NANOFABRICATION TECHNOLOGIES AND NOVEL DEVICE STRUCTURES FOR NANOSCALE CMOS

by

Yang Kyu Choi

Memorandum No. UCB/ERL M01/38

19 December 2001

COVER

NANOFABRICATION TECHNOLOGIES AND NOVEL DEVICE STRUCTURES FOR NANOSCALE CMOS

by

Yang Kyu Choi

Memorandum No. UCB/ERL M01/38

19 December 2001

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Nanofabrication Technologies and Novel Device Structures for Nanoscale CMOS

by

Yang Kyu Choi

B.S. (Seoul National University, Korea) 1989
M.S. (Seoul National University, Korea) 1991
M.S. (University of California, Berkeley) 1999

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 Chenming Hu, Chair
Professor Tsu-Jae King
Professor Steven Louie

Fall 2001

The dissertation of Yang Kyu Choi is approved:


Dec. 4, 2001

Chair Date


12/12/01

Date


12/11/2001

Date

University of California, Berkeley

Fall 2001

Nanofabrication Technologies and Novel Device Structures for Nanoscale CMOS

by

Yang Kyu Choi

Abstract

Nanofabrication Technologies and Novel Device Structures for Nanoscale CMOS

by

Yang Kyu Choi

Doctor of Philosophy in Engineering –
Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Chenming Hu, Chair

This dissertation investigates new patterning technologies and novel device structures for sub-20nm complementary metal-oxide-semiconductor (CMOS). Ashing-trimming and spacer lithography technology for patterning sub-20nm features are investigated.

Ultra-thin body (UTB) metal-oxide-semiconductor field-effect transistors (MOSFETs) are demonstrated and they show excellent suppression of short-channel effects. One of its challenges is the large series resistance of the thin-body silicon. To overcome this difficulty, resist and poly-silicon etch-back process and selective germanium deposition are developed for raised source and drain. Devices with sub-30nm gate lengths, $750\mu A/\mu m$ of NMOS drive current, and $350\mu m/\mu m$ of PMOS drive current are demonstrated.

Thin body silicon can cause a change of sub-bands structure. As a result, threshold voltage shift and mobility enhancement are observed in UTB devices.

Threshold voltage shift of UTB CMOS is modeled analytically and the model is verified with measured data. Mobility enhancement in the thin body is also examined.

Double-gate structure can provide more robustness against the short-channel effects. Simplified planar double-gate FinFETs are fabricated with two different patterning approaches: e-beam lithography and spacer lithography. Spacer lithography technology achieves twice the device density within a given pitch, which is limited by optical or e-beam lithography. It provides more uniform fin width, and ultimately narrower fins than what can be produced with conventional lithography. Devices with features below 60nm and drive current above $1000\mu A/\mu m$ (NMOS) and $760\mu A/\mu m$ (PMOS) have been demonstrated. Chemical-mechanical polishing (CMP) process is developed to overcome process challenges coming from the vertical device structures of FinFETs.

E-beam lithography with subsequent ashing-trimming has produced a 10nm silicon fin width and a sub-20nm gate length, which is the world record smallest transistor. Its NMOS drive current is $730\mu A/\mu m$ and PMOS drive current is $550\mu A/\mu m$. Selective germanium is utilized to fabricate raised source and drain which minimize the parasitic series resistance and improve the drive current.

The dissertation abstract of Yang Kyu Choi is approved:

Chenming Hu Dec. 4, 2001

Professor Chenming Hu Date

Committee Chair

To my family, for their love and encouragement

Table of Contents

Acknowledgements

First of all, I would like to express my deepest gratitude to my advisor Professor Chenming Hu for his guidance, encouragement and support through all the years of my study at University of California-Berkeley. It has been my great fortune to work with him. His keen insight and vast knowledge has guided me through every stage of my work. I am deeply grateful to Professor Tsu-Jae King for not only serving as the chair committee in my qualifying examination but also her insightful advice. The quality of my manuscripts had been dramatically improved for publication with her careful proofreading and comments. I would like to thank Professor Jeffrey Bokor for his advice, invaluable feedbacks, and encouragement during the group meeting.

I would like to thank Professor Steven Louie for serving on my qualifying exam committee and dissertation committee even though he was in the sabbatical period. I would like to express my appreciation for the efforts of Professor Andrew Neureuther, who served on my qualifications.

I thank Kazuya Asano for his co-work and the evaluation of ashing and trimming techniques. I really appreciate for his friendship. I am indebted to Dr. Donggun Park's instructions and invaluable helps in the Microlab. I wish to thank Professor Vivek Subramanian for his instructions on e-beam machine at Stanford University. I appreciate Hideki Takeuchi for his help and discussion on Tystar19, Lam5 etcher, and process issues. Especially, I would like to thank for Hideki Takeuchi and Pushkar Ranade for allowing me to use their experimental results of SiGe. I thank Dr. Yoo-Chan Jeon for

sharing his experience on Tystar19 and Professor Yong-Sang Kim for his helpful discussions and friendship. I am very grateful to Leland Chang, Kevin Yang, Charles Kuo, and Ryan Malloy for their proofreading and feedbacks. Especially, I appreciate Leland Chang for his help in simulation and invaluable discussion.

I would like to thank to Nick Lindert and Stephen Tang for their collaboration to make FinFETs and Leland Chang and Patrick Xuan for their arrangement of e-beam lithography at the Lawrence Berkeley National Laboratory. I thank also to Erik Anderson for his e-beam lithography. I appreciate Daewon for his help in microfabrication and measurement. I would like to thank former and present residents in 373 Cory Hall, Cathy Huang, Sriram Balasubramanian, Yee-Chia Yeo, Igor Polishchuk, Kevin Cao, Mark Cao, Min She, Qiang Lu, Wen-Chin Lee, Ya-Chin King, Jeongsoo Lee, and Kyoungsub Shin, Shiying Xiang, Pin Su, Jane (Xuemei) Xi, and Qing Ji.

I am extremely thankful to Judy Fong for her proofreading on my dissertation and many helps, and Ruth Gjerde for providing all the answers to any of my questions on department policy throughout my study at University of California-Berkeley. I would like to thank to all Microlab staffs for their helps.

I would like to sincerely thank my parents and brothers for their everlasting love and support through my life. I would like to represent my deep thanks to my parents-in-law, brothers-in-law and sisters-in-law for their love.

Finally, to my dear wife, Kyung Choon, I sincerely thank her for everlasting love, sacrificial help, and taking care of my two sons. She has encouraged me and always been on my side. I would like to thank my sons for their love.

Chapter 1

Introduction

1.1 Scaling

Technological advancements in silicon MOSFETs have been achieved over the past three decades primarily through the scaling of device dimensions [1.1]-[1.3] to attain continued improvement in circuit speed and reduction in size. General scaling trends of CMOS (Complementary Metal Oxide Semiconductor) technology are shown in Fig. 1.1 [1.4][1.5].

Three major scaling parameters: power supply voltage, threshold voltage, and gate oxide thickness reduce, as the gate length decreases as shown in Fig.1.1. However, off-state leakage current increases as the gate length decreases as shown in Fig. 1.1. In addition, the threshold voltage cannot be scaled in proportion to the power supply voltage. This is because the drain current does not drop immediately to zero below threshold, but decreases exponentially with a slope in the logarithmic scale inversely proportional to the thermal energy, kT.

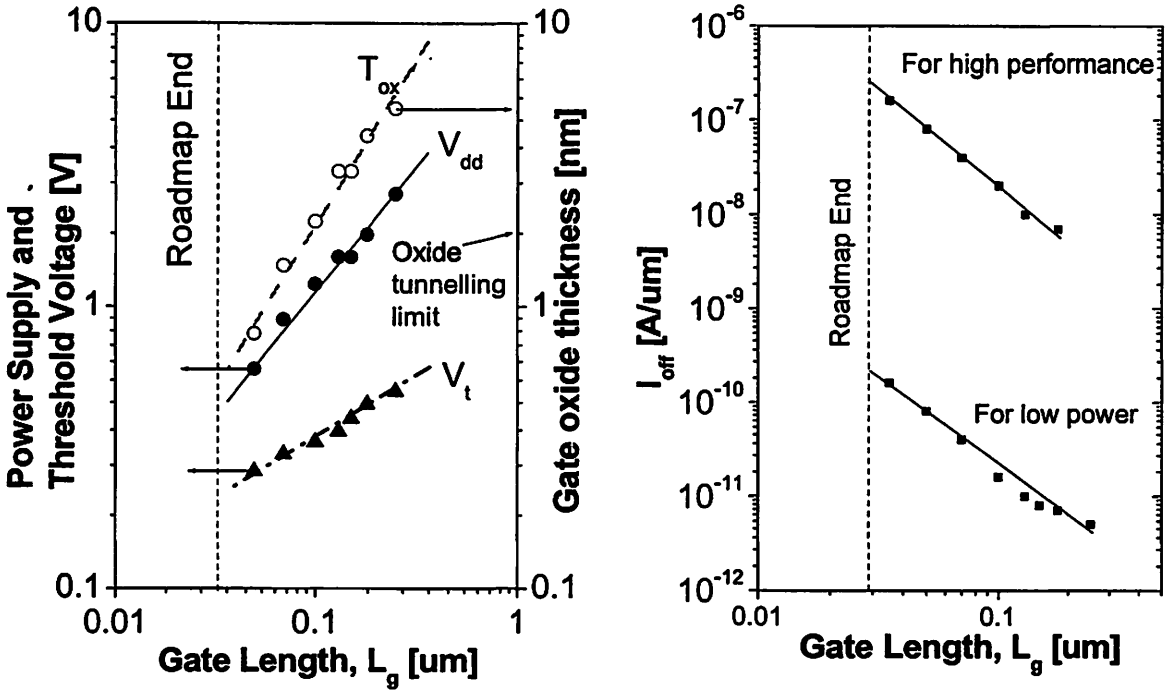


Figure 1.1 Scaling trends of power supply voltage, threshold voltage, and gate oxide thickness versus CMOS gate length [1.5].

Off-state leakage current at zero gate bias should not exceed a few tens of nanoamperes for a chip with an integration level of 100 million transistors. This constraint limits the threshold voltage to a minimum of approximately 0.2V as shown in Fig. 1.2. An ideal subthreshold slope of *60mV/dec* is highly desirable as it may allow for further reduction of the threshold voltage, thus improving on-state performance.

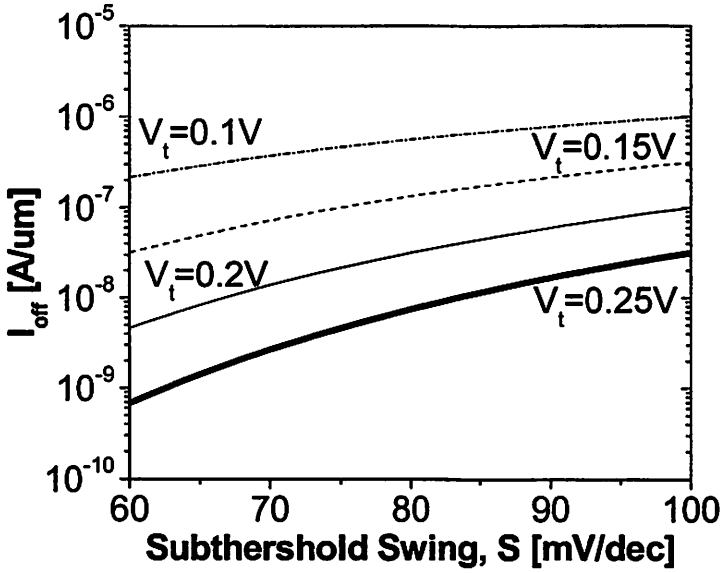


Figure 1.2 Off-state leakage current versus subthreshold swing for different threshold voltages.

As the MOSFET channel lengths are reduced to 30nm and below, suppression of the off-state leakage current becomes an increasingly difficult technological challenge, one that will ultimately limit the scalability of the conventional MOSFET structure. Thus, novel structures have been proposed to extend CMOS scaling beyond the end of the ITRS roadmap.

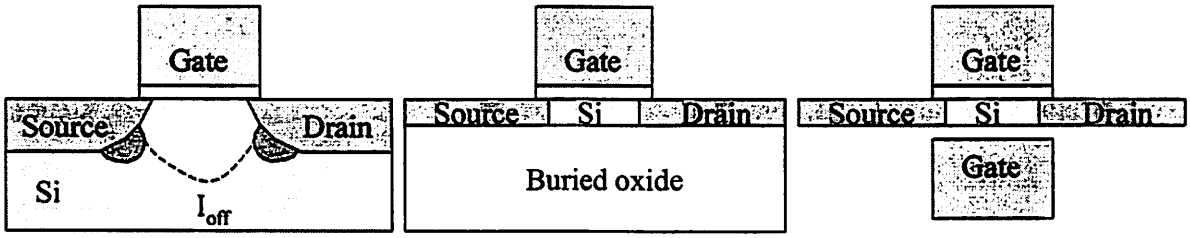
To suppress the short-channel effects, the channel of a MOSFET should be controlled by the gate rather than the drain. Thus, the gate capacitance to the channel should be larger than the drain capacitance. Otherwise, the drain current is determined not by the gate but by the drain, so that the device will behave like a nonlinear resistor instead of a transistor [1.6]. In a conventional bulk MOSFET, the gate capacitance can be increased by reducing the gate oxide thickness. The ITRS [1.5] predicts that an

equivalent oxide thickness (EOT) of 0.5~0.6nm (2 layers of SiO_2) will be needed for 35nm technology generation. In accordance with this rule of thumb, Yu et al. [1.7] reported 35nm gate length MOSFETs with 0.7nm EOT and Chau et al. [1.8] announced a 30nm gate length MOSFETs with 0.8nm T_{ox}. However, for a further scaling beyond 35nm, we cannot arbitrarily scale the oxide thickness as the oxide tunneling current will become too large [1.4]. To alleviate this problem, one solution is to use gate dielectric materials with permittivities higher than that of SiO_2 [1.9]-[1.11]. However, CMOS process compatibility and reliability are still issues for these high-permittivity (high-K) gate dielectrics. It is thus uncertain as to how much scaling is left with traditional bulk silicon technology.

1.2 Novel Device Structures

With an ultra-thin gate dielectric, the channel potential at Si-SiO_2 interface is well controlled by the gate. However, the dominant leakage path is far from this interface, which is least effectively controlled by the gate (Fig. 1.3 (a)). One solution to suppress this leakage current is to eliminate any parts of the channel that are not effectively modulated by the gate - i.e. to remove the bottom portion of the silicon. The ultra-thin body single-gate MOSFET (Fig. 1.3 (b)) and the ultra-thin body double-gate MOSFET (Fig 1.3 (c)) have thus been proposed to allow for device scaling beyond the roadmap. Both are distinctly different from their bulk-Si counterpart in that no current conduction path between the source and drain is far removed from a controlling gate electrode. The gate voltage can therefore effectively control the electric potential throughout the channel,

without the need for a high channel dopant concentration and high gate capacitance. The depth of the source and drain junctions is naturally limited to the thin body thickness so that the formation of ultra-shallow source and drain junctions is not an issue. As a result, many challenges in the scaling of bulk-Si MOSFETs can be circumvented through the adoption of thin-body transistor structures.



(a) Traditional bulk-Si MOSFET (b) Ultra-thin body MOSFET (c) Ultra-thin body double-gate

Figure 1.3 Cross-sectional schematic diagrams showing several novel device structures.

Thin body SOI MOSFETs are highly attractive for suppression of the short-channel effects, but they can also introduce a large series resistance due to the thin silicon film. This problem is analogous to that of ultra-shallow junctions in bulk devices because of the conflicting demands of junction depth and low series resistance. One way to decouple junction depth and series resistance is to use a raised source and drain structures [1.12]-[1.13]. Two novel processes to form this structure are proposed in this dissertation : an etched-back raised poly-Si S/D and a selectively deposited raised Ge S/D.

1.3 Novel Process Technologies

Defining sub-50nm features is a difficult task. One option, e-beam lithography, has successfully produced 15nm gates [1.14]. However, the throughput of e-beam lithography using positive resists is too low even for research even though its resolution is high. In contrast, the throughput of chemically amplified resists is high but its resolution is not good enough. So, even with e-beam lithography, it is challenging to obtain sub-30nm patterns. Another option, extreme-ultra-violet (EUV) lithography has generated 38nm period patterns [1.15], but is not readily available yet. Two novel nano-lithography technologies for sub-30nm pattering are proposed in this dissertation: resist ashing plus hard mask oxide trimming and spacer lithography.

Resist ashing with oxygen plasma was first developed for making sub-micron devices from g-line lithography about 13 years ago [1.16]. Since then, it has been widely used to produce smaller features than the resolution limit of lithography [1.17]-[1.19]. Oxide hard mask trimming with diluted HF is also relatively straightforward. The combination of these two techniques makes it possible to fabricate sub-30nm line widths from 0.5um lines defined by i-line lithography.

Similar to Horstmann et al. [1.20], a spacer lithography process technology using a sacrificial layer and a chemical vapor deposition (CVD) spacer layer has also been developed with conventional dry etching. A sacrificial layer is initially patterned with i-line lithography and a conventional anisotropic etch. Then, another thin CVD layer is deposited and etched to form ring-like sidewalls, spacers. These spacers are used as a mask and transferred to the substrate using an anisotropic etch. Thus, the pattern width is defined not by optical or e-beam lithography but by the CVD deposition thickness. In

addition, this spacer lithography technology can yield critical dimension (CD) variations of minimum-sized features that are much smaller than can be achieved by optical or e-beam lithography. Furthermore, it can provide a doubling of device density for a given lithography pitch. One drawback, however, is that only one line width is available in such a process. This can be overcome by combining a conventional masking process with the spacer process. This spacer lithography technology is used to pattern Si-fin structures for double-gate MOSFETs (FinFETs), in which fine line widths and pitch are especially important [1.21][1.22].

To fabricate a self-aligned raised source and drain (S/D) MOSFETs on an ultra-thin body, a resist etched-back poly-silicon process and a selective germanium deposition process are proposed in this dissertation. Despite process complexity, the resist etch-back process is implemented reliably, can be relatively insensitive to surface cleanness, and does not require new equipment. Selective Ge deposition by LPCVD is itself a low temperature process while dopant activation in Ge can also be performed at low temperature. Thus it is compatible with metal gate and high-K gate dielectric technology. Another benefit of the selective Ge process is that it provides an *in-situ* clean for the removal of native oxide due to germane gas flow [1.23]. In this work, the resist etch-back process with poly-Si is used for making the raised S/D on UTB MOSFETs while selective Ge deposition by LPCVD is used for making the raised S/D on UTB MOSFETs and FinFETs. In addition, a nickel germanide process is investigated.

Finally, chemical-mechanical polishing (CMP) [1.24][1.25] is evaluated for gate planarization in the FinFET structure, which suffers from vertical topography. This

provides a depth-of-focus (DOF) margin in lithography and etching process window without stringers and residues. Spacer FinFETs fabricated with CMP are demonstrated.

1.4 References

[1.1] R.H. Dennard, F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassous and A. R. LeBlanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE Journal of Solid-State Circuits*, vol.9, p.256-268, 1974.

[1.2] C. Hu, "MOSFET scaling in the next decade and beyond," *Semiconductor International*, p.105-114, June 1994.

[1.3] B. Davari, "CMOS technology scaling: 0.1 μm and beyond," *IEDM Technical Digest*, p.555-558, 1996.

[1.4] Y. Taur, E.J. Nowak, "CMOS devices below 0.1 um: how high will performance go?", *IEEE IEDM Technical Digest*, p.215-218, 1997.

[1.5] *International Technology Roadmap for Semiconductors*, Semiconductor Industry Association, 2000. http://public.itrs.net/Home.htm

[1.6] C. Hu, "SOI and Nanoscale MOSFETs", *59th Device Research Conference*, p.3-4, 2001.

[1.7] B. Yu, H. Wang, Q. Xiang, J. X. An, J. Jeon, and M.-R. Lin, "Scaling Towards 35nm Gate Length CMOS", *Symposium on VLSI Technology*, p.9-10, 2001.

[1.8] R. Chau, J. Kavalieros, B. Roberds, R. Schenker, D. Lionberger, D. Barlage, B. Doyle, R. Arghavani, A. Murthy, G. Dewey, "30nm Physical Gate Length CMOS

Transistors with 1.0 ps n-MOS and 1.7 ps p-MOS Gate Delays", *IEEE IEDM Technical Digest*, p.45-48, 2001.

[1.9] L. Kang, K. Onishi, Y. Jeon, B. H. Lee, C. Kang, W.-J. Qi, R. Nieh, S. Gopalan, R. Choi, J.C. Lee, "MOSFET devices with polysilicon on single-layer HfO_2 high-K dielectrics", *IEEE IEDM Technical Digest*, p.35-38,2000.

[1.10] A. Chin, Y.H. Wu, S.B. Chen, C.C. Liao, W.J. Chen, "High quality La_2O_3 and Al_2O_3 gate dielectrics with equivalent oxide thickness 5-10 A", *Symposium on VLSI Technology*, p.16-17, 2000.

[1.11] C.H. Lee, H.F. Luan, W.P. Bai, S.J. Lee, T.S. Jeon, Y. Senzaki, D. Roberts, D.L. Kwong, "MOS characteristics of ultra thin rapid thermal CVD ZrO_2 and Zr silicate gate dielectrics", *IEEE IEDM Technical Digest*, p.27-30, 2000.

[1.12] K. Miyano, I. Mizushima, A. Hokazono, K. Ohuchi, Y. Tsunashima, "Low thermal budget elevated source/drain technology utilizing novel solid phase epitaxy and selective vapor phase etching", *IEEE IEDM Technical Digest*, p. 433-6, 2000.

[1.13] A. Srivastava, J. Sun, K. Bellur, R.F. Bartholomew, P. O'Neil, S.M. Celik, C.M. Osburn, N.A. Masnari, M.C. Ozturk, R. Westhoff, B. Fowler, "A 0.18 um CMOS technology for elevated source/drain MOSFETs using selective silicon epitaxy", *Proceedings of the Sixth International Symposium on Ultralarge Scale Integration Science and Technology (ULSI Science and Technology)*, p.571-85, 1997.

[1.14] J. Kedzierski, P. Xuan, E. Anderson, J. Bokor, T.-J. King, C. Hu, "Complementary silicide source/drain thin-body MOSFETs for the 20nm gate length regime," *IEDM Technical Digest*, p. 57-60, 2000.

[1.15] H.H. Solak, D. He, W. Li, S. Singh-Gasson, F. Cerrina, "Exposure of 38nm period grating patterns with extreme ultraviolet interferometric lithography", *Applied Physics Letter,* vol.75, p.2328-30, 1999.

[1.16] J. Chung, M. Jeng, J. E. Moon, A.T. Wu, T.Y. Chan, P.K. Ko, C. Hu, "Deep-submicrometer MOS device fabrication using a photoresist-ashing technique," *IEEE Electron Device Letters*, vol.9, p. 186-8, 1988.

[1.17] M. Tanabe, T. Matsuno, N. Kashiwagi, H. Sakai, K. Inoue, and A. Tamura, "0.1um AlGaAs/InGaAs high electron mobility transistor fabrication by the new method of thinned resist pattern reversed by metal," *Journal of Vacuum Science & Technology B*, vol.14, p.3248-51, 1996.

[1.18] M. Ono, M. Saito, T. Yoshitomi, C. Fiegna, T. Ohguro, H.S. Momose, H. Iwai, "Fabrication of sub-50-nm gate length n-metal-oxide-semiconductor field effect transistors and their electrical characteristics," *Journal of Vacuum Science & Technology B*, vol.13, p.1740-3, 1995.

[1.19] D.Y. Jeon, G.M. Chin, K.F. Lee, R.H. Yan, E. Westerwick, M. Cerullo, "Gate technology for 0.1-um Si complementary metal-oxide-semiconductor using g-line exposure and deep ultraviolet hardening," *Journal of Vacuum Science & Technology B*), vol.12, p. 2800-4, 1994.

[1.20] J.T. Horstmann, U. Hilleringmann, and K.F. Goser, "Matching Analysis of Deposition Defined 50-nm MOSFET's", *IEEE Transactions on Electron Devices,* vol.45, p.299-306, 1998.

[1.21] X. Huang, W.-C. Lee, C. Kuo, D. Hisamoto, L. Cang, J. Kedzierski, E. Anderson, H. Takeuchi, Y.-K. Choi, K. Asano., V. Subramanian, T.-J. King, J. Bokor, C. Hu, "Sub 50-nm FinFET : PMOS," *IEDM Technical Digest,* p. 67-70, 1999

[1.22] S. Tang, L. Chang, N. Lindert, Y.-K. Choi, W.-C. Lee, X. Huang, V. Subramanian, J. Bokor, T.-J. King, C. Hu, "FinFET-A Quasi-Planar Double-Gate MOSFET", *International Solid State Circuits Conference*, p.118-9, 2001.

[1.23] M. Moslehi, "Low-temperature in-situ dry cleaning process for epitaxial layer multiprocessing," *Proceedings of the SPIE - The International Society for Optical Engineering,* vol.1393, p.90-108, 1991

[1.24] L.H. Blake, E. Mendel, "Chemical-mechanical polishing of silicon", *Solid State Technology*, vol.13, p.42-6, 1970.

[1.25] H.-S. Kim, D.-H. Ko, D.-L. Bae, K. Fujihara, H.-K. Kang, " The formation of Ti-polycide gate structure with high thermal stability using chemical-mechanical polishing (CMP) planarization technology", *IEEE Electron Device Letters*, vol.20, p.86-8, 1999.

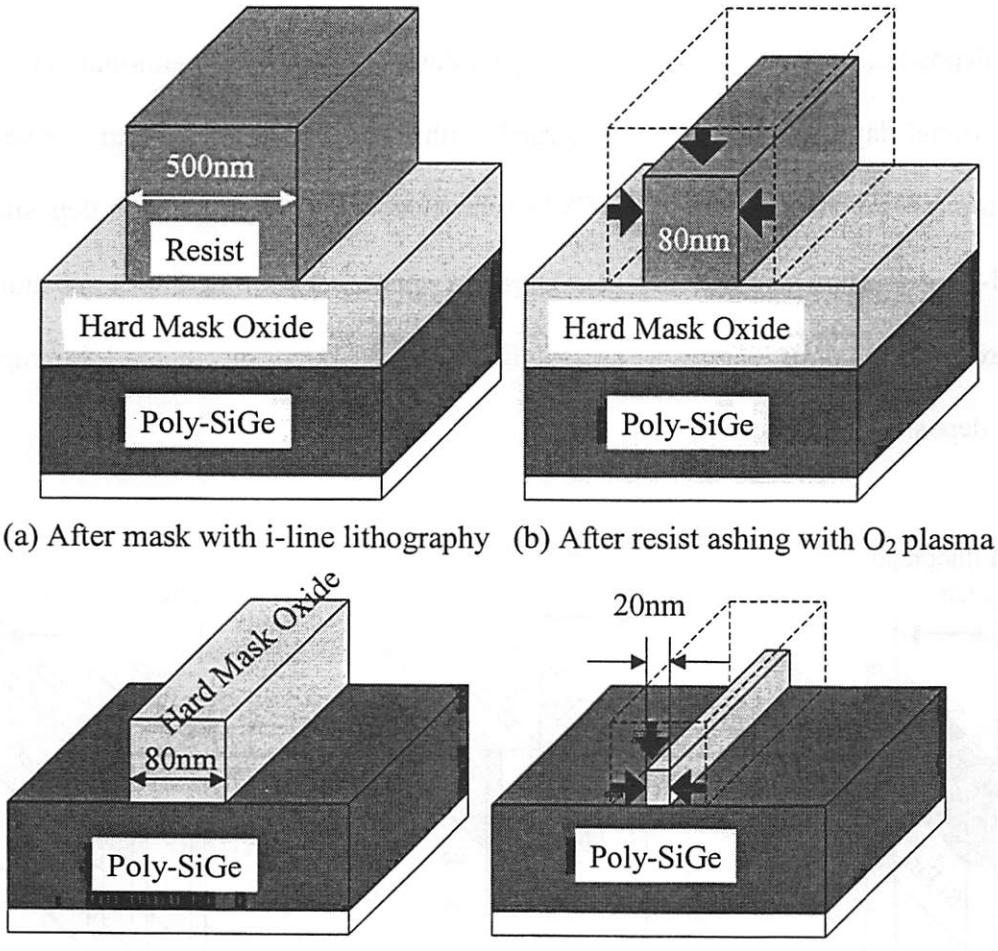
Chapter 2

Nanofabrication Technologies

2.1 Introduction

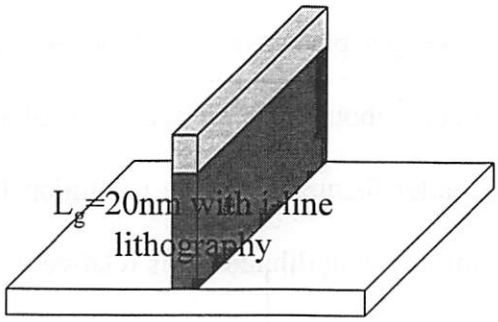
Currently, gate lengths of MOSFETs for advanced research are below 50nm. Making such a small feature is not an easy task in general. E-beam lithography has produced 15nm gates [2.1] and extreme-ultra-violet (EUV) lithography has generated 38nm period patterns [2.2]. When e-beam lithography employs some positive resists such as poly-methyl methacrylate (PMMA) providing high resolution, its throughput is too low even for research. In contrast, throughput of chemically amplified resists such as SAL-601, SNR-2000, and UVHS-II is high but its resolution is not good enough. So, even with e-beam lithography, it is challenging to obtain sub-30nm patterns. EUV lithography is not readily available yet. Two novel nano-lithography technologies for sub-30nm pattering are proposed and demonstrated: combination of resist ashing and hard mask oxide trimming, and spacer lithography.

Resist ashing with oxygen plasma was first developed for making sub-micron devices from g-line lithography about 13 years ago [2.3]. Since then, it has been rather widely used to produce smaller features than the resolution limit of lithography [2.4]-[2.6]. Oxide hard mask trimming with diluted HF is relatively straightforward, but I have found no report which describes it, let alone report on its use in the sub-30nm regime. Combination of these two techniques makes it possible to fabricate sub-30nm line widths using i-line lithography. Figure 2.1 shows how ashing-trimming works for pattern reduction from 500nm to 20mn [2-7]-[2.9].



(a) After mask with i-line lithography (b) After resist ashing with O_2 plasma

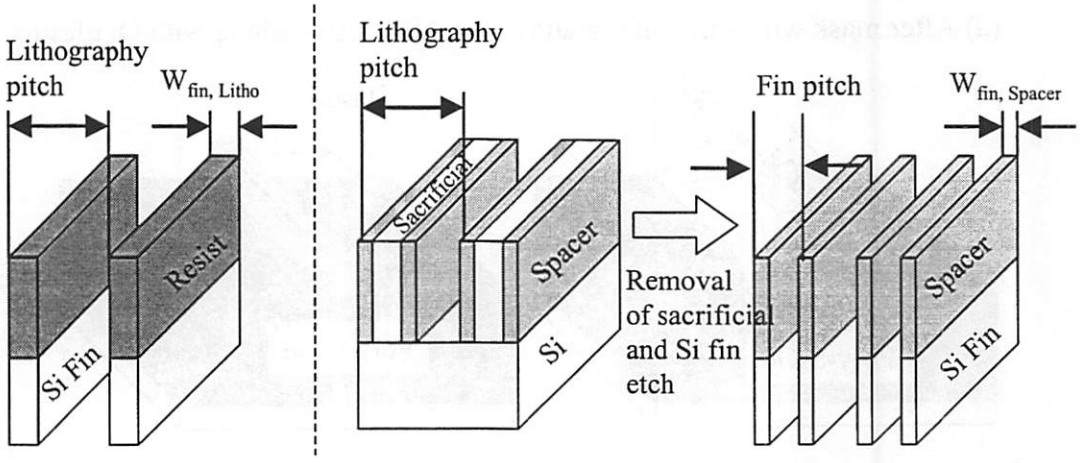
(c) After oxide hard mask etching (d) After hard mask oxide trimming with HF



(e) After gate poly etching

Fig. 2.1 The process sequence of resist ashing and oxide hard mask trimming.

A spacer lithography process technology using a sacrificial layer and a chemical vapor deposition (CVD) spacer layer had been developed with conventional dry etching. A sacrificial layer was initially patterned with i-line lithography and conventional anisotropic etch. Then another thin CVD layer that would be spacers is deposited and etched-back as shown in Fig. 2.2 (b). These spacers are transferred to a substrate with anisotropic etch. Thus, pattern width is defined not by optical or e-beam lithography but by the deposition thickness of CVD layer.



(a) Conventional lithography (b) Spacer lithography

Figure 2.2 Schematic diagrams of conventional lithography and spacer lithography.

Johnson et al. [2.11] made 250nm poly-Si spacer gate and To et al. [2.12] reported 90nm poly-Si spacer gate, respectively. But those minimum feature sizes are not small enough for nano-scale CMOS devices. Horstmann et al. [2.10] reported 50nm gate length. However, it provided only one gate length per wafer, and it is not practical for actual implementation to mass production.

For the FinFET, the short-channel effects can be suppressed by employing a body thickness (Si-fin width) which is approximately half of gate length L_g [2.13]-[2.15]. This is clearly impossible to accomplish with standard lithography technologies when L_g is at the limit of lithography. One drawback of ashing-trimming technology is poor uniformity. Critical dimension (CD) uniformity is not acceptable even for patterns defined by e-beam lithography. Uniformity is especially critical for the FinFET because variation in fin width (W_{fin}) can cause a change in channel potential and sub-bands structures, which governs short-channel behavior and quantum confinement effects of inversion charges.

The minimum-sized features in spacer lithography are defined not by photolithography but by the CVD film thickness. Therefore the spacer lithography technology yields CD variations of minimum-sized features which are much smaller than achieved by optical or e-beam lithography. It also provides a doubling of device density for a given lithography pitch as shown in Fig. 2.2. This spacer lithography technology is used to pattern Si-fin structures for double-gate MOSFETs (FinFETs). Higher Si-fin pitch than can be achieved with lithography is desirable in FinFETs, because multiple fins are demanded to increase the effective channel width [2.13][2.16]. A high fin density is also required to obtain large transistor drive current with good layout-area efficiency.

One drawback [2.10]-[2.12] of the prior spacer technology is that only one line width (gate length) is available. But by combining a conventional masking process and the spacer process in a novel manner, this limitation is overcome.

2.2 Resist Ashing and Hard Mask Oxide Trimming

2.2.1 I-line Resist Ashing

Sample wafers are exposed with an i-line stepper. The thickness of the positive i-line resist is 1.1um and baked at 90°C for 1 minute before exposure and at 120°C for 1min after exposure, respectively. The developed resist patterns are ashed in a conventional oxygen-plasma asher, Technics PE II. Oxygen pressure is 260mTorr with a flow rate of 51.1sccm.

The ashing rate of the i-line resist without hard baking is shown in Fig.1. The vertical ashing rate is the rate of reduction of the resist thickness, while the horizontal ashing rate is the rate of reduction of the line width. Both ashing rates change linearly with the ashing power and are independent of the initial line width. After hard baking of the resist at 120°C and 20 minutes, the ashing rate decreased by about one third. Hard baking before ashing is recommended when a slow ashing rate is required for a tight CD control.

The ratio of the horizontal ashing rate to the vertical ashing rate is about 1.2 : 1. This ratio is quite desirable. In the case of a lower ratio, resist thickness becomes too thin after ashing to act subsequently as an etching mask. Higher ratio would make the aspect

ratio of the resist profile too high after ashing. And the resist might fall down during etching.

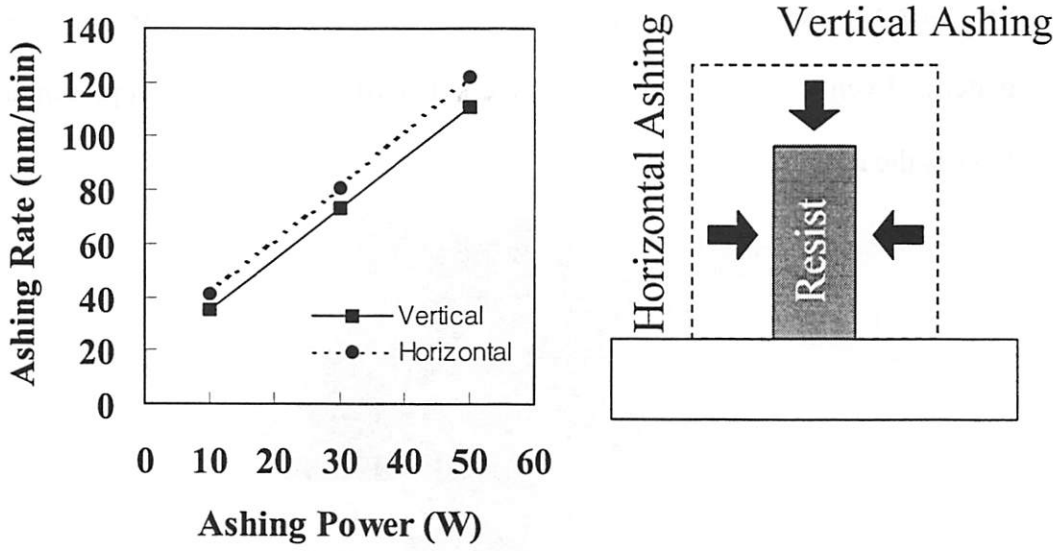


Fig. 2.3 Vertical ashing rate and horizontal ashing rate are linear functions of the microwave power and almost the same.

Ashing does not change a line-edge roughness of the resist as shown in Fig. 2.4. The initial line-edge roughness is very important for obtaining a more straight-line resist profile. If its line-edge roughness is more than the target width after ashing-trimming, the final profile can be seriously notched or, in the worst case, broken. In the case of i-line lithography, line-edge roughness depends strongly on the pattern fidelity on the mask and time delay between a resist development after exposure and ashing. Fig. 2.5 (b) shows that line-edge roughness is very large for a time delay of one week between resist development and plasma ashing. Oxygen in air may degrade a resist quality. It needs more study and investigation. When that time delay is smaller than three hours, a line-edge roughness was not observed as shown in Fig. 2.5 (a). For this condition, 500nm

initial line width was reduced down to 20nm directly without a hard mask oxide trimming. The slope angle at the narrow width resist increases slightly after ashing [2.8]. The top of the resist may be rounded in the end. This can happen earlier in narrow lines than in wide patterns. Even if the top of the resist is rounded off, etching does not present a problem as long as the resist is thick enough.

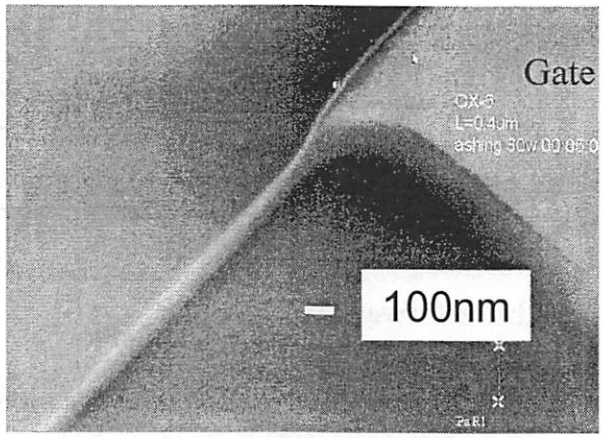
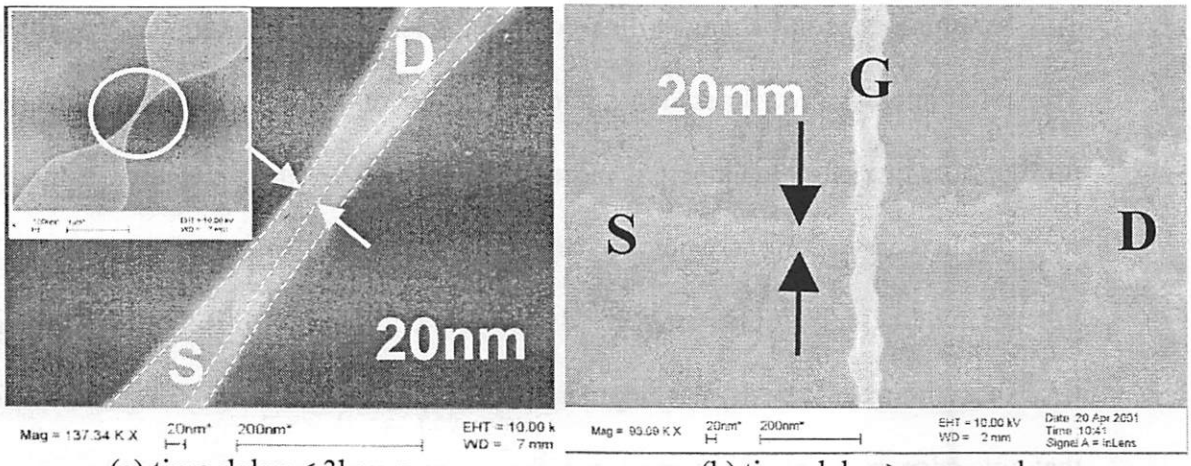


Figure 2.4 Tilted SEM picture of an i-line resist after ashing. It is reduced down to 80nm from 500nm.



(a) time delay < 3hours (b) time delay > one week.

Figure 2.5 Tilted SEM photographs showing 20nm width resist after ashing (a) and 20nm width gate and Si fin profile after gate etch (b). A long time delay between resist development and ashing produces large line-edge roughness.

2.2.2 E-beam Resist Ashing

Two chemically amplified resists, SNR-2000 and SAL-601, are evaluated for an investigation of ashing properties. Ashing of the e-beam resists is done in the same asher, Technics PE II. In the case of e-beam resist patterns, only a small amount of ashing compared to i-line patterns is needed because the initial line width is 100nm or less. Ashing power is fixed at 5W which is the lowest power to sustain a stable plasma. In Table 2.1, the ashing rates of the e-beam resists are shown. For SAL-601 and SNR-2000, the ashing rates are almost the same.

Resist (thickness)	Ashing Rate (nm/min)
SNR (200nm)	25
SNR (400nm)	22
SAL (150nm)	30

Table 2.1 Ashing rate of SAL-601 and SNR-2000 at 5W.

One interesting phenomenon is a resist hardening caused by scanning electron microscopy (SEM) during SEM inspection. Since SAL-601 and SNR-2000 are negative resists, they are hardened by the exposure to electron beam (energy is less than 1KeV) during SEM. After SEM, ashing rate of the resist patterns exposed to the e-beam decreases to two thirds of those not exposed. A SEM photograph of 17nm SNR-2000 resist is shown in Fig. 2.6 after ashing at 5W. It is reduced down to 17nm from 80nm.

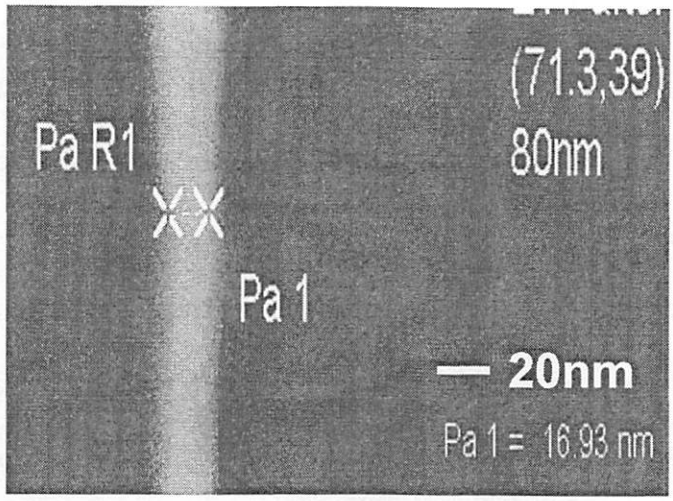


Figure 2.6 Tilted SEM photograph of 17nm SNR-2000 e-beam resist after ashing. Its initial line width is 80nm.

2.2.3 Hard Mask Oxide Trimming

The concept of oxide hard mask trimming is similar to the resist ashing. After ashing, a resist is transferred to hard mask oxide with an anisotropic plasma etch. Lam research model 9400 TCP etcher is used for an anisotropic etching. Two different etch recipes are developed and summarized in Table 2.2.

	Top RF power (W)	Bottom RF power (W)	Pressure (mTorr)	CF_4 (sccm)	CHF_3 (sccm)	Ar (sccm)	Etch rate (nm/min)
Recipe 1	200	40	13	100	0	0	120
Recipe 2	200	40	20	0	90	200	120

Table 2.2 Hard mask oxide plasma etch recipe.

Recipe 1 produces a vertical etch profile as shown in Fig. 2. 7 (a) and its selectivity of oxide to poly-Si is 1:1. During etching a hard mask oxide with recipe 1, ions are recoiled at a sidewall of hard mask oxide and a micro-trench is made along an edge of pattern as shown in Fig 2.7 (a). This recipe may cause more serious problems

when it is applied to UTB device because the selectivity is poor, thus it results in cutting the thin body. Recipe 2 produces a sloped etch profile as shown in Fig. 2.7 (b) and its selectivity of oxide to poly-Si is 3~4:1. It does not make any micro-trench. However, it is not available to get narrow line widths because of CD gain after hard mask oxide etch due to the sloped profile as shown in Fig. 2.7 (b). Thus two step etch is desirable, which is composed of recipe 1 and recipe 2. Recipe 1 is used to etch 90% of hard mask oxide thickness at first and recipe 2 is used to etch remaining 10% of hard mask oxide, with 20% over etch at last. Since CD gain is negligible in the last etch step, a narrow line is defined and the micro-trench is not observed.

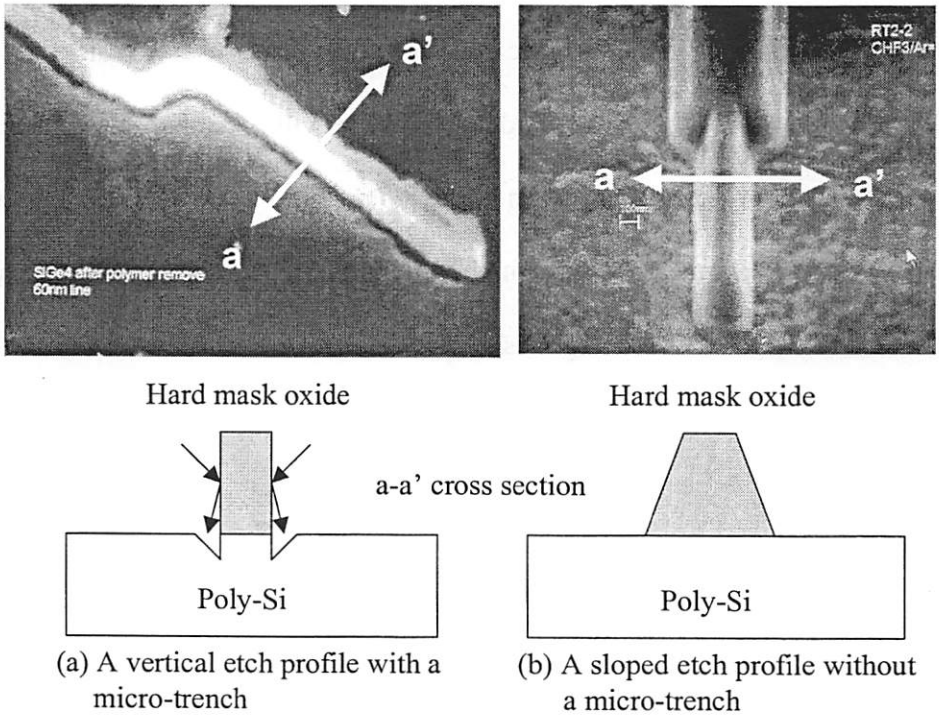
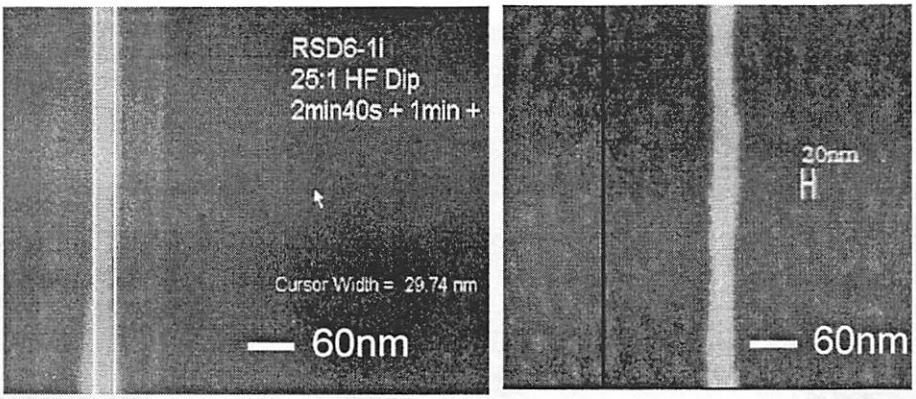


Figure 2.7 Tilted SEM photographs and schematic diagrams. (a) A hard mask oxide profile obtained with recipe 1 and (b) a hard mask oxide profile obtained with recipe 2.

Then, the patterned hard mask oxide is isotropically etched to a desired smaller size in diluted HF solution. This etch technique is named "hard mask oxide trimming". Typical etch rate of high temperature oxide (HTO) by LPCVD is about 30nm/min in 25:1 HF experimentally. The etch rate in diluted HF strongly depends on the oxide deposition conditions which determine the stoichiometry and annealing conditions to make a densified oxide. After hard mask oxide trimming, the gate can be etched quite easily without any photo-resist. HTO is deposited on a gate poly-Si or poly-SiGe with a thickness of a 120nm. The hard mask oxide is especially useful when the resist thickness is too thin to be a good etching mask.

Fig. 2.8 (a) shows a 30nm width of hard mask oxide, which is isotropically trimmed with (25:1) HF solution. The line starts as a 500nm i-line resist. It is ashed down to 80nm and transferred into the hard mask oxide with two step plasma etch. The hard mask oxide is then trimmed to 30nm from 80nm. Fig. 4 shows the top view of a poly-SiGe gate with a hard mask oxide, which is anisotropically etched with the trimmed hard mask oxide serving as the etching mask.



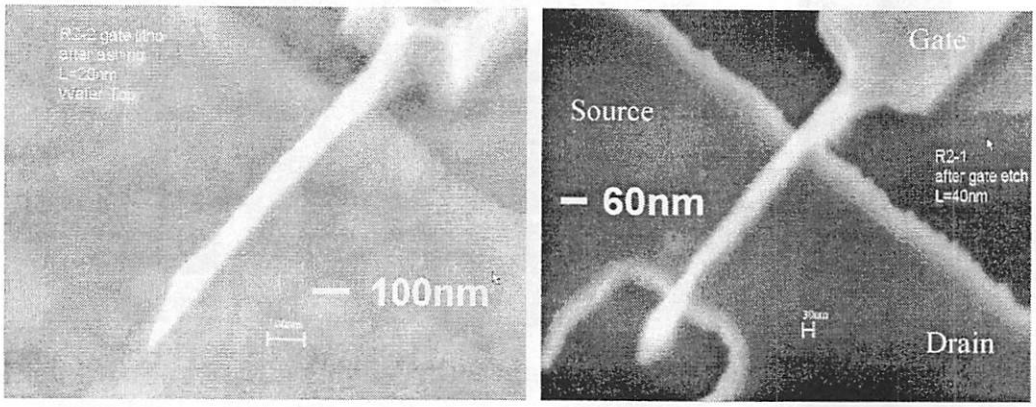
(a) A 30nm hard mask oxide after trimming

(b) A 30nm gate poly-SiGe after SiGe etch with a hard mask oxide

Figure 2.8 SEM photographs showing (a) a 30nm hard mask oxide and (b) a 30nm gate poly-SiGe with a hard mask oxide. They started with a 500nm i-line resist.

The poly-SiGe etch is performed with Lam 9400 TCP etcher as shown in Fig. 2.8 (b). The etching condition is 50sccm of Cl_2 and 150sccm of HBr, 15mTorr of pressure, 300W of RF top power, and 120W of RF bottom power for main etching. 200sccm of HBr, 5sccm of O_2, 35mTorr of pressure, 250W of RF top power, and 150W of RF bottom power is used for 100% overetch. 7sec main etching and 60sec overetch is required to etch 180nm poly-SiGe.

Fig. 2.9 (a) shows a 20nm width of isotropically trimmed hard mask oxide from a 80nm width, which is etched with 80nm width of SNR-2000 resist defined by e-beam lithography. Fig. 2.9 (b) shows a 20nm gate length poly-SiGe profile after anisotropic SiGe plasma with a 20nm width of hard mask oxide.



(a) A 20nm hard mask oxide after trimming

(b) A 20nm gate poly-SiGe after SiGe etch with a hard mask oxide

Figure 2.9 SEM photographs showing (a) a 20nm hard mask oxide and (b) a 20nm gate poly-SiGe with the hard mask oxide. They start with a 80nm SNR-2000 (e-beam) resist.

2.3 Spacer Lithography Technology

A spacer lithography process technology using a sacrificial layer and a CVD spacer layer has been developed, and is demonstrated to achieve sub-7nm structures with conventional dry etching. The minimum-sized features are defined not by the photolithography but by the CVD film thickness. Therefore, the spacer lithography technology yields critical dimension (CD) variations of minimum-sized features which are much smaller than achieved by optical or e-beam lithography as shown in Fig. 2.10. It was applied for making extremely narrow Si fins of FinFETs and for making a gate of ultra-thin body (UTB) MOSFET. It also doubles the fin pattern density in a given pitch, which is limited by optical and e-beam lithography in FinFETs.

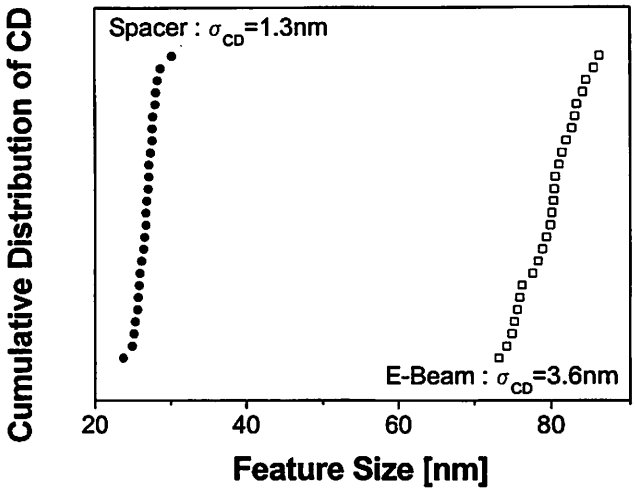


Figure 2.10 Measured CD uniformity of Si fin after Si fin etch and gate resist after e-beam lithography across a 4 inch wafer. CD uniformity by the pacer lithography is better than that by the e-beam lithography.

2.3.1 Spacer Lithography for Formation of Si Fin in FinFETs

In the FinFET, the fin width W_{fin} as shown in Fig. 2.13 (b) must be narrower than the gate length L_g [2-13]-[2.15] in order to suppress the short-channel effects. This means that narrow fins beyond the lithographic limit are needed. Tight control of CD variation is required, because a small variation in W_{fin} can cause significant change in device characteristics. A high fin density is also required to obtain large transistor drive current with good layout-area efficiency. Spacer lithography technology is proposed to solve the aforementioned problems. It provides for a doubling of fin density, which doubles the drive current for a given lithography pitch, as shown in Fig. 2.2 (a) and (b). All masking processes used in this work are performed with i-line optical lithography, because its

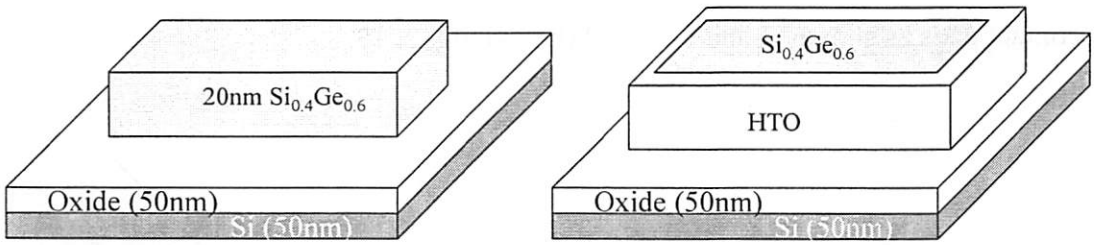
throughput is much better than e-beam lithography and the spacer lithography technology does not require very high resolution lithography.

(100) SOI wafers are used as the starting material. The SOI Si film is reduced from 100nm to 50nm by thermal oxidation and a hard mask oxide was thermally grown to a thickness of 50nm to protect the Si-fin during the subsequent gate poly-SiGe etch. 200nm $Si_{0.4}Ge_{0.6}$ is deposited by LPCVD on the oxide hard mask and patterned into sacrificial structures (to support the spacers) with optical lithography and plasma etching as shown in Fig. 2.11 (a). The process conditions of the anisotropic $Si_{0.4}Ge_{0.6}$ etch are as follows: 50sccm of Cl_2, 150 sccm of HBr, 15mTorr of pressure, 300W of RF top power, and 150W of RF bottom power in a Lam Research 9400 TCP etcher. The etch rate is 1.1um/min. All anisotropic plasma etches are performed with this etcher.

A vertical profile of sacrificial SiGe is very crucial to the spacer lithography technique because sloped $Si_{0.4}Ge_{0.6}$ sidewalls lead to sloped HTO spacers resulting in increased final fin widths. It is also important to completely remove polymers after the $Si_{0.4}Ge_{0.6}$ plasma etch. The estimated residual polymer thickness adhering to the sidewalls of $Si_{0.4}Ge_{0.6}$ is 20~30nm, which significantly enlarges a minimum-sized feature. As a post-etch treatment for the removal of polymers, the following consecutive steps are used: (100:1) HF 10sec, photo-resist strip with oxygen plasma, (100:1) HF 10sec, and piranha ((4:1) H_2SO_4: H_2O_2 @ 120°C) cleaning.

10nm HTO is then deposited by LPCVD over the patterned sacrificial $Si_{0.4}Ge_{0.6}$ layer. The thickness of HTO at the sidewalls of the sacrificial $Si_{0.4}Ge_{0.6}$ structures determines the final fin width. An extremely narrow fin width, beyond the lithographic limit as well as very uniform fin width can therefore be obtained with this spacer

lithography process. A subsequent anisotropic HTO spacer etch removes the HTO film on top of the sacrificial $Si_{0.4}Ge_{0.6}$ structure, as shown in Fig. 2. 11 (b), and it generates an even number of spacers (fins). 100% HTO overetch is applied to eliminate any spacer tails at the bottom of the $Si_{0.4}Ge_{0.6}$, because even a little spacer tail results in a broadened fin width. The process conditions for the HTO spacer etch are: 100sccm of CF_4, 13mTorr of pressure, 200W of RF top power, and 40W of RF bottom power. The etch rate is 120nm/min. This recipe produces ring-like spacers as shown in Fig. 2.12 (a) and Fig. 2.14.



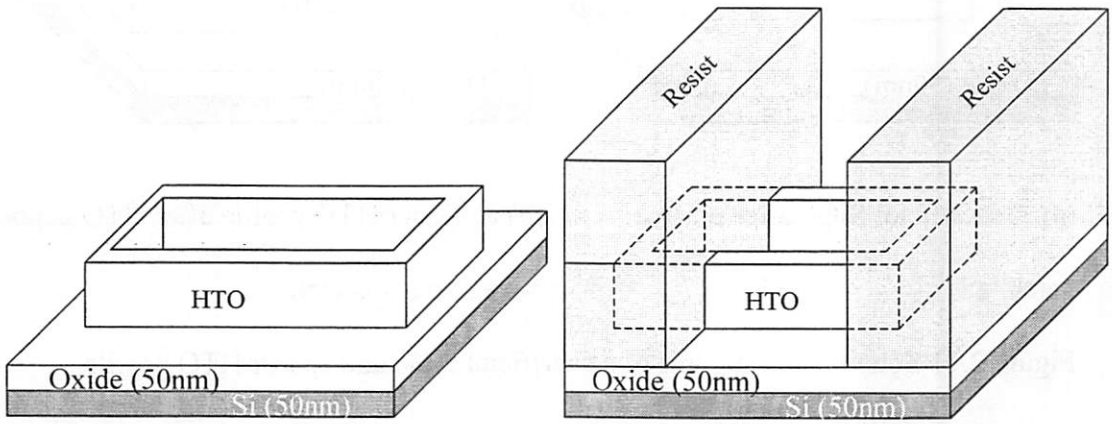
(a) A sacrificial SiGe after plasma etch

(b) A spacer HTO profile after HTO deposition and plasma etch

Figure 2.11 Schematic diagrams of a sacrificial SiGe and spacer HTO profile.

The $Si_{0.4}Ge_{0.6}$ structures are then removed by dry etching with 200sccm of HBr, 5sccm of O_2, 35mTorr of pressure, 250W of RF top power, and 120W of RF bottom power. The etch rate is 800nm/min. This anisotropic etch does not result in any loss of the thermally grown oxide because of high selectivity of poly-SiGe to oxide (400:1). $Si_{0.4}Ge_{0.6}$ residues after the plasma etch were removed with (5:1:1) H_2O:NH_4OH:H_2O_2 at 75°C [2.17]. HTO, thermally grown oxide, and Si are not etched significantly in this

solution. The final HTO spacer profile is shown in Fig. 2.12 (a) and Fig. 2.14. After removal of SiGe, some of HTO spacers can be collapsed down during spin dry process (2000 rpm) as shown in the inset of Fig. 2.14. For a wide process window, non-spin dry process is required for drying a wafer after DI water rinse. Optical lithography is used to define large S/D contact pads as shown in Fig. 2.12 (b) and Fig. 2.15. Therefore, the narrowest Si fins are defined by HTO spacers and S/D contact regions are defined by optical lithography at the same time as shown in Fig. 2.12 (b) and Fig. 2.15. One drawback of the spacer technique is that only one line width is provided [2.10]-[2.12]. Various fin widths are achieved by using photo-resist to define the fins as well as the S/D contact pads as shown in and Fig. 2.12 (b) and Fig. 2. 13.



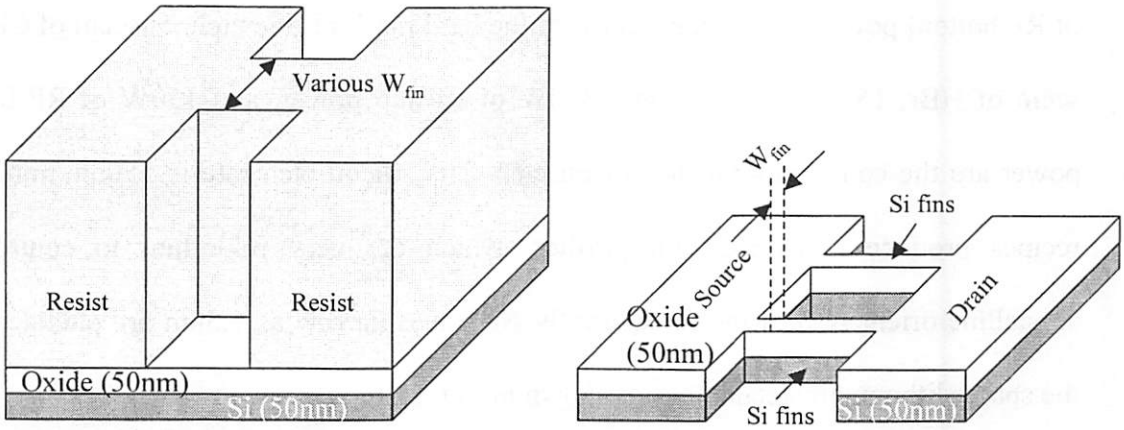
(a) A spacer HTO after removal of SiGe (b) A spacer HTO for Si fin and resists for S/D pads after S/D pad mask

Figure 2.12 Schematic diagrams of a spacer HTO and resists.

The anisotropic silicon fin etch consists of two steps: a hard mask oxide and a silicon fin etch. 100sccm of CF_4, 13mTorr of pressure, 200W of RF top power, and 40W

of RF bottom power are the conditions for the hard mask nitride etch. 50sccm of Cl_2, 150 sccm of HBr, 15mTorr of pressure, 300W of RF top power, and 150W of RF bottom power are the conditions for the silicon etch. The silicon etch rate is 550nm/min. Both recipes produce non-sloped fin profiles, which are very important to control the crystalline orientation of the fin sidewalls. Si fins as narrow as 6.5nm are obtained with the spacer lithography technology as shown in Fig. 2.16.

Unfortunately, a 6.5nm fin disappears after 3nm sacrificial oxidation. When a silicon film thickness is thinner than 10nm, an oxidation rate is significantly faster compared to an oxidation rate in the bulk-Si. This new phenomena is not reported yet. It needs more study. Detailed oxidation rate for a different silicon thickness will be shown in the Figure 4.9. For a reliable process and wider process window, spacer HTO film is increased to 30nm from 10nm. Due to the 70% of HTO step coverage at the 200nm height of $Si_{0.4}Ge_{0.6}$, the thickness of HTO at the sidewalls is 20nm for a 30nm deposited HTO. Thus 20nm Si fin is made and Fig. 2.17 (a) shows a cross-sectional 20nm Si fin obtained with a tunneling electron microscopy (TEM). The asymmetrical shape of Si fin in Fig. 2.17 (a) comes from the fact that HTO spacer is not symmetric, i.e., its shape of the tip is close to an asymmetrical trapezoid rather than a rectangle. A left side HTO (more vertical) in Fig. 2.17 (a) is adhered to a sacrificial $Si_{0.4}Ge_{0.6}$ and a right side HTO is not. The bowed fin profile in Fig. 2.17 (a) is caused by a faster oxidation rate at the middle of the fin, which is sandwiched by a top oxide and bottom buried oxide during a 5nm sacrificial oxidation. It is verified with a process simulator, TSUPREM4 as shown in Fig. 2.17 (b).



(a) A resist for various fin widths (b) Si fin with hard masks after oxide and Si etch

Figure 2.13 Schematic diagrams of a resist for various fin widths and Si fins formed by spacers after a hard mask oxide and Si.

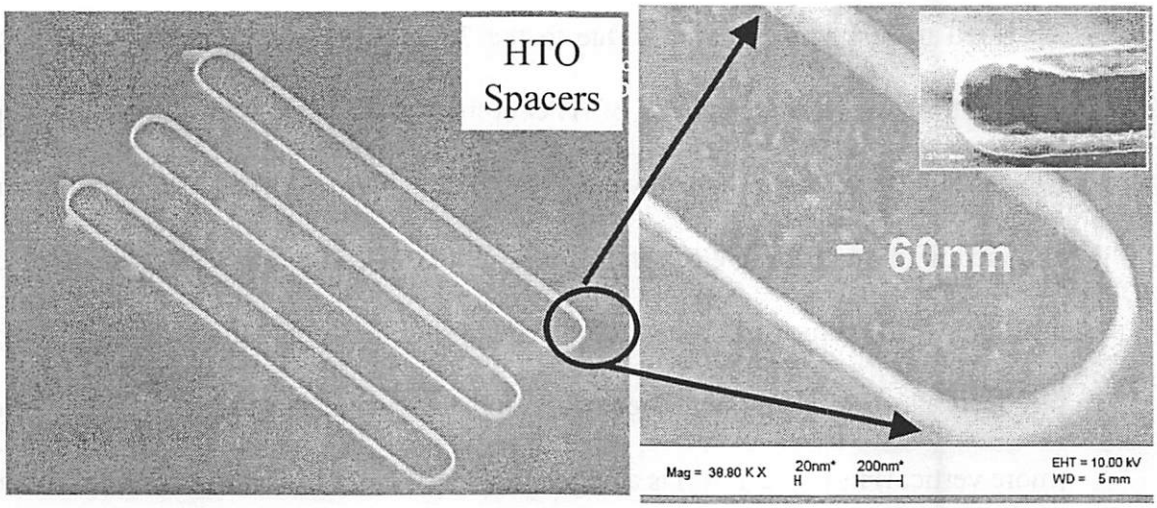


Figure 2.14 Tilted SEM photographs of HTO spacer profile after removal of sacrificial SiGe. HTO film width is 20nm in the right picture. Inset shows a collapse-downed HTO spacer after spin-dry (2000 rpm) process.

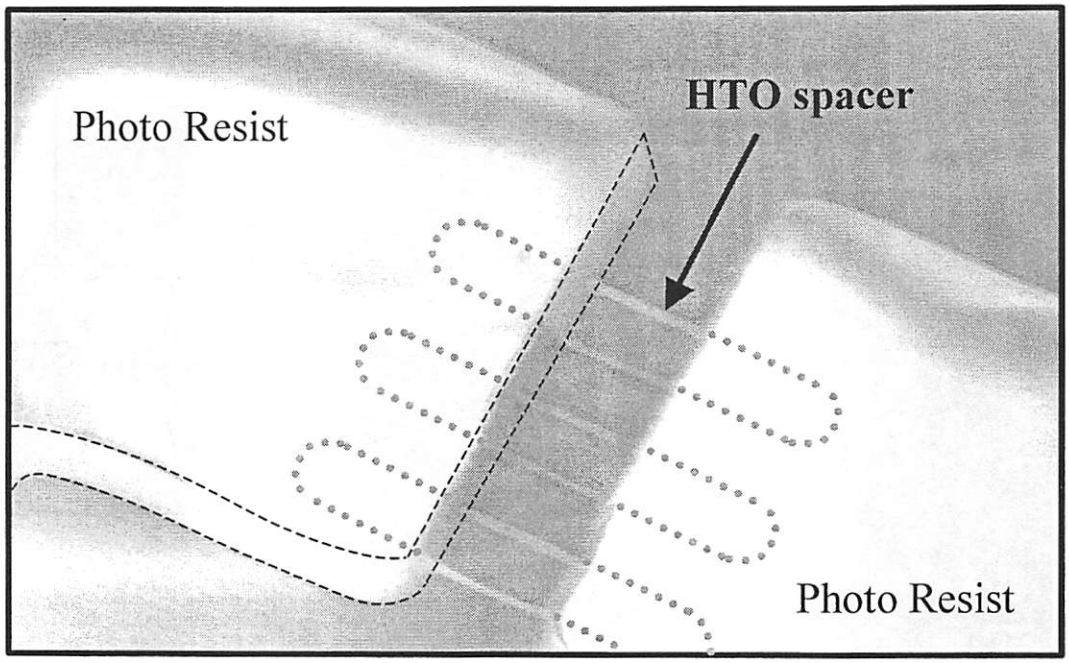


Figure 2.15 Tilted SEM photographs of HTO spacers for narrow fins and resists for S/D contact pads.

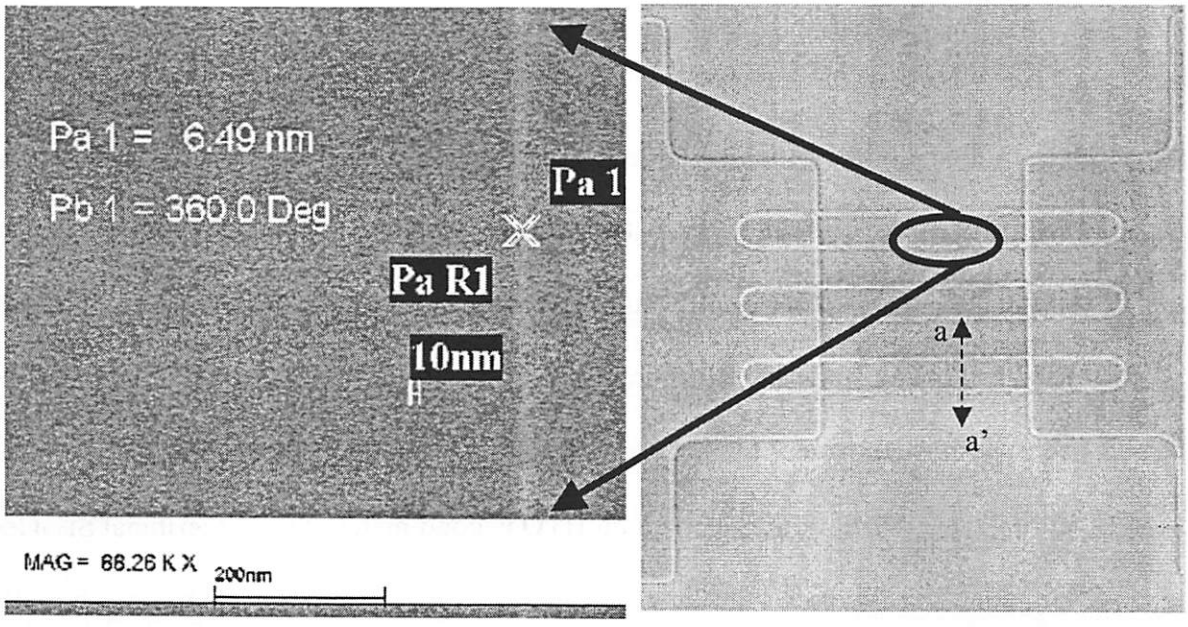
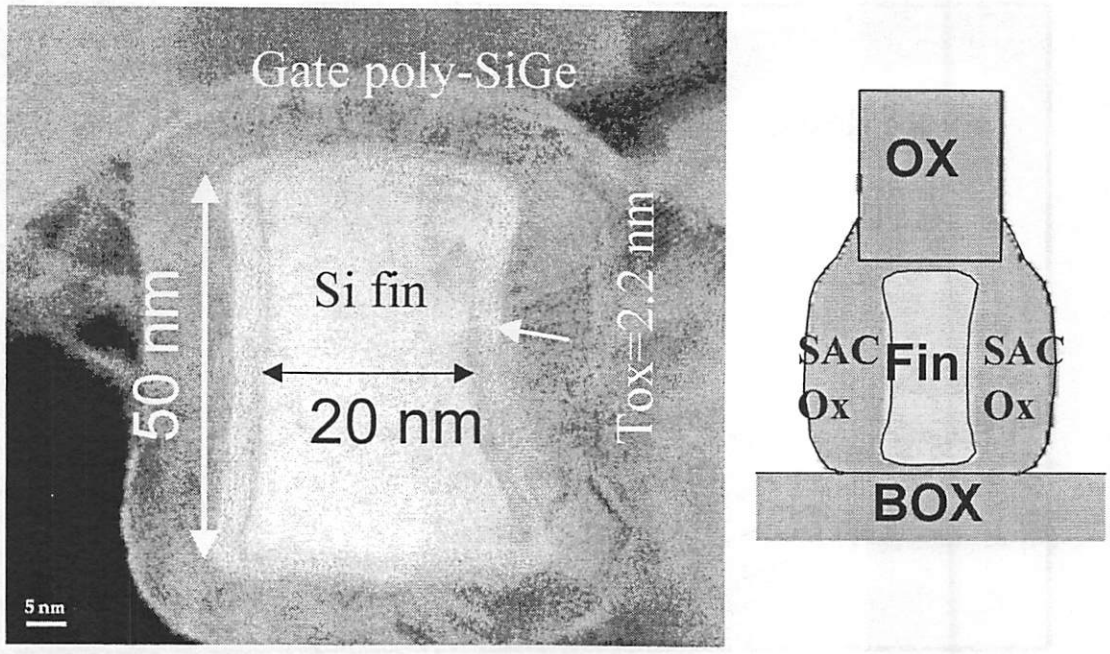


Figure 2.16 SEM photographs of 6.5nm Si fin width after etching a hard mask nitride, pad oxide, and Si.



(a) A 20nm fin width

(b) A bowed fin profile after sacrificial oxidation with TSUPREM4 simulator

Figure 2.17 A cross-sectional (a-a' direction in Fig. 2. 16) TEM photograph of 20nm Si fin width after gate patterning.

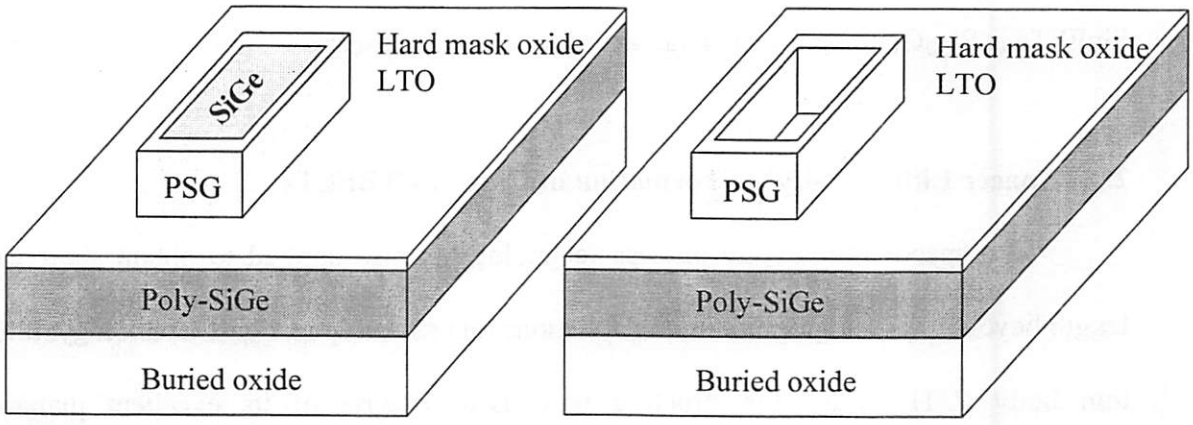
For a wider process window, $Si_{0.2}Ge_{0.8}$ is preferred rather than $Si_{0.4}Ge_{0.6}$ because its residues are dissolved in H_2O:NH_4OH:H_2O_2 (5:1:1) at 75°C quickly [2.17] after plasma etch. Phospho-silicate glass (PSG), low temperature oxide (LTO), and HTO are used for spacer material. A desirable layer as spacer material is LTO rather than HTO because its deposition temperature is low, thus additional SiGe oxidation is avoided during LTO deposition. Even though HTO is good material for a sacrificial $Si_{0.4}Ge_{0.6}$, it is not good for $Si_{0.2}Ge_{0.8}$ because it can be oxidized during HTO deposition (800°C). Thereby, LTO (450°C) is preferred for a $Si_{0.2}Ge_{0.8}$. One problem of LTO deposition is that a deposition rate (10nm/min) is uncontrollably fast for an application to spacer

lithography. At least 2 minutes are required for gas stabilization in 4 inch LPCVD furnace. There are two ways to reduce a LTO deposition rate: lowering the deposition temperature and diluting silane (SiH_4) gas with oxygen (O_2). The former will degrade a LTO quality, and be more porous. Low LTO deposition rate (~2nm/min) is achieved with the conditions of SiH_4=5sccm, O_2=70sccm, 300mTorr, and 450°C. One benefit of PSG is in a high selectivity to a thermal oxide in a diluted HF. Before gate formation, spacers should be removed to reduce the step height of fin. In this process sequence, those spacers are removed with a diluted HF during removal of a sacrificial oxide. If spacers have a low selectivity to a thermal oxide, a buried oxide of SOI is more etched because of extended etch time of HF. Thus, an undercut is made and a large overlap capacitance is induced. Therefore, PSG is applied to spacer layer. If phosphine (PH_3) is introduced to the above condition for forming PSG layer, the deposition rate is increased up to 8nm/min again. The best combination for the spacer lithography to form Si fin in FinFETs is $Si_{0.2}Ge_{0.8}$ for a sacrificial layer and LTO for spacers.

2.3.2 Spacer Lithography for Formation of Gate in UTBFETs

The spacer lithography process technology is also applied to obtain short gate length beyond the lithographic limit. To demonstrate such a spacer gate technology, ultra-thin body (UTB) transistor structure is chosen because of its excellent planarity [2.18][2.19]. The formation of "stringers" of the gate hard mask can be problematic for non-planar structures. Up through SiGe gate deposition and LTO deposition for a hard mask, the device fabrication process is similar to that reported in [2.18][2.19]. 100nm of (100) SOI films is reduced to less than 10 nm by multiple thermal oxidations. Isolation is simply achieved with a thin body silicon etch. Thermal oxide was grown for a gate

dielectric. P+ in-situ doped poly- $Si_{0.5}Ge_{0.5}$ is deposited with a thickness of 200nm on the gate oxide. A hard mask oxide LTO is deposited with a thickness of 100nm on the gate poly-$Si_{0.5}Ge_{0.5}$. LTO is adopted because of its low deposition temperature (450°C) instead of nitride (800°C) in order to avoid boron penetration. As for the spacer lithography process for gate formation, 200nm sacrificial $Si_{0.4}Ge_{0.6}$ is deposited and patterned as shown in Fig. 2.18 (a). 30nm PSG spacers are then formed, similarly. Since the CF_4-based etch recipe (recipe 1 in Table 2.2) makes a micro-trench along the $Si_{0.4}Ge_{0.6}$ pattern, this pattern is transferred to the underlying ultra-thin Si body during the subsequent gate etch, and cut out ultra-thin body. Thereafter, the channel and S/D are disconnected. Recipe 2 in Table 2.2 is used for etching PSG, and leaves a ring-like PSG spacer as shown in Fig. 2.18 (b). However, this recipe generates a sloped profile, which results in an enlargement of the final gate length from 30nm to 47nm.

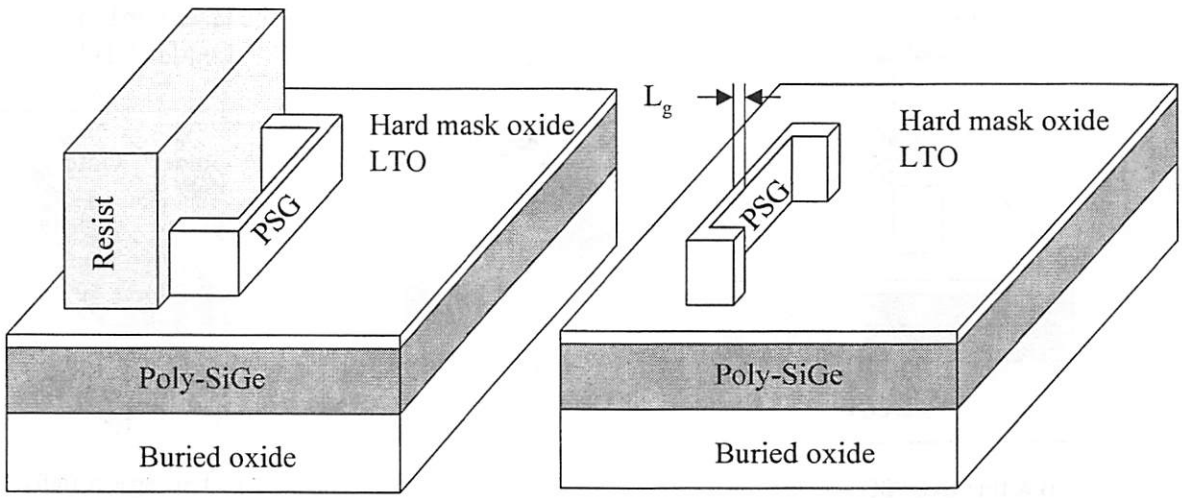


(a) A sacrificial SiGe and etched PSG spacer

(b) A ring-like PSG spacer after removing SiGe

Figure 2.18 Schematic diagrams of a sacrificial SiGe and etched PSG spacer.

Because the spacer lithography process always produces lines in pairs, one of the spacer lines needs to be removed in order to obtain a single gate. Another resist mask (dummy gate mask) is used for this purpose as shown in Fig. 2.19 (a). This dummy gate mask is not necessary for certain gate layouts as in the case of shorted NMOS and PMOS gate electrodes for CMOS inverter. A diluted (25:1) HF removed the PSG spacer not protected by the photo-resist as shown in Fig. 2.19 (b). The etch selectivity of PSG to LTO is higher than 10 in that diluted HF.



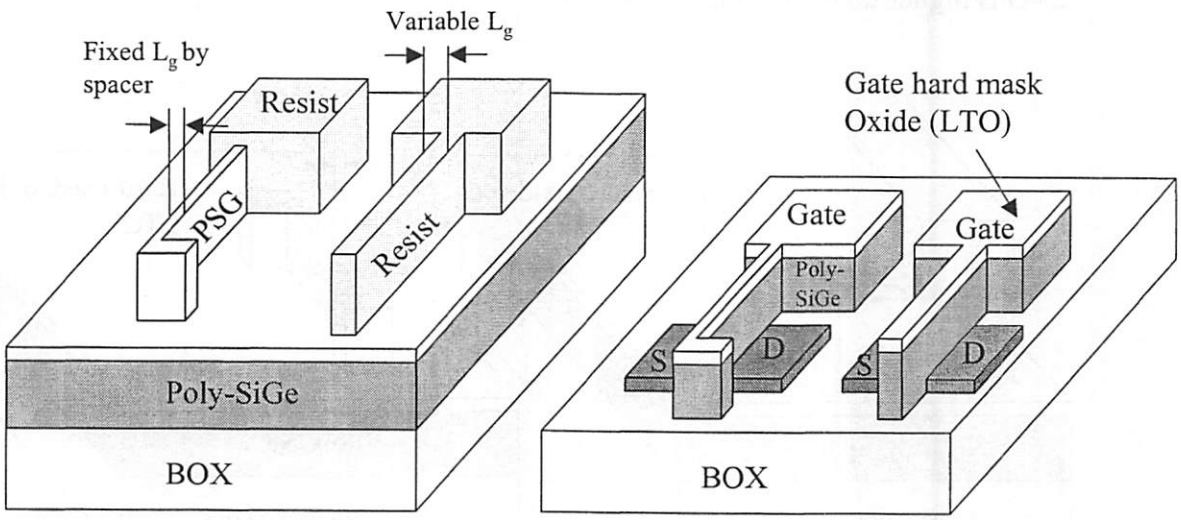
(a) A ring-like PSG spacer with a dummy gate resist

(b) A line-like PSG spacer after removing the other side of PSG with a diluted HF

Figure 2.19 Schematic diagrams of a dummy gate resist and a line-like PSG spacer after removing PSG spacer uncovered with the resist.

A gate contact pad mask is used to define contact pads and also to provide variable channel lengths as shown in Fig. 2.20 (a) and (b). The gate length of the transistor on the right in Fig. 2.20 (b) can be varied using conventional lithography. In this process, the minimum channel length is provided by the spacer, and longer and

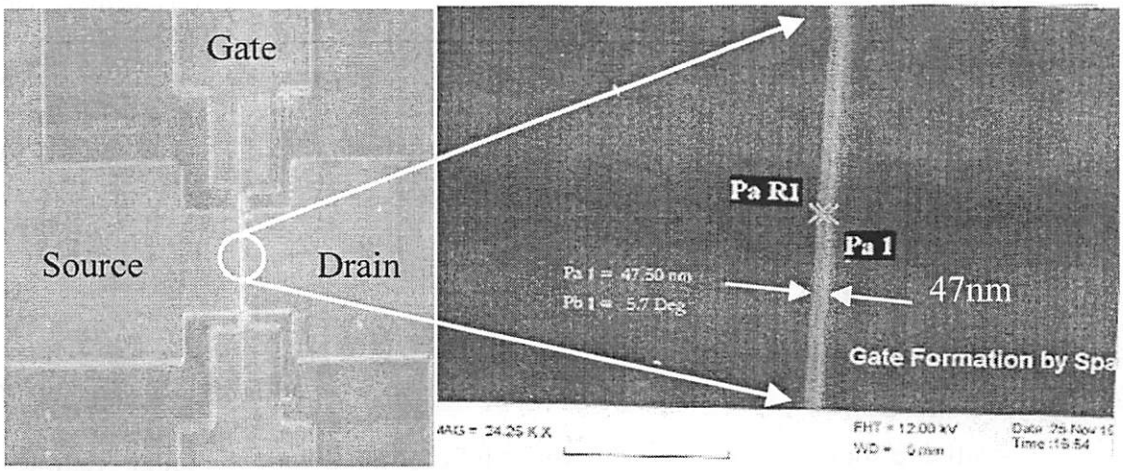
variable channel lengths are provided by conventional lithography. Fig. 2.21 (a) shows top view of an etched gate poly-SiGe profile on an ultra-thin body Si with a hard mask LTO, and Fig. 2.21 (b) shows 47nm gate poly-SiGe profile defined by PSG spacer. The residual rectangle images in Fig. 2.21 (a) is caused by a recess of a hard mask LTO because of a non high-selectivity of LTO to PSG during PSG etch using diluted HF.



(a) A line-like PSG spacer with resists for a gate contact pad and a various gate length before gate etch

(b) A minimum-sized gate by PSG spacer (left) and a various long gate by resist (right) after gate etch

Figure 2.20 Schematic diagrams of (a) a line-like PSG spacer with a resist for gate contact pad (left), and a resist for a various and longer gate length (right) and (b) a minimum sized gate by PSG spacer (left), and a various and longer gate by resist (right).



(a) (b)

Figure 2.21 Optical microscope photograph (left) and SEM photograph (right). 47nm gate length was defined by a line-like PSG spacer.

2.4 Conclusion

Resist ashing is a convenient technique to extend the lithography line width limit for i-line and e-beam lithography. Oxide hard mask trimming is also useful technique to obtain small features beyond the limit of conventional lithography. Combination of these two techniques can produce sub-20nm patterns for research purposes.

A spacer lithography technology is developed for defining the Si fin in FinFETs and defining the gate in UTB MOSFETs. A 6.5nm width of Si fin is successfully defined, which is the smallest features ever reported for a FinFET. Sub-50nm long gates of UTBFETs are also patterned. The spacer technology provides minimum-sized features

beyond the limit of optical and e-beam lithography and better CD uniformity than optical and e-beam lithography. It also doubles the pattern density in a given pitch, which is limited by optical and e-beam lithography. Process details are discussed and reported.

2.5 References

[2.1] J. Kedzierski, P. Xuan, E. Anderson, J. Bokor, T.-J. King, C. Hu, "Complementary silicide source/drain thin-body MOSFETs for the 20nm gate length regime," *IEDM Technical Digest,* p. 57-60, 2000.

[2.2] H.H. Solak, D. He, W. Li, S. Singh-Gasson, F. Cerrina, "Exposure of 38nm period grating patterns with extreme ultraviolet interferometric lithography", *Applied Physics Letter,* vol.75, p.2328-30, 1999.

[2.3] J. Chung, M. Jeng, J. E. Moon, A.T. Wu, T.Y. Chan, P.K. Ko, C. Hu, "Deep-submicrometer MOS device fabrication using a photoresist-ashing technique," *IEEE Electron Device Letters*, vol.9, p. 186-8, 1988.

[2.4] M. Tanabe, T. Matsuno, N. Kashiwagi, H. Sakai, K. Inoue, and A. Tamura, "0.1um AlGaAs/InGaAs high electron mobility transistor fabrication by the new method of thinned resist pattern reversed by metal," *Journal of Vacuum Science & Technology B*, vol.14, p.3248-51, 1996.

[2.5] M. Ono, M. Saito, T. Yoshitomi, C. Fiegna, T. Ohguro, H.S. Momose, H. Iwai, "Fabrication of sub-50-nm gate length n-metal-oxide-semiconductor field effect transistors and their electrical characteristics," *Journal of Vacuum Science & Technology B*, vol.13, p.1740-3, 1995.

[2.6] D.Y. Jeon, G.M. Chin, K.F. Lee, R.H. Yan, E. Westerwick, M. Cerullo, "Gate technology for 0.1-um Si complementary metal-oxide-semiconductor using g-line exposure and deep ultraviolet hardening," *Journal of Vacuum Science & Technology B*), vol.12, p. 2800-4, 1994.

[2.7] K. Asano, Y.-K. Choi, T.-J. King, C. Hu, "Patterning Sub-30-nm MOSFET Gate with I-line Lithography," *IEEE Transactions on Electron Devices,* vol. 48, p.1004-6, 2001

[2.8] K. Asano, Master Thesis, "Development of 25nm Raised Source and Drain MOSFET", Department of Electrical Engineering and Computer Science, *University of California at Berkeley*, 1999.

[2.9] Y.-K. Choi, Master Thesis, "Ultra-thin-body SOI MOSFET for Deep-sub-tenth Micron Era", Department of Electrical Engineering and Computer Science, *University of California at Berkeley*, 1999.

[2.10] J.T. Horstmann, U. Hilleringmann, and K.F. Goser, "Matching Analysis of Deposition Defined 50-nm MOSFET's", *IEEE Transactions on Electron Devices,* vol.45, p.299-306, 1998.

[2.11] R.A. Johnson, S.D. Kasa, P.R. Houssaye, G.A. Garcia, I. Lagnado, P.M. Asbeck, "Novel Polysilicon Sidewall Gate Silicon-on-Saphire MOSFET for Power Amplifier Applications," *IEEE SOI Conference*, p.136-37, 1997

[2.12] K.H. To, J. Woo, "High Performance Sub-0.1um SOI Polysilicon Spacer gate MOSFETs Using Large Angle Tilted Implant for Drain Engineering," *IEEE SOI Conference*, p. 94-95, 1999.

[2.13] X. Huang, W.-C. Lee, C. Kuo, D. Hisamoto, L. Cang, J. Kedzierski, E. Anderson, H. Takeuchi, Y.-K. Choi, K. Asano, V. Subramanian, T.-J. King, J. Bokor, C. Hu, "Sub 50-nm FinFET : PMOS," *IEDM Technical Digest,* p. 67-70, 1999

[2.14] L. Chang, S. Tang, T.-J. King, J. Bokor, C. Hu, "Gate Length Scaling and Threshold Voltage Control of Double-Gate MOSFETs," *IEDM Technical Digest,* p. 719-22, 2000

[2.15] N. Lindert, Y.-K. Choi, L. Chang, E. Anderson, W. Lee, and T.-J. King, J. Bokor, and C. Hu, "Quasi-Planar NMOS FinFETs with Sub-100nm Gate Lengths", *59th Device Research Conference,* p.26-27, 2001.

[2.16] D. Hisamoto, W.-C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T.-J. King, J. Bokor, C. Hu, "FinFET-A Self-Aligned Double-Gate MOSFET Scalable to 20nm," *IEEE Transactions on Electron Devices,* vol.47, p.2320-5, 2000.

[2.17] 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", *Journal of the Electronic Materials*, vol.21, p.805-810, 1992.

[2.18] Y.-K. Choi, K. Asano, N. Lindert, V. Subramanian, T.-J. King, J. Bokor, C. Hu, "Ultra-thin-body SOI MOSFET for Deep-sub-tenth Micron Era," *IEEE Electron Device Letters,* vol. 21, p.254-255, 2000.

[2.19] Y.-K. Choi, D. Ha, T.-J. King, C. Hu, "Nanoscale ultrathin body PMOSFETs with raised selective germanium source/drain", *IEEE Electron Device Letters*, vol.22, p.447-8, 2001.

Chapter 3

Process Technologies for Novel Device Structures

3.1 Introduction

Over the past three decades, technological advancements in MOSFETs have been achieved primarily through the scale-down of device dimensions [3.1]-[3.3], which improves circuit speed and reduces die size. As the MOSFET channel length is reduced to 50nm and below, the suppression of off-state leakage current becomes an increasingly difficult technological challenge that will ultimately limit the scalability of the conventional MOSFET structure.

Thin silicon body using a silicon-on-insulator (SOI) technology is a promising structure because of its ability to suppress off-state leakage current [3.4]-[3.6]. One of its technological challenges is the large series resistance of the thin body layer. Raised sources and drains (S/D) fabricated by selective silicon epitaxial growth [3.7]-[3.9], metal silicide [3.10][3.11], selective germanium or SiGe deposition [3.12]-[3.14] have been

implemented in attempts to reduce that resistance. However, selective silicon epitaxial growth on thin body silicon has been problematic. This growth is not compatible with high-K gate dielectric and metal gate technologies because it is a high temperature process [3.15]. It is also very difficult and complicated to implement material having a lower barrier height for NMOS and PMOS S/D junction at the same time in Schottky barrier UTB MOSFETs. [3.16]. Metal contact formation on the Schottky barrier UTB MOSFET is another challenge.

In this work, a resist etch-back process with poly-Si [3.6] and selective germanium deposition by LPCVD [3.12][3.13] to make raised S/D on thin-body silicon are proposed and demonstrated. The resist etch-back process has been used to make planarized surface by using non-conformality of photo-resists and blanket etch-back. Even though the resist etch-back process is complicated, it is reliable and less sensitive to surface cleanness than silicon selective epitaxial growth. Selective Ge deposition by LPCVD is a low temperature process (350°C), and dopants in Ge S/D are activated at lower temperatures than in Si S/D. Thus it is compatible with metal gate and high-k gate dielectric tecnologies. Another benefit of a selective Ge process is that the germane gas provides an in-situ removal of native oxide [3.17]. The resist etch-back process with poly-Si is used to make a raised S/D on UTB MOSFETs and a selective Ge deposition by LPCVD is used to make a raised S/D on UTB MOSFETs and FinFETs. In addition, a nickel-germanide process designed to reduce sheet resistance is evaluated.

FinFETs suffer from a narrow process window caused by non-planar and vertical structures. To overcome this difficulty, chemical-mechanical polishing (CMP)

[3.18][3.19] to make a planarized gate has been evaluated in the Microlab of the University of California at Berkeley and used. CMP provides a large depth-of-focus (DOF) margin in lithography and an etching process window without stringers and residues. Spacer FinFETs with CMP are demonstrated for the first time [3.34].

3.2 Resist Etch-Back for Raised Poly-Si Source and Drain

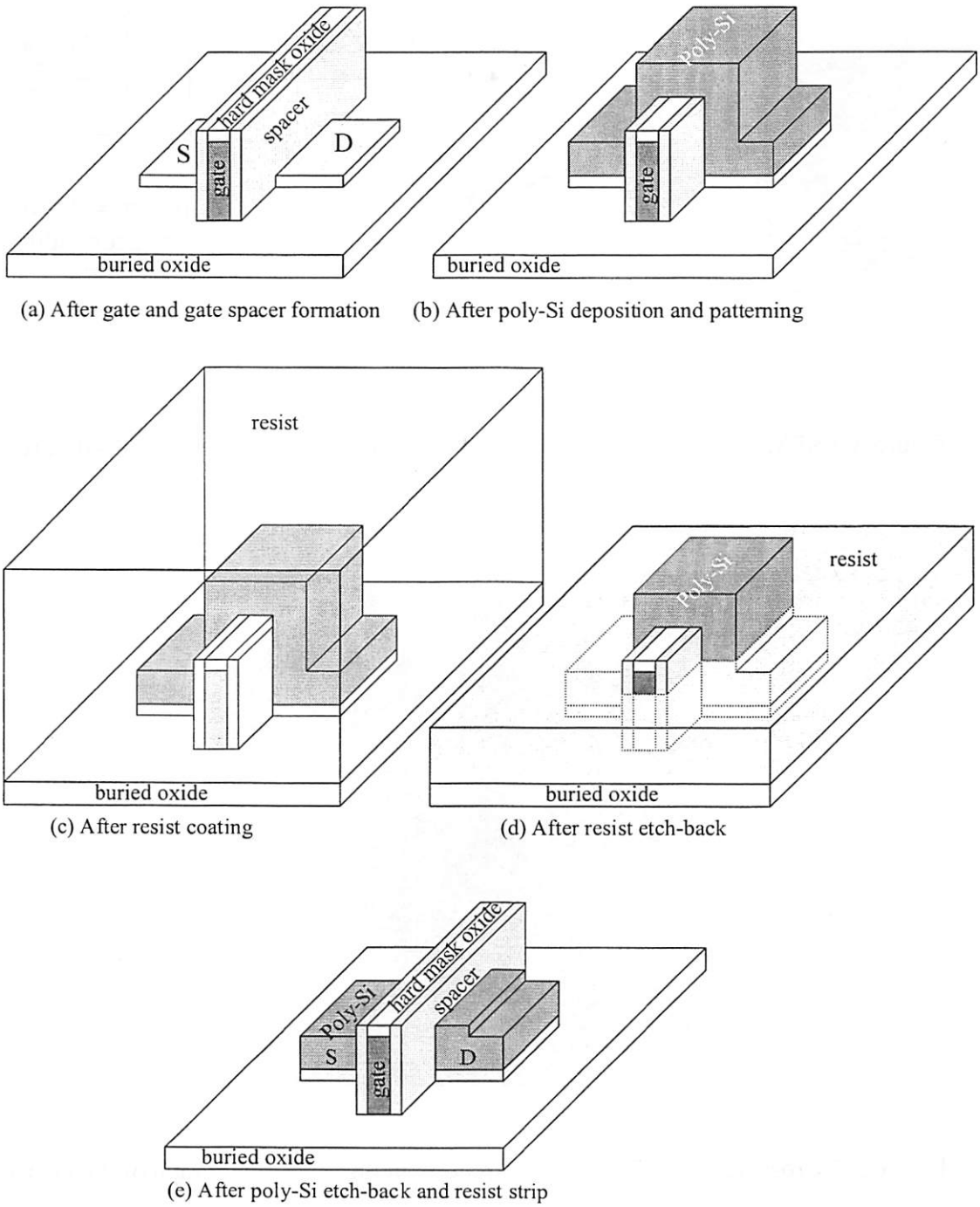
A resist etch-back process is used to make a self-aligned raised poly-Si S/D on UTB MOSFETs. A key idea of a resist etch-back process is to use the non-conformality of resist after spin coating [3.20]. The resist is thin over the high features on the wafer and thick over low features on the wafer. Detailed process flows have been reported [3.6][3.21]. The starting material is a (100) 100nm SOI p-type wafer. The 100nm SOI body is reduced to thinner than 20nm by multiple thermal oxidations. The active pattern is defined with conventional i-line lithography and plasma etch. All masking steps are done by i-line lithography in this work. Gate oxide with a thickness of 2.4nm is grown on the thin body. Then, 180nm of in-situ boron doped $Si_{0.5}Ge_{0.5}$ and 120nm of high temperature oxide (HTO) hard mask are deposited by LPCVD, consecutively. Sub-40nm gate lengths are defined by ashing-trimming with i-line lithography and plasma etch.

HTO is deposited with a thickness of a 10nm for LDD implantation. Additional HTO with a thickness of 20nm is deposited on the first deposited HTO to make spacers for N+ / P+ heavily doped junctions and to reduce overlap capacitance between the gate and the raised poly-Si S/D. After the spacer HTO etch, the $Si_{0.5}Ge_{0.5}$ gate is shielded by the HTO hard mask on the top and the spacer HTO on the sidewalls, as shown in Fig. 3.1 (a).

Next, in-situ doped poly-Si is deposited by LPCVD on the cleaned thin body (20nm), which is supposed to be S/D with a thickness of 100nm. It is etched anisotropically and patterned with the same mask that is used to make the thin body active area in the previous step, as shown in Fig. 3.1 (b). A cap HTO with a thickness of 10nm is deposited on the patterned poly-Si (not shown in Fig. 3.1). The cap oxide is used for widening the process window of plasma etch-back. It serves as a hard mask for the poly-Si etch when the etched-back resist is not thick enough to protect poly-Si which should not be etched. The next step is to coat the wafer with resist. The required resist thickness is comparable to the height of the gate stack because one purpose of resist etch-back is to easily eliminate the resist on the top of the gate. Thus 300nm thickness of the resist is coated at 3000 rpm, as shown in Fig. 3.1 (c). Due to the high viscosity of conventional resist, resist which is thinner than 800nm can not be achieved even with a high spin speed (8000 rpm). I-line resist is chosen because of the high selectivity of the resist to poly-Si. It is mixed with a thinner to reduce the resist thickness to 300nm. The mixing ratio is 3:1 (thinner : i-line resist).

The resist etch-back process is broken into 3 steps: resist etch-back, cap oxide etch, and poly–Si etch. All etching processes are done with the Lam Research model 9400 TCP etcher. The detailed processes are summarized in Table 3.1. Plasma etching 40% of the resist removes the resist on top of the gate oxide hard mask, as shown in Fig. 3.1 (d). Subsequent etches remove the cap oxide and poly-Si which are exposed by the resist etch-back, and separate the source and drain, as shown in Fig. 3.1 (e). All oxides shielding the gate protect the gate poly-SiGe during the etch-back. The cap oxide protects the raised poly-Si, which is supposed to be source and drain even though the resist is

consumed during the etch-back process. After removal of the remaining resist, a self-aligned raised poly-Si S/D is achieved, as shown in Fig. 3.1 (e), Fig. 3.2, and Fig. 3.3.



(a) After gate and gate spacer formation (b) After poly-Si deposition and patterning

(c) After resist coating (d) After resist etch-back

(e) After poly-Si etch-back and resist strip

Figure 3.1 Schematic diagrams of the resist and poly-Si etch-back process.

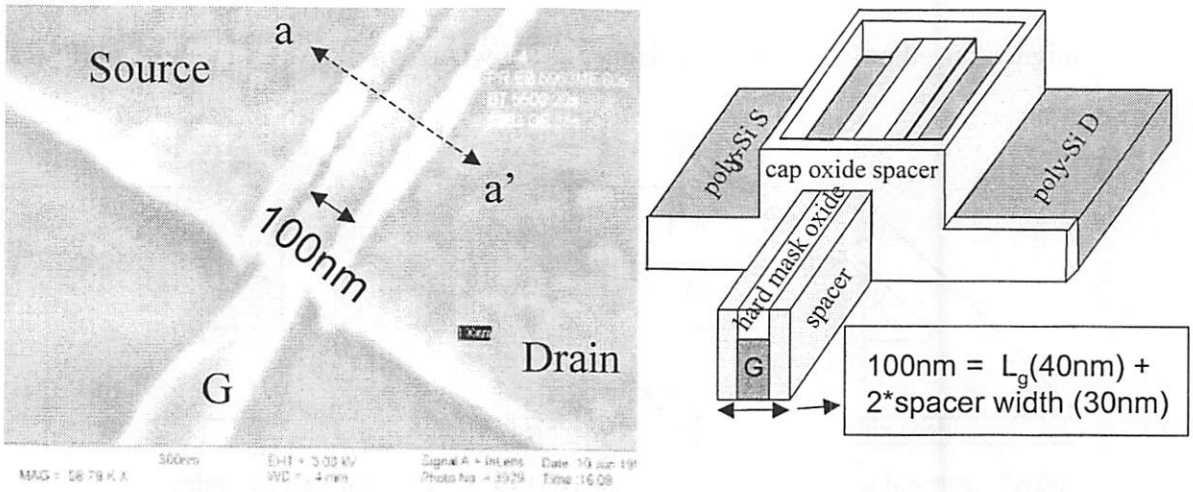


Figure 3.2 SEM photograph and schematic diagram of the etched-back raised poly-Si S/D.

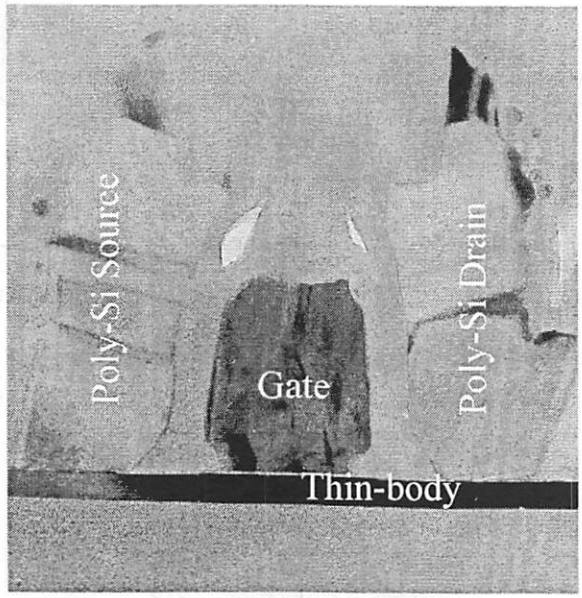


Figure 3.3 Cross-sectional TEM (a-a' direction in Fig. 3.2) photograph of the etched-back poly-Si raised S/D. A spacer oxide tail is seen at the corner of the gate and the thin body.

	Resist etch-back	Oxide etch-back	Poly etch-back
Pressure (mTorr)	35	20	35
RF top power (W)	250	200	250
RF bottom power (W)	120	40	120
CHF_3 (sccm)	0	90	0
Cl_2 (sccm)	0	0	0
HBR (sccm)	200	0	200
O_2 (sccm)	5	0	5
Ar (sccm)	0	200	0

Table 3.1 Comparison table of the recipe of the resist etch-back process

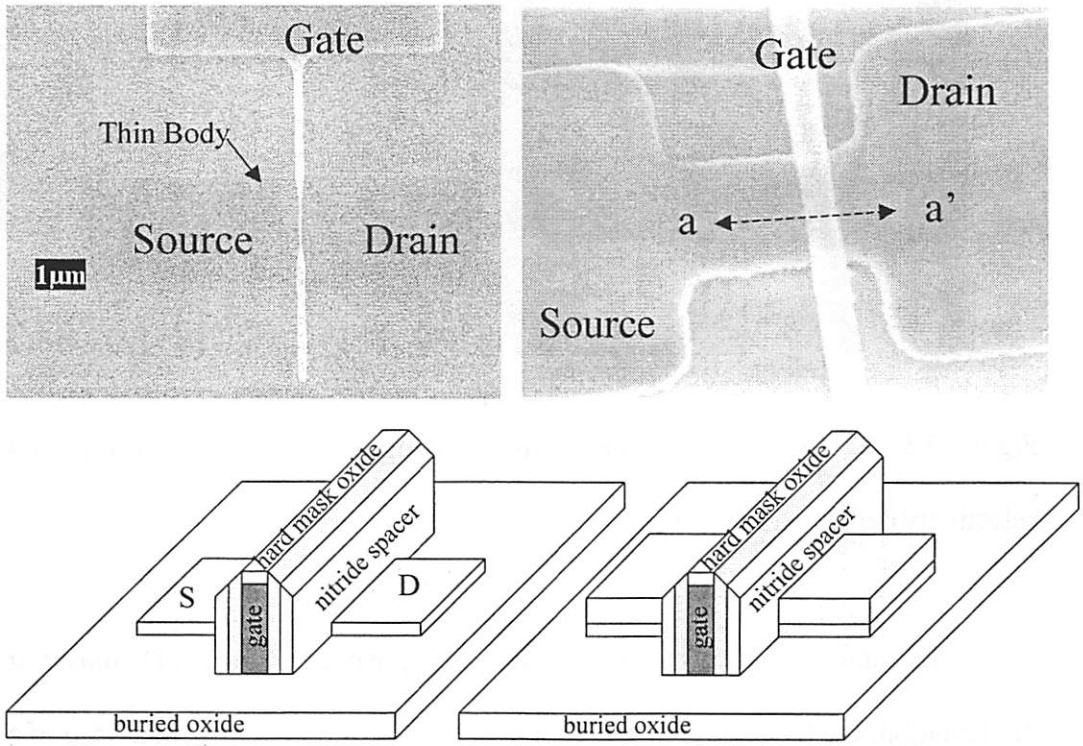
Even though a resist etch-back process for application to the raised poly-S/D on UTBFETs is complicated, it does not require novel or advanced equipment. Therefore, it is very useful at the research level. It has many applications for planarization of interlayer dielectrics and for low temperature technologies. One drawback of resist etched-back raised poly-Si S/D is the large overlap capacitance between the gate and the raised poly-Si S/D. A selective Si epitaxial growth or selective Ge deposition by LPCVD would be a good alternative to solve this problem. However, a selective Si epitaxial growth on thin body Si is very difficult because of the surface cleanness issue. Thus selective Ge deposition is an attractive solution because it has lower capacitance, process complexity, and temperature than the alternatives.

3.3 Selective Ge Deposition for Raised Poly-Ge Source and Drain

Selectively deposited Ge is used to make raised S/D's on the thin body of UTB MOSFETs. A selective germanium deposition is performed on a cleaned silicon surface by LPCVD furnace. The process conditions for undoped Ge are 350°C of deposition temperature, 300mTorr of pressure, and 200sccm of GeH_4 flow. The process flow for a selectively deposited raised Ge S/D before poly-Si S/D deposition is the same as that for an etched-back raised poly-Si S/D. The device structure immediately before a selective Ge deposition is shown in Fig. 3. 4 (a).

Due to a low etching selectivity of oxide to Si, it is very difficult to make a vertical oxide spacer without a tailing profile because oxide overetch at the corner of the gate and the thin body is not allowed [3.21]. Therefore bi-layer spacers which are composed of inside HTO and outside nitride are used. The gate is shielded by a hard mask oxide on the top and bi-layer spacers on the sidewalls. To remove the native oxide on the thin body, (25:1) HF can be used before loading wafers to a LPCVD furnace for germanium deposition. This is clearly not an in-situ cleaning process. However, germane gas in the furnace provides *in-situ* cleaning to remove the native oxide [3.17]. Thus a selective Ge deposition process is less sensitive to surface cleanness than a selective Si-epitaxial growth. Ge is not grown on a gate hard mask oxide (HTO), a sidewall nitride, or a buried oxide, so deposition selectivity is very high. No Ge spots or clusters are found on the buried oxide when viewed by an optical microscope over a 4 inch wafer or an SEM

over a few dies, as shown in Fig. 3.4. Fig. 3.5 shows a cross-sectional TEM picture of the selective Ge profile on a 3nm thin body.



(a) After gate and gate spacer formation (b) After a selective Ge deposition

Figure 3.4 SEM photographs and schematic diagrams for raised S/D profile before and after Ge deposition. SEM pictures are taken at the different transistors between before Ge deposition and after Ge deposition.

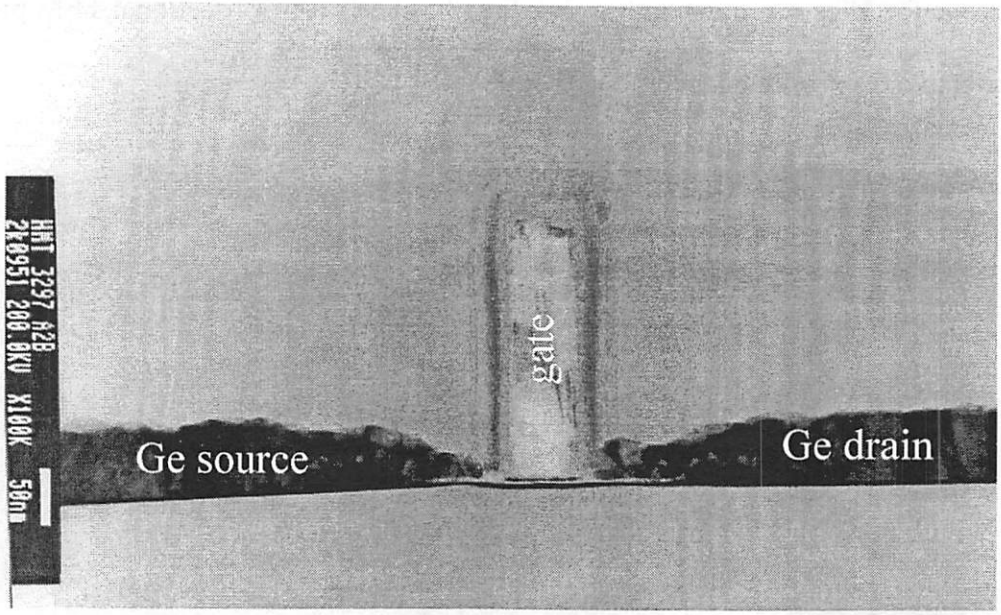
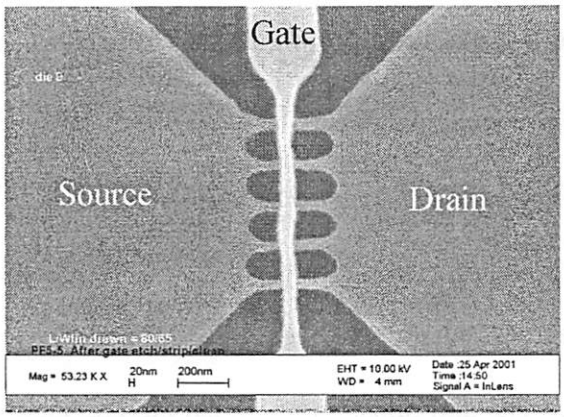


Figure 3.5 A cross-sectional TEM image along a-a' direction in Fig. 3.4 shows a selectively deposited raised Ge S/D.

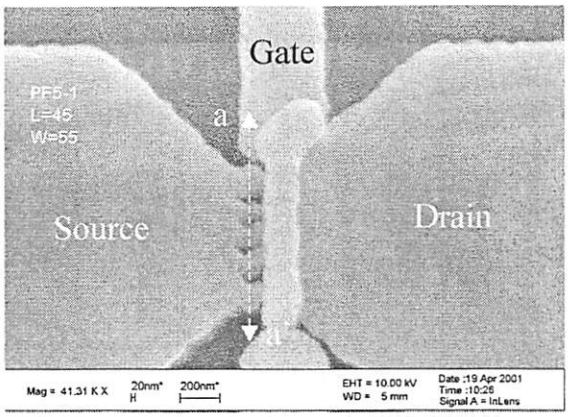
To make CMOS, undoped Ge is deposited. After S/D masking and ion implantation, the Ge is heavily doped. Phosphorus is implanted with a dose of $5x10^{15}cm^{-2}$ for n-type (N+) S/D and boron is implanted with a dose of $5x10^{15}cm^{-2}$ for p-type (P+) S/D. Phosphorus is used rather than arsenic because there is a smaller diffusivity difference between phosphorus and boron. The dopants in the S/D are activated with rapid thermal annealing (RTA) at 650°C to 750°C.

A selective Ge deposition is also used to make the raised S/D on the fins of FinFETs. A selective Ge is successfully deposited at the tops and sidewalls of fins. The crystal orientation is (100) on the top surfaces and (110) on the sidewalls of fins.

Apparently, there is no deposition-rate dependence on crystal orientation, as shown Fig. 3.7 and 3.8. Ge is isotropically and selectively deposited, as shown Fig. 3.8.

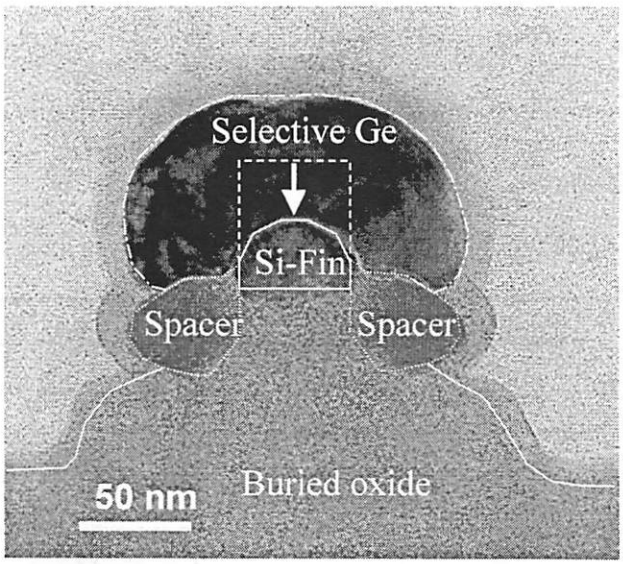


(a) A gate and fin profile before a selective Ge deposition

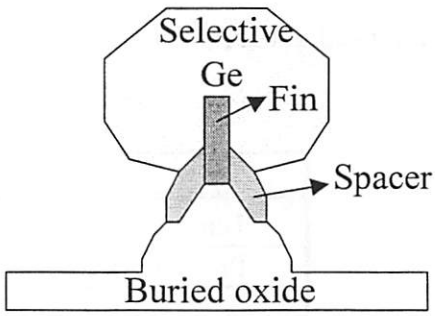


(b) A gate and fin profile after a selective Ge deposition

Figure 3.7 SEM photographs of a selectively deposited Ge on Si fin in FinFETs.



(a) TEM photograph of selectively deposited raised Ge on a fin (a-a' direction in Fig. 3.7)



(b) Schematic diagram of selectively deposited raised Ge on a fin

Figure 3.8 TEM photograph and schematic of a selectively deposited Ge on Si fin. The original fin (dashed line) is reduced by the overetch of spacers.

If the deposition temperature is below 410°C, high selectivity is maintained [3.22]. There are no experimental data between 350°C and 410°C yet. Deposition selectivity is has been shown to be very high at 350°C. In-situ doped Ge has been deposited in order to investigate the selectivity dependence on doping gas. Phosphine (PH_3) is used for heavily n-type (N+) doped Ge and diborane (B_2H_6) is used for heavily p-type (P+) doped Ge. In the case of in-situ boron-doped Ge, selectivity disappeared and deposition rate increased. This phenomenon may be caused by a diborane gas diluted with 90% silane (SiH_4), which may serve as seeds for Ge deposition on the oxide and nitride. Detailed process conditions for undoped Ge, boron-doped Ge, and phosphorus-doped Ge are summarized in Table 3.2. In the case of in-situ phosphorus doped Ge deposition, Ge is not deposited at all when the ratio of the exposed Si area to the exposed oxide area is 1:10000. However, when that ratio is 1:400, selectivity is recovered and the deposition rate is decreased. Two different wafers (1:400 vs. 1:10000) were loaded to the furnace at the same time to verify this phenomenon. The same results were reproduced. Further studies are needed to analyze this phenomenon.

	Undoped Ge	N+ Ge	P+ Ge
Temperature (°C)	350	350	350
Pressure (mTorr)	300	300	300
GeH_4	200	200	200
PH_3 (sccm)	0	5	0
B_2H_6 (sccm)	0	0	40
Deposition rate (nm/min)	10.2	4.1	14.9

Table 3.2 Summary of Ge deposition conditions.

Heavily doped p-type (P+) Ge shows a lower sheet resistance than heavily doped p-type (P+) Si because of higher levels of boron activation and higher hole mobilities in Ge [3.23]-[3.26]. The annealing temperature for boron activation in $Si_{1-x}Ge_x$ is very low, as shown in Fig. 3.9. As the annealing temperature rises above 850°C, the sheet resistance of B-doped Ge (x=1) is rapidly increased to more than 100kΩ/□ (not shown in Fig. 3.9). Fig. 3.9 shows that the sheet resistance decreases as the Ge fraction in poly-$Si_{1-x}Ge_x$ increases when the annealing temperature is below 600°C, and it is consistent with previous reports [3.25][3.26]. However, this trend is reversed in the case of phosphorus-doped films. The sheet resistance is high for larger Ge fraction (x > 0.45) in poly-$Si_{1-x}Ge_x$ due to significant reduction in phosphorous activation [3.25][3.26]. In the case of phosphorus-doped Ge (x=1) films, the sheet resistance is larger than 1kΩ/□, as shown in Fig. 3.10. The sheet resistance of heavily arsenic-doped Ge is comparable to that of heavily arsenic-doped Si, as shown in Fig. 3.10.

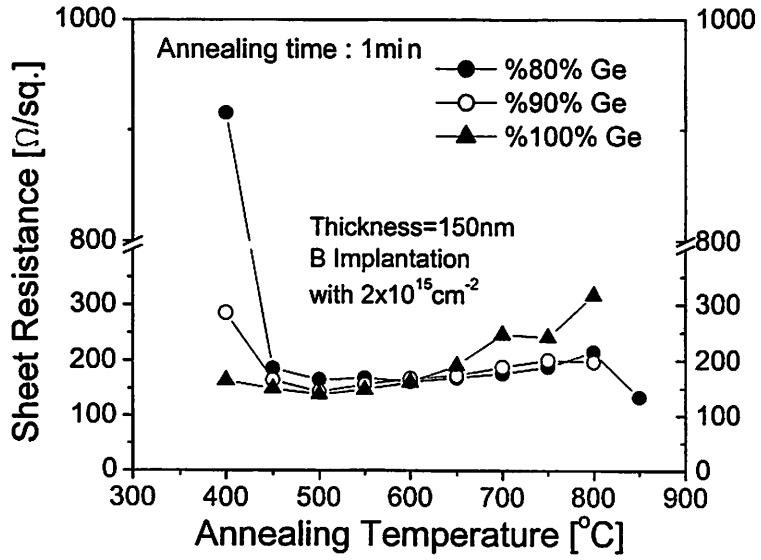


Figure 3.9 Sheet resistance of 150nm-thick boron doped poly-$Si_{1-x}Ge_x$ for different annealing temperatures [3.27].

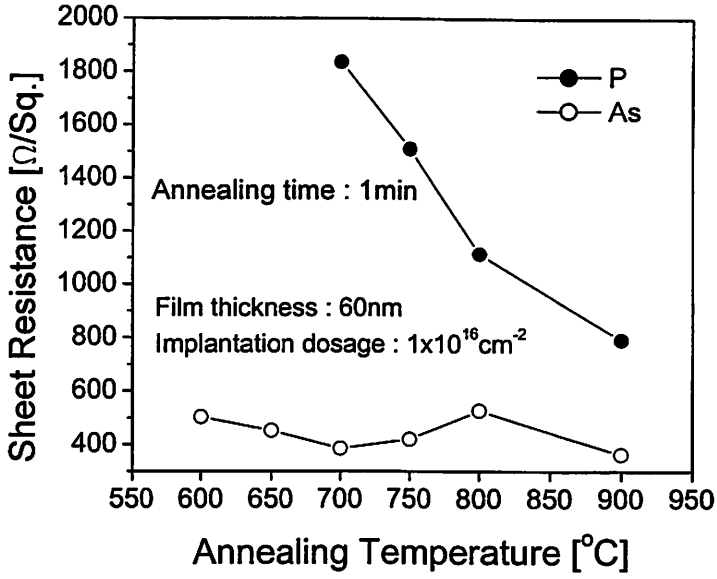
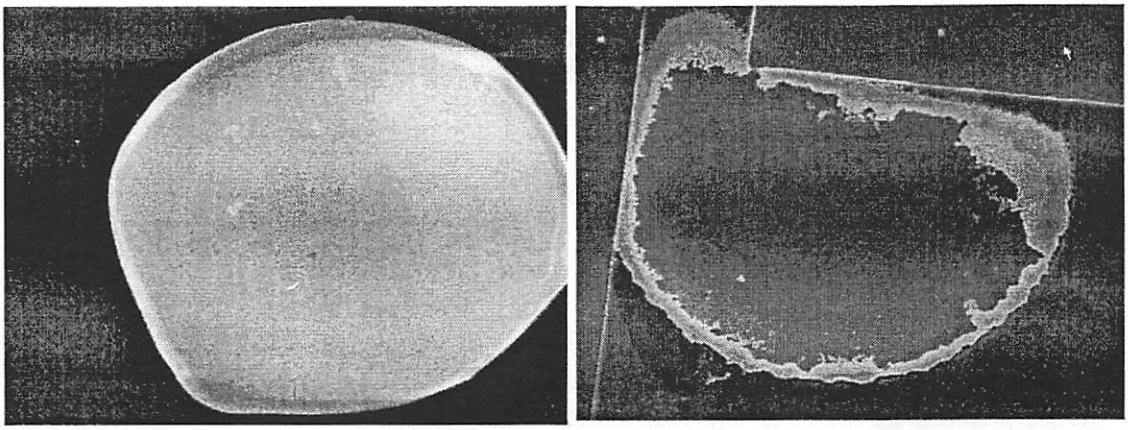


Figure 3.10 Sheet resistance of 60nm-thick phosphorus- and arsenic-doped poly-Ge for different annealing temperatures [3.28].

An abnormal phenomenon found in the selectively deposited Ge raised S/D's of UTBFETs. After rapid thermal annealing (RTA) at 900°C for 1min, the selectively deposited Ge on the thin bodies of UTBFETs, which is capped with 30nm of LTO, is swollen and burst-out, as shown in Fig. 3.11. This phenomenon was not found in a test bulk wafer with the same process conditions. It is likely that the defect is related to the thin body structure. In the UTBFETs, the drain current was less than 10pA/um after RTA at 800 °C for 1min. The mechanism for this low current is not understood yet. The channel is destroyed due to the burst. However, NMOS and PMOS UTBFETs with selectively deposited raised Ge S/D's show nice performances after RTA at 650°C ~750 °C for 1~3min [3.12][3.13].



(a) SEM picture of bubble after RTA at 900°C for 1min

(b) SEM picture of burst-out bubble after RTA at 900°C for 1min

Figure 3.11 SEM photographs of bubble defect and burst-out defect after RTA at 900°C for 1min.

Nickel-germanide is investigated for the purpose of obtaining a low parasitic series resistance. Undoped Ge is deposited with a thickness of 45nm on an oxide test wafer and patterned. The sheet resistance is 9.5kΩ. Nickel is sputtered in the CPA 9900 sputtering system. The detailed sputtering conditions are 15mTorr of pressure, 1.0kW of power, and 80cm/min of belt speed for 20nm of nickel. A different RTA condition with N_2 ambient for 1min is applied for an investigation of temperature dependence on a series resistance of nickel-germanide. Fig. 3.12 shows the sheet resistance of nickel-germanide versus annealing temperature. Diluted (10:1) HCl is used to selectively remove the 20nm Ni on the oxide. After 10min of dipping, all Ni on oxide is removed and the change of the series resistance in nickel-germanide is less than 2.5%, which shows that the nickel-germanide is not damaged and its etching selectivity is very high. After HCl etching, the

nickel-germanide profile is inspected by an optical microscope, which shows that the Ni is completely removed from the exposed buried oxide.

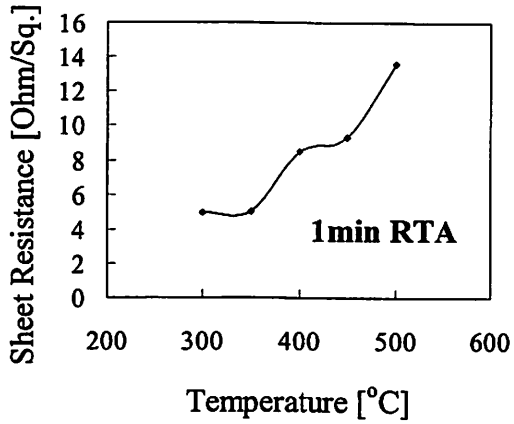
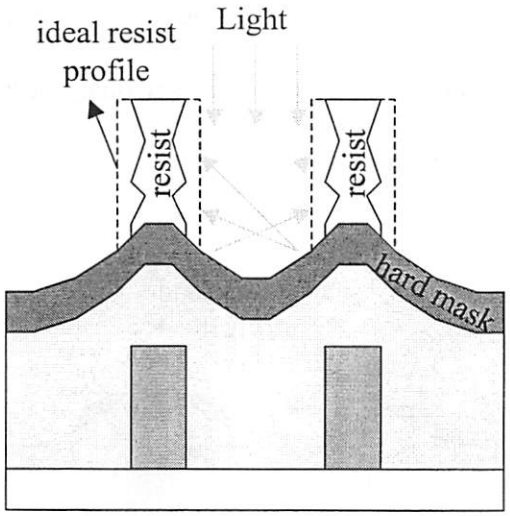


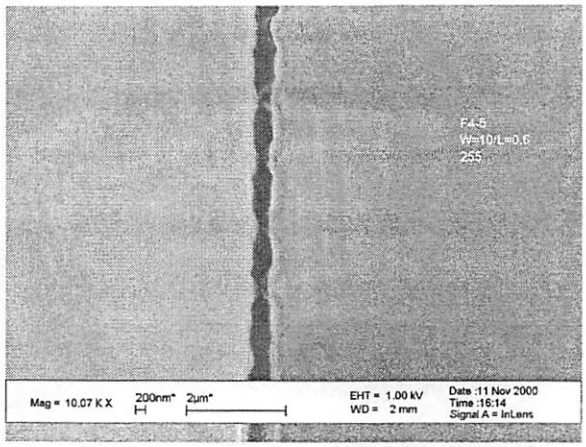
Figure 3.12 The sheet resistance of nickel-germanide increases as RTA temperature increases.

3.4 Chemical-Mechanical Polishing

The process window of lithography (e.g. the depth-of-focus (DOF) margin) is seriously degraded in a vertical structure device. When a gate is exposed with an i-line stepper, the resist profile of the gate is notched very much by reflected light at the curved surface. The non-planarized and curved surfaces are the results of the step height of the fins of FinFETs, as shown in Fig. 3.13. Even without topography of substrate patterns, a notch is made by the rough grain structures of poly-SiGe, as shown in Fig. 3.14. The notched gate causes line-edge roughness, which degrades device characteristics [3.29][3.30].

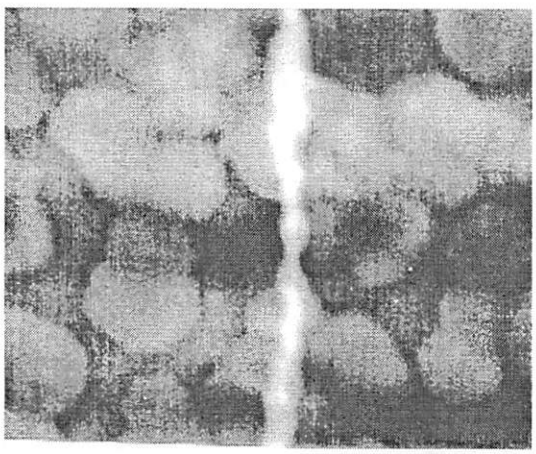


(a) A schematic diagram of cross-sectional view of notched resist

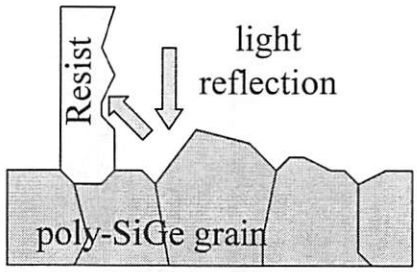


(b) Top view SEM photograph of notched gate

Figure 3.13 Cross-sectional schematic diagram and SEM top view of a notched gate after i-line lithography.



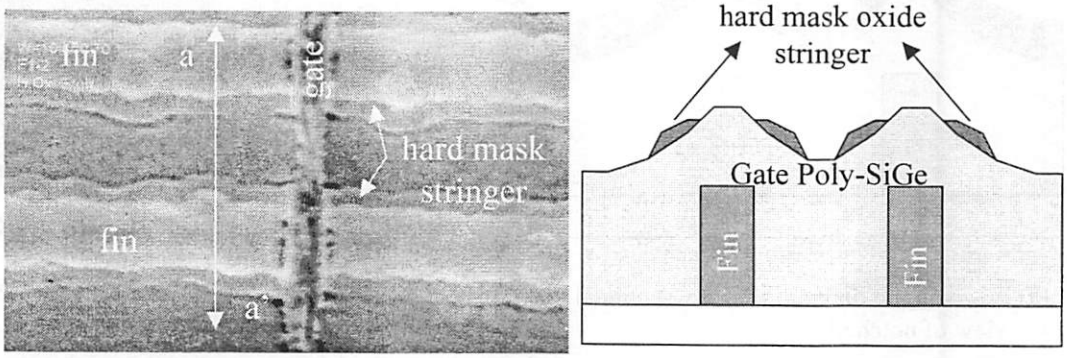
(a) SEM photograph of a resist notched by a rough grain structure of poly-SiGe



(b) Schematic of a resist notched by a rough grain structure of poly-SiGe

Figure 3.14 Cross-sectional schematic diagram and SEM top view of a notched gate after i-line lithography.

Notched resist profiles and stringers of gate hard mask oxide (Fig. 3.15) are transferred to the gate SiGe after gate plasma etch, producing notched gate poly-SiGe profiles and stingers of poly-SiGe, as shown in Fig. 3.16. The vertical structure seriously narrows the etching process window because of stingers and residues.



(a) SEM of photograph of stringers of gate hard mask oxide

(b) Drawing picture of stringers of gate hard mask oxide along a-a' direction

Fig. 3.15 SEM photograph and schematic of stringers of gate hard mask oxide.

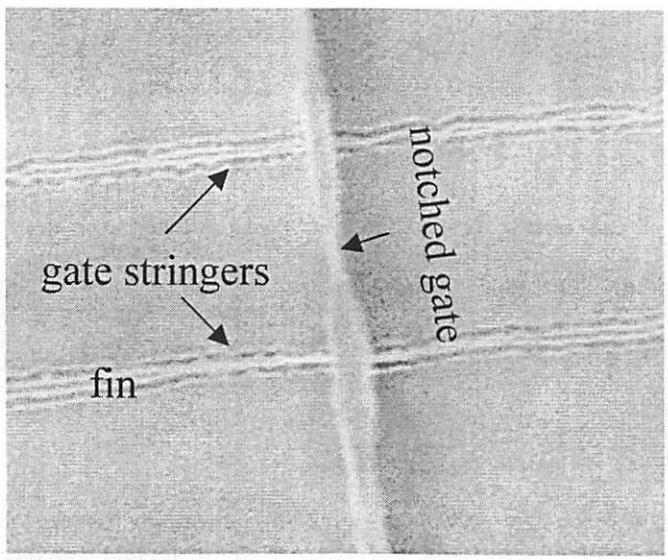


Figure 3.16 SEM photograph of a notched gate and gate hard mask oxide stringers.

Chemical-mechanical polishing (CMP) is proposed to solve the aforementioned problems and to obtain wide process windows for lithography and plasma etching. The IC1000/SUBA IV composite pad is used in a CMP experiment. The slurry used in this work consists of DI water, KOH, and oxide particles (approximately 200um diameter). The main purpose of CMP is to make a planarized undoped $Si_{0.5}Ge_{0.5}$ gate over fins of FinFETs. Detailed CMP recipes are summarized in Table 3.3.

Step	1	2	3	4
Time (sec)	15	20	Total polishing time-20	5
Down force (psi)	0	5	8	0
Table Rotation (rpm)	24	24	24	24
Chuck Rotation (rpm)	6	6	6	6
Back Pressure (psi)	-2	-1	1	-2
Table temperature (°C)	30	30	30	30
Slurry 1 (ml/min)	100	100	100	0

Table 3.3 Detailed recipes of CMP for poly-$Si_{0.5}Ge_{0.5}$.

Fig. 3.17 shows how the thickness of poly-$Si_{0.5}Ge_{0.5}$ decreases as polishing time increases. When the polishing time is over than 3min, the wafer is polished in 3 minute intervals to keep better uniformity. For example, 14 min of CMP is divided into four intervals of 3 min plus one interval of 2min. The remaining film uniformity over 4-inch wafer worsens as the polishing time increases (a cumulative slop decreases as the polishing time increases) as shown in Fig. 3.17. Measured thicknesses are collected before and after CMP at 49 points across a wafer.

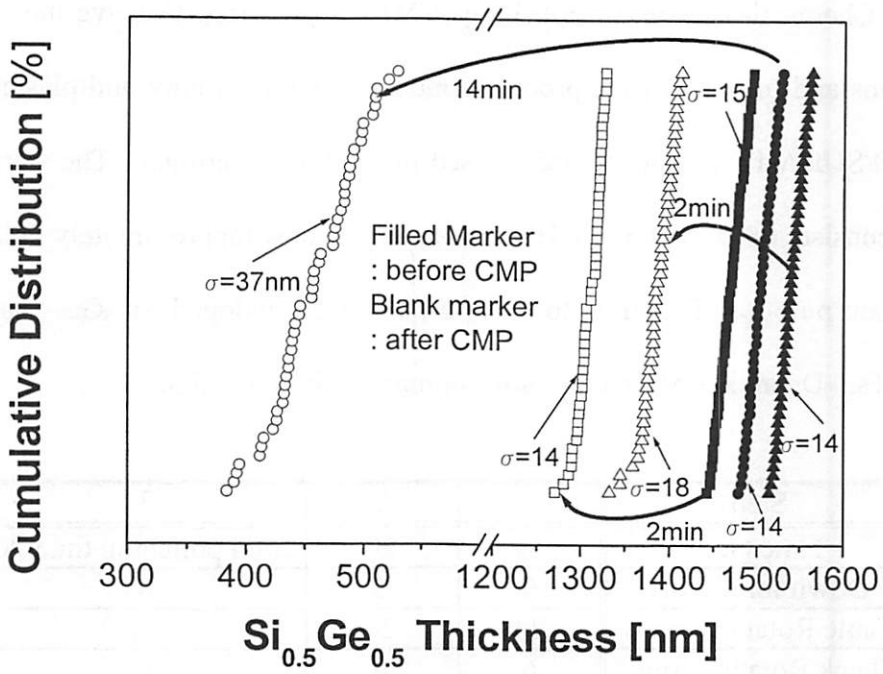


Figure 3.17 Poly-$Si_{0.5}Ge_{0.5}$ thicknesses before and after CMP. The standard deviation of film uniformity increases as polishing time increases.

As shown in Fig. 3.18, the polishing rate is high at the edge of the wafer. The enhanced polishing rate is caused by the fact that the downward pressure is high at the edge of the wafer. Film uniformity is also observed to be poor as the polishing time is increased.

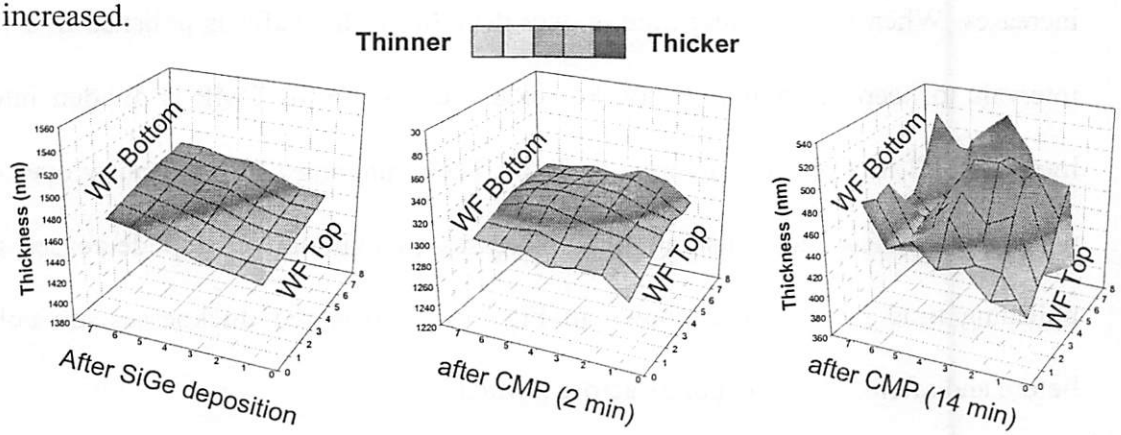


Figure 3.18 Remaining film uniformity before CMP and after CMP.

The polishing rate of poly-$Si_{0.5}Ge_{0.5}$ is lower than that of poly-Si. Ge in $Si_{0.5}Ge_{0.5}$ is not etched with KOH, which is an ingredient of the slurry. Thus, the polishing rate is expected to be reduced as the Ge content in a SiGe film increases. There are no previous reports for SiGe CMP yet. Measurements of polishing rate versus Ge content in $Si_{1-x}Ge_x$ films should be performed in the future. If this trend is reproducible, there are many applications for CMOS device fabrication (e.g. shallow trench isolation (STI) to relieve stress, and planarization of interlayer dielectric) and micro-electro-mechanical-Systems (MEMS) fabrication.

Experiments attempting to increase the polishing rate have been done. The easiest way to increase the rate is to increase the downforce. The upper limit of the downforce is the maximum force that does not break the wafer, and the lower limit is the minimum force to hold it. The polishing rates for downforces of 7, 8, and 9 psi are measured. Fig. 3.19 shows that the polishing rate increases as the downforce increases, as predicted. There is no significant variation in the standard deviation for different downforces.

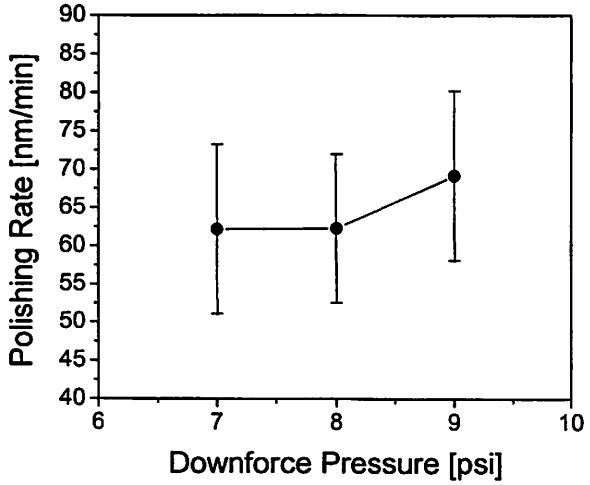


Figure 3.19 Polishing rate versus downforce. The standard deviation is independent of the downforce.

The polishing rate for different materials with the same recipe is investigated as a reference experiment. Thermal oxide and LTO show the lowest polishing rates (52nm/min and 63nm/min, respectively). The polishing rate of PSG is three times faster than that of the thermal oxide. Figure 3.20 shows that the polishing rate of poly-Si is seven times faster than that of poly-$Si_{0.5}Ge_{0.5}$.

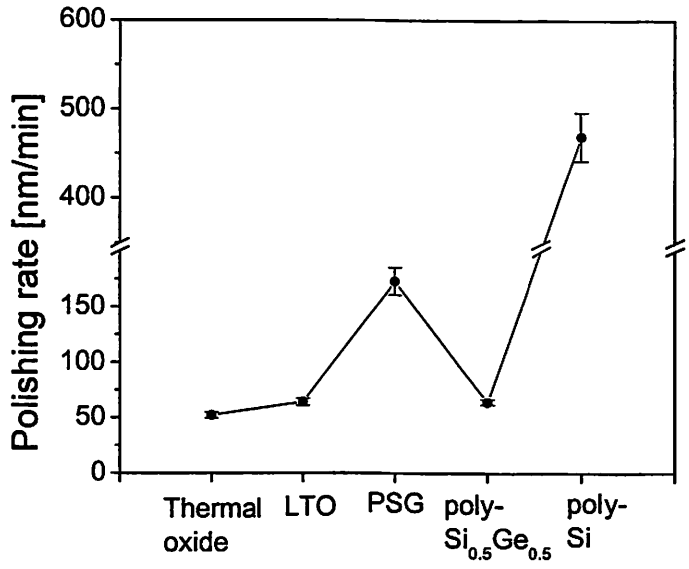


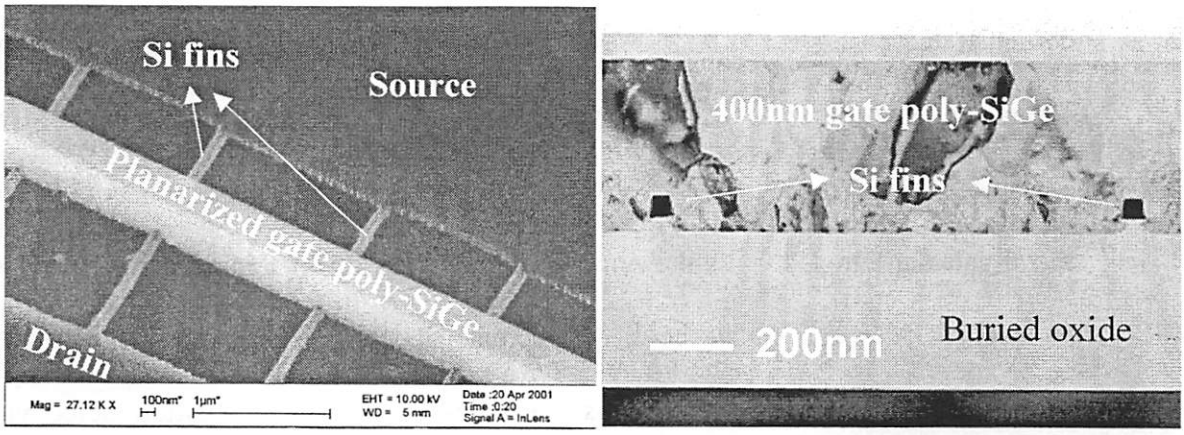
Figure 3.20 Polishing rates of thermal oxide, LTO, PSG, poly-$Si_{0.5}Ge_{0.5}$, and poly-Si.

KOH is included in the slurry as an etchant of silicon. After poly-$Si_{0.5}Ge_{0.5}$ CMP, KOH may contaminate or degrade the quality of gate oxide. Multiple post-cleanings are used to eliminate the possibility of KOH contamination [3.25][3.26]. Detailed procedures are summarized in Table 3.4.

Step	Cleaning after CMP
DI Rinse	1min
NH_4OH	1min
DI Rinse	6min
Piranha (H_2O_2:H_2SO_4 [1:5]) at 120°C	1min
DI Rinse	6min
(5:1) HF	10sec
DI Rinse	6min
SC-1(H_2O:H_2O_2:NH_4OH_4 [5:1:1]) at 65°C	5min
DI Rinse	6min

Table 3.4 Post-cleaning of poly- $Si_{0.5}Ge_{0.5}$ CMP by KOH-based slurry.

CMP is used to make planarized poly-SiGe gates over Si fins in FinFETs. For this, 1.3um-thick poly-SiGe is deposited and polished down to 0.4um by CMP for 14min. Fig. 3.21 shows the completely planarized gate profile. The DOF margin is improved from 0.2um to 0.6um with CMP because there is no gate notching. No hard mask stringers or gate poly-SiGe stringers are left as shown in Fig. 3.21 (a).



(a) Tilted SEM view of a poly-SiGe gate planarized by CMP after gate patterning

(b) Cross-sectional TEM view of a poly-SiGe gate planarized by CMP after gate patterning

Figure 3.21 SEM and TEM photographs of a SiGe gate planarized by CMP.

3.5 Conclusion

Three process technologies that allow fabrication of novel device structures are proposed and demonstrated: resist etch-back, selective Ge deposition, and chemical-mechanical polishing (CMP). A resist etch-back is applied to make a raised poly-Si source and drain (S/D) on a thin body SOI device. The poly-Si raised S/D are separated by resist and poly-Si plasma etch-back. The functional key to this etched-back process is the non-conformailty of a coated resist. UTB MOSFETs with etched-back and raised poly-Si S/D's are demonstrated successfully [3.6]. The etch–back process is a low temperature process and can be used for many applications that demand a low temperature planarization technology.

Ge is selectively deposited on thin body silicon by a conventional LPCVD furnace. It is used to make a raised Ge S/D on the thin body of UTBFETs (horizontal device) and the narrow fin of FinFETs (vertical device). Working devices have been demonstrated [3.12][3.13][3.33][3.34]. Raised S/D's formed by selective germanium deposition have much lower gate overlap capacitance than etched-back raised poly-Si S/D's [3.6]. Furthermore, the selective germanium deposition process is compatible with metal gates and high-K gate dielectrics because of its lower temperatures for deposition and RTA. A sheet resistance behavior that depends on annealing temperature is investigated for boron-, phosphorus-, arsenic-doped Ge. Deposition selectivity and opened-silicon area effects in a wafer are also investigated for in-situ boron- and phosphorus-doped Ge. Nickel-germanide ($NiGe_x$) is formed to lower the sheet resistance

of the S/D and etching process uses diluted (10:1) HCl, which provides high selectivity to $NiGe_x$, Ni, and oxide.

The CMP process is used to make a completely planarized poly-SiGe gate over Si fins in FinFETs. The process windows of lithography and plasma etching are significantly improved by the CMP. The notched gate resist problem is significantly improved since no stringers or residues of the gate hard mask oxide or the gate poly-SiGe are left after CMP. The CMP polishing rate of poly-SiGe film is investigated and is shown to be that of 1/7 poly-Si. FinFETs with implementation of CMP are demonstrated for the first time and show excellent device performances [3.34][3.35].

3.6 References

[3.1] R.H. Dennard, F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassous and A. R. LeBlanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE Journal of Solid-State Circuits*, vol.9, p.256-268, 1974.

[3.2] C. Hu, "MOSFET scaling in the next decade and beyond," *Semiconductor International*, p.105-114, June 1994.

[3.3] B. Davari, "CMOS technology scaling: 0.1 μm and beyond," *IEDM Technical Digest*, p.555-558, 1996.

[3.4] B. Yu, Y.-J. Tung, S. Tang, E. Hui, T.-J. King, and C. Hu, "Ultra-thin-body Silicon-On-Insulator MOSFET's for terabit-scale integration", *International Semiconductor Device Research Symposium*, p. 623-626, 1997.

[3.5] V. Subramanian, J. Kedzierski, N. Lindert, H. Tam, Y. Su, J. McHale, K. Cao, T.-J. King, J. Bokor, C. Hu, "A bulk-Si-compatible ultrathin-body SOI technology for sub-100 nm MOSFETs", *57th Device Research Conference,* p.28-9, 1999.

[3.6] Y.-K. Choi, K. Asano, N. Lindert, V. Subramanian, T.-J. King, J. Bokor, C. Hu, "Ultra-thin-body SOI MOSFET for Deep-sub-tenth Micron Era," *IEEE Electron Device Letters,* vol. 21, p.254-255, 2000.

[3.7] S. Nagao, K. Higashitani, Y. Akasaka, H. Nakata, "Applications of selective silicon epitaxial growth for CMOS technology", *IEEE Transactions on Electron Devices*, vol.33, p.1738-44, 1986.

[3.8] J. Borland, R. Wise, Y. Oka, M. Gangani, S. Fong, Y. Matsumoto, "Silicon epitaxial growth for advanced device structures", *Solid State Technology*, vol.31, p.111-19, 1988.

[3.9] A. Srivastava, J. Sun, K. Bellur, R.F. Bartholomew, P. O'Neil, S.M. Celik, C.M. Osburn, N.A. Masnari, M.C. Ozturk, R. Westhoff, B. Fowler, "A 0.18 um CMOS technology for elevated source/drain MOSFETs using selective silicon epitaxy", *Proceedings of the Sixth International Symposium on Ultralarge Scale Integration Science and Technology (ULSI Science and Technology)*, p.571-85, 1997.

[3.10] C. Wang, J.P. Snyder, J.R. Tucker, "Sub-40 nm PtSi Schottky source/drain metal-oxide-semiconductor field-effect transistors", *Applied Physics Letters*, vol.74, p.1174-6, 1999

[3.11] J.R. Tucker, "Schottky barrier MOSFETs for silicon nanoelectronics", *Advanced Workshop on Frontiers in Electronic (WOFE '97 Proceedings)*, p.97-100, 1997.

[3.12] Y.-K. Choi, D. Ha, T.-J. King, C. Hu, "Nanoscale ultrathin body PMOSFETs with raised selective germanium source/drain", *IEEE Electron Device Letters*, vol.22, p.447-8, 2001.

[3.13] Y.-K. Choi; Y.-C. Jeon, P. Ranade, H. Takenuchi, T.-J. King, J. Bokor, C. Hu "30 nm ultra-thin-body SOI MOSFET with selectively deposited Ge raised S/D", *58th Device Research Conference*, p.23-4, 2000

[3.14] T. Uchino, T. Shiba, K. Ohnishi, A. Miyauchi, M. Nakata, Y. Inoue, T. Suzuki, "A raised source/drain technology using in-situ P-doped SiGe and B-doped Si for 0.1-um CMOS ULSIs", *IEDM Technical Digest*, p.479-82, 1997

[3.15] N. Lindert, Doctoral Thesis, "Process Development and Device Design for Continued MOSFET Scaling", Department of Electrical Engineering and Computer Science, *University of California at Berkeley*, 2001.

[3.16] J. Kedzierski, Doctoral Thesis, "Complementary Silicide Thin-Body Silicon-on-Insulator CMOS Devices", Department of Electrical Engineering and Computer Science, *University of California at Berkeley*, 2001.

[3.17] M. Moslehi, "Low-temperature in-situ dry cleaning process for epitaxial layer multiprocessing," *Proceedings of the SPIE - The International Society for Optical Engineering*, vol.1393, p.90-108, 1991

[3.18] L.H. Blake, E. Mendel, "Chemical-mechanical polishing of silicon", *Solid State Technology*, vol.13, p.42-6, 1970.

[3.19] H.-S. Kim, D.-H. Ko; D.-L. Bae, K. Fujihara, H.-K. Kang, " The formation of Ti-polycide gate structure with high thermal stability using chemical-mechanical polishing (CMP) planarization technology", *IEEE Electron Device Letters*, vol.20, p.86-8, 1999.

[3.20] T.R. White, W.J. Ciosek, E.J. Prinz, C.F. King, R. Blumenthal, C.W. Stager, B.M. Somero, M.P. Woo, U. Sharma, R.W. Fiordalice, J.L. Klein, "Extending resist etch back planarization to 0.5 um logic and ASIC circuits", *Proceedings of the SPIE - The International Society for Optical Engineering*, vol.2090, p.93-102, 1993.

[3.21] Y.-K. Choi, Master Thesis, "Ultra-thin-body SOI MOSFET for Deep-sub-tenth Micron Era", Department of Electrical Engineering and Computer Science, *University of California at Berkeley*, 1999.

[3.22] H. Ishii, Y. Takahashi, J. Murota, "Selective Ge deposition on Si using thermal decomposition of GeH_4", *Applied Physics Letters*, vol.47, p.863-5, 1985.

[3.23] S.M. Sze, *Physics of Semiconductor Devices*, John Wiley & Sons, New York, 2nd Edition, 1981

[3.24] P.-E. Hellberg, A. Gagnor, S.-L. Zhang, C.S. Petersson, "Boron-doped polycrystalline Si_xGe_{1-x} films: dopant activation and solid solubility", *Journal of the Electrochemical Society*, vol.144, p.3968-73, 1997.

[3.25] T.-J. King, J.P. McVittie, K.C. Saraswat, "Electrical Properties of Heavily Doped Polycrystalline Silicon-Germanium Films", *IEEE Transactions on Electron Devices*, vol.41, p. 228-32, 1994.

[3.26] D.S. Bang, M. Cao, A. Wang, and K. C. Saraswat, "Resistivity of boron and phosphorous doped polycrystalline $Si_{1-x}Ge_x$ films", *Applied Physics Letter*, vol.66, p.195-97, 1995.

[3.27] This experiment was performed by Hideki Takeuchi at the Microlab of the University of California-Berkeley. It will be submitted to a journal.

[3.28] This experiment was performed by Pushkar Ranade at the Microlab of the University of California-Berkeley. It will be submitted to a journal.

[3.29] C.H. Diaz, H.-J. Tao; Y.-C. Ku; A. Yen, K. Young, "An experimentally validated analytical model for gate line-edge roughness (LER) effects on technology scaling", *IEEE Electron Device Letters*, vol.22, p.287-9, 2001.

[3.30] P. Oldiges, Q. Lin, K. Petrillo, M. Sanchez, M. Ieong, M. Hargrove, "Modeling line edge roughness effects in sub 100 nanometer gate length devices", *2000 International Conference on Simulation Semiconductor Processes and Devices,* p.131-4, 2000.

[3.31] F. Tardif, I. Constant, T. Lardin, O. Demolliens, M. Fayolle, Y. Gobil, J. Palleau, J. Torres, "Cleaning after silicon oxide CMP", *Microelectronic Engineering*, vol.37-38, p.285-91, 1997.

[3.32] Y.-J. Seo, W.-S. Lee, S.-Y. Kim, J.-S. Park, E.-G. Chang, "Optimization of post-CMP cleaning process for elimination of CMP slurry-induced metallic contaminations", *Journal of Materials Science: Materials in Electronics*, vol.12, p.411-15,2001

[3.33] N. Lindert, Y.-K. Choi, L. Chang, E. Anderson, W.-C. Lee, T.-J. King, J. Bokor, C. Hu, "Quasi-Planar FinFETs with Selectively Grown Germanium Raised Source/Drain", *IEEE International SOI Conference*, p.111-2, 2001

[3.34] Y.-K. Choi, N. Lindert, P. Xuan, S. Tang, D. Ha, E. Anderson, T.-J. King, J. Bokor, C. Hu, "Sub-20nm CMOS FinFET Technologies", *IEEE IEDM Technical Digest*, p.421-24, 2001.

[3.35] Y.-K. Choi, T.-J. King, C, Hu, "Nanoscale CMOS Spacer FinFET for the Terabit Era", *Electon Device Letters*, to be published, 2002

Chapter 4

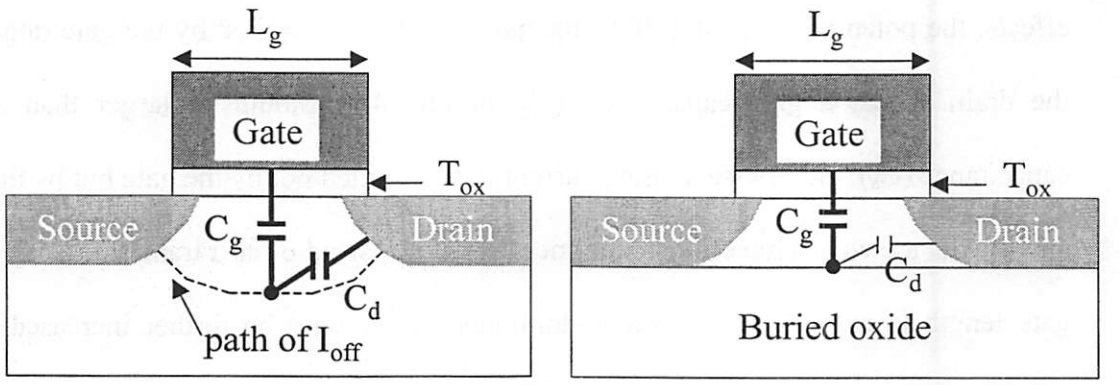
Ultra-Thin Body MOSFETs

4.1 Introduction

Technological advancements in MOSFET have been achieved over the past three decades primarily through the scaling of device dimensions [4.1]-[4.3] in order to attain continued improvement in circuit speed and reduction in size. To suppress short-channel effects, the potential in a MOSFET channel should be controlled by the gate rather than the drain. Thus, a gate capacitance (C_g, in Fig. 4.1) should be larger than a drain capacitance (C_d). Otherwise, a drain current is determined not by the gate but by the drain, and the device will behave like a nonlinear resistor instead of as a transistor [4.4]. As the gate length is reduced, C_d becomes dominant. C_g should be further increased with a reduction of gate oxide thickness (T_{ox}). The ITRS [4.5] predicts that an equivalent oxide thickness (EOT) as a gate dielectric is 0.5~0.6nm (2 mono-layers of SiO_2) for a 35nm technology generation. Yu et al. [4.6] reported a 35nm gate length MOSFETs with a 0.7nm EOT, and Chau et al. [4.7] announced a 30nm gate length MOSFETs with 0.8nm T_{ox}. A 2nm oxide is considered the lower limit for oxide scaling because of gate tunneling current ($1A/cm^2$ *at* $V_g=2V$) [4.8]. One solution proposed to alleviate oxide tunneling

current is to use gate dielectric materials with permittivities higher than that of SiO_2 [4.9]-[4.11]. CMOS process compatibility and reliability are issues for high-permittivity (high-K) gate dielectrics. For scaling beyond a 35nm technology node, how thin of an oxide thickness is needed? How much scaling is left? A channel potential at the interface of Si-SiO_2 is well controlled by the gate with an ultra-thin gate dielectric. The main leakage path, as shown in Fig. 4.1, crosses through the middle of L_g and far away from the gate where gate control of channel potential is weak. The suppression of off-state leakage current is an increasingly difficult technological challenge-one that will ultimately limit the scalability of the conventional MOSFET structure.

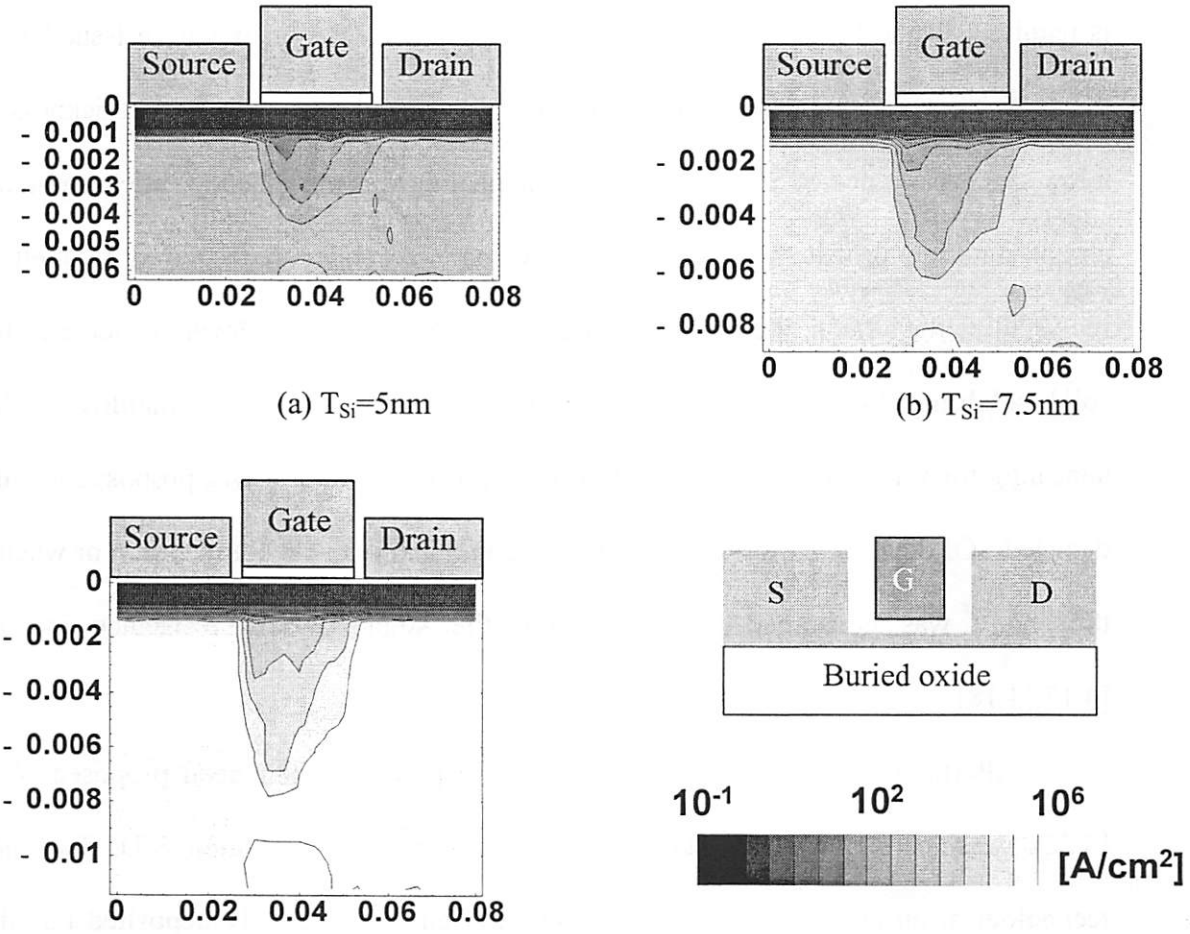
One solution to suppress off-state leakage current is to eliminate the silicon which is least effectively modulated by the gate (i.e. to remove a bottom part of silicon as shown in Fig. 4.1).



(a) Conventional bulk MOSFET (b) Thin body MOSFET

Figure 4.1 Schematic diagrams of conventional bulk silicon MOSFET (a) and thin body silicon-on-insulator MOSFET (b)

A NMOS leakage current density profile was obtained from a 2-D device simulator [4.14] for $L_g = 25nm$, $T_{ox} = 1.5nm$, and $N_{body} = 1x10^{15}cm^{-3}$ as shown in Fig. 4.2 for $V_g=0V$ and $V_d=0.7$V. For reduced body thicknesses, the off-state leakage current significantly decreases. Fig. 4.2 shows that the highest leakage current flows through the bottom of the thin body. Thus, an off-state leakage current is suppressed by thinning the silicon body, thereby accommodating further scaling beyond the end of the roadmap without requiring aggressive gate oxide scaling.



(a) T_{Si}=5nm

(b) T_{Si}=7.5nm

(c) T_{Si}=10nm

Figure 4.2 NMOS Leakage current density for a different body thickness ($T_{Si}=5nm$, $7.5nm$, $10nm$) at $V_g=0V$, $V_d=0.7V$. L_g is 25nm, T_{ox} is 1.5nm, and N_{body} is $1x10^{15}cm^{-3}$.

Even though UTB MOSFETs is highly attractive in terms of overwhelming suppression of short-channel effects, they are undesirable because of the large sheet resistance of thin body. Junction depth (X_j) has been scaled down with the gate length ($X_j \sim 0.4*L_g$) in order to minimize sub-surface leakage current. The formation of ultra-shallow junctions is a significant technological challenge because of conflicting demands of a junction depth as shallow as possible and a sheet resistance as low as possible to maintain a high drive current. A way to decouple the junction depth and sheet resistance is required. In the UTB devices, the shallow junction formation is not an issue because the depth of source and drain junctions is naturally limited to the thin body thickness. For a low sheet resistance of S/D junction, laser annealing was proposed in order to obtain an abrupt junction profile because only a few nano-seconds are required to melt and recrystallize the silicon, thereby achieving high dopant activation levels which are above solid solubility [4.15][4.16]. There are no reports describing availability of laser annealing for UTB. On the other hand, a self-aligned silicidation was proposed for fully-depleted SOI devices. However, when the silicon thickness was below 50nm or when the thin body was consumed by silicidation, the silicided S/D resistance increased [4.17][4.18].

Alternatively, a self-aligned elevated S/D processes had been proposed [4.19]-[4.22]. A selective silicon epitaxial growth was used for elevated S/D. Two novel technologies: an etched-back raised poly-Si S/D and a selectively deposited raised Ge S/D were proposed and demonstrated in the chapter 3. Each has its own benefits and drawbacks. Device performances for two novel technologies are discussed in this chapter.

Traditional CMOS devices have been developed and scaled down by adopting high channel doping, a thin gate oxide, shallow S/D junctions. The UTB structure provides excellent short channel effects with a less aggressively scaled gate oxide thickness and ultra-shallow junctions of S/D which are naturally limited by the body thickness. In bulk MOSFETs, a high channel doping results in mobility degradation, large S/D junction capacitance, and junction leakage. The UTB device does not demand high channel doping to suppress short-channel effects. It has been highly recommended that nanoscale MOSFETs be designed with a thin, undoped silicon channel in order to avoid random dopant fluctuations which can cause a variation of threshold voltage [4.26][4.27]. A threshold voltage in thin body SOI devices is preferably controlled by gate work function engineering [4.28][4.29]. Details of process flow and device characteristics for UTB MOSFETs with raised S/D are shown in this chapter.

4.2 Simulations

Fully depleted Silicon-On-Insulator (SOI) technology has the advantages of lower junction capacitance and better subthreshold swing [4.30]. However, the conventional fully depleted SOI MOSFET is known to have worse short-channel effects than bulk MOSFETs and partially depleted SOI MOSFETs [4.31]. Fig. 4.3 shows that the UTB SOI device can be scaled down to 18nm with < 5nm body thickness. The ultra-thin-body (UTB) structure eliminates the leakage paths between source and drain [3]. Nearly all the leakage current at $Vg=0$ in the $Tsi=7nm$ flows along the bottom 2nm of the body, which is least strongly controlled by the gate. Therefore eliminating this 2nm (i.e. making $Tsi=5nm$) reduces the leakage current by 30 times. Through a reduction of 4nm body

thickness, off-state leakage current decreased by 1000 times as shown Fig. 4.3. The data in Fig. 4.3 were obtained through device simulation (MEDICI) assuming a low and uniform body doping ($10^{15}cm^{-3}$) and simple Gaussian S/D doping profiles (peak concentration = $10^{20}cm^{-3}$, surface doping concentration under the gate edge = $7.1x10^{16}$ cm^{-3}). An energy balance model without quantum effect consideration was used. Work function for gate electrode was assumed to be 4.74eV. It meets the goals of $<3nA/um$ I_{off} and $>600uA/um$ I_{on} for $L_g=18nm$ with $1.5nm$ gate oxide.

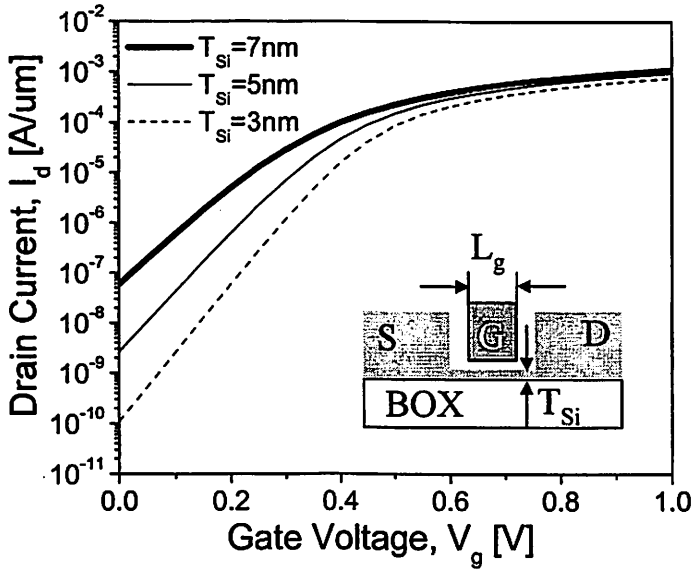


Figure 4.3 Impact of body thickness on the I_d-V_g characteristics of UTB device. ($T_{ox}=1.5nm$, $N_{body}=1x10^{15}cm^{-3}$, $V_d=1V$).

When the gate length is scaled down, the power supply voltage should be decreased as well to keep the device power and electric field strength within reasonable limits. However, the threshold voltage has not been scaled in proportional to the power supply voltage. This comes from the fact that subthreshold slope is mainly governed by thermally activated diffusion and is independent of power supply voltage and channel length. For further device scaling, subthreshold swing should be kept as small as possible.

As the body thickness decreases, subthreshold swing reduces significantly and converges to the ideal value of *60mV/dec* as shown in Fig. 4.4. It is insensitive to the body doping concentration; therefore, it should be controlled by the body thickness instead of the doping concentration.

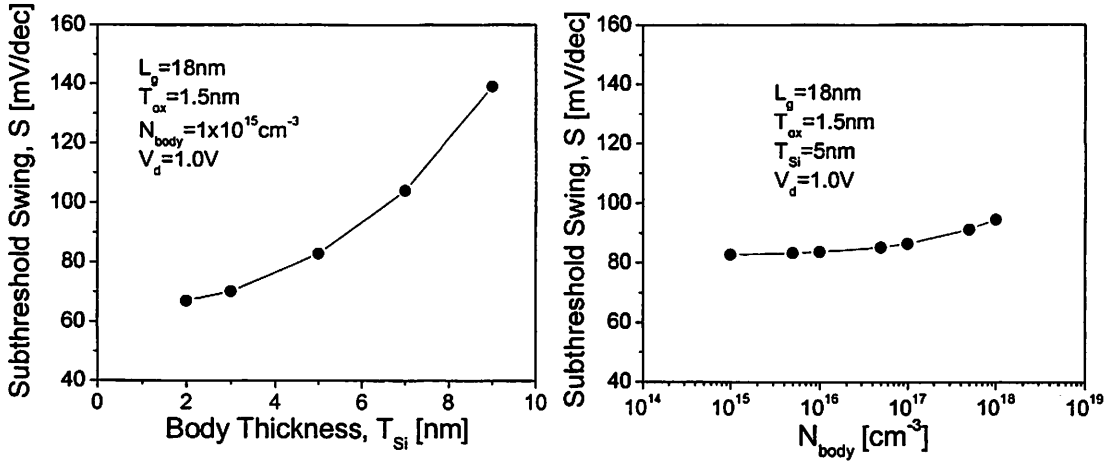


Figure 4.4 Subthreshold swing dependence on body thickness and body doping concentration.

A behavior of the threshold voltage for different body thickness was simulated. The threshold voltage shift by an inversion charge confinement with a quantum mechanical effect was not considered in this work when the body doping concentration is above $1x10^{18}cm^{-3}$ [4.32][4.33]. Fig. 4.5 shows that a change of body doping concentration is not effective to adjust the threshold voltage in UTB devices. Higher doping concentrations degrade mobility [4.34] and cause random dopant fluctuations, resulting in the statistical variation of the threshold voltage [4.35]. Undoped or low doping is highly preferred. The threshold voltage adjustment with gate work function engineering had been proposed. A metal gate with a tunable technology [4.36] and dual

gate work function [4.37][4.38] was announced. An approach with p-type doped $Si_{1-x}Ge_x$ by changing Ge fraction in $Si_{1-x}Ge_x$ was reported [4.39]-[4.41].

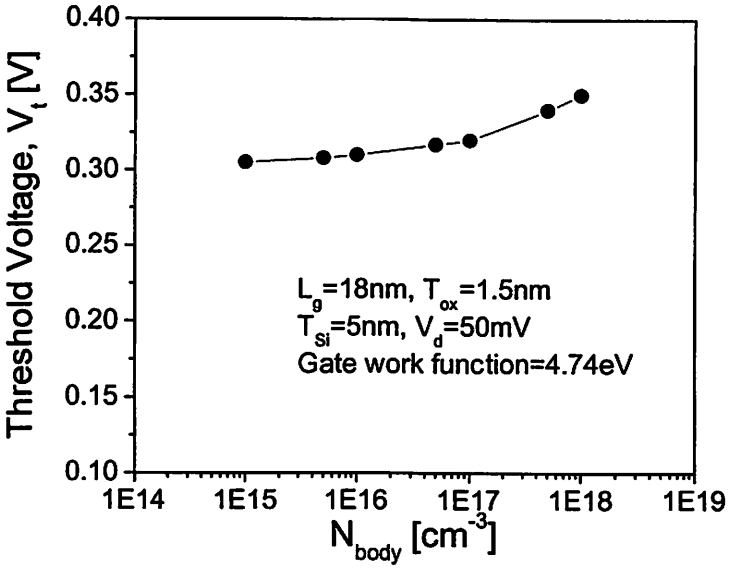


Figure 4.5 Threshold voltage behavior of NMOS relying on a body doping concentration.

As transistor sizes are scaled down, there is increasing concern that series resistance may limit the ultimate performance of the scaled devices [4.42][4.43]. If we designate I_{d0} as the drain current without the effect of R_{sd} ($=R_s+R_d$), the drain current including R_{sd} will be

$$I_d = \frac{V_d}{R_{ch}+R_{sd}} = \frac{V_d / R_{ch}}{1+R_{sd}/R_{ch}} = \frac{I_{d0}}{1+R_{sd}I_{d0}/V_d}$$

where R_{ch} is the intrinsic channel resistance. R_{sd} is composed of a bias dependent part in series with a bias-independent part. Even though this equation cannot tell the behavior of drive current depending on the series resistance explicitly, it can be qualitatively deduced that the drain current decreases as the series resistance increases. In the proposed UTB MOSFETs, this series resistance can be reduced by shortening the length of the S/D

extension region, which is achieved by using thinner spacers as shown in Fig. 3.1 and Fig. 3.4. However, it increases gate-to source and gate-to-drain overlap capacitances. For small D in the inset of Fig. 4.7, the overlap capacitance rises up rapidly as D decreases. When D is large, it varies very slowly with D. So there should be a value of D at which the performance reaches the optimized point [4.44]. To investigate how series resistance reduces the drive current quantitatively, it was simulated with the aid of 2-D MEDICI simulator [4.14] by changing the gap, D between the gate and source. Fig. 4.7 shows that drive current was enhanced when the gap was reduced as predicted.

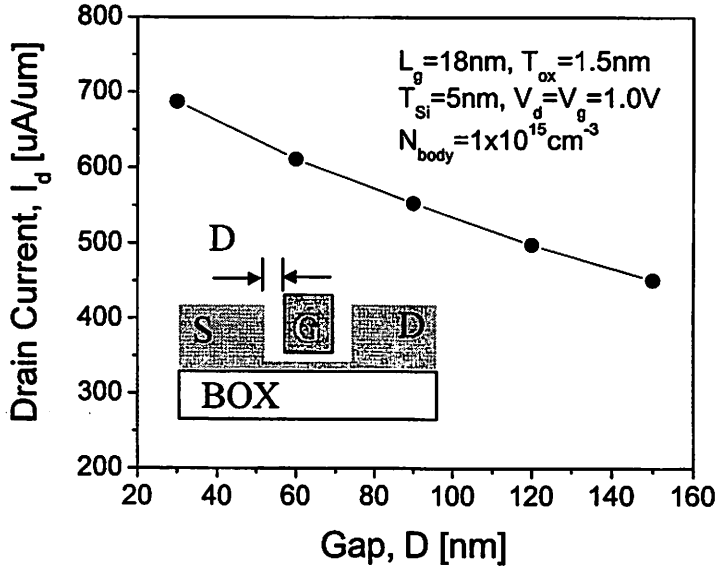


Figure 4.7 Drive current dependence of a series resistance of source and drain.

4.3 Process Developments

The starting material was (100) SOI p-type (N_{body} =$1x10^{15}cm^{-3}$) wafer with a 400nm buried oxide produced by SOITEC, Inc with Smart Cut and Unibond technology [4.45]. The average of initial film thickness was 100nm and the standard deviation is

1.13nm. The initial film uniformity is very critical to make a uniformed body, which governs device characteristics [4.46]. A thickness of 100nm was reduced to 25nm with wet oxidation at 850°C. And it was reduced further from 20nm to 3nm with multiple dry oxidations at 900°C. For uniform oxidation, the flat zone of wafers faced up during all oxidations. Fig. 4.8 shows thinned body uniformity across 4-inch wafer. It was deduced that an oxidation rate was faster at the bottom (near the flat zone) of wafer. After body thinning oxidation, the standard deviation was 1.36nm, which did not worsen the uniformity of initial silicon film thickness. Silicon film thickness was measured with Nanometrics 210 XP Scanning UV and calibrated with TEM. For a more uniform body, it is highly recommended that all wafers should be loaded alternatively between face-up and face-down for each oxidation.

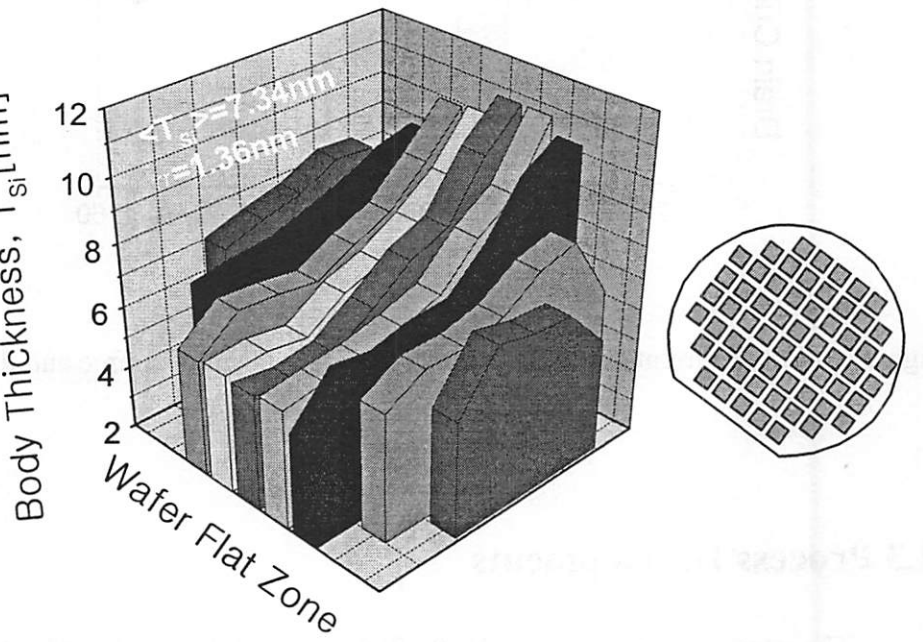


Figure 4.8 Uniformity of silicon film thickness after body thinning oxidation across 4inch wafer. One pixel denotes one die.

One interesting phenomena found was an anomalous oxidation. Oxidation rate was fast when the body thickness is thinner than 10nm as shown in Fig. 4.9. Oxidation rate was extracted by using the fact that the thickness of silicon consumed is 44% times of thickness of SiO_2 formed. Fig. 4.10 shows that the variation of oxide thickness increases as the body thickness decreases. Commercial process simulators such as TSUPREM4 and SILVACO did not show this abnormal oxidation behavior, and there were no reports about it. Further study is required to understand and model it.

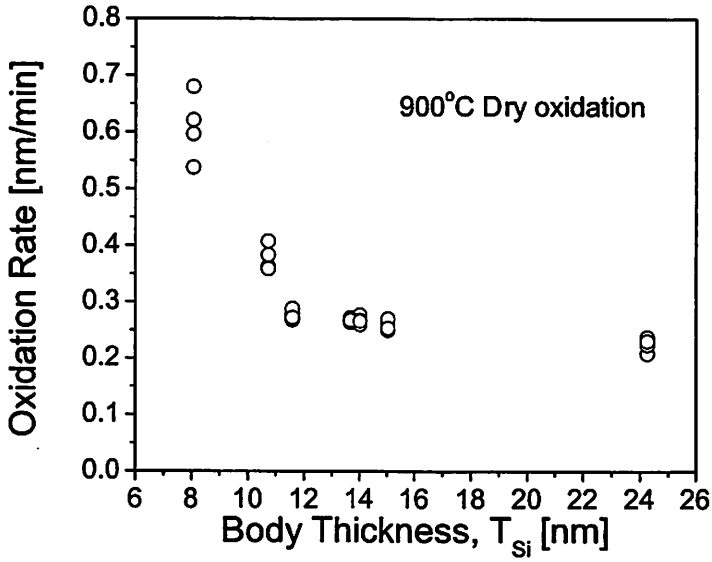


Figure 4.9 Oxidation rate vs. the thin thickness. One circle is the 5 die average value in a 4-inch wafer.

To isolate devices, the thinned silicon body was etched with HBr plasma. Details of etch recipe are 13mTorr of pressure, 70W of top RF power, 10W of bottom RF power, and 100sccm of CF_4 for removal of native oxide removal and 15mTorr of pressure, 70W of top RF power, 30W of bottom RF power, 50sccm of HBr, 1sccm of O_2, and 50sccm of He for silicon etch. The ideal features of CMOS device isolation are perfect planarity,

defect free levels, low leakage current, and process simplicity [4.47]. An isolation technology with a simple plasma etch in UTBFETs satisfies all requirements of ideal isolation. Since the thin body thickness was thinner than 20nm, a planarization technology was not needed. Leakage current between devices and S/D junction leakage are kept tremendously low with SOI structures. It dose not use any high stress film such as a nitride and high temperature process which can potentially cause defects. A single lithography step and silicon mesa etch are good enough for device isolation in UTBFETs. In contrast, a conventional shallow trench isolation (STI) is composed at least 4 steps: silicon trench formation, trench filling with CVD oxide, CMP, channel stop implantation [4.48]. To improve the isolation properties such as leakage current [4.49] [4.50], dishing [4.51], and corner rounding [4.52], more process steps are required. Thus, isolation technology of UTBFETs with a simple silicon mesa etch will become a great benefit in terms of cost reduction and higher yield.

A gate oxide was thermally grown at 750°C for 12minutes, which produced a thickness of 2.1nm. In-situ N_2 annealing was used to improve the gate oxide quality at 900°C and for 30minutes. Loading and unloading temperature is kept below 500°C in order to minimize additional oxide growth at the beginning stage of the gate oxidation. Three different gate materials: in-situ n-type (N+) doped poly-Si, p-type (P+) doped $Si_{0.5}Ge_{0.5}$, and TiN, are used for tuning the threshold voltage. For poly-Si gate and poly-$Si_{0.5}Ge_{0.5}$ gate, silicon and silicon-germanium film are deposited on the gate oxide with LPCVD furnace. For TiN gate, a thickness of 3nm TiN was sputtered with a CPA 9900 sputtering system for metal gate. The detailed process condition includes of 1kW of power, 20mTorr of pressure ($Ar:N_2=15:5$), and a 80cm/min of displacement speed. In-

situ N+ poly-Si was deposited on TiN with a thickness of 200nm. A hard mask oxide was deposited on the gate poly-Si or $Si_{0.5}Ge_{0.5}$ with a thickness of 100nm to 200nm. A LTO or HTO was used as the gate hard mask oxide. The main purpose of the gate hard mask oxide is to provide a pattern reduction for trimming with HF, to avoid counter-doping during N+ or P+ implantation for S/D, to shield the gate for a subsequent selective Ge deposition or to use for resist and poly-Si etch-back as described in the chapter 3. For P+ $Si_{0.5}Ge_{0.5}$, a LTO is recommended to minimize boron penetration because HTO deposition temperature was 800°C, and its temperature stabilization can take more than a few hours in the worst case. The gate pattern was delineated with i-line lithography or spacer technology as described in the chapter 2. Ashing-trimming reduced a 0.5um line width to sub-30nm for a technology using i-line lithography. Except for TiN, detailed gate etch processes are described in the chapter 2 for a gate hard mask oxide etch, trimming, and poly-Si/poly-$Si_{0.5}Ge_{0.5}$ plasma etch. The important thing is that the micro-trench in Fig. 4.10 should be avoided because a portion of the channel will be disconnected with S/D area. For TiN etch, 35mTorr of pressure, 250W of top RF power, 120W bottom of RF tower, 200sccm of HBr, and 5sccm of O_2 are used, resulting in an etch rate was 50nm/min and a selectivity of TiN to thermal oxide of 20:1.

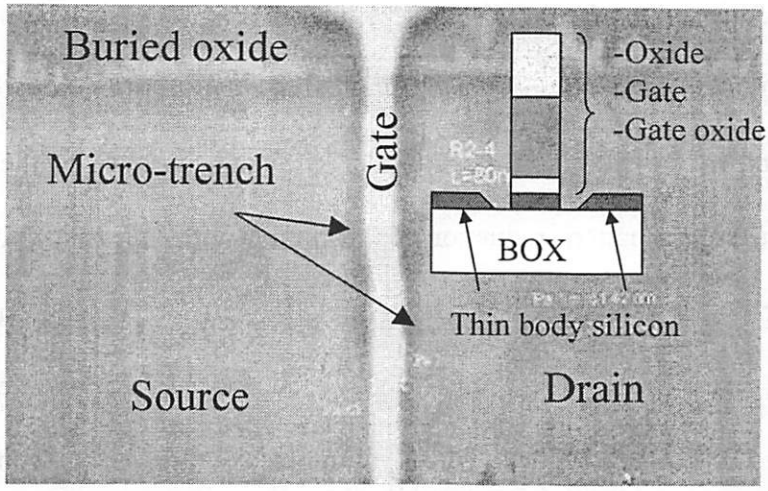
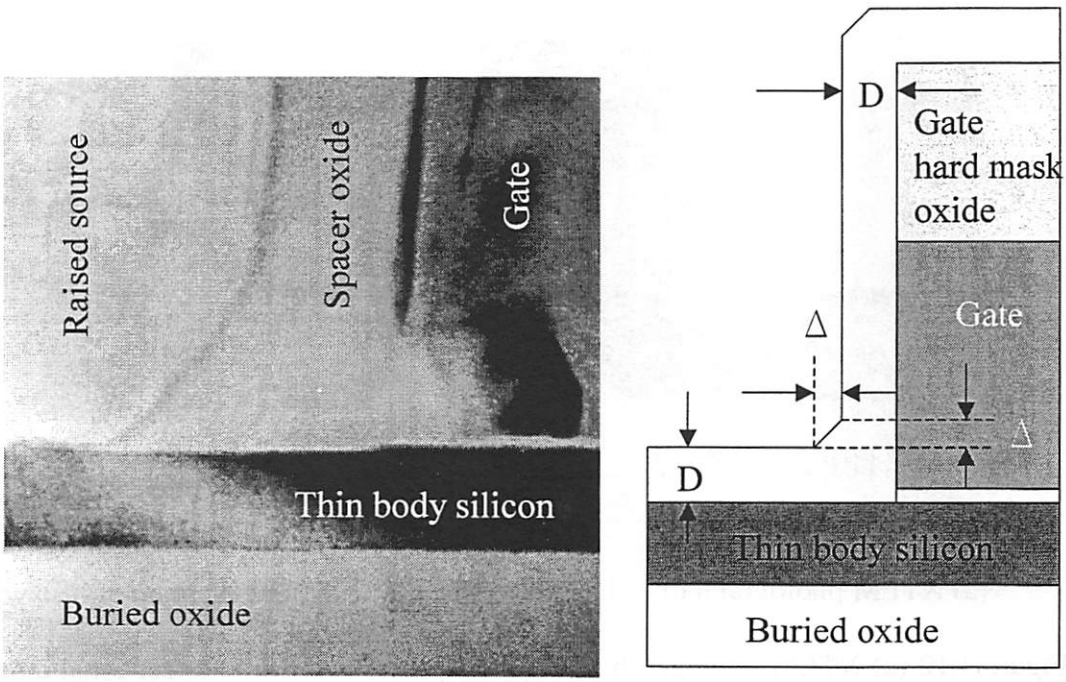


Figure 4.10 Top view of SEM photograph showing the gate and thin body silicon profile after the gate plasma etch with CF_4 causing micro-trench. Inset is the cross-sectional diagram of the SEM photograph.

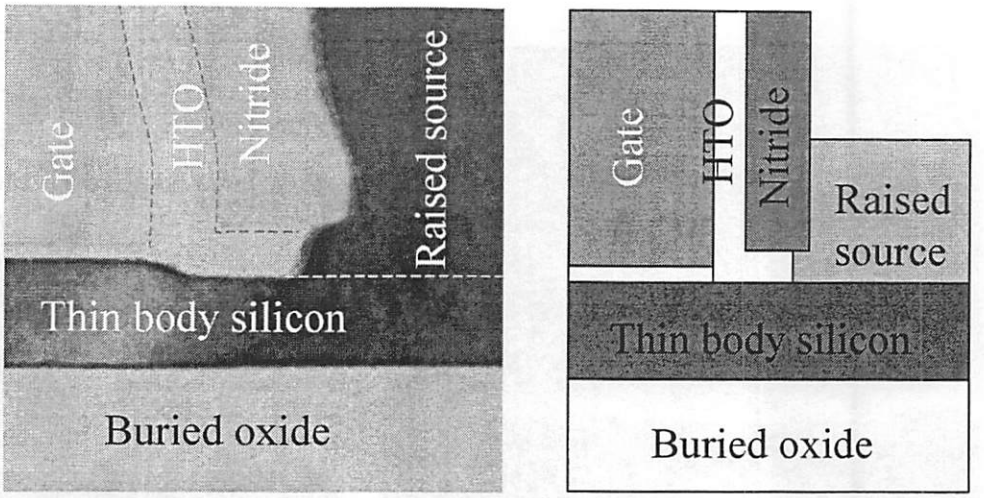
Before S/D implantation, spacers are formed to avoid a short between the gate and raised S/D. Removal of spacer tails is very crucial to reduce the series resistance. In the first trial, a 30nm HTO was deposited as spacer material and etched by CF_4 plasma. Selectivity of HTO to the thin silicon body is close to 1:1. Any overetch of spacer HTO cannot be allowed in order to minimize damage of the thin silicon body. Since, the HTO film thickness in Fig. 4.11 (b) is thicker $(D+\Delta)$ than the deposited thickness (D) at the corner of the gate edge and the thin body silicon, the spacer tail can exist without an overetch.



(a) XTEM photograph of tailed spacer (b) Schematic diagram of tailed spacer

Figure 4.11 (a) XTEM photograph after spacer etch and (b) cross-sectional diagram of tailed spacer before spacer etch.

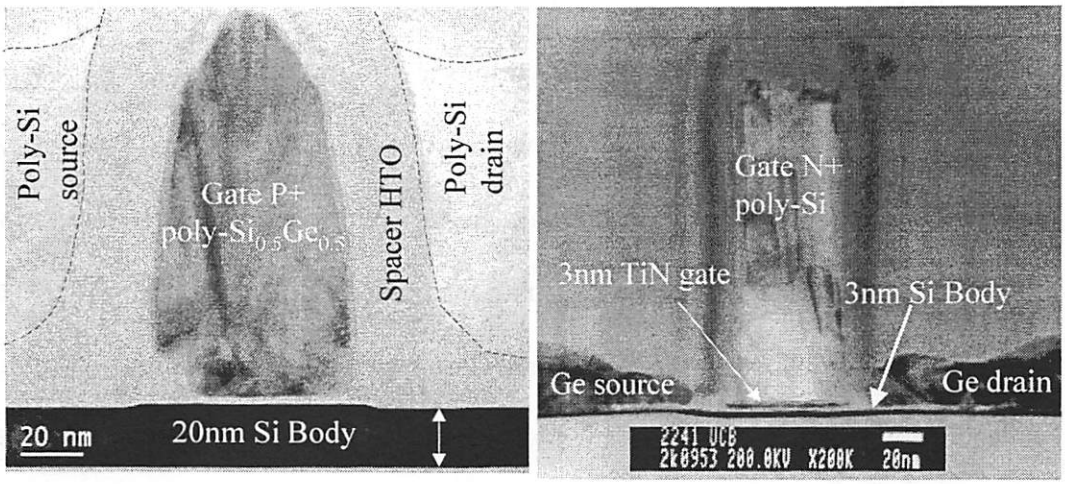
A bi-layer spacer structure is proposed to solve this problem. HTO is deposited by LPCVD with a thickness of 10nm after gate patterning. Next, 20nm of nitride is deposited by LPCVD on the HTO. The outer nitride is etched with Cl_2 based plasma. The process conditions of the plasma nitride etch is 15mTorr of pressure, 150W of top RF power, 75W of bottom RF power, 50sccm of Cl_2, and 150sccm of HBr. The etch rate of nitride is 16nm/min. 40% overetch is used to etch back the nitride spacer. The selectivity of nitride to HTO is 3:1. The exposed HTO is removed with (25:1) HF and an undercut is formed in order to reduce the series resistance as shown in Fig. 4.12 (a) and (b).



(a) XTEM photograph of bi-layer spacers (b) Schematic diagram of bi-layer spacers.

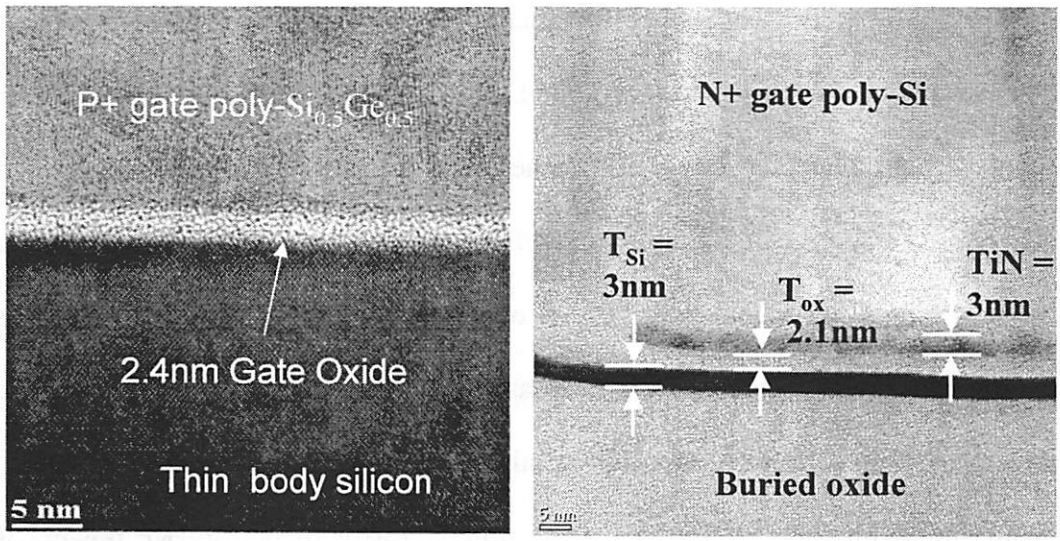
Figure 4.12 (a) XTEM photograph of bi-layer spacers and (b) cross-sectional diagram of bi-layer spacers without the tail [4.53].

After mono-spacer (HTO) or bi-spacer (HTO/Si_3N_4) formation, raised S/D is formed with a resist etch-back process or a selective Ge deposition as described in the chapter 3. A heavy dose implantation ($5x10^{15}$ cm^{-2} and *0* degree) is used for N+/P+ S/D using phosphorus and boron after S/D mask, respectively. Fig. 4.13 shows XTEM photographs of UTBFETs with poly-Si raised S/D and P+ $Si_{0.5}Ge_{0.5}$ gate as well as poly-Ge S/D and TiN gate.



(a) UTBFETs with etched-back raised poly-Si S/D

(b) UTBFETs with selective deposited raised poly-Ge S/D

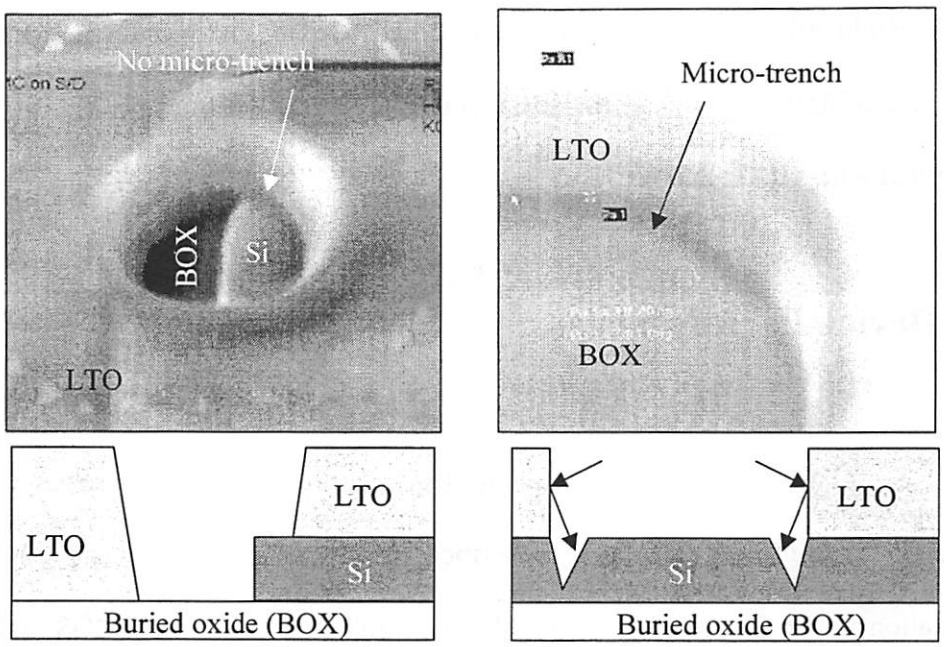


(c) XTEM close-up of (a)

(d) XTEM close-up of (b)

Figure 4.13 XTEM photographs of UTBFETs with raised S/D and different gate material. (a) raised S/D is formed by resist and poly-Si etch-back and (b) raised S/D is formed by a selective Ge deposition [4.13][4.54].

RTA at 900°C for 10sec to 1min is used to activate dopants in poly-Si S/D in N_2 ambient. RTA at 650°C for 20sec is used for Ge S/D. In the first lot, a RTA is repeatedly used to find the optimized annealing conditions of dopant activation. Thus, n-type doped poly-silicon is used for interconnection instead of aluminum. RTA at 900°C for 1min is the optimized condition for phosphorus doped poly-Si using 30nm spacers. 400nm of LPCVD LTO is deposited to avoid shorting transistors and interconnection lines. Contact holes for interconnections are opened with conventional lithography and plasma etching. CF_4 based oxide etch in Lam 5 suffered from a low selectivity and caused micro-trench problems as shown in Fig. 4.14 (a), possibly resulting in poor step coverage of aluminum at the micro-trench. This can cause junction spiking, and damage the p-n junction. One way to circumvent this is to use diluted HF for the overetch. However, this widens the contact hole diameter since HF etches isotropically. The enlarged size results in incomplete coverage of aluminum over the contact holes as shown in Fig. 4.15, creating a serious reliability issues. Thus, a new oxide etch recipe based on CHF_3/Ar is used to produce a sloped contact hole profile as seen in Fig. 4.14 (a), no micro-trenches, and relatively high selectivity of oxide to silicon (3:1) can be realized. The recipe calls for 20mTorr of pressure, 450W of top RF power, 750W of bottom RF power, 30sccm of CHF_3, and 200sccm of Ar. The oxide etch rate is 540nm/min and does not relying on oxide quality. There was no significant difference of etch rate among thermally grown oxide, HTO, and LTO.



(a) Tilted SEM view of metal contact hole without micro-trench

(b) Tilted SEM view of metal contact hole with micro-trench

Figure 4.14 Tilted SEM views of (a) sloped metal contact hole without micro-trench and (b) non-sloped metal contact hole with micro-trench.

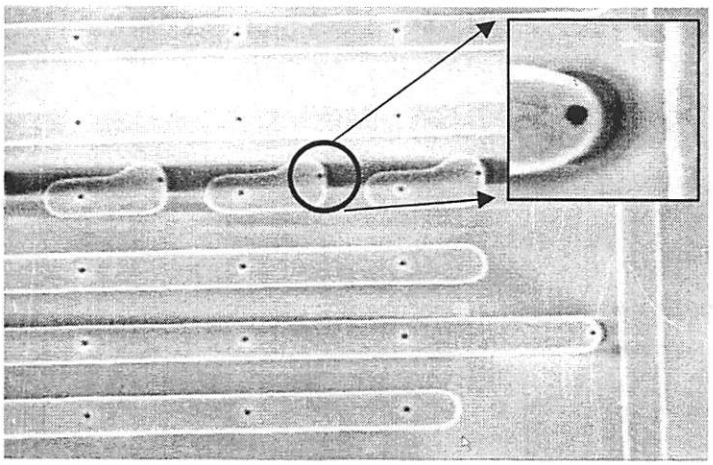


Figure 4.15 Top SEM view of Al interconnection line. The inset shows that the enlarged contact hole by HF makes Al line not to cover that contact hole completely.

Right after HF cleaning to remove native oxide, aluminum is sputtered with a thickness of 450nm. After patterning metal interconnection lines (in Fig. 4.15), sintering is applied with 400°C and N_2:H_2=9:1.

4.4 Device Performances

Simulations show that off-state leakage current strongly relies on the body thickness of UTB MOSFETs as described in the section 4.2. This prediction by the simulation is verified by experimentally measured data. Off-state PMOS leakage current significantly decreases with reduced body thickness as shown in Fig. 4.16.

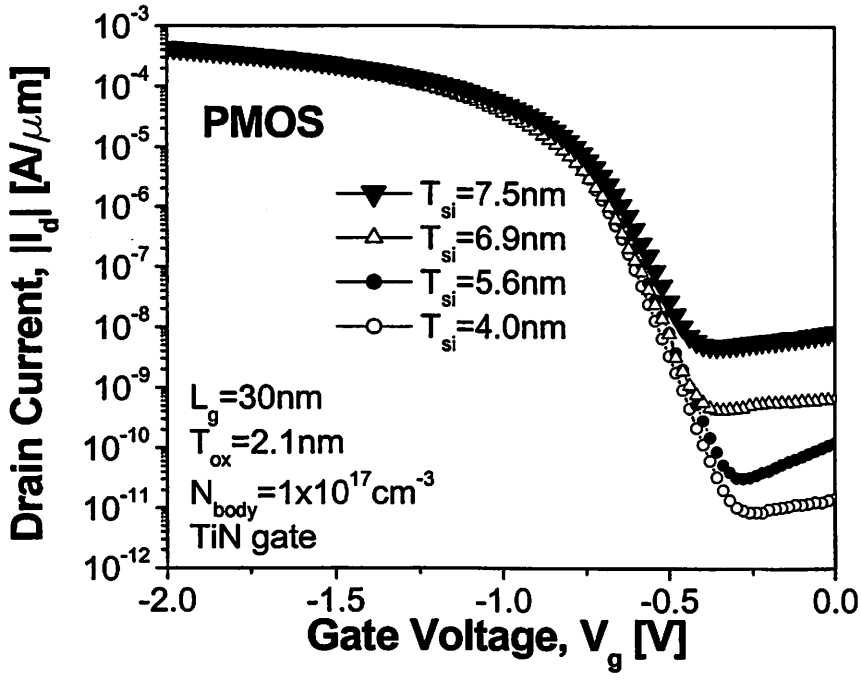


Figure 4.16 Measured off-state leakage current decreased as the body thickness reduced in UTBFETs. Threshold voltage shift is also found [4.54].

As significant threshold voltage shift is observed when the body doping concentration is higher than $1x10^{18}cm^{-3}$ in bulk devices [4.32][4.33][4.55]. However, this threshold voltage shift is observed even with a low body doping [4.56]. An ultra-thin body surrounded by a front gate oxide and back buried oxide produces a two dimensional quantum well in the body, resulting in a change in sub-band structures for low or moderate body doping.

Measured I-V characteristics of NMOS and PMOS devices are shown in Fig. 4.16. NMOS and PMOS transistors are fabricated on separate wafers. The NMOS device in Fig. 4.16 features a p-type body (initially doped $N_{body}=1x10^{15}cm^{-3}$), 20nm body thickness ($T_{Si}$), 2.4nm gate oxide ($T_{ox}$), P+ heavily doped $Si_{0.5}Ge_{0.5}$, and 80nm gate length (L_g). The on-state drain current is *750uA/um* at $V_g-V_t=V_d=1V$ and off-state leakage current below *1pA/um*. The threshold voltage is 1.4V, which is higher than the expected. This higher threshold voltage is caused by the boron penetration from the P+ $Si_{0.5}Ge_{0.5}$ gate. It is believed to have happened during HTO deposition because of a 15 hour temperature stabilization period caused by equipment trouble. For a body doping that is higher than $1x10^{19}cm^{-3}$ and caused by the boron penetration, the threshold voltage shift is believed to be ~0.6-0.8V, taking into account inversion charge confinement effects [4.55]. For PMOS, device features include n-type doped body ($N_{body}=1x10^{17}cm^{-3}$), 4nm body thickness ($T_{Si}$), 2.1nm gate oxide ($T_{ox}$), and TiN gate, and 30nm gate length (L_g). The on-state drain current is *350uA/um* at $V_g-V_t=V_d=1V$ and off-state leakage current below *10pA/um*.

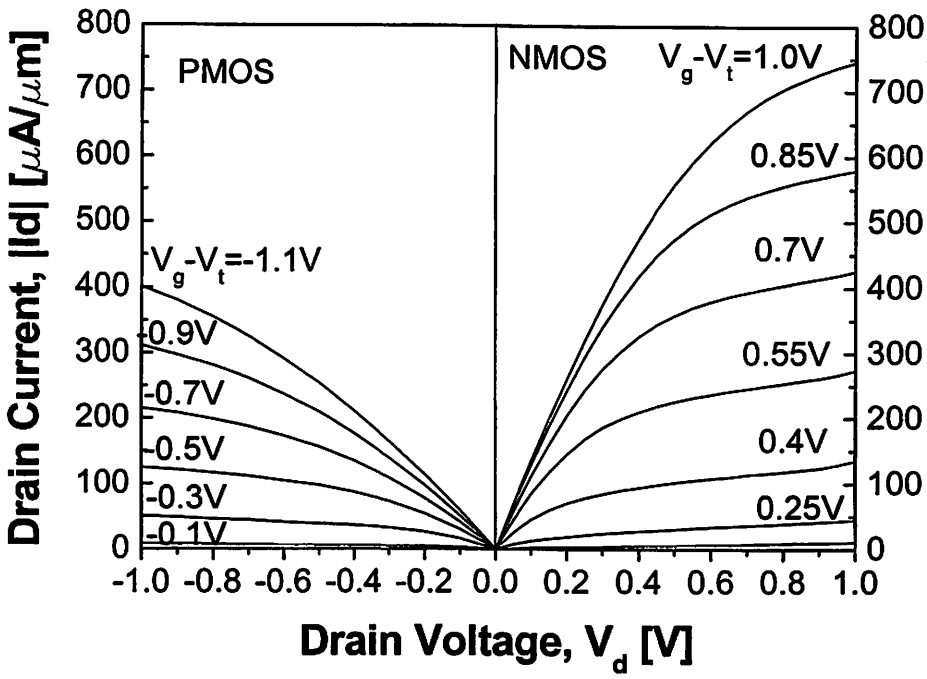


Figure 4.17 Measured I_d-V_d characteristics for a separate NMOS and PMOS.

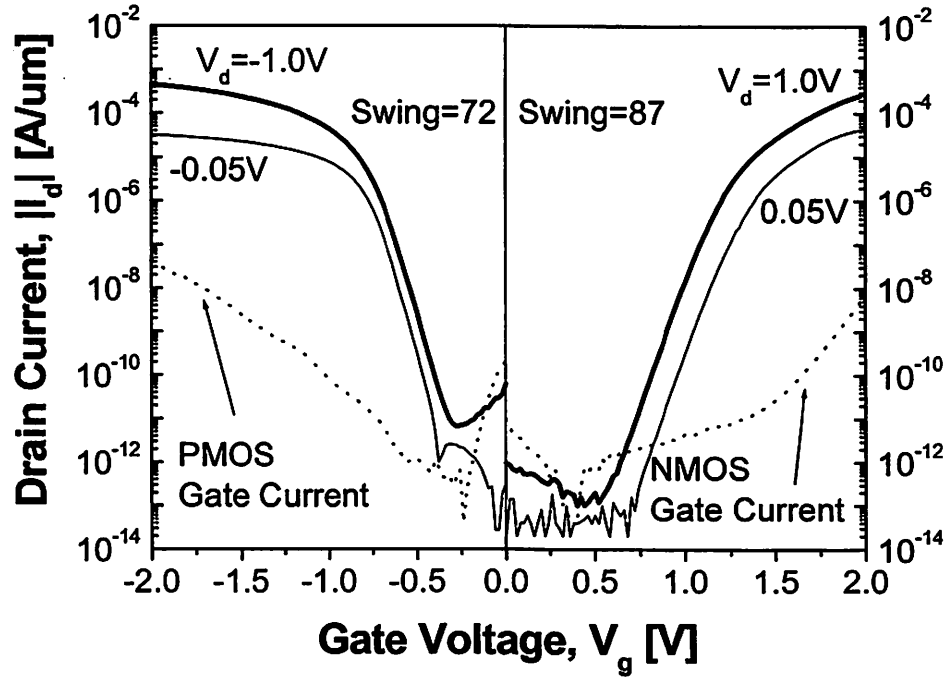


Figure 4.18 Measured I_d-V_g characteristics for a separate NMOS and PMOS.

Different gate materials such as P+ $Si_{0.5}Ge_{0.5}$, N+ poly-Si, and TiN are applied to UTBFETs in order to investigate the threshold voltage dependence on gate work function. Figure 4.19 and 4.20 shows NMOS and PMOS threshold voltage. P+ heavily doped $Si_{0.5}Ge_{0.5}$ without boron penetration is expected to obtain a reasonable threshold voltage. A TiN gate with initial low doped-body is predicted to achieve lower PMOS threshold voltage. For NMOS, a reasonable threshold voltage is not obtained yet. Threshold voltage adjustment in UTB CMOS is still unresolved.

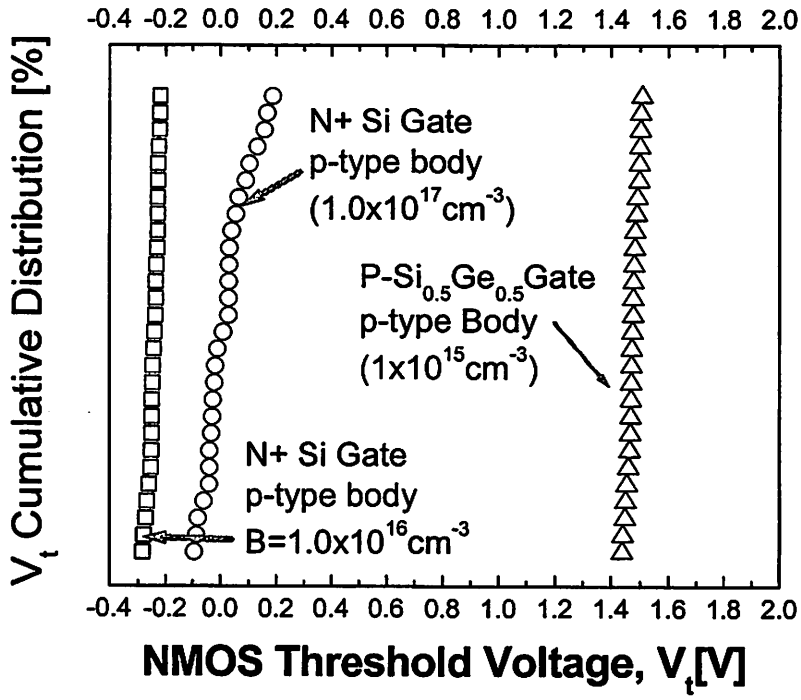


Figure 4.19 NMOS threshold voltage for N+ Si gate and P+ $Si_{0.5}Ge_{0.5}$ gate with different body doping, $1x10^{16}cm^{-3}$, $1x10^{17}cm^{-3}$.

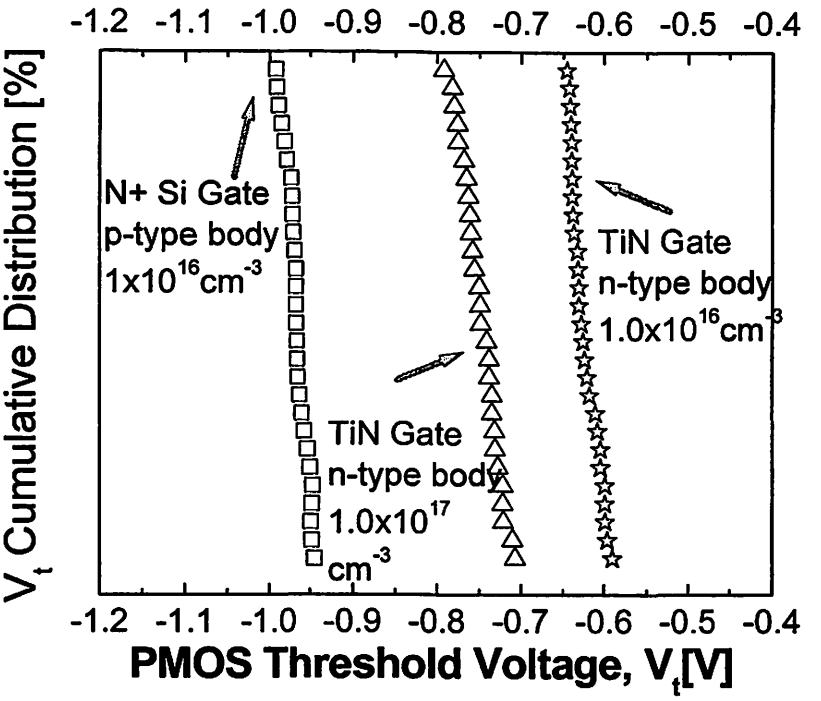
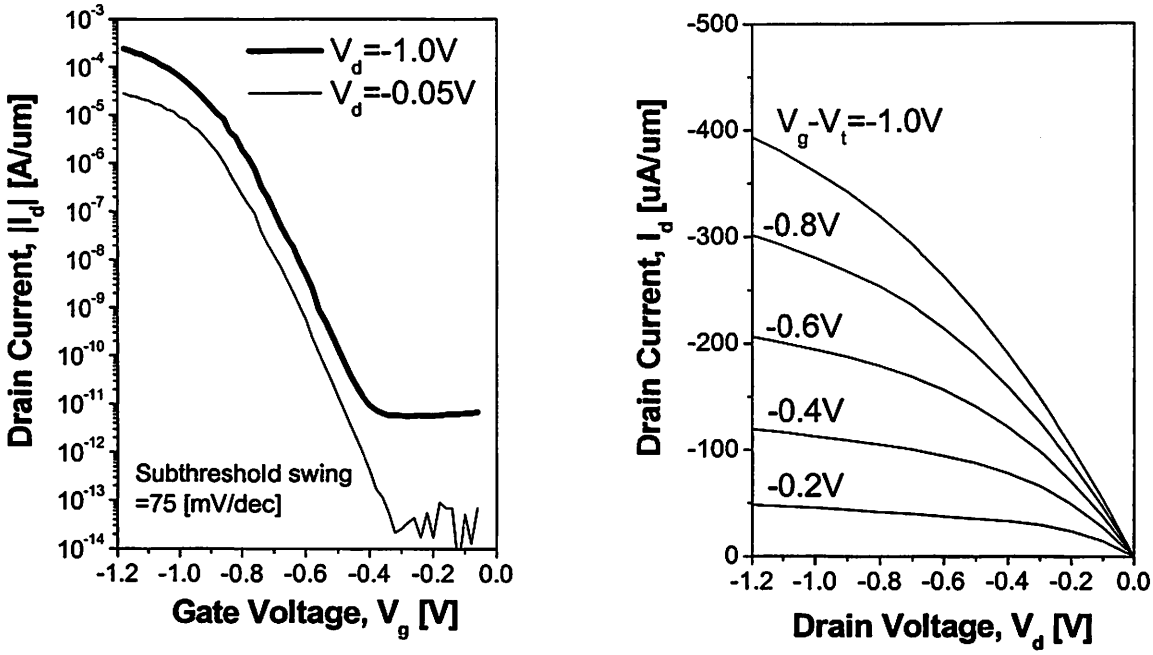
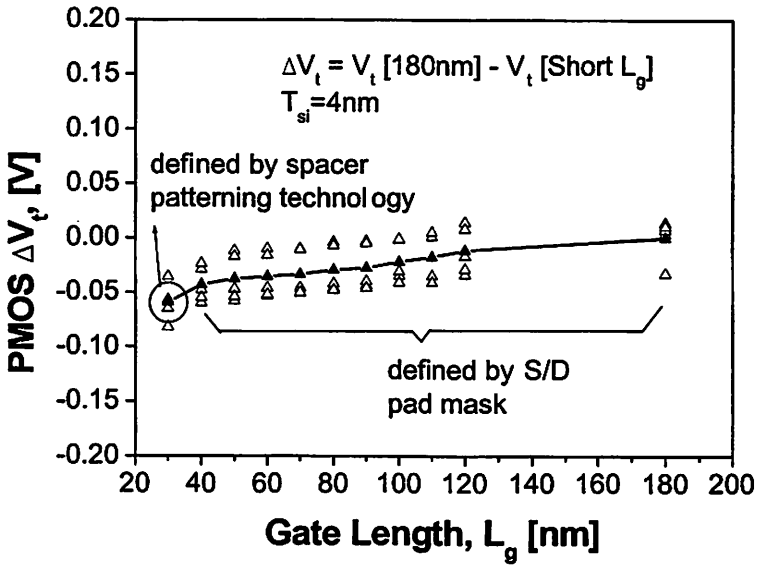


Figure 4.20 PMOS threshold voltage for N+ Si gate and TiN gate with different body doping, $1x10^{16}cm^{-3}$, $1x10^{17}cm^{-3}$.

The gate patterns in this work are defined with conventional i-line lithography except for spacer gate and were reduced down to sub-30nm by utilizing ashing-trimming. Some gates are also defined with spacer lithography technology as described in the chapter 2. Its minimum channel length is 47nm in the top view as shown in Fig. 2.21. In the bottom of gate, the gate length is believed to be sub-30nm due to a lateral undercut as shown in Fig. 4.13 (a). N-type doped poly-Si is used for the gate and the thin body is doped with phosphorus ($1x10^{16}cm^{-3}$). Typical I-V and threshold voltage roll-off characteristic are shown in Fig. 4.21.



(a) (b)



(c)

Figure 4.21 (a) measured I_d-V_g, (b) measured I_d-V_d, and (c) threshold voltage characteristics of PMOS defined by spacer lithography.

Subthreshold characteristics and short-channel effects are investigated in NMOS (N+ poly-Si gate, T_{ox}=*2.1nm*) and PMOS (TiN gate, T_{ox}=*2.1nm*) shown in Fig. 4.22 to Fig. 4.25. V_t roll-off is more severe for thicker and lower doped-bodies. For a 2.1nm gate oxide thickness, the ratio of gate length to body thickness (L_g/T_{Si}) is larger than 4. This ratio can be reduced for thinner gate oxide. Subthreshold swing, DIBL, and off-state leakage current satisfy the criterion of 100mV/dec, 100mV/V, and 1nA/um, respectively as shown in Fig. 4.23 to Fig. 4.25. These data are collected from a range of grate lengths (30nm to 180nm) and body thicknesses (4nm to 12nm).

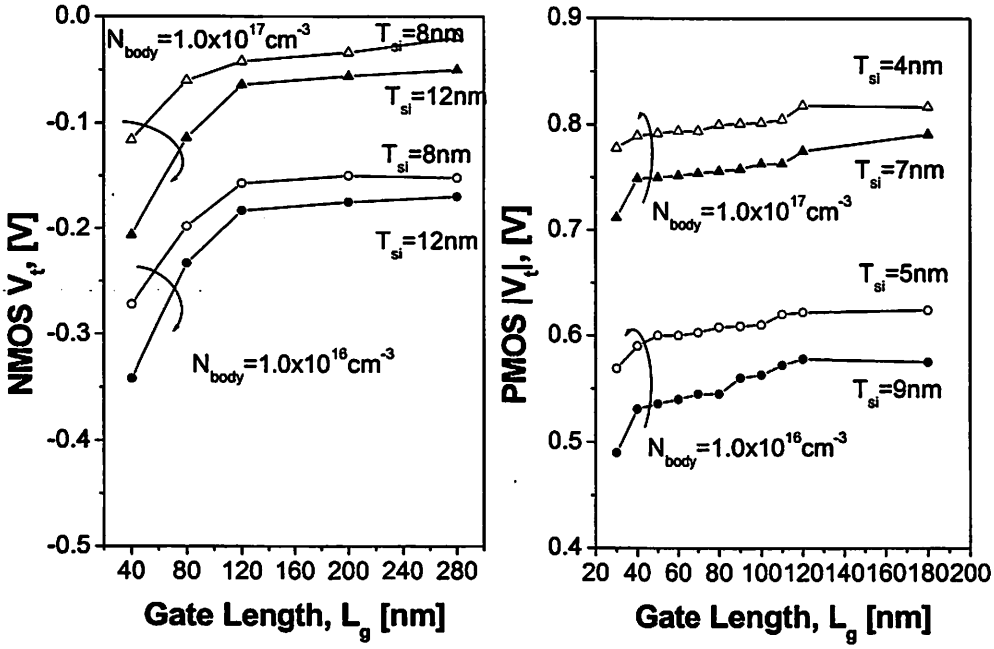


Figure 4.22 Threshold voltage roll-off characteristics of NMOS with N+ Si gate and PMOS with TiN gate.

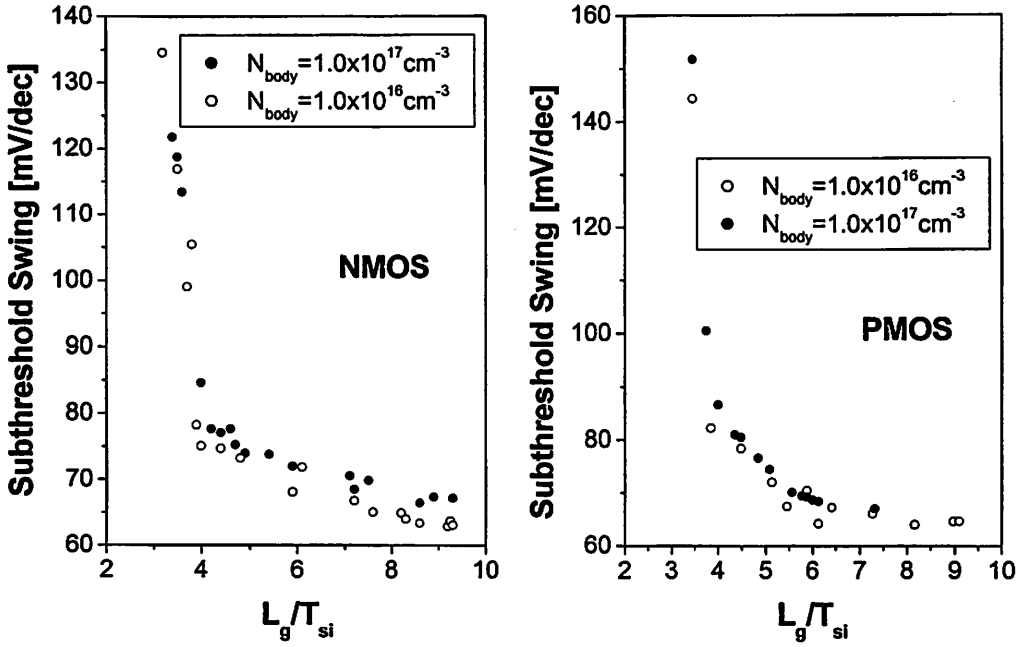


Figure 4.23 Subthreshold swing versus L_g/T_{Si}.

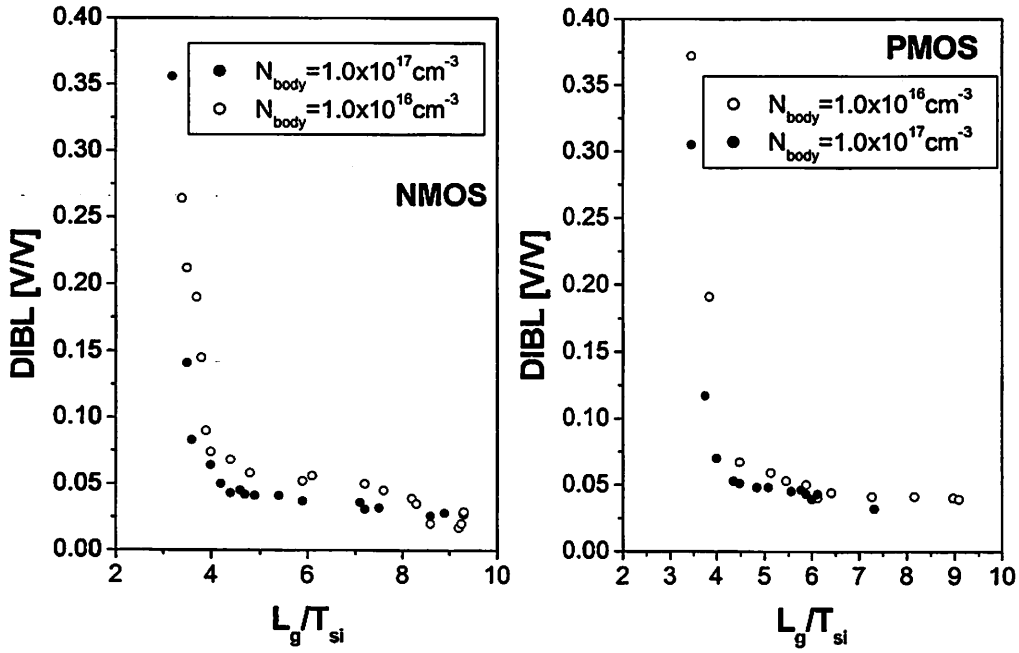


Figure 4.24 Drain induced barrier lowering (DIBL) versus L_g/T_{Si}.

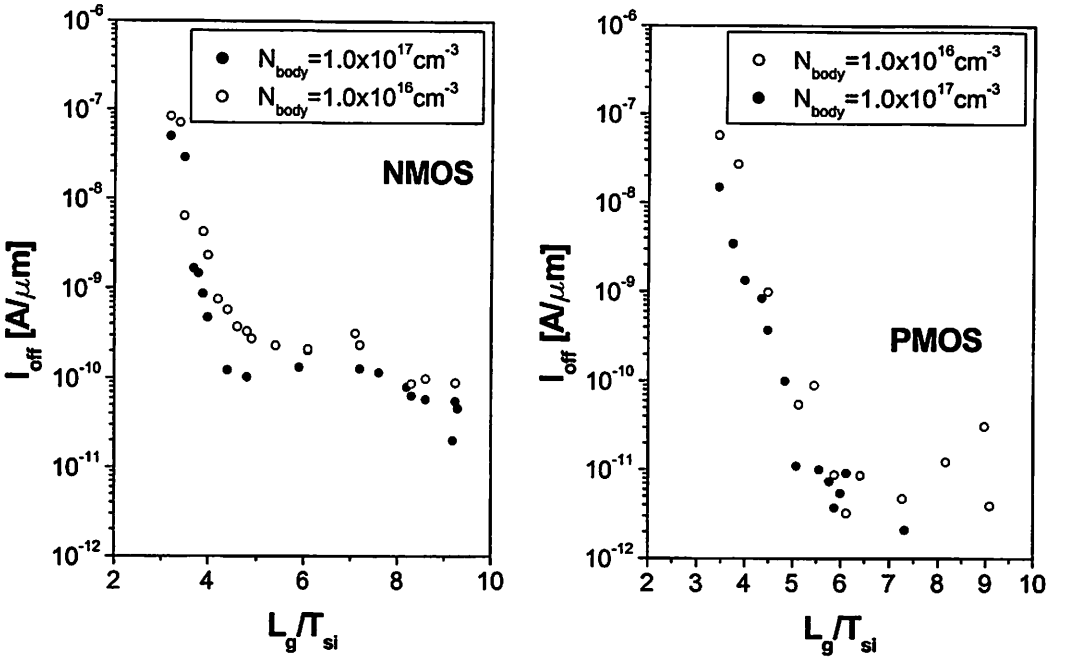
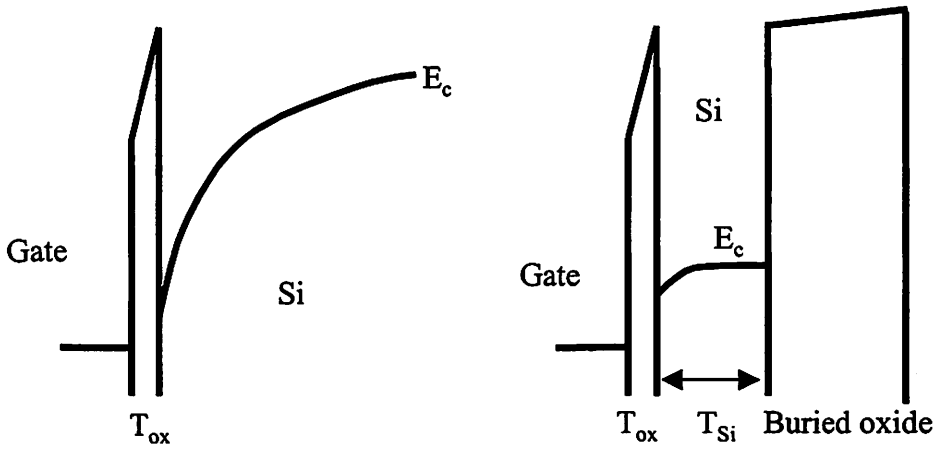


Figure 4.25 Off-state leakage current versus L_g/T_{Si}.

One benefit of UTB MOSFETs is that the gate oxide tunneling current decreases for thinner bodies as shown in Fig. 4.27. Previous experimental results have shown that gate current in SOI devices is affected by the electric field distribution in the body [4.57]. As the body thickness is decreased, the vertical electric field is reduced. In particular, the electric field near the bottom of the inversion layer is dramatically reduced [4.58] as shown in Fig. 4.26. Gate leakage is suppressed due to the reduced vertical electric field.



(a) Bulk MOSFET (b) UTB MOSFET

Figure 4.26 Gate oxide (T_{ox}=*2.1nm*) leakage current for different body thickness.

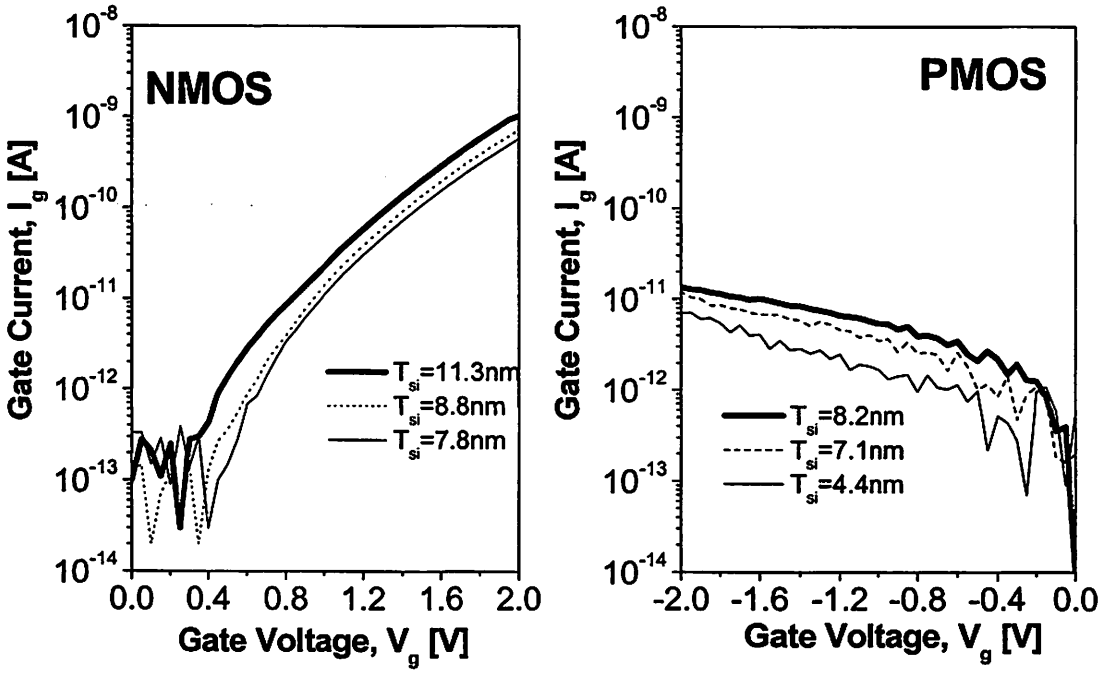


Figure 4.27 Measure gate current (T_{ox}=*2.1nm*) for different body thicknesses.

Electron mobility in UTBFETs is extracted from the capacitance measurement of a large-sized transistor as shown in Fig. 4.28. It is close to the universal mobility.

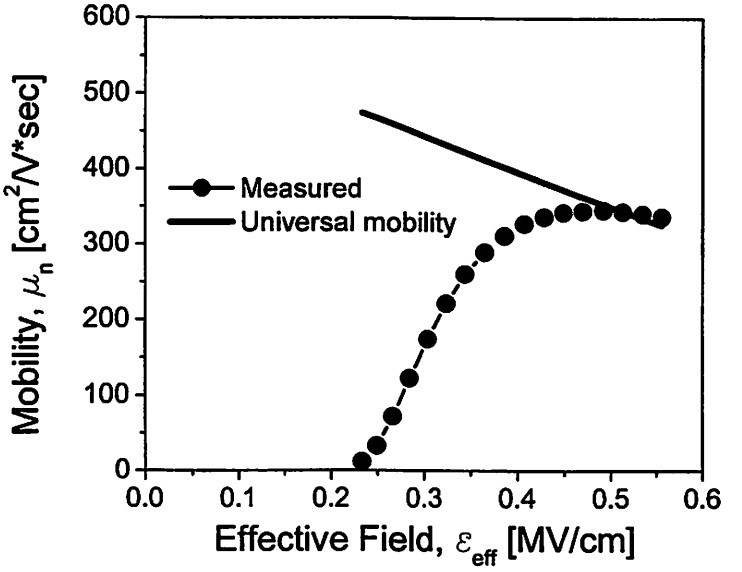


Figure 4.28 Electron mobility versus effective electric field.

4.5 Conclusion

An ultra-thin body field effect transistor (UTBFETs) with raised S/D is proposed. Working devices with sub-30nm gate length and 4nm body thickness are successfully demonstrated. Resist etch-back of poly-silicon S/D and selectively deposited germanium S/D by LPCVD, are two approaches for the raised S/D. Processing details ranging from body thinning process to metallization are discussed and solutions are presented to avoid technical problems. The great benefits of UTBFETs are in the simple process and low cost of fabrication. Simulations show that UTBFETs can be scaled down to 18nm with

5nm body thickness and can be used with thicker gate oxides (1.5nm) for sub-20nm gate lengths. UTBFETs show excellent short-channel effects, low off-state current, and high on-state drive current. The UTB device structure has many features in common with today's bulk MOSFET, which makes it easier for industry to introduce into manufacturing. Companies such as Intel and TSMC are starting to consider UTBFETs. They will be one of the promising structures when looking beyond the end of ITRS roadmap.

4.6 References

[4.1] R.H. Dennard, F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassous and A. R. LeBlanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE Journal of Solid-State Circuits*, vol.9, p.256-268, 1974.

[4.2] C. Hu, "MOSFET scaling in the next decade and beyond," *Semiconductor International*, p.105-114, June 1994.

[4.3] B. Davari, "CMOS technology scaling: 0.1 μm and beyond," *IEDM Technical Digest*, p.555-558, 1996.

[4.4] C. Hu, "SOI and Nanoscale MOSFETs", *59th Device Research Conference*, p.3-4, 2001.

[4.5] *International Technology Roadmap for Semiconductors*, Semiconductor Industry Association, 2000. http://public.itrs.net/Home.htm

[4.6] B. Yu, H. Wang, Q. Xiang, J. X. An, J. Jeon, and M.-R. Lin, "Scaling Towards 35nm Gate Length CMOS", *Symposium on VLSI Technology*, p.9-10, 2001.

[4.7] R. Chau, J. Kavalieros, B. Roberds, R. Schenker, D. Lionberger, D. Barlage, B. Doyle, R. Arghavani, A. Murthy, G. Dewey, "30nm Physical Gate Length CMOS Transistors with 1.0 ps n-MOS and 1.7 ps p-MOS Gate Delays", *IEEE IEDM Technical Digest*, p.45-48, 2001.

[4.8] Y. Taur, E.J. Nowak, "CMOS devices below 0.1um: how high will performance go?", *IEEE IEDM Technical Digest*, p.215-218, 1997.

[4.9] L. Kang, K. Onishi, Y. Jeon, B. H. Lee, C. Kang, W.-J. Qi, R. Nieh, S. Gopalan, R. Choi, J.C. Lee, "MOSFET devices with polysilicon on single-layer HfO_2 high-K dielectrics", *IEEE IEDM Technical Digest*, p.35-38,2000.

[4.10] A. Chin, Y.H. Wu, S.B. Chen, C.C. Liao, W.J. Chen, "High quality La_2O_3 and Al_2O_3 gate dielectrics with equivalent oxide thickness 5-10 A", *Symposium on VLSI Technology*, p.16-17, 2000.

[4.11] C.H. Lee, H.F. Luan, W.P. Bai, S.J. Lee, T.S. Jeon, Y. Senzaki, D. Roberts, D.L. Kwong, "MOS characteristics of ultra thin rapid thermal CVD ZrO_2 and Zr silicate gate dielectrics", *IEEE IEDM Technical Digest*, p.27-30, 2000.

[4.12] B. Yu, Y.-J. Tung, S. Tang, E. Hui, T.-J. King, and C. Hu, "Ultra-thin-body Silicon-On-Insulator MOSFET's for terabit-scale integration", *International Semiconductor Device Research Symposium*, p. 623-626, 1997.

[4.13] Y.-K. Choi, K. Asano, N. Lindert, V. Subramanian, T.-J. King, J. Bokor, C. Hu, "Ultra-thin-body SOI MOSFET for Deep-sub-tenth Micron Era," *IEEE Electron Device Letters*, vol. 21, p.254-255, 2000.

[4.14] MEDICI version 4.1, Avant! Corporation, 1998.

[4.15] K. Goto, T. Yamamoto, T. Kubo, M. Kase, Y. Wang, T. Lin, S. Talwar, T. Sugii, "Ultra-low contact resistance for deca-nm MOSFETs by laser annealing", *IEEE IEDM Technical Digest*, p.931-3, 1999

[4.16] S. Talwar, G. Verma, K.H. Weiner, "Ultra-shallow, abrupt, and highly-activated junctions by low-energy ion implantation and laser annealing", *Proceedings of International Conference on Ion Implantation Technology*, vol.2, p.1171-4, 1998.

[4.17] H.I. Liu, J.A. Burns, C.L. Keast, P.W. Wyatt, "Thin silicide development for fully-depleted SOI CMOS technology", *IEEE Transactions on Electron Devices*, vol.45, p.1099-104, 1998.

[4.18] L.T. Su, M.J. Sherony, H. Hu, J.E. Chung, D.A. Antoniadis, "Optimization of series resistance in sub-0.2um SOI MOSFET's", *IEEE Electron Device Letters*, vol.15, p.363-5, 1994.

[4.19] M. Chan, F. Assaderaghi, S.A. Parke, C. Hu, P.K. Ko, "Recessed-channel structure for fabricating ultrathin SOI MOSFET with low series resistance", *IEEE Electron Device Letters*, vol.15, p.22-4, 1994.

[4.20] K. Miyano, I. Mizushima, A. Hokazono, K. Ohuchi, Y. Tsunashima, "Low thermal budget elevated source/drain technology utilizing novel solid phase epitaxy and selective vapor phase etching", *IEEE IEDM Technical Digest*, p. 433-6, 2000.

[4.21] J. Lyu, B.-G. Park, K. Chun, J. D. Lee, "A novel 0.1 um MOSFET structure with inverted sidewall and recessed channel", *IEEE Electron Device Letters*, vol.17, p.157-9, 1996.

[4.22] W.-H. Lee, Y.-J. Park, J. D. Lee, "A new 0.25-um recessed-channel MOSFET with selectively halo-doped channel and deep graded source/drain", *IEEE Electron Device Letters*, vol.14, p.578-80, 1993.

[4.23] S. Nagao, K. Higashitani, Y. Akasaka, H. Nakata, "Applications of selective silicon epitaxial growth for CMOS technology", *IEEE Transactions on Electron Devices*, vol.33, p.1738-44, 1986.

[4.24] J. Borland, R. Wise, Y. Oka, M. Gangani, S. Fong, Y. Matsumoto, "Silicon epitaxial growth for advanced device structures", *Solid State Technology*, vol.31, p.111-19, 1988.

[4.25] A. Srivastava, J. Sun, K. Bellur, R.F. Bartholomew, P. O'Neil, S.M. Celik, C.M. Osburn, N.A. Masnari, M.C. Ozturk, R. Westhoff, B. Fowler, "A 0.18 um CMOS technology for elevated source/drain MOSFETs using selective silicon epitaxy", *Proceedings of the Sixth International Symposium on Ultralarge Scale Integration Science and Technology (ULSI Science and Technology)*, p.571-85, 1997.

[4.26] R.W. Keyes, "Effect of randomness in the distribution of impurity ions on FET thresholds in integrated electronics", *IEEE Journal of Solid-State Circuits*, vol.SC-10, p.245-7, 1975.

[4.27] H.-S. Wong, Y. Taur, "Three-dimensional "atomistic" simulation of discrete random dopant distribution effects in sub-0.1 um MOSFET's", *IEEE IEDM Technical Digest*, p.705-8, 1993.

[4.28] D. Hisamoto, W.-C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T.-J. King, J. Bokor, C. Hu, "FinFET-A Self-Aligned Double-Gate

MOSFET Scalable to 20nm," *IEEE Transactions on Electron Devices,* vol.47, p.2320-5, 2000.

[4.29] X. Huang, W.-C. Lee, C. Kuo, D. Hisamoto, L. Cang, J. Kedzierski, E. Anderson, H. Takeuchi, Y.-K. Choi, K. Asano., V. Subramanian, T.-J. King, J. Bokor, C. Hu, "Sub-50 nm P-channel FinFET", *IEEE Transactions on Electron Devices*, vol.48, p.880-6, 2001.

[4.30] C. Hu, "Silicon-On-Insulator for High Speed Ultra Large Scale Integration," *Japan Journal of Applied Physics*, vol. 33, p. 365-369, 1994.

[4.31] L. T. Su, J. B. Jacobs, J. E. Chung, and D. A. Antoniadis, "Short-Channel Effects in Deep-Submicrometer SOI MOSFETs," *IEEE International SOI conference Proceedings*, p.112-113, 1993.

[4.32] M. J. van Dort, P. H. Woerlee, A. J. Walker, C. A. H. Juffermans, H. Lifka, "Influence of High Substrate Doping Levels on the Threshold Voltage and the Mobility of Deep-Submicrometer MOSFETs," *IEEE Transactions on Electron Deivces*, vol.39, p.932-938, 1992.

[4.33] C.-Y. Hu, S. Banerjee, K. Sadra, B. G. Streetman, R. Sivan, "Quantization Effects in Inversion Layers of pMOSFETs on Si (100) Subtrate," *IEEE Electron Device Letters*, vol.17, p. 276-278, 1996.

[4.34] S.-C. Takagi, A. Toriumi, M. Iwase, H. Tango, "On the Universality of Inversion Layer Mobility in Si MOSFET's: Part I - Effects of Substrate Impurity Concentration", *IEEE Transactions on Electron Deivces*, vol.41, p.2357-62, 1994.

[4.35] K. Takeuchi, T. Tatsumi, A. Furukawa, "Channel engineering for the reduction of random-dopant-placement-induced threshold voltage fluctuation", *IEEE IEDM Technical Digest*, p. 841-4, 1997.

[4.36] H. Wakabayashi, Y. Saito, K. Takeuchi, T. Mogami, T. Kunio, "A novel W/TiNx metal gate CMOS technology using nitrogen-concentration-controlled TiNx film", *IEEE IEDM Technical Digest*, p.253-6, 1999.

[4.37] Q. Lu, R. Lin, P. Ranade, T.-J. King, C. Hu, "Metal gate work function adjustment for future CMOS Technology", *Symposium on VLSI Technology*, p.45-46, 2001.

[4.38] I. Polishchuk, P. Ranade, T.-J. King, C. Hu, "Dual Work Function Metal Gate CMOS Technology Using Metal Interdiffusion", *IEEE Electron Device Letters*, vol.22, p. 444-6, 2001.

[4.39] T.-J. King, J. P. McVittie, and K. C. Saraswat, "Electrical Properties of Heavily Doped Polycrystalline Silicon-Germanium Films," *IEEE Transactions on Electron Devices*, vol. 41, p.228-231, 1994.

[4.40] P.-E. Hellberg, S.-L. Zhang, C.S. Petersson, "Work Function of Boron-Doped Polycrystalline Si_xGe_{1-x} Film," *IEEE Trans. on Electron Devices*, vol. 18, p.456-8, 1997.

[4.41] C. Salm, D.T. van Veen, D.J. Gravesteijn, J Holleman, P.H. Worlee, "Diffusion and Electrical Properties of Boron and Arsenic Doped Poly-Si and Poly-Ge_xSi_{1-x} (x~0.3) as Gate Material for Sub-0.25um Complementary Metal Oxide Semiconductor applications," *Journal of Electrochemical Society*, vol.144, p.3665-73, 1997.

[4.42] E. Gondro, P. Klein, F. Schuler, "An analytical source-and-drain series resistance model of quarter micron MOSFETs and its influence on circuit simulation", *Proceedings of the 1999 IEEE International Symposium on Circuits and Systems*, p.206-9, 1999.

[4.43] P. Keys, H.-J. Gossmann, K.K. Ng, C.S. Rafferty, "Series resistance limits for 0.05 um MOSFETs", *Superlattices and Microstructures*, vol.27, p.125-36, 2000.

[4.44] S. Zimin; L. Litian; L. Zhijian, "Optimization of MOSFETs with polysilicon-elevated source/drain", *5th International Conference on Solid-State and Integrated Circuit Technology*, p.188-90, 1998.

[4.45] http://www.soitec.com/index..htm

[4.46] Y.-K. Choi, D. Ha, T.-J. King, C. Hu, "Ultra-Thin Body PMOSFETs with selectively deposited Ge Source/Drain", *Symposium on VLSI Technology*, p.19-20, 2001.

[4.47] J.M. Pimbley, M. Ghezzo, H.G. Parks, D.M. Brown, *Advanced CMOS Process Technology*, vol.19, Academic Press, Inc., San Diego, 1989.

[4.48] S. Nag, A. Chatterjee, K. Taylor, I. Ali, S. O'Brien, S. Aur, J.D. Luttmer, I.C. Chen, "Comparative evaluation of gap-fill dielectrics in shallow trench isolation for sub-0.25 um technologies", *IEEE IEDM Technical Digest*, p.841-5, 1996.

[4.49] P.M. Fahey, S.R. Mader, S.R. Stiffler, R.L. Mohler, J.D. Mis, J.A. Slinkman, "Stress-induced dislocations in silicon integrated circuits", *IBM Journal of Research and Development*, vol.36, p.158-82, 1992.

[4.50] K. Ishimaru, F. Matsuoka, M. Takahashi, M. Nishigohri, Y. Okayama, Y. Unno,M. Yabuki, K. Umezawa, N. Tsuchiya, O. Fujii, M. Kinugawa, "Mechanical stress

induced MOSFET punch-through and process optimization for deep submicron TEOS-O_3 filled STI device", *Symposium on VLSI Technology*, p.123-4, 1997.

[4.51] B.H. Roh, Y.H. Cho, Y.G. Shin, G.H. Hong; S.D. Gwun, K.Y. Lee, H.G. Kang, K.N. Kim, J.W. Park, "Easily manufacturable shallow trench isolation for gigabit dynamic random access memory", *Japanese Journal of Applied Physics (Part 1)*, vol.35, p.4618-23, 1996.

[4.52] C.P. Chang, C.S. Pai, F.H. Baumann, C.T. Liu, C.S. Rafferty, M.R. Pinto, E.J. Lloyd, M. Bude, F.P. Klemens, J.F. Miner, K.P. Cheung, J.I. Colonell, W.Y.C. Lai, H. Vaidya, S.J. Hillenius, R.C. Liu, J.T. Clemens, "A highly manufacturable corner rounding solution for 0.18um shallow trench isolation", *IEEE IEDM Technical Digest*, p.661-4, 1997.

[4.53] N. Lindert, Doctoral Thesis, "Process Development and Device Design for Continued MOSFET Scaling", Department of Electrical Engineering and Computer Science, *University of California at Berkeley*, 2001.

[4.54] Y.-K. Choi, D. Ha, T.-J. King, C. Hu, "Nanoscale ultrathin body PMOSFETs with raised selective germanium source/drain", *IEEE Electron Device Letters*, vol.22, p.447-8, 2001.

[4.55] M. J. van Dort, P. H. Woerlee, A. J. Walker, "A simple Model for Quantisation Effects in heavily-doped silicon MOSFETs at Inversion Conditions", *Solid-State Electronics*, vol.37, p.411-4, 1994.

[4.56] Y.-K. Choi, D. Ha, T.-J. King, C. Hu, "Threshold voltage Shift by Quantum Confinement in Ultra-thin Body Device", *59th Device Research Conference*, p.85-6, 2001.

[4.57] L.T. Su, F. Hao, J.E. Chung, D.A. Antoniadis, "Hot-carrier effects in fully-depleted SOI nMOSFETs", *IEEE IEDM Technical digest*, p.349-52, 1992.

[4.58] L. Chang, K.J. Yang, Y.-C. Yeo, Y.-K. Choi, T.-J. King, C. Hu, "Reduction of Direct-Tunneling Gate Leakage Current in Double-Gate and Ultra-Thin Body MOSFETs", *IEEE IEDM Technical digest*, p.99-102, 2001.

4.7 Appendices

4.7.1 Process Flows of UTBFETs with Etched-Back Poly-Si S/D

Step	Process Name	Process Specification	Equipment	Comment
1.0	Wafers	4" p-type SOI		T_{Si}=100nm
2.0	Body Thinning			
2.01	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
2.02	1st body thinning	Wet, SWETOXB, 850°C, 3hour, 180nm	Tylan 2	
2.03	Measuement	Body thickness measurement	Nanoduv	
2.04	Implant mask	Resist coating : coat = prog#01/bake = prog#01 Exposure Development : bake = prog#01/develop = prog#01 Hard bake at 120°C, 30min	Svgcoat gcaws Svgdev Vwr	for implantation
2.05	Channel implantation	Split	Implanter	foundry company
2.06	Resist Strip	O_2 ashing, 300W, 5min	Technics-c	
2.07	Post cleaning	Piranha, 120°C, 20min	Sink 8	
2.08	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
2.09	2nd body thinning	Dry, SGATEOX, 900°C, 50min, 18nm Multiple oxidations up to the target of T_{Si}	Tylan 5	
3.0	Active Formation			
3.01	1st active mask	Resist coating : coat = prog#01/bake =prog #01 Exposure (focus and expose test) Development : bake = prog#01/develop = prog#01	Svgcoat Gcaws Svgdev	
3.02	Thin body etch	BT:TP=70W,BP=10W,13mT,CF_4=100, 10sec ME:TP=70W,BP=30W,15mT,HBr:He:O_2=50:50:1,20sec	Lam5 Lam5	time etch
3.03	Resist Strip	O_2 ashing, 300W, 5min	Technics-c	
3.04	Post cleaning	Piranha, 120°C, 20min	Sink 8	
3.05	Measurement	Remaind BOX measurement	Nanoduv	Rox=400nm
4.0	Gate Formation			
4.01	Pre cleaning	Piranha, 120°C, 20min (25:1) HF, 30sec	Sink 6 Sink 6	
4.02	Gate oxidation	Dry, THIN_ANN, 750°C, O_2, 12min/900°C, N_2, 30min	Tylan 6	
4.03	In-situ P+ $Si_{0.5}Ge_{0.5}$ Recipe: SIGEVAR.019	Nucleation : 550°C, 300mT, SiH_4=200,30sec Deposition : 450°C, 300mT, SiH_4=124,GeH_4=40 B_2H_6=40 (enter 80), 40min	Tystar 19	
4.04	LTO deposition	LTO deposition, 11SULTOA, 8min, 150nm	Tystar 11	
4.05	Gate mask	Resist coating : coat = prog#01/bake =prog #01 Exposure (focus and expose test) Development : bake = prog#01/develop = prog#01	Svgcoat Gcaws Svgdev	
4.06	Ashing	O_2 ashing, 30W, 7min	Technics-c	
4.07	Measurement	CD measuremt and additive ashing	Leo/Tech.-c	
4.08	LTO Etch	1st ME:TP=200W,BP=40W,13mT,CF_4=100, 70sec 2nd ME:TP=200W,BP=40W,20mT CHF_3:Ar=90:200, 15sec	Lam5	
4.09	Resist strip	(100:1) HF, 20sec O_2 ashing, 300W, 5min (100:1) HF, 20sec	Sink 7 Technics-c Sink 7	Polymer removal
4.10	Post cleaning	Piranha, 120°C, 20min	Sink 8	
4.11	Measurement	CD measuremt and Inspection	Leo	

Step	Process Name	Process Specification	Equipment	Comment
4.12	Trimming	LTO Trimming : (100:1) HF, 4min	Sink 7	
4.13	Measurement	CD measuremt and additive trimming	Leo/Tech.-c	
4.14	Poly-$Si_{0.5}Ge_{0.5}$ etch	BT : TP=200W,BP=40W,20mT CHF_3:Ar=90:200, 10sec ME : TP=300W,BP=150W,15mT Cl_2:HBr=50:150, EDP-2sec OE : TP=250W,BP=120W,15mT HBr:O_2=200:5, 30sec	Lam5	
4.15	Post cleaning	(100:1) HF, 10sec Piranha, 120°C, 20min	Sink 7 Sink 8	
4.16	Measurement	CD Measurement and inspection	Leo	
5.0	Spacer Foramtion (Bi-layer)			
5.01	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
5.02	HTO deposition	9VHTOA, N_2O=90,DCS=18,300mT,800°C,10nm,12min	Tystar 9	
5.03	Measurement	HTO thickness measurement	Nanoduv	
5.04	Pre cleaning	Piranha, 120oC, 20min	Sink 6	
5.05	Nitride deposition	9SNITA,5min, 20nm	Tystar 9	
5.06	Measurement	Nitride thickness measurement	Nanoduv	
5.07	Spacer nitride etch	BT:TP=70W,BP=10W,13mT,CF_4=100, 10sec ME:TP=150W,BP=75W,15mT,Cl_2:HBr=50:150,80sec	Lam5	time etch time etch
5.08	Post cleaning	Piranha, 120°C, 20min	sink 8	
6.0	Etch-Back Raised S/D Formation			
6.01	Pre cleaning	Piranha, 120°C, 20min (25:1) HF, 50sec	Sink 6 Sink 6	
6.02	N+ poly-Si deposition	11SDPLYI, 2hour, 200nm	tylan11	
6.03	Measurement	Poly-Si thickness measurement	Nanoduv	
6.04	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
6.05	Cap HTO deposition	9VHTOA, N_2O=90,DCS=18,300mT,800°C,10nm,12min	Tystar 9	
6.06	Thin resist coating	Coating, 3000rpm, 45sec, 300nm	Spinner1	
6.07	Measuement	Resist thickness measurement	Nanoduv	
6.08	Etch-Back	Resist E/B:TP=250W,BP=120W,35mT,HBr:O_2=200:5 SiO_2 E/B:TP=200W,BP=40W,20mT,CHF_3:Ar=90:200 Poly-Si E/B:TP=250W,BP=120W,35mT,HBr:O_2=200:5	Lam5	
6.09	Resist strip	O_2 ashing, 300W, 5min	Technics-c	
6.10	Post cleaning	Piranha, 120°C, 20min	sink 8	
6.11	2nd active mask	Resist coating : coat = prog#01/bake =prog #01 Exposure (focus and expose test) Development : bake = prog#01/develop = prog#01	Svgcoat Gcaws Svgdev	
6.12	Poly-Si S/D etch	BT : TP=200W,BP=40W,13mT,CF_4=100,15sec ME : TP=300W,BP=150W,15mT Cl_2:HBr=50:150, EDP + 100% OE	Lam5	
6.13	Resist strip	(100:1) HF, 20sec O_2 ashing, 300W, 5min (100:1) HF, 20sec	Sink 7 Technics-c Sink 7	Polymer removal
6.14	Post cleaning	Piranha, 120°C, 20min	Sink 7	
6.15	Inspection	SEM Inspection	Leo	For CMOS, N+/P+ S/D mask/implant

Step	Process Name	Process Specification	Equipment	Comment
7.0	Metallization			
7.01	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
7.02	LTO deposition	11SULTOA, 450°C, 22min, 400nm	Tystar 11	
7.03	Measurement	LTO thickness maesurement	Nanoduv	
7.04	RTA annealing	900°C, 1min, N_2, 400nm	Heatpulse 3	
7.05	Contact mask	Resist coating : coat = prog#01/bake =prog #01 Exposure (focus and expose test) Development : bake = prog#01/develop = prog#01	Svgcoat Gcaws Svgdev	
7.06	Hard bake	120°C, 30min	Vwr	
7.07	Contact etch	ME:TP=450W,BP=750W,20mT,CHF3=100 Ar=200, 85sec	Lam5	time etch
7.08	Resist strip	(100:1) HF, 20sec O_2 ashing, 300W, 5min (100:1) HF, 20sec	Sink 7 Technics-c Sink 7	Polymer removal
7.09	Post cleaning	Piranha, 120°C, 20min	sink 8	
7.10	Pre cleaning	Piranha, 120°C, 20min (25:1) HF, 30sec	Sink 6 Sink 6	
7.11	Al sputtering	Ar:300cc, 6mT, 15cm/min, one pass, 450nm	Cpa	
7.12	Metal mask	Resist coating : coat = prog#01/bake =prog #01 Exposure (focus and expose test) Development : bake = prog#01/develop = prog#01	Svgcoat Gcaws Svgdev	0.8*exposure time
7.13	Hard bake	120°C, 30min	Vwr	
7.14	Al etch	Al etchant, manual edn point detection with eye	sink 8	
7.15	Resist strip	O_2 ashing, 300W, 5min	Technics-c	
7.16	DI Rinse	3 cycle DI rinse	Sink 8	
7.17	Sintering	VSINT400, 400°C,30min, N_2:H_2=10:1	Tylan 13	

4.7.2 Process Flows of UTBFETs with Selective Ge S/D

Step	Process Name	Process Specification	Equipment	Comment
5.06		Up to 5.06, process flows are the exactly the same		
6.0	Selective Ge S/D Formation			
6.01	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
		(25:1) HF, 50sec	Sink 6	
6.02	Ge deposition	SELDEP.019, 350°C, 300mT, GeH_4=200, 10min, 100nm	Tystar19	
6.03	Measurement	Ge thickness measurement and selectivity check	Nanoduv/Leo	
6.04	DI rinse	3 cycle DI rinse	Sink 6	
6.05	Cap LTO deposition	11SULTOA, 450°C, 90sec, 30nm	Tystar 11	
6.06	N+ S/D mask	Resist coating : coat = prog#01/bake =prog #01	Svgcoat	
		Exposure (focus and expose test)	Gcaws	
		Development : bake = prog#01/develop = prog#01	Svgdev	
6.07	Hard bake	120°C, 30min	Vwr	
6.08	N+ S/D implant	Phosphorus, $5x10^{15}cm^{-2}$, 60KeV,0°		
6.09	Resist strip	O_2 ashing, 300W, 5min	Technics-c	
6.10	Post cleaning	Piranha, 120°C, 20min	Sink 8	
6.11	P+ S/D mask	Resist coating : coat = prog#01/bake =prog #01	Svgcoat	
		Exposure (focus and expose test)	Gcaws	
		Development : bake = prog#01/develop = prog#01	Svgdev	
6.12	Hard bake	120°C, 30min	Vwr	
6.13	P+ S/D implant	Boron, $5x10^{15}cm^{-2}$, 60KeV,0°		
6.14	Resist strip	O_2 ashing, 300W, 5min	Technics-c	
6.15	Post cleaning	Piranha, 120°C, 20min	Sink 8	
6.16	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
6.17	LTO deposition	11SULTOA, 450°C, 20min, 360nm	Tystar 11	
6.18	Measurement	LTO thickness maesurement	Nanoduv	
6.19	RTA annealing	650~750°C, 1min, N_2	Heatpulse 3	
6.20	Contact mask	From contact mask to sintering, the process are the		
		same except for piranha cleaning after contact etch.		

4.7.3 Process Flows of Spacer UTBFETs

Step	Process Name	Process Specification	Equipment	Comment
4.03		Up to 4.03, process flows are the exactly the same		
4.04	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
4.04	LTO deposition	LTO deposition, 11SULTOA, 6min 30sec, 100nm	Tystar 11	
4.05	Measurement	LTO thickness measurement	Nanoduv	
4.06	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
4.07	Sacrificial $Si_{0.4}Ge_{0.6}$ deposition Recipe: SIGEVAR.019	Nucleation : 550°C, 300mT, SiH_4=200,30sec Deposition : 450°C, 300mT, SiH_4=124,GeH_4=80 18min, 200nm	Tystar 19	
4.08	Measurement	SiGe thickness measurement	Nanoduv	
4.09	Spacer mask	Resist coating : coat = prog#01/bake = prog#01 Exposure Development : bake = prog#01/develop = prog#01	Svgcoat gcaws Svgdev	
4.10	Poly-$Si_{0.4}Ge_{0.6}$ etch (Sacrificaial layer)	BT : TP=200W,BP=40W,20mT CHF_3:Ar=90:200, 10sec ME : TP=300W,BP=150W,15mT Cl_2:HBr=50:150, EDP-2sec OE : TP=250W,BP=120W,15mT HBr:O_2=200:5, 30sec	Lam5	
4.11	Resist strip	(100:1) HF, 10sec O_2 ashing, 300W, 5min (100:1) HF, 10sec	Sink 7 Technics-c Sink 7	Polymer removal
4.12	Post cleaning	Piranha, 120°C, 20min	sink 8	
4.13	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
4.14	Spacer PSG deposition	SLTO.V, PH_3=11.3, 450°C, SiH_4:O_2=5:70, 10min, 30nm	Tylan12	
4.15	Measurement	PSG thickness measurement	Nanoduv	
4.16	PSG and SiGe etch	PSG:TP=200W,BP=40W,20mT CHF_3:Ar=90:200,20sec SiGe:TP=250W,BP=120W,35mT, HBr:O_2=200:5,1min	Lam5	
4.17	SiGe residue removal	H_2O:NH_4OH:H_2O_2=(5:1:1) at 75°C, 10sec		
4.18	Inspection	SEM inspection	Leo	
4.19	Dummy mask	Resist coating : coat = prog#01/bake = prog#01 Exposure Development : bake = prog#01/develop = prog#01	Svgcoat gcaws Svgdev	
4.20	Hard bake	120°C, 30min	Vwr	
4.21	PSG wet etch	(25:1) HF, 20sec	Sink 7	
4.22	Resist strip	O_2 ashing, 300W, 5min	Technics-c	
4.23	Post cleaning	Piranha, 120°C, 20min	Sink 8	
4.24	Gate pad mask	Resist coating : coat = prog#01/bake = prog#01 Exposure Development : bake = prog#01/develop = prog#01	Svgcoat gcaws Svgdev	
4.25	Ashing	O_2 ashing, 30W, 7min	Technics-c	
4.26	Measurement	CD measuremt and additive ashing	Leo/Tech.-c	
4.27	LTO Etch	1st ME:TP=200W,BP=40W,13mT,CF_4=100, 40sec 2nd ME:TP=200W,BP=40W,20mT CHF_3:Ar=90:200, 30sec	Lam5	

Step	Process Name	Process Specification	Equipment	Comment
4.28	Resist strip	(100:1) HF, 5sec O_2 ashing, 300W, 5min (100:1) HF,5sec	Sink 7 Technics-c Sink 7	Polymer removal
4.29	Post cleaning	Piranha, 120°C, 20min	Sink 8	
4.30	Measurement	CD measurement and Inspection	Leo	
4.31	Poly-$Si_{0.5}Ge_{0.5}$ etch (Gate)	BT : TP=200W,BP=40W,20mT CHF_3:Ar=90:200, 10sec ME : TP=300W,BP=150W,15mT Cl_2:HBr=50:150, EDP-2sec OE : TP=250W,BP=120W,15mT HBr:O_2=200:5, 30sec	Lam5	
4.32	Post cleaning	(100:1) HF, 10sec Piranha, 120°C, 20min	Sink 7 Sink 8	
4.33	Measurement	CD Measurement and inspection	Leo	
5.01	Pre cleaning	From 5.01, the remaining process flows are the same	Leo	

Chapter 5

Quantum Effects

5.1 Introduction

The continued rapid advancement of MOSFET technology will require the development of novel device structures in order to improve device performance without sacrificing short-channel performance for sub-100 nm technology nodes [5.1]. Single and double gate thin body MOSFETs are promising device structures which have been successfully demonstrated to achieve improved sub-threshold characteristics, better short-channel behavior, reduced parasitic capacitance, compared with bulk MOSFETs [5.2]-[5.6]. Figure 4.13 shows the cross-sectional TEM (transmission electron microscopy) micrograph of a single gate ultra-thin body (UTB) MOSFET. In UTB MOSFETs with body thickness < 10 nm, the inversion charge layer should be treated quantum mechanically as a 2-dimensional electron gas which is confined in a narrow trapezoidal potential well formed by the ultra-thin body and the two adjacent oxide layers (gate oxide and buried oxide), as shown in Figure 5.2. When the width of this potential well is comparable to the inversion layer thickness, the band structure and the density of states (DOS) will be significantly changed, so that carrier confinement effects such as energy quantization should be considered. Therefore, a classical treatment which assumes a 3-

dimensional electron gas is no longer valid and a quantum mechanical treatment of the 2-dimensional electron gas should be employed. As the body thickness T_{Si} decreases, the sub-band splitting in conduction and valence bands becomes more significant, thus increasing the threshold voltage and enhancing mobility for a thinner body than a critical thickness. The quantum mechanical (QM) confinement of charge carriers has been studied numerically and experimentally in heavily doped ($> 10^{18}$ cm^{-3}) bulk MOSFETs as well as in lightly or moderately doped ($< 10^{18}$ cm^{-3}) ultra-thin (< 10 nm) body MOSFETs [5.7]-[5.10]. However, previous studies have been based on numerical simulations, which are not appropriate for efficient circuit simulation and which cannot provide explicit expressions for the effects of varying device parameters such as SOI thickness and body doping concentration on threshold voltage.

In this chapter, the impact of quantum confinement on the threshold voltage and hole mobility in UTB MOSFETs is studied, and an analytical model for the threshold voltage shift is proposed. The ultra-thin body and adjacent silicon dioxide (gate and buried oxide) layers form a thin potential well, which causes sub-band splitting (between 2-fold and 4-fold valleys of the conduction band and between light and heavy hole sub-bands of the valence band). The quantum confinement of inversion charges results in a smaller density of states, so that more energy-band bending is required to attain a desired inversion-charge density as compared with a bulk silicon device. Thus, an increase in threshold voltage is expected for ultra-thin body devices even with a lightly or moderately doped body. As the body thickness T_{Si} is reduced, heavier effective mass (confinement effective mass: m_z) of electrons or holes is preferably residing at the lower energy level, therefore, its occupation numbers at the lower energy level increases. As a

result, mobility is also expected to be improved because light effective mass (conductivity effective mass: m_c) of carriers has high mobility and the inter-band scattering rate reduces.

5.2 Threshold Voltage Shift by Quantum Confinement

5.2.1 Theoretical Background

An inversion layer is formed in a p-type semiconductor NMOSFETs when the energy bands at an interface of a gate oxide and silicon substrate are bent so strongly that the majority of carriers near the interface are electrons even though the carriers in the bulk are holes. The converse is true for n-type silicon (PMOSFETs). The energy levels in an inversion layer are represented by series of sub-bands, each of which is a 2-dimensional continuum of levels. They are associated with the two degrees of freedom parallel to the interface. The 2-dimensional electron gas distributions within inversion layers of NMOSFETs have been evaluated by solving the coupled Schrödinger equation and Poisson equation self-consistently based on the effective-mass approximation [5.11]. The band banding at a semiconductor surface can be characterized by an electrostatic potential $\psi(z)$. In the effective-mass approximation, the 3-dimensional Schrödinger equation is decoupled into a 1-dimensional equation that describes the envelope function perpendicular to the interface, $\xi(z)$, that constrains Bloch waves traveling parallel to the

interface (*x-y* plane). Thus, the electronic wave function for *j*th sub-band is the product of the Bloch function at the bottom of the conduction band and an envelope function [5.12]

$$\varphi_j(x,y,z) = \xi_j e^{i\theta} e^{ik_x x + ik_y y}, \tag{5.1}$$

where θ depends on k_x and k_y, and $\xi_j(z)$ is the solution of

$$\frac{d^2\xi_j}{dz^2} + \frac{2m_z}{\hbar^2}\left[E_j + e\psi(z)\right]\xi_j(z) = 0. \tag{5.2}$$

The boundary conditions which are applied assume that the energy eigen function vanishes at the interface of gate oxide and silicon substrate and the $\xi(\infty) = 0$. It is assumed that the potential barrier to electrons in the potential well is infinitely high at the interface of the gate oxide and silicon substrate. Actually the barrier is approximately 3.1eV in the case of SiO_2. However, this assumption is no longer valid when the impact of wave function penetration is significant, i.e. the gate oxide thickness is very thin. For a simple self-consistent calculation, the approximation is used. The eigenvalue E_j obtained from the solution of Eq. (5.2), is given by

$$E_{ej}(\vec{k}) = E_j + \hbar^2 k_x^2 / 2m_x + \hbar^2 k_y^2 / 2m_y, \tag{5.3}$$

where m_x and m_y are the effective masses for motion parallel to the surface and m_z is the effective mass for motion perpendicular to the surface.

The potential $\psi(z)$ which appears in Eq. (5.2) is the solution of Poisson's equation

$$\frac{d^2\psi(z)}{dz^2} = -[\rho_{depl} - e\sum_j N_j \xi_j^2(z)]/\varepsilon_{Si}, \tag{5.4}$$

where N_j is the carrier concentration in the *jth* sub-band, given by

$$N_j = \int_E^{\infty} N(E) f(E) dE \tag{5.5}$$

f(E) in Eq. (5.5) is the Fermi-Dirac distribution function and $N(E) = 4\pi g_i \sqrt{m_x m_y}/h^2$. In the heavily doped body bulk MOSFETs, the potential, $\psi(z)$ can be approximated as triangular potential ($-\mathcal{E}_s z$). It leads to the Airy equation [5.13] with solutions

$$\xi_j(z) = Ai\{(2m_{iz_j} e\mathcal{E}_s/\hbar^2)^{1/3}[z-(E_j/e\mathcal{E}_s)]\}, \qquad (5.6a)$$

$$E_{ij} \approx \left[\frac{3hq\mathcal{E}_s}{4\sqrt{2m_{iz}}}(j-1/4)\right]^{2/3}, \; j=1,2,3,\ldots\ldots \qquad (5.6b)$$

The subscript *i* stands for the corresponding energy valleys, i.e. 2-fold valley and 4-fold valley for the conduction bands and light hole band, heavy hole band for the valence band, simply. Eq. (5.6a) and (5.6b) are still valid in the UTB MOSFETs because of its similarity of the potential shape as shown in Fig. 5.2.

When the device is in moderate to strong inversion, the Airy function does not describe the ground state eigenfunction accurately, because of the perturbation of the potential due to the inversion layer charges [5.14]. Fortunately, Stern [5.11] introduced a variational approach that had been shown to provide a good estimate of the wave function of the lowest sub-band. The ground state wave function and eigen energy was represented by

$$\xi_0(z) = \frac{b^{3/2}}{\sqrt{2}} z \cdot \exp(-bz/2), \qquad (5.7a)$$

$$E_0 = 2\left(\frac{e^2\hbar}{\sqrt{m_z}\varepsilon_{Si}}\right)^{2/3} \cdot \left(N_{depl} + \frac{55}{96}N_{inv}\right) \cdot \left(N_{depl} + \frac{11}{32}N_{inv}\right)^{-1/3}, \qquad (5.7b)$$

where $b = \left(12m_z e^2(N_{depl} + 11/32N_{inv})/\varepsilon_{Si}\hbar^2\right)^{1/3} - 4N_A/3(N_{depl} + 11/32N_{inv})$, N_A is the net acceptor concentration [5.11]. The ground state energy can be obtained by adjusting *b*

iteratively. When the quantum confinement effect is very large (i.e. large band bending), most carriers are occupying the ground state. In this regime, the variational approach can provide a simple and accurate estimate of the ground state energy even for a strong inversion without numerical simulation because the contributions from higher sub-bands are negligible. Figure 5.1 shows the inversion charge distribution generated by Eq. (5.7a). Inversion carrier distribution is strongly affected by b, which is proportional to $m_z^{1/3}$. Since the electron confinement effective mass (m_z) is larger than the hole confinement effective mass (m_z), the peak of the spatial distribution of holes lies deeper than that for electrons as shown in Fig. 5.1. This model gives us simple physical insight, allowing quantum confinement to be understood and represented by the effective mass of the inversion carriers.

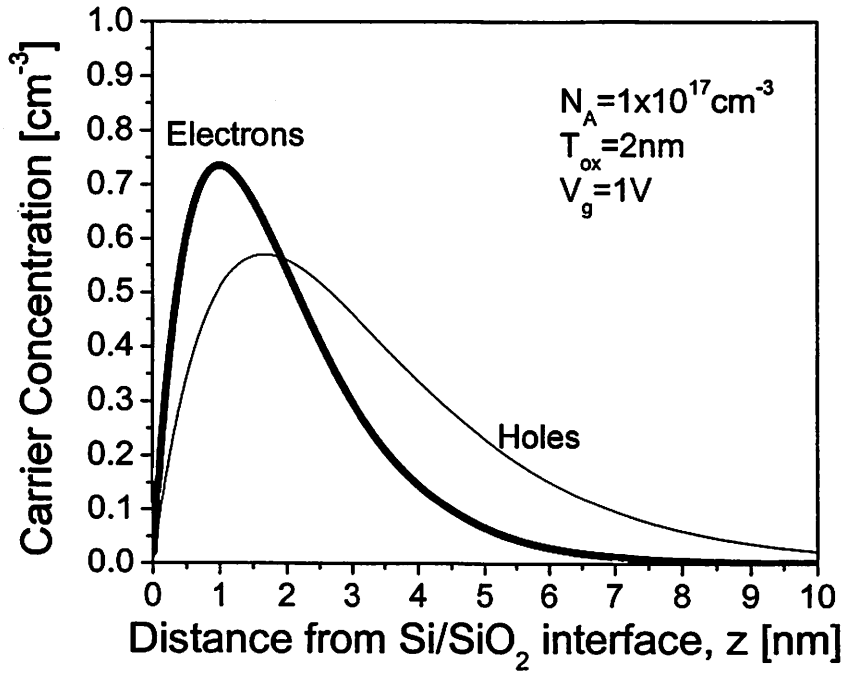


Figure 5.1 Inversion carrier concentration for electrons and holes in bulk device.

The above self-consistent calculation has been performed on the basis of the Hartree approximation [5.15], i.e. without consideration of electron-electron interaction in the conduction bands or hole-hole interaction in the valence bands. As the spatial extent of the electron density (hole density in PMOS) in NMOS becomes smaller, the exchange-correlation effect should be taken into account [5.16]. It has been shown that the ground state level was significantly lowered (22meV at N_{sub}=1x10^{14}cm^{-3}) which energy levels of higher sub-bands were affected only slightly [5.16][5.17]. This shift due to exchange-correlation is expected to be more significant as the body doping increases in bulk devices or the body thickness T_{Si} decreases in UTB devices because carrier-to-carrier interactions increase as more carriers reside at the ground state energy level due to increasing quantum confinement. A competing factor which affects the sub-band system is the image potential, which compensates many-body effects [5.18]. When the effect of wave function penetration is included, it has been reported that the ground state energy level is lowered to 20meV, which should be considered seriously for high doping bulk device and UTB device in the case of thinner oxide for more accurate calculation [5.19].

5.2.2 Analytical Modeling of Threshold Voltage Shift

The actual inversion charge layer profile can be determined by solving two coupled equations self-consistently, namely the Poisson equation and Schrödinger equation using Fermi–Dirac statistics [5.20]-[5.24]. However, these calculations are computationally intensive, and hence are not appropriate to incorporate into circuit simulations. In the sub-threshold regime, i.e. when the inversion charge density is low,

the energy band bending depends on the depletion charge. Then, it is possible to decouple the two equations assuming a certain potential well approximation.

Figure 5.2 shows the energy band diagrams for (a) bulk and (b) UTB MOSFETs. For the UTB MOSFET, a potential well is formed by the two oxide layers (gate oxide and buried oxide) and the ultra-thin body silicon. Due to the light body doping concentration and the thin body thickness, it is trapezoidal rather than triangular as for a bulk MOSFET with high doing concentration. Numerical simulation results using Schred [5.25] show that the trapezoidal well approximation is valid in the sub-threshold regime, as shown in Figure 5.3.

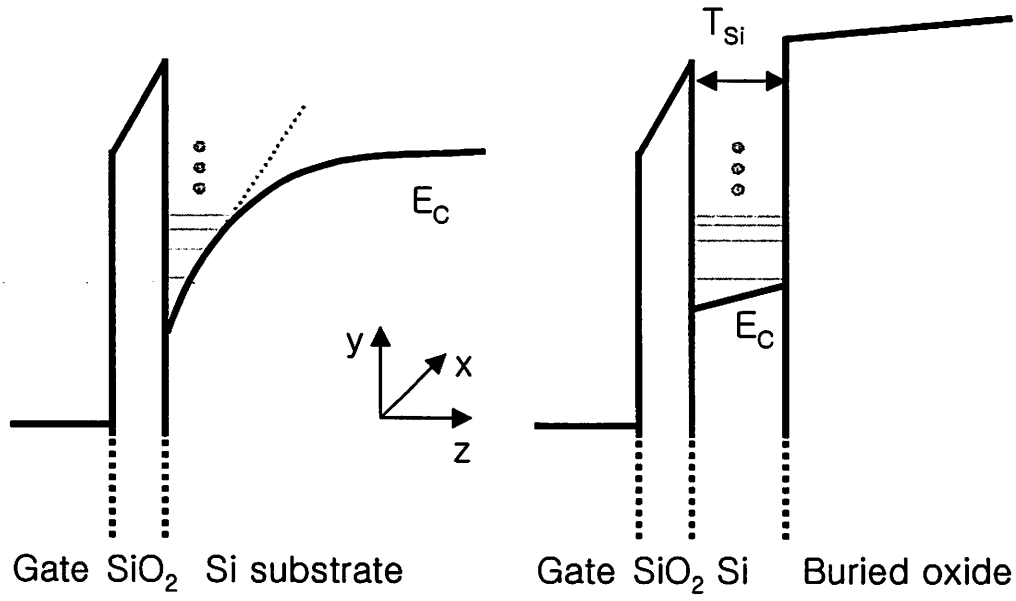


Figure 5.2 Energy band diagrams and eigen-energy states in (a) bulk MOSFET with heavily doped body and (b) ultra-thin body (UTB) MOSFET with lightly doped body.

For UTB MOSFET, the potential well is formed by two oxide layers (gate oxide and buried oxide) and the conduction band of silicon. The potential wells for bulk and UTB MOSFET can be approximated as triangular and trapezoidal, respectively.

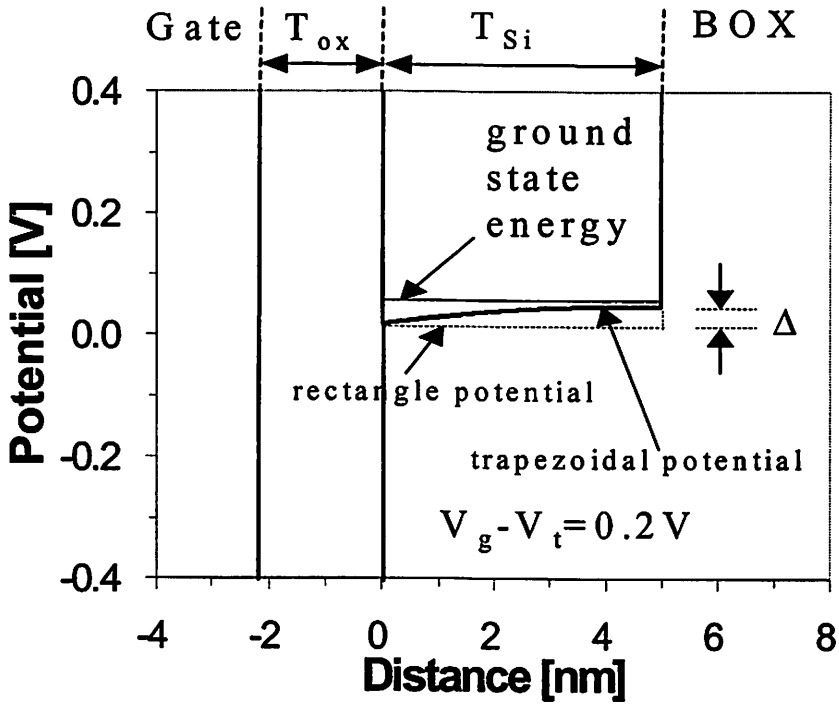


Figure 5.3 Numerical simulation results for the ground state energy for $V_g - V_t = 0.2\ V$ using the simulator Schred [5.25]. The thicknesses of the gate oxide and body is 2.1 nm and 5 nm, respectively. The doping concentration is 10^{16} cm^{-3}.

This trapezoidal well can be approximated as a rectangle well when the lowest eigen energy is above the conduction band edge because gate the oxide barrier height is much greater than the band bending Δ in Figure 5.3. For an accurate model, this effect should be treated perturbatively. However, it is not appropriate for simple analytical modeling and circuit simulation. Assuming an infinite potential barrier at $x=0$ and $x=T_{Si}$,

the Schrödinger equation can be solved with the boundary conditions that the wave functions are zero at both oxide interfaces; i.e. wave function penetration into the gate and buried oxides is assumed to be negligible. The energy eigenvalues are given by

$$E_{ij} = \frac{j^2 h^2}{8 m_{zi} T_{Si}^2} \quad , \quad j = 1, 2, 3, \ldots \tag{5.8}$$

where h is Plank's constant (6.63 x 10^{-34} J-s), T_{Si} is the MOSFET body thickness and m_{zi} is the effective mass perpendicular to the interface. When the potential is assumed to be triangular, the energy eigenvalues are given by Eq. (5.6b).

The 2-dimensional nature of hole inversion layers has not received as much attention as has the case of electron inversion layers. This is due to the coupling between the valence bands and the inadequacy of the effective mass approximation for warped valence band structure. Hu et al. [5.24] reported hole quantization in PMOSFETs using the simple effective mass approximation considering heavy hole and light hole sub-bands. The 1-dimensional Schrödinger equation for holes is the same as Eq. (5.2). For effective masses in the z-direction, m_{zi} (i : light or heavy hole), the warped energy surfaces of heavy hole and light hole sub-bands are considered as [5.14][5.24]

$$E_h = \left(\frac{\hbar^2 k^2}{2m_0}\right)\left\{A \mp \left[B^2 + C^2(k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2)k^{-4}\right]^{1/2}\right\}, \tag{5.9}$$

where $k^2 = k_x^2 + k_y^2 + k_z^2$, m_0 is the free electron mass, A, B, C are constants with a certain uncertainty determined from cyclotron resonance experiments [5.26]. Here the values of A, B, *and* C are chosen to be 4.22, 0.78, and 4.8, respectively [5.14]. The + and – sign of the square bracket refer to light and heavy holes, respectively. The effective masses m_{zi} for heavy holes and light holes are taken near the top of valence band along the z-direction (i.e. k_x=k_y=0) and can be deduced as

$$m_{zi} = m_0 /(A \mp B), \tag{5.10}$$

The density-of-states mass is expressed as

$$m_{di} = \left(\frac{m_0}{A}\right) \cdot \int_0^{2\pi} [A \mp \sqrt{(B^2 + C^2 \cos^2\theta \sin^2\theta}]^{-1} d\theta, \tag{5.11}$$

All holes effective masses used in this work are summarized in Table 5.1. The electron masses were cited from [5.11] and hole effective mass are calculated with Eq. (5.10) and Eq. (5.11). The effective mass for carriers in both the heavy hole sub-bands are degenerate, similar to that for electrons in the 2-fold valley and 4-fold valley. In addition, the heavy hole/split-off band hybrid which is shifted by 44meV due to spin-orbit coupling is ignored to simplify the calculation.

		Density of States Effective Mass (m_{di})	Z-directional Effective Mass (m_{zi})	Degeneracy (g_i)
Conduction Band	Two-fold Valley	0.19 m_0	0.92 m_0	2
	Four-fold Valley	0.42 m_0	0.19 m_0	4
Valence Band	Light Hole Sub-band	0.25 m_0	0.20 m_0	1
	Heavy Hole Sub-band	0.65 m_0	0.29 m_0	1

Table 5.1 Effective masses and degeneracy for conduction and valence bands in (100) silicon.

As the body thickness (T_{Si}) of a (100) oriented UTB MOSFET decreases and becomes comparable to the inversion layer thickness, the potential well causes significant sub-band splitting between two-fold and four-fold valleys of the conduction band for NMOS, and between light hole and heavy hole sub-bands of the valence band for PMOS, respectively. For example, in the case of the NMOSFET, the energy levels of the lower valley (two-fold degeneracy, $m_{zi} = 0.92\ m_0$) and the higher valley (four-fold degeneracy, $m_{zi} = 0.19\ m_0$) are designated as $E_0, E_1, E_2, \ldots$ and $E_0', E_1', E_2', \ldots$, respectively.

Because of the confinement of charge carrier normal to the surface in UTB MOSFET, it is necessary to treat them as 2-dimensional carriers quantum mechanically. For 2-D density of states, the number of electron states per unit area with an energy between E and $E + dE$ is expressed as follows:

$$N(E)dE = \frac{4\pi \cdot g_i \cdot \sqrt{m_x m_y}}{h^2} dE, \tag{5.12}$$

where g_i is the degeneracy of the sub-band, and $m_{di} \equiv \sqrt{m_x m_y}$ is the effective mass of density of states.

The total inversion charge (electrons) per unit area (Q^{QM}) is calculated by summing over all the sub-bands in two-fold and four-fold valleys as follows:

$$\begin{aligned} Q^{QM} &= \sum_j \int_{E_{min}}^{\infty} N(E) f(E) dE \\ &= \frac{4\pi qkT}{h^2} \left\{ g \cdot \sqrt{m_x m_y} \sum_j \ln[1 + e^{(E_f - E_{min})/kT}] + g' \cdot \sqrt{m_x' m_y'} \sum_j \ln[1 + e^{(E_f - E_{min}')/kT}] \right\} \end{aligned} \tag{5.13}$$

where $f(E)$ is the Fermi – Dirac distribution function, and E_{min} is the minimum energy for the jth sub-band and defined as $E_{min}=E_c - q\psi_S + E_j$, where ψ_S is the surface potential. Note that the above equation can be further simplified by the fact that, because of the light body doping concentration, the Fermi level is located at least several kT below the lowest sub-band, thus $ln(1+e^{(E_f-E_{min})/kT})$ can be approximated as $e^{(E_f-E_{min})/kT}$. Also, using $E_f - E_c = kT\ln(n_i^2 / N_c N_{body})$, Eq. 5.13 can be expressed as

$$Q^{QM} = \frac{4\pi qkT \cdot n_i^2}{h^2 \cdot N_c N_{body}} \left\{ 2 \cdot \sqrt{m_x m_y} \sum_j e^{-E_j/kT} + 4 \cdot \sqrt{m_x' m_y'} \sum_j e^{-E_j'/kT} \right\} \qquad (5.14)$$

where N_C is the effective density of states in the conduction band, n_i is the intrinsic carrier density, N_{body} is the body doping concentration and the eigen energy (E_j and E_j') for each valley is given by Eq. (5.8).

Figure 5.4 shows the fraction of inversion charge residing in the lowest sub-band, calculated using Eq. (5.14), as a function of body thickness for NMOS and PMOS. When the body thickness (T_{Si}) decreases below 10 nm, a larger fraction of the inversion charge occupies the lowest energy state. Also, this change in occupancy occurs at a thicker T_{Si} for the NMOSFET than for the PMOSFET due to the larger difference in the effective mass (m_{zi}) between two-fold and four-fold valleys than between light-hole and heavy-hole subbands. Quantum-mechanical (QM) confinement is more significant in electrons than in holes in the given body thickness. This indicates that the threshold

voltage shift will be larger for NMOS than for PMOS, which is consistent with previous reports [5.27].

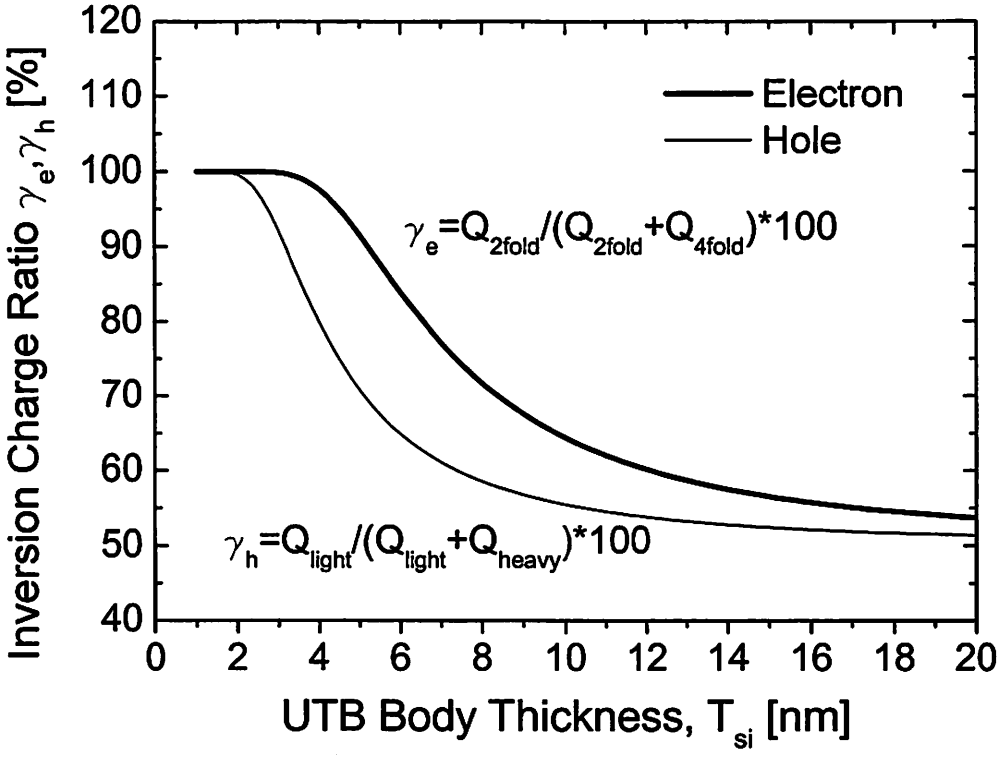


Figure 5.4. Calculated fraction of inversion charge residing at the lowest energy level as a function of body thickness for NMOSFET and PMOSFET using Eq. (5.14).

According to the classical definition of threshold voltage (V_t), the inversion charge per unit area (Q^{CL}) for the sub-threshold region is given by [5.28]

$$Q^{CL} = \frac{kT \cdot n_i^2}{\mathcal{E}_S \cdot N_{body}} e^{\frac{q\psi_s}{kT}} \tag{5.15}$$

where $\mathcal{E}_s$ is the surface electric field. Figure 5.5 shows the inversion charge densities (Q^{QM} to Q^{CL}) for NMOS as functions of body thickness (T_{Si}) with and without

considering quantum confinement effects calculated using Eq. (5.14) and Eq. (5.15). When T_{Si} is larger than 15 nm, the energy spacing between the sub-bands is negligible compared to kT at room temperature (~26 meV). Therefore, a large number of sub-bands are occupied and there is little difference between Q^{QM} and Q^{CL}. However, when T_{Si} is smaller than 15 nm, the sub-band spacing is greater than kT. Due to sub-band splitting, the density of states (DOS) is reduced as the body thickness (T_{Si}) decreases. As a result, Q^{QM} is much smaller than Q^{CL}. Thus, in order to obtain the same amount of inversion charge density, additional energy band bending is necessary for UTB MOSFETs. Using Eq. (5.14) and (5.15), the additional band bending required ($\Delta\psi_s$) can be calculated as

$$\Delta\Psi_s = \frac{kT}{q}\ln\left(\frac{Q^{CL}}{Q^{QM}}\right) \tag{5.16}$$

The threshold voltage shift for NMOSFETs due to QM confinement can be expressed as

$$\begin{aligned}\Delta V_t^{QM} &= \frac{dV_g}{d\psi_s}\Delta\psi_s = m\Delta\psi_s \\ &= \left(1+3\frac{T_{ox}}{T_{Si}}\right)\frac{kT}{q}\ln\left(\frac{h^2 N_C}{4\pi q\mathcal{E}_s}\sum_{ij} g_i m_{di} e^{-E_j/kT}\right)\end{aligned} \quad \text{[5.36]}, \tag{5.17}$$

where m is the body effect factor which is defined as $1 + 3\,(T_{ox}/W_{dmax})$ [5.28], T_{ox} is the gate oxide thickness, $\mathcal{E}_s$ is the surface electric field (defined as $\mathcal{E}_s = qN_{body}T_{Si} / \varepsilon_{si}$), and m_{di} is density of states effective mass. For PMOSFETs, the threshold voltage shift can be calculated using Eq. (5.17) with the appropriate values for g_i and m_{di} (as indicated in Table 5.1), E_j, and N_V (instead of N_C).

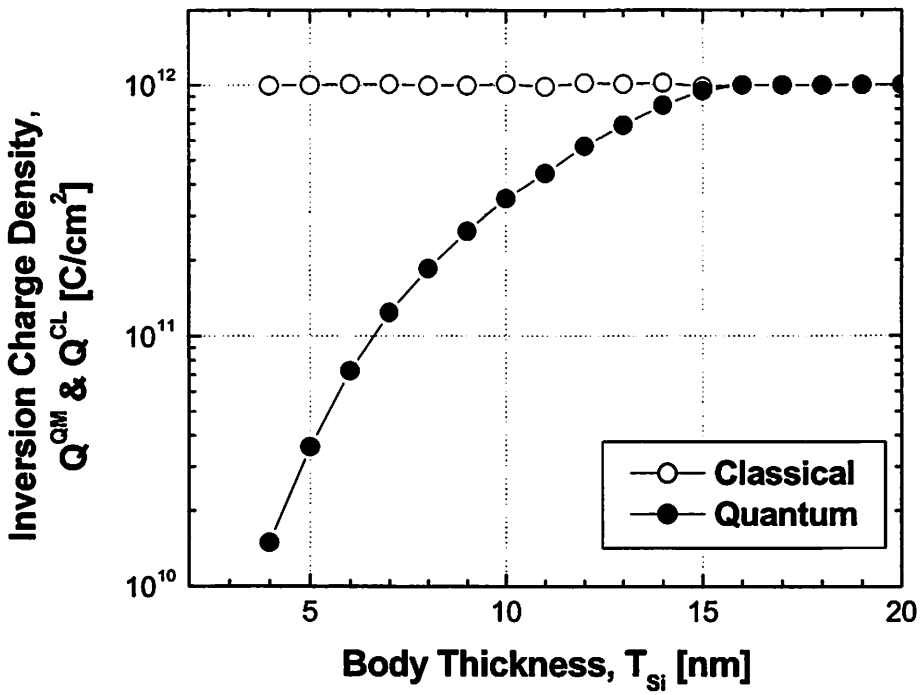


Figure 5.5 Inversion charge density for NMOSFET as a function of body thickness (T_{Si}) with and without considering quantum confinement. For T_{Si} < 15 nm, additional band bending is required for a UTB MOSFET to obtain the same amount of inversion charge density as in a bulk MOSFET.

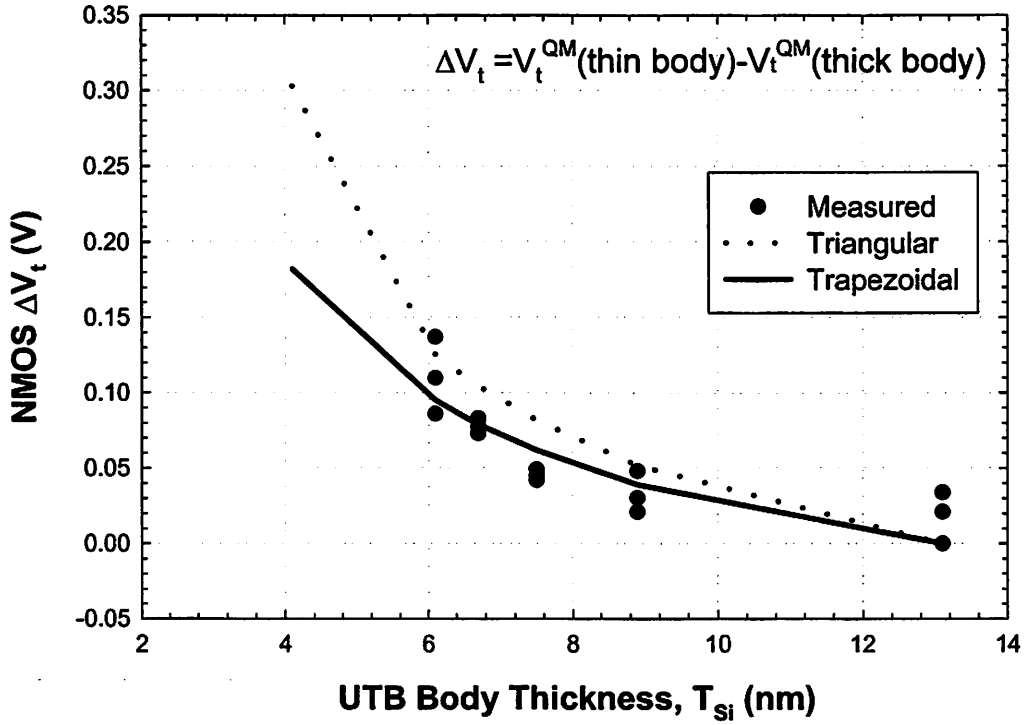
5.2.3 Experimental Results and Discussions

In order to illustrate the validity of the model, the predicted threshold voltage shift is compared with experimental data. UTB MOSFETs were fabricated with body thickness T_{Si} ranging from 20 nm to 4 nm, measured using a Nanometrics 210 XP Scanning UV Microspectrophotometer and calibrated using cross-sectional TEM (transmission electron microscopy) analysis. The gate oxide thickness is 2.1 nm and the 60nm gate length is chosen in order to minimize short-channel effects. Although 60 nm

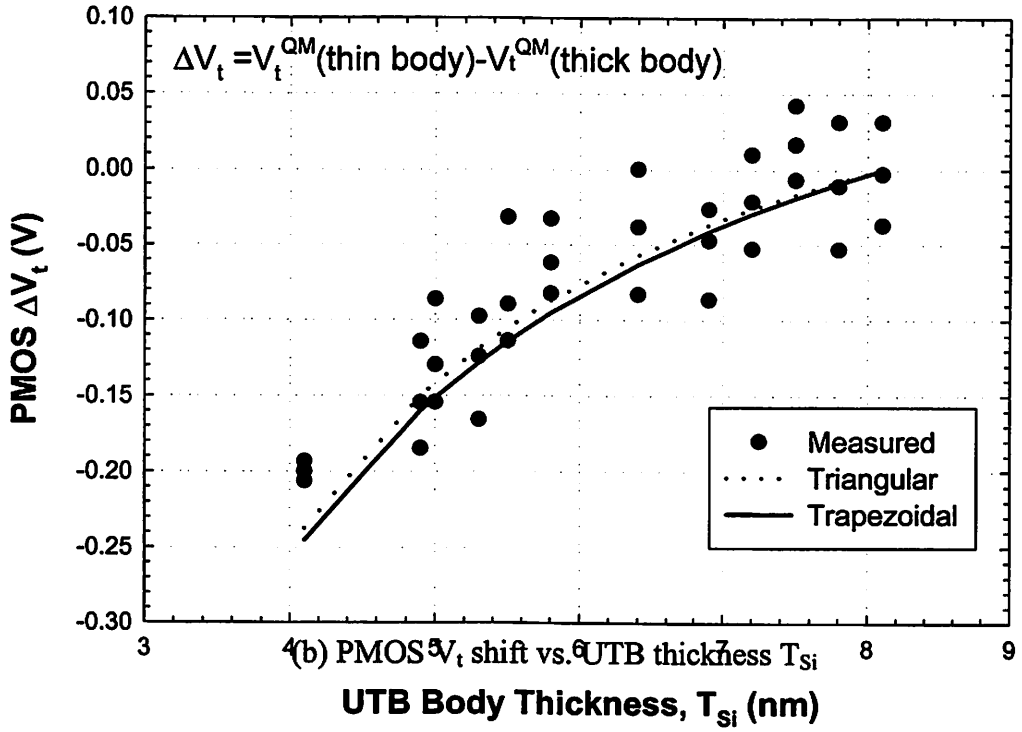
gate length is very small, short-channel effects are successfully suppressed when T_{Si} is thinner than 10 nm [5.5]. The detailed process was reported previously [5.5]. The threshold voltage (V_t) is defined as the gate voltage when the drain current (I_d) is *100 nA/um* when the applied drain voltage (V_d) is 0.05 V. After measuring the threshold voltage of each MOSFET, the threshold voltage shift is calculated using $\Delta V_t = V_t^{QM}(thin\ body) - V_t^{QM}(T_{Si}=20\ nm)$ [5.36].

Figures 5.5 (a) and (b) show the threshold voltage shift (ΔV_t) for NMOS and PMOS devices as functions of body thickness (T_{Si}) for the measured data (solid circle), the analytical model based on a trapezoidal potential (solid line), and the model based on a triangular potential approximation (dotted line). Eq. (5.6b) was inserted into Eq. (5.17) for the triangular approximation case and Eq. (5.8) was inserted into Eq. (5.17) for the trapezoidal approximation case. As the body thickness decreased below 10 nm, the threshold voltages of both NMOS and PMOS start to increase as explained in the previous section. The threshold voltage shift for NMOS due to thin body thickness begins earlier than that for PMOS. For example, when the body thickness is 7 nm, the threshold voltage increases by 75 mV for NMOS and 35 mV for PMOS as compared to thicker (> 20nm) body devices. This is because the fraction of inversion charge that occupies the lowest energy states increases faster in the NMOSFET than in the PMOSFET, as explained in Figure 5.4. Both models fit the measurement data quite well. There is a larger difference between the two models for NMOS than for PMOS because of a difference of effective mass (m_{zi}). The eigen energy is inversely proportional to the effective mass ($\propto m_{Zi}^{-1}$) in the trapezoidal well approximation (Eq. (5.8)), but to one third of the effective mass ($\propto m_{Zi}^{-1/3}$) in the triangular well approximation (Eq. (5.6b)). This

difference in dependence on effective mass is more apparent for NMOS devices because the difference of m_{zi} between 2-fold valley and 4-fold valley is larger than between light-hole band and heavy-hole band.



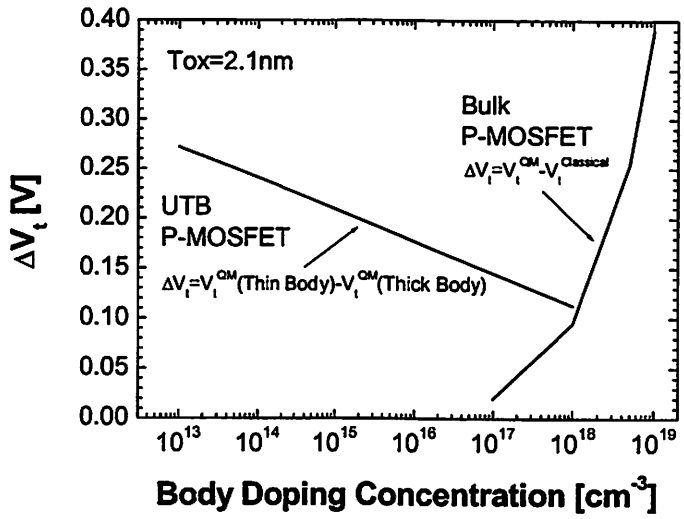
(a) NMOS V_t shift vs. UTB thickness T_{Si}



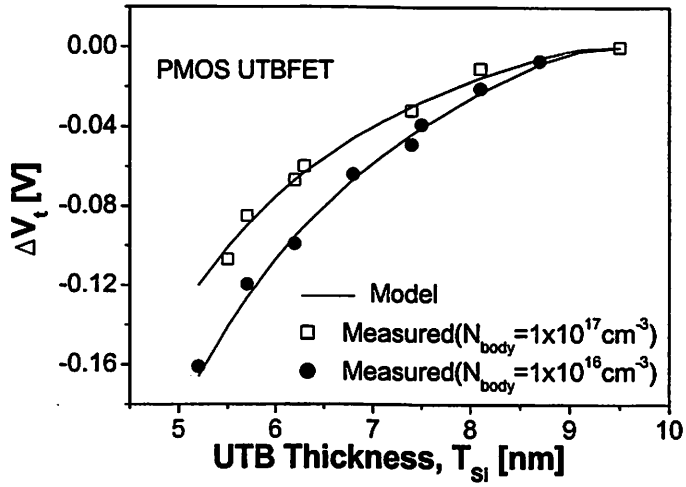
Figures 5.6 The threshold voltage shift (ΔV_t) for (a) NMOS and (b) PMOS as a function of body thickness (T_{Si}) for measured data (solid circle), the analytical model based on a trapezoidal potential well (solid line), and the model based on triangular potential approximation (dashed line), respectively.

Figure 5.7 (a) shows the predicted threshold voltage shift (ΔV_t) for bulk and UTB PMOSFETs, based on the analytical model, as a function of body doping concentration (N_{body}). Although short-channel effects can be suppressed without body doping in the UTB device, some body doping may be required to adjust the threshold voltage. The predicted threshold voltage shift is compared with measurement data for different body doping concentrations (N_{body} = 10^{16} and 10^{17} cm^{-3}) in Figure 5.7 (b). In contrast to the situation for the bulk MOSFET, a lower body doping concentration (N_{body}) results in a

larger threshold voltage shift in the UTB MOSFET. This phenomenon is due to the fact that potential well is more like trapezoidal in lightly doped UTB and triangular in relatively heavily doped UTB. That is, for the lightly doped body UTB MOSFETs, the electric field is very small (not zero) and the ground state is slightly above the conduction band edge at the buried oxide interface, as shown in Figure 5.2. Therefore the potential well can be approximated as trapezoidal and the lowest eigen-energy is a strong function of the body thickness ($\propto T_{Si}^{-2}$). However, in the case of the heavily doped body, the electric field is large and the potential well can be approximated as triangular rather than trapezoidal, so the lowest eigen-energy is a weak function of the body thickness ($\propto T_{Si}^{2/3}$). Thus, a small change of body thickness results in a relatively larger threshold voltage shift in lightly doped UTB than in heavily doped UTB.



(a) Threshold voltage shift by quantum effect in PMOS UTB and bulk device



(b) Measured threshold voltage shift by quantum effect in PMOS UTB device.

Figure 5.7 (a) Threshold voltage shift (ΔV_t) for the bulk and UTB PMOSFET as a function of body doping concentration (N_{body}), based on the analytical model. (b) Measured threshold voltage shift in UTB PMOSFET with different body doping concentrations ($N_{body} = 10^{16}$ & 10^{17} cm^{-3}). ΔV_t is larger in lightly doped body UTB MOSFETs.

In UTB MOSFETs, threshold voltage variation (σV_t) can be induced by fluctuations in both the body thickness and dopant concentration [5.28]-[5.32]. An extra charge sheet ΔQ can be added to represent dopant number deviation, ΔV_t is expressed as

$$\Delta V_t^{dopant} = \left(\frac{\Delta Q}{C_{ox}}\right), \tag{5.18}$$

where $\Delta Q = \frac{q\sqrt{N_{body}(z)LW\Delta z}}{LW}$, L and W are the channel length and width, respectively and N_{body}(z) is the doping concentration [5.29]. The standard deviation of the threshold voltage (σV_t^{dopant}) can be obtained by summing the contributions of the charge sheets

$$(\sigma V_t^{dopant})^2 = \sum (\Delta V_t^{dopant})^2 . \tag{5.19}$$

The result is

$$\sigma V_t^{dopant} = (q / C_{ox})\sqrt{N_{body} T_{Si} / LW} . \tag{5.20}$$

This model makes it possible to estimate the threshold voltage variation for arbitrary vertical channel dopant distributions very easily. Fig. 5.8 shows the threshold voltage shift by quantum effect (σV_t^{QM}) and by statistical dopant fluctuation (σV_t^{dopant}) where T_{Si}=5 nm, T_{ox}=2.1 nm, L=20 nm, and W=100 nm. σV_t^{QM} was calculated by using Eq. (5.17) and σV_t^{dopant} was obtained with Eq. (5.20). Both effects are competing each other, which affect the total threshold voltage shift oppositely as shown in Fig. 5.7. Decreasing the body doping concentration decreases the standard deviation of threshold voltage due to dopant fluctuations so that the uniformity of body thickness dominates the total threshold voltage variation at moderate and low body doping concentrations.

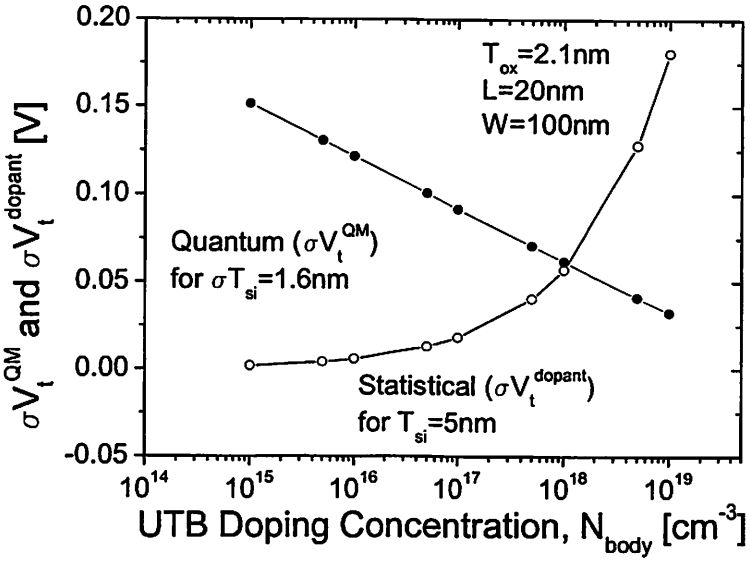
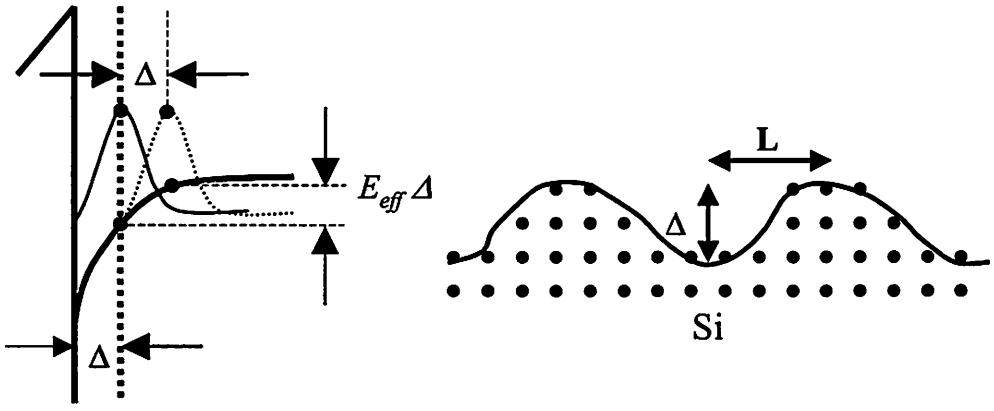


Figure 5.8 Standard deviation of threshold voltage induced by fluctuations in body thickness (solid circle) and body dopant concentration (open circle).

The remaining factor affecting the threshold voltage is the channel surface roughness. Assuming that the interface is shifted with respect to its reference position by a distance Δ in part of the channel as shown in Fig 5.9 (b), the carrier wave function will be shifted by Δ (dashed line in Fig. 5.9. (a)). As a result, the potential felt by the centroid of the carriers will change by $E_{eff}\Delta$ as shown in Fig. 5.9 (a). This change gives rise to an increased DOS for inversion charges, which results in a reduced the threshold voltage shift by quantum confinement. Thus, the rough surface is expected to reduce the threshold voltage shift by quantum effects. However, this roughness will not affect the threshold voltage shift by the statistical fluctuation of dopants significantly because Δ is usually the order of one or two silicon monolayers.



(a) Energy band diagram at rough Si-SiO_2 interface (b) Cross-section of rough surface

Figure 5.9 Effect of a shift Δ of Si-SiO_2 interface on the inversion charge wavefunction. The potential energy at the centroid is raised by approximately $E_{eff}\Delta$. Δ is the root-mean-square (rms) value of rough surface and L is the surface correlation length.

5.3 Hole Mobility Enhancement by Quantum Effects

Takagi et al. [5.33] reported that electron mobility was enhanced as the body thickness was decreased in single gate UTB devices with numerical simulation. It was decreased monotonously when $T_{Si} < 20nm$, however, it was increased when $T_{Si} < 5nm$. As T_{Si} becomes thinner than 20nm, the SOI physical thickness starts to limit the extent of the wave function of electrons. As a result, electron mobility slightly decreased because of the decrease in the inversion layer thickness. When T_{Si} becomes thinner than $5nm$, the sub-band energy of the 4-fold valleys is lifted up, thus the occupancy of the 2-fold valleys and resultant total mobility increase. Shoji et al. [5.34] also showed the same trend with a relaxation time approximation and a one-dimensional self-consistent calculation in single

gate UTB devices. They demonstrated that the electron mobility enhancement was still observed in double gate thin body devices using numerical calculations [5.35]. However, there have not yet been any reports of simulation or experimental studies for hole mobility behavior in UTB devices. This is due in part to the complex nature of the valence band such as the coupling between the valence bands and the inadequacy of the effective mass approximation for the warped valence band structure. As a result, hole mobility does not lend itself readily to simplified theories or models. However, the PMOS threshold voltage shift by quantum effects predicted using a simple effective mass approximation with consideration of heavy hole and light hole sub-bands, was well matched with measured data in the previous section. Using the same approach, which is very similar to that for conduction bands, experimental hole mobility behavior, which depends on the body thickness changes, is explained.

The effective hole mobility was extracted from the linear I_{ds}-V_{gs} data with a compensation of series resistance, from several devices at each body thickness. Higher mobility compared with the hole mobility of bulk PMOSFETs is obtained with lower body doping concentration and higher drive current for a relatively thick gate oxide (2.1nm) in L_g=*30nm* regime. Fig. 5.10 shows the hole-mobility dependence on UTB thickness with the maximum, minimum and average mobility values indicated. The data show that mobility decreases with body thickness T_{Si} reduction down to 5 nm, and then increases as the body thickness decreases further. This trend is similar to that expected for n-channel UTBFETs [5.33][5.34]. When $5nm < T_{si} < 8nm$, the total amount of inversion charge is decreased because of the increased threshold voltage due to the quantum-mechanical confinement, and the gate capacitance is increased due to the

reduced inversion layer capacitance. As a result, hole mobility gradually decreases. The enhancement in hole mobility seen for $T_{si} < 5nm$ is caused by an increase in the fraction of light holes in the lowest sub-band, which has higher mobility than the other sub-bands as shown in Fig. 5.4. Since the quantization energy is proportional to $1/m_{eff}^{a}T_{Si}^{b}$ (a, b >0) in Eq. (5.6b) and (5.8), the bands are shifted away form each other as T_{Si} decreases. This results in decrease in effective mass and inter-bands scattering rate and an enhancement in mobility. Since the total mobility is determined by averaging sub-band mobilities weighted by their fractions γ_i (in Fig. 5.4) [5.34]

$$\mu = \sum_i \gamma_i \mu_i \tag{5.21}$$

the higher mobility is attributed to the rapid increase of the fraction of holes in the lowest sub-band which has a higher mobility than other sub-bands for $T_{Si} < 5$nm.

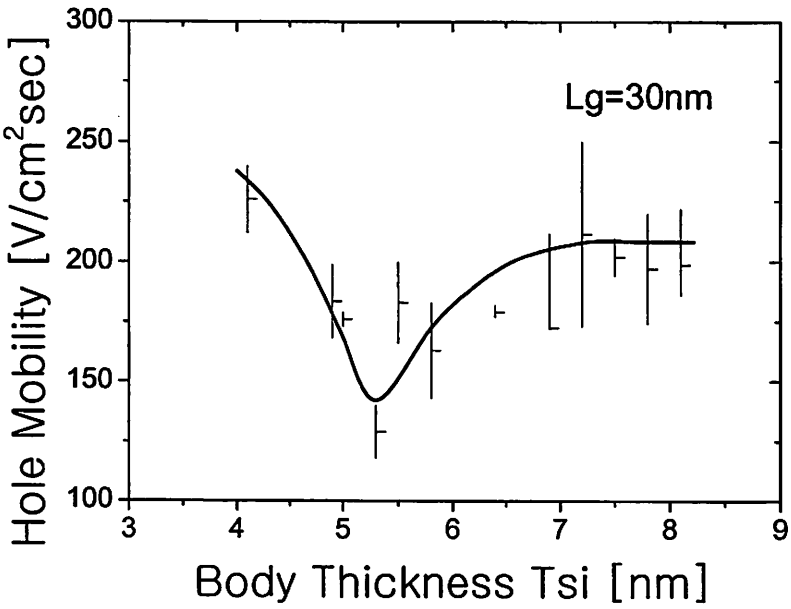


Figure 5. 10 Hole mobility is enhanced when the body thickness is thinner than 5nm.

5.4 Conclusion

The theoretical background has been surveyed and new analytical models to explain the threshold voltage shift and the mobility enhancement in thin body devices have been proposed and successfully demonstrated. These models are based on the trapezoidal potential approximation valid for light body doping and ultra-thin body thickness. To prove the validity of the model, experimental data for NMOS and PMOS UTBFETs were compared to the analytical model. The UTB MOSFETs with lightly doped body show larger threshold voltage shift with decreasing body thickness than with heavily doped body. This is the opposite trend compared with the threshold voltage shift in bulk devices. This difference is explained by a simple analytical model for the first time. Also, the threshold voltage variation due to ultra-thin body thickness variation is more important than that due to dopant fluctuations in the body for a moderately doped and ultra-thin body SOI.

Enhancement in hole mobility is experimentally observed for body thickness below 5nm and explained for the first time.

5.5 References

[5.1] *International Technology Roadmap for Semiconductor*, 2001.

[5.2] H.-S. P. Wong, D. J. Frank, P. M. Solomon, C. H. J. Wann, and J. J. Welser, "Nanoscale CMOS", *Proceeding of the IEEE*, vol.87, p.537-70, 1999.

[5.3] D. Frank, S. Laux, and M. Fischetti, "Monte Carlo Simulation of a 30 nm Dual-gate MOSFET: How far can Si go?," *IEEE IEDM Technical Digest*, p.553-6, 1992.

[5.4] X. Huang, W.-C. Lee, C. Kuo, D. Hisamoto, L. Chang, J. Kedzierski, E. Anderson, H. Takeuchi, Y.-K. Choi, K. Asano, V. Subramanian, T.-J. King, J. Bokor, and C. Hu, "Sub 50 nm FinFET: PMOS," *IEE IEDM Technical Digest*, p.67-70, 1999.

[5.5] Y.-K. Choi, D. Ha, T.-J. King, and C. Hu, "Ultra-thin Body PMOSFETs with Selectively Deposited Ge Source/Drain", *Symposium on VLSI Technology*, p.16-7, 2001.

[5.6] S.-H. Oh, J. M. Hergenrother, T. Nigam, D. Monroe, F. P. Klemens, A. Kornblit, W. M. Mansfield, M. R. Baker, D. L. Barr, F. H. Baumann, K. J. Bolan, T. Boone, N. A. Ciampa, R. A. Cirelli, D. J. Eaglesham, E. J. Ferry, A. T. Fiory, J. Frackoviak, J. P. Garno, H. J. Gossmann, J. L. Grazul, M. L. Green, S. J. Hillenius, R. W. Johnson, R. C. Keller, C. A. King, R. N. Kleiman, J. T.-C. Lee, J. F. Miner, M. D. Morris, C. S. Rafferty, J. M. Rosamilia, K. Short, T.W. Sorsch, A. G. Timko, G. R. Weber, G. D. Wilk, and J. D. Plummer, "50 nm Vertical Replacement Gate (VRG) pMOSFETs," *IEEE IEDM Technical Digest*, p. 65-8, 2000.

[5.7] H. Majima, H. Ishikuro, and T. Hiramoto, "Threshold voltage increase by quantum mechanical narrow channel effect in ultra-narrow MOSFETs," *IEEE IEDM Technical Digest*, p.379-82, 1999.

[5.8] Y. Omura, S. Horiguchi, M. Tabe, and K. Kishi, "Quantum-mechanical effects on the threshold voltage of ultrathin-SOI nMOSFETs," *IEEE Electron Device Letters*, vol.14, p.569-71, 1993.

[5.9] Y. Ma, Z. Li, L. Liu, Z. Yu, "MOS structure threshold voltage model by rigorously considering quantum mechanical effect", *Proceedings of 22nd International Conference on Microelectronics*, vol.1, p.225-8, 2000.

[5.10] T. Janik and B. Majkusiak, "Analysis of the MOS Transistor Based on the Self-Consistent Solution to the Schrodinger and Poisson Equations and on the Local Mobility Model," *IEEE Transactions on Electron Devices*, vol.45, p.1263-71, 1998.

[5.11] F. Stern, "Self-Consistent Results for n-type Si Inversion Layers," *Physical Review B*, vol.5, p.4891-9, 1972.

[5.12] F. Stern, and W. E. Howard, "Properties of Semiconductor Surface Inversion Layers in the Electric Quantum Limit," *Physical Review*, vol.163, p.816-35, 1967.

[5.13] Handbook of Mathematical Functions, edited by M. Abramowitz and I.A. Stegun (U.S. GPO, Washington D.C., 1964), Chap. 10.

[5.14] C. Moglestue, "Self consistent calculation of electron and hole inversion charges at the silicon-silicon dioxide interface", Journal of Applied Physics, vol.59, p.3175-83, 1986.

[5.15] N. W. Ashcroft, N.D. Mermin, Solid State Physics, Saunders College Publishing, 1976.

[5.16] F. Stern, "Electron Exchange Energy in Si Inversion Layers", *Physical Review Letters*, vol.30, p.278-80, 1973.

[5.17] T. Ando, "Density-Functional Calculation of Subband Structure of Semiconductor Surfaces", Surface Science, vol.58, p.128-34, 1976.

[5.18] T. Ando, A.B. Fowler, F. Stern, "Electronic properties of two-dimensional systems", *Review of Modern Physics*, vol.54, p.437-672, 1982.

[5.19] L. Chang, K.J. Yang, Y.-C. Yeo, Y.-K. Choi, T.-J. King, C. Hu, "Reduction of Direct-Tunneling Gate Leakage Current in Double-Gate and Ultra-Thin Body MOSFETs", *IEEE IEDM Technical Digest,* p.99-102, 2001.

[5.20] M. J van Dort, P. H. Woerlee, and A. J. Walker, "A Simple Model for Quantization Effects in Heavily-doped Silicon MOSFET's at Inversion Conditions," *Solid-State Electronics*, vol.37, p.411-4, 1994.

[5.21] M. J. van Dort, P. H. Woerlee, A. J. Walker, C. A. H. Juffermans, and H. Lifka, "Influence of High Substrate Doping Levels on the Threshold Voltage and the Mobility of Deep-Submicrometer MOSFETs," *IEEE Transactions on Electron Devices*, vol.39, p.932-8, 1992.

[5.22] G. Chindalore, S.A. Hareland, S. Jallepalli, A.F. Tasch, C.M. Maziar, V.K.F Chia, S. Smith, "Experimental determination of threshold voltage shifts due to quantum mechanical effects in MOS electron and hole inversion layers", *IEEE Electron Device Letters,* vol.18, p.206-8, 1997.

[5.23] S.A. Hareland, S. Jallepalli, W.-K. Shih, H. Wang; G.L. Chindalore, A.F. Tasch, C.M. Maziar, "A physically-based model for quantization effects in hole inversion layers", *IEEE Transactions on Electron Devices*, vol.45, p.179-86, 1998.

[5.24] C.-Y. Hu, S. Banerjee, "Quantization Effects in inversion layers of PMOSFET's on Si (100) Substrates", *IEEE Electron Device Letters*, vol.17, p.276179-78, 1996.

[5.25] Schred 2.0, Numerical Simulator, *Purdue University*, http://nanohub.purdue.edu

[5.26] G. Dresslhaus, A.F. Kip, C. Kittel, "Cyclotron resonance of electrons and holes in silicon and germanium crystals", *Physical Review*, vol.98, p.368-84, 1955.

[5.27] S. Jallepalli, J. Bude, W.K Shih, M. R. Pinto, C. M. Maziar, and A. F. Tasch, "Electron and hole quantization and their impact on deep submicron silicon p- and n-MOSFET characteristics", *IEEE Transactions on Electron Devices*, vol.44, p.297-303, 1997.

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

[5.29] K. Takeuchi, T. Tatsumi, A. Furukawa, "Channel engineering for the reduction of random-dopant-placement-induced threshold voltage fluctuation", *IEEE IEDM Technical Digest*, p. 841-4, 1997.

[5.30] K. Takeuchi, "Channel Size Dependence of Dopant Induced Threshold Voltage Fluctuation," *Symposium on VLSI Technology,* p.72-3, 1998.

[5.31] T. Mizuno, "Novel Statistical Fluctuations of Dopant Concentration and its Influence on Scaled MOS Device Performance", *International Workshop on Statistical Metrology*, p.16-19, 1997.

[5.32] T. Mizuno, J.I. Okamura, A. Toriumi, "Experimental Study of Threshold Voltage Fluctuation Due to Statistical Variation of Channel Dopant Number in MOSFET's", *IEEE Transactions on Electron Devices*, vol.41, p.2216-21, 1994.

[5.33] S.-I. Takagi, J. Koga, A. Toriumi, "Subband Structure Engineering for Performance Enhancement of Si MOSFETs," IEEE *IEDM Technical Digest,* p.219-22, 1997.

[5.34] M. Shoji, S. Horiguchi, "Phonon-limited inversion layer electron mobility in extremely thin Si layer of silicon-on-insulator metal-oxide-semiconductor field-effect transistor", *Journal of Applied Physics*, vol.82, p.6096-6101, 1997.

[5.35] M. Shoji, S. Horiguchi, "Electronic structures and phonon-limited electron mobility double-gate silicon-on-insulator Si inversion layers", *Journal of Applied Physics*, vol.85, p.2722-31, 1999.

[5.36] Y.-K. Choi, D. Ha, T.-J. Ling, C. Hu, "Threshold Voltage Shift by Quantum Confinement in Ultra-thin Body Device", *59th Device Research Conference*, p.85-6, 2001.

Chapter 6

FinFETs

6.1 Introduction

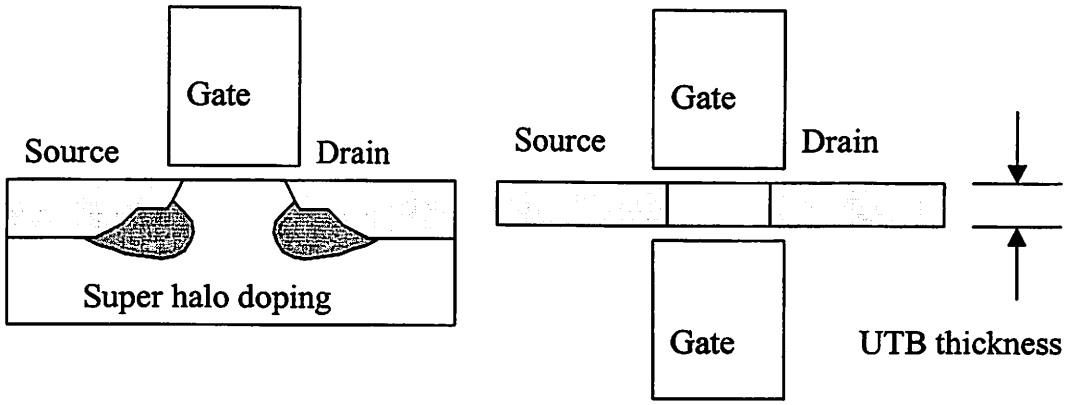
The ultra-thin body (UTB) single gate MOSFET structure is promising for silicon device scaling down to sub-20nm gate lengths [6.1]. As shown in Chapter 4, the device exhibits excellent drive current, well-controlled short-channel effects, and very low off-state leakage current. However, it cannot be scaled down to the ultimate limit of silicon devices because of drain field penetration into the channel through the buried oxide. The double-gate (DG) UTB MOSFET is a more promising candidate for device scaling into the sub-20nm regime [6.2] [6.3].

According to Brew's scaling theory [6.4], the channel doping concentration in bulk MOSFETs should be continuously increased to suppress short-channel effects - up to beyond $10^{18}cm^{-3}$ for gate lengths below 0.1um [6.5][6.6]. Unfortunately, heavy doping concentration (the super halo structure shown in Fig. 6.1 (a)) degrades device performance due to decreased mobility, increased junction capacitance, and increased junction leakage. Conventional bulk CMOS devices require aggressive gate oxide scaling, which increases the gate oxide leakage current, and ultra shallow S/D junctions to suppress short-channel effects. One way to circumvent to the limit of oxide thickness

scaling is to change the gate insulator to another material (high-K dielectric) such that the effective capacitive thickness can be reduced without increasing the tunneling current [6.7]-[6.9]. This technology, however, is not yet mature. Another approach is to change the device structure so that the MOSFET can be scaled further even with a relatively thicker oxide. The double-gate MOSFETs [6.10]-[6.18] (Fig. 6 1 (b)) is one such example. The double-gate device is electrostatically much more robust than a conventional single-gate MOSFET because the gate shields the channel from both sides and reduces short-channel effects. In a conventional CMOS bulk device, the substrate also plays such a role, but it results in a tradeoff between the degree of shielding and the reduction of the subthreshold slope. In the double-gate MOSFET, this tradeoff does not exist, so gates can be strongly coupled to the channel to increase transconductance [6.19]. The relative scaling advantage of the double-gate MOSFET is roughly a factor of two.

As shown in Fig. 6.1 (b) and Fig. 6.2, the double gate device effectively suppresses short-channel effects by using a thin silicon channel that eliminates leakage paths far from the gates. In contrast, bulk devices need extremely shallow source/drain junctions. It is very difficult to achieve junction depths below 10nm. One solution is to use an induced inversion layer as the source/drain. Noda et al. [6.20] reported that short-channel effects could be effectively suppressed down to a gate length of 80nm by this technique. Kawaura et al. [6.21][6.22] reported a dual-gate structure (electrically variable shallow junction MOSFET: EJ-MOSFET) with an another gate that induces source/drain regions. However, this approach will result in a large gate-to-source/drain overlap capacitance and transistor size is not small enough. In a double-gate structure built on a thin silicon body, the junction depth of the source/drain is automatically limited by body

thickness, and this does not require novel shallow junction technology. Thus, the UTB double-gate alleviates the need for ultra-thin gate oxides, heavy channel doping, and shallow junctions.



(a) (b)

Figure 6.1 Schematic diagrams of (a) a conventional bulk MOSFET structure and (b) a double-gate MOSFET with ultra-thin body.

As discussed in Chapter 4, the single-gate UTB MOSFETs can adequately suppress off-state leakage current for gate lengths above *18nm* with a *5nm* body and a *1.5nm* gate oxide thickness. Beyond *18nm*, it is no longer effective. Fig. 6.2 compares a single-gate device with a *5nm* UTB and a double gate device with a *10nm* UTB which are equivalent to the first order because each gate controls 5nm of body thickness. With the aid of 2-D device simulator MEDICI, off-state leakage current is obtained for both structures. However, the double-gate MOSFET controls the channel more effectively and better suppresses off-state leakage current than the single-gate. The reason for this result

is that the electric field from the drain electrode penetrates the buried oxide and affects the channel potential in the single-gate device.

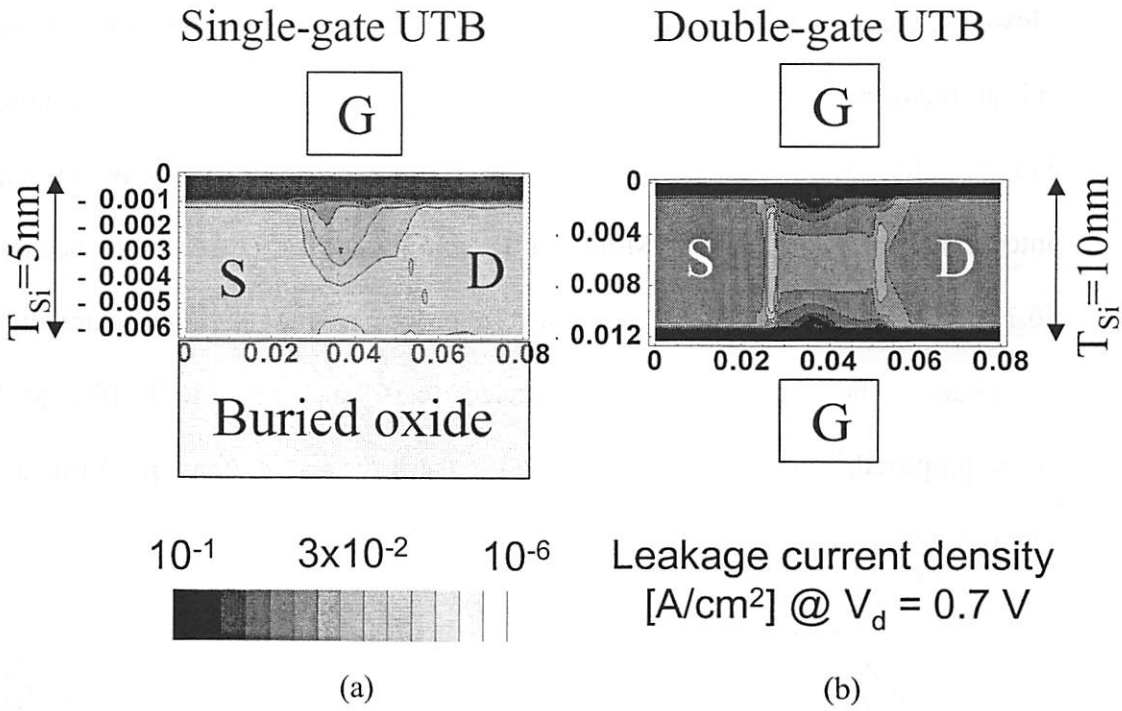


Fig. 6.2 Comparison of leakage current density for (a) single-gate MOSFETs with 5nm UTB and (b) double-gate MOSFETs with 10nm UTB (V_g=0V, L_g=25nm, T_{ox}=1.5nm, N_{body}=$1.0x10^{15}cm^{-3}$).

In the past, many different methods have been proposed and demonstrated [6.10]-[6.18] to fabricate double-gate devices. However, many suffer from process complexity and technical challenges. Fig. 6.3 shows three possible orientations of the double-gate structure. The vertical double-gate structure [6.10][6.11] has a large gate overlap capacitance and provides only one gate length. The horizontal double-gate [6.12][6.13]

also has a large gate overlap capacitance, process complexity, and difficulties even though the top and bottom gate were self-aligned. Wafer bonding or double gate lithography [6.14] has been proposed to create the horizontal double-gate structure. However, these approaches suffer from an inability to align the two gates. To overcome this problem, back side lithography through the transparent substrate was proposed [6.15], but, it could not provide gate lengths below 50nm because of light wave dispersion and interference in the substrate. More exotic self-aligned processes have been proposed [6.12][6.13], but these suffer from severe process complexity and parasitic overlap capacitance. Thus, lateral double-gate devices [6.16] such as FinFETs [6.17][6.18] have been proposed, which show excellent scalability and output performance which circumventing process complexity.

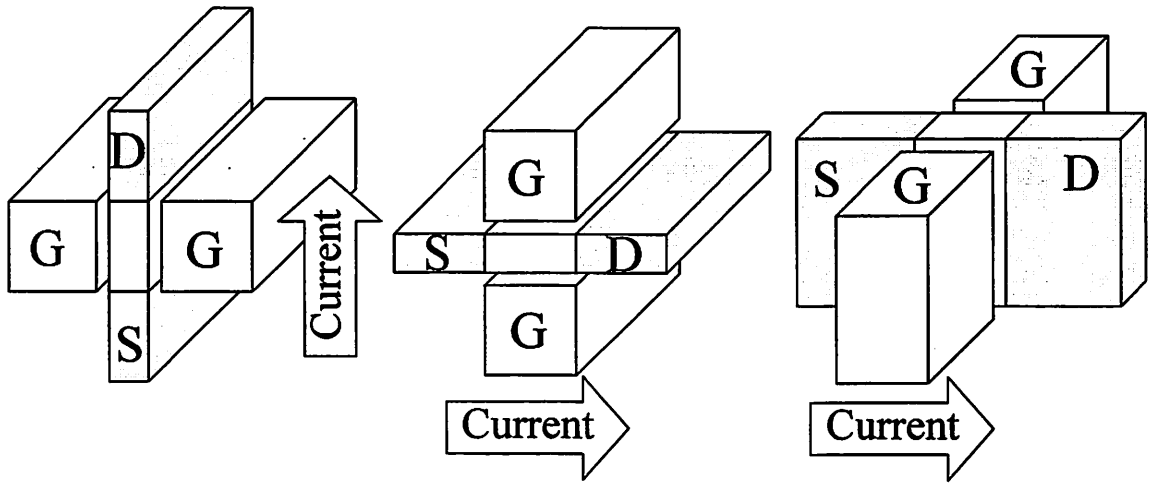


Figure 6.3 Three different double-gate structures. (a) vertical double-gate (b) horizontal double-gate (c) lateral double-gate.

The FinFET uses a single poly-SiGe gate deposited over a silicon fin to form perfectly aligned gates straddling the sidewalls of the fin as shown in Fig. 6.4. The height of the silicon fin (T_{Si} in Fig. 6.4) becomes the transistor width, W. Thus, the total width of the FinFET is twice the fin height ($W=2T_{Si}$). The fin width (W_{fin} in Fig. 6.4) becomes the body thickness that is traditionally associated with the double-gate structure. Current flows along the sidewalls of the fin as shown in Fig. 6.4.

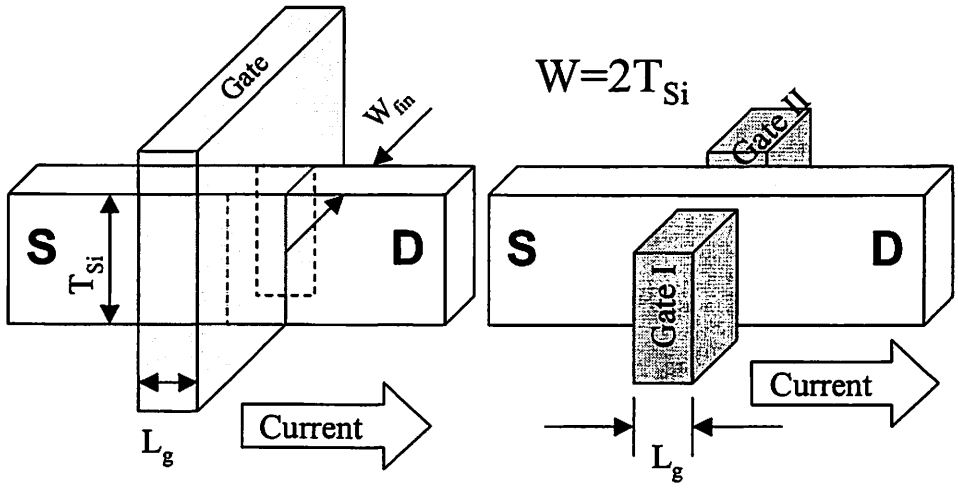
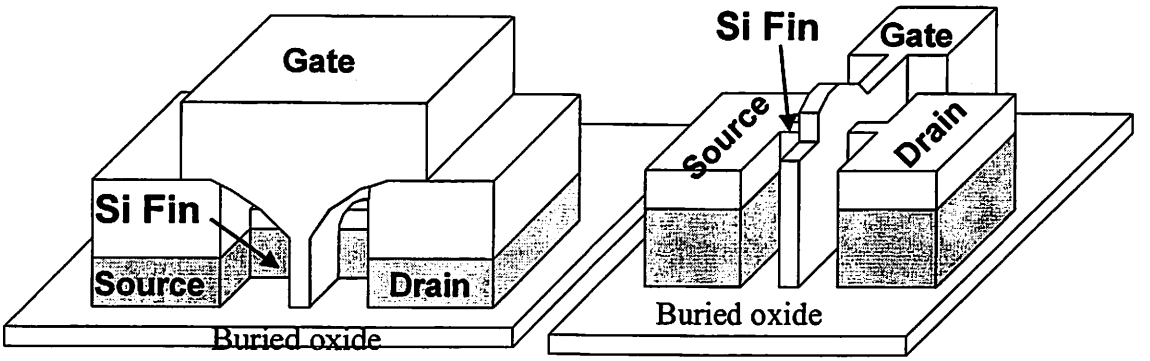


Figure 6.4 3-dimensional view of the FinFET structure.

The original FinFET device [6.17][6.18] and new, improved structure [6.23] are compared in Fig. 6.5. The new FinFET is fabricated by a gate-first, S/D-last process, while the original FinFET was composed of a S/D-first, gate-last process. Therefore, it is more compatible with standard SOI processes. It provides less gate overlap capacitance and its layout is closer to the standard CMOS structure.



(a) (b)

Figure 6.5 3-dimensional view of (a) original FinFET device [6.17][6.18] and (b) new FinFET device [6.23].

The fin width (W_{fin} in Fig. 6.4) is the most important process variable because it governs off-state leakage current and alleviates short-channel effects [6.2][6.3][6.17][6.18][6.23][6.24][6.25] as the ultra-thin body does in the single–gate UTBFETs [6.1]. Channel mobility can also be sensitive to this fin width [6.26]. Fig. 6.6 shows that the leakage current density increases dramatically as the body thickness is increased because gate control of the channel is worsened. The simulation shows that a 10nm reduction of W_{fin} lowers the off-state leakage current up to 1000 times. Channel mobility [6.26], threshold voltage [6.27]-[6.29], and subthreshold characteristics can be sensitive to the fin width [6.30].

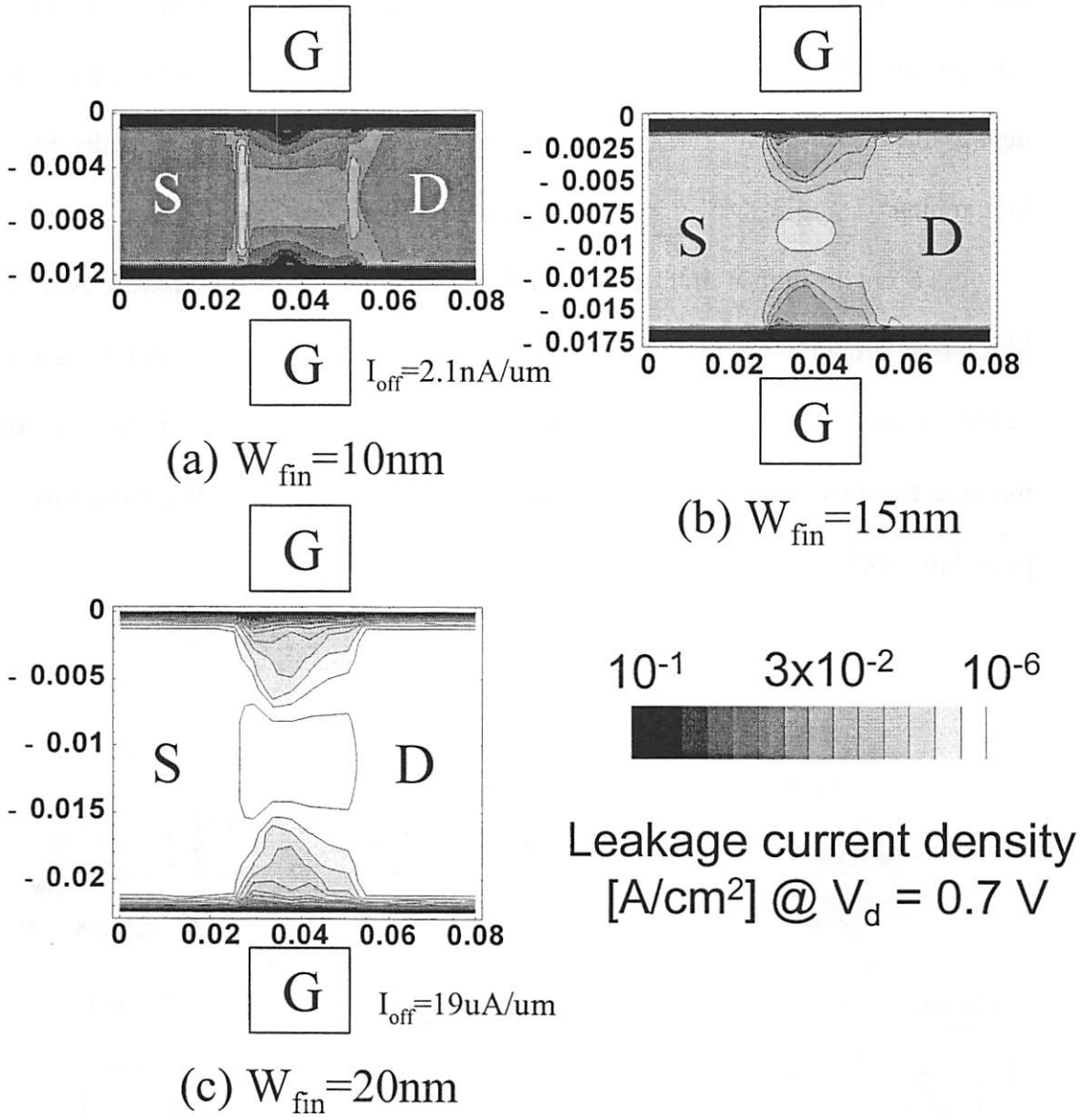


Figure 6.6 Narrowing the fin width eliminates paths far from the gate, thus improving short-channel effects and allowing for reduced gate lengths. (V_g=0V, L_g=25nm, T_{ox}=1.5nm, N_{body}=1.0x10^{15}cm^{-3}) [2].

Definition of the fin width using optical or e-beam lithography inevitably leads to CD variation. A spacer lithography technology is thus proposed to overcome this fin

width variation. In practice, it shows CD uniformity for better than that of e-beam lithography [6.31] (Fig. 2.10) because the fin is defined not by the lithography but by the deposited CVD film thickness. CVD film thickness uniformity is generally better than lithography CD uniformity in standard silicon processing.

To obtain higher drive current, a large channel width is required. Although the FinFET is a double-gate structure, it is similar to a conventional planar MOSFET in layout as shown in Fig. 6.7. The only difference is that the active layer consists of multiple fins instead of a single box. To make a large channel width, more fins must be placed in parallel.

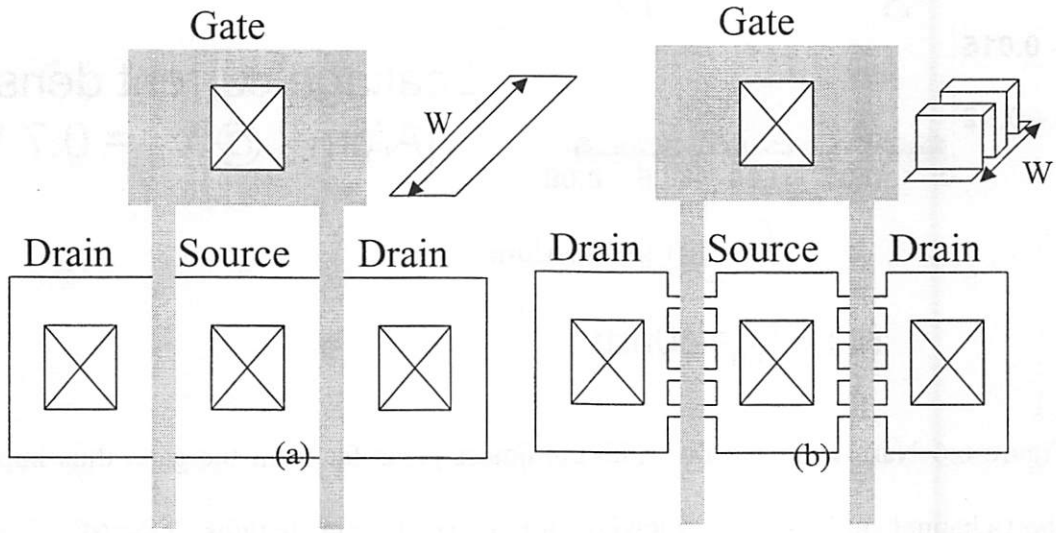


Figure 6.7. Typical layout of (a) conventional MOSFET and (b) FinFET

A layout efficiency factor γ can be defined as :

$$\gamma = \frac{W_{FinFET}}{W_{Conventional}} = \frac{2T_{Si} \cdot N}{(W_{fin} + S) \cdot N} = \frac{2T_{Si}}{(W_{fin} + S)} = \frac{2T_{Si}}{Pitch}, \tag{6.1}$$

where W_{FinFET} is the total channel width of the FinFET, $W_{Conventional}$ is the total channel width of a conventional MOSFET, T_{Si} is the fin height, W_{fin} is the fin width, S is the fin-to-fin space, and N is the total number of fins as shown in Fig. 6.8.

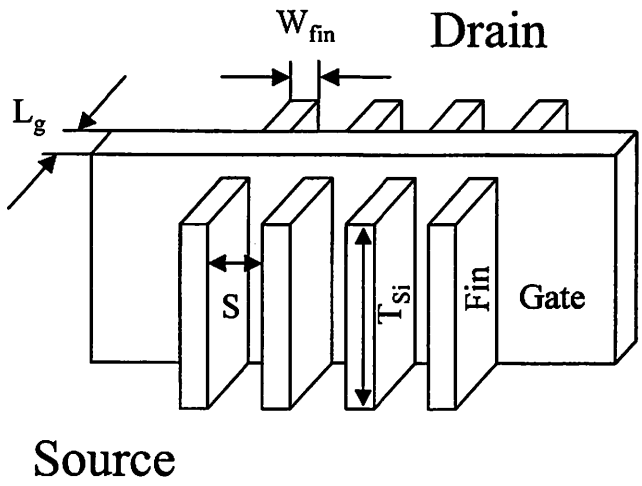


Figure 6.8 3-dimensional view of FinFET with multiple fins.

To match the layout efficiency of a planar double-gate device, γ should be at least two. This means that the fin height (T_{Si}) should be larger than the pitch in Eq. (6.1). However, the fin pitch is limited by lithography whereas the fin height is limited by process capability. A taller fin results in a larger step height and this reduces the lithography depth-of-focus (DOF) and causes resist notching. It also narrows the plasma etch window such that stringers or residues may be left behind (Chapter 3).

A second benefit of spacer lithography technology is that it yields twice the pattern density and hence provides twice the drive current. A FinFET defined by spacer

lithography technology is thus named "Spacer FinFET". A CMP process is proposed to increase the process window (as in the chapter 3), and device results are discussed.

As the fin width is reduced, the series resistance increases. If the gate is misaligned to the source and drain pads, the output characteristics can be changed significantly. An intentionally designed offset between the gate and the source allows for experimental investigation of the extent of drive current reduction due to parasitic resistance. One way to overcome this problem is to adopt a self-aligned contact (SAC) process for making metal contact to the source and drain pads. However, process complexity renders such a technique difficult. Thus, a selective Ge deposition process is proposed [6.32] to alleviate the effect of misalignment and to reduce the series resistance.

One drawback of the FinFET is the rough surface of the channel, which is caused by line edge roughness in the resist during lithography and plasma etch damage (in Fig. 6.9). To alleviate etch damage at the sidewall of fin, argon annealing was applied after a sacrificial oxidation step to smooth the surface.

In this work, the new FinFET structure is fabricated using two technologies: e-beam lithography and spacer lithography. By using e-beam lithography and subsequent ashing-trimming, the world record smallest double-gate CMOS device (L_g=*sub-20nm* and W_{fin}=*10nm*) is demonstrated. The device characteristics of both structures are discussed.

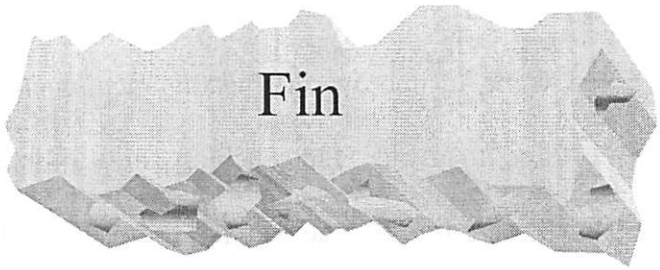


Figure 6.9 The silicon fin surface is rough due to lithography and plasma etch. The uneven profile along the fin is caused by a line edge roughness of the resist in lithography while further roughness is caused by etch damage.

6.2 FinFET by E-Beam Lithography

6.2.1 Process Details

The process flow used here is similar to that in previous work, in which quasi-planar FinFETs were fabricated with dual resist e-beam lithography [6.23]. (100) SOI wafers with initial film thickness of 100nm Si and 400nm buried oxide are thermally oxidized to thin the silicon film down to 50nm. 50nm LTO is deposited over this film to serve as a hard mask oxide. 400nm undoped $Si_{0.5}Ge_{0.5}$ and 100nm LTO are the deposited and patterned to form the zero-level-alignment key for subsequent e-beam and optical lithography. The role of the LTO is to serve as a hard mask to protect that alignment key during silicon fin and gate poly-silicon etch. Next, phosphorus is implanted at $1.0x10^{13}cm^{-2}$, 30KeV, and 0^{o} to dope the channel. 200nm SAL-601 resist is then coated for e-beam lithography. In previous work, field stitching was used to pattern a large die

area. However, this resulted in large die-to-die misalignment. Furthermore, a dual–resist process [6.23][6.33] was used to define narrow fins and large-area patterns simultaneously. In this work, to improve the misalignment accuracy, a small field size (500um x 500um) is used to avoid field stitching. In addition, single-resist e-beam lithography is used, thus requiring that the wafers be metallized to form probe pads. A 60nm initial line width (drawn size of 48nm) could be reduced down to below 10nm with ashing-trimming technology [6.34]. Ashing for SAL-601 is performed at 3W for 20sec with O_2 plasma to reduce a 60nm line width to 20nm. The plasma etch recipe to remove the hard mask oxide is 100sccm of CF_4, 13mTorr of pressure, 200W of top RF power, and 40W of bottom RF power, yielding an etch rate of 120nm/min for thermally grown oxide. After stripping the resist, the hard mask oxide is trimmed down from 20nm to 10nm with (100:1) HF. The plasma etch recipe used to remove 50nm silicon is 50sccm of Cl_2, 150sccm of HBr, 15mTorr of pressure, 300W of RF top power, and 150W of RF bottom power, which gives an etch rate of 550nm/min. As mentioned in Chapter 2, this recipe is desirable because it produces a vertical etch profile. Fig. 6.10 shows a schematic structure after silicon fin etch and Fig. 6.11 shows an SEM photograph of a sub-10nm width fin.

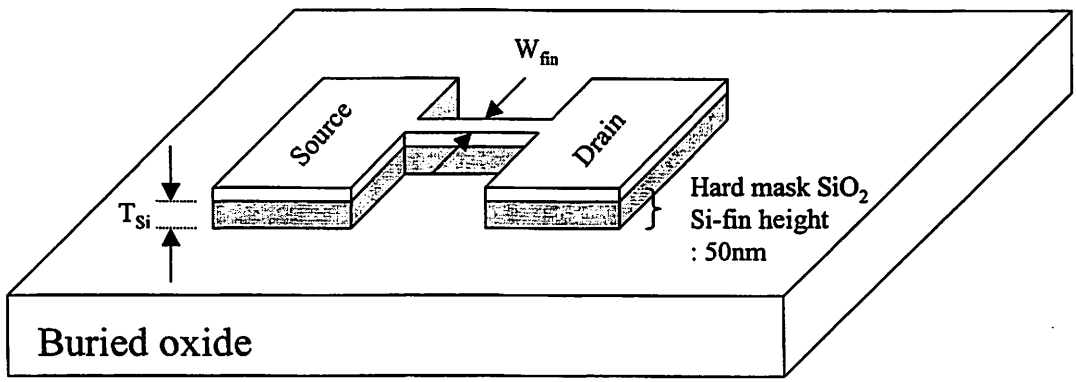


Figure 6.10 FinFET structure after silicon fin etch.

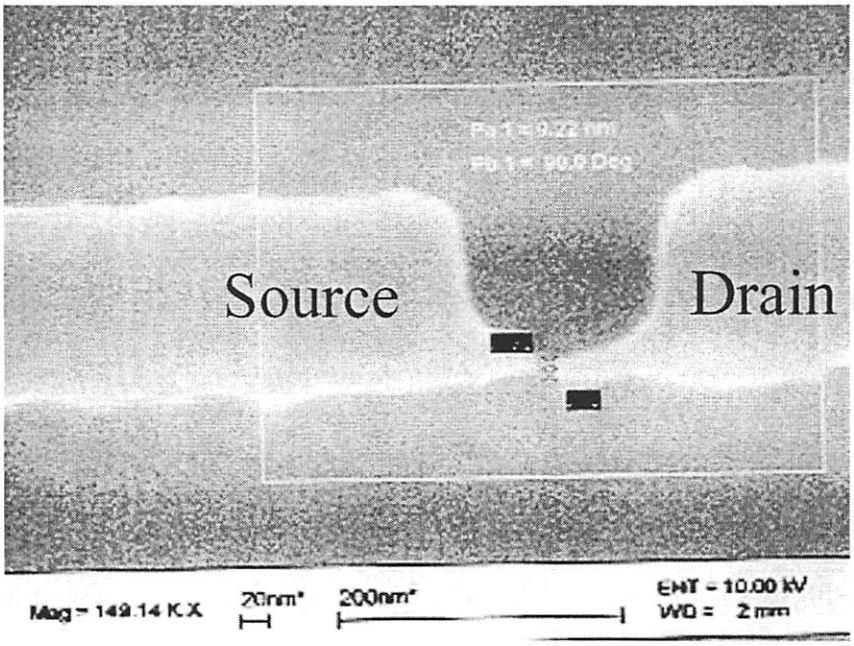


Figure 6.11 SEM top view of a sub-10nm width silicon fin.

After the silicon fins are etched, 3nm sacrificial oxide is thermally grown at 900°C for 3min and removed without seriously undercutting the buried oxide to improve the fin sidewall surface prior to gate oxidation. The gate oxide is grown at 750°C for 12min to yield a thickness of 2.1nm. In-situ boron-doped $Si_{0.6}Ge_{0.4}$ is deposited with a thickness of 240nm as the gate material, which can be used to modulate the threshold voltage by changing the germanium mole fraction [6.35]-[6.37]. A 700nm LTO hard mask oxide is deposited and densified at 900°C for 30min. The purpose of LTO densification is to minimize hard mask oxide recess and to make sure it remained on the top surface of the gate after gate spacer LTO and nitride overetch to allow for a raised S/D process. It may, however, cause boron penetration from the p-type $Si_{0.6}Ge_{0.4}$ into the channel. E-beam

lithography is then performed with 200nm SAL-601 resist. The initial line width is 80nm (drawn size of 56nm) after lithography. Similarly to the fin level, the 80nm line width is reduced down to 20nm with ashing-trimming technology. SAL-601 resist ashing is done at 3W for 25sec with O_2 plasma, which reduces the line width from 80nm to 30nm. The plasma etch recipe of the gate hard mask oxide is 100sccm of CF_4, 13mTorr of pressure, 200W of top RF power, and 40W of bottom RF power (etch rate of 120nm/min for thermally grown oxide). After resist strip, oxide trimming reduces the line width from 30nm to 20nm using (100:1) HF. An extra benefit of hard mask trimming is that stringers along the fin step height are eliminated, which can leave gate poly-$Si_{0.6}Ge_{0.4}$ stringers along the fin as shown in Fig. 3.16. The gate poly-$Si_{0.6}Ge_{0.4}$ is etched using a two-part etch. The main etch provides a vertical etching profile but relatively low selectivity (poly-$Si_{0.6}Ge_{0.4}$: Oxide=20 : 1). Conditions used are 50sccm of Cl_2, 150sccm of HBr, 15mTorr of pressure, 300W of RF top power, and 150W of RF bottom power (etch rate of 1.1um/min). The overetch provided a sloped etch profile, but higher selectivity (poly-$Si_{0.6}Ge_{0.4}$: Oxide=400 : 1). It also formed a notched T-shaped gate as shown in Fig. 4.11 (a) and Fig. 4.13 (a). As a result, the actual gate length is believed to be shorter than 15nm. Figure 6.12 shows the schematic structure of the fin and gate after gate poly-$Si_{0.6}Ge_{0.4}$ etch. Fig. 6. 13 shows a SEM photograph of the 10nm fin width and 20nm top-gate length.

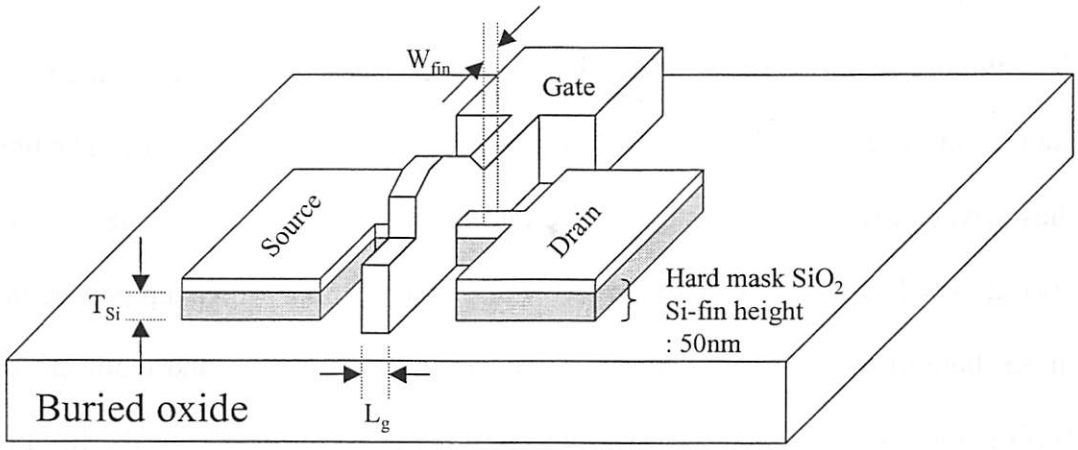


Figure 6.12 FinFET structure using e-beam lithography

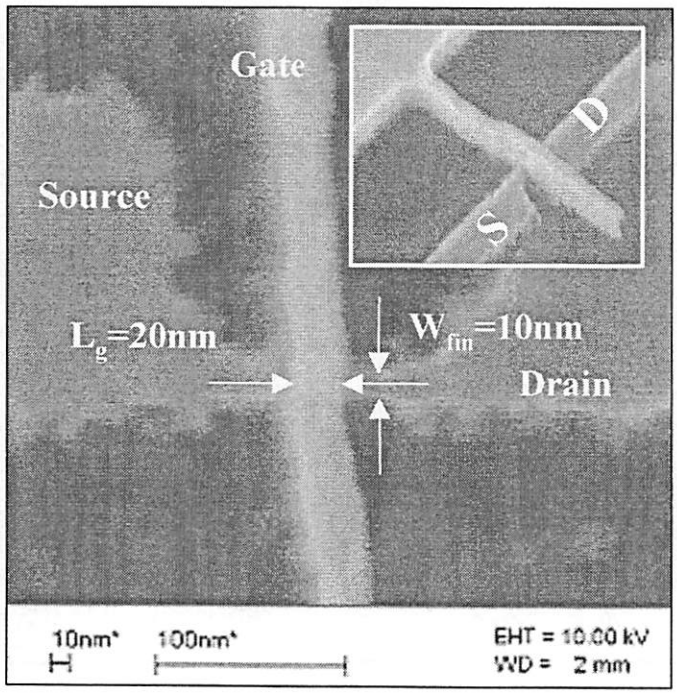


Figure 6.13 SEM top view of a FinFET with 10nm fin width and 20nm top gate length.

After gate formation, a bi-layer spacer of 20nm nitride on 10nm HTO is deposited. The purpose of this bi-layer spacer is to provide a wide etch process window. Due to equipment limitations, the selectivity of the oxide to silicon was only 3:1. The thin HTO thus serves only as an etch-stop for the nitride plasma etch. The nitride etch recipe is 50sccm of Cl_2, 150sccm of HBr, 15mTorr of pressure, 150W of RF top power, and 75W of RF bottom power, which gives a selectivity between nitride and oxide of 3:1. The HTO is then removed with HF, thus avoiding plasma etch damage to the fin. Fig. 6.14 shows the bi-layer spacer profile.

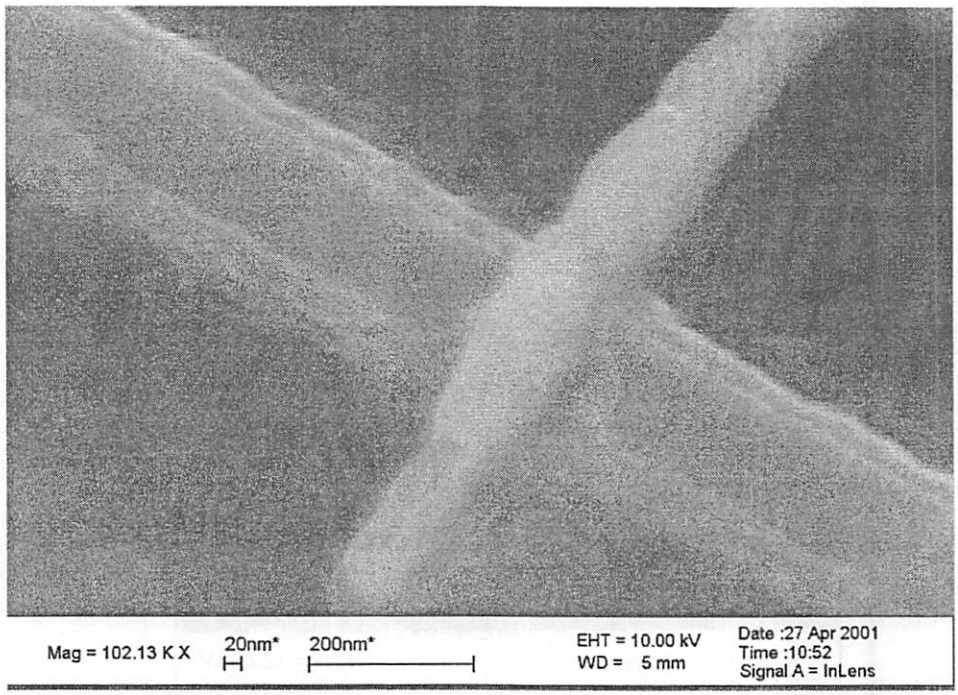


Figure 6.14 Bi-layer spacer profile after nitride plasma and HF wet etch.

The bi-layer spacer thickness is chosen to provide proper separation of the S/D implant profile from the gate. Using N+ and P+ implantation masks, phosphorus is implanted at $5.0x10^{15}cm^{-2}$, 30KeV, and 0^{o} while boron is implanted at $5.0x10^{15}cm^{-2}$, 10KeV, and 0^{o}. 600°C for 15 hours furnace anneal is used to recrystallize any portion of the silicon fin that might have been amorphized by the heavy dose implantation. S/D dopants are activated by RTA at 900°C for 1min in N_2 ambient. 50nm of selective Ge is deposited on the narrow fin to reduce the series resistance for one wafer (NMOS only) after RTA. Fig. 6.15 depicts the device structure after spacer formation and selective Ge deposition. A cross-sectional TEM and top view SEM of the selectively deposited Ge are shown in Fig. 3.7 and Fig. 3.8. The detailed process conditions for selective Ge deposition was described in Chapter 3.

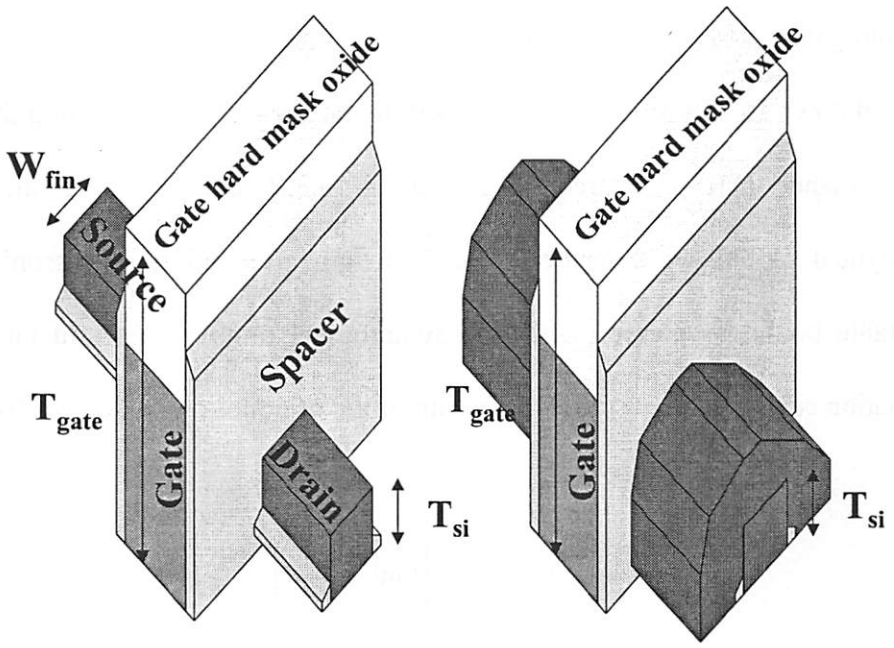


Figure 6.15 Schematic structures after spacer formation and selective Ge deposition.

After selective Ge deposition, a 30nm capping LTO is deposited. A second phosphorus implantation is performed at $5.0x10^{15}cm^{-2}$, 45KeV, and 0° and a second RTA is used at 750°C for 1min in N_2 ambient. Then, 150nm LTO is deposited to shield the gate from the metal interconnection lines. After metal contact hole masking, the LTO is removed with 20mTorr of pressure, 450W of top RF power, 750W of bottom RF power, 30sccm of CHF_3, and 200sccm of Ar. HF is used to remove the remaining oxide as an overetch after resist strip. Aluminum is then sputtered and patterned. Sintering is performed at 400°C for 30min in (10:1) N_2:H_2 ambient.

6.2.2 Device Characteristics of E-Beam FinFETs

In a FinFET device, short-channel effects can be suppressed when $W_{fin} < 0.7L_g$ for a 2nm gate oxide thickness [6.17]. Device simulations show that the gate length can be scaled down to 10nm with a 5nm fin width and 1nm $T_{ox,eq}$ [6.2] using the criteria of high performance ITRS I_{off} target. Frank et al. [6.38] and Suzuki et al. [6.39] proposed analytical models to describe the scale length after solving Poisson's equations with suitable boundary conditions in the symmetrical double gate structure. The following equation can be used to study the scaling limits of double gate devices [6.38],

$$1 = \frac{\varepsilon_{Si}}{\varepsilon_I} \tan\left(\frac{\pi T_{ox}}{\Lambda_1}\right) \tan\left(\frac{\pi W_{fin}}{2\Lambda_1}\right), \qquad (6.2)$$

where Λ_1 is scale length. The minimum channel length occurs when L_g/Λ_1 =1.5~2 depending on the chosen criteria for L_{min} [6.38][6.40]-[6.42]. Fig. 6.16 plots Eq. (6.2) at L_g/Λ_1 = 1.5 to serve as a guideline to determine the gate oxide thickness and the fin width

requirement for a given gate length. Filled circles in Fig. 6.16 represent working FinFET devices. These are a little bit far from the predicted model. Better short-channel effects than the prediction may come from the omega (Ω) shaped gate (in Fig 6.30), which suppresses more effectively the electric field from the drain.

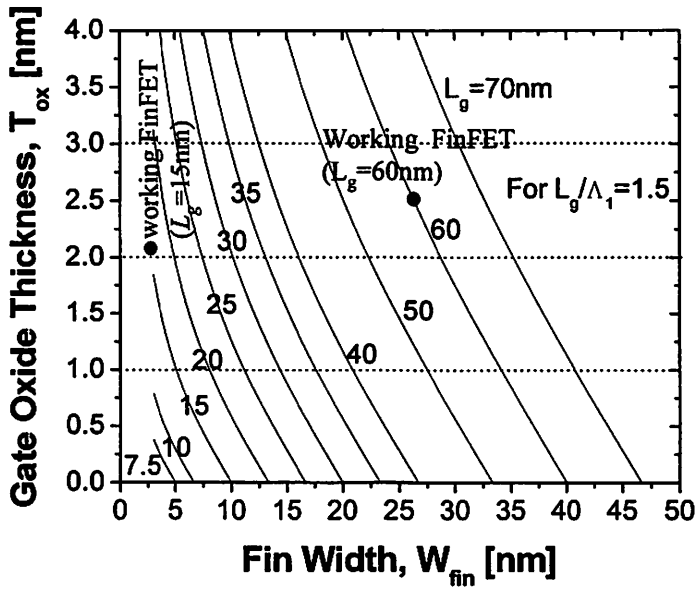
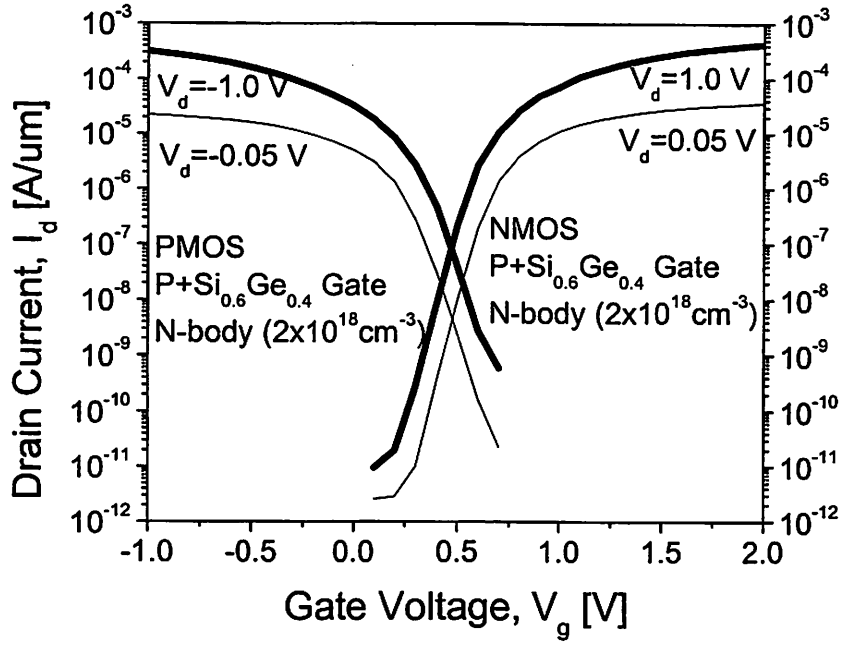
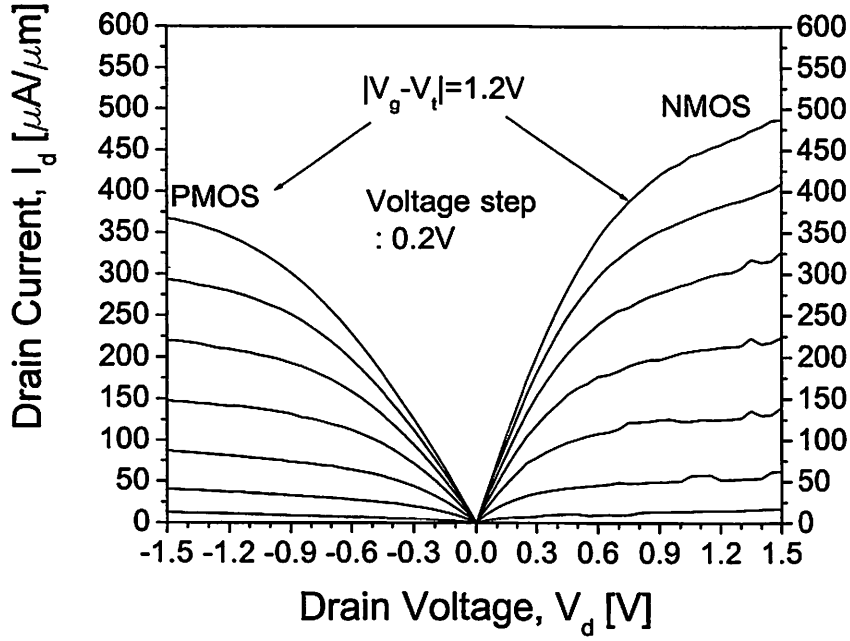


Figure 6.16 Contour plot of gate length versus a gate oxide thickness and silicon body thickness (W_{fin} in FinFET) for double gate MOSFETs (Adapted from [6.38]).

Typical measured I-V characteristics are shown in Fig. 6.17 for sub-20nm gate length and 10nm fin width with a 2.1nm gate oxide thickness. NMOS drive current is 365uA/um and PMOS drive current is 270uA/um at $|V_g-V_t|=|V_d|=1V$. The threshold voltage is defined at 200nA/um of drain current. All currents are normalized with the conservative definition of channel width: $W=2*T_{Si}$ in Fig. 6.12. Using an alternative channel width definition ($W=T_{Si}$ in Fig. 6.12), the NMOS drive current is 730uA/um and the PMOS current is 540uA/um.



(a)



(b)

Figure 6.17 Measured I-V characteristics. (a) I_d-V_g and (b) I_d-V_d characteristics for L_g=20nm, W_{fin}=10nm, p+ $Si_{0.6}Ge_{0.4}$ gate, T_{ox}=2.1nm, and N_{body}=$2.0x10^{18}cm^{-3}$ (n-type).

The off-state current is 70nA/um at the intersection of NMOS and PMOS current and $|V_d|=1.0V$.

Threshold voltage (V_t) roll-off characteristics are shown in Fig. 6.18. As expected, a narrower fin width shows better V_t roll-off. The subthreshold and DIBL (Drain Induced Barrier Lowering) are also improved as the fin width reduces as shown in Fig. 6.19 and Fig. 6.20. Worse short-channel effects in PMOS devices are observed because the effective channel length is shorter. This is due to the increased diffusivity of boron in the heavily dopant P+ S/D junction as compared to that of phosphorus in the heavily doped N+ S/D junction.

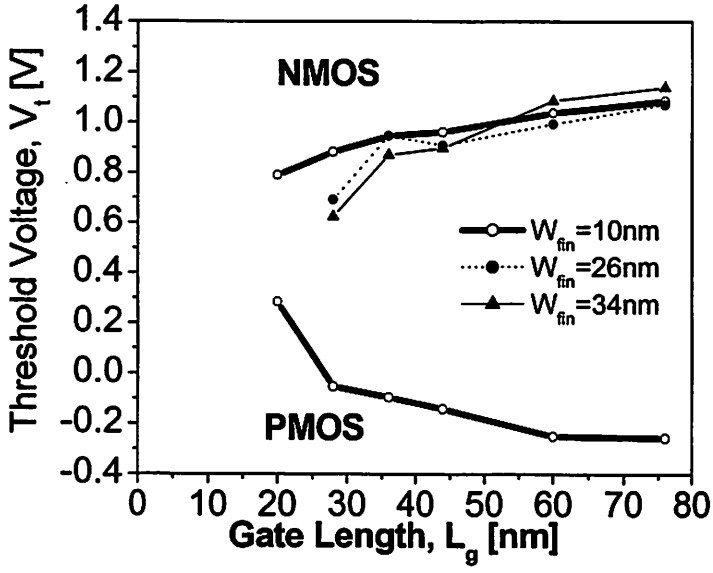


Figure 6. 18 Threshold voltage roll-off characteristics. A narrower fin width shows better roll-off behavior.

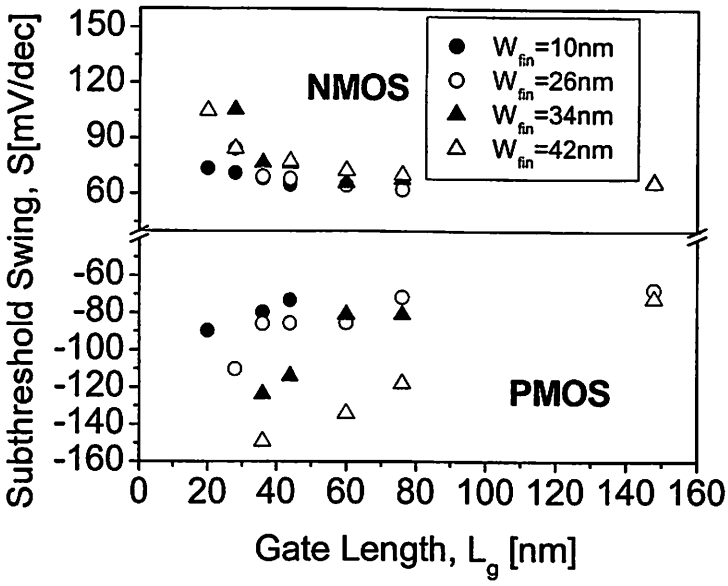


Figure 6. 19 Subthreshold swing versus gate length. A narrower fin width shows lower subthreshold swing.

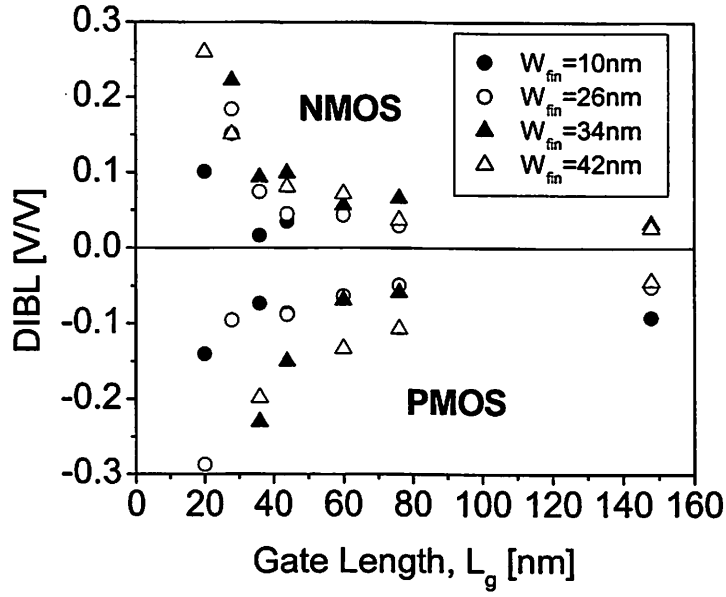


Figure 6. 20 DIBL versus gate length. A narrower fin width shows lower DIBL.

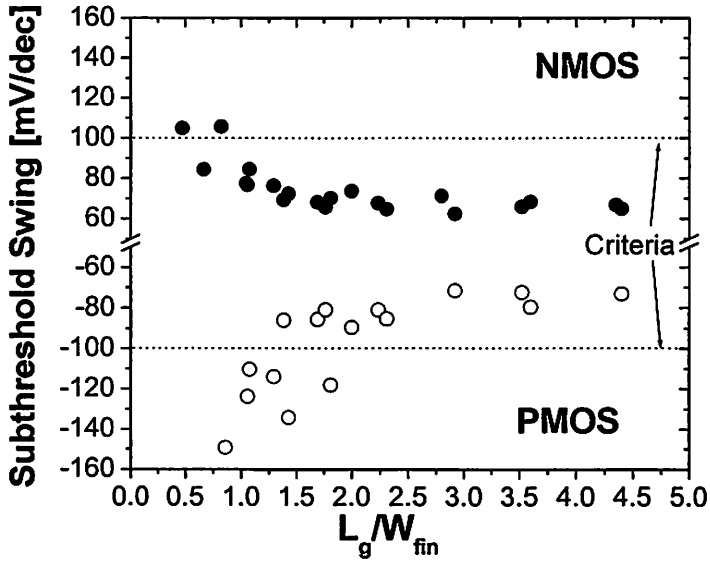


Figure 6. 21 Subthreshold swing versus L_g/W_{fin} for $T_{ox}=2.1nm$. When L_g/W_{fin} is larger than 1.5, both the NMOS and PMOS satisfy the criteria of subthrshold swing ≤ ± 100mV/dec.

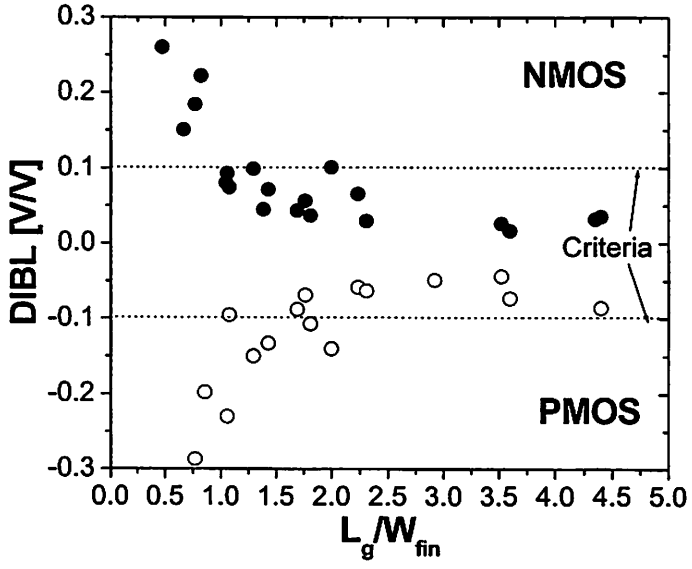


Figure 6. 22 DIBL versus L_g/W_{fin} for $T_{ox}=2.1nm$. When L_g/W_{fin} is larger than 1.5, both the NMOS and PMOS satisfy the criteria of the DIBL ≤ ± 0.1V/V.

Fig. 6. 21 and Fig. 6.22 are plotted using the data from Fig. 6.19 and Fig. 6.20. The subthreshold swing and DIBL satisfy standard criteria, ± 100mV/dec and ± 0.1V/V respectively for NMOS and PMOS when L_g/W_{fin} is larger than 1.5 with T_{ox}=2.1nm. These data reproduce the same short-channel behavior as in the previous work [6.17][6.23].

By utilizing selective Ge deposition with LPCVD, a raised S/D structure is formed on the narrow fins, which reduces the series resistance. Fig. 6.23 shows that the drain currant is enhanced by 28% after selective Ge deposition.

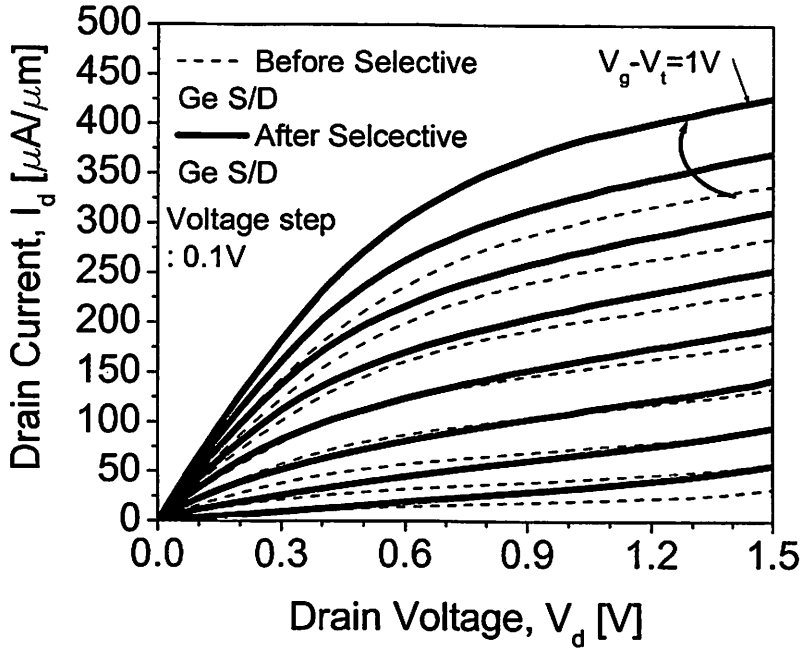


Figure 6.23 Drain current comparison before and after selective Ge deposition. The drain current is increased by 28% after Ge deposition. Data is measured on the same transistor (L_g=*90nm*/W_{fin}=*70nm*).

Mobility is extracted by using large size transistors, which are composed of 150 fins, 26nm fin width (drawn size of 64nm), and 14um gate length. Effective field can be defined as:

$$\mathcal{E} = \eta Q_{inv} / \varepsilon_{Si} = \eta \int C_{ox} dV_g / \varepsilon_{Si}, \quad (6.3)$$

where the inversion charges density is obtained by integrating the measured gate capacitance over the applied gate voltage. η is the fitting parameter and the best values for bulk MOSFETs are $\eta=1/2$ for electrons and $\eta=1/3$ for holes [6.63]. In the double-gate devices, η is predicted to be smaller than that of bulk MOSFET because of the lower electric field. Fig. 6.24 shows that hole mobility, which is close to universal curve. The electron mobility of UTBFETs in Fig. 4.27 is very close to the universal mobility and NMOS drive current is comparable to bulk CMOS's. However, NMOS drive current of FinFETs compared with bulk CMOS is relatively smaller. It can be deduced that the electron mobility is more degraded in an etched surface. Petti et al. [6.43] reported significant electron mobility degradation and an insensitivity of hole mobility to the sidewall surface channel, which was formed by plasma etch. This is believed to be caused by sidewall roughness. There are many reports that more accurately describe mobility dependence on surface roughness scattering [6.44]-[6.47] using the correlation length in Fig. 5.9. However, the phenomenon can be easily understood using a simple argument [6.48]. If the interface in a certain region of the channel is shifted with respect to its average position by a quantity Δ, the average potential felt by the centroid of the carriers

will change by $E_{eff}\Delta$ as shown in Fig. 5.9. It is well known that a potential change gives rise to scattering and consequently acts as a perturbation during carrier transport. The scattering rate is proportional to the square of the perturbing potential according to Fermi golden rule. Thus, mobility due to surface roughness is expressed as:

$$\mu_{SR} \propto (E_{eff}\Delta)^{-2} \tag{6.4}$$

Eq. (6.4) shows that mobility degradation due to surface roughness is large at high effective field even with little roughness (Δ). Petti's experimental results [4.43] showed that electron mobility rapidly decreased at the high effective field while hole mobility still followed the universal curve. This is likely because the inversion charge centroid of electrons is closer to the silicon-dielectric interface than that of holes since the heavy electron effective mass results in a strong quantum confinement effect as discussed in Chapter 5.

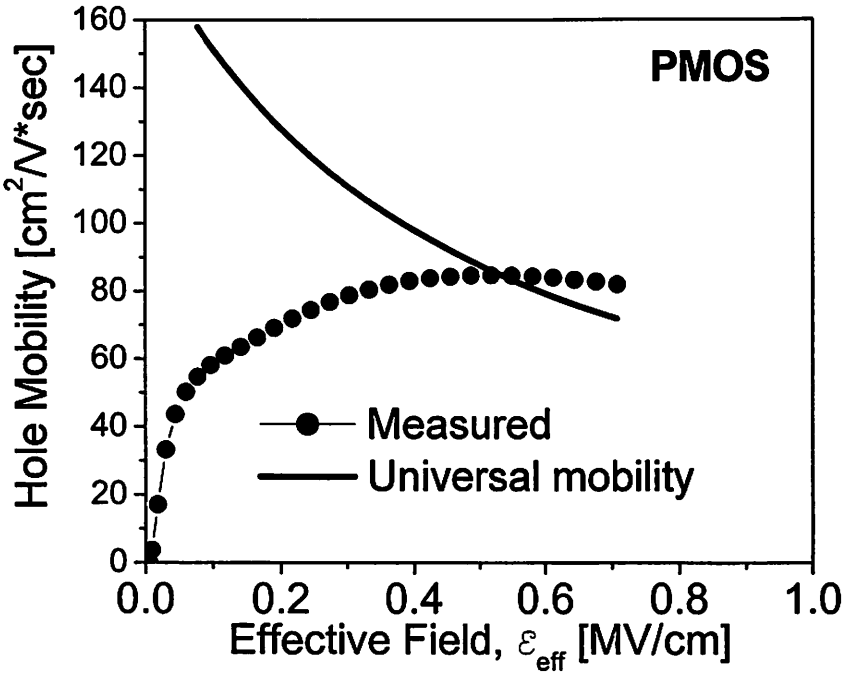


Figure 6.24 Measured hole mobility versus universal mobility (number of fins=150, W_{fin}=26nm, T_{Si}=*50nm*, L_g=*14um*, T_{ox}=*2.1nm*, and N_{body}=$2.0x10^{18}cm^{-3}$). η=*1/4* is used for holes.

The inversion charge centroid for electrons and holes in the FinFET structure is obtained to verify the above arguments with a numerical simulator [6.49] as shown in Fig. 6.25. As predicted, the average inversion charge centroid is closer to the channel surface for electrons.

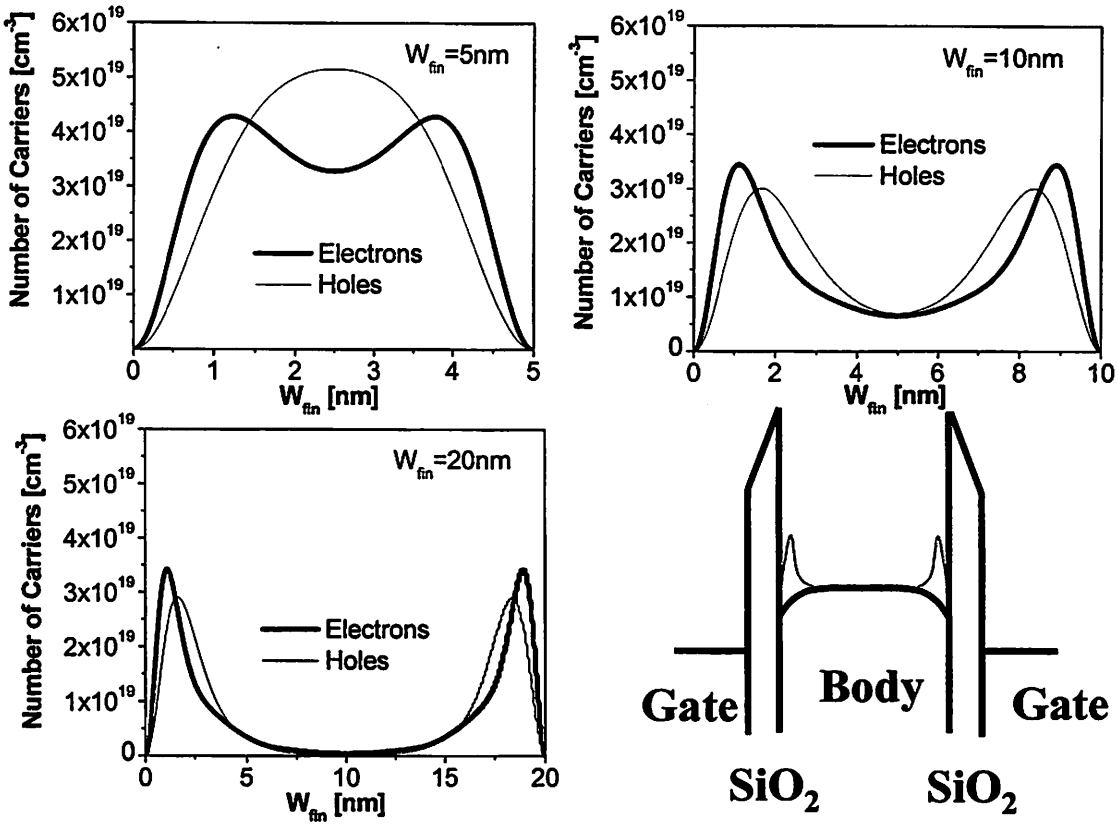


Figure 6.25. Inversion charge centroid versus fin width and energy band diagram of double-gate FinFETs. Data is obtained using Schred [6.49] (T_{ox}=*1nm*, N_{body}=$2.0x10^{18}cm^{-3}$, and V_g=*0.8V*).

The rough surface is observed in the cross-sectional TEM photographs of fin in Fig.6.26. The top surface formed by thermal oxidation is smooth (no contrast difference) whereas sidewall surface formed by plasma etch is rough (contrast difference).

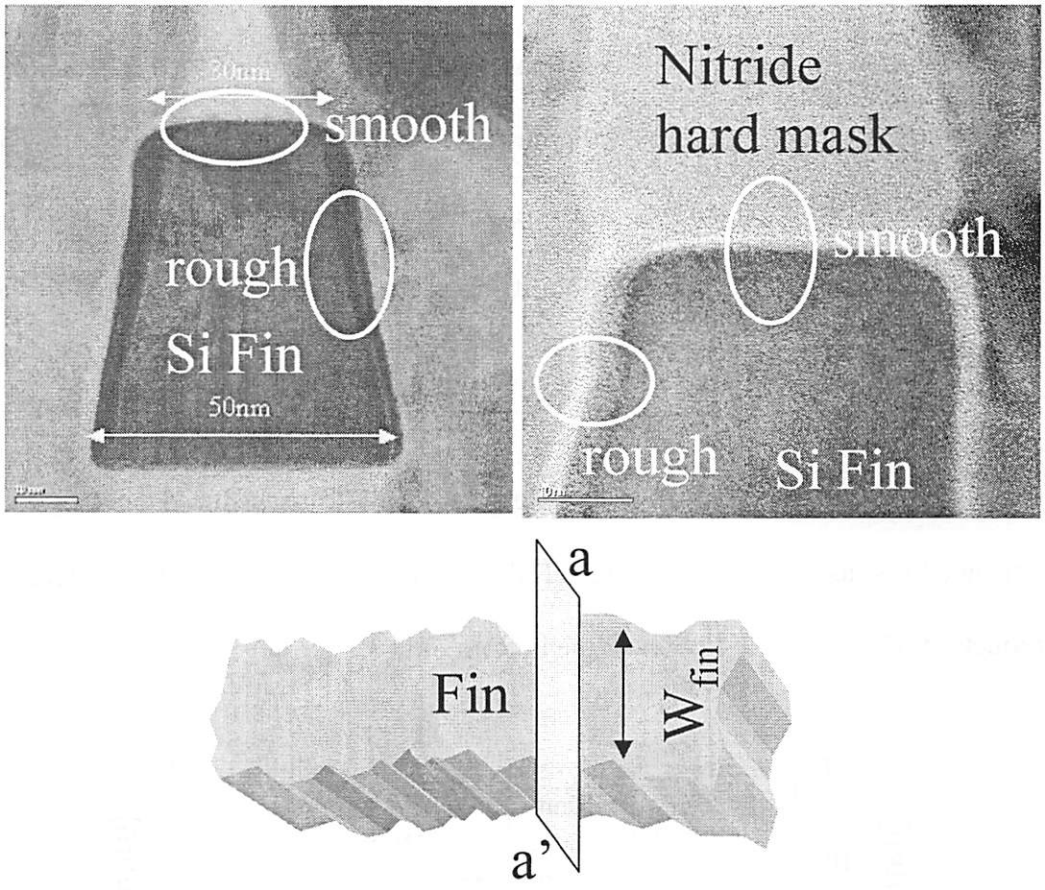


Figure 6. 26 Cross-sectional TEM photographs of a fin along the a-a' direction. The top surface is smooth because the film is thermally oxidized and the sidewall surface is rough because it is formed by lithography and plasma etch.

Stress can also affect the mobility of electrons and holes [6.50]-[6.55]. Normally, this stress arises during the local oxidation of silicon (LOCOS) process or shallow trench isolation (STI) process. Huang et al. [6.54] reported that transistors isolated by a silicon mesa etch showed the similar mobility behavior to standard bulk silicon devices while transistors isolated by a modified LOCOS process showed different mobility behavior. Under tensile stress, electron and hole mobility were both enhanced [6.50][6.55]. Under compressive stress, electron mobility was decreased [6.52][6.54][6.55] while hole

mobility was increased [6.50][6.52][6.54] or decreased [6.55]. However, this stress effect should be more important in a horizontal structure such as the UTBFET than in a vertical structure such as the FinFET. Furthermore, Ref. [6.54] showed that mobility change was independent of the applied gate voltage, which was not observed in [6.43]. Thus, in the FinFET, degradation of electron mobility is likely more strongly affected by the surface roughness than by stress.

Gate leakage current in double-gate FinFETs is reduced as the fin width is narrowed just as in the single-gate UTBFETs (in Fig. 4.26) [6.56][6.57]. This is due to a reduction of the electric field near the bottom of inversion layer.

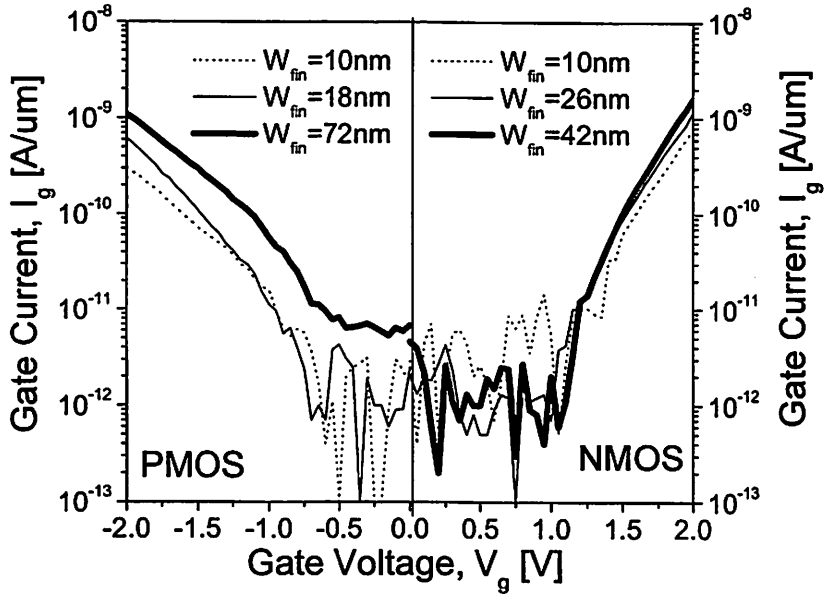


Figure 6.27. Measured gate current versus gate voltage for different fin widths. Narrower fin (thinner body) shows lower gate leakage current. (T_{ox}=2.1nm, N_{body}=2.0x10^{18}cm^{-3}, P+ $Si_{0.6}Ge_{0.4}$ gate)

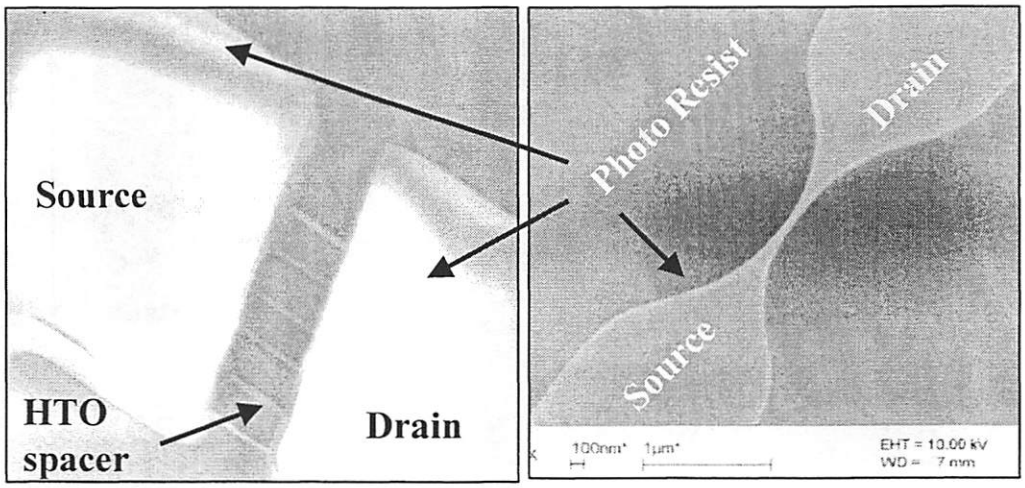
6.3 FinFET by Spacer Lithography

Control of the fin width is crucial for consistent suppression of short-channel effects. L_g/W_{fin} must be greater than *1.5* to control short-channel effects. This means that the fin width needs to be smaller than the gate length. This would be clearly impossible to accomplish with conventional lithography technologies when L_g is at the limit of lithography. Uniformity of the fin width is also critical to obtain consistent device characteristics. Thus, a new technology, spacer lithography, is proposed to make fins as narrow as possible and as uniform as possible [6.24][6.31]. This technique can also double the fin density for a given lithographic pitch, which results in increased current per unit area.

6.3.1 Process Details

The starting material was a (100) SOI p-type ($N_{body}=1x10^{15}cm^{-3}$) substrate with 100nm of silicon and 400nm of buried oxide. The process flow described in section 2.3.1 was not optimized to obtain fully working device because of topography issues. In this section, changes made to this process flow are described. In the previous design, thermal oxide was to be used for a hard mask to protect the silicon fin during the subsequent gate etch. However, the hard mask was already removed with HF during the sacrificial oxide removal as shown in Fig. 2.17 (a). Thus, in the new process, nitride is used as the hard mask. After reducing the SOI silicon from 100nm to 50nm by thermal oxidation, a thin pad oxide (4nm) is grown. Then, 50nm of nitride is deposited on the pad oxide by LPCVD. The role of pad oxide is to relieve the stress between the nitride and silicon

films. 200nm of sacrificial $Si_{0.4}Ge_{0.6}$ is then deposited by LPCVD on the nitride hard mask and patterned (to support the spacers) with optical lithography and plasma etching. 10nm high temperature oxide (HTO) is then deposited by LPCVD over the patterned sacrificial $Si_{0.4}Ge_{0.6}$ layer. The thickness of the HTO along the sidewalls of the sacrificial $Si_{0.4}Ge_{0.6}$ structures determines the final fin width. An extremely small beyond the lithographic limit, as well as very uniform fin width can therefore be obtained with this spacer lithography process. This is because the LPCVD deposition thickness can be controlled up to nanometer scale and can be more uniform across a wafer than the CD variation of lithography. A subsequent anisotropic dry etch removes the HTO film on top of the sacrificial $Si_{0.4}Ge_{0.6}$ structure to generate an even number of spacers. The sacrificial $Si_{0.4}Ge_{0.6}$ is removed with (5:1:1) H_2O:NH_4OH:H_2O_2 at 75°C [11]. HTO, thermally grown oxide, nitride, and Si are not etched significantly in this solution. Optical lithography is then used to define large S/D contact pads. In this same lithography layer, wide and variable fins can be defined. As a result, the narrowest fin is patterned with hard mask HTO spacers while large patterns are defined with photo-resist as shown in Fig. 6.28. In this manner, a drawback of spacer lithography techniques (only one linewidth available) is overcome.



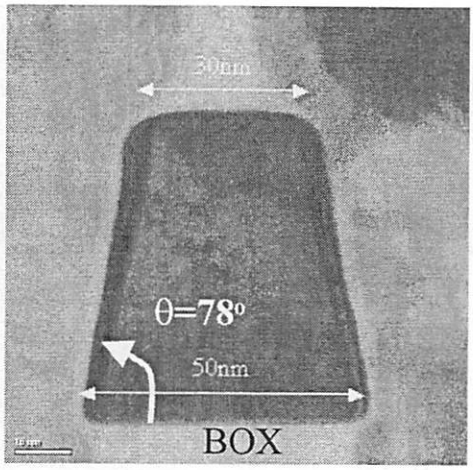
(a) (b)

Figure 6.28 Tilted SEM photographs of multi-fins and single fins. (a) The narrowest multi-fins are defined by HTO spacers while S/D pads are defined by resist and (b) wide and variable single-fins can be defined by the same resist in (a).

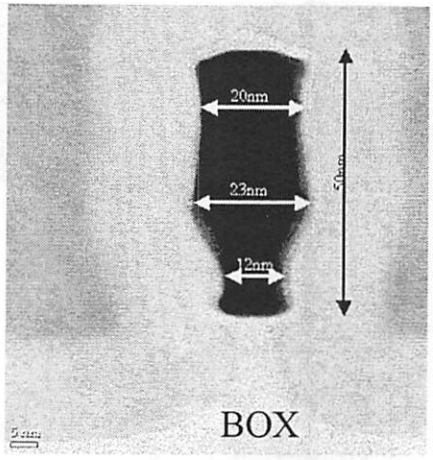
All hard mask oxide etches are performed with CF_4-based etches. For etching silicon, three different processes are used. The detailed process conditions are summarized in Table 6.1. Fig. 6.29 shows that each etching condition produces a different cross-sectional fin profile.

	Cl_2	HBr	CHF_3
Pressure (mTorr)	15	35	20
RF top power (W)	300	250	200
RF bottom power (W)	150	120	40
CHF_3 (sccm)	0	0	90
Cl_2 (sccm)	50	0	0
HBR (sccm)	150	200	200
O_2 (sccm)	0	5	0
Ar (sccm)	0	200	0

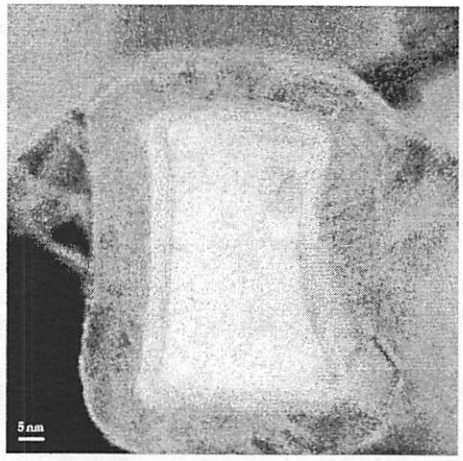
Table 6.1. Etch recipes used for silicon fin etch.



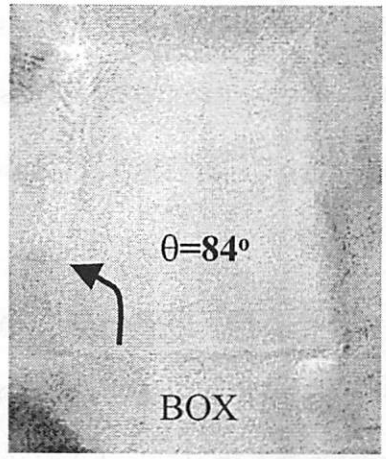
Cl_2 (60% of T_{Si})
HBr (100% of T_{Si}) [6.33]



Cl_2 (100% of T_{Si})
HBr (50% of T_{Si}) [6.33]



Cl_2 (130% of T_{Si})



CHF_3 (130% of T_{Si})

Figure 6.29 Cross-sectional TEM photographs of silicon fin. T_{Si} is the silicon film thickness, which is the same for all samples. Detailed etch recipes are summarized in Table 6.1

After fin etch, a sacrificial oxidation step is then used to remove etch damage. 10nm of thermal oxide was grown in at 900°C for 12min in O_2. Just prior to gate oxidation, the sacrificial oxide is removed with HF. This also removes the HTO spacers and undercuts the buried oxide beneath the fin as shown in Fig. 6.29. Then, argon anneal

at 900°C and 1min is used to cure etch damage from the silicon fin etch. 2.5nm gate oxide is then grown at 750°C for 14min. An in-situ 900°C for 30min N_2 anneal is used to improve the gate oxide quality. 1.5um of undoped $Si_{0.6}Ge_{0.4}$ by LPCVD is deposited as the gate material. SiGe is chosen to achieve the appropriate threshold voltage using gate work-function engineering [6.35]-[6.37]. Because gate planarization is performed with a subsequent CMP step, a very thick film is deposited. The 1.5um $Si_{0.6}Ge_{0.4}$ film is reduced to 400nm using CMP. The process details of the CMP and post-cleaning steps are described in Chapter 3. To investigate its effect on the threshold voltage, the $Si_{0.6}Ge_{0.4}$ gate doping is split using phosphorus and boron implantation after masks. The implant conditions are 150KeV, $5.0x10^{15}cm^{-2}$ and 0° for phosphorus and 60KeV, $5.0x10^{15}cm^{-2}$ and 0° for boron. After a 100nm LTO deposition as a hard mask, the gate electrode is then patterned over the fins using conventional lithography and etch processes. A 20nm gate length (in Fig. 6.32) is formed with ashing-trimming technology, which was described in the chapter 2.

A cross-sectional TEM photograph of the silicon fin with a nitride hard mask is shown in Fig. 6.30. Variations in gate oxide thickness are observed due to silicon crystal orientation and shown in Fig. 6. 31. In addition, the oxide is thinnest at the corner of the fin due to stress effects.

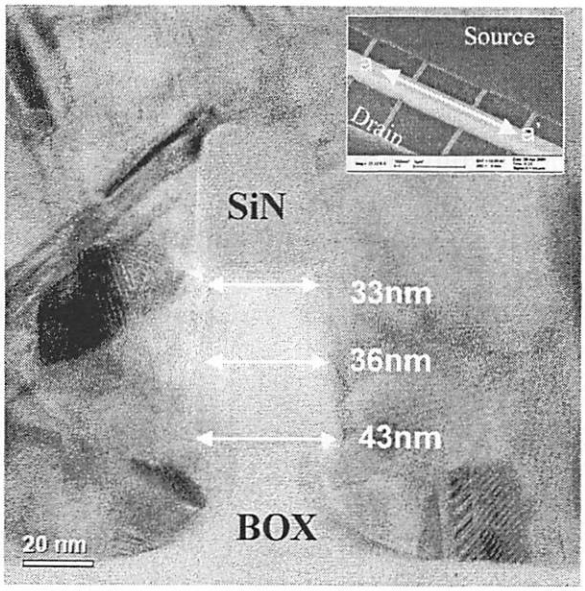
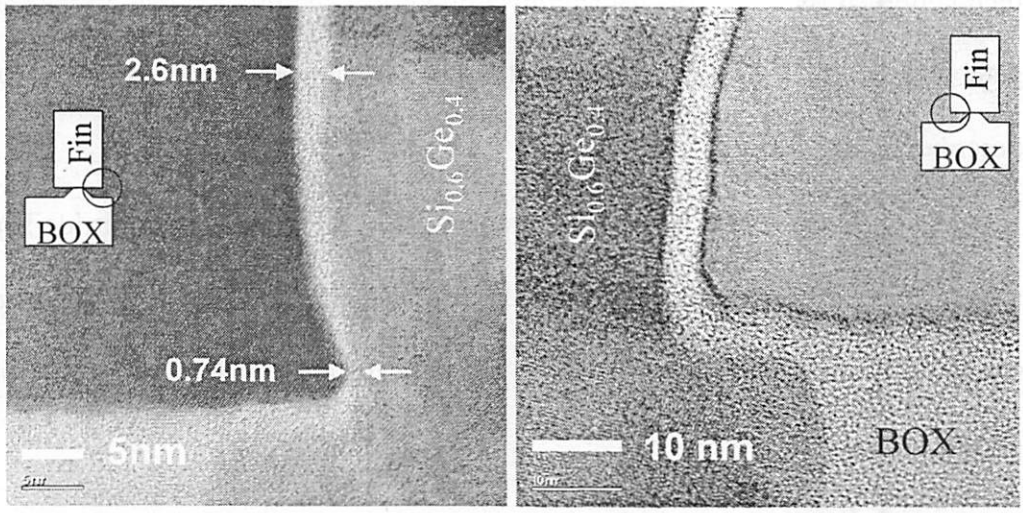


Figure 6.30 Cross-sectional TEM photograph of the silicon fin with a nitride hard mask along the a-a' direction in the inset.



(a) (b)

Figure 6.31 Cross-sectional TEM photographs of the gate oxide. At the sharp corner in (a), the gate oxide thickness is thinner because the growth rate is decreased by high stress.

This can cause serious reliability problems and device degradation. The corner profile needs to be controlled by the fin etch recipe. A vertical and rectangle fin profile is highly demanded. The gate electrode profile with single-fin and multi-fins are shown in Fig. 6.32 and Fig. 6.33.

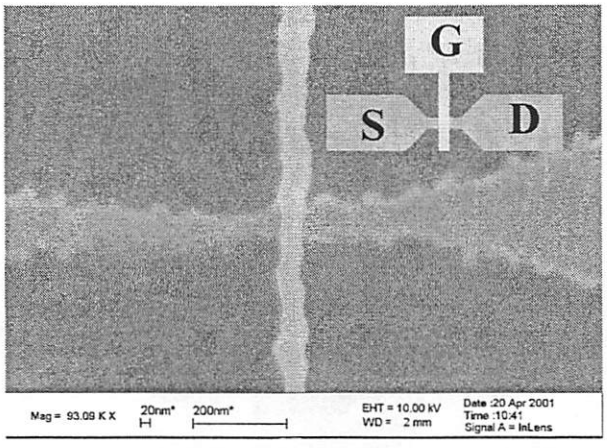


Figure 6.32 SEM top view of the gate (L_g=*20nm*) and a single-fin (W_{fin}=*40nm*).

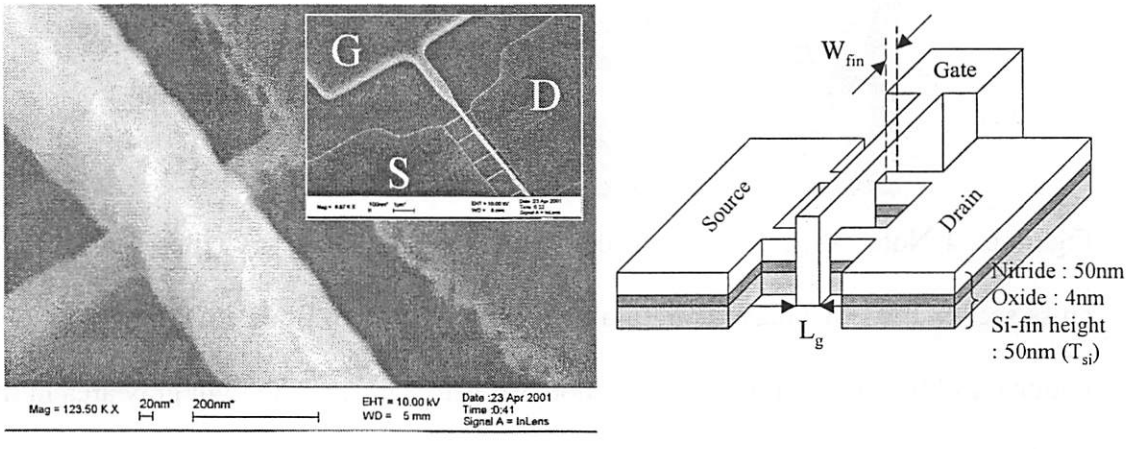


Figure 6.33 SEM tilted view of the gate (L_g=*60nm*) in a multi-fin (W_{fin}=*40nm*) structure. The gate is completely planarized by CMP.

After detection of the end-point signal with Cl_2-based recipe during the gate $Si_{0.6}Ge_{0.4}$ etch, a 40% overetch is used with a HBr-based recipe which resulted in a notched profile (T-shaped gate) both on the planar surface as well as along the sidewall of the fin. For TEM analysis, the silicon fin is cut at a tilted angle as shown in Fig. 6.34 (a). Figure 6.34 (b) shows that a notched region (in the circle of Fig. 6.34 (c) and 6.34 (d)) is made along the vertical sidewall of the fin. The 40nm vertical fin width is broadened to 55nm because of the tilted sample cut as shown in Fig. 6.34 (b).

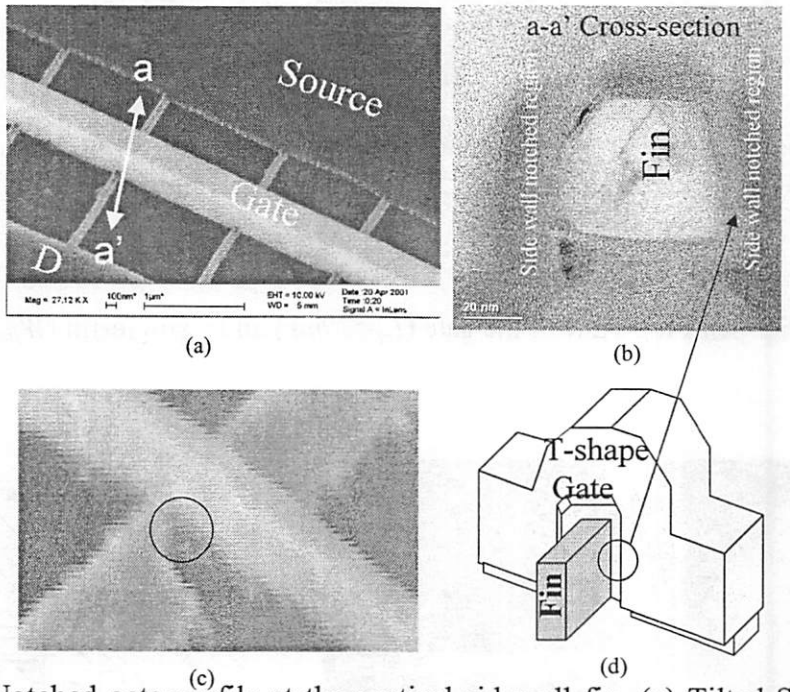


Figure 6.34 Notched gate profile at the vertical sidewall fin. (a) Tilted SEM image, (b) cross-sectional TEM image along a-a' direction in (a), (c) tilted SEM image of notched region (circle), and (d) drawing of the notched gate (circle). The light gray area in (b) is a notched region at the vertical sidewall.

The rest of the spacer process is the similar to the e-beam process described earlier in this chapter. The hard mask nitride on the fin is removed during the spacer overetch. The final spacer profile is shown in Fig. 6.35. Then, S/D doping using ion implantation is done after N+ and P+ S/D masking. Phosphorus is implanted at $5x10^{15}cm^{-2}$, 30KeV, and 0° for N+ S/D while boron is implanted at $5x10^{15}cm^{-2}$, 10KeV, and 0° for P+ S/D. A RTA step at 900°C for 1min in N_2 ambient is used to activate the implanted dopants in the S/D region. The remaining gate hard mask (LTO) protects the gate from being counter-doped. Finally a 400°C forming gas anneal is used for passivation. Metallization is not used in this run.

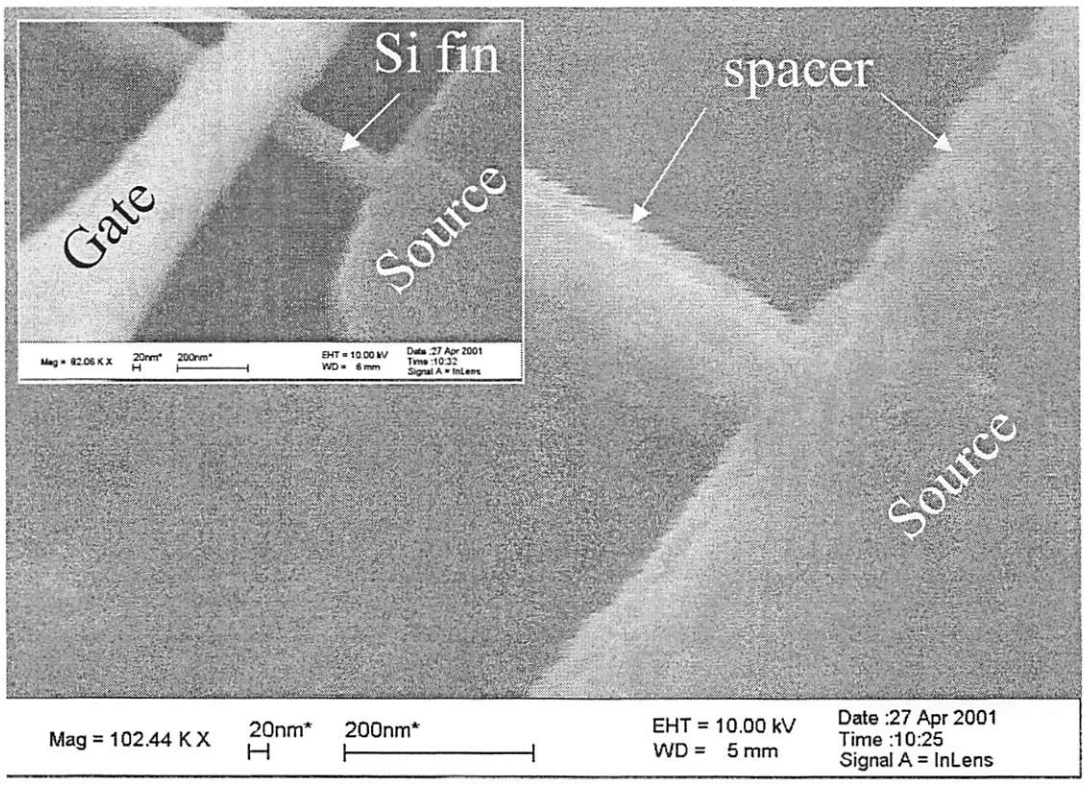
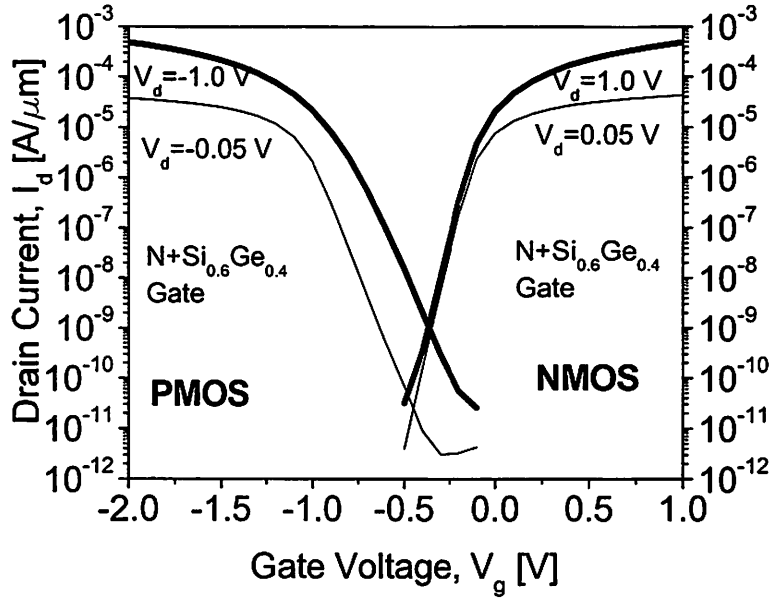


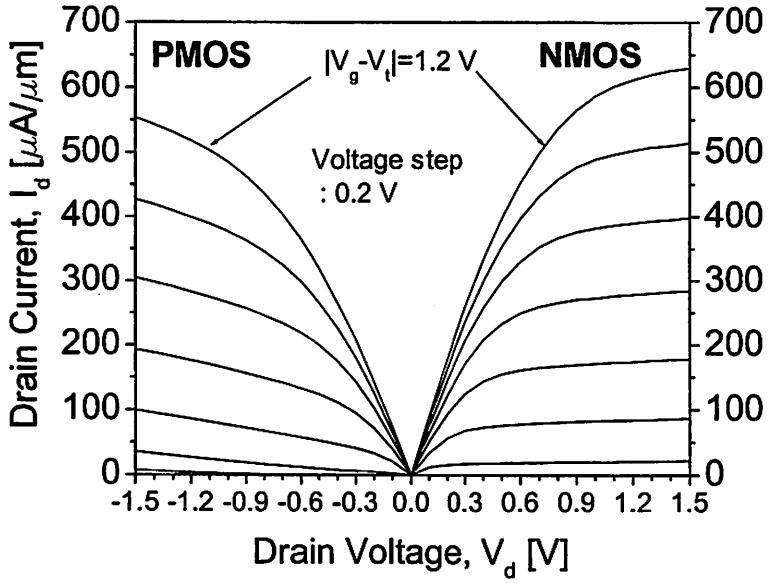
Figure 6.35 Tilted SEM photograph of spacer profile.

6.3.2 Device Characteristics of Spacer FinFETs

Measured current-voltage (I-V) characteristics for a 60nm FinFET with a 40nm fin width and 2.5nm gate oxide are shown in Fig. 6.35. This is a single-fin device, which is defined by conventional lithography with a S/D pad mask (no spacer lithography) as shown in Fig. 6.27. NMOS drive current is 500uA/um and PMOS drive current is -380uA/um at $|V_g-V_t|$=1V and $|V_d|$=1V. All currents are normalized with a conservative channel width definition ($W=2*T_{Si}$ in Fig. 6.12). With a new channel width definition ($W=T_{Si}$), the NMOS current is 1mA/um and the PMOS current is 760uA/um. The higher current in FinFETs by spacer lithography than in FinFETs by e-beam lithography (Fig. 6.17) comes from the wider fin width, which has lower series resistance. For 6-fin devices defined by spacer lithography, typical I-V characteristics are shown in Fig. 6.37. With normalization by the conservative channel definition, the NMOS current is 395uA/um and the PMOS is –340uA/um at the same bias conditions. With the new channel width definition, they will be doubled.

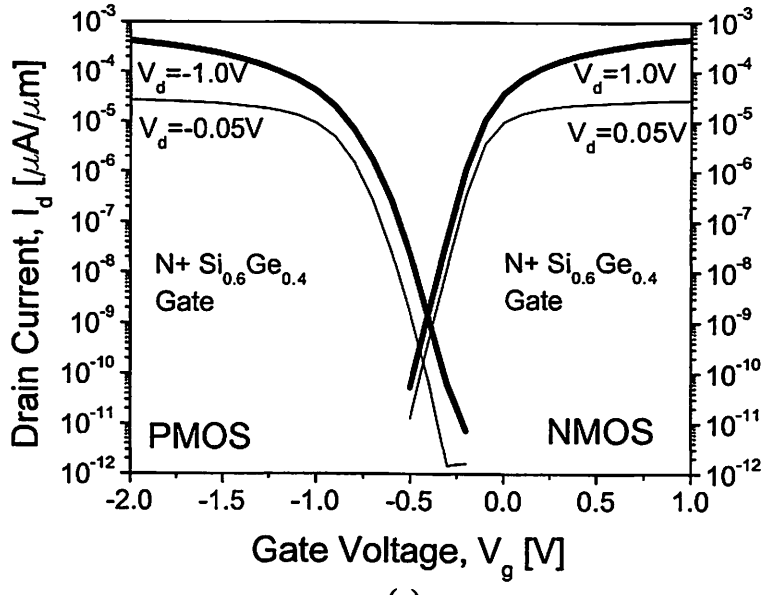


(a)

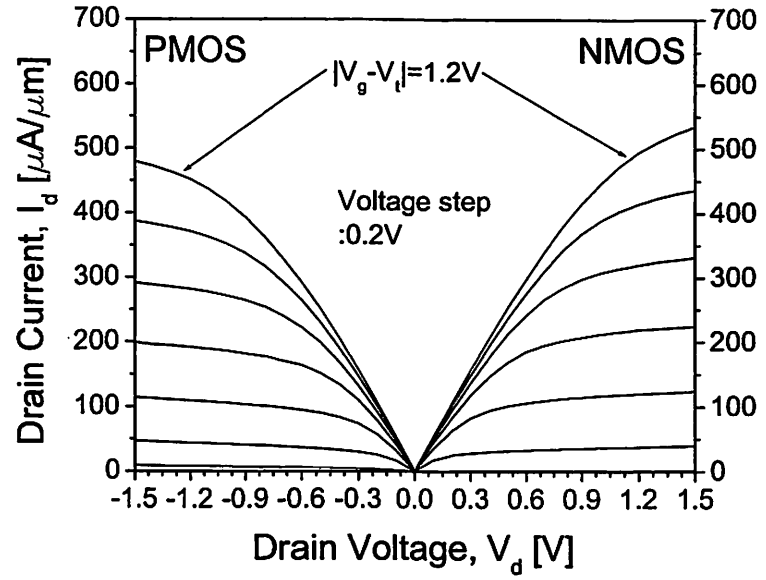


(b)

Figure 6. 36 Measured I-V characteristics for a single fin device. (a) I_d-V_g and (b) I_d-V_d characteristics for L_g=60nm, W_{fin}=40nm, n+ $Si_{0.6}Ge_{0.4}$ gate, T_{ox}=2.5nm, and N_{body}=$2.0x10^{17}cm^{-3}$ (n-type).



(a)



(b)

Figure 6. 37 Measured I-V characteristics for a multi-fin (6) device. (a) I_d-V_g and (b) I_d-V_d characteristics for L_g=60nm, W_{fin}=40nm, n+ $Si_{0.6}Ge_{0.4}$ gate, T_{ox}=2.5nm, and N_{body}=2.0x10^{17}cm^{-3} (n-type).

The threshold roll-off characteristics are shown in Fig. 6.38. Figure 6.39 shows the subthreshold swing and DIBL versus the gate length. The PMOS short-channel effects are slightly worse than the NMOS because the diffusivity of boron in the S/D is larger than that of phosphorus. The minimum gate length is 60nm, which is consistent with the $L_g/W_{fin} = 1.5$ requirement. In these spacer FinFETs, the short-channel effects are suppressed more than in e-beam FinFETs despites a slightly thicker gate oxide (2.5nm versus 2.1nm). This is because a large undercut underneath the fin is made with a relatively long etch time in HF to remove the sacrificial oxide as shown in Fig. 6.30. Park [6.58] reported that short-channel effects would be better suppressed with a Pi (Π)-shaped gate structure. The gate profile wrapping around the fin shields the electric field from the drain at the back of the channel region. The gate structure of e-beam FinFETs is close to the Pi (Π) shape while the gate structure of spacer FinFETs is close to an Omega (Ω) shape which should be even more effective than the pi-shaped gate in controlling short-channel effects at the cost of increased overlap capacitance.

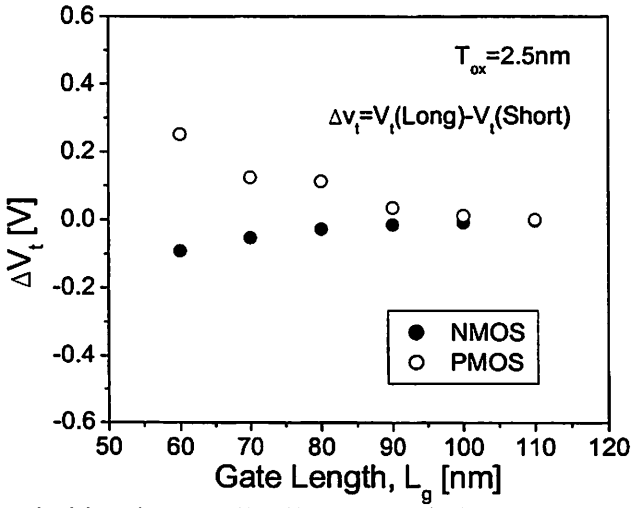


Figure 6.38 Threshold voltage roll-off characteristics. Each data point is the average of five devices (W_{fin}=40nm, $n+$ $Si_{0.6}Ge_{0.4}$ gate, T_{ox}=2.5nm, and $N_{body}=2.0x10^{17}cm^{-3}$ (n-type)).

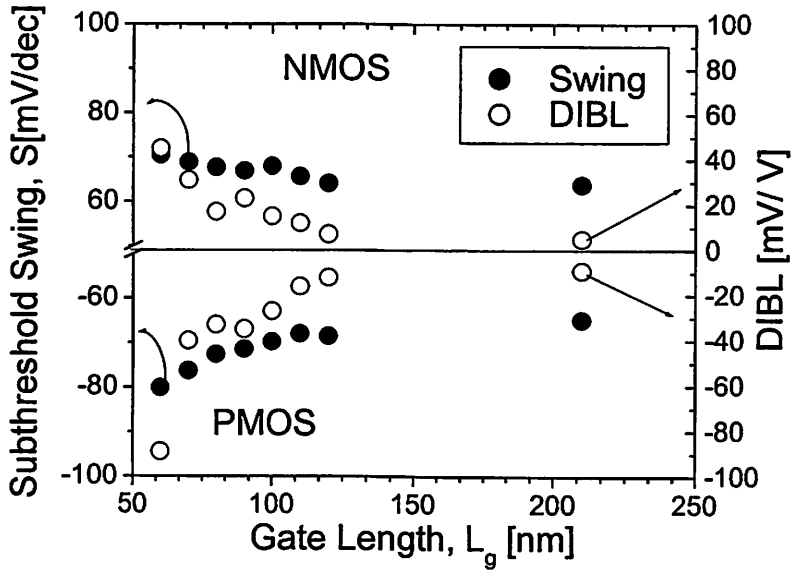


Figure 6.39 Subthreshold swing and DIBL versus the gate length. Each data point is the average of five devices (W_{fin}=40nm, n+ $Si_{0.6}Ge_{0.4}$ gate, T_{ox}=2.5nm, and $N_{body}=2.0x10^{17}cm^{-3}$ (n-type)).

The threshold voltage dependence on gate material and body doping is also investigated. Undoped $Si_{0.6}Ge_{0.4}$ is deposited as the gate material and then doped with ion implantation as described in the section 6.3.1. SOI wafers were initially doped with B ($1.0x10^{15}cm^{-3}$) to form a p-type body while phosphorus implantation formed an n-type body. Argon anneal is used to cure any etch damage on the fin sidewall, which was created during plasma etching. For PMOS devices with a boron doped (P+) $Si_{0.6}Ge_{0.4}$ gate, linear-like resistor behavior is observed. This is due to boron penetration, which results in heavily counter-doping of the body. The doping level may be higher than ($1.0x10^{19}cm^{-3}$ as estimated from resistance measurements. NMOS devices with boron-doped (P+) $Si_{0.6}Ge_{0.4}$ gates show very high threshold voltage (> 1.5V), but still functional

as transistors. This high threshold voltage is caused by heavy doping of the body (from boron penetration) and quantum confinement of the inversion charges. The threshold voltage shift from quantum confinement can be larger than 0.7V at $1.0x10^{19}cm^{-3}$ [6.59]-[6.61]. Only n-type heavily doped $Si_{0.6}Ge_{0.4}$ gated devices show a reasonable threshold voltage as shown in Fig. 6.40. The data obtained here show that achieving the correct threshold voltages for NMOS and PMOS will remain a challenge.

Fukuda et al. reported that argon annealing could reduce interface trap density [6.62]. In this experiment, only devices that has undergone argon annealing yields the predicted value of the threshold voltage. This amounts to a threshold voltage shift of 0.2V~0.6V due to the argon anneal. It can be deduced that the interface trap density is not negligible in the etched body. Plasma hydrogenation is applied to FinFETs and UTBFETs, but the threshold voltage change in FinFETs was 0.2V [6.23] and there was no change in UTBFETs. This also supports the theory that there is much etch damage in the etched body.

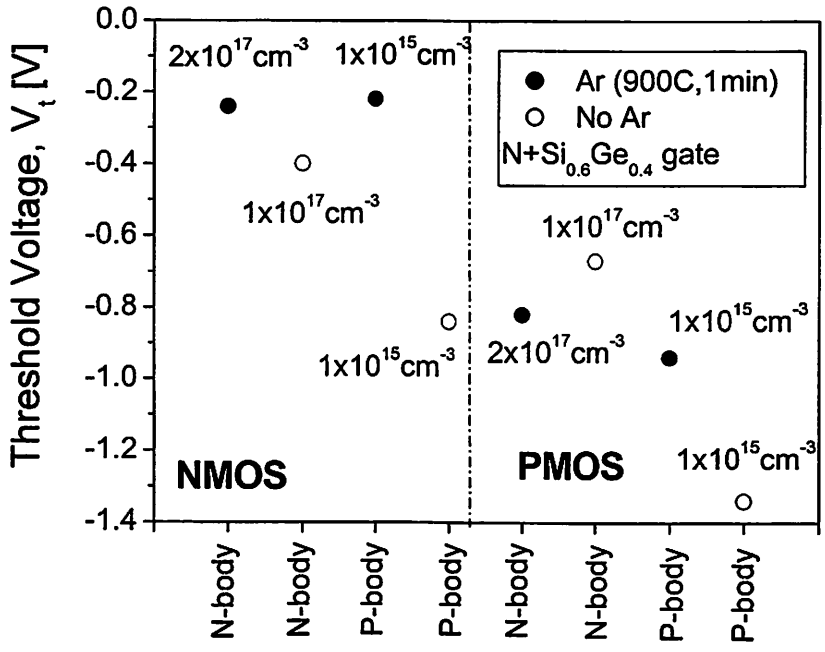


Figure 6.40 Threshold voltage dependence on body type (N-type vs. P-type) and doping concentration. P-type body was initially doped with boron ($1.0x10^{15}cm^{-3}$) and N-type is doped with phosphorus implantation ($1.0x10^{17}cm^{-3}$, $2.0x10^{17}cm^{-3}$). Argon anneal (900°C for 1min) is applied to minimize the etch damage after sacrificial oxidation. (L_g=*60nm*, W_{fin}=*40nm*, *n+* $Si_{0.6}Ge_{0.4}$ gate, and T_{ox}=*2.5nm*). Each data point is the average of five devices.

Since the gate is not self-aligned to the wide S/D pads as shown in shown as shown in the inset of Fig.6.30, series resistance of the narrow fin can degrade device performance. Test structures are designed to measure the total resistance of the fin as a function of the total length of the fin ('D' in Fig. 6.41) with fixed fin height and fin width. Figure 6.42 shows that the total resistance of the fin is proportional to the fin length, D.

The extracted doping concentration at the fin is $1.5x10^{20}cm^{-3}$ for a boron doped S/D and $1.2x10^{20}cm^{-3}$ for a phosphorus doped S/D. To investigate the drive current dependence on the extension length (S_2, the gap between the gate and the source), a misalignment off-set test structure is designed as shown in the inset of Fig. 6.42. The total gap ($D=S_1+S_2+L_g$) between the source pad and drain pad is fixed while the gate is intentionally misaligned with a stepping distance of 20nm. Figure 6.42 shows that the drive current increases as the extension length S_2 decreases. This clearly shows that a raised S/D process such as selective Ge deposition or selective silicon epitaxial growth should be used to minimize the effects of misalignment and series resistance of the extension length (S_2).

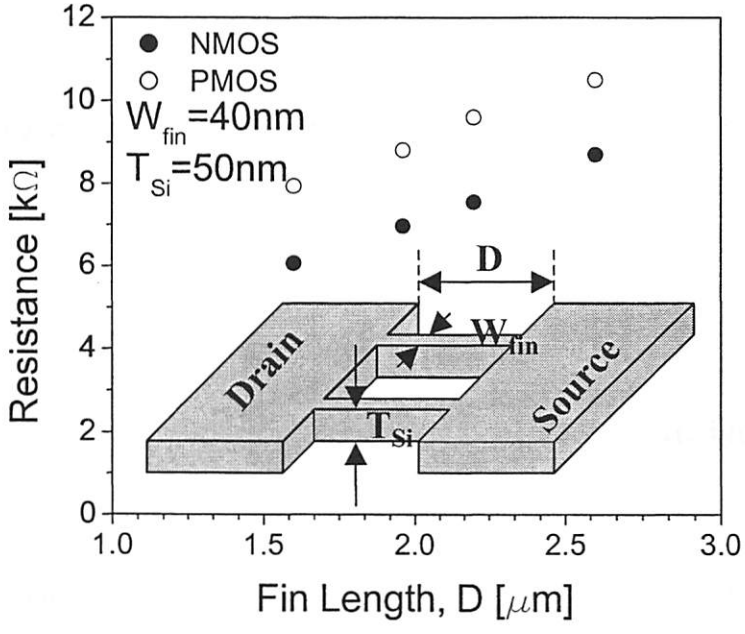


Figure 6.41 The total resistance of the fin is proportional to the fin length, D.

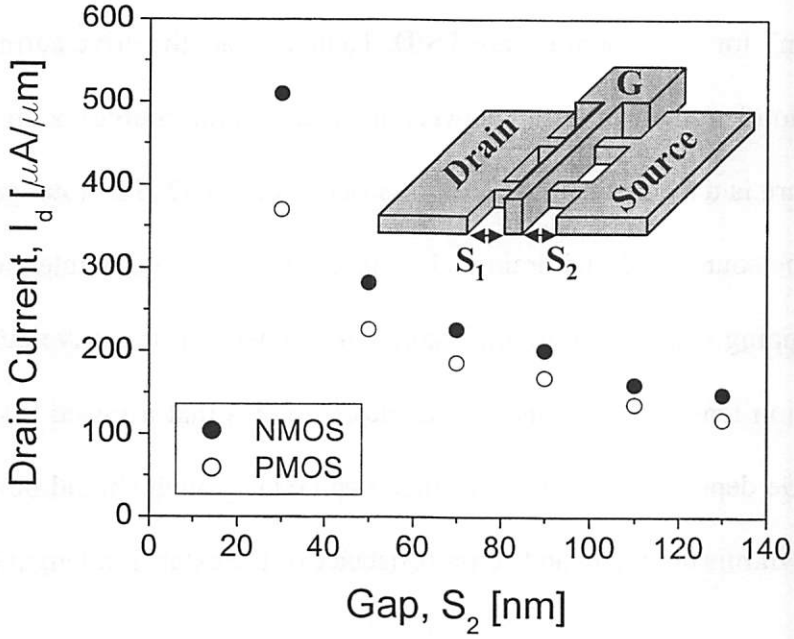


Figure 6.42 Drive current dependence on the gap (S_2) between the gate and the source.

6.4 Conclusion

The double-gate structure as represented by the FinFET appears to offer greater scalability down to 10nm gate length or perhaps even below. Quasi-planar and simplified CMOS FinFETs are proposed and demonstrated. Sub-20nm gate lengths with 10nm fins (world record smallest double-gate CMOS) are successfully demonstrated. Two different lithography technologies: e-beam and spacer, are developed. A selective Ge deposition for a raised S/D is also developed and used, which enhances the drive current by up to

28% in FinFETs. CMP technology for gate planarization is used to provide wide lithography and etch process windows. Both e-beam and spacer FinFETs show excellent off-state leakage and on-state drive current. NMOS drive current is relatively low, which may be related to surface roughness of the fin sidewall. Argon anneal is successfully used to cure etch damage after sacrificial oxidation. Adjustment of the threshold voltage is still unsolved in the thin body device. Extension resistance is found to significantly reduce the drive current. Planar FinFET is currently the most attractive double-gate MOSFET structure and should allow for continued device scaling into the "no known solution" and "beyond roadmap" regimes in the ITRS roadmap. It appears to us that the continued evolution of CMOS integrated circuit technology into this regime will not be impeded by basic limitations underlying transistor technology. The implication of this is that "Moore's law" may continue for yet another 15-20 years before the ultimate device limits for CMOS are reached. FinFET technologies invented at the University of California-Berkeley have already been transferred to industrial research favorites at TSMC, AMD, Motorola, and IBM.

6.5 References

[6.1] Y.-K. Choi, K. Asano, N. Lindert, V. Subramanian, T.-J. King, J. Bokor, C. Hu, "Ultra-thin-body SOI MOSFET for Deep-sub-tenth Micron Era," *IEEE Electron Device Letters,* vol. 21, p.254-255, 2000.

[6.2] L. Chang, S. Tang, T.-J. King, J. Bokor, C. Hu, "Gate Length Scaling and Threshold Voltage Control of Double-Gate MOSFETs", *IEEE IEDM Technical Digest*, p. 719-22, 2000

[6.3] M. Ieong, H.-S.P. Wong, Y. Taur, P. Oldiges, D. Frank, "DC and AC performance analysis of 25 nm symmetric/asymmetric double-gate, back-gate and bulk CMOS", *International Conference on Simulation Semiconductor Processes and Devices*, p.147-50, 2000.

[6.4] J. Brews, W. Fichtner, E.H. Nicollian, S.M. Sze, "Generalized guide for MOSFET miniaturization", *IEEE Electron Device Letters*, vol.1, p.2-4, 1980.

[6.5] R.-H. Yan, A. Ourmazd, K.F. Lee, "Scaling the Si MOSFET: from bulk to SOI to bulk", *IEEE Transactions on Electron Devices*, vol.39, p.1704-10, 1992.

[6.6] Y. Taur, S. Wind, Y.J. Mii, Y. Lii, D. Moy, K.A. Jenkins, C.L. Chen, P.J. Coane, D. Klaus, J. Bucchignano, M. Rosenfield, M.G.R. Thomson, M. Polcari, "High performance 0.1 um CMOS devices with 1.5 V power supply", *IEEE IEDM Technical Digest*, p.127-30, 1993.

[6.7] L. Kang, K. Onishi, Y. Jeon, B. H. Lee, C. Kang, W.-J. Qi, R. Nieh, S. Gopalan, R. Choi, J.C. Lee, "MOSFET devices with polysilicon on single-layer HfO_2 high-K dielectrics", *IEEE IEDM Technical Digest*, p.35-38,2000.

[6.8] A. Chin, Y.H. Wu, S.B. Chen, C.C. Liao, W.J. Chen, "High quality La_2O_3 and Al_2O_3 gate dielectrics with equivalent oxide thickness 5-10 A", *Symposium on VLSI Technology*, p.16-17, 2000.

[6.14] K.P. Anish Kumar, J. K.O. Sin, C.T. Nguyen, P.K. Ko, "Kink-Free Polycrystalline Silicon Double-Gate Elevated-Channel Thin-Film Transistors", *IEEE Transactions on Electron Devices*, vol.45, p.2514-20, 1998.

[6.15] B.E. Roberds, E.J. Whang, A. Rudolph, B.S. Doyle, "Investigation of a Novel Self-Aligned Dual Gate MOSFET Structure", *Proceedings of IEEE International SOI Conference*, p.109-10, 1998.

[6.16] D. Hisamoto, T. Kaga, E. Takeda, "Impact of the vertical SOI 'DELTA' structure on planar device technology", *IEEE Transactions on Electron Devices*, vol.38, p.1419-24, 1991.

[6.17] X. Huang, W.-C. Lee, C. Kuo, D. Hisamoto, L. Chang, J. Kedzierski, E. Anderson, H. Takeuchi, Y.-K. Choi, K. Asano, V. Subramanian, T.-J. King, J. Bokor, C. Hu, "Sub-50 nm P-channel FinFET", *IEEE Transactions on Electron Devices*, vol.48, p.880-6, 2001.

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

[6.19] H.-S.P. Wong, D.J. Frank, P.M. Solomon, C.H.J. Wann, J.J. Welser, "Nanoscale CMOS", *Proceedings of the IEEE*, vol.87, p.537-70, 1999.

[6.20] H. Noda, F. Murai, and S. Kimura, "Threshold voltage controlled 0.1um MOSFET utilizing inversion layer as extremely shallow source/drain", *IEEE IEDM Technical Digest*, p.123-6, 1993.

[6.21] H. Kawaura, T. Sakamoto, T. Baba, Y. Ochiai, J. Fujita, J. Sone, "Transistor characteristics of 14nm-gate-length EJ-MOSFET's", *IEEE Transactions on Electron Devices*, vol.47, p.856-60, 2000.

[6.22] H. Kawaura, T. Toshisugu, T. Baba, "Observation of source-to-drain direct tunneling current in 8nm gate electrically variable shallow junction metal-oxide-semiconductor field-effect-transistors", *Applied Physics Letters*, vol.76, p.3810-12, 2000.

[6.23] N. Lindert, L. Chang, Y.-K. Choi, E. H. Anderson, W.-C. Lee, T.-J. King, C. Hu, "Sub-60-nm Quasi-Planar FinFETs Fabricated Using a Simplified Process", *IEEE Electron Device Letters*, vol.22, p.487-9, 2001.

[6.24] Y.-K. Choi, N. Lindert, P. Xuan, S. Tang, D. Ha, E. Anderson, T.-J. King, J. Bokor, C. Hu, "Sub-20nm CMOS FinFET Technologies", *IEEE IEDM Technical Digest*, p.421-24, 2001.

[6.25] B. Majkusiak, T. Janik, J. Walczak, "Semiconductor Thickness Effects in the Double-Gate SOI MOSFET", *IEEE Transactions on Electron Devices*, vol.45, p.1127-34, 1998.

[6.26] M. Shoji, S. Horiguchi, "Electronic structures and phonon-limited electron mobility double-gate silicon-on-insulator Si inversion layers", *Journal of Applied Physics*, vol.85, p.2722-31, 1999.

[6.27] H. Majima, H. Ishikuro, T. Hiramoto, "Threshold Voltage Increase by Quantum Mechanical Narrow Channel Effect in Ultra-Narrow MOSFETs", *IEEE IEDM Technical Digest*, p.379-82, 2001.

[6.28] Y.-K. Choi, D. Ha, T.-J. King, and C. Hu, "Ultra-thin Body PMOSFETs with Selectively Deposited Ge Source/Drain", *Symposium on VLSI Technology*, p.16-7, 2001.

[6.29] Y.-K. Choi, D. Ha, T.-J. King, C. Hu, "Threshold voltage Shift by Quantum Confinement in Ultra-thin Body Device", *59th Device Research Conference*, p.85-6, 2001.

[6.30] T. Ouisse, "Self-consistent quantum-mechanical calculations in ultrathin silicon-on-insulator structures", Journal of Applied Physics, vol.76, p.5989-95, 1994.

[6.31] Y.-K. Choi, T.-J. King, C. Hu, "Nanoscale CMOS Spacer FinFET for the Terabit Era", *IEEE Electron Device Letters*, to be published, 2002.

[6.32] N. Lindert, Y.-K. Choi, L. Chang, E. Anderson, W.-C. Lee, T.-J. King, J. Bokor, C. Hu, "Quasi-Planar FinFETs with Selectively Grown Germanium Raised Source/Drain", *IEEE International SOI Conference*, p.111-2, 2001.

[6.33] N. Lindert, Doctoral Thesis, "Process Development and Device Design for Continued MOSFET Scaling", Department of Electrical Engineering and Computer Science, *University of California at Berkeley*, 2001.

[6.34] K. Asano, Y.-K. Choi, T.-J. King, C. Hu, "Patterning Sub-30-nm MOSFET Gate with I-line Lithography," *IEEE Transactions on Electron Devices,* vol. 48, p.1004-6, 2001.

[6.35] T.-J. King, J. P. McVittie, and K. C. Saraswat, "Electrical Properties of Heavily Doped Polycrystalline Silicon-Germanium Films," *IEEE Trans. on Electron Devices*, vol. 41, p.228-231, 1994.

[6.36] P.-E. Hellberg, S.-L. Zhang, C.S. Petersson, "Work Function of Boron-Doped Polycrystalline Si_xGe_{1-x} Film," *IEEE Trans. on Electron Devices*, vol. 18, p.456-8, 1997.

[6.37] C. Salm, D.T. van Veen, D.J. Gravesteijn, J Holleman, P.H. Worlee, "Diffusion and Electrical Properties of Boron and Arsenic Doped Poly-Si and Poly-Ge_xSi_{1-x} (x~0.3) as Gate Material for Sub-0.25um Complementary Metal Oxide Semiconductor applications," *Journal of Electrochemical Society*, vol.144, p.3665-73, 1997.

[6.38] D.J. Frank, Y. Taur, H.-S. P. Wong, "Generalized Scale Length for Two-Dimensional Effects in MOSFET's", *IEEE Electron Device Letters*, vol.19, p.385-7, 1998.

[6.39] K. Suzuki, T. Tanaka, Y. Tosaka, H. Horie, Y. Arimoto, "Scaling Theory for Double-Gate SOI MOSFET's", *IEEE transactions on Electron Devices*, vol.40, p.2326-9, 1993.

[6.40] R.-H. Yan, A. Ourmazd, K.F. Lee, "Scaling the Si MOSFET: From bulk to SOI to bulk", *IEEE transactions on Electron Devices*, vol.39, p.1704-10, 1992.

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

[6.42] Z.-H. Liu, C. Hu, J.-H. Huang, T.-Y. Chan, M.-C. Jeng, P.K. Ko, Y.C. Cheng, "Threshold voltage model for deep-submicrometer MOSFET's", *IEEE Transactions on Electron Devices*, vol.40, p.86-95, 1993.

[6.43] C.J. Petti, J.P. McVittie, J.D. Plummer, "Characterization of Surface Mobility on the Sidewalls of Dry-Etched Trenches, *IEEE IEDM Technical Digest*, p.104-7, 1988.

[6.44] A. Harstein, T.H. Ning, A.B. Fowler, "Electron Scattering in Silicon Inversion Layers by Oxide and Surface Roughness", *Surface Science*, vol.58, p.178-81, 1976.

[6.45] Y.C. Cheng, E.A. Sullivan, "On the role of scattering by surface roughness in silicon inversion layers", *Surface Science*, vol.34, p.178-81, 1973.

[6.46] D. Ferry, *Semiconductors*, Macmillan Publishing Company, New York, 1991.

[6.47] F. Gamiz, J.B. Roldan, P. Cartujo-Cassinello, J.A. Lopez-Villanueva, P. Cartujo, "Role of surface-roughness scattering in double gate silicon-oninsulator inversion layers", *Journal of Applied Physics*, vol.89, p.1764-70, 2001.

[6.48] G. Mazzoni, A.L. Lacaita, L.M. Perron, A. Pirovano, "On Surface Roughness-Limited Mobility in Highly Doped n-MOSFET's", *IEEE Transactions on Electron Devices*, vol.46, p.1423-8, 1999.

[6.49] Schred 2.0, Numerical Simulator, *Purdue University*, http://nanohub.purdue.edu

[6.50] S. Tiwari, M.V. Fischetti, P.M. Mooney, J.J. Welser, "Hole mobility Improvement in Silicon-on-Insulator and Bulk Silicon Transistors Using Local Strain", *IEEE IEDM Technical Digest*, p.939-41, 1997.

[6.51] G. Scott, J. Lutze, M. Rubin, F. Nouri, M. Manley, "NMOS Drive Current Reduction Caused by Transistor Layout and Trench Isolation Induced Stress", *IEEE IEDM Technical Digest*, p.827-30, 1999.

[6.52] W.G. En, D.-H. Ju, D. Chan, S. Chan, O. Karisson, "Reduction of STI/active Stress on 0.18um SOI Devices Through Modification of STI Process", *Proceedings of IEEE International SOI Conference*, p.85-6, 2001.

[6.53] C.-L. Huang, H. Soleimani, G. Grula, "LOCOS Process-Induced Stress Effect on CMOS SOI Characteristics", *Proceedings of IEEE International SOI Conference*, p.82-3, 1996.

[6.54] C.-L. Huang, H. Soleimani, G. Grula, N.D. Arora, D. Antoniadis, "Isolation Process Dependence of Channel Mobility in Thin-Film SOI Devices", *IEEE Electron Device Letters*, vol.17, p.291-3, 1996.

[6.55] C.-L. Huang, H. Soleimani, G.J. Grula, J.W. Sleight, A. Villani, H. Ali, D.A. Antoniadis, "LOCOS-Induced Stress Effects on Thin-Film SOI Devices", *IEEE Transactions on Electron Devices*, vol.44, p.646-50, 1997.

[6.56] L.T. Su, F. Hao, J.E. Chung, D.A. Antoniadis, "Hot-carrier effects in fully-depleted SOI nMOSFETs", *IEEE IEDM Technical digest*, p.349-52, 1992.

[6.57] L. Chang, K.J. Yang, Y.-C. Yeo, Y.-K. Choi, T.-J. King, C. Hu, "Reduction of Direct-Tunneling Gate Leakage Current in Double-Gate and Ultra-Thin Body MOSFETs, *IEEE IEDM Technical digest*, p.99-102, 2001.

[6.58] J.-T. Park, J.-P. Colinge, C.H. Diz, "Pi-Gate SOI MOSFET", *IEEE Electron Device Letters*, vol.22, p.405-6, 2001.

[6.59] S. Jallepalli, J. Bude, W.-K. Shih, M. R. Pinto, C.M. Maziar, A.F. Tasch, "Electron and hole quantization and their impact on deep submicron silicon p- and n-MOSFET characteristics", *IEEE Transactions on Electron Devices*, vol.44, p.297-303, 1997.

[6.60] G. Chindalore, S.A. Hareland, S. Jallepalli, A.F. Tasch, C.M. Maziar, V.K.F. Chia, S. Smith, "Experimental determination of threshold voltage shifts due to quantum mechanical effects in MOS electron and hole inversion layers", *IEEE Electron Device Letters*, vol.18, p.206-8, 1997.

[6.61] Y. Ma, Z. Li, L. Liu, Z. Yu, "MOS structure threshold voltage model by rigorously considering quantum mechanical effect", *Proceedings of 22nd International Conference on Microelectronics*, vol.1, p.225-8, 2000.

[6.62] H. Fukuda, M. Yasuda, T. Iwabuchi, S. Kaneko, T. Ueno, I. Ohdomari, "Process dependence of the SiO_2/Si(100) interface trap density of ultrathin SiO_2 films", *Journal of Applied Physics*, vol.72, p.1906-11, 1992.

[6.63] J.T. Watt, J.D. Plummer, "Universal mobility-field curves for electrons and holes in MOS inversion layers", *Symposium on VLSI Technology*, p.81-2, 1987.

6.6 Appendices

6.6.1 Process Flows of E-Beam FinFETs

Step	Process Name	Process Specification	Equipment	Comment
1.0	Wafers	4" p-type SOI		T_{Si}=100nm
2.0	Body Thinning and Alignmnet key formation			
2.01	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
2.02	Body thinning	Wet, SWETOXB, 850°C, 2hour, 120nm	Tylan 2	
2.03	Oxide removal	(10:1) HF, 10min	Sink 6	
2.04	Measuement	Body thickness measurement	Nanoduv	T_{Si}=50nm
2.05	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
2.06	LTO deposition	LTO deposition, 11SULTOA, 2min 40sec, 50nm	Tystar 11	
2.07	Measuement	LTO thickness measurement	Nanoduv	
2.08	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
2.09	$Si_{0.5}Ge_{0.5}$ deposition Recipe: SIGEVAR.019	Nucleation : 550°C, 300mT, SiH_4=200,30sec Deposition : 500°C, 300mT, SiH_4=186, GeH_4=33 1 hour	Tysatr 19	
2.10	Measuement	SiGe thickness measurement	Nanoduv	
2.11	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
2.12	LTO deposition	LTO deposition, 11SULTOA, 5min 20sec, 100nm	Tystar 11	
2.13	Alignment key mask	Resist coating : coat = prog#01/bake = prog#01 Exposure Development : bake = prog#01/develop = prog#01	Svgcoat gcaws Svgdev	
2.14	Poly-$Si_{0.5}Ge_{0.5}$ etch	BT : TP=200W,BP=40W,20mT, CF4=100sccm, 63sec ME : TP=300W,BP=150W,15mT Cl_2:HBr=50:150, EDP OE : TP=250W,BP=120W,15mT HBr:O_2=200:5, 30sec	Lam5	
2.15	Resist strip	(100:1) HF, 20sec O_2 ashing, 300W, 5min (100:1) HF, 20sec	Sink 7 Technics-c Sink 7	Polymer removal
2.16	Post cleaning	Piranha, 120°C, 20min	Sink 8	
2.17	Channel implantation	phosphorus/1.0x$10^{13}$$cm^{-2}$/30KeV/0°	Implanter	foundry company
3.0	Active Formation			
3.01	Fin mask	E-beam lithography SAL-601, 200nm	e-beam nanowriter	at LBL
3.02	Ashing	O_2 ashing, 3W, 20sec	Technics-c	
3.03	Measurement	CD measuremt and additive ashing	Leo/Tech.-c	
3.04	LTO Etch	TP=200W,BP=40W,13mT,CF_4=100, 30sec	Lam5	
3.05	Resist strip	(100:1) HF, 10sec O_2 ashing, 300W, 5min (100:1) HF, 10sec	Sink 7 Technics-c Sink 7	Polymer removal
3.06	Post cleaning	Piranha, 120°C, 20min	Sink 8	
3.07	Measurement	CD measuremt and Inspection	Leo	

Step	Process Name	Process Specification	Equipment	Comment
3.08	Trimming	LTO Trimming : (100:1) HF	Sink 7	
3.09	Measurement	CD measuremt and additive trimming	Leo/Tech.-c	
3.10	Silicon fin etch	TP=200W,BP=40W,13mT,CF_4=100, 30sec ME : TP=300W,BP=150W,15mT Cl_2:HBr=50:150, EDP+5sec	Lam5	
3.11	Post cleaning	(100:1) HF, 10sec Piranha, 120°C, 20min	Sink 7 Sink 8	
3.12	Measurement	CD Measurement and inspection	Leo	
3.13	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
3.14	Sacrificial oxidation	Dry, SGATEOX, 3min/900°C, T_{ox}=3nm	Tylan6	
3.15	Measurement	Sacrificial oxide thickness measurement	Nanoduv	
4.0	Gate Formation			
4.01	Pre cleaning	Piranha, 120°C, 20min (25:1) HF, 30sec	Sink 6 Sink 6	
4.02	Gate oxidation	Dry, THIN_ANN, 750°C, O_2, 12min/900°C, N_2, 30min	Tylan 6	
4.03	In-situ P+ $Si_{0.6}Ge_{0.4}$ Recipe: SIGEVAR.019	Nucleation : 550°C, 300mT, SiH_4=200,30sec Deposition : 450°C, 300mT, SiH_4=124,GeH_4=36 B_2H_6=40 (enter 80), 40min, 240nm	Tystar 19	
4.04	LTO deposition	LTO deposition, 11SULTOA, 3min 40sec, 70nm	Tystar 11	
4.05	Gate mask	E-beam lithography SAL-601, 200nm	e-beam nanowriter	at LBL
4.06	Ashing	O_2 ashing, 3W, 25sec	Technics-c	
4.07	Measurement	CD measuremt and additive ashing	Leo/Tech.-c	
4.08	LTO Etch	TP=200W,BP=40W,13mT,CF_4=100, 35sec	Lam5	
4.09	Resist strip	(100:1) HF, 10sec O_2 ashing, 300W, 5min (100:1) HF, 10sec	Sink 7 Technics-c Sink 7	Polymer removal
4.10	Post cleaning	Piranha, 120°C, 20min	Sink 8	
4.11	Measurement	CD measuremt and Inspection	Leo	
4.12	Trimming	LTO Trimming : (100:1) HF	Sink 7	
4.13	Measurement	CD measuremt and additive trimming	Leo/Tech.-c	
4.14	P+ Si0.6Ge0.4 gate etch	TP=200W,BP=40W,13mT,CF_4=100, 10sec ME : TP=300W,BP=150W,15mT Cl_2:HBr=50:150, EDP OE : TP=250W,BP=120W,15mT HBr:O_2=200:5, 20sec	Lam5	
4.15	Post cleaning	(100:1) HF, 10sec Piranha, 120°C, 20min	Sink 7 Sink 8	
4.16	Measurement	CD Measurement and inspection	Leo	

Step	Process Name	Process Specification	Equipment	Comment
5.0	Spacer Foramtion (Bi-layer)			
5.01	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
5.02	HTO deposition	9VHTOA, N_2O=90,DCS=18,300mT,800°C,10nm,12min	Tystar 9	
5.03	Measurement	HTO thickness measurement	Nanoduv	
5.04	Pre cleaning	Piranha, 120oC, 20min	Sink 6	
5.05	Nitride deposition	9SNITA,5min, 20nm	Tystar 9	
5.06	Measurement	Nitride thickness measurement	Nanoduv	
5.07	Spacer nitride etch	BT:TP=70W,BP=10W,13mT,CF_4=100, 10sec	Lam5	time etch
5.08		ME:TP=150W,BP=75W,15mT,Cl_2:HBr=50:150,80sec		time etch
5.09	Post cleaning	Piranha, 120°C, 20min	sink 8	
6.0	N+ & P+ S/D Formation			
6.01	N+ S/D mask	Resist coating : coat = prog#01/bake =prog #01	Svgcoat	
		Exposure (focus and expose test)	Gcaws	
		Development : bake = prog#01/develop = prog#01	Svgdev	
6.02	Hard bake	120°C, 30min	Vwr	
6.03	N+ S/D implantation	Phosphorus/5.0x$10^{15}$$cm^{-2}$/30KeV/0°	Implanter	foundry company
6.04	Resist strip	O_2 ashing, 300W, 5min	Technics-c	
6.05	Post cleaning	Piranha, 120°C, 20min	sink 8	
6.06	P+ S/D mask	Resist coating : coat = prog#01/bake =prog #01	Svgcoat	
		Exposure (focus and expose test)	Gcaws	
		Development : bake = prog#01/develop = prog#01	Svgdev	
6.07	Hard bake	120°C, 30min	Vwr	
6.08	P+ S/D implantation	Boron/5.0x$10^{15}$$cm^{-2}$/10KeV/0°	Implanter	foundry company
6.09	Resist strip	O_2 ashing, 300W, 5min	Technics-c	
6.10	Post cleaning	Piranha, 120°C, 20min	sink 8	
6.11	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
6.12	Recrystalization	N2ANN550, N_2, 600°C, 15 hours	Tylan7	
6.13	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
6.14	RTA	N_2, 900°C, 1min	Heatpulse3	
7.0	Selcetive Ge Deposition (Optional)			
7.01	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
7.02	Ge deposition	SELDEP.019, 350°C, 300mT, GeH_4=200, 10min, 100nm	Tystar19	
7.03	Measurement	Ge thickness measurement and selectivity check	Nanoduv/Leo	
7.04	DI rinse	3 cycle DI rinse	Sink 6	
7.05	Cap LTO deposition	11SULTOA, 450°C, 90sec, 30nm	Tystar 11	
7.06	Measurement	LTO thickness measurement	Nanoduv	
7.07	N+ S/D mask	Resist coating : coat = prog#01/bake =prog #01	Svgcoat	
		Exposure (focus and expose test)	Gcaws	
		Development : bake = prog#01/develop = prog#01	Svgdev	

Step	Process Name	Process Specification	Equipment	Comment
7.08	Hard bake	120°C, 30min	Vwr	
7.09	N+ S/D implantation	Phosphorus/5.0x$10^{15}$$cm^{-2}$/45KeV/0°	Implanter	foundry company
7.10	Resist strip	O_2 ashing, 300W, 5min	Technics-c	
7.11	Post cleaning	Piranha, 120°C, 20min	sink 8	
7.12	P+ S/D mask	Resist coating : coat = prog#01/bake =prog #01 Exposure (focus and expose test) Development : bake = prog#01/develop = prog#01	Svgcoat Gcaws Svgdev	
7.13	Hard bake	120°C, 30min	Vwr	
7.14	P+ S/D implantation	Boron/5.0x$10^{15}$$cm^{-2}$/15KeV/0°	Implanter	foundry company
7.15	Resist strip	O_2 ashing, 300W, 5min	Technics-c	
7.16	Post cleaning	Piranha, 120°C, 20min	sink 8	
7.17	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
7.18	RTA	N_2, 750°C, 1min	Heatpulse3	
8.0	Metallization			
8.01	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
8.02	LTO deposition	11SULTOA, 450°C, 8min, 150nm	Tystar 11	
8.03	Measurement	LTO thickness maesurement	Nanoduv	
8.04	Contact mask	Resist coating : coat = prog#01/bake =prog #01 Exposure (focus and expose test) Development : bake = prog#01/develop = prog#01	Svgcoat Gcaws Svgdev	
8.05	Hard bake	120°C, 30min	Vwr	
8.06	Contact etch	ME:TP=450W,BP=750W,20mT,CHF3=100 Ar=200, 35sec	Lam5	time etch
8.07	Resist strip	(100:1) HF, 20sec O_2 ashing, 300W, 5min (100:1) HF, 20sec	Sink 7 Technics-c Sink 7	Polymer removal
8.08	Post cleaning	Piranha, 120°C, 20min	sink 8	For Ge S/D, piranha skip
8.09	Pre cleaning	Piranha, 120°C, 20min (25:1) HF, 30sec	Sink 6 Sink 6	For Ge S/D, piranha skip
8.10	Al sputtering	Ar:300cc, 6mT, 15cm/min, one pass, 450nm	Cpa	
8.11	Metal mask	Resist coating : coat = prog#01/bake =prog #01 Exposure (focus and expose test) Development : bake = prog#01/develop = prog#01	Svgcoat Gcaws Svgdev	0.8*exposure time
8.12	Hard bake	120°C, 30min	Vwr	
8.13	Al etch	Al etchant, manual edn point detection with eye	sink 8	
8.14	Resist strip	O_2 ashing, 300W, 5min	Technics-c	
8.15	DI Rinse	3 cycle DI rinse	Sink 8	
8.16	Sintering	VSINT400, 400°C,30min, N_2:H_2=10:1	Tylan 13	

6.6.2 Process Flows of Spacer FinFETs

Step	Process Name	Process Specification	Equipment	Comment
1.0	Wafers	4" p-type SOI		T_{Si}=100nm
2.0	Body Thinning and Fin Formation by Spacers			
2.01	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
2.02	Body thinning	Wet, SWETOXB, 850°C, 2hour, 120nm	Tylan 2	
2.03	Oxide removal	(10:1) HF, 10min	Sink 6	
2.04	Measuement	Body thickness measurement	Nanoduv	T_{Si}=50nm
2.05	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
2.06	Pad oxidation	Dry, SGATEOX, 4min/900°C, T_{ox}=4nm	Tylan6	
2.07	Measuement	Pad oxide thickness measurement	Nanoduv	
2.08	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
2.09	Nitride deposition	9SNITA, 12min 30sec, 50nm	Tystar 11	
2.10	Measuement	Nitride thickness measurement	Nanoduv	
2.11	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
2.12	Sacrificial $Si_{0.4}Ge_{0.6}$ deposition Recipe: SIGEVAR.019	Nucleation : 550°C, 300mT, SiH_4=200,30sec Deposition : 450°C, 300mT, SiH_4=124,GeH_4=80 18min, 200nm	Tystar 19	
2.13	Measurement	SiGe thickness measurement	Nanoduv	
2.14	Spacer mask	Resist coating : coat = prog#01/bake = prog#01 Exposure Development : bake = prog#01/develop = prog#01	Svgcoat gcaws Svgdev	
2.15	Poly-$Si_{0.4}Ge_{0.6}$ etch (Sacrificaial layer)	BT : TP=200W,BP=40W,20mT CHF_3:Ar=90:200, 10sec ME : TP=300W,BP=150W,15mT Cl_2:HBr=50:150, EDP+10sec	Lam5	
2.16	Resist strip	(100:1) HF, 10sec O_2 ashing, 300W, 5min (100:1) HF, 10sec	Sink 7 Technics-c Sink 7	Polymer removal
2.17	Post cleaning	Piranha, 120°C, 20min	sink 8	
2.18	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
2.19	Spacer HTO deposition	9VHTOA, N_2O=90, DCS=18, 300mT, 800°C, 39min, 30nm	Tylan9	
2.20	Measurement	HTO thickness measurement	Nanoduv	
2.21	Alignment Key Protection Mask	Resist coating : coat = prog#01/bake = prog#01 Exposure Development : bake = prog#01/develop = prog#01	Svgcoat gcaws Svgdev	To leave SiGe pattern as an alignment key
2.22	HTO and SiGe etch	PSG:TP=200W,BP=40W,20mT, CF4=100, 30sec SiGe:TP=250W,BP=120W,35mT, HBr:O_2=200:5 1min	Lam5	

Step	Process Name	Process Specification	Equipment	Comment
2.23	Resist strip	(100:1) HF, 10sec O_2 ashing, 300W, 5min (100:1) HF, 10sec	Sink 7 Technics-c Sink 7	Polymer removal
2.24	Post cleaning	Piranha, 120°C, 20min	Sink 8	
2.25	SiGe residue removal	H_2O:NH_4OH:H_2O_2=(5:1:1) at 75°C, 10sec		
2.26	Inspection	SEM inspection	Leo	
2.27	Active S/D Mask	Resist coating : coat = prog#01/bake = prog#01 Exposure Development : bake = prog#01/develop = prog#01	Svgcoat gcaws Svgdev	
2.28	Ashing	O_2 ashing, 30W, 7min	Technics-c	
2.29	Measurement	CD measuremt and additive ashing	Leo/Tech.-c	
2.30	Fin etch	1st ME:TP=200W, BP=40W, 13mT, CF_4=100, 40sec 2nd ME:TP=300W, BP=150W, 15mT Cl_2:HBr=50:150, 10sec	Lam5	
2.31	Resist strip	(100:1) HF, 10sec O_2 ashing, 300W, 5min (100:1) HF, 10sec	Sink 7 Technics-c Sink 7	Polymer removal
2.32	Post cleaning	Piranha, 120°C, 20min	Sink 8	
2.33	Measurement	CD measuremt	Leo	
2.34	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
2.35	Sacrificial oxidation	Dry, SGATEOX, 3min/900°C, T_{ox}=3nm	Tylan6	
2.36	Measurement	Sacrificial oxide thickness measurement	Nanoduv	
3.0	Gate Formation			
3.01	Pre cleaning	Piranha, 120°C, 20min (25:1) HF, 30sec	Sink 6 Sink 6	
3.02	Gate oxidation	Dry, THIN_ANN, 750°C, O_2, 14min/900°C, N_2, 30min	Tylan 6	
3.03	In-situ P+ $Si_{0.6}Ge_{0.4}$ Recipe: SIGEVAR.019	Nucleation : 550°C, 300mT, SiH_4=200,30sec Deposition : 500°C, 300mT, SiH_4=167,GeH_4=53 1 hour 40min, 1.5um	Tystar 19	
3.04	CMP	14min/8 psi(Down force)/24 rpm(table)/6 rpm(chuck) 1 psi (back pressure)/30°C/ 100 ml/min(Slurry)	Cmp	T_{SiGe}=400nm
3.05	Post cleaning	DI rinse 1min NH4OH 1min DI rinse 1min Piranha, 120°C, 1min DI rinse 1min (5:1) HF 10sec DI rinse 1min (5:1:1) H_2O:NH_4OH:H_2O_2 at 65°C, 5min DI rinse 1min	Manual Sink7 Sink7 Sink8 Sink8 Sink8 Sink8 Manual Sink8	
3.06	N+ S/D mask	Resist coating : coat = prog#01/bake =prog #01 Exposure (focus and expose test) Development : bake = prog#01/develop = prog#01	Svgcoat Gcaws Svgdev	

Step	Process Name	Process Specification	Equipment	Comment
3.07	Hard bake	120°C, 30min	Vwr	
3.08	N+ S/D implantation	Phosphorus/5.0x10^{15}cm^{-2}/150KeV/0°	Implanter	foundry company
3.09	Resist strip	O_2 ashing, 300W, 5min	Technics-c	
3.10	Post cleaning	Piranha, 120°C, 20min	sink 8	
3.11	P+ S/D mask	Resist coating : coat = prog#01/bake =prog #01 Exposure (focus and expose test) Development : bake = prog#01/develop = prog#01	Svgcoat Gcaws Svgdev	
3.12	Hard bake	120°C, 30min	Vwr	
3.13	P+ S/D implantation	Boron/5.0x10^{15}cm^{-2}/60KeV/0°	Implanter	foundry company
3.14	Resist strip	O_2 ashing, 300W, 5min	Technics-c	
3.15	Post cleaning	Piranha, 120°C, 20min	sink 8	
3.16	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
3.17	LTO deposition	LTO deposition, 11SULTOA, 5min 20sec, 100nm	Tystar 11	
3.18	Measurement	LTO thickness measurement	Nanoduv	
3.19	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
3.20	Annealing	HIN2ANNL, N_2, 900°C, 30min	Tylan7	
3.21	Gate mask	Resist coating : coat = prog#01/bake =prog #01 Exposure (focus and expose test) Development : bake = prog#01/develop = prog#01	Svgcoat Gcaws Svgdev	
3.22	Ashing	O_2 ashing, 30W, 7min	Technics-c	
3.23	Measurement	CD measuremt and additive ashing	Leo/Tech.-c	
3.24	LTO Etch	TP=200W,BP=40W,13mT,CF_4=100, 50sec	Lam5	
3.25	Resist strip	(100:1) HF, 10sec O_2 ashing, 300W, 5min (100:1) HF, 10sec	Sink 7 Technics-c Sink 7	Polymer removal
3.26	Post cleaning	Piranha, 120°C, 20min	Sink 8	
3.27	Measurement	CD measuremt and Inspection	Leo	
3.28	Trimming	LTO Trimming : (100:1) HF	Sink 7	
3.29	Measurement	CD measuremt and additive trimming	Leo/Tech.-c	
3.30	P+ Si0.6Ge0.4 gate etch	TP=200W,BP=40W,13mT,CF_4=100, 10sec ME : TP=300W,BP=150W,15mT Cl_2:HBr=50:150, EDP OE : TP=250W,BP=120W,15mT HBr:O_2=200:5, 20sec	Lam5	
3.31	Post cleaning	(100:1) HF, 10sec Piranha, 120°C, 20min	Sink 7 Sink 8	
3.32	Measurement	CD Measurement and inspection	Leo	
4.0	Spacer Foramtion (Bi-layer)			
4.01	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
4.02	HTO deposition	9VHTOA, N_2O=90,DCS=18,300mT,800°C,10nm 12min	Tystar 9	
4.03	Measurement	HTO thickness measurement	Nanoduv	

Step	Process Name	Process Specification	Equipment	Comment
4.04	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
4.05	Nitride deposition	9SNITA,5min, 20nm	Tystar 9	
4.06	Measurement	Nitride thickness measurement	Nanoduv	
4.07	Spacer nitride etch	BT:TP=70W,BP=10W,13mT,CF_4=100, 10sec ME:TP=150W,BP=75W,15mT,Cl_2:HBr=50:150,2min	Lam5	time etch time etch
4.08	Post cleaning	Piranha, 120°C, 20min	sink 8	
5.0	N+ & P+ S/D Formation			
5.01	N+ S/D mask	Resist coating : coat = prog#01/bake =prog #01 Exposure (focus and expose test) Development : bake = prog#01/develop = prog#01	Svgcoat Gcaws Svgdev	
5.02	Hard bake	120°C, 30min	Vwr	
5.03	N+ S/D implantation	Phosphorus/5.0x$10^{15}$$cm^{-2}$/30KeV/0°	Implanter	foundry company
5.04	Resist strip	O_2 ashing, 300W, 5min	Technics-c	
5.05	Post cleaning	Piranha, 120°C, 20min	sink 8	
5.06	P+ S/D mask	Resist coating : coat = prog#01/bake =prog #01 Exposure (focus and expose test) Development : bake = prog#01/develop = prog#01	Svgcoat Gcaws Svgdev	
5.07	Hard bake	120°C, 30min	Vwr	
5.08	P+ S/D implantation	Boron/5.0x$10^{15}$$cm^{-2}$/10KeV/0°	Implanter	foundry company
5.09	Resist strip	O_2 ashing, 300W, 5min	Technics-c	
5.10	Post cleaning	Piranha, 120°C, 20min	sink 8	
5.11	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
5.12	Recrystalization	N2ANN550, N_2, 600°C, 15 hours	Tylan7	
5.13	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
5.14	RTA	N_2, 9000°C, 1min	Heatpulse3	
5.15	Pre cleaning	Piranha, 120°C, 20min	Sink 6	
5.16	Sintering	VSINT400, 400°C,30min, N_2:H_2=10:1	Tylan 13	

Chapter 7

Conclusion

7.1 Summary

The silicon-based microelectronics industry has been growing exponentially according to Moore's law for past three decades. An essential key to this progress has been miniaturization [7.1]-[7.3]. The fabrication technology for CMOS devices has recently entered into the nanoscale era, and a fundamental change in device architecture is necessary to continue the projected scaling trends. Thus, novel structures is required to overcome the current scaling issues in nanoscale CMOS and break through the current ITRS roadmap.

Ideal MOSFETs have high drive current when the gate electrode is biased to turn the transistor on and low leakage current when the gate electrode is biased to turn the transistor off. As the MOSFET channel length is reduced to 30nm and below, the suppression of off-state leakage current becomes an increasingly difficult technological challenge; one that will ultimately limit the scalability of the conventional MOSFET structure.

Ultra-thin body (UTB) single-gate MOSFET [7.5][7.6] and a double-gate FinFET [7.7]-[7.9] structures are proposed to suppress short-channel effects and extend CMOS scaling to the ultimate limit of silicon. To support device fabrication of UTBFET and FinFETs, nano-lithography technologies and novel process technologies have been proposed and demonstrated. To define sub-20nm patterns, two lithographic technologies: ashing-trimming [7.10] and spacer lithography [7.11][7.14] have been demonstrated. The combination of resist ashing with oxygen plasma and hard mask oxide trimming with diluted HF reduced a 500nm wide-pattern to less than 20nm wide. Spacer technology provides sub-lithographic features and very uniform pattern widths because the patterns are defined not by conventional lithography but by the deposited thin hard mask film, which is used to define the gate in UTBFETs and the fin in FinFETs. In the case of FinFETs, spacer lithography technology doubles the device density.

Three process technologies: resist etch-back, selective germanium deposition, and chemical-mechanical polishing (CMP), are proposed and demonstrated. A resist etch-back is applied to make a raised poly-Si S/D on a thin body SOI device [7.5]. Germanium is selectively deposited on a thin silicon body with a conventional LPCVD furnace. It is applied to make a raised Ge S/D on the thin body of a UTBFET (horizontal device) [7.6] and on the narrow fin of a FinFET (vertical device) [7.12]. This process is compatible with a metal gate and high-K gate dielectric because of the low temperature process for deposition and RTA. A CMP process is used to make a planarized gate over single- or multiple-Si fins of FinFETs and a completely planarized poly-SiGe surface is achieved [7.11][7.14]. The process windows for lithography and plasma etch are significantly

improved with CMP. A notched gate resist profile is not observed, and stringers or residues of the gate hard mask oxide and gate poly-SiGe are not found after CMP.

Ultra-thin body field effect transistors (UTBFETs) with a raised S/D are proposed. Two methods to make the raised S/D are to etch back the resist and the poly-silicon and to deposit germanium selectively. The greatest benefits of this process are its simplicity and low cost. Simulations showed that UTBFETs could be scaled down to 18nm with 5nm body thickness and even with relatively thick gate oxide (1.5nm) for a sub-20nm gate length. UTBFETs showed excellent short-channel effects and low off-state current and high on-state drive current. Companies such as TSMC and Intel are looking to UTBFETs and fabricating them for the future.

Quantum-mechanical effects such as the threshold voltage shift and mobility enhancement due to the confinement of inversion charges have been observed at room temperature. An analytical model to explain the threshold voltage shift in the UTBFET has been proposed and compared with the experimental data for the first time [7.5][7.13]. Enhancement in hole mobility has been experimentally observed for body thickness below 5nm and provided an explanation [7.6].

Quasi-planar and simplified CMOS FinFETs with e-beam lithography and spacer lithography are proposed and demonstrated for the first time [7.14]. A 10nm fin with sub-20nm gate length has been successfully fabricated (This is the world's smallest double-gate CMOS device). Both e-beam FinFETs and spacer FinFETs show excellent off-state leakage current and on-state drive current. Planar FinFETs are the most attractive

structures for the double gates, and they allow aggressive device scaling to the regimes of "no known solution" and "beyond roadmap" in the ITRS roadmap [7.4]. These structures will enable device scaling to continue at its historical rate for another 15 years, to the limit of silicon-based semiconductors. FinFETs developed at the University of California-Berkeley have been transferred to TSMC, AMD, Motorola, and IBM. Further research for FinFETs is now being pursued by large companies in the industry.

7.2 Suggestions for Future Work

7.2.1 Mobility Enhancement

One drawback of FinFETs is that the channel is formed by lithography and etch. Thus, there may be lots of etch damage and large CD variation. To improve the smoothness of the sidewall of the silicon fin, three methods are proposed. They are hydrogen annealing at high temperature, sacrificial oxidation with diluted oxygen at high temperature (>1000°C), and argon annealing (900°C, >10sec) after sacrificial oxidation. Also, etch roughness can be alleviated by changing the lithography condition and etch recipe used to form the silicon fin.

7.2.2 Threshold Voltage Adjustment

In-situ boron doped $Si_{0.4}Ge_{0.6}$ gives a reasonable NMOS and PMOS threshold voltages in simulations. However, experimental devices suffered from counter-doping of the body due to boron penetration. Boron penetration can be minimized by adoption of

oxy-nitride as the gate dielectric material and a reduction of the total thermal budget after the gate formation. Alternatively, a metal gate can be used. Dual metal gate processes such as Mo gate with nitrogen implantation [7.15] and metal inter-diffusion of nickel-Ti [7.16] are good candidates and TiN gate [7.17] with suitable body doping is acceptable. However, these approaches are complicated and reliability issues are not verified yet. Thus, asymmetrical double-gate is a promising structure to tune the threshold voltage because the threshold voltage is adjusted by the ratio of the silicon body thickness to the gate oxide thickness [7.18], as expressed by

$$V_t = \left(\frac{\frac{C_{bg}}{C_{ox}}}{1+\frac{C_{bg}}{C_{ox}}} \right) E_g \qquad (7.1)$$

where C_{ox} is the front gate capacitance, C_{bg} is the back gate capacitance, and E_g is the energy band gap of gate material. Detailed process flows are proposed in the following section.

7.2.3 Asymmetrical Double-Gate FinFETs

The FinFET structure can be used to make an asymmetrical double-gate easily by using a large angle tilted implantation. After fin formation and gate oxidation, undoped poly-silicon is deposited as shown in Fig. 7.1 (a). To minimize the counter-doped area, the tilted angle for implantation should be as large as possible. The top part of the poly-silicon at the fin is etched by etch-back process in order to minimize the counter-diffusion through the poly-silicon as shown in Fig. 7.1 (b). TiN is deposited and connects the N+ poly-silicon and P+ poly-silicon because dopant diffusivity through TiN is

negligible as shown in Fig 7.1 (c). Finally, thicker N+ poly-silicon is deposited again for interconnection and planarized with CMP as shown in Fig. 7.1 (d).

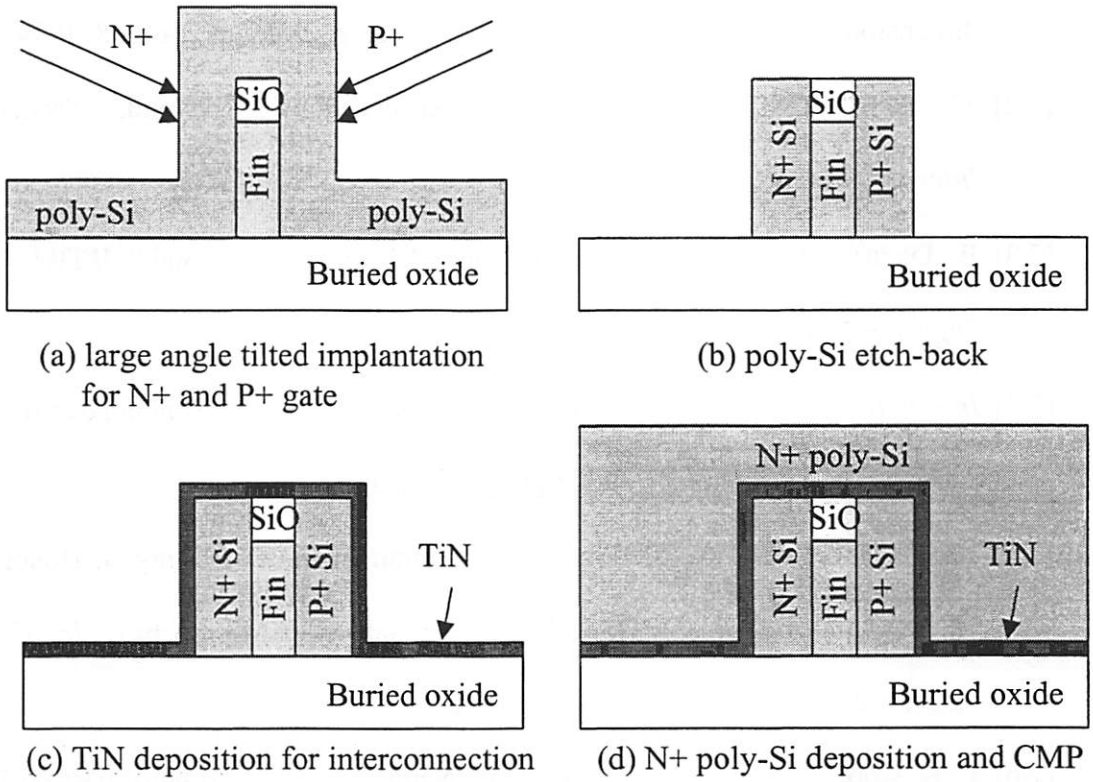


Figure 7.1 Cross-sectional diagrams showing the process flow of asymmetrical double-gate FinFETs.

The critical issues for double-gate FinFETs are to adjust the threshold voltage to a CMOS-compatible value, to obtain multiple threshold voltages for circuit implementation, and to improve the NMOS on-state drive current. With significant progress in the above three challenges, double-gate FinFETs will break through the 10nm-barrier.

7.3 References

[7.1] R.H. Dennard, F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassous and A. R. LeBlanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE Journal of Solid-State Circuits*, vol.9, p.256-268, 1974.

[7.2] C. Hu, "MOSFET scaling in the next decade and beyond," *Semiconductor International*, p.105-114, June 1994.

[7.3] B. Davari, "CMOS technology scaling: 0.1 μm and beyond," *IEDM Technical Digest*, p.555-558, 1996.

[7.4] *International Technology Roadmap for Semiconductors*, Semiconductor Industry Association, 2000. http://public.itrs.net/Home.htm

[7.5] Y.-K. Choi, K. Asano, N. Lindert, V. Subramanian, T.-J. King, J. Bokor, C. Hu, "Ultra-thin-body SOI MOSFET for Deep-sub-tenth Micron Era," *IEEE Electron Device Letters,* vol. 21, p.254-255, 2000.

[7.6] Y.-K. Choi, D. Ha, T.-J. King, C. Hu, "Nanoscale ultrathin body PMOSFETs with raised selective germanium source/drain", *IEEE Electron Device Letters*, vol.22, p.447-8, 2001.

[7.7] D. Hisamoto, T. Kaga, E. Takeda, "Impact of the vertical SOI 'DELTA' structure on planar device technology", *IEEE Transactions on Electron Devices*, vol.38, p.1419-24, 1991.

[7.8] X. Huang, W.-C. Lee, C. Kuo, D. Hisamoto, L. Chang, J. Kedzierski, E. Anderson, H. Takeuchi, Y.-K. Choi, K. Asano, V. Subramanian, T.-J. King, J. Bokor, C. Hu, "Sub-50 nm P-channel FinFET", *IEEE Transactions on Electron Devices*, vol.48, p.880-6, 2001.

[7.9] N. Lindert, L. Chang, Y.-K. Choi, E. H. Anderson, W.-C. Lee, T.-J. King, C. Hu, "Sub-60-nm Quasi-Planar FinFETs Fabricated Using a Simplified Process", *IEEE Electron Device Letters*, vol.22, p.487-9, 2001.

[7.10] K. Asano, Y.-K. Choi, T.-J. King, C. Hu, "Patterning Sub-30-nm MOSFET Gate with I-line Lithography," *IEEE Transactions on Electron Devices,* vol. 48, p.1004-6, 2001.

[7.11] Y.-K. Choi, T.-J. King, C. Hu, "Nanoscale CMOS Spacer FinFET for the Terabit Era", *IEEE Electron Device Letters*, to be published, 2002.

[7.12] N. Lindert, Y.-K. Choi, L. Chang, E. Anderson, W.-C. Lee, T.-J. King, J. Bokor, C. Hu, "Quasi-Planar FinFETs with Selectively Grown Germanium Raised Source/Drain", *IEEE International SOI Conference*, p.111-2, 2001.

[7.13] Y.-K. Choi, D. Ha, T.-J. King, C. Hu, "Threshold voltage Shift by Quantum Confinement in Ultra-thin Body Device", *59th Device Research Conference*, p.85-6, 2001.

[7.14] Y.-K. Choi, N. Lindert, P. Xuan, S. Tang, D. Ha, E. Anderson, T.-J. King, J. Bokor, C. Hu, "Sub-20nm CMOS FinFET Technologies", *IEEE IEDM Technical Digest*, p.421-24, 2001.

[7.15] P. Ranade, H. Takeuchi, T. -J. King and C. Hu, "Work Function Engineering of Molybdenum Gate Electrodes by Nitrogen Implantation", *Electrochemical and Solid State Letters*, vol.4, p.G85-G87, 2001.

[7.16] I. Polishchuk, P. Ranade, T.-J. King, C. Hu, "Dual Work Function Metal Gate CMOS Technology Using Metal Interdiffusion", *IEEE Electron Device Letters*, vol.22, p. 444-6, 2001.

[7.17] H. Wakabayashi, Y. Saito, K. Takeuchi, T. Mogami, T. Kunio, "A novel W/TiN_x metal gate CMOS technology using nitrogen-concentration-controlled TiN_x film", *IEEE IEDM Technical Digest*, p.253-6, 1999.

[7.18] S. Tang, Doctoral Thesis, "Dynamic Threshold MOSFETs for Future Integrated Circuits", Department of Electrical Engineering and Computer Sciences, *University of California at Berkeley*, 2001.