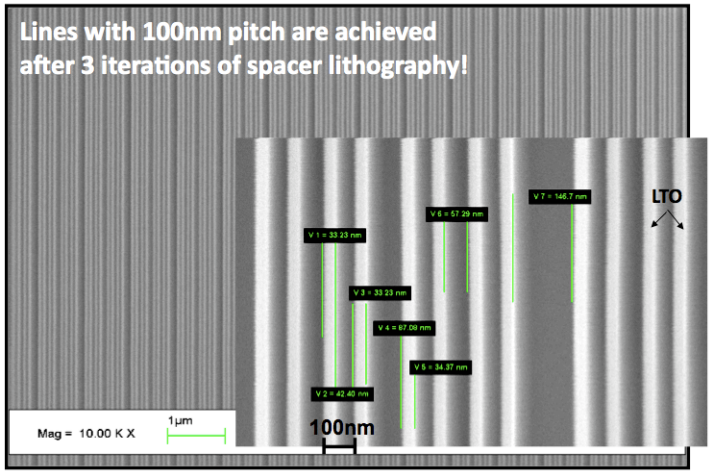


Figure 5.4. Plan-view scanning electron micrograph of LTO spacers formed after 3 iterations of spacer lithography.

5.3.2 Negative Spacer Lithography Process

Selective epitaxial growth is another approach that can be used to form the semiconductor stripes of corrugated substrate. In this case, uniform oxide trenches instead of silicon stripes must be formed. A negative spacer lithography process [5] is used to define these oxide trenches with high uniformity, as illustrated in Fig. 5.5. The negative spacer lithography process is similar to that of conventional spacer lithography, except that the sacrificial material is used as the hard mask.

Starting with a patterned silicon-germanium (SiGe) sacrificial layer (anisotropically etched by Cl_2 and HBr plasma), LTO spacers are formed by CVD and anisotropic etch-back (using CHF_3 plasma). Then another SiGe sacrificial layer is deposited and anisotropically etched to expose the spacers. The LTO spacers are then selectively removed. After transferring the SiGe trenches to underlying thermal oxide layer (in CHF_3 plasma), uniform oxide trenches are formed, as shown in Fig. 5.6a.

Note that the sacrificial material needs to be chosen carefully, since the last step of this negative spacer lithography process is to remove the sacrificial material selectively to oxide as well as silicon. A solution of ammonium hydroxide (NH_4OH), hydrogen peroxide (H_2O_2), and de-ionized water (DI H_2O) has been found to selectively etch SiGe over both oxide and silicon [6]. With the optimal composition of the solution (30% NH_4OH : H_2O_2 : H_2O = 1:1:5) at 75°C, the SiGe is removed selectively (Fig. 5.6b).

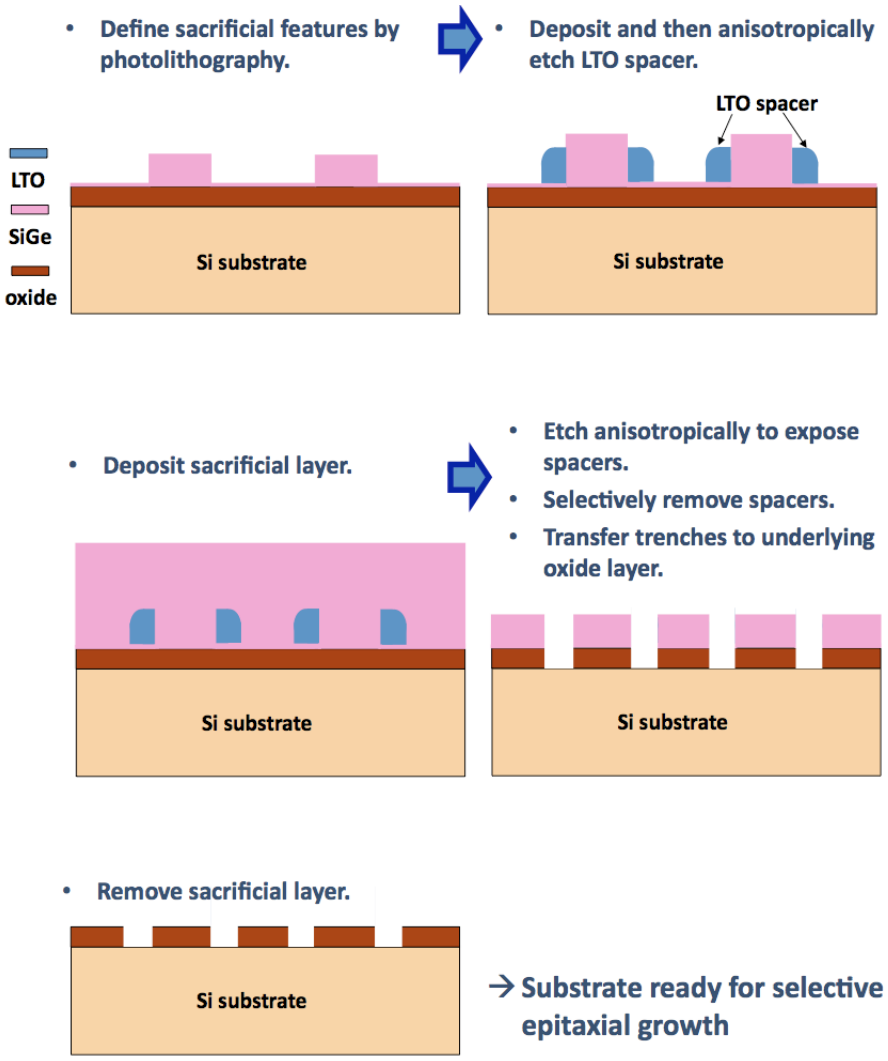
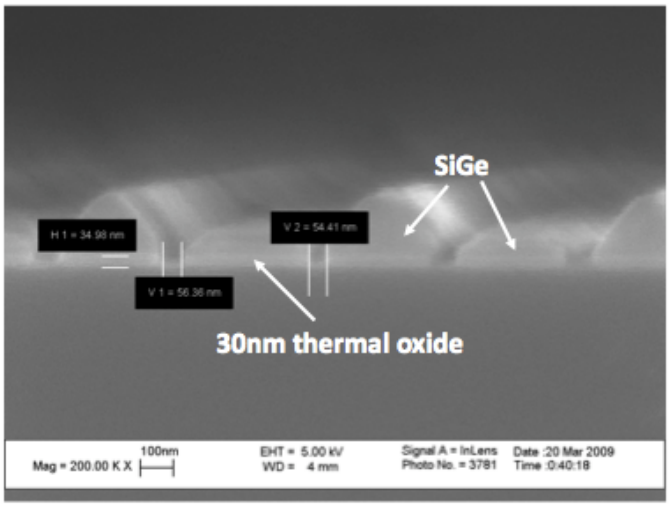
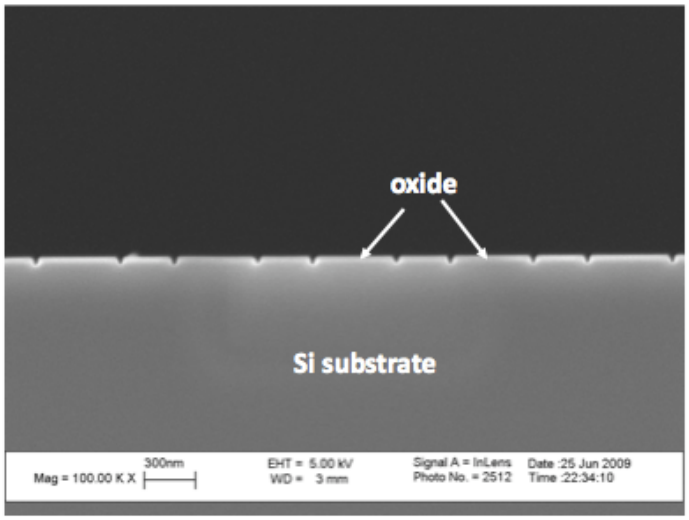


Figure 5.5. Illustration of negative spacer lithography process.



(a)



(b)

Figure 5.6. Cross-sectional-view scanning electron micrograph of oxide trenches formed after negative spacer lithography. (a) Before SiGe sacrificial layer removal. (b) After SiGe sacrificial layer removal.

5.4 Control of Very Shallow Trench Isolation (VSTI) Oxide Recess

5.4.1 Introduction

Starting with the corrugated substrate, the tri-gate structure can be achieved by recessing the isolation oxide in-between the stripes prior to gate-stack formation. Typically, the oxide recess is achieved with a timed etch, so that non-uniformity in etch rate results in commensurate non-uniformity in etch depth. This is because the etch rate of the oxide is uniform throughout its depth. In order to improve the uniformity/controllability of a timed oxide etch process, the oxide etch rate must be enhanced within a depth range that reaches down to the desired etch depth x_D.

Shallow argon (Ar) ion implantation has previously been used to enhance the etch rate within the surface region of a SiO_2 film in order to achieve well-tapered wet-etch profiles [7]. Deep, high-dose ion implantation recently has been reported to enhance the etch rate of thick (>100nm) SiO_2 films in aqueous HF and also in vapor HF [8]. In this section, it is demonstrated that moderate-dose Ar ion implantation can be used to selectively enhance the etch rate of a silicon dioxide (SiO_2) film down to a precise depth, to allow for improved control of oxide etch depth in an integrated-circuit device fabrication process. Lower argon implantation energy than previous reported was used to achieve sharper transition between the etch-rate-enhanced and the unaffected oxide region. An empirical model, fit to experimental data, is used to explain the correlation between the implantation conditions (dose, energy) and the etch-rate enhancement

parameters. It is further shown that this model gives consistent predictions of the etch rate enhancement as the previously published model based on nuclear deposited energy.

5.4.2 Experimental Results

Ion implantation causes structural damage (destruction of Si–O bonds) via nuclear collisions. If the nuclear energy loss exceeds ~10^{23} eV/cm^3 (corresponding to the total Si–O bonding energy in a unit volume of SiO_2), then the wet etch rate of SiO_2 is significantly enhanced [9]. An Ar^+ implant can be used to induce damage above the threshold level required for etch-rate enhancement, down to a certain depth x_D determined by the implant conditions. If the implanted film is subjected to a short etch, such that the etch depth x_d is less than x_D, then the average etch rate enhancement factor is simply $S \equiv x_d/x_0 = r_d/r_0$, where x_0 is the etch depth for an unimplanted film, r_d is the average etch rate of the damaged portion of the implanted film, and r_0 is the average etch rate of undamaged oxide. The short-etch depth data and calculated enhancement factors for etching of thermal SiO_2 films in dilute HF solution (10:1 H_2O:HF) are tabulated in Table 5.1, for various Ar^+ implant conditions.

If the implanted film is subjected to a long etch, such that the etch depth x_d is greater than x_D for the implanted oxide, then it can be shown that $\Delta x \equiv x_d - x_0 = x_D(1 - r_0/r_d)$. Thus x_D can be determined from the difference in long-etch depth for implanted oxide *vs.* unimplanted oxide, with $r_0/r_d = 1/S$ determined from short-etch depth data. The long-etch depth data and calculated values of x_D are also tabulated in Table 5.1.

Ar^+ implant conditions	Short Etch Experiment		Long Etch Experiment	
	Etch depth (Å)	Enhancement factor S	Etch depth (Å)	Damage depth x_D (Å)
1×10^{14} cm^{-2} @ 25 keV	207.0	4.14	651.9	464.0
2×10^{14} cm^{-2} @ 25 keV	247.8	4.96	703.6	505.6
2×10^{14} cm^{-2} @ 35 keV	240.6	4.81	832.7	672.4
None	50.0	-	300.0	-

Table 5.1. Oxide etch depth data and calculated etch-rate enhancement factors and depths, obtained from thermally grown oxide samples (formed by 1050°C wet oxidation) etched in 10:1 H_2O:HF solution. The short-etch and long-etch times were selected to remove 50Å and 300Å of unimplanted oxide, respectively. The oxide thickness was measured by OptiProbe™, before and after etching, to determine the etch depth.

Fig. 5.7 plots x_D against Ar^+ implant energy, for the two doses studied in this work. It can be seen that the depth of etch-rate enhancement is strongly dependent on implant energy, and mildly dependent on implant dose.

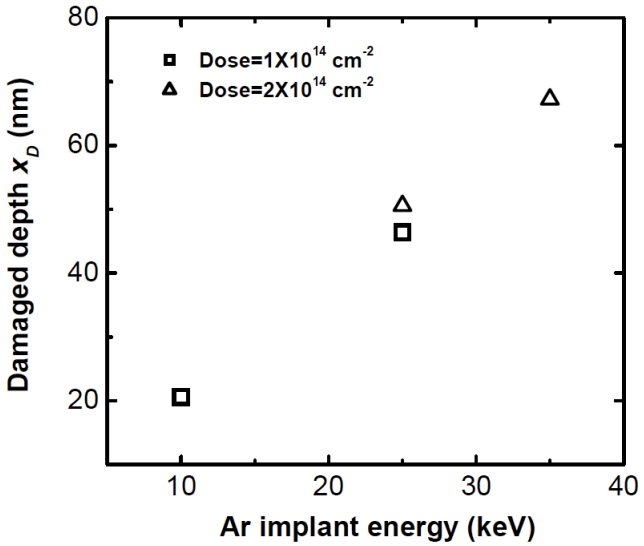
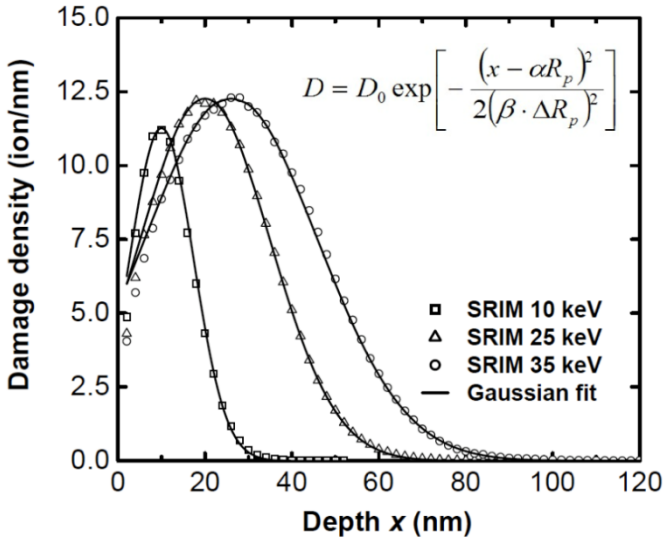


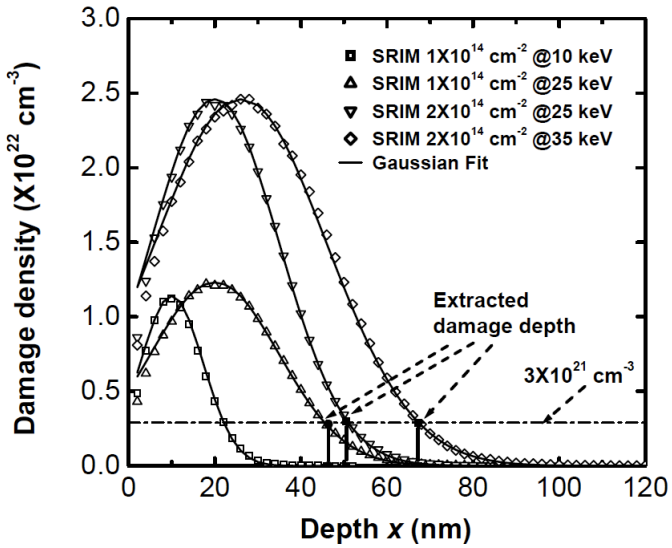
Figure 5.7. Depth of enhanced oxide etch rate (determined from experimental data) *vs*. Ar+ implantation energy, for different doses.

5.4.3 Implant Damage Simulation and Model

In order to clarify the effects of Ar^+ implant energy and dose on x_D and S, SRIM simulations [10] were carried out with full damage cascades. The etch rate enhancement is believed to be due to the broken bonds in SiO_2 as a result of the implantation [8, 9], therefore, it is intuitive to look at the correlation between the damage concentration and the etch rate and depth. The damage profile, *i.e.*, the distribution of displacement events caused by each ion can be empirically modelled by a Gaussian profile using tabulated values of projected range (R_p) and straggle (ΔR_p), as shown in Fig. 5.8a. The damage concentration profile is simply obtained by multiplying this profile by the implanted ion dose. The damage concentration at depth x_D is the threshold level required to significantly enhance the oxide etch rate, and is found consistently to be $\sim 3\times10^{21}$ cm^{-3} (Fig. 5.8b). Assuming that ~75% of the Ar ion energy is lost due to nuclear collisions [8] within the distance x_D, the critical nuclear deposited energy density is $\sim 10^{22} - 10^{23}$ eV/cm^3, which is consistent with previous reports [8, 9]. From the empirical model (Fig. 5.8a inset), it can be derived that x_D is roughly proportional to R_p, which in turn is proportional to the implant energy. Thus, x_D should increase approximately linearly with implant energy as seen in Fig. 5.7. It can also be derived from the model that x_D is a weak function of the peak damage concentration, which is directly proportional to the implant dose. Thus, x_D should show a weak dependence on implant dose as seen in Fig. 5.7.



(a)



(b)

Figure 5.8. Implantation-induced damage distributions obtained from SRIM simulations. (a) The number of displacements per ion can be empirically modeled by a Gaussian profile (inset): D_0 is the peak number of displacement events induced by the ion implantation; R_P and ΔR_P are the projected range and straggle of the implanted Ar distribution, respectively. The fitting parameters are α = 0.67 and β = 1.3. (b) Damage concentration profiles obtained by multiplying the distributions in (a) by the ion dose: the concentration at depth x_D is the threshold level required to significantly enhance the oxide etch rate, and is seen to be ~3×10^{21} cm^{-3}.

Fig. 5.9 shows that the etch-rate enhancement factor S increases linearly with the average damage concentration (calculated based on the empirical model) within the depth x_D. Thus, for a given implant energy (hence x_D), the selectivity of the oxide etch process increases with the Ar^+ implant dose, as seen from the experimental data in Table 5.1. It should be noted that S is expected to saturate at a maximum value of ~5 [9] because not more than ~15% of the Si–O bonds can be broken by Ar ion implantation [11].

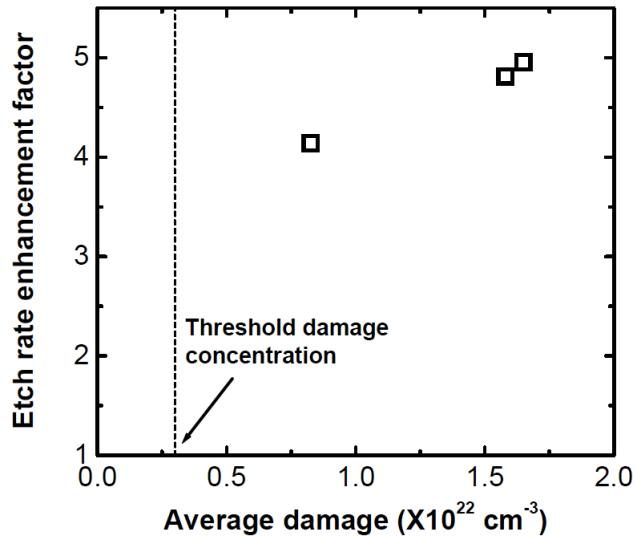


Figure 5.9. The etch-rate enhancement factor increases linearly with the average damage concentration, which can be controlled by adjusting the implant dose.

It also can be seen that a lower implant energy, which has smaller straggle ΔR_P, results in a sharper transition from the etch rate enhanced region to the unaffected oxide region. Fig. 5.10 shows the characteristic width $\beta \cdot \Delta R_P$ for different implant energies. The actual width of the transition region depends on the dose as well as the choice of cut-off damage concentration, and can be a few times the characteristic width. By using a lower implant energy, a narrower transition region and a weaker dose dependence of the etch depth can be achieved, thereby allowing for more precise control of the etch depth.

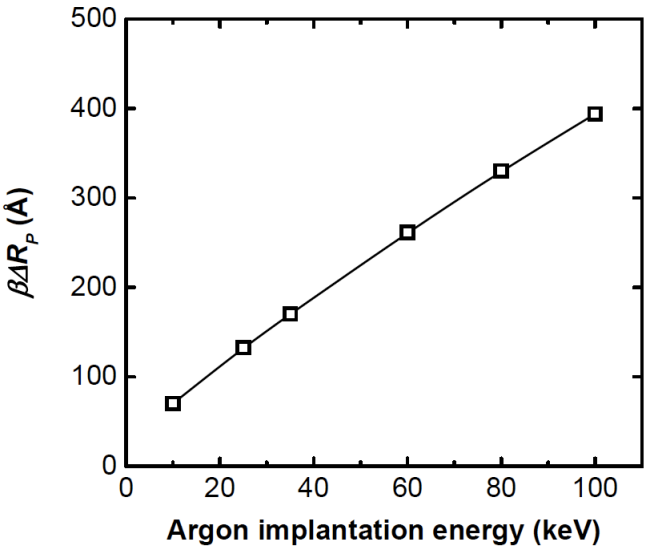


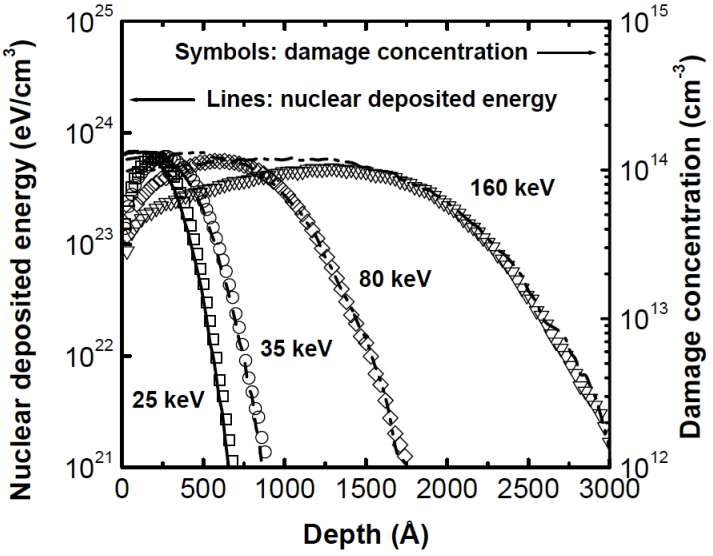
Figure 5.10. The characteristic width of the transition region for different argon implant energies. Lower implant energy results in a sharper transition, thereby a more precise control of the etch depth.

5.4.4 Discussion

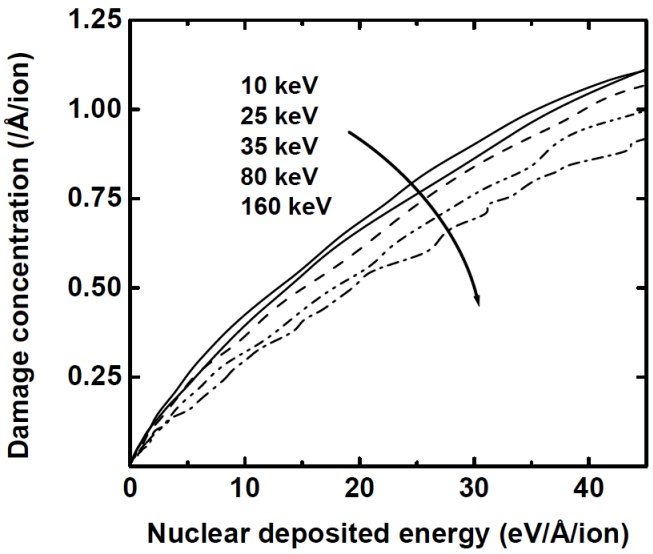
Due to process non-uniformity, the etch time used in practice is typically greater than that required to achieve the nominal etch depth x_D, by a fixed percentage p. Thus, the etch depth can exceed x_D by the nominal amount px_D in some areas. By selectively enhancing the etch rate by a factor S down to the target etch depth x_D, the nominal etch time is reduced by the factor S, hence the over-etch time is reduced by the factor S. Thus, the non-uniformity in etch depth is reduced by the factor S, *i.e.* the nominal variation in etch depth is reduced to px_D/S.

It should be noted that moderate-dose Ar^+ implantation can also cause significant damage to Si. This damage can be healed by moderate thermal annealing to epitaxially re-crystallize the Si from the underlying undamaged Si. For example, in a tri-gate bulk MOSFET fabrication process, after shallow-trench isolation, Ar+ implantation and oxide-recess etching, a re-crystallization anneal should be applied before gate stack formation.

The previously published model describes the etch rate enhancement mechanism in terms of the nuclear deposited energy. In order to compare the two models, the nuclear deposited energy and the damage concentration simulated using SRIM are plotted as a function of the depth for a few different implant energies in Fig. 5.11a. Note that the two quantities are on different scales (left and right axes, respectively). The peak damage location roughly corresponds to the point where the nuclear deposited damage begins to fall off. Beyond this point the nuclear deposited energy and the damage concentration correlate well, and follow a slightly sub-linear relation with fairly weak dependence on the implant energy (Fig. 5.11b). Therefore, it can be expected that the two models can give consistent prediction of the etch rate enhancement depth with some empirical parameter adjustment. The above result is readily explained by the Kinchin-Pease model [12].



(a)



(b)

5.11. (a). The simulated (using SRIM) nuclear deposited energy (left axis) and the damage concentration (right axis) vs. oxide depth for different implant energies. The nuclear deposited energy begin to fall off around the peak damage concentration location, beyond which (b) the two follows a power-law relation, implying that the two models give consistent predictions of the etch depth with adjusted empirical parameters.

To summarize, Ar ion implantation can be used to selectively enhance the wet etch rate of SiO_2 down to a precise depth, to allow for improved control of a timed oxide etch process. The depth of etch-rate enhancement is determined primarily by the implant energy, whereas the etch-rate enhancement factor is controlled by the implant dose. Enhancement factors close to 5 can be achieved to depths in the range 50nm to 100nm with a moderate-dose ($2\times10^{14}/cm^2$) Ar+ implant. Lower implant energy results in more precise control of the etch depth. This technique can be used to improve process uniformity in fabrication of 3-D semiconductor device structures for reduced variations in circuit performance.

5.5 Device Results

N-channel conventional (non-segmented-channel) MOSFETs and prototype tri-gate bulk MOSFETs with L_G = 40nm were fabricated using planar and corrugated p-type (001) bulk-Si wafers, respectively. The corrugated substrates were prepared using a spacer lithography process and dry etching to form stripes of 18nm width (from XTEM analysis) and 90nm average stripe pitch, HDP-CVD to fill the very shallow trenches with SiO_2, and CMP to planarize the surface. A standard STI process was used to define the active device regions, after which standard well and anti-punch-through implants (B) were performed. The wet cleaning process prior to formation of the gate dielectric (SiO_xN_y, 1.7nm EOT) included a dilute-HF dip which caused the STI surfaces to be recessed only slightly (<5nm, from XTEM analysis) below the Si channel surfaces.

Thus, the fabricated segmented bulk MOSFETs are essentially planar devices. A standard Boron channel implant was performed after gate dielectric formation. After gate-electrode formation (100nm-thick n+ poly-Si), standard source/drain (S/D) extension and halo implants were performed. Afterwards, gate-sidewall spacers were formed and deep-S/D implants (Arsenic) were performed. After a spike anneal to activate the implanted dopants (X_J = 23nm), Ni was deposited and annealed to form self-aligned silicide. Device fabrication was completed with standard back-end process steps. Fig. 5.12 shows an XTEM image of a completed MOSFET. Measured segmented MOSFET characteristics (Fig. 5.13) show good SCE control. It should be noted that on-state current I_{ON} is limited by parasitic series resistance due to a non-optimized S/D silicidation process.

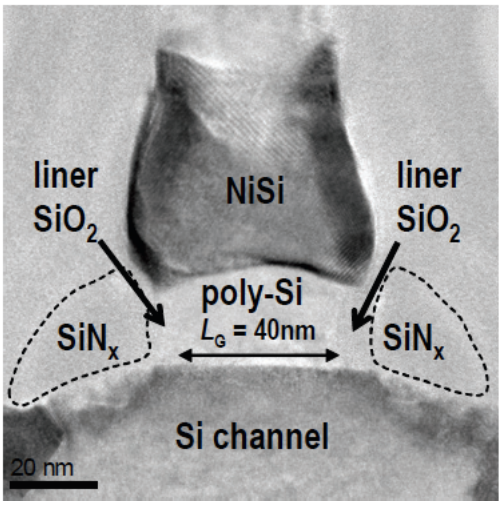
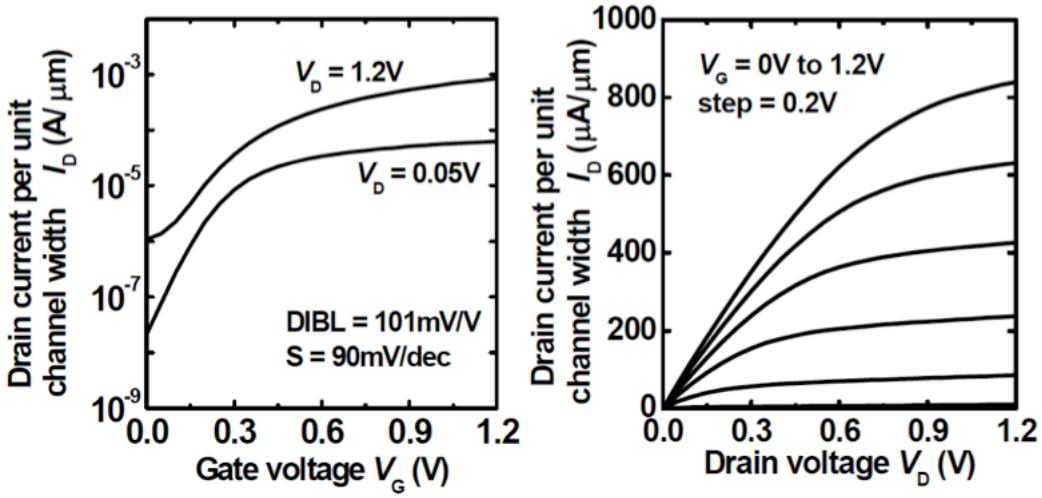
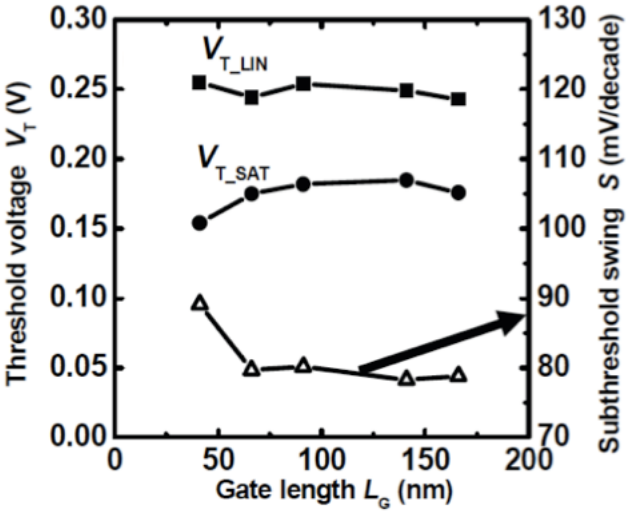


Figure 5.12. Cross-sectional TEM image along the channel of a fabricated tri-gate bulk MOSFET, showing a physical gate length L_G = 40nm.



(a)



(b)

Figure 5.13. (a) Measured *I-V* characteristics of a fabricated tri-gate bulk MOSFET. L_G=40nm, EOT=1.7nm, W_{STRIPE}=18nm, $W_{SPACING}$=72nm, H_{STRIPE}<5nm. (b) Measured V_T and subthreshold swing *vs.* L_G for the fabricated segmented bulk MOSFETs.

Since the width of each channel segment is precisely the same in a segmented bulk MOSFET (because the active area mask is drawn to ensure that the channel region consists of an integer number of stripes), there is minimal V_T variation from stripe to stripe and hence with channel width. Thus, the narrow width effect is eliminated (Fig. 5.14).

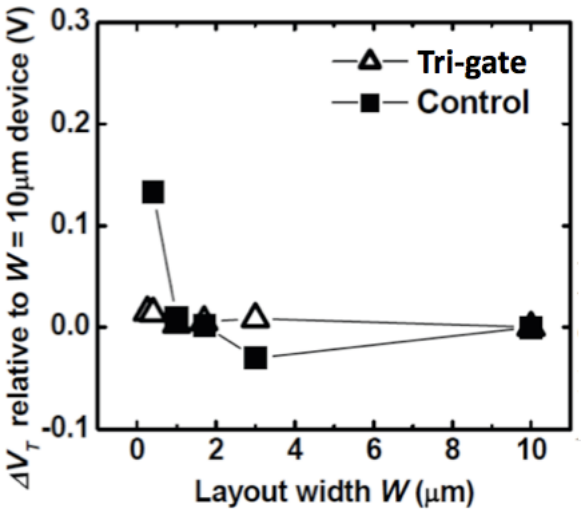


Figure 5.14. Measured V_T *vs.* device layout width for fabricated n-channel FETs. L_G=200nm, EOT=1.7nm.

5.6 Summary

Tri-gate bulk MOSFETs can be fabricated in a relatively straightforward manner using a conventional process flow, starting with a corrugated substrate.

The geometrically regular features on the corrugated substrate make it possible to use spacer lithography to achieve long stripes of uniform width with very fine pitch. A 3-iteration spacer lithography process is demonstrated to be effective to form silicon stripes of uniform width with small pitch. A negative spacer lithography process is also demonstrated to provide a corrugated substrate with equally wide oxide trenches for selective epitaxial growth of semiconductor stripes.

Argon ion implantation can be used to selectively and dramatically enhance the etch rate of a silicon dioxide film down to a precise depth, to allow for improved control of etch depth in a timed oxide etch process. An empirical model based on damage concentration, fit to experimental data, is used to explain the correlation between the implantation conditions (dose, energy) and the etch-rate enhancement parameters. The model is shown to yield consistent predictions of the etch rate enhancement as the existing model based on nuclear deposited energy.

N-channel conventional (non-segmented-channel) MOSFETs and prototype segmented bulk MOSFETs with L_G = 40nm were fabricated. The device results show that the segmented bulk MOSFET has good electrostatic integrity as well as minimal V_T dependence on active area width, as expected.

5.7 References

[1] T.-J. King Liu, and L. Chang, "Transistor scaling to the limit," *Moore's Law: Beyond Planar Silicon CMOS and into the Nano Era*, Springer-Verlag, 2008.

[2] Y.-K. Choi, T.-J. King, and C. Hu, "A spacer patterning technology for nanoscale CMOS," *IEEE Trans. Electron Devices*, Vol. 49, pp. 436-441, March 2002.

[3] Y.-K. Choi, J. S. Lee, J. Zhu, G. A. Somorjai, L. P. Lee, and J. Bokor, "Sublithographic nanofabrication technology for nanocatalysts and DNA chips," *J. Vac. Sci. Tech. B*, pp. 2951-2955, 2003.

[4] R. Rooyackers, E. Augendre, B. Degroote, N. Collaert, A. Nackaerts, A. Dixit, T. Vandeweyer, B. Pawlak, M. Ercken, E. Kunnen, G. Dilliway, F. Leys, R. Loo, M. Jurczak, and S. Biesemans, "Doubling or quadrupling MuGFET fin integration scheme with higher pattern fidelity, lower CD variation and higher layout efficiency," *IEDM Tech. Dig.*, pp. 993-996, 2006.

[5] A. Carlson, and T.-J. King Liu, "Negative and iterated spacer lithography processes for low variability and ultra-dense integration," *Proc. SPIE*, vol. 6924, 2008.

[6] F. S. Johnson, D. S. Miles, D. T. Grider, and J. J. Wortman, "Selective chemical etching of polycrystalline SiGe alloys with respect to Si and SiO^2," *J. Electronic Materials*, vol. 21, pp. 805-810, 1992.

[7] J.C. North, T. E. McGahan, D. W. Rice, A. C. Adams, "Tapered windows in phosphorus-doped SiO_2 by ion implantation," *IEEE Trans. Electron Devices*, vol. 25, no. 7, pp. 809-812, 1978.

[8] R. Charavel and J.-P. Raskin, "Etch rate modification of SiO_2 by ion damage," *Electrochemical and Solid-State Letters*, vol. 9, no. 7, pp. G235-G237, 2006.

[9] A. Hiraiwa, H. Usui, and K. Yagi, "Novel characterization of implant damage in SiO_2 by nuclear-deposited energy," *Applied Physics Letters*, vol. 54, no. 12, pp. 1101-1108, 1989.

[10] J. F. Ziegler, J. P. Biersack, and U. Littmark, "The Stopping and Range of Ions in Solids, SRIM-2003," *http://www.srim.org*, 2003.

[11] B. Garrido, J. Samitier, S. Bota, C. Dominguez, J. Montserrat, and J. R. Morante, "Structural damage and defects created in SiO_2 films by Ar ion implantation," *J. Non-crystalline Solids*, vol. 187, p. 101, 1995.

[12] J. D. Plummer, M. D. Deal, P. B. Griffin, "Silicon VLSI Technology: Fundamentals, Practice and Modeling", Prentice Hall, 2000.

Chapter 6

Conclusion

Transistor scaling has become increasingly difficult in the sub-100 nm regime. Increased leakage current [1] and variability [2] in transistor performance present formidable challenges for continued scaling of bulk-Si CMOS technology. Improvements in transistor architecture and process control are needed to reduce variability.

6.1 Contributions of This Work

This work contributes specifically to the following aspects of reducing variability in transistor performance: improvements in advanced transistor architecture design; understanding of scalability and variability of different MOSFET structures; and investigation of process technology for reduced variability.

The benefit of using a spacer lithography process to mitigate the effect of gate line edge roughness (LER) is assessed using statistical 3-D device simulations. The simulation results indicate that spacer gate lithography is a scalable technology which can dramatically reduce LER-induced variation in transistor performance [3].

A quasi-planar tri-gate bulk MOSFET structure is proposed to provide an evolutionary pathway for continued CMOS scaling [4]. The combination of retrograde channel doping with a multi-gate structure provides for superior electrostatic integrity. The tri-gate bulk MOSFET offers the advantages of a conventional planar bulk MOSFET (low substrate cost, capability for dynamic threshold voltage control, established compact model), along with the advantages of a multi-gate MOSFET (improved performance and scalability).

An analytical equation for the scale length of tri-gate and planar ground-plane bulk MOSFETs is derived, and its sensitivity to device design parameters is investigated [5]. This provides a straightforward way to assess the scaling limit of different transistor designs.

Design optimization and performance of two different multi-gate bulk MOSFET designs (tri-gate bulk MOSFETs [4] vs. bulk FinFETs [6]) are compared to provide guidance for multi-gate bulk MOSFET design.

The impacts of process-induced systematic and random variations on transistor performance are investigated for three candidate device structures to extend transistor scaling to its ultimate limit. The tri-gate bulk MOSFET design shows the least variability as well as the best nominal performance.

The fabrication process flow and the most critical processes (formation of corrugated substrate and precise control of oxide recess) for tri-gate bulk MOSFET fabrication are discussed. Iterative spacer lithography and negative spacer lithography are proven to be effective for defining lines or trenches, respectively, at dense pitch. Ar ion implantation can be used to selectively enhance the wet etch rate of silicon dioxide down to a precise depth, to allow for improved control of a timed oxide etch process. The prototype device results show that the segmented bulk MOSFET (multi-stripe tri-gate bulk MOSFET) has good electrostatic integrity as well as minimal threshold voltage dependence on active area width, as expected.

6.2 Future Directions

Fabrication of advanced transistor structures (*e.g.* tri-gate bulk MOSFET and other candidate structures for sustained CMOS scaling) in a well-controlled fabrication facility is needed to collect statistically meaningful data and explore the pros and cons of different transistor designs.

Although this work only discusses the tri-gate bulk silicon MOSFET structure, the corrugated substrate can be used as a platform to explore post-silicon devices (Ge and III-V channel materials) and to improve the performance and scalability of other transistor designs (accumulation-mode MOSFETs, junction FETs, bipolar transistors, etc.).

6.3 References

[1] S. Borkar, “Circuit techniques for subthreshold leakage avoidance, control, and tolerance,” *IEDM Tech. Dig.*, pp. 421-424, 2004.

[2] H. Masuda, S. Ohkawa, A. Kurokawa, and M. Aoki, "Challenge: variability characterization and modeling for 65- to 90-nm process," *Custom Integrated Circuits Conf*., pp. 593-599, 2005.

[3] X. Sun, and T.-J. King Liu, "Spacer gate lithography for reduced variability due to line edge roughness," *IEEE Trans. Semiconductor Manufacturing*, vol. 23, issue 2, pp. 311-315, 2010.

[4] X. Sun, Q. Lu, V. Moroz, H. Takeuchi, G. Gebara, J. Wetzel, S. Ikeda, C. Shin, and T.-J. King Liu, "Tri-gate bulk MOSFET design for CMOS scaling to the end of the roadmap," *IEEE Electron Device Letters*, pp. 491-493, May 2008.

[5] X. Sun, and T.-J. King Liu, "Scale length assessment of the tri-gate bulk MOSFET design," *IEEE Trans. Electron Devices*, vol. 56, no. 11, pp. 2840-2842, 2009.

[6] C.-H. Lee, J. M. Yoon, C. Lee, H. M. Yang, K. N. Kim, T. Y. Kim, H. S. Kang, Y. J. Ahn, D. Park, and K. Kim, "Novel body tied FinFET cell array transistor DRAM with negative word line operation for sub 60nm technology and beyond," *Symp. VLSI Tech. Dig.,* pp. 130-131, 2004.