Compact Modeling of Nanoscale CMOS

Chung-Hsun Lin

Electrical Engineering and Computer Sciences
University of California at Berkeley

December 19, 2007

Compact Modeling of Nanoscale CMOS

by

Chung-Hsun Lin

B.S. (National Taiwan University, Taiwan) 1999
M.S. (National Taiwan University, Taiwan) 2001

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering – Electrical Engineering
and Computer Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Chenming Hu, Chair
Professor Ali M. Niknejad
Professor Oscar Dubon

Fall 2007

The dissertation of Chung-Hsun Lin is approved:

_____

Prof. Chenming Hu, Chair   Date

_____

Prof. Ali M. Niknejad   Date

_____

Prof. Oscar Dubon   Date

University of California, Berkeley

Fall 2007

Compact Modeling of Nanoscale CMOS

Copyright 2007

by

Chung-Hsun Lin

Abstract

Compact Modeling of Nanoscale CMOS

by

Chung-Hsun Lin

Doctor of Philosophy in Engineering –
Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Chenming Hu, Chair

The scaling of bulk CMOS technology results in incredible reduction of cost-per-computation with higher computational performance. It is conventionally achieved through the combination of thinner gate oxide, higher effective channel doping and shallower source/drain junction depth. However, these techniques are rapidly approaching material and process limits. The multi-gate FET such as FinFET has emerged as the most promising candidate to extend CMOS scaling beyond the 22nm technology node. The strong electrostatic control over the channel originating from the use of multiple gates reduces the coupling from source and drain. It enables the multi-gate transistor to be scaled beyond bulk planar CMOS for a given dielectric thickness.

A compact model serves as a link between process technology and circuit design. It maintains a fine balance between accuracy and simplicity. An accurate model based on physics allows the process engineers and circuit designers to make projections beyond the available silicon data for scaled dimensions and also enables fast circuit and device co-optimization. It is thus necessary to develop a compact model of multi-gate FETs for

technology and circuit development in the short term and for product design in the longer term.

Full scale multi-gate FET compact models are developed. Unique device physics in the multi-gate MOSFET due to extra gates are studied and investigated. Modeling methodologies are proposed to incorporate these unique multi-gate physics in the compact model. Different flavors of the multi-gate FETs are modeled in two categories: the symmetric/common-gate multi-gate FETs and the independent/asymmetric multi-gate FETs. The complete multi-gate compact models are verified with TCAD simulation results and experimental data.

The performance comparison and design concepts of multi-gate-based logic and memory circuits are studied using the BSIM-MG model. The impact of the process variation can be tuned out by using back-gate tuning. The impact of back-gate length and misalignment on the threshold voltage is discussed.

In the sub-45nm CMOS technology regime, the impact of device variations on circuit functionality becomes critical. The scaling of the device geometry makes device characteristics more sensitive to the fluctuation of process steps. A novel methodology for generating Performance Aware Models (PAM) cards is presented for accurately predicting the statistical variations of VLSI circuit performance due to process variation. The PAM cards also improve the accuracy of Monte Carlo simulations by reconciling the physical and electrical-test variances.

_____

Professor Chenming Hu

Dissertation Committee Chair

2

To my family, for the their love, encouragement, and support

Table of Contents

Acknowledgements

First and foremost, I would like to express my gratitude to Professor Chenming Hu for being my Ph.D. research advisor during my Ph.D. study at U.C. Berkeley. His enthusiasm, guidance and wealth of technical knowledge provided excellent inspiration and insight to my research. I always felt excited and refreshed after our group and individual meetings. He was also an outstanding mentor while I was a teaching assistant for EE130. I am very grateful for his valuable suggestions and help regarding many other aspects of my life, especially being my witness at my wedding at San Francisco City Hall. It is my great pleasure and honor to have known him and his family.

I would also like to thank Professor Ali Niknejad for co-directing the BSIM project, serving as the chair of my qualifying exam committee, and the member of my dissertation committee. Our technical discussions were always rich and fruitful. His insights on RF applications and device modeling have inspired me a lot during the past years. Thanks to Professor Vivek Subramanian for serving as the member in my qualifying exam committee, providing valuable discussion on device physics, and mentoring the Cal VIEW consultant task. I am also grateful to Professor Oscar Dubon for serving as the member in my qualifying exam committee and dissertation committee. Thanks to Professor Tsu-Jae King Liu for her helpful discussion regarding my research.

I am grateful to Professor Mansun Chan of Hong Kong UST for serving as my research co-advisor during my first year at Berkeley when Professor Hu was on sabbatical leave. Thanks to him for leaving me his nice car to me when he went back to Hong Kong.

I am indebted to many former and current members of the device group in 373

Chapter 1

Introduction

1.1 CMOS Technology Scaling and Challenges

Throughout recent history, silicon-based mircoelectronics has experienced tremendous growth and performance improvements since the innovative concept of integrated circuit (IC) was invented by J. Kilby in late 1950's [1.1]. The computational power is enhanced at a tremendous rate with chpeaper cost, resulting in incredible reduction of cost-per-computation with higher computational performance. In 1965, Gordon Moore made a very famous and important oberservation that the complexity of ICs approximately doubles every year (Moore later refined the period to two years) [1.2]. This obeservation is the well known "Moore's Law" [1.3]. Over the past four decades, the

scaling of the conventional metal-oxide-semiconductor field-effect transistor (MOSFET) has been accomplished with technology innovations and led the device dimensions well into the nanometer era as shown in Fig. 1.1.
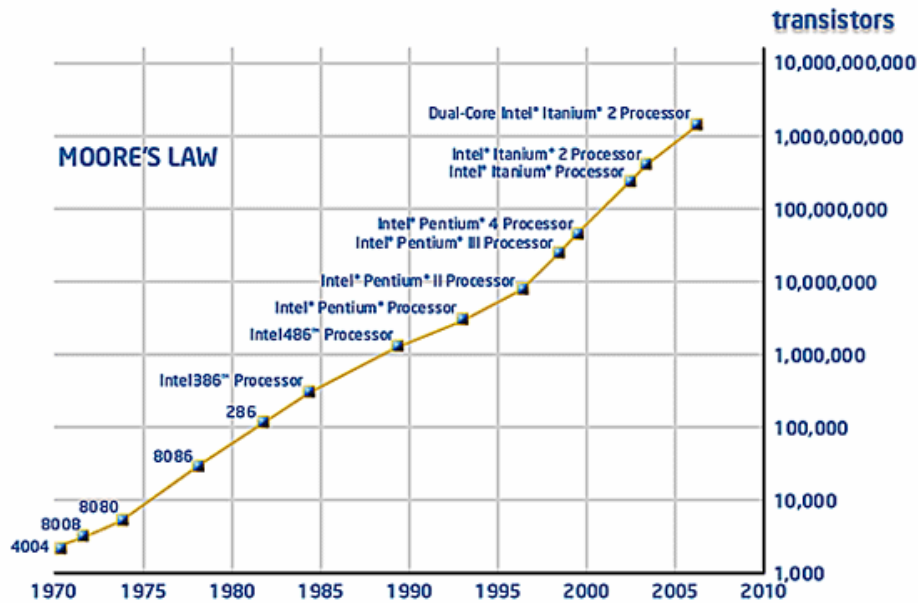
Fig. 1.1 Moore's Law has delivered exponential increases in the number of transistors integrated into microprocessors and other leading platform ingredients. [Source: Intel Corporation]

The scaling of MOSFET is conventionally achieved through the combination of thinner gate oxide, higher effective channel doping, and shallower source/drain junction depth [1.4-1.5]. However, the performance improvement by scaling the dimension of MOSFET is approaching a limit. The channel potential control from the gate electrode

degrades as the gate length decreases. The potential penetration from the drain electrode leads to severe short channel effects (SCE), such as threshold voltage roll-off, subthreshold swing degradation, and drain-induced barrier lowering (DIBL). SCE leads to significant increase of off-state drain current, which will limit the scaling of MOSFET due to significant leakage power consumption. The thinner gate oxide leads to an exponentially increased leakage current due to direct tunneling through the dielectric [1.6]. The higher effective channel doping degrades the carrier mobility and increases source/drain junction leakage. The higher channel doping also induces doping fluctuation, and thus threshold voltage variation, in the nanoscale transistor [1.7]. The formaiton of shallower source/drain junction depth is limited due to the thermal budget of dopant activation [1.8].

To overcome the above issues, the mobility-enhanced technology and high-K/metal-gate are introduced to improve the on-state drive current without degrading off-state leakage current [1.9, 1.10]. To further maintain the performance improvement by scaling the device dimension, the advanced transistor structures such as ultra-thin-body (UTB) MOSFET and multi-gate (MG) MOSFET are expected to be introduced in the future technology nodes.

1.2 Advanced MOSFET Structures

As mentioned in the previous section, the off-state leakage current increases as gate length decreases. The leakage path is located far from the dielectric/channel interface, which is least effectively controlled by the gate electrode. Advanced MOSFET structures utilize the thin body to suppress the off-state leakage path [1.11, 1.12]. The short channel behavior is controlled by the thickness of the thin body instead of the channel doping. Therefore, the channel can be lightly doped. The lightly doped channel of advanced MOSFET structures provides several advantages compared to the conventional planar MOSFET:

(i)   Improved mobility due to lower vertical field and less columbic scattering

(ii)   Improved subthreshold swing due to better control of SCE

(iii) Less parasitic junction capacitance

However, the fabrication of such advanced structures has difficulties of process integration. Maintaining uniformity of thin-body thickness over the entire wafer is very challenging. The series resistance could be potentially problematic due to the nature of ultra-thin film.

Multi-gate FETs such as FinFETs have emerged as the most promising candidates to extend the CMOS scaling beyond 22nm technology node [1.13]. FinFET is known to be

the most manufacturable thin-body devices due to self-aligned gate electrodes which are compatible with conventional planar bulk CMOS process. The FinFET can be built on either SOI or bulk silicon substrate. The strong electrostatic control over the channel originating from the use of multiple gates reduces the coupling between source and drain in the subthreshold region and it enables the multi-gate transistor to be scaled beyond bulk planar CMOS for a given dielectric thickness. Fig. 1.2 shows examples of advanced thin-body MOSFET structures.
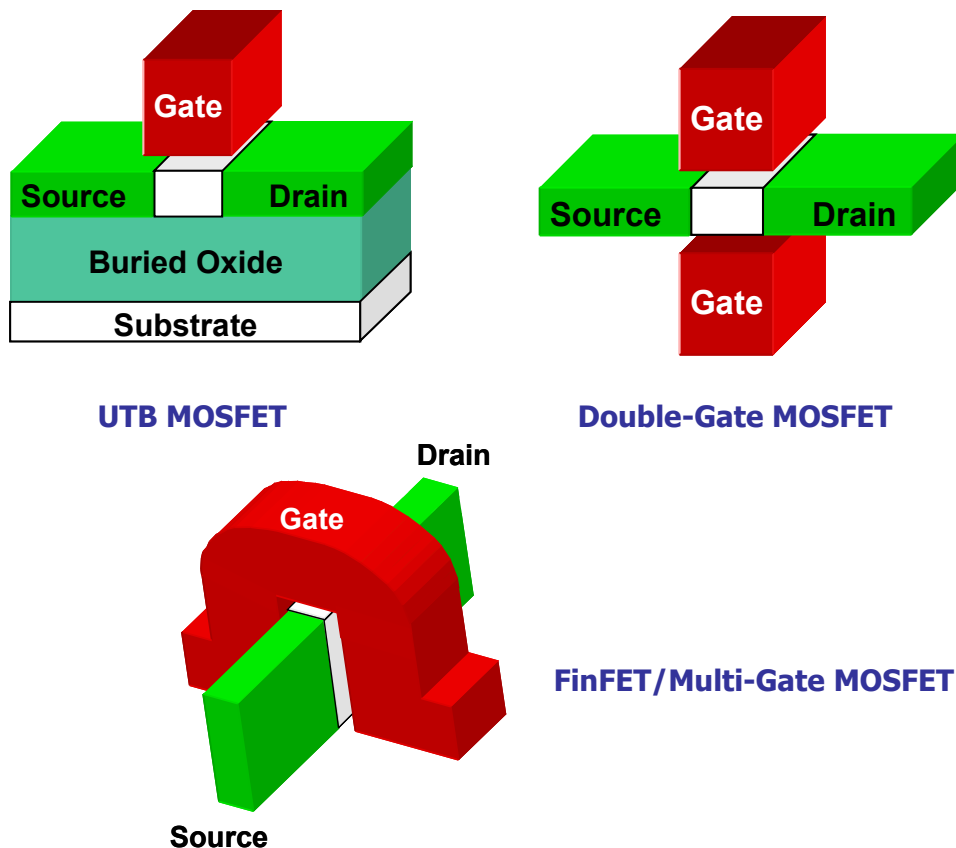
Fig. 1.2 Examples of advanced thin-body MOSFET structures

1.3 Compact Modeling of Multi-Gate MOSFETs

A compact model serves as a link between process technology and circuit design. It is a concise mathematical description of the complex device physics in the transistor. A compact model maintains a fine balance between accuracy and simplicity. An accurate model based on physics allows the process engineer and circuit designer to make projections beyond the available silicon data (scalability) for scaled dimensions and also enables fast circuit and device co-optimization. The simplifications in the physics enable very fast analysis of device/circuit behavior compared to the much slower numerical based TCAD simulations. It is thus necessary to develop a compact model of multi-gate FETs for technology/circuit development in the short term and for product design in the longer term.

One of the biggest challenges in modeling multi-gate FETs is the need to model several flavors of multi-gate FETs. The silicon body can be controlled by either two gates or three gates or four gates. The gates can all be electrically interconnected or they can be biased independently. Multi-gate FETs can be built on SOI or bulk silicon. It is important to obtain a versatile model which can model all the different types of multi-gate FETs without making the model computationally intensive. Fig. 1.3 shows the development cycle for a compact model [1.14].

```
                              ┌─────────────────┐
                              │      START       │
                              └─────────────────┘
                                       │
                                       ▼
                    ┌──────────────────────────────────┐
                    │  Examine Device Behavior          │
                    │    ●  Embedded physics (TCAD)      │
                    │    ●  Silicon based data           │
                    │       (Measurement)                │
                    └──────────────────────────────────┘
                                       │
                                       ▼
                    ┌──────────────────────────────────┐
                    │  Compact Model Formulation        │
                    │    ●  Develop model equations      │
                    │       (I-V, C-V)                   │
                    │    ●  Add model parameters for     │
                    │       flexibility                  │
                    │    ●  Develop extraction routine   │
                    └──────────────────────────────────┘
                                       │
                                       ▼
           No       ┌──────────────────────────────────┐
                    │  Model Execution                  │
                    │    ●  Numerically robust?          │
                    │    ●  Computationally efficient?   │
                    └──────────────────────────────────┘
                                       │
                                       ▼
           No       ┌──────────────────────────────────┐
                    │  Silicon Data Fitting             │
                    │    ●  Accurate description?        │
                    │    ●  Scalability over geometry,   │
                    │       bias, and temperature?       │
                    └──────────────────────────────────┘
                                       │
                                       ▼
                              ┌─────────────────┐
                              │       END        │
                              └─────────────────┘
```
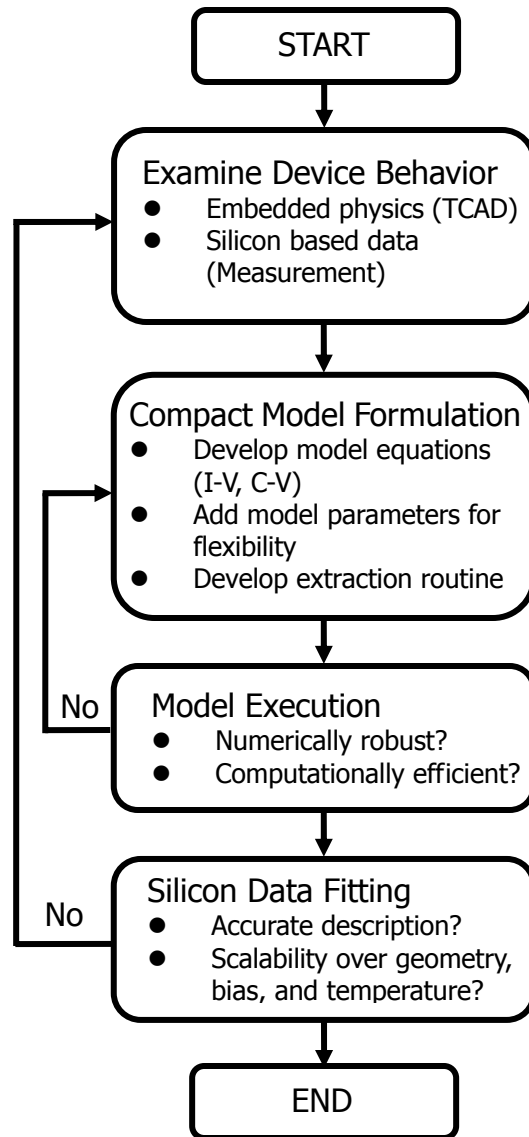
Fig. 1.3 The development cycle for a compact model [1.14]

1.4 Process Variations in Compact Models

In the sub-45nm CMOS technology regime, the impact of device variations on circuit

functionality becomes critial. The scaling of the device geometry makes device characteristics more sensitive to the fluctuation of process steps. The control of the critical diemension such as gate length continues to be a difficult challenge since the physical gate length is considerably smaller than the lithography printed line width in the nanoscale regime. The fluctuations during the processing lead to variability in device characteristics both within the die and between the dies. This leads to greater variance of device/circuit performance around the nominal technology node of circuit design. Statistical modeling techniques are essential for estimating circuit yields, designing manufacturable and robust systems. Therefore, statistical compact modeling for considering process variation becomes more important than ever.

1.5 Objectives

In this dissertation, compact models are developed for multi-gate MOSFET and statistical process variations. The unique device physics of multi-gate MOSFET are studied and modeled for inclusion in the multi-gate compact model. The circuit design concept of the advanced MOSFET structures are studied and discussed using the developed compact model.

Chapter 2 presents the proposed modeling methodologies and approaches to

incorporate these unique multi-gate physics in the compact model. The proposed models are verified with 2-D and 3-D TCAD simulation results.

In Chapter 3, the core formulation and physical effects of real device modeled in BSIM-CMG (Berkeley Short-channel IGFET Model – Common Multi-Gate) are presented. The core model agrees with TCAD simulation very well without using any fitting parameter. The complete model is verified against experimental data for two different FinFET technologies – SOI FinFETs and bulk FinFETs.

Compact model for the independent multi-gate device (BSIM-IMG, Berkeley Short-channel IGFET Model – Independent Multi-Gate) is introduced in Chapter 4. The model is used to study FinFET based SRAM cells and device variation tuning using back gate bias, highlighting its use for both circuit and technology development. The impact of back-gate length and misalignment of planar double-gate (DG) MOSFET on threshold voltage is also discussed.

In Chapter 5, the circuit design issues and concepts of MG-based circuits are discussed. The circuit design trade-off of two main types of DG MOSFETs: (i) the symmetric/common DG (SDG) and (ii) the asymmetric/independent DG (ADG) are discussed. The performance of FinFET technology in digital circuit applications is studied compared to the planar SOI technology under various device parameter

variations.

A novel methodology for generating Performance Aware (Corner/Distribution) Models (PAM) cards is presented in Chapter 6. More accurate and application-specific (for speed, power, gain, etc) model cards can be easily generated at any distribution levels (such as $+2\sigma$, -1σ). The accuracy improvement of generatd PAM cards is demonstrated by applying it to different scale of logic circuits. The PAM cards also improve the accuracy of Monte Carlo simulation by reconciling the physical and electrical-test (ET) variances.

An overall summary of this dissertation is presented in Chapter 7. Key research contributions and suggestions for future research direction are highlighted.

1.6 References

[1.1]   J. S. Kilby, "Miniaturized electronic circuits," U.S. Patent #3,138,743, issued June 23, 1964.

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

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

[1.4]   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 J. Solid-State Circuits, vol. 9, pp. 256-268, 1974.

[1.5]   J. R. Brews, W. Fichtner, E. H. Nicollian, and S. M. Sze, "Generalized Guide for MOSFET Miniaturization," IEEE Electron Device Lett., vol. 1, pp. 2-4, 1980.

[1.6]   S.-H. Lo, D. A. Buchanan, Y. Taur, and W. Wang, "Quantum-Mechanical Modeling of Electron Tunneling Current from the Inversion Layer of Ultra-Thin-Oxide nMOSFET's," IEEE Electron Device Lett., vol. 18, no. 5, 209-211, 1997.

[1.7]   T. Tanaka, T. Usuki, T. Futatsugi, Y. Momiyama, and T. Sugii, "Vth Fluctuation Induced by Statistical Variation of Pocket Dopant Profile," IEEE International Electron Devices Meeting, pp. 271-274, 2000.

[1.8]   A. Hokazono, K. Ohuchi, M. Takayanagi, Y. Watanabe, S. Magoshi, Y. Kato, T. Shimizu, S. Mori, H. Oguma, T. Sasaki, H. Yoshimura, K. Miyano, N. Yasutake, H. Suto, K. Adachi, H. Fukui, T. Watanabe, N. Tamaoki, Y. Toyoshima, and H. Ishiuchi, "14nm Gate Length CMOSFETs utilizing Low Thermal Budget Process

with Poly-SiGe and Ni Salicide," IEEE International Electron Devices Meeting, pp. 639-642, 2002.

[1.9]   T. Ghani, M. Armstrong, C. Auth, M. Bost, P. Charvat, G. Glass, T. Hoffmann, K. Johnson, C. Kenyon, J. Klaus, B. McIntyre, K. Mistry, A. Murthy, J. Sandford, M. Silberstein, S. Sivakumar, P. Smith, K. Zawadzki, S. Thompson, and M. Bohr, "A 90nm High Volume Manufacturing Logic Technology Featuring Novel 45nm Gate Length Strained Silicon CMOS Transistors," IEEE International Electron Devices Meeting, pp. 978-980, 2003.

[1.10]   Robert Chau, Suman Datta, Mark Doczy, Brian Doyle, Jack Kavalieros, and Matthew Metz, "High-K/Metal Gate Stack and Its MOSFET Characteristics," IEEE Electron Device Lett., vol. 25, no.6, pp.408-410, 2004

[1.11]   L. Chang, Y.-K. Choi, J. Kedzierski, N. Lindert, P. Xuan, J. Bokor, C. Hu, and T.-J. King, "Moore's Law Lives On: Ultra-Thin Body SOI and FinFET CMOS Transistors Look to Continue Moore's Law for Many Years to Come," IEEE Circuits and Devices Magazine, vol. 19, iss. 1, January 2003, pp. 35-42.

[1.12]   D.J. Frank, R. H. Dennard, E. Nowak, P. M. Solomon, Y. Taur and H.-S.P. Wong, "Device Scaling Limits of Si MOSFETs and Their Application Dependencies," Proceedings of IEEE, vol. 89, pp. 259-288, March 2001.

[1.13] D. Hisamoto, Wen-Chin Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E.

Anderson, Tsu-Jae King, J. Bokor and Chenming Hu, "FinFET – A Self-Aligned

Double Gate MOSFET Scalable to 20nm," IEEE Trans. Electron Devices, vol. 47,

pp. 2320-2325, Dec. 2000.

[1.14] M. Dunga, Ph.D. Dissertation, University of California at Berkeley, Fall 2007.

Chapter 2

Modeling of Physical Effects in Multi-Gate MOSFETs

2.1 Motivation

The continuous evolution and enhancement of bulk CMOS technology has fueled the growth of the microelectronics industry for the past several decades. When we reach the end of the technology roadmap for the classical CMOS, multi-gate CMOS structures will take up the baton. The multi-gate structure can be scaled to the shortest channel length for a given oxide thickness due to the enhanced electrostatic control from the multiple gates. It is thus necessary to develop a compact model of multi-gate MOSFETs for technology/circuits development in the short term and for product design in the longer term.

There are several important unique device physics in the multi-gate MOSFET due to extra gates from sidewall. The prediction and completion of a compact model relies on the inclusion of these unique device physics in the model with appropriate flexibility. In this chapter, the quantum mechanical effects, short channel effects, and corner effects in the multi-gate MOSFET are studied. Modeling methodologies and approaches are proposed to incorporate these unique multi-gate physics in the compact model.

The quantum effect is already important in advanced planar single gate transistors. In the multi-gate devices, the additional confinement of the thin body makes the quantum effect an indispensable part of the model. Several quantum model approaches for the common double-gate structures are available in the literature. However, these approaches are limited to the carrier distribution and electrostatic potential profile, but not extended to device characteristics, such as I-V and C-V. In section 2.2, we developed an accurate quantum mechanical compact model for multi-gate MOSFET. Both V_{th} shift in the subthreshold and strong inversion regions and gate capacitance degradation in the strong inversion region due to QM are corrected simultaneously. The model can predict the complicated QM effect dependence on various device parameters.

The multi-gate structure has better short channel behaviors due to enhanced electrostatic control from the multiple gates. Critical geometry parameters which

determine device short-channel behaviors include gate length, fin thickness, fin height, oxide thickness, and channel doping. It is very important to include all these parameters in the short channel model to give correct scalability over a wide range of device parameters. In section 2.3, the short channel model is developed based on the quasi 2-D potential profile in the conduction channel. The degree of SCE (V_{th} roll-off, drain-induced-barrier-lowering (DIBL), and subthreshold slope degradation) depends on strength of gate control which is modeled by a characteristic field penetration length ($\lambda = f(T_{ox}, T_{si})$) derived from quasi 2-D Poisson's equation. The SCE model shows excellent agreements with 2-D TCAD simulation results without the use of any fitting parameters. The SCE model implementation captures V_{th} roll-off, DIBL and subthreshold slope degradation for short channel multi-gate FETs simultaneously.

The corner effect is important for extending the double-gate model to cover triple gate (and all-around gate) structures. These effects are known to dominate the subthreshold and weak inversion current. Therefore accuracy is of paramount importance. The corner effect is studied using the 3-D TCAD device simulation. Section 2.4 shows the proposed modeling methodology by introducing the "cap" transistor.

16

2.2 Quantum Mechanical Effects

Carrier energy quantization has become significant in the state-of-the-art MOSFETs due to increased vertical E-field. The energy quantization and the shift of the inversion charge centroid will delay the formation of inversion charge (threshold voltage (V_{th}) shift) and reduce the current driving capability (increase the effective oxide thickness). The quantization effect is more complicated in double-gate (DG) MOSFETs than in bulk MOSFETs due to the extra structural confinement by the body thickness (T_{si}). The quantum mechanical (QM) effect has been included in the compact models by introducing the effective oxide thickness for a bias-dependent reduction of the gate capacitance, and bias-independent correction for the V_{th} shift, separately [2.1]. However, the lack of predictivity of this approach is particularly undesirable for the DG MOSFETs compact model since less DG MOSFETs Si data are available than single-gate MOSFETs. Several groups have investigated the influence of T_{si} on the quantized carrier distribution and threshold voltage shift numerically [2.2] and analytically [2.3-2.5] in DG MOSFETs. However, these approaches are limited to the carrier distribution and electrostatic potential profile, but not extended to device characteristics, such as I-V and C-V.

While the channel carriers are confined in one-dimension, the subband splitting due to the field-induced electrical confinement (EC, Fig. 2.1(a)) has significant impact on device

characteristics. The energy quantization is more complicated in the DG MOSFETs due to

the extra structural confinement (SC) in the subthreshold region as shown in Fig. 2.1(b).
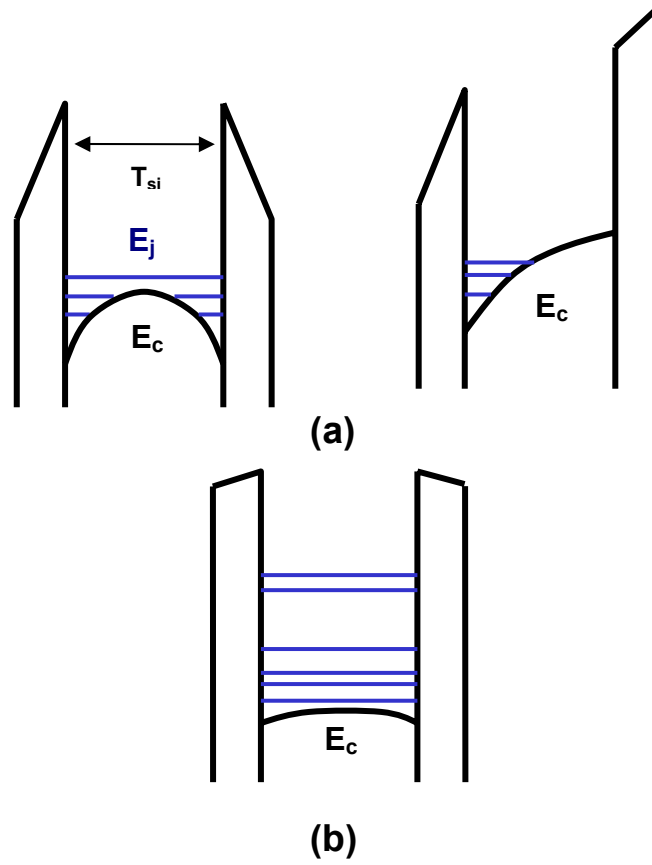
Fig. 2.1 Energy-band diagrams of DG-MOSFET under different operation modes illustrate two mechanisms of carrier energy quantization in DG MOSFETs: (a) electrical confinement (EC); (b) structural confinement (SC).

The quantum confinement in DG MOSFETs is affected by the gate work function, gate

insulator dielectric thickness, body thickness, substrate doping, and gate bias. In order to

study these complicated device parameter dependences of QM effects, a self-consistent

1-D Schrödinger solver, SCHRED 2.1, is used for simulating QM effects in DG structure

[2.6].

Fig. 2.2 shows the simulation results of the ratio of QM-corrected and classical charge

density over different substrate doping concentrations for n-type DG MOSFETs.
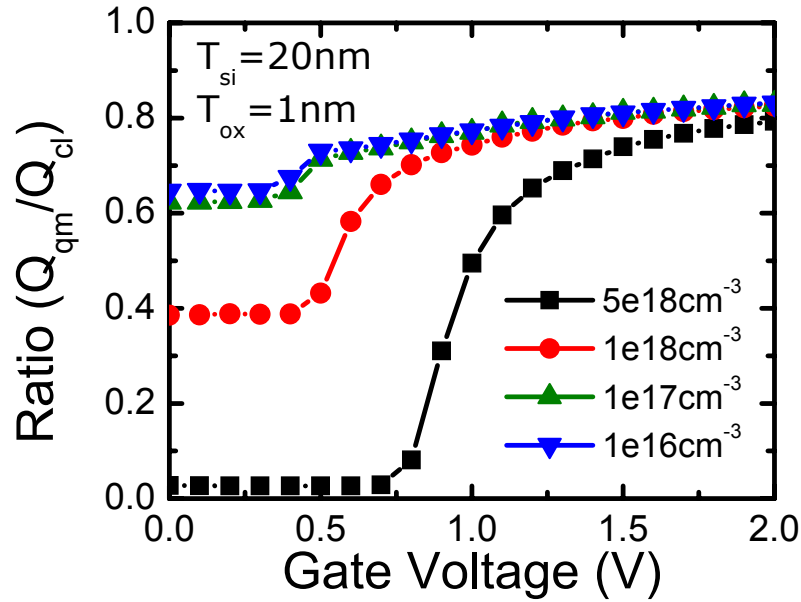
Fig. 2.2 The doping concentration dependence of QM effects. QM effects is very significant in heavily-doped device due to larger vertical E-field.

In the subthreshold region, the ratio of charge reflects the V_{th} shift, while in the strong

inversion region, the ratio of charge indicates the amount of the gate capacitance

degradation. In the lightly doped or undoped DG MOSFETs, QM effect is less significant

due to weaker EC (smaller E-field). The SC is weak in this structure due to relatively

thick T_{si}. The dependence of body thickness on QM effect is shown in Fig. 2.3. The QM

effect is less significant in the heavily doped thinner body device (T_{si}=5nm) due to

weaker EC (electrical coupling from two gates).

Fig. 2.4 shows the inversion carrier distribution vs. the body thickness at the strong inversion region in the symmetric DG MOSFETs. The body doping is 10^{15}cm^{-3} and the midgap workfunction is used in the simulation. The DG MOSFETs exhibit a more uniform inversion carrier density when the body thickness is scaling down (T_{si}=5nm). The two finite charge centroids overlap so strongly that the maximum of the inversion carrier concentration is located in the middle of the body film. The inversion carriers travel away from the oxide/Si interface. Note that the mobility characteristics also change since there is less surface roughness scattering experienced by the inversion carriers. The effective mobility in the DG MOSFETs is affected by the average position of the inversion carriers from the oxide/Si interface. The average position of the inversion carrier can be expressed as

$$X_{dc} = \frac{\int_0^{Tsi/2} x n(x) \cdot dx}{\int_0^{Tsi/2} n(x) \cdot dx} \tag{2.1}$$

Fig. 2.5 shows the extracted X_{dc} vs. gate bias for the devices with different body thickness. The X_{dc} is clamped at ~1/3 body thickness in the subthreshold region due to the strong overlap of the two finite charge centroids. The X_{dc} decreases as the gate bias increases due to stronger vertical electric field.
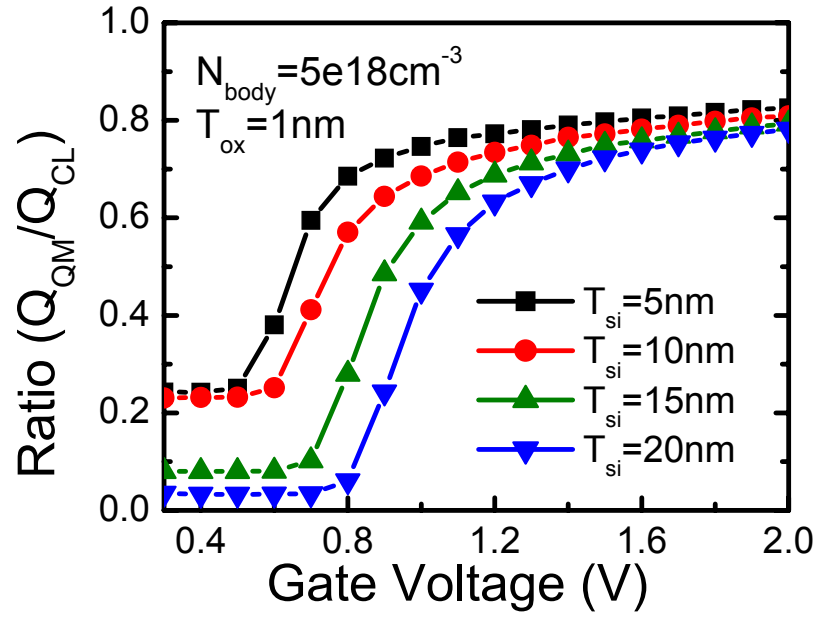
Fig. 2.3 The body thickness dependence of QM effects. QM effects is weak in the heavily doped thinner body device due to smaller vertical E-field.
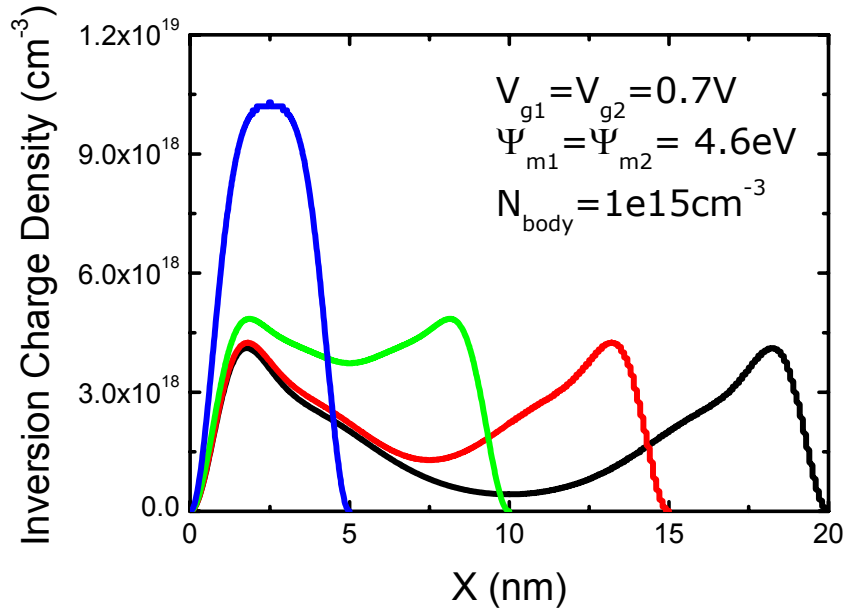

Fig. 2.4 Impact of the body thickness on the inversion carrier concentration distribution for a given fixed gate bias.
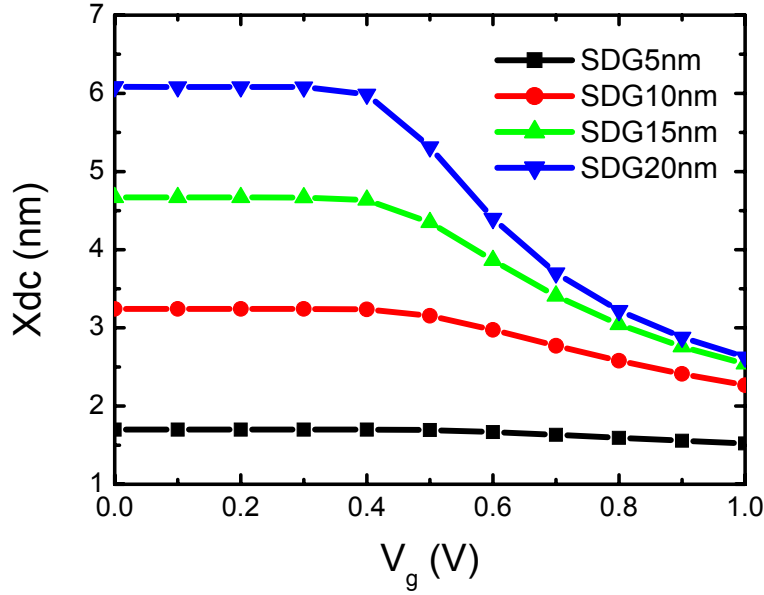
Fig. 2.5 The extracted charge centroids of the symmetric DG-MOSFETs with different body thickness.

Conventional compact models handle QM correction by introducing the effective oxide thickness for a bias-dependent reduction of the gate capacitance, and bias-independent correction for the V_{th} shift, separately [2.1]. The charge centroid theory gives the universal bias-dependent reduction of the gate capacitance and has been verified by the self-consistent Schrödinger solver.

We have developed a surface-potential based multiple-gate MOSFET compact model, where the surface potential is solved by Poisson's equation and Gauss's law [2.7]. A bias-dependent ground-state subband energy in the unprimed valley (E_0) is added in the

surface potential calculation. Trivedi et al. have calculated the E_0 using the vibrational approach to include both structural and electrical confinement in DG MOSFETs [2.4]. However, the model only considers the threshold voltage shift in intrinsic undoped body device. A modification was made for E_0 calculation to extend the model to strong inversion region with the finite doped body. Drift-diffusion equation is then employed to obtain a model for drain current in terms of surface potential. The V_{th} shift and gate capacitance degradation are inherently captured without the need of any explicit individual modifications as long as the surface potential is modeled correctly.

In the new surface potential calculation, a QM correction term ($\Delta\psi_{QM}$) is added in the inversion carrier term when solving Poisson's equation.

$$
V_g = V_{fb} + \phi_s + \frac{\varepsilon_{si}}{C_{ox}}
$$
$$
\cdot \sqrt{\frac{2qn_i}{\varepsilon_{si}} \cdot \left(\frac{e^{\frac{q(\phi_s - \Delta\psi_{QM})}{kT}} - e^{\frac{q(\phi_0 - \Delta\psi_{QM})}{kT}}}{q/kT} \cdot e^{\frac{-q\varphi_B}{kT}} \cdot e^{\frac{-qV_{ch}}{kT}} + e^{\frac{q\varphi_B}{kT}} \cdot (\phi_s - \phi_0) \right)} \tag{2.2}
$$

where $\Delta\psi_{QM} = E_0/q$. E_0 is the ground state subband energy obtained by solving Schrödinger equation via the vibrational approach [2.4]

$$
E_0 \approx \frac{\hbar^2}{2m_x} \left[\left(\frac{\pi}{T_{si}} \right)^2 + b_0{}^2 \left(3 - \frac{4}{3} \frac{1}{\left[\left(\frac{b_0 T_{si}}{\pi} \right)^2 + 1 \right]} \right) \right] \tag{2.3}
$$

b_0 represents the vertical electric field dependence. Both structural and electrical confinements are included in the ground state subband energy. The first term of the right hand side represents the structural confinement. The thinner the body thickness is, the higher ground state subband energy. The second term of the right hand side represents the structural confinement. The b_0 represents the strength of the vertical electrical field, which is expressed as

$$b_0 \approx \sqrt[3]{\frac{3}{4}\frac{2m_x q E_x}{\hbar^2}} \tag{2.4}$$

where E_x is the vertical field calculated from the surface potential. Note that the vertical field is accurate from subthreshold region to strong inversion region, consequently the QM correction is accurate for all regions of the operations

$$E_x = \frac{V_{gs} - V_{fb} - \phi_s}{\varepsilon_{Si}/\varepsilon_{ox} \cdot T_{ox}}. \tag{2.5}$$

Fig. 2.6 shows the ground state subband energy vs. gate bias in the symmetric/common double-gate MOSFET obtained from Eq. (2.3). The ground state subband energy is constant in the subthreshold region due the constant E_x in the subthreshold region in SDG. However, the ground state subband energy increases as gate bias increases in the strong inversion region. The bias-dependent correction of ground state subband energy accurately predicts the constant V_{th} shift in the subthreshold region and bias-dependent V_{th} shift (gate capacitance degradation) in the strong inversion region.
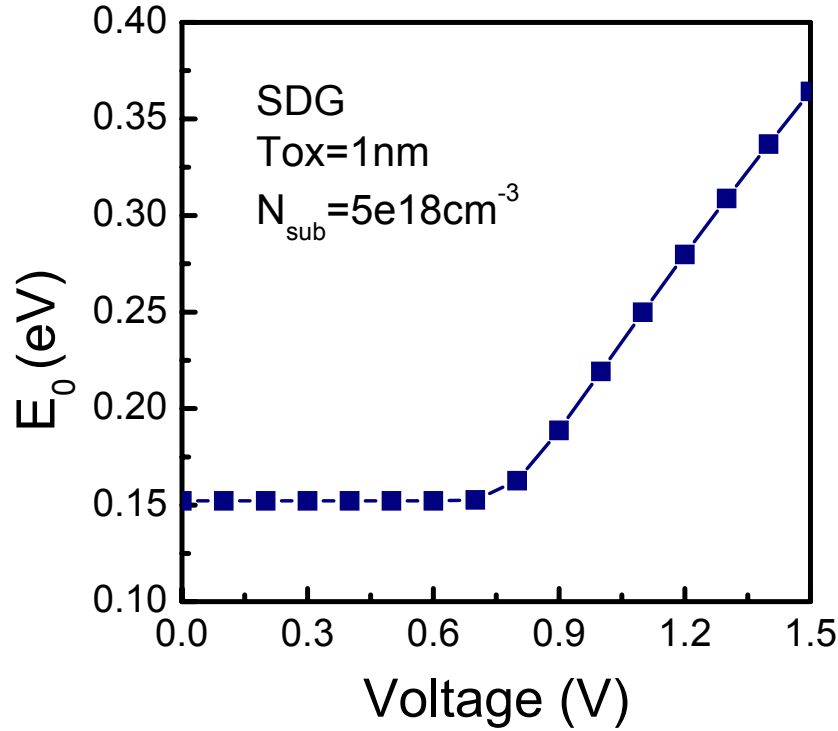
Fig. 2.6 The predicted ground state subband energy in the SDG MOSFET. The E_0 is constant in the subthreshold region and increases as gate bias increases.

Fig. 2.7 shows the surface potential versus V_g for a long channel n-type DG MOSFET where source/drain is grounded. A uniformly doped body ($5 \times 10^{18} cm^{-3}$) is assumed. The model matches 2-D device simulation very well from subthreshold region to strong inversion region. The increase of the surface potential due to the QM effect is well predicted by the model. Larger surface potential is needed to achieve the same level of inversion. The model can accurately predict surface potential over a wide range of body doping concentration as shown in Fig. 2.8 ($10^{16} cm^{-3}$).
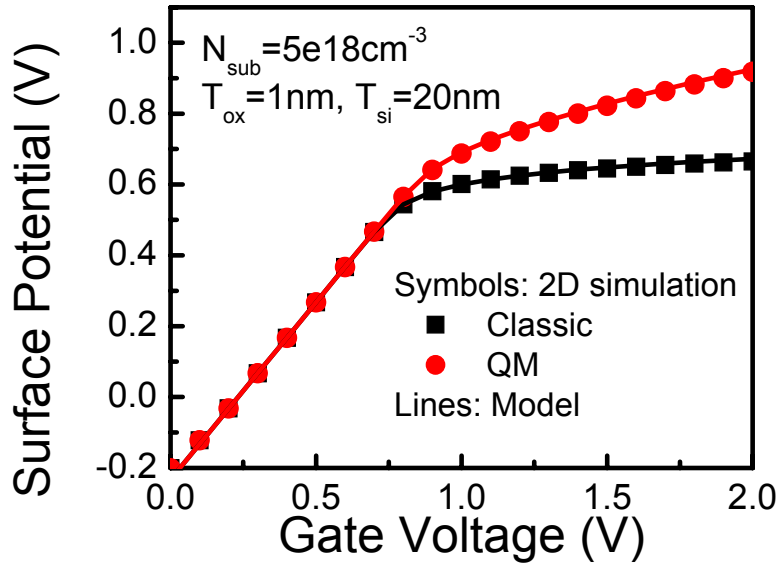
Fig. 2.7 QM-corrected surface potential solution matches 2-D simulation well. The QM-induced V_{th} shift in subthreshold region and C_{ox} degradation in strong inversion region are simultaneously predicted as long as QM-corrected surface potential is accurate
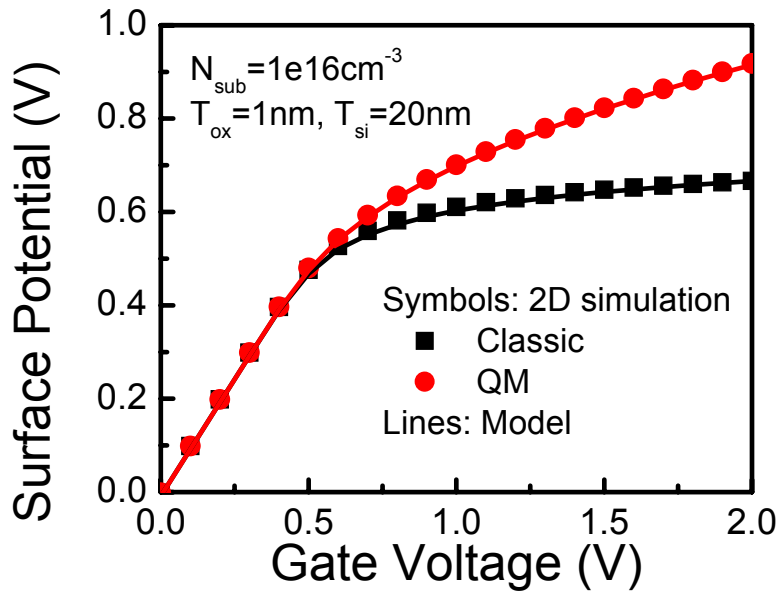

Fig. 2.8 Surface potential vs. gate voltage for undoped body device. Model can accurately predict the doping dependence.

The Drift-diffusion equation is then employed to obtain a model for drain current in terms of QM-corrected surface potential. TAURUS 2-D device simulations are used to verify the QM model. Fig. 2.9 shows the I_d-V_g characteristics of the n-type symmetrical/common-gate DG MOSFET for both classical and QM models. The drain is biased in linear region (V_{ds}=50mV). The oxide thickness of 1nm and midgap work-function gate material are used in simulation. The substrate doping is 5×10^{18}cm^{-3}. The QM model predicts the V_{th} shift accurately compared to the 2-D device simulation results. The model can predict V_{ds} dependence accurately as shown in Fig. 2.10.
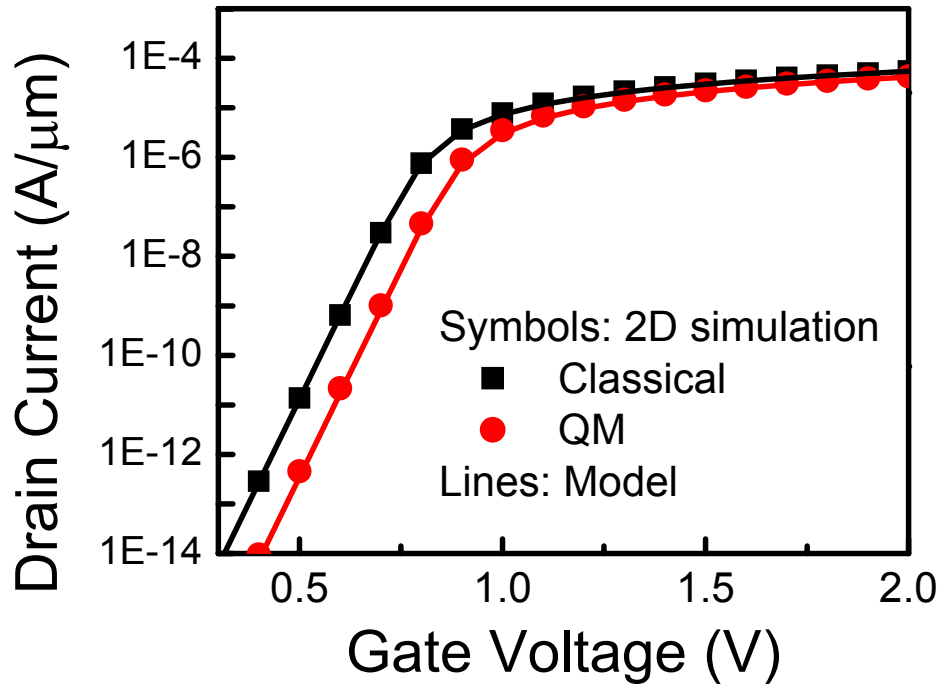

Fig. 2.9 The classical and QM-corrected I_d-V_g characteristics for n-typed symmetrical/common-gate DG MOSFET.

(a)

(b)
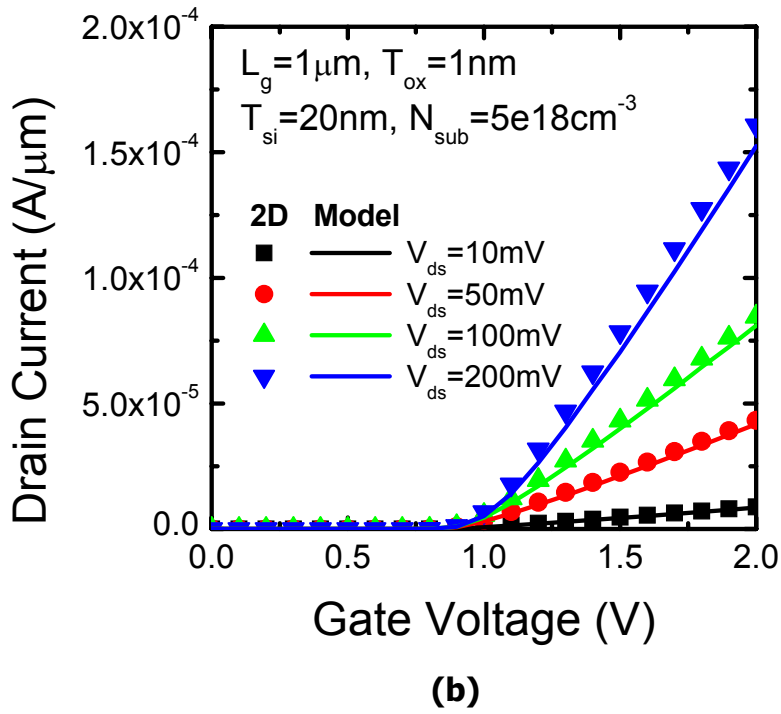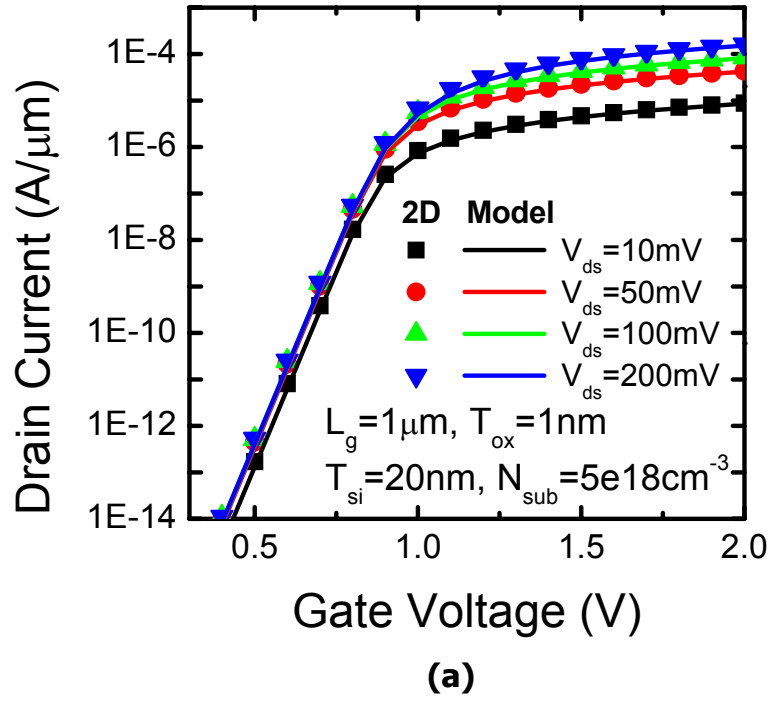
Fig. 2.10 The QM-corrected drain current model predicts V_{ds} dependence of I_d-V_g characteristics well, shown in (a) log scale and (b) linear scale.

The ratio of QM-corrected and classical drain current is shown in Fig. 2.11. The ratio in the subthreshold region reflects the effect of the V_{th} shift. The ratio in the strong inversion region is mainly caused by the gate capacitance degradation due to finite charge centroid. It clearly shows that the QM model predicts both the V_{th} shift and gate capacitance degradation without introducing the definition of an effective oxide thickness. The model predicts QM effects accurately in all regimes of operation since the QM-corrected surface potential is modeled accurately against the QM simulation.
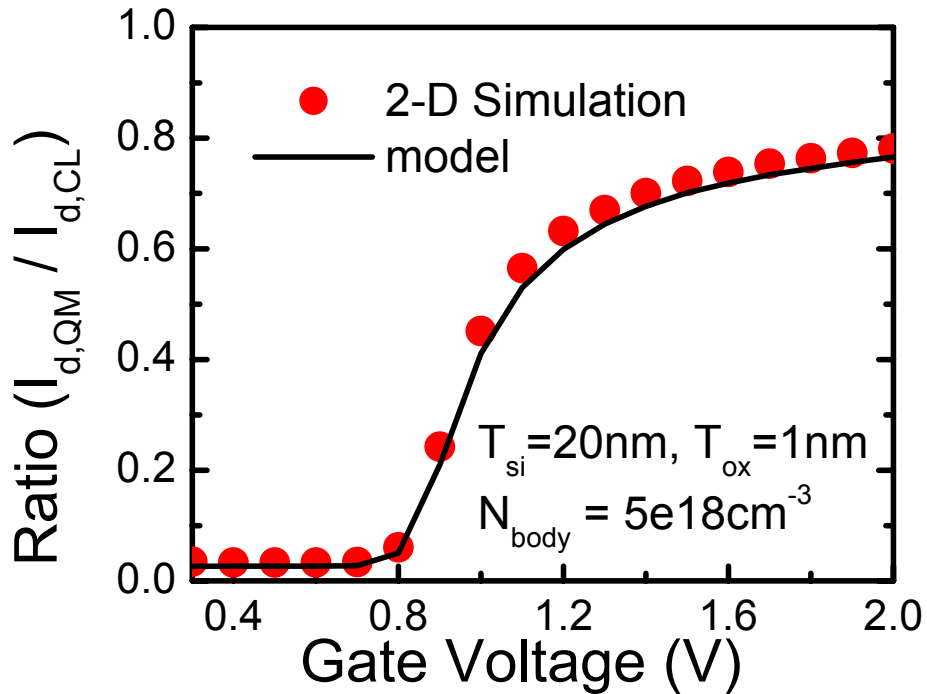
Fig. 2.11 The ratio of QM-corrected and classical drain current. The model captures the threshold voltage shift and gate capacitance degradation simultaneously through accurate surface potential calculation.

The predictivity of model is important for MG technology and circuits development since less Si data is available compared to single-gate MOSFETs. The QM model can predict the complicated QM effect dependences on various device parameters, such as N_{body}, T_{si}, T_{ox}, etc. Fig. 2.12 shows that the predicted V_{th} shift over a wide range of body doping concentration matches the 2-D simulation results very well. EC dominates in higher doping concentration region, while SC dominates in lower doping concentration region.

Fig. 2.12 V_{th} shift can be modeled accurately over a wide range of body doping. EC dominates in higher doping region, while SC dominates in lower doping region.

2.3 Short Channel Effects

Short channel effects (SCE) such as drain-induced barrier lowering (DIBL), threshold voltage (V_{th}) roll-off and sub-threshold slope degradation are modeled. SCE are essentially 2-D effects where the drain significantly affects the potential barrier at the source due to its close proximity to the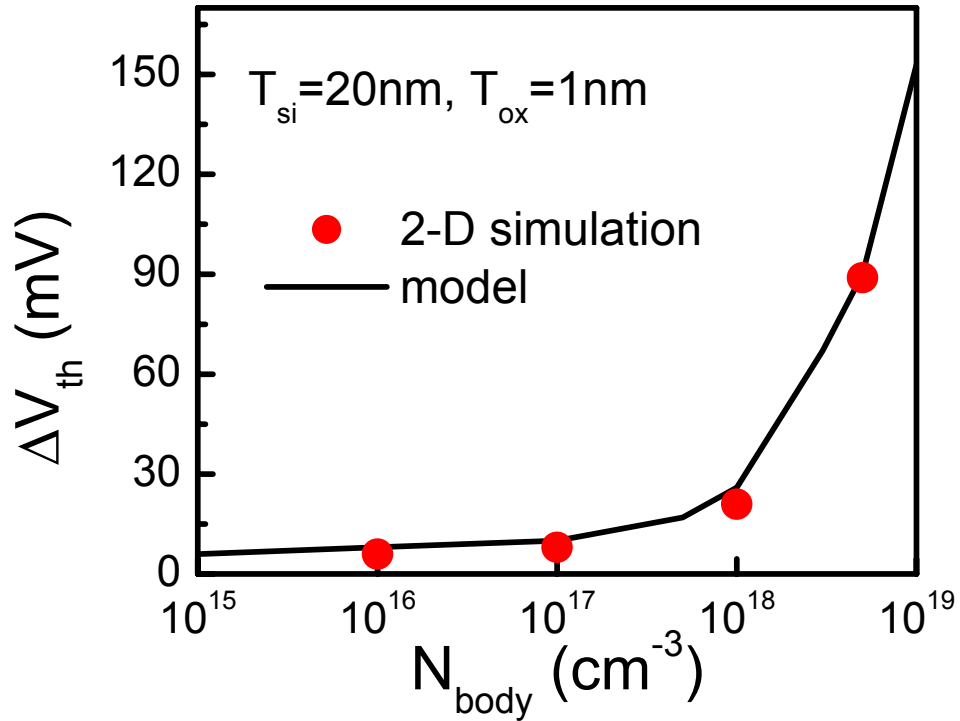 source region in a short channel transistor. A good SCE must be scalable over a wide range of the device parameters such as gate length L, gate insulator thickness T_{ox}, fin thickness T_{fin}, fin height H_{fin} and channel doping N_A.

There have been numerous efforts in the past to model the SCE in DG-FETs. Invariably all the methods to model SCE solve the 2-D Poisson's equation in sub-threshold region inside the silicon body with varying degree of simplifying assumptions [2.8-2.11]. In [2.8] the authors solve the complete 2-D boundary value problem by expressing the potential through a infinite sum of sin functions. The solution is numerically complex and will be hard to extend to the case of multi-gate FETs with triple or quadruple gates. In [2.9] the authors first solve a 1-D Poisson equation in the channel length direction and then solve the 2-D Poisson equation. The solution in this case is independent of V_{ds} and hence modeling DIBL is a challenge in this approach. Another approach assumes a parabolic potential function perpendicular to the silicon-insulator interface and solves the 2-D Poisson's equation [2.10-2.11]. This

approach maintains a balance between the model accuracy and model computation time and hence is used to develop a SCE model for DG-FETs.

The SCE is determined by minimum potential barrier ($\psi_{c(min)}$) seen by the carrier entering at the source end. In the DG-FET, the minimum potential barrier height is located at the center plane of the body. By using the parabolic potential assumption [2.12], the $\psi_{c(min)}$ is derived as

$$\psi_c\left(\min\right)=V_{SL}-\frac{V_{ds}^{\;2}\cdot e^{-L/2\lambda}}{\left(e^{L/\lambda}-e^{-L/\lambda}\right)\sqrt{Z_0 Z_L}}+2\sqrt{Z_0 Z_L}\,\frac{\sinh\left(L/2\lambda\right)}{\sinh\left(L/\lambda\right)} \tag{2.6}$$

where V_{SL} is equivalent to the center potential for a long channel transistor

$$V_{SL}=V_{gs}-V_{fb}-\frac{qN_A}{\varepsilon_{Si}}\lambda^2 \tag{2.7}$$

where λ is a characteristic field penetration length defined as

$$\lambda=\sqrt{\frac{\varepsilon_{Si}}{2\varepsilon_{ox}}\cdot\left(1+\frac{\varepsilon_{ox}T_{si}}{4\varepsilon_{Si}T_{ox}}\right)\cdot T_{si}\cdot T_{ox}} \tag{2.8}$$

where $Z_0=V_{bi}-V_{SL}$, $Z_L=V_{bi}-V_{SL}+V_{ds}$ and V_{bi} the built-in potential at the source end. For a long channel transistor, $\psi_{c(min)}\sim V_{SL}$.

The compact model is traditionally developed in a long channel framework. To obtain the correct minimum barrier height for a short channel DG-FET at a certain V_{gs} using the long channel surface potential model, one can use the long channel surface potential model with an effective V_{gs}' where

$$V'_{gs} = V_{gs} + \psi_{c(min)} - V_{SL} \tag{2.9}$$

This method can be easily implemented into any compact model by replacing V_{gs} with effective V'_{gs}. Note that the SCE model in Eq. (2.9) not only captures the V_{th} roll-off and DIBL but also captures the degradation of subthreshold slope simultaneously since the correction term is a function of gate voltage.

The SCE model is verified against 2-D TCAD simulations. Drain current is calculated using the SCE and I-V model for a wide range of channel lengths ranging from a long channel $L = 1\mu m$ DG-FET to a short channel DG-FET of $L = 30nm$. Fig. 2.13 shows the normalized I_{ds} ($I_{ds} \times L/W$) calculated from the model against the current obtained from 2-D TCAD simulations. Good agreement in the SCE behavior such as V_{th} roll-off and subthreshold slope degradation is observed between the model and 2-D TCAD.

Good scalability of the SCE model is highly desirable. A scalable model allows one to stretch the limited available silicon data to perform technology projections. The scalability of the SCE model is examined extensively with respect to physical parameters such as L, T_{Si} and T_{ox}. Fig. 2.14 shows the V_{th} roll-off extracted from the model and 2-D TCAD for different T_{Si}. As expected, DG-FETs with smaller T_{Si} have smaller drain field penetration and hence the model predicts smaller V_{th} roll-off for thinner T_{Si}.
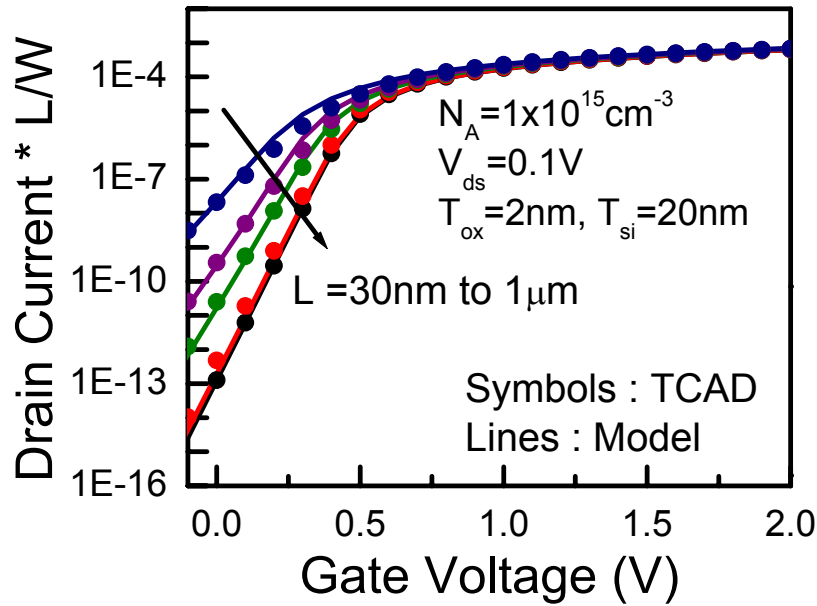
Fig. 2.13 Normalized drain current for DG-FETs with different channel lengths, from L = 1µm down to L = 30nm. All the important SCE are captured by the model and agree well with 2-D TCAD simulations.
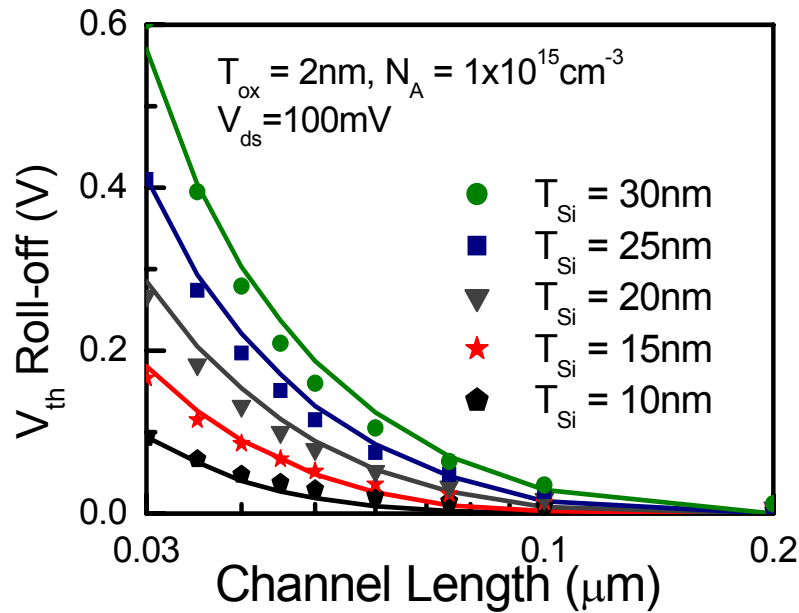

Fig. 2.14 Threshold voltage roll-off of DG-FETs for different T_{Si}. Model (lines) demonstrates excellent scalability when compared against 2-D TCAD simulations (symbols).

Thus far, only DG-FETs, i.e. multi-gate FETs with gate electrodes only on two opposite sides of Si body, have been analyzed. Multi-gate FETs can have three or four gate electrodes to further improve the electrostatic control. In multi-gate FETs with three or more gates, the physical location of the minimum potential barrier (or the path for maximum drain leakage current) is different from DG-FET. The extra electrostatic control from vertical ends (top gate or bottom gate) reduces the short channel effect. The V_{th} roll-off decreases as fin height (H_{fin}) decreases. The most leaky channel path is located at the center bottom of the fin where the electrostatic control from the gate is the weakest as shown in Fig. 2.15. The potential barrier at this most leaky path decreases as fin height increases, resulting in an H_{fin} dependence of short channel effects.
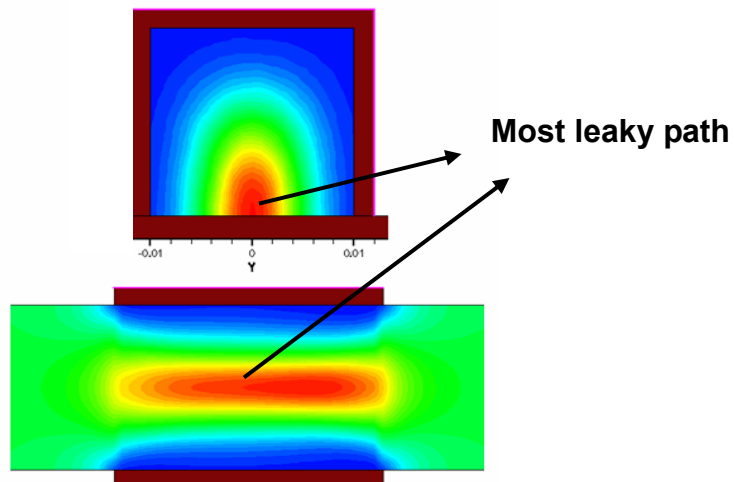

Fig. 2.15 Leakage current path is different in triple-gate FETs due to 3D effects. The most leaky path located at the bottom center of the fin.

Pei et al. has proposed a 3-D analytical electrostatic potential based on the solution of 3-D Laplace's equation in subthreshold region to predict the short channel behavior of FinFET [2.13]. In Eq. (2.9), the DG-FET short channel effects are modeled by a characteristic field penetration length. To model the fin height dependence on short channel effects, a new characteristic field penetration length λ_{Hfin} is introduced.

$$\lambda_{Hfin} = \sqrt{\frac{\varepsilon_{Si}}{4\varepsilon_{ox}} \cdot \left(1 + \frac{\varepsilon_{ox} H_{fin}}{2\varepsilon_{Si} T_{ox}}\right) \cdot H_{fin} \cdot T_{ox}} \qquad (2.10)$$

The new effective characteristic length used in the short channel model is defined to include the effect of fin height

$$\lambda_{eff} = \frac{1}{\sqrt{\left(\frac{1}{\lambda}\right)^2 + \left(\frac{a}{\lambda_{Hfin}}\right)^2}} \qquad (2.11)$$

where $a = 0$ for DG-FET, $a = 0.5$ for triple-gate FinFET, $a = 1$ for surrounding-gate FinFET. Note that in the case of triple-gate FET, one can also use different oxide thickness in λ and λ_{Hfin} to model the thick SiO_2 layer (hard mask) on top of the fin. Fig. 2.16 shows the comparison of predicted V_{th} roll-off between compact model and the TCAD simulations. Fin height dependence for the short channel effect is verified using 3-D TCAD simulation. The V_{th} roll-off increases as fin height increases for a given fin thickness. The model agrees with the TCAD simulation results very well.
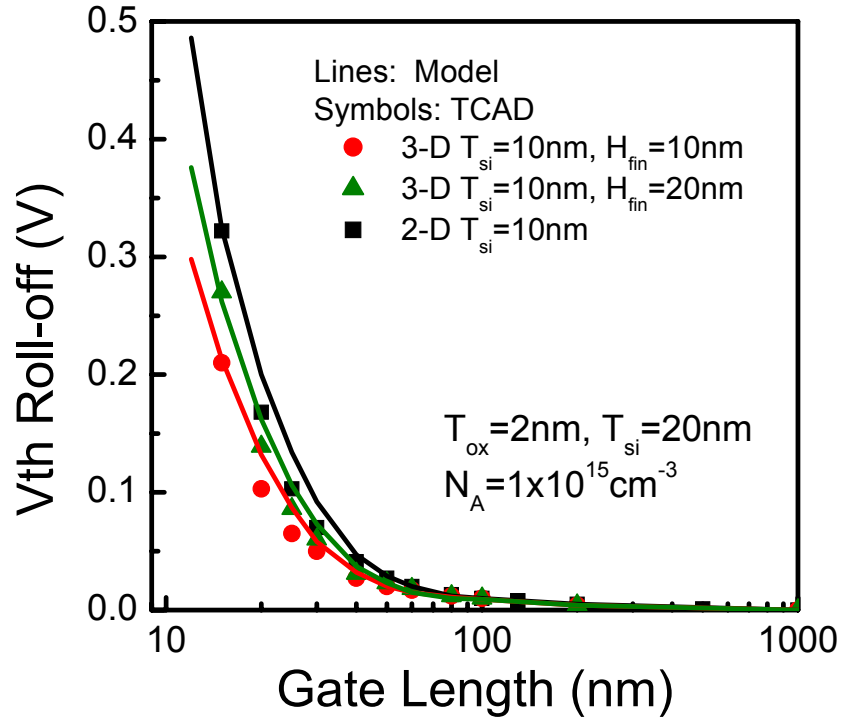
Fig. 2.16 SCE model exhibits fin height scalability for triple-gate FinFETs (symbols: 3-D TCAD, lines: model)

2.4 Corner Effects

In the multi-gate MOSFET, the triple-gate device exerts superior electrostatic control over the channel due to the presence of the top gate. However, the subthreshold swing and threshold behavior are different in the triple-gate device at the top corners. According to 3-D device simulations, the corner conduction (at top portion) will dominate the subthreshold leakage current if the channel doping is high. The corner effect can be

suppressed with rounded corners, thin gate oxides, and lower channel doping [2.14-2.16].

To address the corner effect in the model, a new corner effect model is needed for the

inversion charge calculation with threshold voltage modification.

In order to calculate the inversion charge densities and potential in the silicon fin, the

3-D device simulation was carried out using DESSIS [2.17]. The simulation structure is

shown in the Fig. 2.17. A set of devices with physical gate length 100 nm, various

substrate doping levels ranging from 10^{16} cm^{-3} to 10^{18} cm^{-3}, a 1.1 nm thick gate oxide, 30

nm fin width, 30 nm and 60 nm fin height. We assume the mid-gap metal gate material.

To study the corner effect, we investigate the 2-D cross section of the device in the center

of the channel.
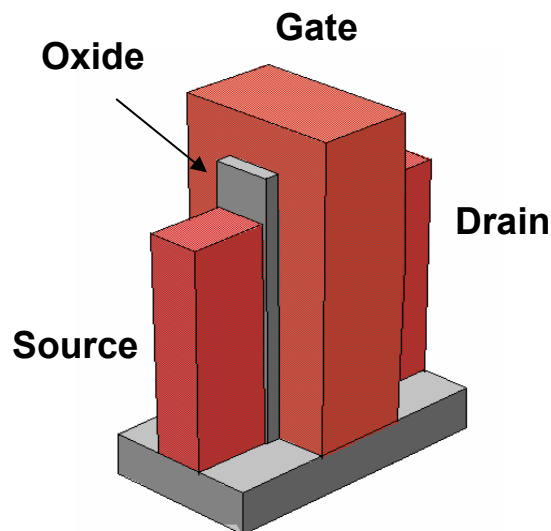
Fig. 2.17 3-D FinFET structure used for TCAD simulations

Fig. 2.18 shows the 2-D charge distribution in the 30nm×30nm fin cross section bias at subthreshold region. The substrate doping is 10^{18}cm^{-3}. Note that the inversion charge is built up significantly at the top corners of the silicon fin, which corresponds to a smaller threshold voltage of the top corner portion. Similar results are observed in the device with taller fin (60nm). To suppress the corner effect, lower substrate doping or rounded corner shape is required. Fig. 2.19 shows 2-D charge distribution of the FinFET with hard mask on top of the Si fin. The corner effect is suppressed in the device with lightly doped substrate. Note that even though there is hard mask on top of the channel, there is still significant inversion charge built in the top corner portion of the heavily doped device.
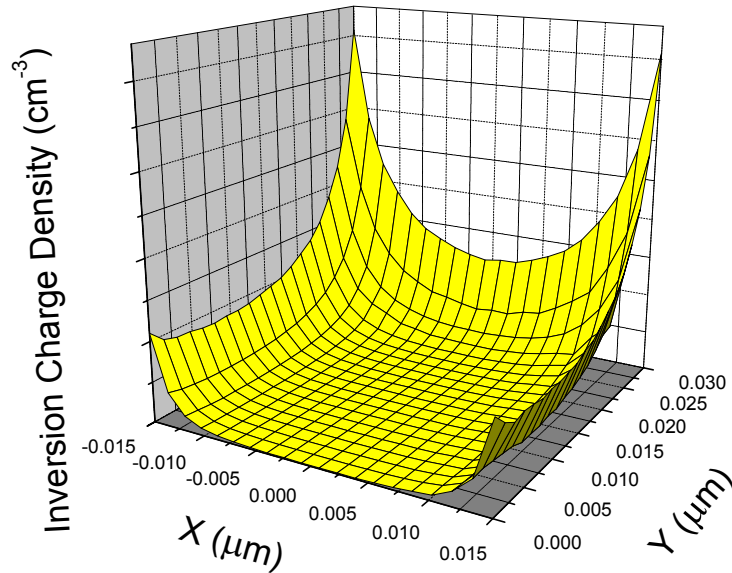
Fig. 2.18 2-D inversion charge distribution in the Si fin biased at subthreshold region.
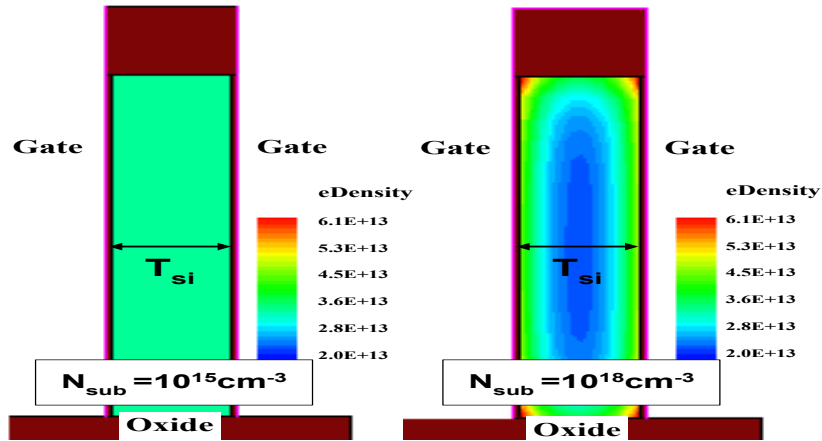
Fig. 2.19 The electron density distribution from the 3-D ISE device simulator of L_g=100nm nFinFETs. Corner effect is suppressed in the lightly doped channel device (left).

In order to model the corner effect, a "cap" transistor model is introduced. The height of cap transistor is half of the fin width independent of fin height, which can be explained by the charge sharing concept as shown in Fig. 2.20.
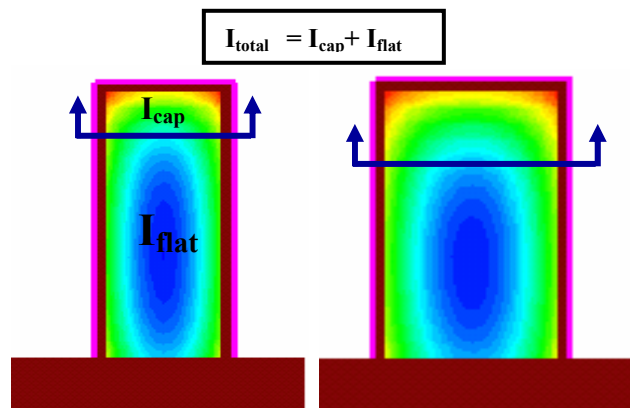
Fig. 2.20 2-D inversion charge distribution of three triple-gate MG-FETs with different fin thickness (same fin height). The height of cap transistor equals to half of fin width independent of fin height due to charge sharing by vertical and sidewall gates.

Therefore, the "cap" transistor can be extracted by test MG structures with different fin heights, but identical fin width. The lower threshold voltage, stronger level of inversion and process induced corner shape variations can be modeled by effective flat-band voltage shift, effective thinner oxide, and effective width using the same core DG model. Fig. 2.21 shows the I_d-V_g characteristics of test MG structures with different fin heights (10 to 50nm), but identical fin width (10nm). Note that the identical drain current in the subthreshold region indicates the "cap" transistor dominates the conduction current due to smaller threshold voltage (stronger electric field).
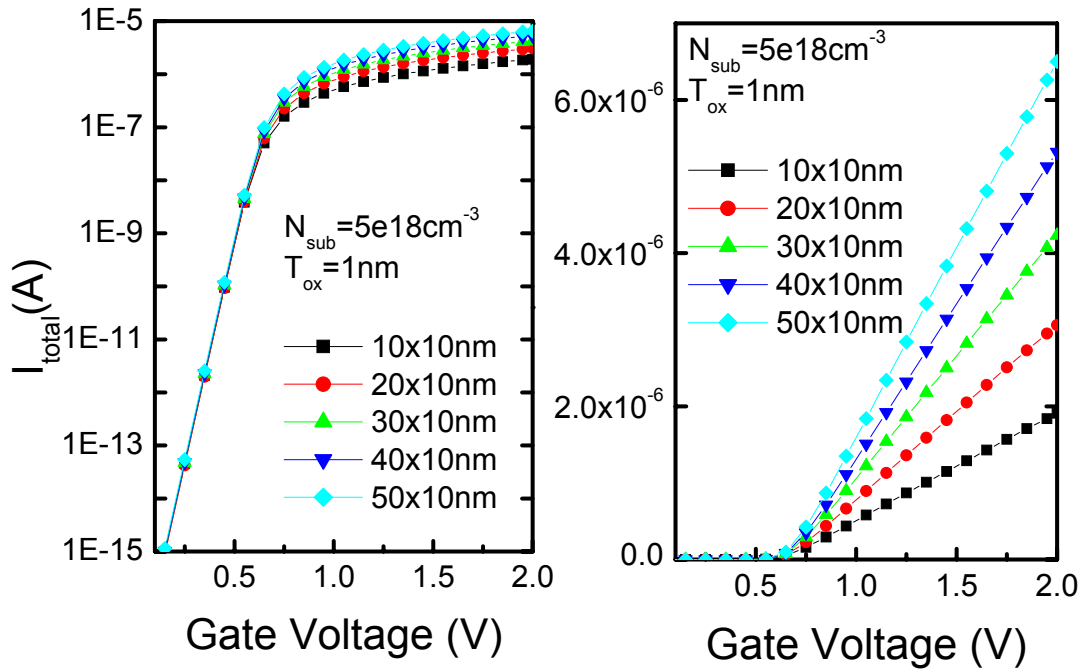
Fig. 2.21 The I_d-V_g characteristics of different MG structures, shown in (a) log scale, (b) linear scale.

The "cap" transistor current is extracted by subtracting the current flow in the flat region from the total conduction current,

$$I_{cap} = I_{total} - 2 \times I_{flat} \cdot \left(H_{fin} - \frac{1}{2} T_{fin} \right).$$

(2.12)

The I_{cap} can be fitted by the flat core DG model with different parameters as shown in Fig. 2.22. The net drain current (cap + flat) of multi-gate FET (H_{fin}=50nm and T_{fin}=10nm) matches the 3-D simulation results well (Fig. 2.23). In the subthreshold region, the cap transistor dominates the drain current, while the current is determined by both the cap and flat transistors in the strong inversion region.
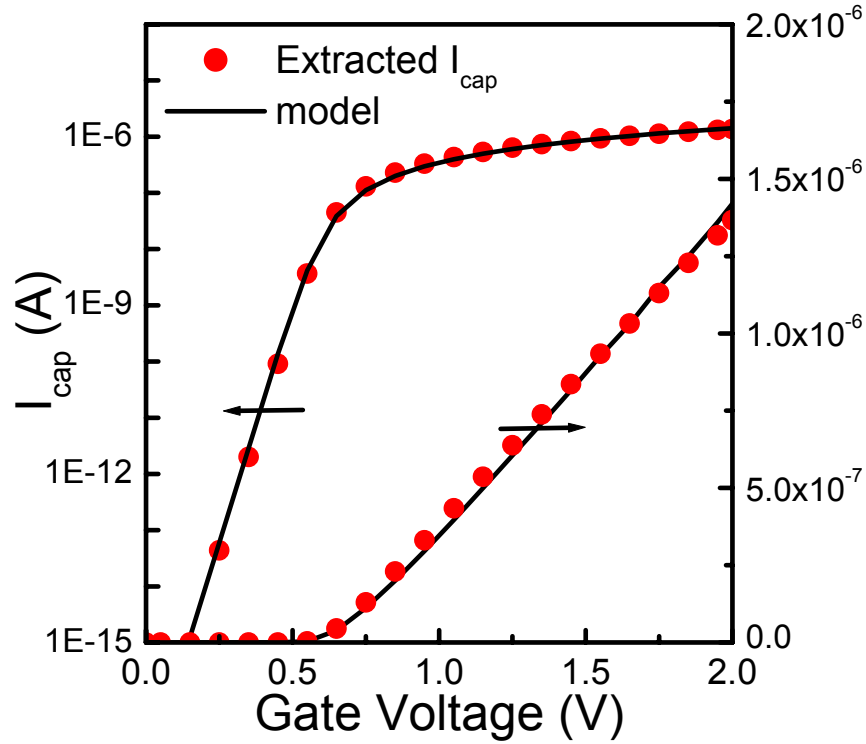
Fig. 2.22 Extracted Icap can be modeled by core DG equation with V_{fb} shift, thinner oxide thickness, and effective width.
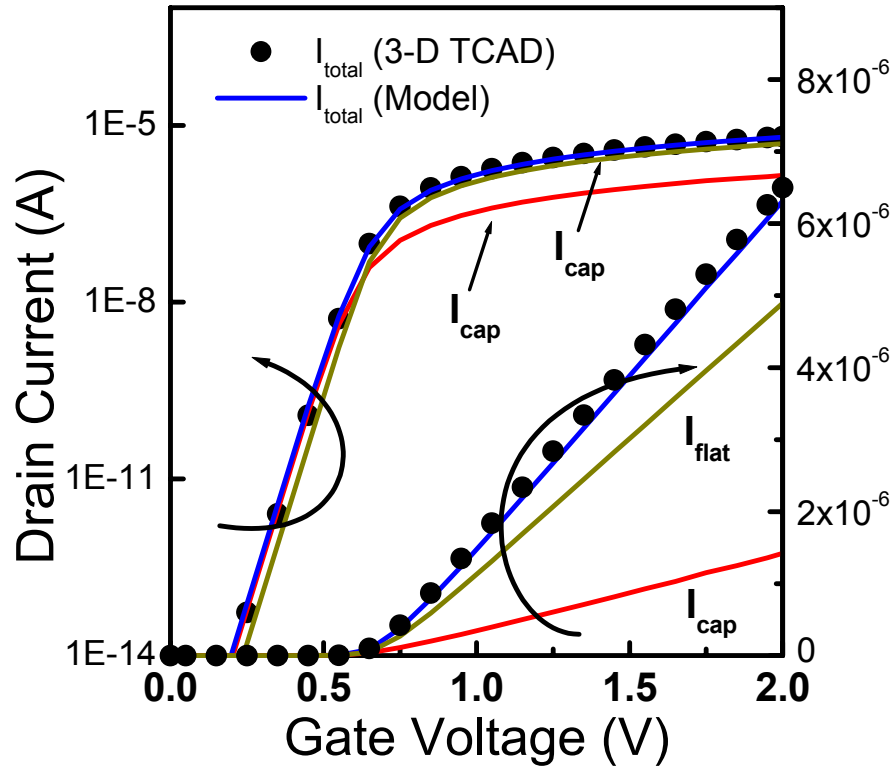
Fig. 2.23 Net current of Icap and Iflat matches 3-D simulation well. The cap transistor dominates the subthreshold region due to lower V_{th}.

2.5 Summary

A bias-dependent QM correction for the surface potential of DG MOSFETs is developed for all regimes of operation. The QM-corrected surface potential agrees with the 2-D simulation results well. Both V_{th} shift in the subthreshold and strong inversion regions and gate capacitance degradation in the strong inversion region due to QM are

corrected simultaneously. The model can predict the complicated QM effect dependence on various device parameters.

The degree of SCE (V_{th} roll-off, drain-induced-barrier-lowering (DIBL), and subthreshold slope degradation) depends on strength of gate control which is modeled by a characteristic field penetration length ($\lambda = f(T_{ox}, T_{si})$) derived from quasi 2-D Poisson's equation. The SCE model shows excellent agreements with 2-D TCAD simulation results without the use of any fitting parameters. Good scalability over T_{ox} and T_{si} down to 30nm channel length (L_g) is clearly visible. The SCE model is extended for considering the triple or more gates structures by making $\lambda = f(T_{ox}, T_{si}, H_{fin})$. The SCE model implementation captures V_{th} roll-off, DIBL and subthreshold slope degradation for short channel multi-gate FETs simultaneously.

According to 3-D device simulations, the corner conduction (at top portion) dominates the subthreshold leakage current if the channel doping is high. The corner effect can be suppressed with rounded corners, thin gate oxides, and lower channel doping. In order to model the corner effect, a cap transistor model is introduced. The height of cap transistor is half of the fin width independent of fin height, which can be explained by the charge sharing concept.

2.6 References

[2.1]   Y. King, H. Fujioka, S. Kamohara, K. Chen, and C. Hu, "DC Electrical Oxide

Thickness model for Quantization of the inversion layer in MOSFETs," Semicond.

Sci. Technol., 13, p. 963 (1998).

[2.2]   B. Majkusiak, T. Janik, and J. Walczak, "Semiconductor Thickness Effects in the

Double-Gate SOI MOSFET," Tran. Electron Devices, 45, p.1127 (1998).

[2.3]   L. Ge, and J. Fossum, "Analytical modeling of quantization and volume inversion

in thinSi-film DG MOSFETs," Tran. Electron Devices, 49, p.287 (2002).

[2.4]   V. Trivedi, and J. Fossum, "Quantum-Mechanical Effects on the Threshold

Voltage of Undoped Double-Gate MOSFETs," Electron Devices Lett., 26, p.579

(2005).

[2.5]   Y. Li and S.-M. Yu, "A unified quantum correction model for nanoscale single-

and double-gate MOSFETs under inversion conditions," Nanotechnology, 15, p.

1009 (2004).

[2.6]   D. Vasileska and Z. Ren, SCHRED-2.1 Manual (2000).

[2.7]   M. V. Dunga, C.-H. Lin, D. D. Lu, W. Xiong, C. R. Cleavelin, P. Patruno, J.-R.

Hwang, F.-L. Yang, A. M. Niknejad and C. Hu, "BSIM-MG: A Versatile

Multi-Gate FET Model for Mixed-Signal Design," Proceedings of the VLSI

Technology Symposium, pp.. 80-81, 2007.

[2.8] X. Liang and Y. Taur, "A 2-D Analytical Solution for SCEs in DG MOSFETs," IEEE Trans. Electron Devices, vol. 51, no. 9, pp. 1385-1391, 2004.

[2.9] Q. Chen, E. M. Harrell, and J. D. Meindl, "A Physical Short-channel Threshold Voltage Model for Uundoped Symmetric Double-Gate MOSFETs," IEEE Trans. Electron Devices, vol. 50, no. 7, pp. 1631-1637, 2003.

[2.10] K. Suzuki, T. Tanaka, Y. Tosaka, H. Horie, and Y. Arimoto, "Scaling Theory for Double-Gate SOI MOSFETs," IEEE Trans. Electron Devices, vol. 40, no. 12, pp. 2326-2329, 1993.

[2.11] K. Suzuki, Y. Tosaka, and T. Sugii, "Analytical Threshold Voltage Model for Short Channel n+/p+ Double-Gate SOI MOSFETs,," IEEE Trans. Electron Devices, vol. 43, no. 5, pp. 732-738, 1996.

[2.12] K. K. Young, "Short-Channel Effect in Fully Depleted SOI MOSFET's," IEEE Trans. Electron Devices, vol. 36, no. 2, pp. 399-402, 1989.

[2.13] G. Pei, J. Kedzierski, P. Oldiges, M. Ieong, and E. Kan, "FinFET Design Considerations Based on 3-D Simulation and Analytical Modeling," IEEE Trans. Electron Devices, vol. 48, no. 8, pp. 1441-1419, 2002.

[2.14] B. Doyle, B. Boyanov, M. Doczy, S. Hareland, B. Jin, J. Kavalieros, T. Linton, R.

Rios, and R. Chau, "Tri-Gate Fully-Depleted CMOS Transistors: Fabrication, Design and Layout," in Proc. on VLSI Tech. Symp., 2003, pp. 133-134.

[2.15] M. Stadele, R. J. Luyken, M. Roosz, M. Specht, W. Rosner, L. Dreeskornfeld, J. Hartwich, F. Hofmann, J. Kretz, E. Landgraf, and L. Risch, "A comprehensive study of corner effects in tri-gate transistors," in Proc. ESSDERC, 2004, pp. 165-168.

[2.16] J. G. Fossum, J.-W. Yang, V. P. Trivedi, "Suppression of corner effects in triple-gate MOSFETs," IEEE EDL Vol. 24, pp. 745-747, 2003.

[2.17] "DESSIS", TCAD release 9.0 Manual, Synopsys Inc., Mountain View, CA

Chapter 3

BSIM-CMG: A Compact Model for Symmetric/ Common-Gate Multi-Gate MOSFETs

3.1 Introduction

The scaling of conventional planar CMOS is expected to become increasingly difficult due to increasing gate leakage and subthreshold leakage [3.1-3.2]. Multi-gate FETs such as FinFETs have emerged as the most promising candidates to extend the CMOS scaling into the sub-25nm regime [3.3-3.4]. The strong electrostatic control over the channel originating from the use of multiple gates reduces the coupling between source and drain

48

in the subthreshold region and it enables the multi-gate transistor to be scaled beyond

bulk planar CMOS for a given dielectric thickness. Numerous efforts are underway to

enable large scale manufacturing of multi-gate FETs. At the same time, circuit designers

are beginning to design and evaluate multi-gate FET circuits.

A compact model serves as a link between process technology and circuit design. It is a

concise mathematical description of the complex device physics in the transistor. A

compact model maintains a fine balance between accuracy and simplicity. An accurate

model based on physics basis allows the process engineer and circuit designer to make

projections beyond the available silicon data (scalability) for scaled dimensions and also

enables fast circuit/device co-optimization. The simplifications in the physics enable very

fast analysis of device/circuit behavior when compared to the much slower numerical

based TCAD simulations. It is thus necessary to develop a compact model of multi-gate

FETs for technology/circuit development in the short term and for product design in the

longer term.

One of the biggest challenges in modeling multi-gate FETs is the need to model

several flavors of multi-gate FETs. The silicon body can be controlled by either two gates

or three gates or four gates. The gates can all be electrically interconnected or they can be

biased independently. Multi-gate FETs can be built on SOI or bulk silicon. Fig. 3.1

illustrates some of the different architectures of multi-gate FETs which need to be accounted in the compact model. It is important to obtain a versatile model which can model all the different types of multi-gate FETs without making the model computationally intensive. One possible technique to handle the different multi-gate FET architectures is to classify them into two categories and introduce a separate model for each category: a symmetric/common gate model and an asymmetric/independent gate model. The term "common-gate" means that all the gates in the multi-gate FET (double-gate or triple-gate or quadruple-gate FinFET) are electrically interconnected and are biased at the same electrical gate voltage. The common-gate model further assumes that the gate work-functions and the dielectric thicknesses on the two, three or four active sides of the fin are the same. However, the carrier mobilities in the inversion layers on the horizontal and vertical active sides of the fin can be different due to different crystal orientations and/or strain. The asymmetric/independent gate model allows different work-functions and dielectric thicknesses on the two sides of the fin. The asymmetric/independent gate model also permits that the two gates can be biased independently.

The existing modeling efforts for the multi-gate FETs are limited to undoped or lightly doped silicon body for double-gate (DG) FETs [3.5-3.7]. A multiple threshold voltage

(a) Triple-gate FinFET on SOI   (b) Double-gate FinFET on SOI

(c) Triple-gate FinFET on bulk Si   (d) Planar double-gate FinFET
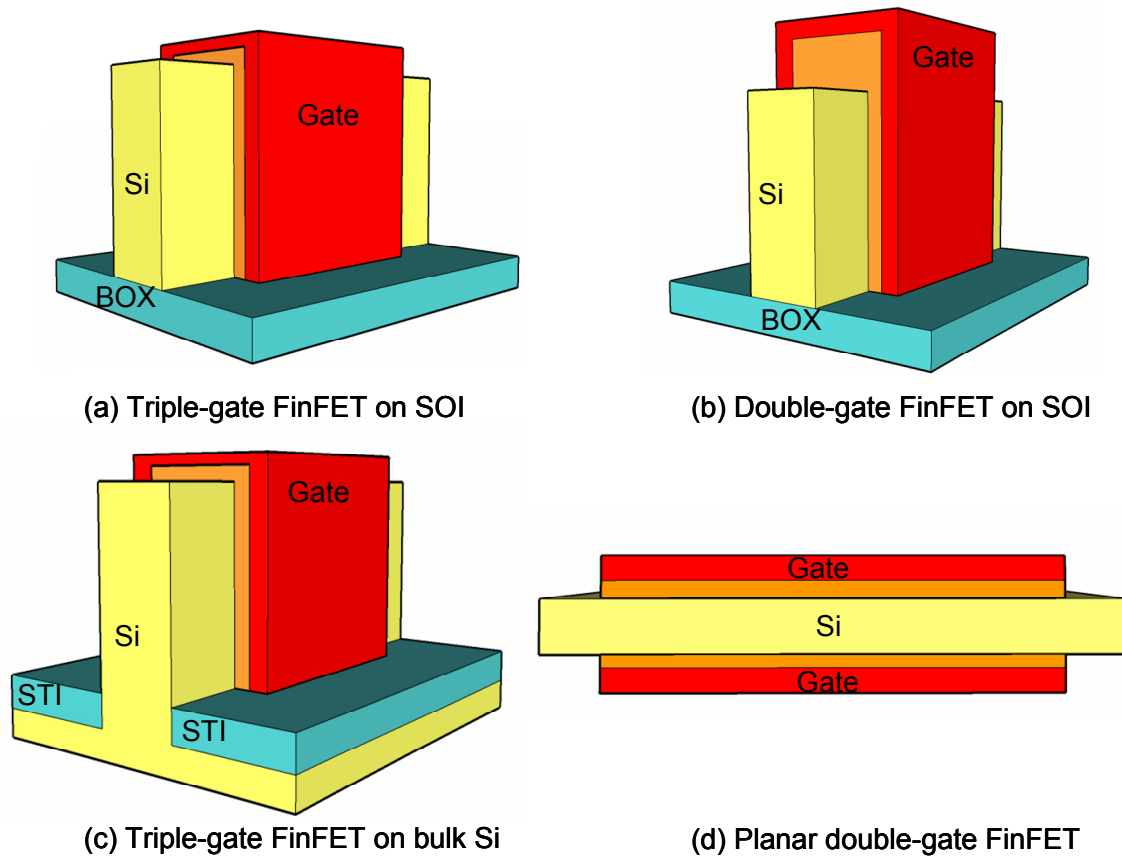
Fig. 3.1 Different possible architectures of multi-gate FET

technology based on symmetric/common-gate multi-gate FETs will likely require a

significant concentration of body dopant for threshold voltage tuning. As a result, we

have developed a core model considering the effect of finite body doping on the electrical

characteristics of a multi-gate FET in the Poisson's equation [3.8]. Combining with

numerous physical effects, such as quantum mechanical effect, short channel effects,

mobility degradation, and carrier velocity saturation, BSIM-CMG (Berkeley

Short-channel IGFET Model – Common Multi-Gate) has been verified with the measured electrical characteristics of FinFETs [3.9]. BSIM-IMG (Berkeley Short-channel IGFET Model – Independent Multi-Gate) has been developed and verified with TCAD simulation results [3.10].

In this chapter, the core formulation and physical effects of real device modeled in BSIM-CMG is presented. The surface potential is calculated analytically from the coupled solution of Poisson's equation and Gauss's law. The calculated surface potential agrees very well with TCAD simulation. The I-V model is obtained using drift-diffusion formulation in terms of surface potential without using any charge-sheet approximation. The intrinsic capacitance model is derived in terms of surface potential as well. Both long channel I-V model and intrinsic capacitance model agree with 2-D device simulation very well without using any fitting parameter. Short channel effects are modeled by a characteristic field penetration length. The complete model is verified against two different FinFET technologies – SOI FinFETs and bulk FinFETs. BSIM-CMG was able to describe the drain current and its derivatives for long and short channel FETs for both technologies.

3.2 Surface Potential Calculation

In BSIM-CMG, all electrical variables such as terminal currents, charges and capacitances are expressed as functions of the surface potentials at the source and the drain end. The calculation of the surface potentials forms the basis of the model. The core model for BSIM-CMG is a long channel double-gate FET model. Numerous physical phenomena observed in an advanced multi-gate FET technology are added to the core model to yield the final model. Fig. 3.2 shows the schematic of the symmetric/common-gate DG-FET under study.
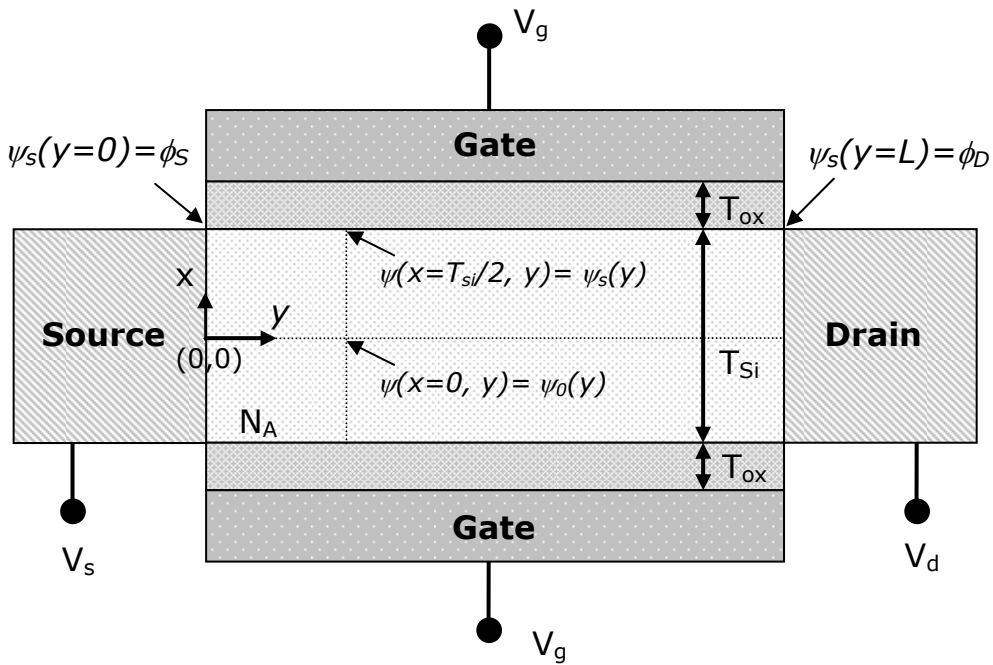

Fig. 3.2 Schematic of the symmetric common-gate DG-FET under study

The electric potential in the body is obtained by solving the Poisson's equation. For a long channel transistor, the gradual channel approximation is used which states that the horizontal electric field is much smaller than the vertical electric field. The use of gradual channel approximation results in a 1-D Poisson's equation (in the vertical dimension).

The 1-D Poisson's equation including both inversion carriers and bulk charge in the body can be written as

$$\frac{\partial^2 \psi(x,y)}{\partial x^2} = \frac{qn_i}{\varepsilon_{si}} \cdot e^{\frac{q(\psi(x,y)-\phi_B-V_{ch}(y))}{kT}} + \frac{qN_A}{\varepsilon_{si}} \tag{3.1}$$

where $\psi(x,y)$ is the electronic potential in the body, $V_{ch}(y)$ is the channel potential $(V_{ch}(0)=0$ and $V_{ch}(L)=V_{ds})$, N_A is the body doping and

$$\phi_B = \frac{kT}{q} \cdot \ln\left(\frac{N_A}{n_i}\right) \tag{3.2}$$

For a lightly doped body, the body doping can be neglected and Eq. (3.1) can be solved easily by integrating it twice and using Gauss's law as the boundary condition [3.5]. However, for moderate to heavy body doping, the doping term cannot be neglected and it complicates the calculation of surface potential as Eq. (3.1) cannot be integrated analytically twice. To overcome this limitation, a perturbation approach is proposed by Dunga et al. to solve the Poisson's equation in presence of significant body doping [3.8].

The computationally efficient correction term of surface potential, ψ_{pert}, is derived, which

is continuous between partial-depletion and full-depletion regimes.

$$\psi_{pert} = MIN\left(\psi_c, \psi_{bulk}\right) \tag{3.2}$$

where ψ_c and ψ_{bulk} the perturbation potential in the partial depletion and full depletion

regimes, respectively.

$$\psi_c = \left(-\sqrt{\frac{q\varepsilon_{Si}N_A}{2C_{ox}}} + \sqrt{\frac{q\varepsilon_{Si}N_A}{2C_{ox}} + V_{gs} - V_{fb}}\right)^2 \tag{3.3}$$

$$\psi_{bulk} = \frac{1}{2}\frac{qN_A}{\varepsilon_{Si}}\left(\frac{T_{si}}{2}\right)^2 \tag{3.4}$$

By using ψ_{pert}, the surface potential in both the regimes is calculated through a single

continuous equation. Through a simple transformation of variables,

$$\beta = \frac{T_{Si}}{2}\sqrt{\frac{q^2}{2\varepsilon_{Si}kT}\frac{n_i^2}{N_A}e^{\frac{q(\psi_0(y)-V_{ch}(y))}{kT}}} \tag{3.5}$$

the unified surface potential (ψ_s) equation used in the core model for BSIM-CMG is

written as

$$f_0(\beta) = \ln(\beta) - \ln\left(\cos(\beta)\right) - \frac{V_{gs} - V_{fb} - V_{ch} - \psi_{pert}}{2\frac{kT}{q}} - \ln\left(\frac{T_{Si}}{2}\sqrt{\frac{q^2}{2\varepsilon_{Si}kT}\frac{n_i^2}{N_A}}\right) +$$

$$\frac{2\varepsilon_{Si}}{T_{Si}C_{ox}}\cdot\sqrt{\beta^2\left(\frac{e^{\frac{q\psi_{pert}}{kT}}}{\cos^2(\beta)} - 1\right) + \frac{\psi_{pert}}{\left(kT/q\right)^2}\left(\psi_{pert} - 2\frac{kT}{q}\ln\left(\cos(\beta)\right)\right)} = 0 \tag{3.6}$$

where β is the only unknown variable. In BSIM-CMG, the transcendental ψ_s equation (Eq.

(3.6)) is solved for β using an analytical approximation instead of iterative methods to make the model numerically robust and computationally efficient.

The surface potential is given by

$$\psi_s = V_{ch} + 2\frac{kT}{q}\left(\ln(\beta) - \ln(\cos(\beta)) - \ln\left(\frac{T_{Si}}{2}\sqrt{\frac{q^2}{2\varepsilon_{Si}kT}\frac{n_i^2}{N_A}} \right) \right) + \psi_{pert}. \tag{3.7}$$

Surface potential at source terminal (ϕ_S) is obtained by solving Eq. (3.6) at the source end, i.e $V_{ch} = 0$. Similarly, the surface potential at the drain end (ϕ_D) is calculated by solving Eq. (3.6) with $V_{ch} = V_{ds}$.

Fig. 3.3 compares the surface potential calculated using the model against TCAD. All TCAD simulations for verification of the core model use gate material with mid-gap workfunction and assume constant carrier mobility. The surface potential is calculated as a function of gate voltage for a wide range of body doping ranging from a light doping of 10^{15}cm^{-3} to heavy doping of 5×10^{18}cm^{-3}. Very good agreement is observed between the model and TCAD for all cases. The transition from partial-depletion regime to full-depletion regime with increasing gate voltage is clearly visible in the heavily doped DG-FET. The error in the analytical approximation of ψ_s is limited to only a few nano-volts as shown in Fig. 3.4.
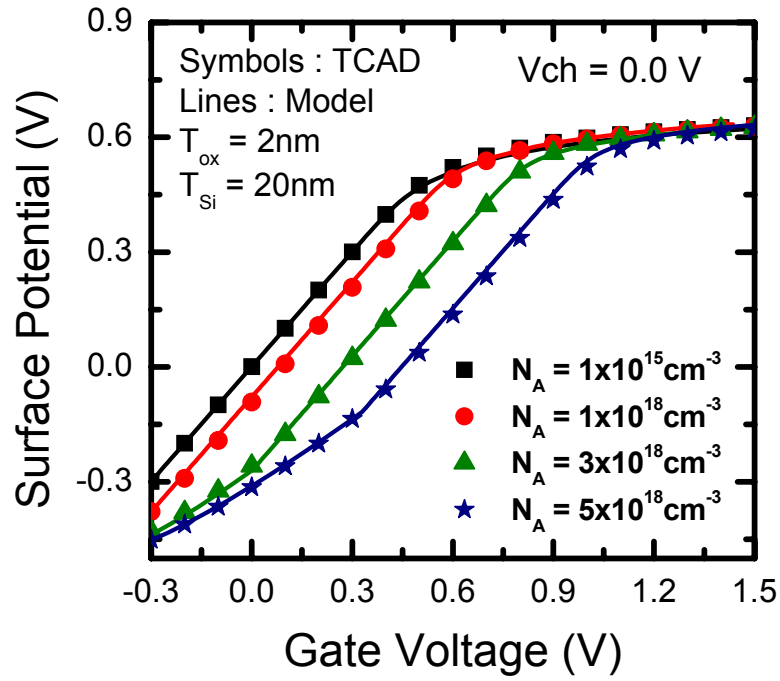
56

Fig. 3.3 Comparison between surface potential solution and TCAD simulation results for different body doping
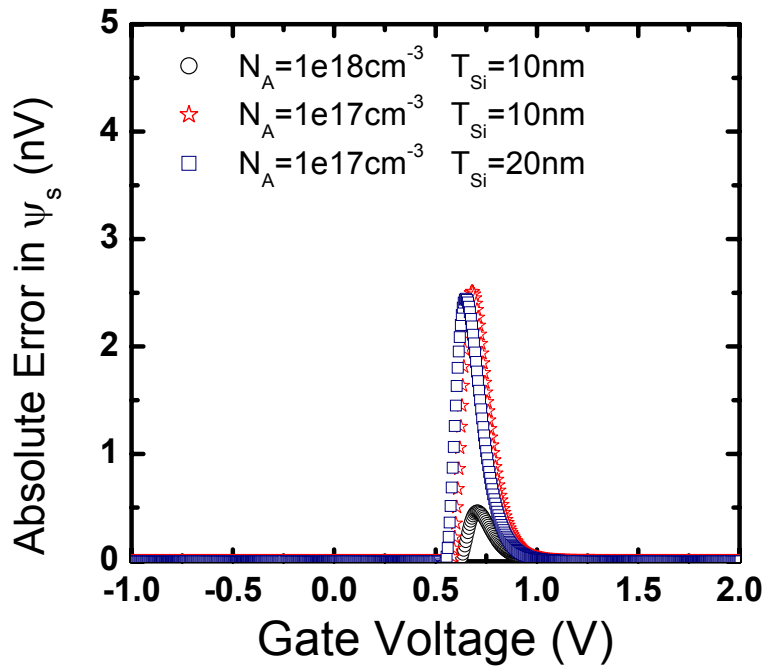

Fig. 3.4 Error in the analytical approximation of ψ_s.. The error is limited to a few nano-volt for wide range of body doping and fin thickness.

Eq. (3.6) yields surface potential for both light and heavy body doping as shown in Fig. 3.3. The perturbation method yields surface potential in both full-depletion and partial-depletion regimes. However, the inclusion of bulk charge in the analysis of a lightly doped body is redundant and it leads to significant overhead in model runtime. For a lightly doped device, Eq. (3.6) can be simplified into

$$\ln\beta - \ln(\cos\beta) - \frac{V_{gs} - V_{fb} - V_{ch}}{2\frac{kT}{q}} - \ln\left(\frac{T_{Si}}{2}\sqrt{\frac{q^2}{2\varepsilon_{Si}kT}\frac{n_i^2}{N_A}}\right) + \frac{2\varepsilon_{Si}}{T_{Si}C_{ox}}\beta\tan\beta = 0 \tag{3.8}$$

Based on this insight, the computational efficiency of BSIM-CMG can be further improved by setting the model parameter PHISMOD=1 if the channel doping concentration is below mid 10^{17}cm^{-3}. Fig. 3.5 shows the normalized computational time of surface potential calculation between the iterative approach and analytical approximation approach. The computational efficiency of analytical approximation approach is ~35% faster compared to iterative approach. The computational efficiency is further improved for another 15% for the lightly doped device.
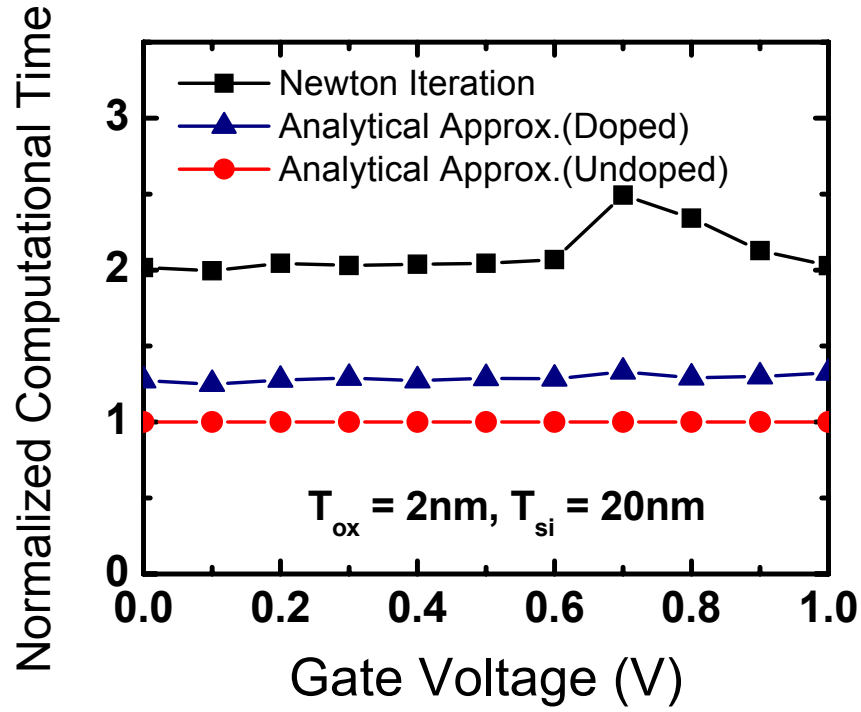
Fig. 3.5 Normalized computational efficiency of the different surface potential solving methodologies

3.3 Drain Current Model

The I-V model is obtained using drift-diffusion formulation without using any charge-sheet approximation [3.11]. The current flowing through the body of a DG-FET can be written as

$$I_d(y) = 2 \cdot \mu \cdot W \cdot Q_{inv}(y) \frac{dV_{ch}(y)}{dy} \tag{3.9}$$

where μ is the carrier mobility (assumed position independent), W is the channel width, $Q_{inv}(y)$ is the inversion charge in one half of the body and the factor of two accounts for the front and back channel currents in a symmetric/common-gate DG-FET. Under the quasi-static operation of transistor, the drain current is identical at any points between source and drain. Eq. (3.9) can be integrated from source to drain yield

$$I_d = 2 \cdot \mu \cdot \frac{W}{L} \cdot \int_0^{V_{ds}} Q_{inv}(y) \cdot \frac{dV_{ch}(y)}{dy} \tag{3.10}$$

where L is the channel length. The inversion charge (Q_{inv}) is simply the difference between the total charge in the body and the bulk charge,

$$Q_{inv}(y) = Q_{total}(y) - Q_{bulk}(y). \tag{3.11}$$

Bulk charge (Q_{bulk}) can be obtained from the perturbation potential ψ_{pert},

$$Q_{bulk} = \sqrt{2q\varepsilon_{Si}N_A\psi_{pert}}. \tag{3.12}$$

Gauss's Law can be used to determine the total charge in the body,

$$Q_{total}(y) = C_{ox} \cdot \left(V_{gs} - V_{fb} - \psi_s(y)\right). \tag{3.13}$$

The gradient in the quasi Fermi potential can be expressed in terms of Q_{inv}

$$\frac{dV_{ch}}{dy} = \frac{d\psi_s}{dy} + \frac{kT}{q}\frac{dQ_{inv}}{dy}\left(\frac{2Q_{bulk} + 5\frac{kT}{q}\frac{\varepsilon_{si}}{T_{si}}}{Q_{inv} + 2Q_{bulk} + 5\frac{kT}{q}\frac{\varepsilon_{si}}{T_{si}}} - \frac{2}{Q_{inv}}\right). \tag{3.14}$$

Analytical expression for drain current can be obtained by using Eqs. (3.11-3.14) in Eq.

(3.10) and integrating the resulting expression from source to drain. The drain current can

be expressed as difference of two terms evaluated at the source and drain ends.

$$I_d = 2 \cdot \mu \cdot \frac{W}{L} \cdot \left(f\left(Q_{invs}\right) - f\left(Q_{invd}\right) \right) \tag{3.15}$$

where Q_{invs} and Q_{invd} are the inversion charge density at the source and drain ends, and

the function f(Q) is

$$f(Q) = \frac{Q^2}{2C_{ox}} + 2\frac{kT}{q}Q - \frac{kT}{q}\left(2Q_{bulk} + 5\frac{kT}{q}\frac{\varepsilon_{Si}}{T_{Si}}\right) \cdot \ln\left(1 + \frac{Q}{5\frac{kT}{q}\frac{\varepsilon_{Si}}{T_{Si}} + 2Q_{bulk}}\right). \tag{3.16}$$

Eqs. (3.15-3.16) predict the drain current for a symmetric/common-gate DG-FET and

constitute the core I-V model for BSIM-CMG.

The accuracy and predictivity of the I-V model is verified against TCAD simulations

without using any fitting parameters. Fig. 3.6 shows the model predicted and TCAD

simulated I_d-V_{gs} and I_d-V_{ds} characteristics for a heavily doped DG-FET ($N_A = 3 \times 10^{18} cm^{-3}$).

BSIM-CMG can predict very accurate drain current in all the regimes of transistor
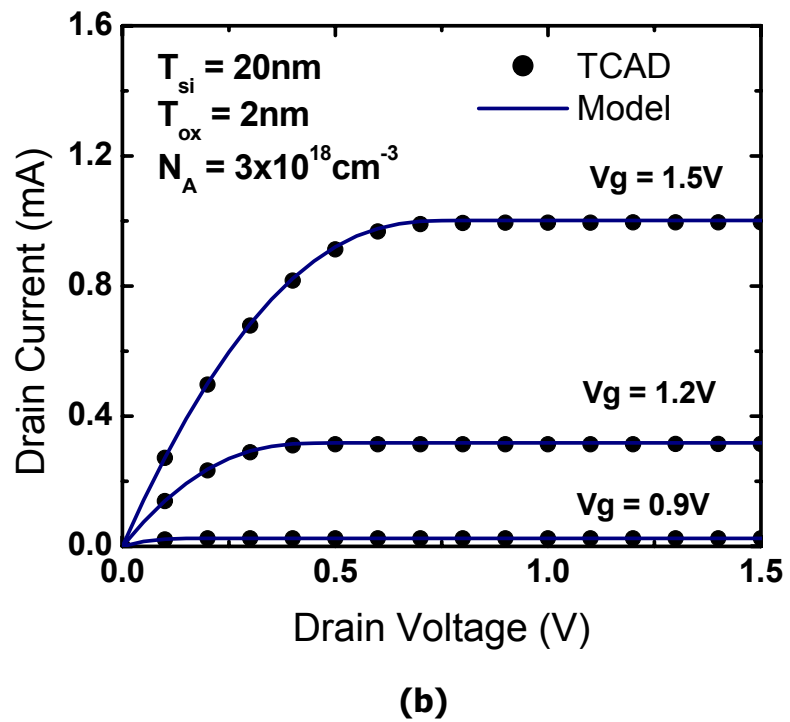
operation: sub-threshold, linear and saturation.

(a)
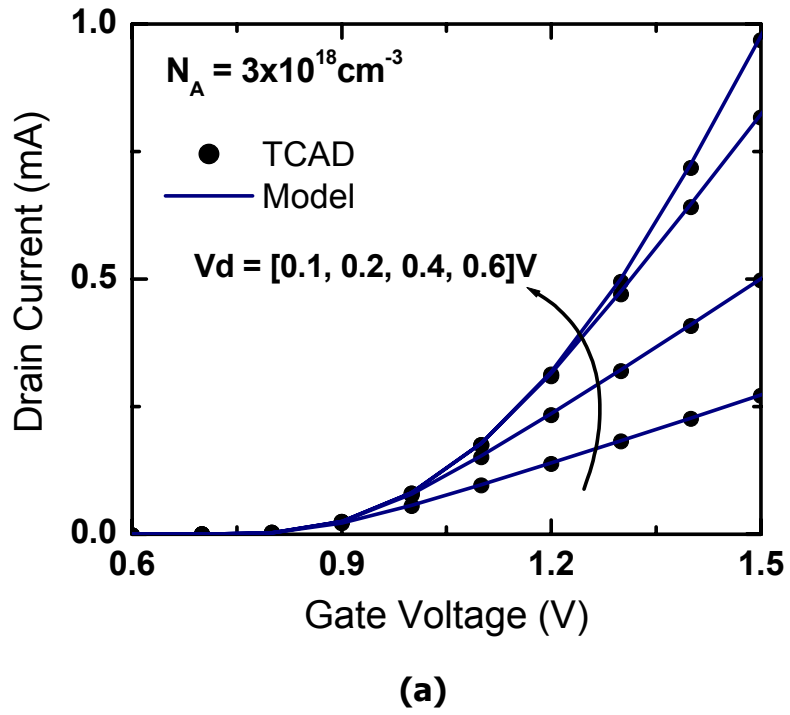

(b)

Fig. 3.6 (a) I_d-V_{gs} and (b)I_d-V_{ds} characteristics of a DG-FET with heavy body doping of $N_A = 3 \times 10^{18} cm^{-3}$. The model (lines) agree very well with 2-D TCAD simulations in all regimes of operation

One main feature of BSIM-CMG core model is the capability of predicting drain current over a wide range of body doping. As shown in Fig. 3.7, the model predicts the accurate drain current in both full-depletion and partial-depletion regimes compared to 2-D TCAD simulation results without using any fitting parameters. The transition from partial depletion to full depletion is very smooth as for the case of $N_A = 1x10^{19}cm^{-3}$.
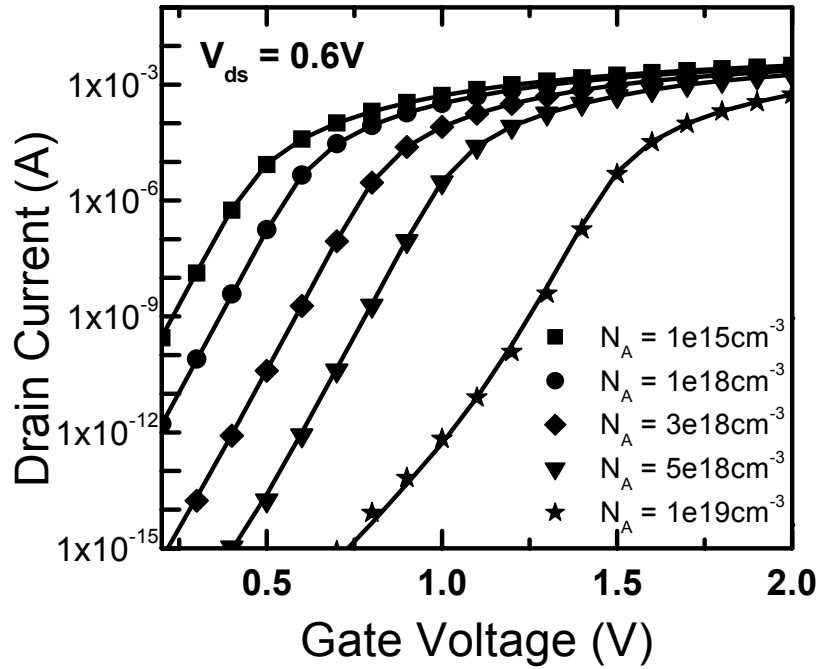

Fig. 3.7 Verification of the I-V model against 2-D TCAD for both lightly doped and heavily doped DG-FETs. I-V model agrees very well with 2D TCAD for large range of body doping.

A unique behavior of lightly doped DG-FET with a thin body is volume inversion. Due to the absence of bulk charge in the body and presence of only few inversion carriers in

the sub-threshold regime, there is negligible potential drop between the surface and center of the body. Fig. 3.8(a) shows the virtually flat potential profile in lightly doped DG-FETs in the subthreshold regime for different body thickness. The potential in the body has a very weak dependence on body thickness. Any small increase in gate voltage in subthreshold regime increases the potential through the entire body causing inversion in the entire body. As a result, the inversion carrier density is nearly constant through the whole conduction channel in the subthreshold regime. This phenomenon is called bulk inversion or volume inversion [3.12]. Since the electronic potential is virtually independent of body thickness, the amount of total inversion carriers in the body is linearly proportional to the body thickness for equal area DG-FETs. As a result, the subthreshold current in lightly doped DG-FETs is a linear function of the body thickness. The I-V model is able to predict this trend correctly as shown in Fig. 3.8(b) where the I_d for a 20nm thick body is ~4x of current flowing in a 5nm thick body in the subthreshold regime. The body doping used in the simulation is $10^{15}cm^{-3}$.
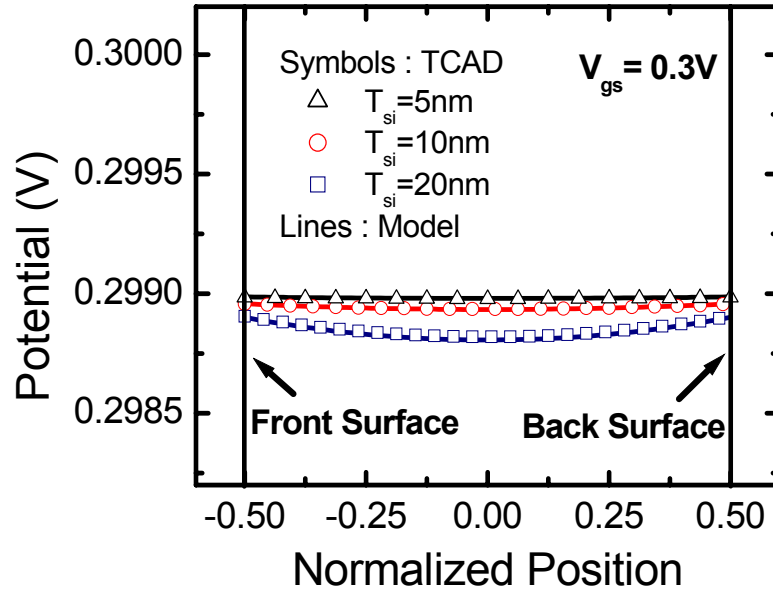
(a)
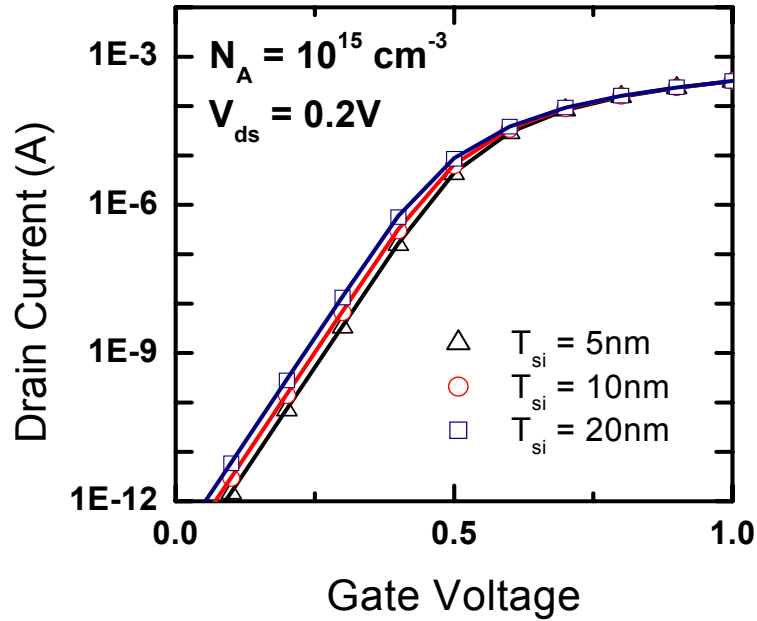
(b)

Fig. 3.8 Volume inversion in lightly doped DG-FETs. (a) Potential profile in the body (b) Subthreshold regime I_d-V_{gs} characteristics for different body thickness showing volume inversion.

3.4 Charge and Capacitance Model

The C-V model defines both the terminal charges and the associated capacitances for the transistor, which are essential for AC and transient analysis. In DG-FET, the charge on the top and bottom gate electrodes equals to the total charge in the body. The total charge is calculated by integrating the charge along the channel. The inversion charge in the body is divided between the source and drain terminals using the Ward-Dutton charge partition approach [3.13]. The terminal charges are formulated in terms of surface potential at source and drain ends.

$$Q_g = 2WLC_{ox}\left(V_{gs} - V_{fb} - \frac{\phi_S + \phi_D}{2} + \frac{(\phi_D - \phi_S)^2}{6(B - \phi_D - \phi_S)}\right)$$

$$Q_d = -2WLC_{ox}\left(\begin{array}{l} \dfrac{V_{gs} - V_{fb} - Q_{bulk}/C_{ox}}{2} - \dfrac{\phi_S + \phi_D}{4} + \dfrac{(\phi_D - \phi_S)^2}{60(B - \phi_D - \phi_S)} \\ + \dfrac{(5B - 4\phi_D - 6\phi_S)(B - 2\phi_D)(\phi_S - \phi_D)}{60(B - \phi_D - \phi_S)^2} \end{array}\right) \tag{3.17}$$

$$Q_s = -\left(Q_g + Q_{bulk} + Q_d\right)$$

where

$$B = 2 \cdot \left(V_{gs} - V_{fb} - \frac{Q_{bulk}}{C_{ox}} + 2\frac{kT}{q}\right). \tag{3.18}$$

The expressions for terminal charges are continuous and are valid over the sub-threshold, linear and saturation regimes of operation. Fig. 3.9 shows the terminal charges calculated using Eq. (3.17) as a function of V_{ds} and V_{gs}. The ratio of the drain charge to source ratio is 40/60 in the saturation region as seen in Fig. 3.9. This is due to

Ward-Dutton charge partition which is physically correct under the quasi-static condition.
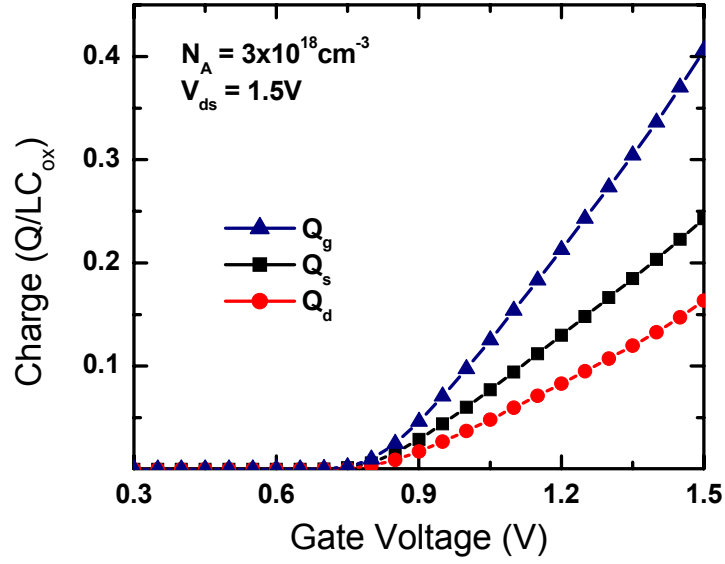
(a)
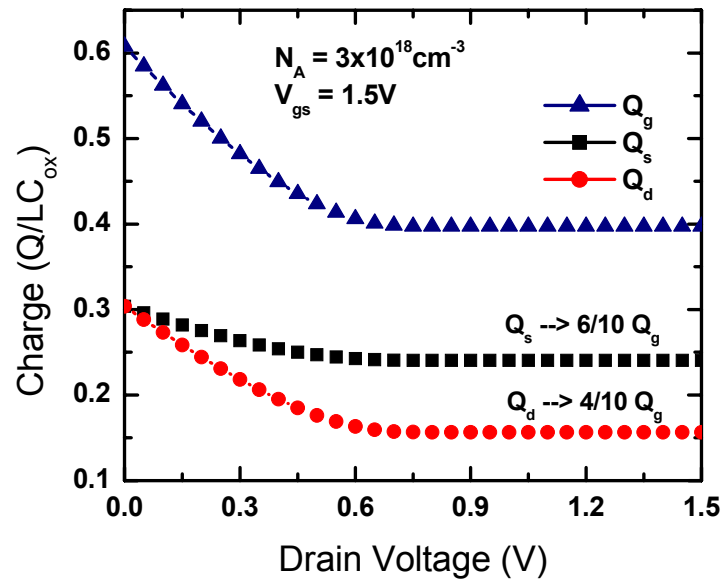
(b)

Fig. 3.9 Terminal charges of DG-FET with $N_A= 3 \times 10^{18} cm^{-3}$ calculated using Eq. (3.31) as a function of (a) V_{gs} and (b) V_{ds}.

Eq. (3.17) forms the C-V model for BSIM-CMG. The terminal charges are used as state variables in the circuit simulation. All the capacitances are derived from the terminal charges to ensure charge conservation. The capacitances are defined as

$$C_{ij} = \frac{\partial Q_i}{\partial V_j} \tag{3.19}$$

where i and j denote the multi-gate FET terminals.

The C-V model is verified against TCAD simulations without using any fitting parameters. Fig. 3.10 shows that the capacitance values from the model are in excellent agreement with TCAD simulated values in all regimes of transistor operation. Fig. 3.10(b) shows that $C_{gs}=C_{gd}$ at $V_{ds} = 0V$. This equality in capacitances at $V_{ds} = 0V$ demonstrates the symmetry of the core model. Model symmetry is important for predicting correct distortion metrics for circuits switching about $V_{ds} = 0V$ especially in the analog and RF domain.
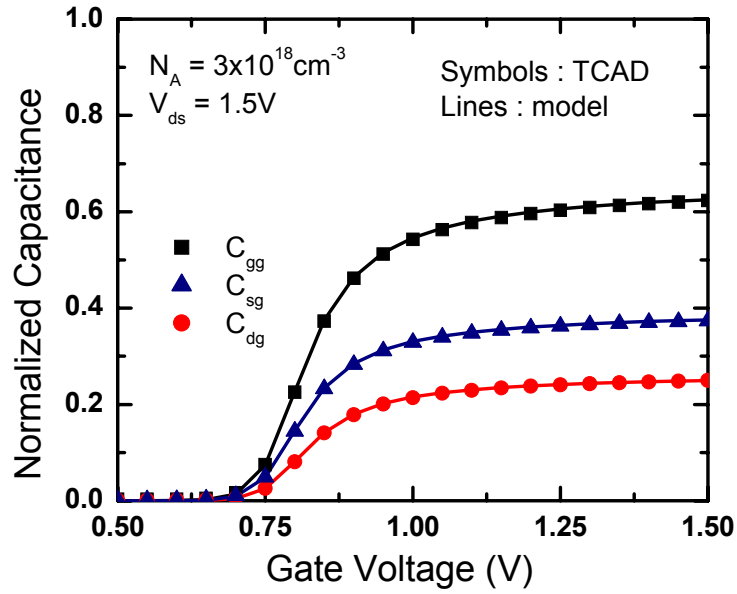
(a)
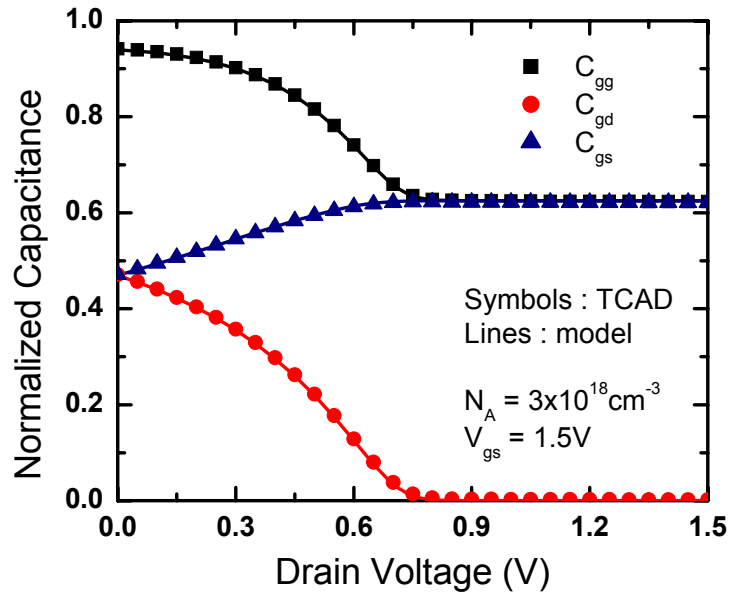
(b)

Fig. 3.10 Capacitances (normalized to $2WLC_{ox}$) calculated from the C-V model and 2-D TCAD as a function of (a) V_{gs} and (b) V_{ds}. Good agreement between the model and TCAD is seen for all the capacitances without the use of any fitting parameter. Model is symmetric at $V_{ds}=0$.

3.5 Physical Effects of Real Device

The core model is only the beginning of any compact model, which is also the case

with BSIM-CMG. BSIM-CMG, in the tradition of BSIM3 and BSIM4, models numerous

physical phenomena that are expected to be important to accurately represent advanced

multi-gate FET technologies. The thin body in multi-gate FETs experiences significant

QME through structural and electrical confinement which are modeled by modifying the

core surface potential equation and the C-V model. SCE such as drain-induced barrier

lowering (DIBL), threshold voltage (V_{th}) roll-off and sub-threshold slope degradation are

modeled. BSIM-CMG models the PDE as well since polysilicon-gated FinFETs may be

used in low-cost memories to enable continued cell size reduction. All the physical effects

included in BSIM-CMG are listed in Table 3.1.

The model symmetry is a vital requirement for certain RF circuits. This multigate

model with numerous physical effects of a real device is implemented in the Verilog-A.

Gummel Symmetry Test (GST) is used to test the symmetry of the model. Fig. 3.11

shows the example of of GST at V_{gs}=1V. BSIM-CMG maintains symmetry after careful

implementation steps.

Table 3.1 List of Physical effects modeled in BSIM-CMG

| 1	Quantum Mechanical Effects
2	Short Channel Effects
3	Vth roll-off
4	Drain Induced Barrier Lowering
5	Subthreshold Slope Degradation
6	Channel Length Modulation
7	S/D Series Resistance
8	Mobility Degradation
9	Poly Depletion
10	Velocity Saturation and Overshoot
11	Gate Induced Drain Leakage
12	Gate Tunneling
13	S/D Junction Leakage
14	Impact Ionization
15	Parasitic Capacitance
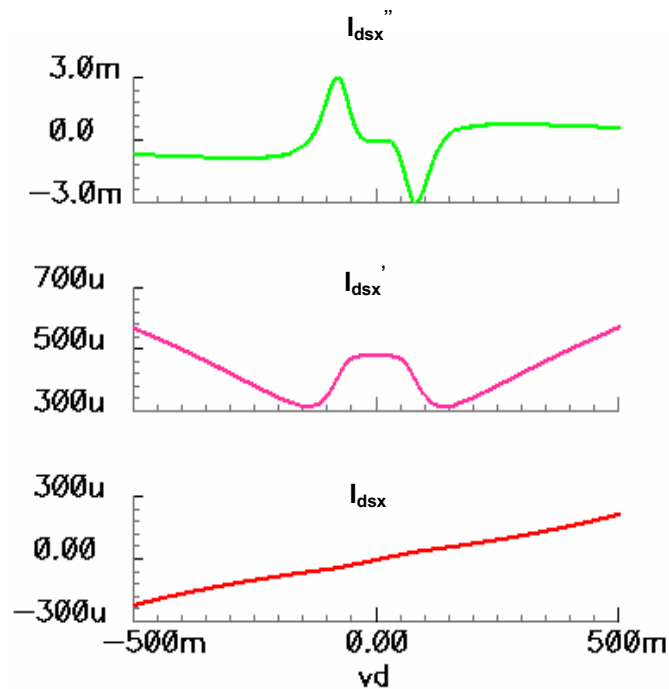
Fig. 3.11 The model passes Gummel symmetry test after physical effects of real devices are added.

3.6 Experimental Verification

BSIM-CMG has been verified against two different FinFET technologies – SOI FinFETs and bulk FinFETs [3.9]. The Verilog-A model is implemented in ICCAP environment for parameter extraction. The model successfully described the measured drain current and its derivatives, transconductance (g_m) and output conductance (g_{ds}), for both long channel and short channel multi-gate FETs.

The SOI FinFETs were fabricated on a lightly doped 60nm thick Si film with 2nm SiO_2 dielectric and a strained TiSiN gate [3.14]. The strained gate strains the channel to enhance the electron mobility, hence increasing the current drive. Measured devices had 20 parallel fins, where each fin is 22nm thick. Figure 3.12 shows the model fitting to the I_d-V_{gs} characteristics and its derivatives for short channel L_g = 90nm device in the linear and saturation regimes. Precise modeling of physical phenomenon such as DIBL, mobility degradation, and GIDL is clearly visible. Fig. 3.13 illustrates the model fitting to the I_d-V_{ds} characteristics and its derivatives for short channel L_g = 90nm device. Model captures the short channel phenomenon such as channel length modulation very well.

BSIM-CMG has also been verified against bulk FinFET measurements. Bulk FinFETs with moderate doping were fabricated with a TiN gate. Measured devices have 25nm thick fins and an EOT of 1.95nm.
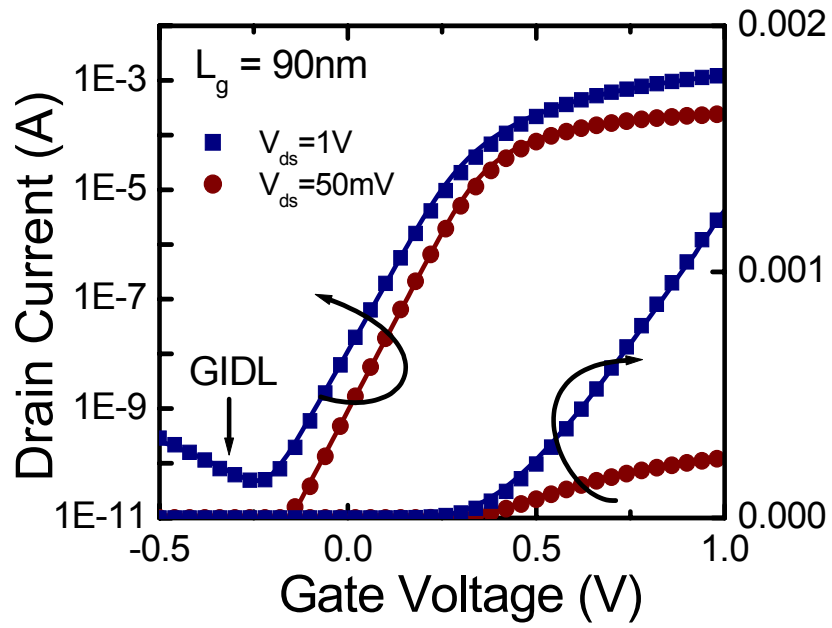
(a)
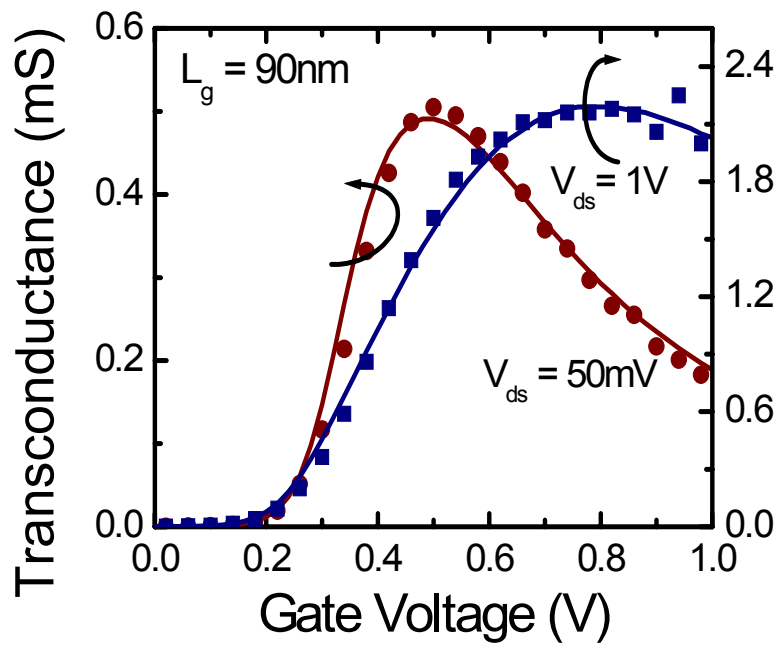
(b)

Fig. 3.12 Model fitting to short channel L = 90nm SOI FinFET measurements. (a) Drain current and (b) transconductace (g_m) as a function of V_{gs} (Symbols: measured data, Lines : BSIM-CMG model).
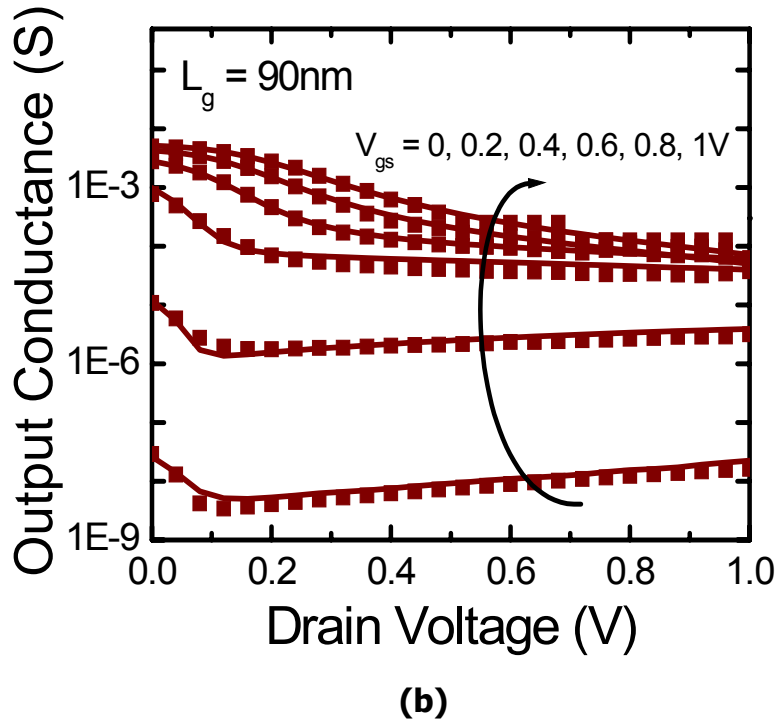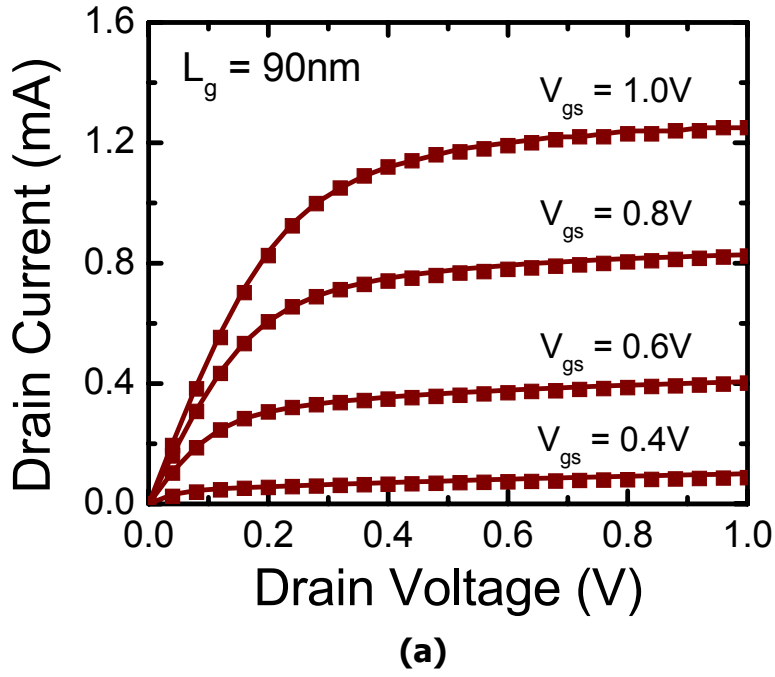
Fig. 3.13 Model fitting to short channel L = 90nm SOI FinFET measurements. (a) Drain current (Id) and (b) transconductace (g_{ds}) as a function of V_{ds} (Symbols: measured data, Lines: BSIM-CMG model).

Fig. 3.14 shows the measured short channel (L_g = 50nm) characteristics and the corresponding BSIM-CMG fitting results. BSIM-CMG is extended to bulk FinFETs by the addition of "bulk" node and substrate current model. Good agreement is observed between the model and the measured data for the long channel (L_g = 0.97 μ m) transistor as well. Derivatives of the drain current, g_m and g_{ds}, are shown for the long channel bulk FinFET in Figure 3.15. The measured bulk current due to impact ionization for the short channel bulk FinFET together with model fitting is shown in Fig. 3.16.

The experimental verification shows that BSIM-CMG accurately captures the characteristics of advanced multi-gate FETs. Triple-gate multi-gate FETs were used for model verification demonstrating the ability of the model to capture phenomena such as corner effect which are unique to tri-gate and quadruple-gate FETs. The model is able to describe both SOI and bulk silicon based multi-gate FET technologies. Accurate description of the drain current and its derivatives warrants the use of BSIM-CMG for both digital and analog design.
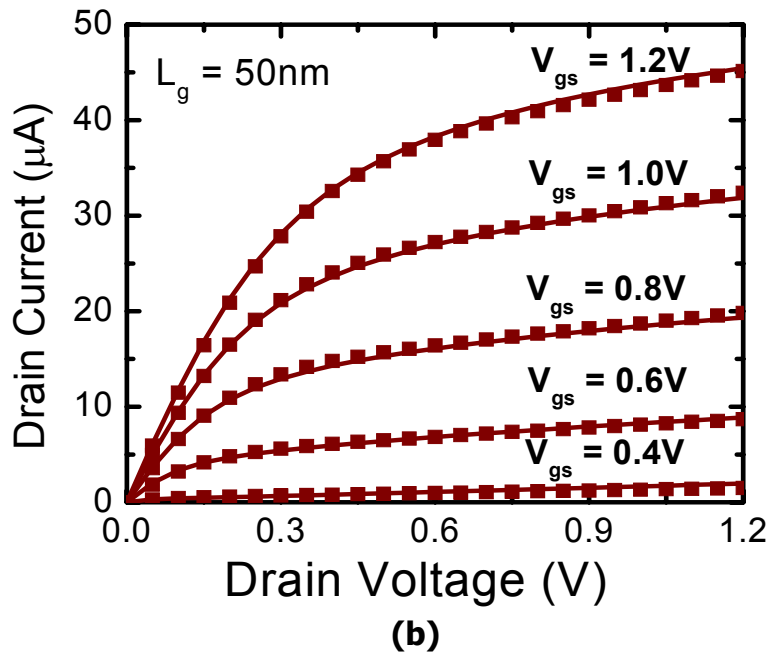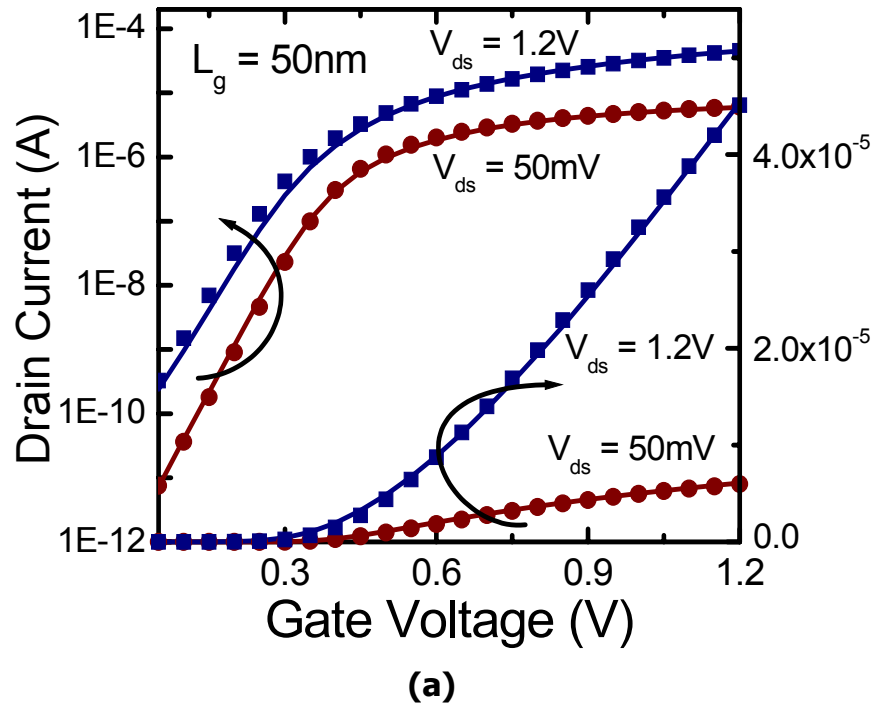
Fig. 3.14 Model fitting to short channel L = 50nm bulk FinFET measurements (a) I_d-V_{gs} and (b) I_d-V_{ds} (Symbols: measured data, Lines: BSIM-CMG model). The model describes short channel bulk FinFET very well.
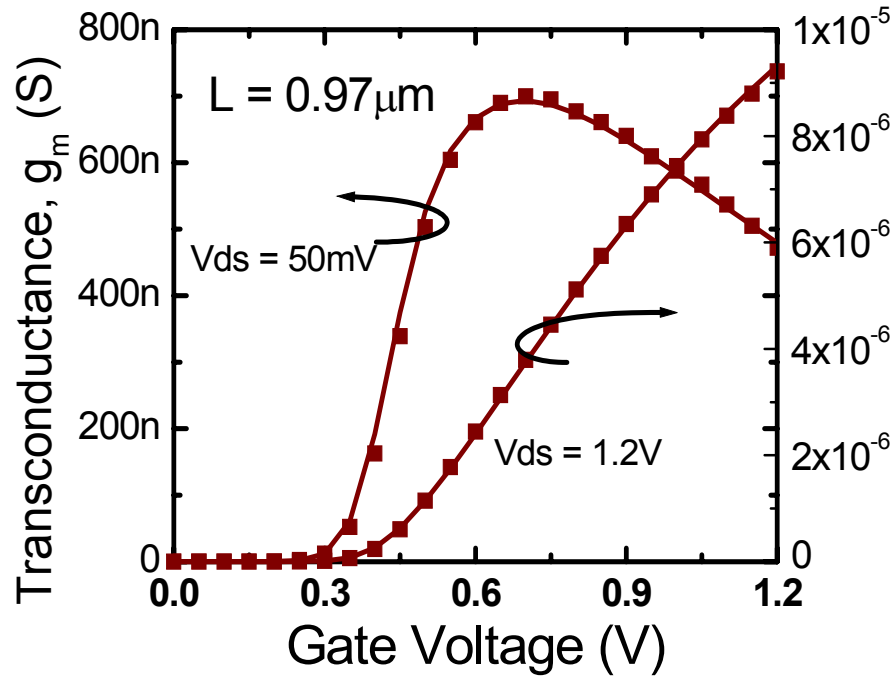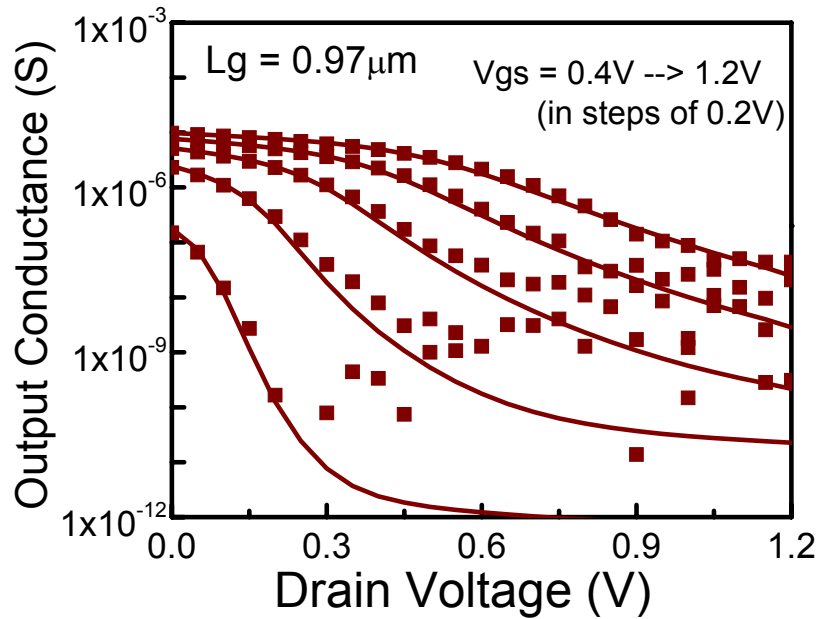
Fig. 3.15 BSIM-CMG model fitting to analog design metrics (a) transconductance g_m and (b) output conductance g_{ds} for a long channel bulk FinFET (L_g =0.97μm). Symbols represent the measured data and lines indicate model fitting results
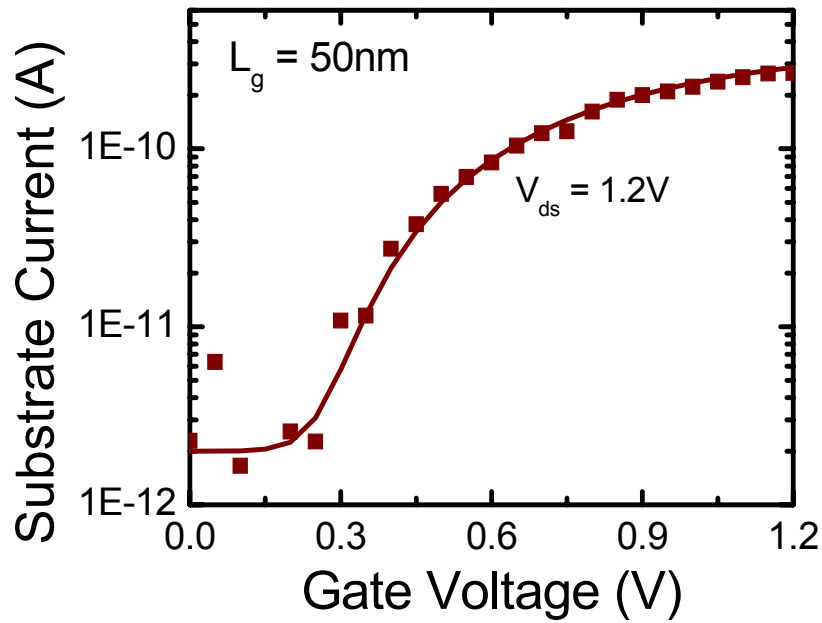
Fig. 3.16 The substrate current model agrees with the measured date well for short channel bulk FinFET.

3.7 Summary

A full-scale compact model for common-gate symmetric multi-gate FETs is developed. The surface potential model together with the I-V and C-V model for DG-FET form the core model for BSIM-CMG. The I-V model exhibits excellent accuracy over a wide range of body doping. The core model is highly predictive and has a high degree of accuracy. The model agrees with TCAD simulations without the use of

any fitting parameters. It demonstrates the inherent physical predictivity and scalability of the model. BSIM-CMG model is experimentally verified against both SOI FinFET and bulk FinFET technologies for both long and short channel transistors. The model describes both analog and digital design metrics very well making it suitable for mixed-signal design applications.

3.8 References

[3.1] Y. Taur, D. A. Buchanan, Wei Chen, D. J. Frank, K. E. Ismail, Shih-Hsien Lo, G. A. Sai-Halasz, R. G. Viswanathan, H.-J. C. Wann, S. J. Wind and Hon-Sum Wong, "CMOS Scaling into the Nanometer Regime," Proceedings of IEEE, vol. 85, pp. 486-504, April 1997.

[3.2] D.J. Frank, R. H. Dennard, E. Nowak, P. M. Solomon, Y. Taur and H.-S.P. Wong, "Device Scaling Limits of Si MOSFETs and Their Application Dependencies," Proceedings of IEEE, vol. 89, pp. 259-288, March 2001.

[3.3] D. Hisamoto, Wen-Chin Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, Tsu-Jae King, J. Bokor and Chenming Hu, "FinFET – A Self-Aligned Double Gate MOSFET Scalable to 20nm," IEEE Trans. Electron Devices, vol. 47, pp. 2320-2325, Dec. 2000.

[3.4] H.-S. P. Wong, D. J. Frank and P. M. Solomon, "Device Design Considerations for Double-Gate, Ground Plane and Single-Gate Ultra-Thin SOI MOSFETs at the 25nm Channel Length Generation," IEDM Tech. Dig., pp. 407-410, 1998.

[3.5] Y. Taur, "Analytic Solutions of Charge and Capacitance in Symmetric and Asymmetric Double-Gate MOSFETs," IEEE Trans. Electron Devices, vol. 48, pp. 2861-2869, Dec. 2001.

[3.6] G. D. J. Smit, A. J. Scholten, N. Serra, R. M. T. Pijper, R. van Langevelde, A. Mercha, G. Gildenblat and D. B. M. Klassen, "PSP-based Compact FinFET Model Describing DC and RF Measurements," IEDM Tech. Dig., pp. 175-178, 2006.

[3.7] J.-M. Sallese, F. Krummenacher, F. Pregaldiny, C. Lallement, A. Roy and C. Enz, "A Design Oriented Charge Based Current Model for Symmetric DG MOSFET and its Correlation with the EKV Formalism," Solid-State Electronics, vol. 49, pp. 485-489, 2005.

[3.8] Mohan V. Dunga, Chung-Hsun Lin, Xumei. Xi, Darsen D. Lu, Ali M. Niknejad and Chenming Hu, "Modeling Advanced FET Technology in a Compact Model," IEEE Trans. Electron Devices, vol. 53, pp. 1971-1978, Sept. 2006.

[3.9] M. V. Dunga, C.-H. Lin, D. D. Lu, W. Xiong, C. R. Cleavelin, P. Patruno, J.-R. Hwang, F.-L. Yang, A. M. Niknejad and C. Hu, "BSIM-MG: A Versatile Multi-Gate FET Model for Mixed-Signal Design," Proceedings of the VLSI Technology Symposium, pp.. 80-81, 2007.

[3.10] Darsen D. Lu, Mohan V. Dunga, Chung-Hsun Lin, Ali M. Niknejad and Chenming Hu, "A Multi-Gate MOSFET Compact Model Featuring Independent-Gate Operation," to be published in IEDM 2007.

[3.11] J. R. Brews, "A Charge Sheet Model of the MOSFET," Solid-State Electronics, vol. 21, pp. 345-355, 1978.

[3.12] F. Balestra, S. Cristoloveanu, M. Benachir, J. Brini, and T. Elewa, "Double-gate Silicon-on-Insulator Transistor with Volume Inversion: A New Device with Greatly Enhanced Performance," IEEE Electron Device Lett., vol. 8, p. 410, 1987.

[3.13] D. Ward and R. Dutton, "A Charge-Oriented Model for MOS Transistor Capacitances," IEEE J. Solid State Circuits, vol. SSC-13, pp. 703-708, Oct. 1978.

[3.14] W. Xiong, K. Shin, C. R. Cleavelin, T. Schulz, K. Schruefer, I. Cayrefourcg, M. Kennard, C. Mazure, P. Patruno, and T.-J. K. Liu, "FinFET Performance Enhancement with Tensile Metal Gates and Strained Silicon on Insulator (sSOI) Substrate," in Proc. of the IEEE 2006 Device Research Conference, Jun. 2006, pp. 39–40.

Chapter 4

Independent Multi-Gate

MOSFETs - Model and Circuit

Implications

4.1 Introduction

The multi-gate MOSFET is a promising device for extending CMOS scaling due to

its ability to suppress short-channel effects and reduce device variations.   One category

is the independent multi-gate FET (IMG-FET) in which two independently-biased gates

are incorporated in FinFET [4.1] or SOI MOSFET [4.2] (Fig. 4.1).   IMG-FET allows

more flexible circuit design than single-gate or common-gate MOSFETs. For example,

in independent-gate SRAM cells, back-gate dynamic feedback resolves the trade-off

between read/write margins [4.3]. In nano-scale CMOS processes where V_{th} variation

between circuit blocks is large, back-gate biasing of IMG-FET can be exploited for

tuning out delay variation. In analog circuits, it can also be used as a single transistor

mixer where a large LO signal and a small RF signal can be applied on different gates

[4.4]. To facilitate circuit design using IMG-FETs, section 4.2 presents a full-fledged,

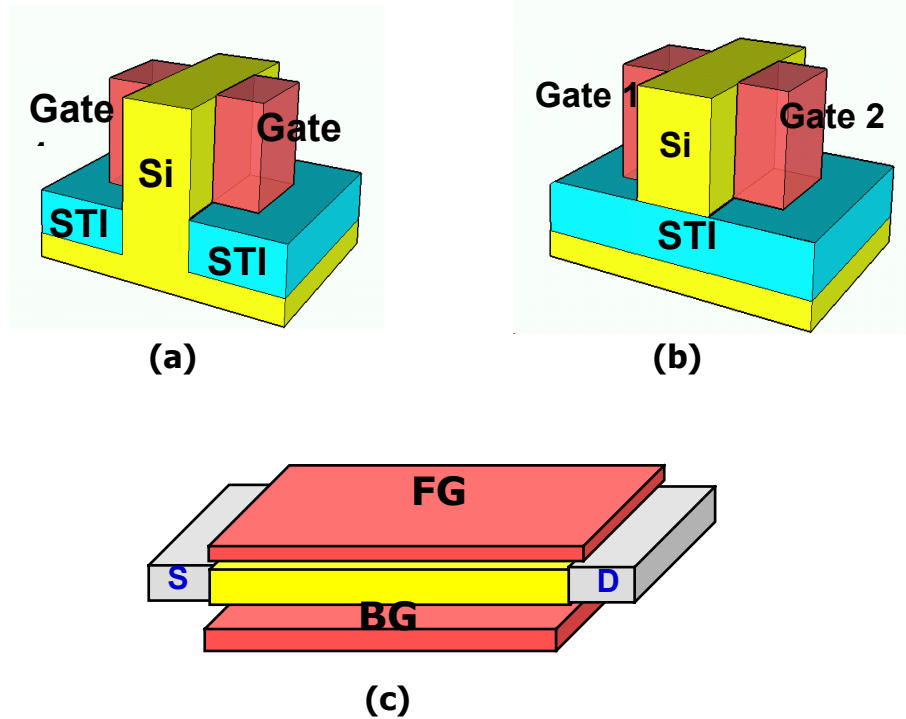accurate, and efficient compact model -- BSIM-IMG.

Fig. 4.1 (a) Bulk FinFET with two independent gates. (b) SOI FinFET with independent gates. (c) Back-gated SOI MOSFET.

The BSIM-IMG is surface potential based model. The surface potential and charge is obtained from an efficient analytical approximation. The core I-V equation is derived from drift-diffusion formulation. The model is verified using 2-D TCAD simulation. Numerous real device effects are added in the core model.

It is important to understand and model the effects of back-gate control in IMG-FETs. The independent multi-gate MOSFET (IMG-FET) [4.1] has various advantages over conventional bulk MOSFETs: (1) with a fully-depleted body, short channel effects (SCE) are effectively suppressed. (2) Dynamic threshold voltage (V_{th}) control allows the reduction of standby power in unused circuit blocks. (3) Device-to-device variation can be tuned out by biasing the back gates at different voltages. (4) V_{th} can be set by biasing the back-gate, allowing the use of an undoped body to minimize random dopant fluctuation. Furthermore, since the back-gate is not defined by a self-aligned process, misalignment will be an issue. In section 4.3, we study these effects through TCAD simulations. An analytical model for back-gate control is developed and implemented in BSIM-IMG.

4.2 BSIM-IMG: A Compact Model for Independent Multi-Gate MOSFETs

The surface-potential-based core model is derived from a lightly-doped double-gate FET structure, as shown in Fig. 4.2. An implicit analytical expression for the front and back surface potentials at the source (ψ_{sf}, ψ_{sb}) and drain (ψ_{df}, ψ_{db}) is known [4.5]. However, since an iterative algorithm is needed to compute the solution of such implicit expression, implementation into a compact model is difficult. Therefore, a robust and efficient analytical approximation for surface potential has been derived.
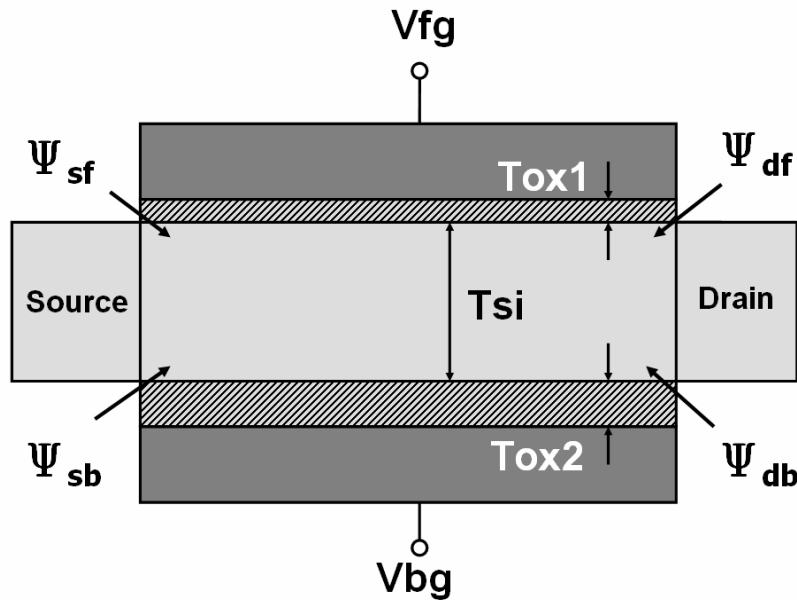

Fig. 4.2 Structure under study and the definition of symbols in the formulation.

Based on the assumption that the inversion charge at the back surface of the IMG-FET is negligible, ψ_{sf}, ψ_{df}, and the total charge densities across the body ($q_{s,tot}$, $q_{d,tot}$) are computed without the use of any iterative algorithms.

The I-V model is derived based on the charge sheet approximation [4.6] at both the front and back surfaces of the IMG-FET,

$$I_{df} = \mu \frac{W}{L} \left(\frac{q_{sf} + q_{df}}{2} + C_{ox1} \frac{kT}{q} \right) \left(\psi_{sf} - \psi_{df} \right)$$

$$I_{db} = \mu \frac{W}{L} \left(\frac{q_{sb} + q_{db}}{2} + C_{ox2} \frac{kT}{q} \right) \left(\psi_{sb} - \psi_{db} \right) \qquad (4.1)$$

$$I_d = I_{df} + I_{db}$$

where q_{sf}, q_{df} are the inversion charge per area in the front half of the body at the source and drain ends, respectively; q_{sb}, q_{db} are those in the back half of the body.

Eq. (4.1) is valid when both the front and back surfaces of the IMG-FET are biased in strong inversion. However, since the analytical approximation of surface potential assumes negligible inversion charge at the back surface, a simplified I-V is derived assuming significant current conduction at the front surface:

$$I_{ds} = \mu \frac{W}{L} \left(\frac{q_{s,tot} + q_{d,tot}}{2} \left(\psi_{df} - \psi_{sf} \right) + \frac{kT}{q} \left(q_{s,tot} - q_{d,tot} \right) \right) \qquad (4.2)$$

where $q_{s,tot}$ and $q_{d,tot}$ denoted the total inversion charge in the body at the source and drain end, respectively. C-V model for BSIM-IMG is derived using the Ward-Dutton charge partition approach and the current continuity relation.

Fig. 4.3-4.6 verifies the core I-V and C-V model against TCAD [4.7] simulations without using any fitting parameters. Fig. 4.3 shows I_{ds}-V_{ds} of an independent-gates FinFET with identical front and back oxide thickness and gate work-function. The I-V calculated from Eq. (4.1) (based on the iterative surface potential solution), Eq. (4.2) (based on the analytical approximation), and TCAD simulations are compared. The analytical approximation introduces only a small error in the I-V. The I_{ds}-V_{fg} calculated from the model also agrees well compared to TCAD as shown in Fig. 4.4. All the remaining results in the section use the I-V model described by Eq. (4.2). In Fig. 4.5 the model accurately predicts the effects of varying T_{ox2} and V_{bg} for a device with unequal front and back gate work-functions and oxide thicknesses. The expressions for the terminal charges are continuous and valid over all regimes of transistor operation. Terminal capacitances are obtained by differentiating the terminal charges with respect to the terminal voltages. Fig. 4.6 shows the variation of capacitances associated with the front gate terminal as a function of drain voltage for an IDG-FET with asymmetric front and back gates and V_{bg} =0V. The C-V model agrees very well with TCAD simulations without using any fitting parameters. Fig. 4.6 also shows that $C_{g1,s}$ and $C_{g1,d}$ are equal at V_{ds}=0, reflecting the inherent source-drain symmetry of the model.

88
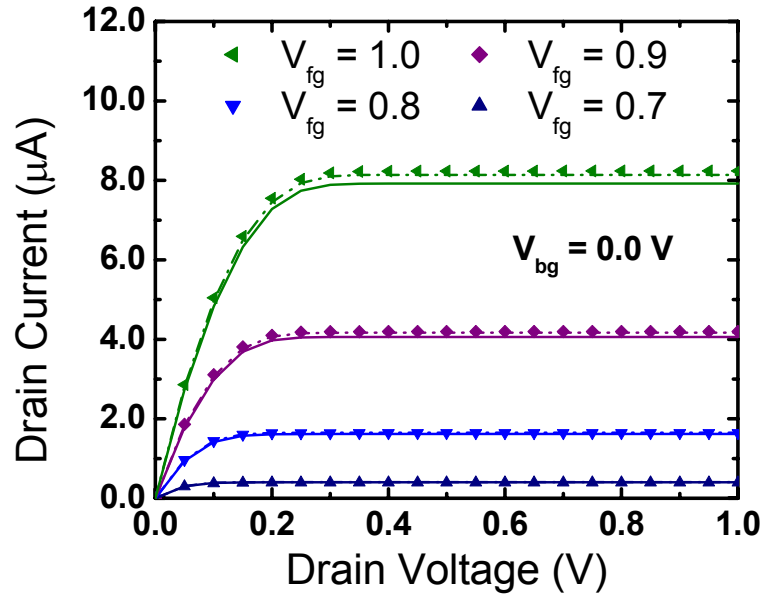
Fig. 4.3 I_{ds}-V_{ds} for varying front gate voltage (V_{fg}). V_{bg}=0, T_{si}=15nm, T_{ox1}=T_{ox2}=2nm. Symbols: TCAD; dashed lines: I-V calculated using Eq. (4.1); solid line: I-V calculated using Eq. (4.2).
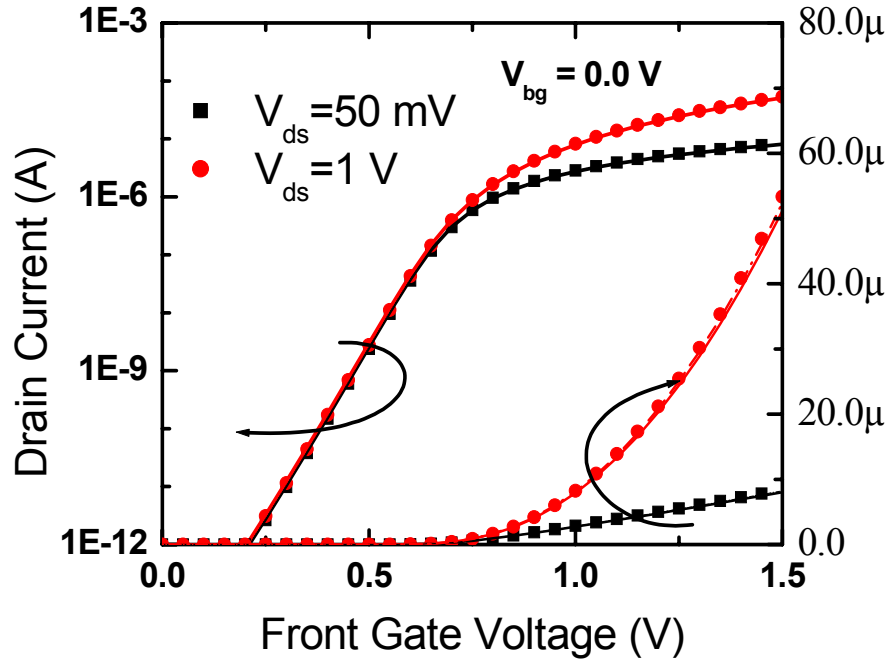
Fig. 4.4 I_{ds}-V_{fg} in both the linear and saturation regions. V_{bg}=0, T_{si}=15nm, T_{ox1}=T_{ox2}=2nm. Symbols: TCAD; dashed lines: I-V calculated using Eq. (4.1); solid line: I-V calculated using Eq. (4.2).
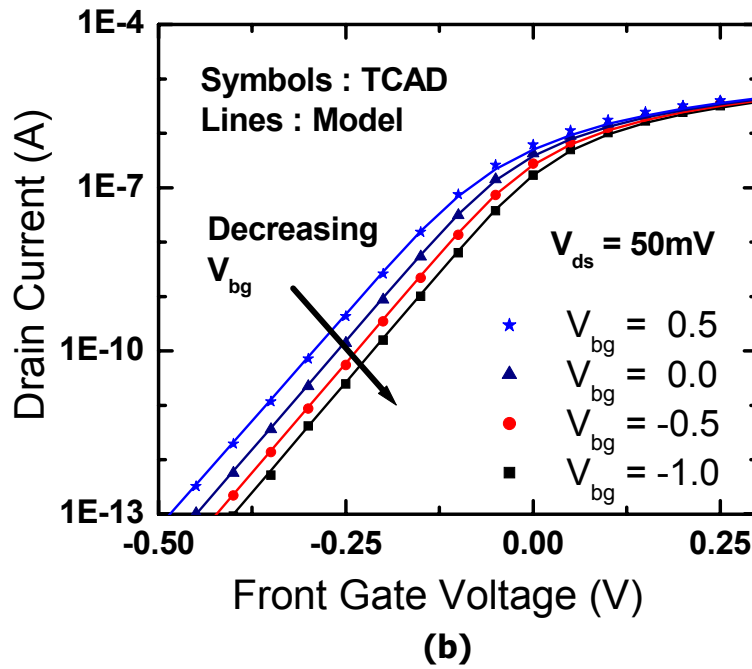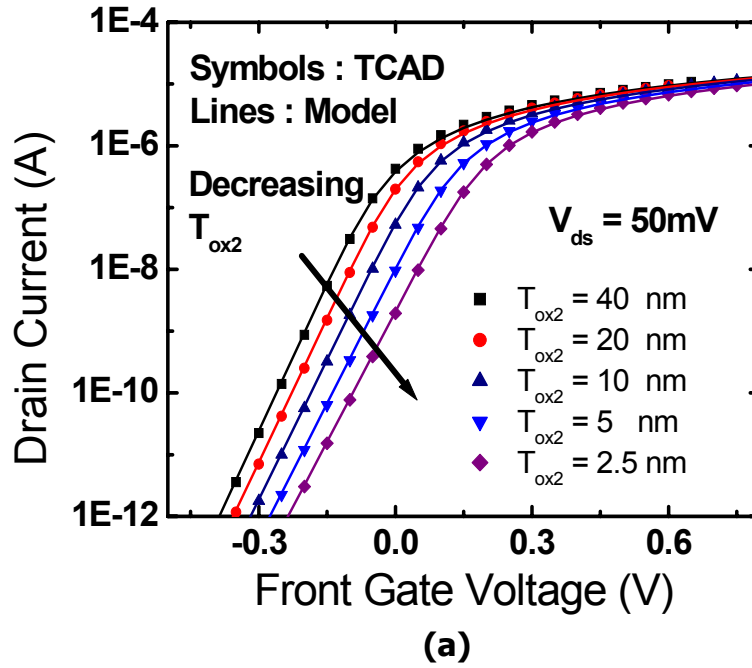
Fig. 4.5 Model predicts the effect of changing (a) back oxide (T_{ox2}), and (b) back gate voltage (V_{bg}) accurately for a device with unequal front and back gate workfunctions and oxide thicknesses.   (T_{ox1} = 1.2nm, T_{ox2} = 40nm, T_{si} = 15nm, V_{bg} = 0 unless specified)
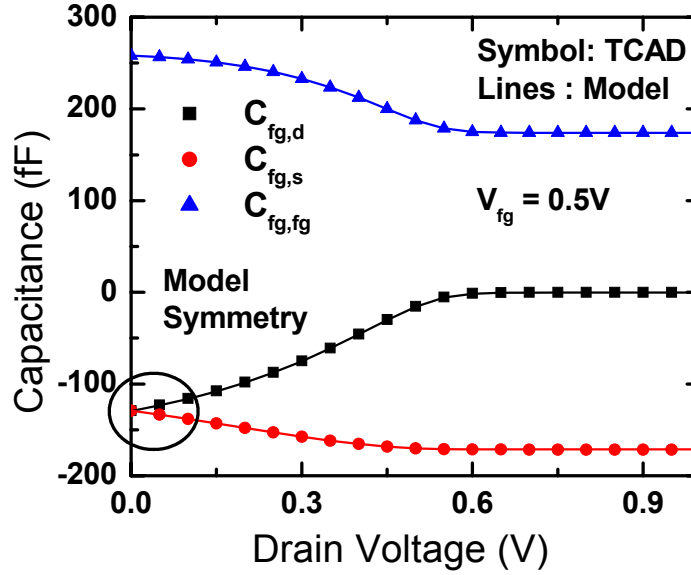
Fig. 4.6 The C-V model agrees with TCAD simulation well from the linear to saturation region without any fitting parameters. $C_{fg,s}$ and $C_{fg,d}$ are equal at $V_{ds}=0$, reflecting the model's symmetry at $V_{ds}=0$. (T_{ox1} = 1.2nm, T_{ox2} = 40nm, T_{si} = 15nm, V_{bg} = 0, N+-P+ Gates)

The model includes important real device effects. Threshold voltage roll-off and DIBL are captured very well, as shown in Fig. 4.7. Sub-threshold slope (SS) degradation at short channel lengths is modeled via a capacitor divider concept,

$$SS = \left(1 + \frac{C_{dsc} + C_{it} + C_{si} \| C_{ox2}}{C_{ox1}}\right) \times \frac{kT}{q} \ln 10 \qquad (4.3)$$

where C_{it} is the effective capacitance due to interface trap charge; C_{dsc} accounts for capacitive coupling from the drain side; $C_{si} = \varepsilon_{si} / T_{si}$; $C_{ox1} = \varepsilon_{ox} / T_{ox1}$; $C_{ox2} = \varepsilon_{ox} / T_{ox2}$.

Mobility degradation due to the vertical electric field is accounted for by expressing the mobility as a function of the average vertical electric field (E_{avg}). E_{avg} in an

IMG-FET is given by

$$E_{avg} = \frac{Q_b + 0.5Q_i}{\varepsilon_s} + E_{bg} \qquad (4.4)$$

where Q_b is the bulk charge per unit area in a fully-depleted IMG-FET; Q_i is the average

inversion charge per unit area; E_{bg} is the electric field at the back surface.

Since body doping may be necessary to provide multiple V_{th} for low power circuit

design, the effect of finite body doping on V_{th} is also modeled. Other physical

phenomena such as quantum mechanical confinement effects, impact ionization, gate

tunneling current, velocity overshoot, and source velocity limit are also accounted for in

BSIM-IMG.
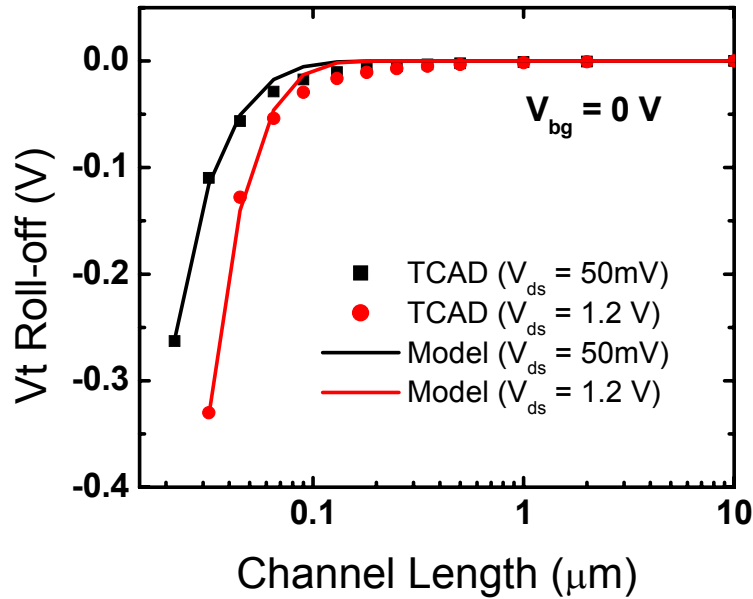
Fig. 4.7 The Vt roll-off Model fits TCAD results well for the back-gated SOI MOSFET. (T_{ox1} = 1.2nm, T_{ox2} = 20nm, T_{si} = 10nm, N+-P+ Gates)

The Gummel symmetry test is applied to the model to verify its symmetry with respect to $V_{ds}=0$. In Fig. 4.8, the lines and symbols (flipped version of the lines) overlay each other, showing perfect symmetry. Symmetry is maintained in the model even after the incorporation of all the real device effects.

Table 4.1 enlists the real device effects modeled in BSIMIMG. The model has been written in Verilog-A and implemented in popular circuit simulators such as SPECTRE and HSPICE.
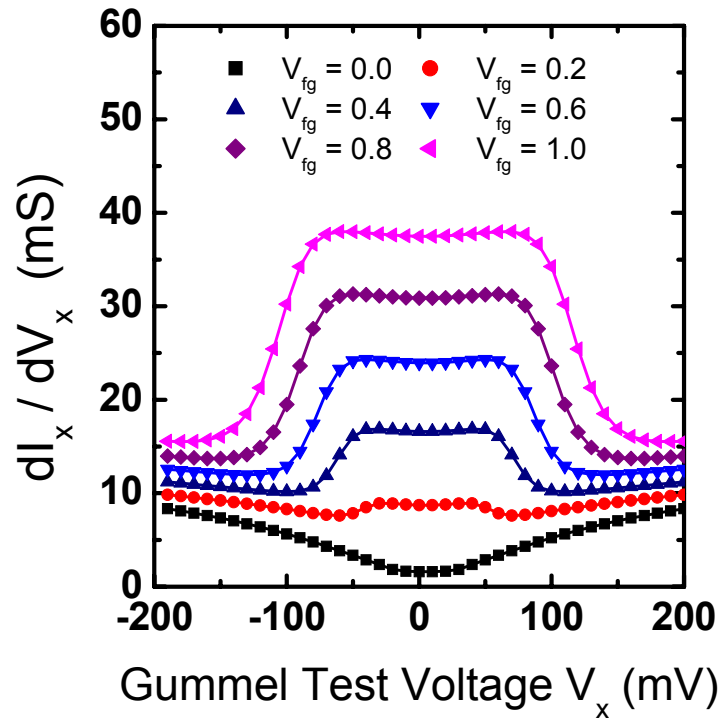
Fig. 4.8 First order Gummel symmetry tests illustrates model symmetry with respect to $V_{ds}=0$. The symbols are flipped versions of the lines, showing dI_x/dV_x is indeed an even function of V_x.

Table 4.1 List of Physical effects modeled in BSIM-IMG

1	Quantum Mechanical Effects
2	Short Channel Effects
3	Vth roll-off
4	Drain Induced Barrier Lowering
5	Subthreshold Slope Degradation
6	Channel Length Modulation
7	S/D Series Resistance
8	Mobility Degradation
9	Velocity Saturation
10	Velocity Overshoot
11	Gate Induced Drain Leakage
12	Gate Tunneling
13	S/D Junction Leakage
14	Impact Ionization
15	Parasitic Capacitance

BSIM-IMG is used to simulate SRAM cells with back-gate dynamic feedback. Fig. 4.9 (a) shows the circuit schematic. The two access transistors are independent-gates FinFETs and are modeled with BSIM-IMG. The four transistors in the inverter pair are common-gate FinFETs and are modeled with BSIM-CMG [4.8]. For the control case, all six transistors are common-gate FinFETs modeled with BSIM-CMG. Fig. 4.9 (b) shows the simulated butterfly curves for the 6-T SRAM cell. The significantly-improved read margin due to back-gate feedback is well-predicted by the model. This result is consistent with mixed-mode TCAD simulation performed in [4.3].

(a)

(b)

Fig. 4.9 Read margin simulation for a FinFET SRAM cell with and without back-gate dynamic feedback. (a) Circuit schematic. (b) Butterfly plots and read margin extraction results. BSIM-IMG model is used for the independent-gates FinFETs and BSIM-CMG [4.7] is used for the common-gate FinFETs.

The model is also used to explore the tuning of device variations through the back-gate bias in IMG-FETs. Under body thickness (T_{si}) variation, we simulate I_{on}, I_{off} of an n-type IMG-FET and the delay per stage in an IMG-FET-based 17-stage ring oscillator (Fig. 4.10) (T_{si} for n-type and p-type devices vary at the same time). With increasing T_{si}, V_{th} increases due to the finite body doping. Therefore, the ring oscillator delay increases (Fig. 4.10(a)), and I_{on} and I_{off} decrease (Fig. 4.10(b)). Delay variation can be tuned out with proper choice of the back-gate voltage (V_{bg}) for both n-type and p-type devices (4.10(a)). The same V_{bg} can also tune out most of the I_{on} (4.10(b)) variation. However the trend of I_{off} is reversed since the sub-threshold slope for thick-body devices is slightly larger and I_{off} is larger
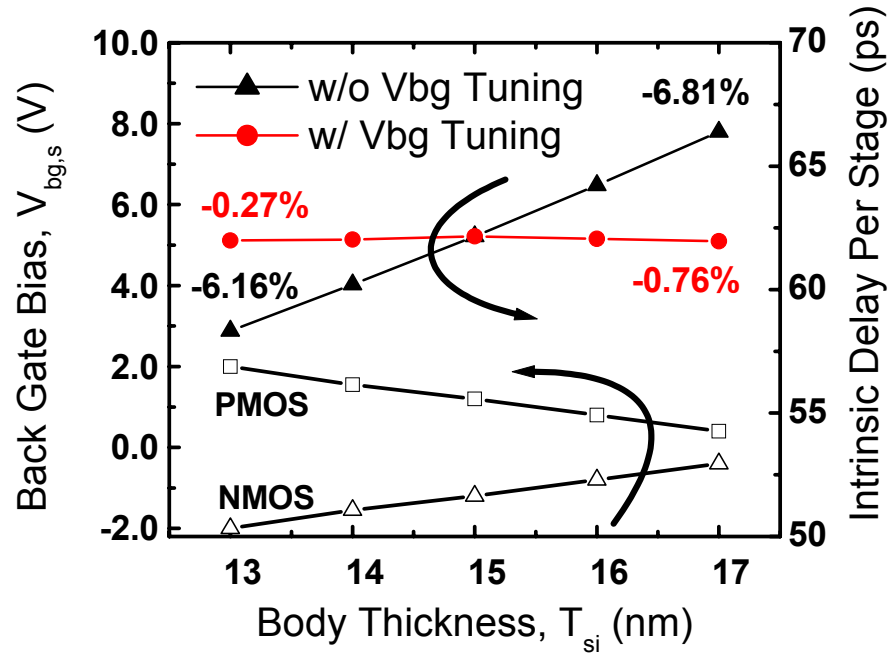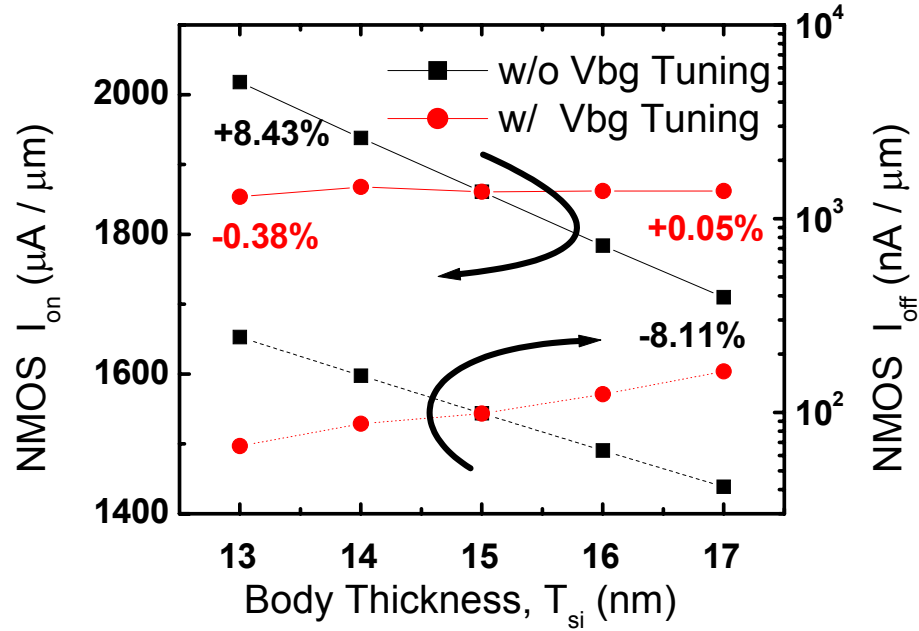
Fig. 4.10 Device variation tuning using back-gated SOI MOSFET. (a) Gate delay variation due to T_{si} variation can be tuned out with V_{bg}. (b) The same V_{bg} tunes out I_{on} nearly completely but reverses the trend of I_{off}. Gate delays are extracted from a 17 stage ring oscillator simulation.

4.3 Back-Gate Control and Misalignment Modeling

The independent multi-gate MOSFET (IMG-FET) [4.2] has various advantages over conventional bulk MOSFETs: (1) With a fully-depleted body, short channel effects (SCE) are effectively suppressed. (2) Dynamic threshold voltage (V_{th}) control allows the reduction of standby power in unused circuit blocks. (3) Device-to-device variation can be tuned out by biasing the back gates at different voltages. (4) V_{th} can be set by biasing the back-gate, allowing the use of an undoped body to minimize random dopant fluctuation. It is therefore important to understand and model the effects of back-gate control in IMG-FETs. Furthermore, since the back-gate is not defined by a self-aligned process, misalignment will be an issue.

In order to verify the model, TCAD simulations are performed on a planar double-gate SOI n-MOSFET (Fig. 4.11). Physical parameters of this structure are listed in Table 4.2. Real device effects such as mobility degradation, quantum mechanical effects, direct electron tunneling through the gate, and band-to-band tunneling are considered. Fig 4.12 and 4.13 shows the drain (I_D) and gate (I_{FG}) current for a 1µm device versus front gate bias ($V_{FG,S}$). In Fig. 4.12, we simulate I_D-$V_{FG,S}$ with and without quantum mechanical effects by changing the simulation options for TCAD simulation.
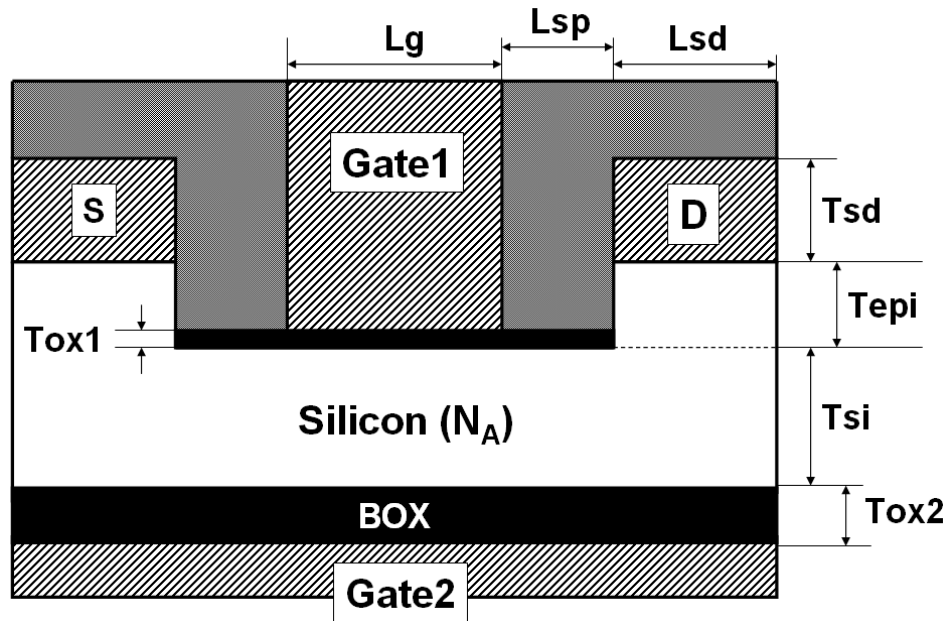
Fig. 4.11 The planar independent double-gate SOI MOSFET structure for this study.

Table 4.2 Device Parameters for TCAD simulation. Φ_{FG} and Φ_{BG} are the front- and back-gate workfunctions. σ_{SD} denotes the Gaussian spread of the source/drain doping [4.9].

Lg	13 nm	EOT1	0.78 nm
Lsp	12 nm	Tox2 = EOT2	4 nm
Lsd	40 nm	Φ_{FG}	4.23 V
Tsd	10 nm	Φ_{BG}	5.17 V
Tepi	5 nm	N_A	10^{16} cm^{-3}
Tsi	8 nm	N_{SD}	10^{20} cm^{-3}
Tox1	1.5 nm	σ_{SD}	4 nm

Both the threshold voltage shift and effective degradation of oxide capacitance due to quantum mechanical effects are modeled very well. Leakage current when $V_{FG,S} < 0$ is dominated by gate direct tunneling current. This current component is also observed in the I_{FG}-$V_{FG,S}$ plot (Fig. 4.13). When $V_{FG,S} < 0$, the gate current is dominated by electron tunneling from the gate to the conduction band of the channel. Since the gate-to-channel voltage ($|V_{GC}|$) near the drain end increases with increasing V_{DS}, this current also increases with V_{DS}.
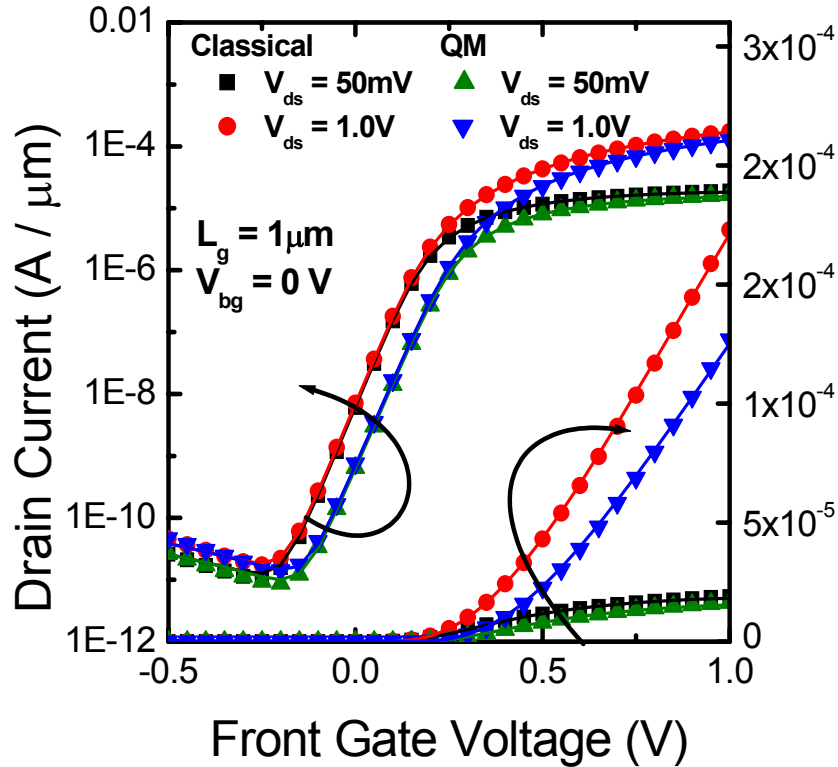
Fig. 4.12 I_D-$V_{FG,S}$ with and without considering quantum mechanical effects. Symbols: TCAD; Lines: Model.
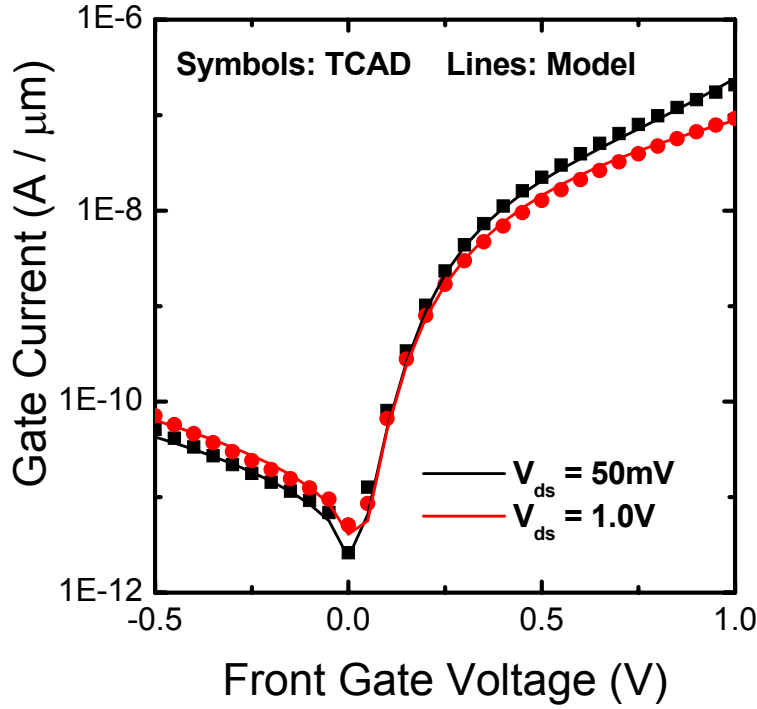
Fig. 4.13 I_{FG}-$V_{FG,S}$ for low (50mV) and high (1.0V) drain biases. Symbols: TCAD; Lines: Model

On the other hand, when $V_{FG,S} > 0$, the gate current mainly consists of electrons tunneling current from the inverted channel into the gate. At high V_{DS}, this current component is smaller due the the smaller amount of inversion charge. Source/drain series resistance and intercept length are extracted from the drain-source conductance close to V_{DS}=0 for devices with different front gate lengths (L_{FG}) and front gate overdrive ($V_{FG,S} - V_{th}$), as shown in Fig. 4.14. An under-lap of 9.8nm is found for this structure. V_{th} roll-off and drain-induced barrier lowering (DIBL) are well-captured by the model, as shown in Fig. 4.15
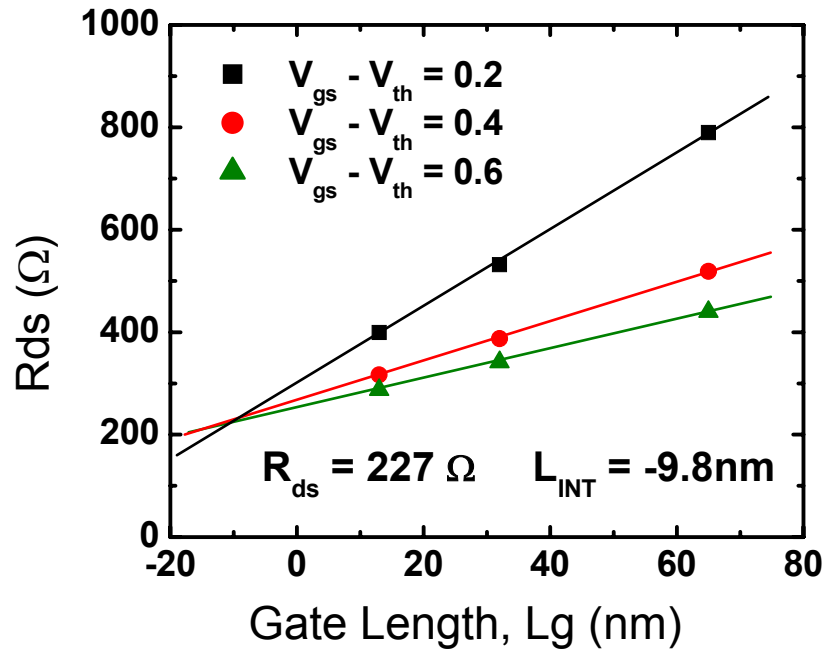
Fig. 4.14 Extraction of Rds and intercept length (L_{INT}).
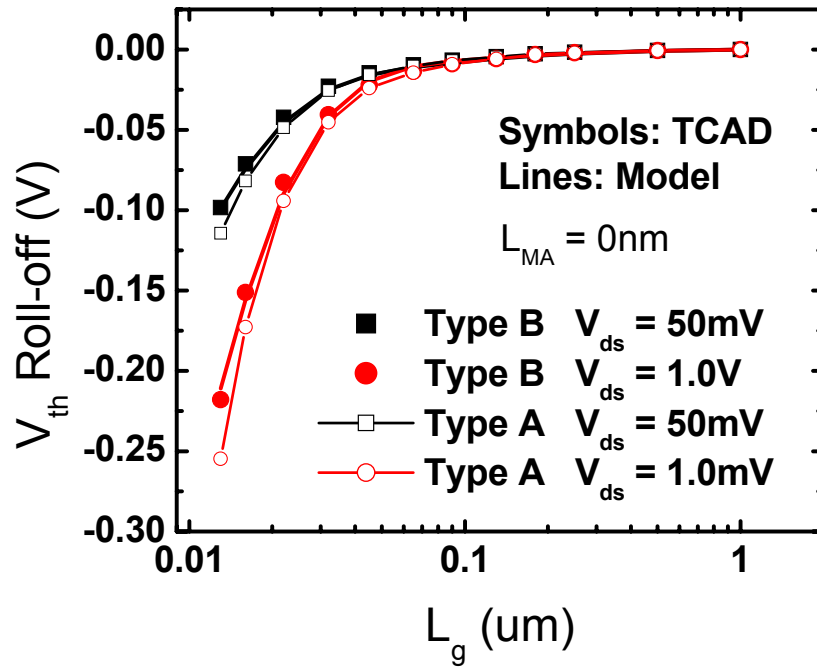
Fig. 4.15 V_{th} roll-off and DIBL for two types of structures.

In a planar SOI-type IMG-FET, the back gate may not be self-aligned to the front gate. Therefore, we perform TCAD simulations to study the effect of varying back-gate length (L_{BG}) and back-gate misalignment (L_{MA}). In Fig. 4.15, we compared V_{th} roll-off for two different types of back gate structure. In type A structure (Fig. 4.16(a)), the back gate is aligned to the front gate. In type B structure (Fig. 4.16(b)), the back gate covers the entire back plane. For type A structure SCE is less severe and overlap capacitance is reduced. However, back-gate misalignment tolerance is less, as we will see in Fig. 4.19.
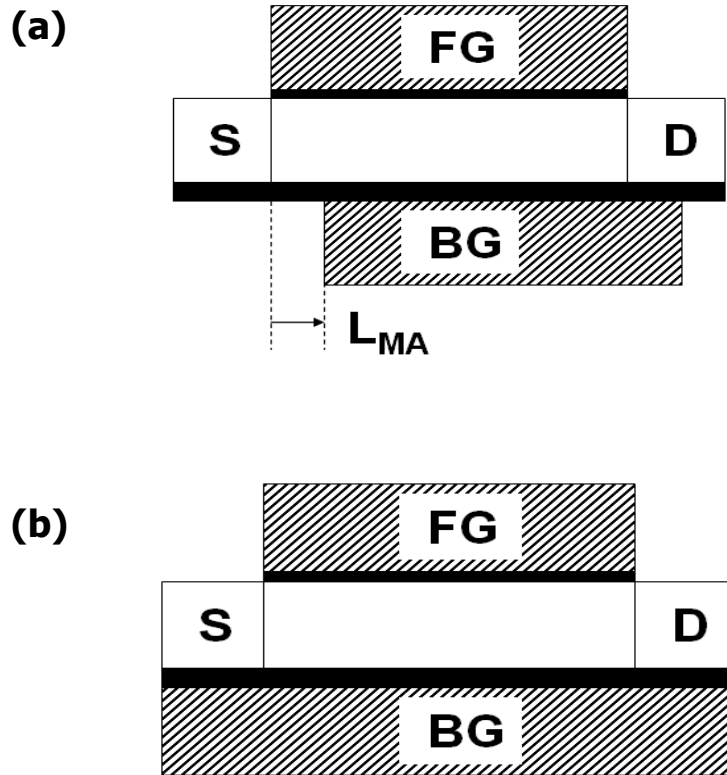
Fig. 4.16 (a) Type A (back-gate aligned) structure with misalignment. (b) Type B (back-gate fully-overlaped) structure.

Fig. 4.17 shows that V_{th} is approximately linearly related to $V_{BG,S}$ when $V_{BG,S}$ is close to

0V. Once $V_{BG,S}$ is less than a critical value, the back surface potential will be pinned and

the back gate will have almost no effect on V_{th} [4.10]. To quantify the amount of

back-gate control, we define the sensitivity of V_{th} on $V_{BG,S}$ as $\gamma = |\, dV_{th}\, /\, dV_{BG,S}\, |$. In this

study, γ is extracted at $V_{BG,S}=0$. The relation of γ to the L_{FG} is shown in Fig. 4.18. As

L_{FG} decreases, the coupling from the drain side degrades γ. However, for very short L_{FG},

γ increases again. This can be explained as follows: back-gate biasing can suppress

SCE. As L_{FG} decreases, SCE becomes large, and the amount of V_{th} increase due to

back-gate SCE suppression becomes larger, and γ is increased.
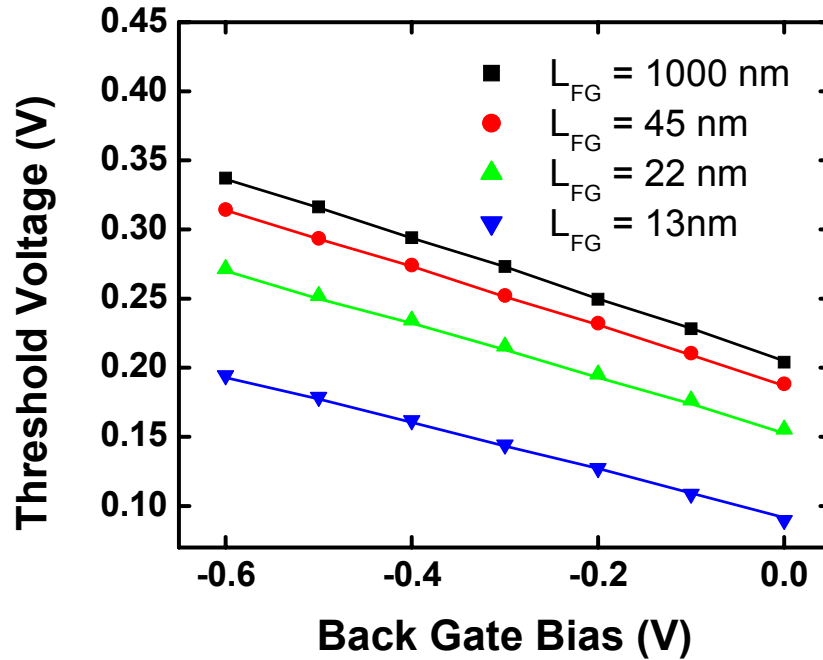
Fig. 4.17 V_{th} as a function of back-gate bias for 4 different gate lengths.
Symbols: TCAD; Lines: Model.
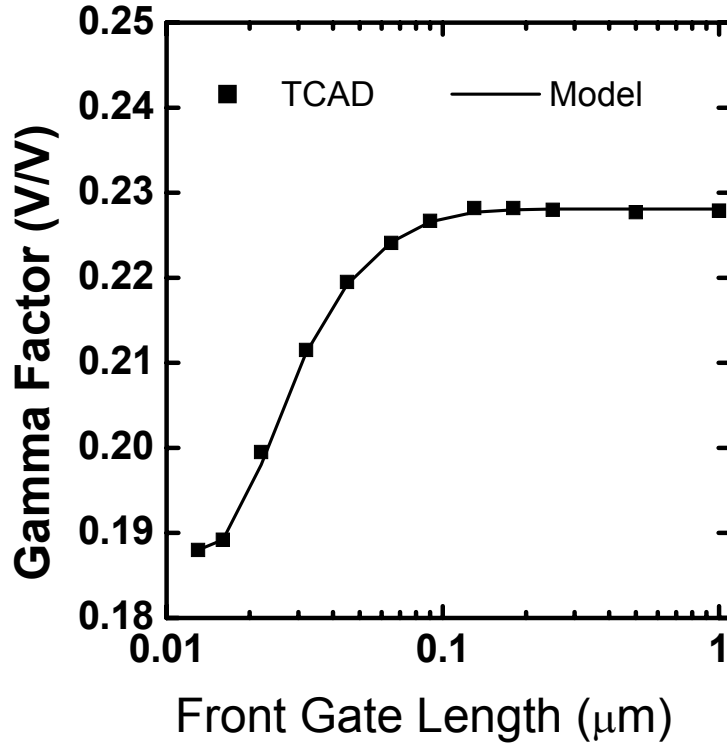
Fig. 4.18 Gamma factor for different front gate lengths.

Fig. 4.19 shows the effect of back-gate misalignment on γ factor. As expected, γ factor is larger for larger L_{BG}, since a larger back-gate area enhances its control over V_{th}. In the linear region (Fig. 4.19(a)), when the back-gate moves away from its aligned position, γ is reduced. However, in the saturation region (Fig. 4.19(b)), since the V_{th} is determined by the electrostatics close to the source end, γ remains larger as long as the back-gate covers the source-channel junction, but decreases if the back-gate is moved towards the drain end.
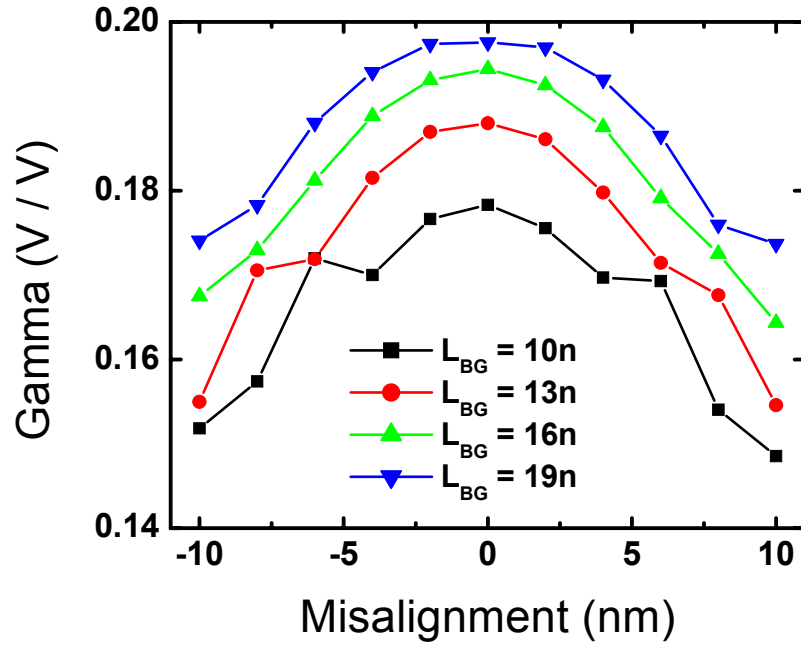
(a)
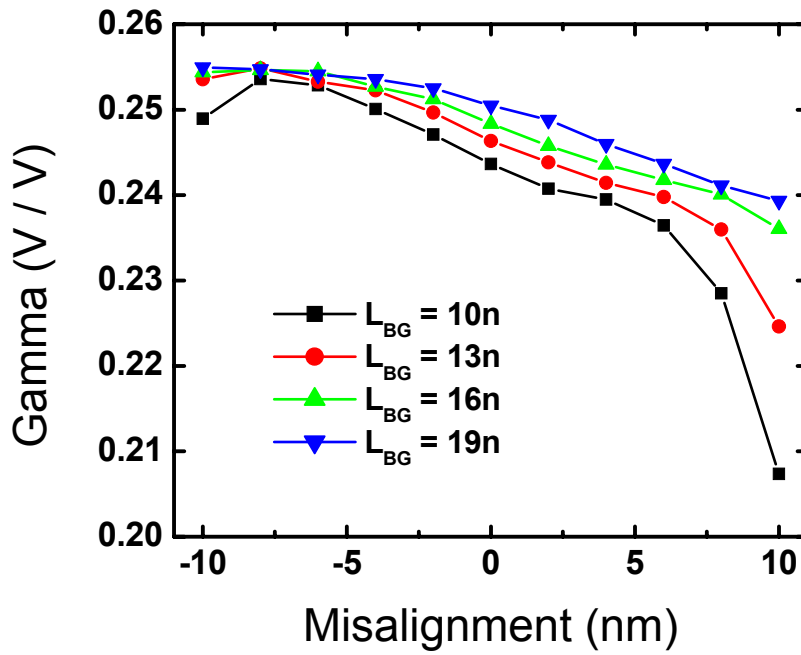
(b)

Fig. 4.19 Gamma factor for different back-gate lengths and misalignments.
(a) V_{DS}=50mV   (b) V_{DS}=1.0V

106

4.4 Summary

We have developed a compact model for independent multi-gate MOSFETs – BSIM-IMG. BSIM-IMG is a surface-potential-based compact model developed for asymmetric multi-gate MOSFETs in the independent-gates operation mode. The surface potential and charge is obtained from an efficient analytical approximation. The core I-V equation is derived from drift-diffusion formulation. Through TCAD simulations, we verified that the model predicts both the I-V and C-V accurately. Numerous real device effects are incorporated in BSIM-IMG. Source/drain symmetry of the model is verified through the Gummel symmetry test. The model is used to study FinFET based SRAM cells and device variation tuning using back gate bias, highlighting its use for both circuit and technology development. The model is under study to allow conduction at both front and back surfaces.

A planar double-gate SOI nMOSFET is simulated with TCAD and fitted using multi-gate compact model, BSIM-IMG. The impact of back-gate length and misalignment on V_{th} is discussed. Misalignment effect is more serious when the back-gate is misplaced towards the drain end.

4.5 References

[4.1]   D. M. Fried, J. S. Duster and K. T. Kornegay, "High-Performance P-Type Independent-Gate FinFETs," IEEE EDL, vol. 25, no. 4, p.199, 2004.

[4.2]   I. Y. Yang, C. Vieri, A. Chandrakasan and D. A. Antoniadis, "Back-Gated CMOS on SOIAS For Dynamic Threshold Voltage Control," IEEE TED, vol. 44, no. 5, p. 822, 1997.

[4.3]   Z. Guo, S. Balasubramanian, R. Zlatanovici, T.-J. King and B. Nikolić, "FinFET-Based SRAM Design," Proc. ISLPED, p. 2, 2005.

[4.4]   L. Mathew, Yang Du, A. V-Y Thean, M. Sadd, A Vandooren, C. Parker, T. Stephens, R. Mora, R. Rai, M. Zavala. D. Sing, S. Kalpat, J. Hughes, R. Shimer, S. Jallepalti, G. Workman, W. Zhang, J.G. Fossum, B. E. White, B-Y. Nguyen. J. Mogab, "CMOS Vertical Multiple Independent Gate Field Effect Transistor (MIGFET)," Proc. SOI Conf., p. 187, 2004.

[4.5]   Y. Taur, "Analytic Solutions of Charge and Capacitance in Symmetric and Asymmetric Double-Gate MOSFETs," IEEE TED, vol. 48, no. 12, p. 2861, 2001.

[4.6]   J. Brews, ''A Charge-Sheet Model of the MOSFET," SSE, vol. 21, p. 345, 1978.

[4.7]   Taurus-Device v. 2003 (Synopsys, Inc.).

[4.8]   Mohan V. Dunga, Chung-Hsun Lin, Darsen D. Lu, Weize Xiong, C. R. Cleavelin, P. Patruno, J.-R. Huang, Fu-Liang Yang, Ali M. Niknejad[1] and Chenming Hu, "BSIM-MG: A Versatile Multi-Gate FET Model for Mixed-Signal Design," Symp. VLSI Tech., p. 60, 2007.

[4.9]   S. Balasubramanian, Ph.D. Dissertation, University of California at Berkeley, Fall 2006

[4.10]   C.-H. Lin, X. Xi, J. He, L. Chang, R. Q. Williams, M. B. Ketchen, W. E. Haensch, M. Dunga, S. Balasubramanian, A. M. Niknejad, M. Chan, and C. Hu, " Compact Modeling of FinFETs Featuring Independent-Gate Operation Mode," Proc. VLSI-TSA, p. 120, 2005.

Chapter 5

Circuit Design Using Multi-Gate Transistors

5.1 Introduction

   The scaling of conventional planar CMOS is expected to become increasingly difficult due to increasing gate leakage and subthreshold leakage. Multi-gate FETs such as FinFETs have emerged as the most promising candidates for extending CMOS scaling into the sub-25nm regime. The strong electrostatic control over the channel originating from the use of multiple gates reduces the coupling between source and drain in the subthreshold region and enables the multi-gate transistor to be scaled beyond bulk planar

CMOS for a given dielectric thickness. Numerous efforts are underway to enable large scale manufacturing of multi-gate FETs. The use of the lightly-doped or undoped body provides the immunity against the dopant fluctuation effect, which induce the threshold voltage variation, and higher carrier mobility. Multi-gate FETs can be built on SOI or bulk silicon. At the same time, circuit designers are beginning to design and evaluate multi-gate FET circuits. In this chapter, the circuit design issues of MG-based circuits are discussed.

Section 5.2 compares the circuit design trade off of two main types of double-gate (DG) MOSFETs: (1) the symmetric/common DG (SDG) device and (2) the asymmetric/independent DG (ADG) device. The term "common-gate" means that all the gates in the multi-gate FET (double-gate or triple-gate or quadruple-gate FinFET) are electrically interconnected and are biased at the same electrical gate voltage. The common-gate model further assumes that the gate work-functions and the dielectric thicknesses on the two, three or four active sides of the fin are the same. The asymmetric/independent gate model allows different work-functions and dielectric thicknesses on the two sides of the fin. The asymmetric/independent gate permits that the two gates can be biased independently. Although the characteristics of SDG and ADG device have been investigated by many groups, the relative circuit performance of these

two devices still remains controversial.

In the conventional common multi-gate transistors (CMG), advanced gate work function engineering is needed for the threshold voltage (V_{th}) control. Dynamic V_{th} control will inevitably be required for future low power circuit design. The independent-gate FinFET (IDG) shows flexible V_{th} control through the use of second-gate bias (the second gate of NMOS and PMOS FinFETs are switched independently). Section 5.3 describes the concept of the dynamic V_{th} control in the compact modeling of FinFET.

As we scaling the device dimensions, process-induced variations cause an increasing spread in the distribution of circuit delay and power, and affecting the robustness of VLSI designs [5]. SRAM has become the focus of technology scaling since embedded SRAM is estimated to occupy nearly 90% of the chip area in the near future [6]. Due to the area-constrained limit, the device fluctuation in the SRAM cell is significant. In section 5.4, we explore the performance of FinFET technology in digital circuit applications at 90nm technology node under various device parameter variations. Comprehensive comparison of FinFET vis-à-vis PD-SOI has been done for logic gates as well as memory structures that are most commonly used in commercial VLSI designs. We also compare the performance of the two technologies at ultra-low voltages for low-power applications.

5.2 Evaluation of Circuit Performance for Double-Gate MOSFETs

As CMOS technology is fast moving toward the scaling limit, the double-gate (DG) MOSFETs is considered the most promising structure to suppress the short channel effect for a given equivalent gate oxide thickness by using two gates to control the channel [5.1]. The use of the lightly-doped or undoped body provides the immunity against the dopant fluctuation effect, which induce the threshold voltage variation, and higher carrier mobility. There are two main types of DG MOSFETs: (1) the symmetric DG (SDG) device with both gates of identical work functions, and (2) the asymmetric DG (ADG) device with different work functions for the gates. Although the characteristics of SDG and ADG device have been investigated by many groups [5.2-5.4], the relative circuit performance of these two devices still remains controversial. In this section, the performance of DG MOSFETs from the circuit-design perspective is examined via simulation using device structures based on the ITRS specification [5.5]. The propagation delay (t_{pd}) and energy dissipation of DG CMOS inverter chains with different number of fan-outs (FO) are investigated. Load capacitors are added to the output node of each inverter to simulate the parasitic wiring capacitance (C_{Int}) between two stages (Fig. 5.1).
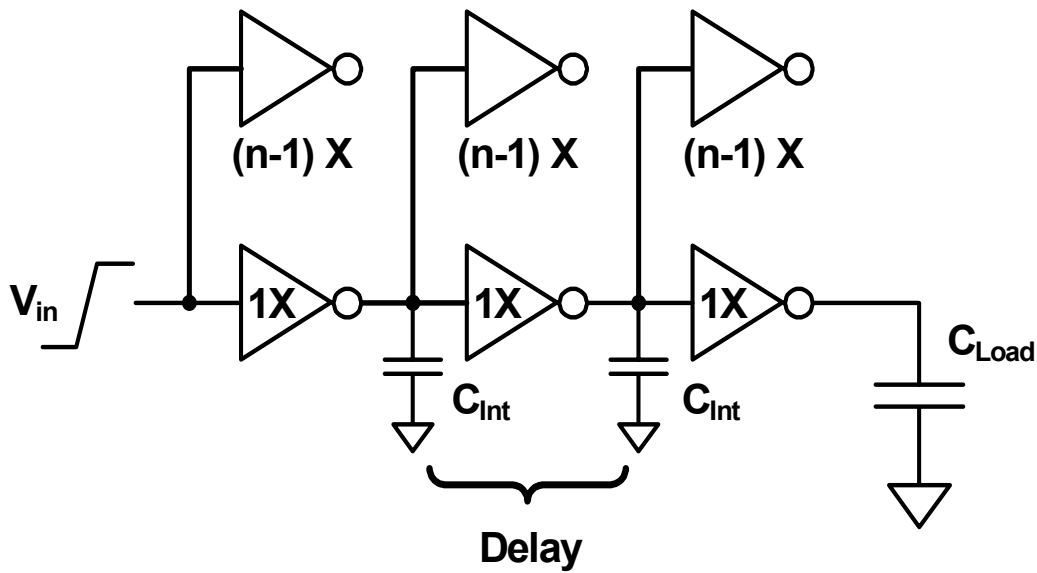
Fig. 5.1 Inverter chain circuit with different number of fan-outs (n) and wiring capacitance used to simulate the performance. Reasonable input rise time and output loading capacitance are assumed.

Fig. 5.2 shows the I_{DS}-V_{GS} characteristics of 25 nm gate length SDG and ADG nMOSFETs simulated by the ISE 2-D device simulator. The oxide thickness and silicon body thickness are 1.1 nm and 7 nm, respectively. The silicon film doping is 10^{15} cm^{-3} and the S/D doping gradient is 2.8 nm/dec. The gate height is 50 nm, and the spacer width is 20 nm. The cross-sectional schematic of the structure used in the simulation is shown in the Fig. 5.3. The gate work function of the SDG is adjusted to achieve an I_{off} equal to that of the ADG. The mobility model used in the simulation consists of impurity scattering, carrier-carrier scattering, transverse field degradation, and high field saturation.

Due to the lower transverse electric field in the silicon film, the SDG has higher mobility than the ADG. For the same threshold voltage, the inversion mobile charge (Q_i) of the SDG is higher than that of the ADG [5.2]. Combining these two factors, the SDG has larger I_{on} driving current than the ADG. Besides the SDG's shown in Fig. 1.2 may provide 2 times larger width than the ADG's shown in Fig. 4.1.
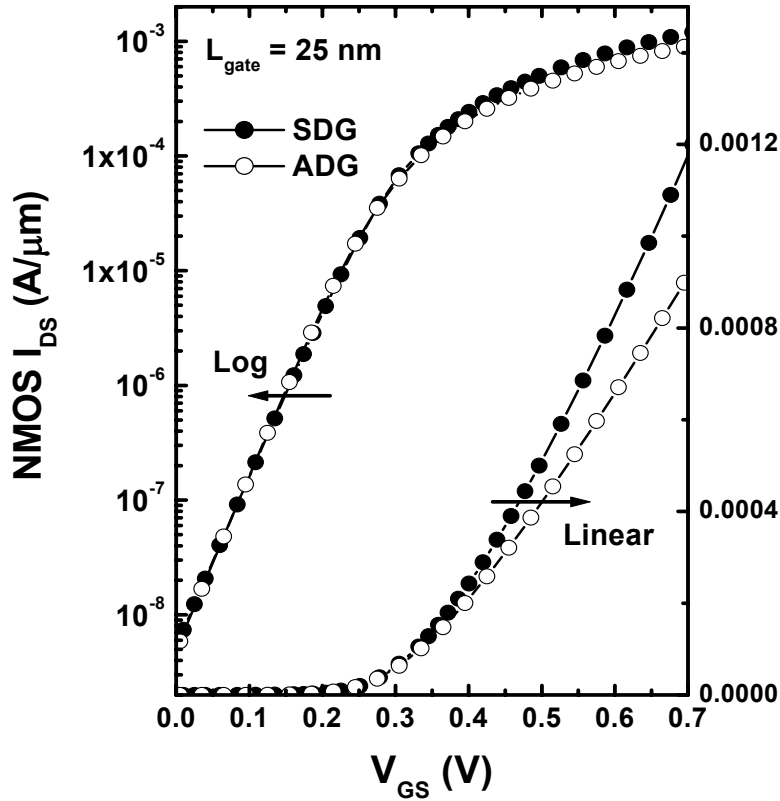

Fig. 5.2 The I_{DS}-V_{GS} characteristics (V_{DS}=0.7V) of 25nm gate length SDG and ADG NMOSFETs in both logarithmic (left), and linear (right) scales. The work function of the SDG device is adjusted to have get I_{off} equal to that of the ADG device. The SDG shows higher I_{on} than ADG due to higher mobility and inversion charge.
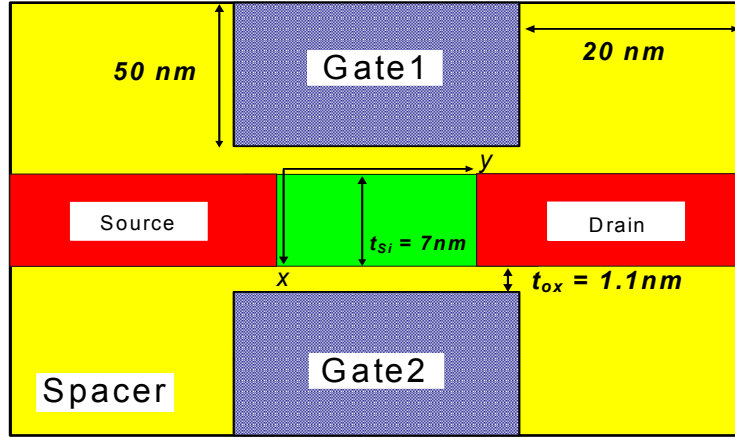
Fig. 5.3 Cross-sectional schematic of the DG MOSFET structure used in the simulation. The sidewall spacers are included to simulate the parasitic capacitance. The source and drain contacts are placed on either side.

The common CV/I metric tends to underestimate the propagation delay and the integration of $C_G \cdot dV_G$ is needed to get the average delay over the whole switching time. Therefore, the mixed-mode 2-D device simulation is employed for more accurate results of circuit performance. The 1st stage shapes the input signal and the 3rd stage provides appropriate loading. The delay was measured at the 2nd stage. Fig. 5.4 shows the intrinsic propagation delay of the SDG and ADG CMOS inverters as a function of fan-out. No wiring capacitance was included in the simulation. Although the absolute value of delay difference is very small in the FO-1 (fan-out=1) condition (which is consistent with [5.6]), the ratio of delay difference is quite significant. Obviously, the SDG outperforms the ADG at the range of 10~20% for the all FO-1~4 circuits, which contradicts previous

results [5.4]. The speed superiority of SDG is mainly from higher driving current, i.e. the higher mobility and larger inversion charge. Therefore, accurate mobility and charge capacitance models are essential in the DG modeling.
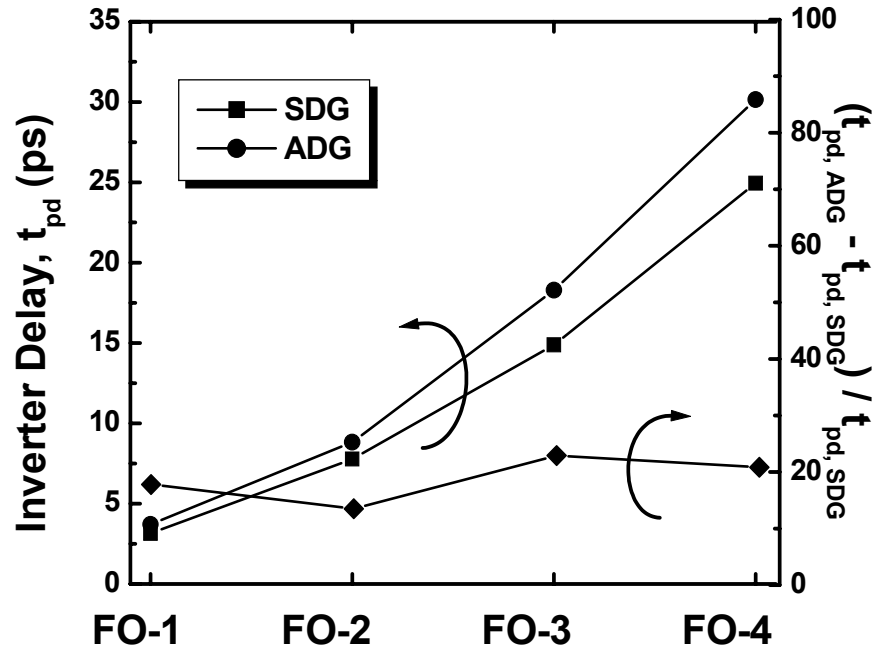
Fig. 5.4 The inverter delay vs. different number of fan-outs for SDG and ADG devices. Although the absolute difference value is small for low fan-out number, the SDG maintains a 10-20% inverter delay improvement over the ADG.

Both the driving current and intrinsic capacitance are related to the inversion charge. It is important to compare the Q_i-V_g performance between the SDG and ADG. By adjusting the gate work functions, the threshold can be reset to equate the Q_i's under the

off-state or subthreshold condition. One can write simple charge-versus-V_g equation

$$\frac{dQ_i}{dV_g} = \frac{d(Q_1 + Q_2)}{dV_g} \approx \frac{C_{ox} \cdot C_{i1}}{C_{ox} + C_{i1}} + \frac{C_{ox} \cdot C_{i2}}{C_{ox} + C_{i2}} \qquad (5.1)$$

The C_{i1} and C_{i2} are negligible in the subthreshold region. For the SDG, the dQ_i/dVg is approximately equal to $2C_{ox}$ after V_g exceeds the threshold voltage. However, for the ADG, the C_{i2} is small and dQ_i/dVg is approximately equal to C_{ox} for thick silicon thickness. The dQ_i/dVg of ADG will be enhanced by the back gate coupling effect if the silicon thickness becomes thinner, but is still smaller than that of SDG [5.2]. Therefore, the SDG has higher inversion charge than the ADG under for the same subthreshold condition.

Fig. 5.5 shows the band diagram of SDG and ADG nMOSFETs under different bias condition. Only one gate turns on at $V_g=V_t$ in ADG. When V_g increases further, both gates will turn on, but the inversion charge of ADG is still smaller than that of SDG. Fig. 5.6 shows the inversion charge distributions for the SDG and ADG device at V_g =0.7V, and V_d=0.05V. The left surface of the ADG is more strongly inverted than both surface of the SDG. However, the total inversion charge in SDG is still higher than that in ADG due to the two inverted surface in the SDG. Note that the inversion charge ratio ($Q_{i,SDG}/Q_{i,ADG}$) is less than 2 due to the channel coupling effect.
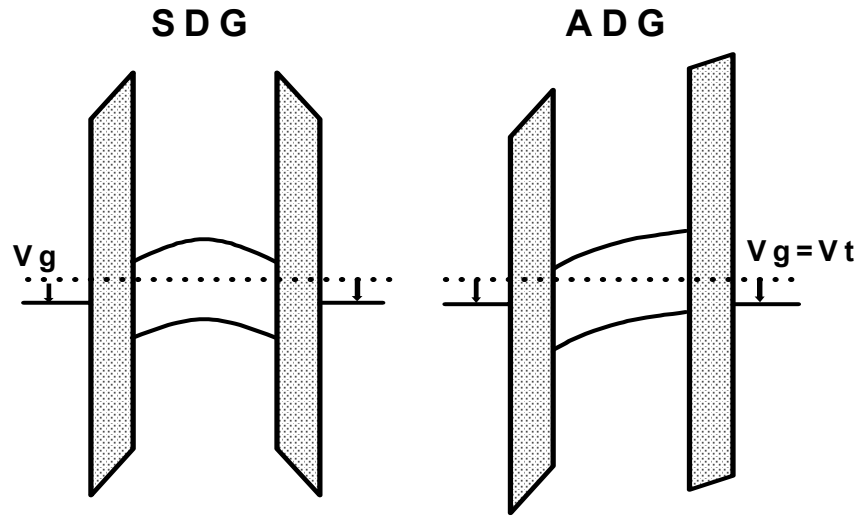
Fig. 5.5 Schematic band diagrams of SDG and ADG nMOSFETs at threshold voltage gate bias. For SDG, the conduction band of silicon body at both surfaces is bent to near the conduction band of the n$^+$ source (dot line). For ADG, however, only the conduction band of silicon body at left surface is bent to the near the conduction band of the n$^+$ source (dot line).
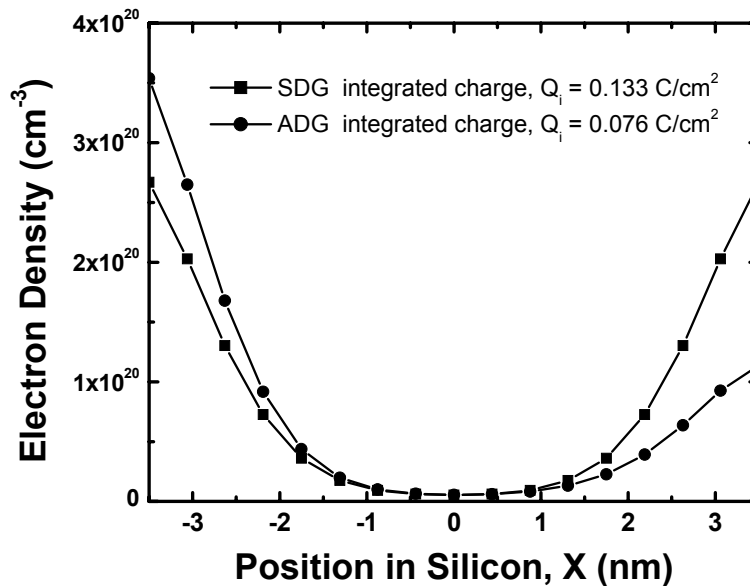
Fig. 5.6 Comparison of electron density in silicon between SDG and ADG at V_g =0.7V, and V_d =0.05V. The ADG is more strongly inverted than both surface of the SDG. However, the total inversion charge in SDG is still higher than that in ADG due to the two inverted surface in the SDG.

The vertical electric field is a critical parameter for mobility performance. Fig. 5.7 shows the vertical electric field distributions for the SDG and ADG device at V_g =0.7V and V_d =0.05V. The vertical electric field of ADG at the left surface is larger than that of SDG, which indicates that the mobility in the ADG is lower than that in the SDG. Therefore, the on-state current of SDG is higher than that of ADG due to the mobility enhancement. Note that the lowest point of the vertical electric field moves left from the right side V_g increases. This point may be used for boundary condition in DG compact model since there is no reference point in DG device. However, the mobility model is under investigation due to the complexity of inversion charge distribution in DG device if the quantum mechanical effect is included into the model.
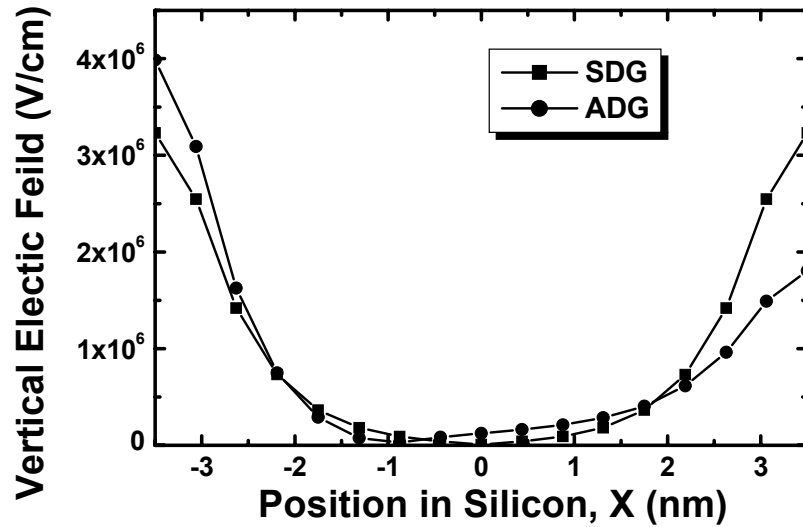
Fig. 5.7 The vertical electric field distribution in the silicon of SDG and ADG device at V_g =0.7V, and V_d =0.05V.

We need to calculate charge distribution accurately in the silicon film and calculate the average position of electrons from the strongly scattering interface.

The FO-4 inverter delay is a standard technology benchmark used to predict the delay of more complex circuits. The energy–delay product is often used as the ultimate quality metric. The energy dissipation vs. inverter delay of the FO-4 inverter chain for SDG and ADG are shown in Fig. 5.8. The total energy dissipation of the inverter is measured as

$$\text{Energy} = \int_{0}^{T} I(t) \cdot V_{dd} \cdot dt \qquad (5.2)$$

where I(t) the sum of the dynamic and standby leakage currents, and T is the clock cycle.

Fig. 5.8 The energy dissipation vs. inverter delay for SDG and ADG devices. For a given amount of energy, the SDG achieves a lower inverter delay due to the higher mobility and absence of input capacitance due to coupling with the back gate.

For a given amount of energy, the SDG achieves a lower inverter delay than the ADG due to the higher mobility. The SDG is more efficient than the ADG in the application of logic circuit.

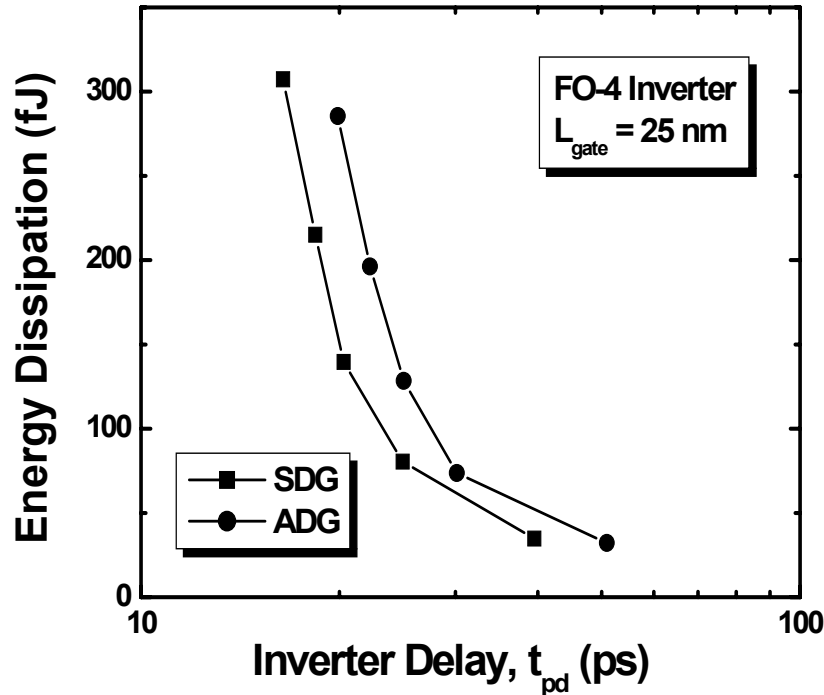The delay equation including the switching resistance, input, and output capacitance is the following [5.7]:

$$t_{pd} = R_{sw} \times \left(C_{out} + FO \times C_{in} + C_{Int} \right) \tag{5.3}$$

R_{sw} is defined as the slope of the delay-versus-wiring-capacitance line, and is inversely proportional to the driving current. C_{out} represents the equivalent capacitance at the output node of the sending stage, and C_{in} represents the equivalent capacitance of the receiving stage to the sending stage. Fig. 5.9 shows the inverter delay versus the wiring capacitance C_{Int} for the FO-4 inverter. The slope (R_{sw}) of SDG is ~30% smaller (better) than the ADG due to larger inversion charge (i.e. higher intrinsic gate capacitance, C_g), and higher mobility. For the highly loaded circuits, where the C_{Int} dominates, the delay ratio is given by the R_{sw} ratio. The intrinsic delay, $R_{sw}(C_{out}+4C_{in})$, of the SDG is 20% better than that of the ADG, which indicates that ($C_{out}+4C_{in}$) for SDG is ~8% higher than ADG. Although the C_g of SDG starts out more than 8% higher than ADG, it is diluted by C_{out} and fringe capacitances. The SDG shows less sensitivity to the wiring capacitance than the ADG.
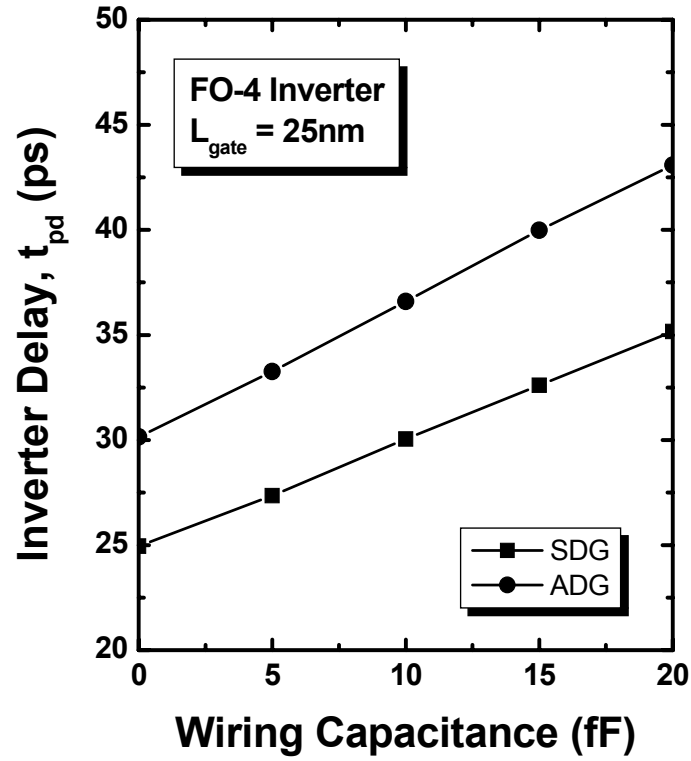
Fig. 5.9 The Inverter delay (t_{pd}) vs. parasitic wiring capacitance C_{Int} for FO-4. The SDG shows less sensitivity to wiring capacitance due to the smaller switching resistance.

In summary, The SDG device shows ~10-20% speed and energy efficiency superiority over the ADG device. Although the ADG structure has the advantage of threshold control flexibility, our evaluation indicates that the SDG SOI CMOS is the most promising candidate for the next generation of high performance CMOS technology

5.3 Circuit Performance Comparison of CMG and IMG

FinFET is considered one of the most promising structure to suppress the short channel effects for a given equivalent gate oxide thickness by using two gates to control the channel [5.1, 5.8]. In the conventional connected multi-gate transistors (CMG), advanced gate work function engineering is needed for the threshold voltage (V_{th}) control [5.9]. Dynamic V_{th} control will inevitably be required for future low power circuit design. The independent-gate FinFET (IDG) shows flexible V_{th} control through the use of second-gate bias [5.10-5.11] (the second gate of NMOS and PMOS FinFETs are switched independently). A detailed description of the device behaviors in the form of compact model is extremely important to explore the advantage of using FinFET for application development. This section describes the concept of the dynamic V_{th} control in the compact modeling of FinFET. The model is implemented into Berkeley SPICE3 and verified with multiple-dimensional device simulator.

V_{th} is usually defined as the gate voltage necessary for obtaining a band bending of $2\phi_B$, which does not apply to FinFET due to the strong coupling effect between the two gates. The surface potential at threshold voltage is smaller than $2\phi_B$ [5.2]. The threshold voltage model of FinFET that works with different operation modes is modified from conventional fully-depleted SOI V_{th} model [5.12]. Multiple-dimensional device simulator

DESSIS is used to investigate the dynamic V_{th} control with different device parameters and to calibrate the V_{th} model. V_{th} is extracted using the gate voltage where the derivative of the transconductance reaches a maximum, i.e. $d^3I_D/dV_G^3=0$, with $V_{DS}=50mV$. Fig. 5.10 shows the V_{th} of first gate (V_{th1}) as a function of V_{G2} for different substrate doping (N_{sub}) and body thickness (T_{Si}). The V_{th1} shift rate ($-\Delta V_{th1}/\Delta V_{G2}$) increases from 0.12V/V to 0.30V/V with decreasing T_{Si} from 50nm to 20nm. The thinner body is more effective in controlling the V_{th} in a fixed biasing range.
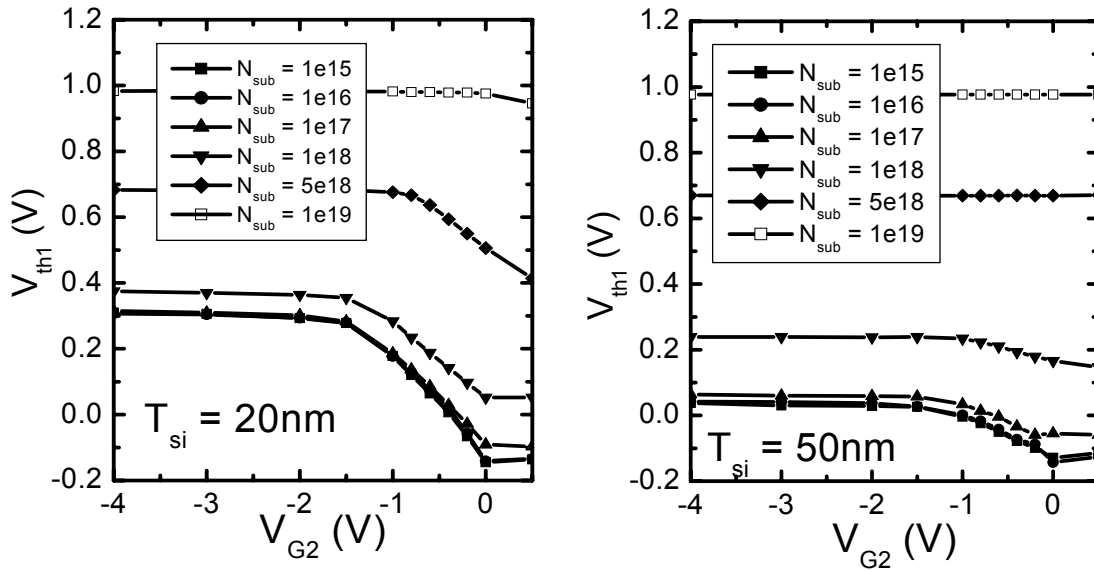
Fig. 5.10 V_{TH1} vs. V_{G2} with different substrate doping at T_{si} = 20nm (left) and T_{si} = 50nm (right). The gate length of the device is 100nm, and T_{ox1}= T_{ox2}= 2nm. The work function of gate 1 and gate 2 are identical (n+ poly).

This coupling effect from the second gate diminishes with increasing N_{sub} due to the transition from fully depleted to partially depleted body. If a large negative V_{G2} is applied, the surface potential of the second gate is pinned, and the V_{th1} will be independent of V_{G2}. Hence, the strong potential coupling between the two gates happens when the second gate is biased at depletion region (Fig. 5.11).
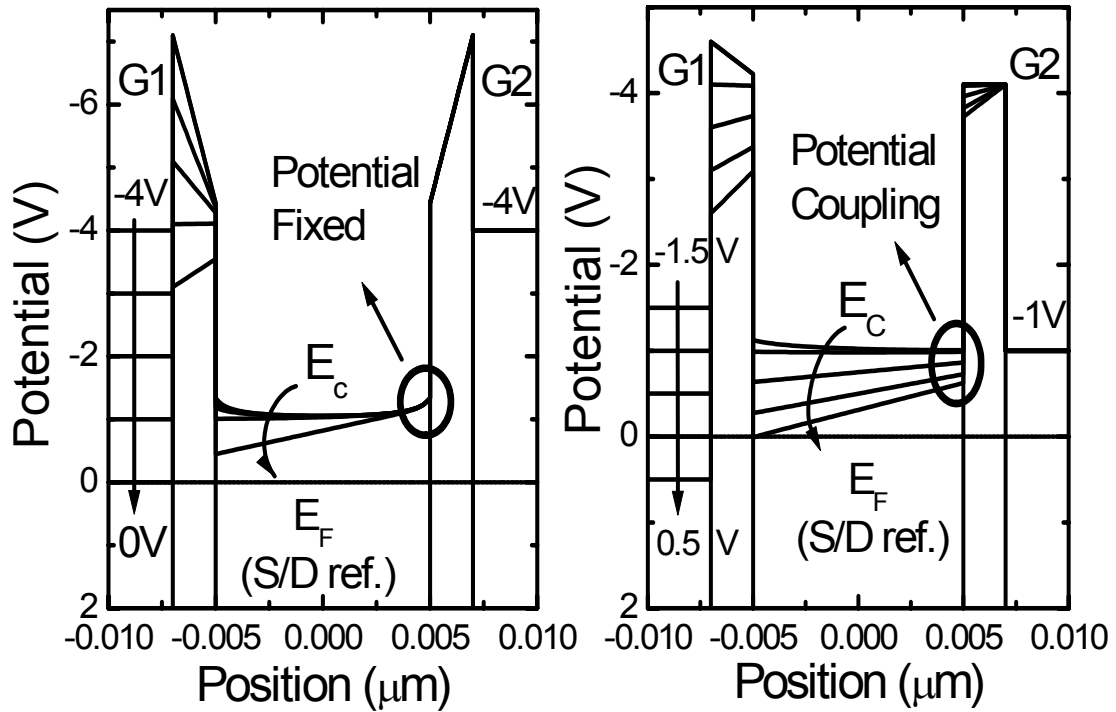
Fig. 5.11 The band diagrams of independent-gate FinFET biased at −4V (second gate accumulated, left) and −1V (second gate depleted). The surface potential of second gate is fixed at large negative bias. There is no coupling between two channels.

Fig. 5.12 shows the V_{th1} as a function of N_{sub} and T_{Si} at a fixed V_{G2} of -1V. With lower

N_{sub}, V_{th1} is independent of N_{sub} and inversely proportional to T_{Si} (volume inversion plus

short channel effect at subthreshold region). However, with higher N_{sub}, V_{th1} increases

with N_{sub} increasing, which is similar to the bulk device. These phenomena should be

included in the model to capture the coupling behaviors in FinFET with different

operation modes.
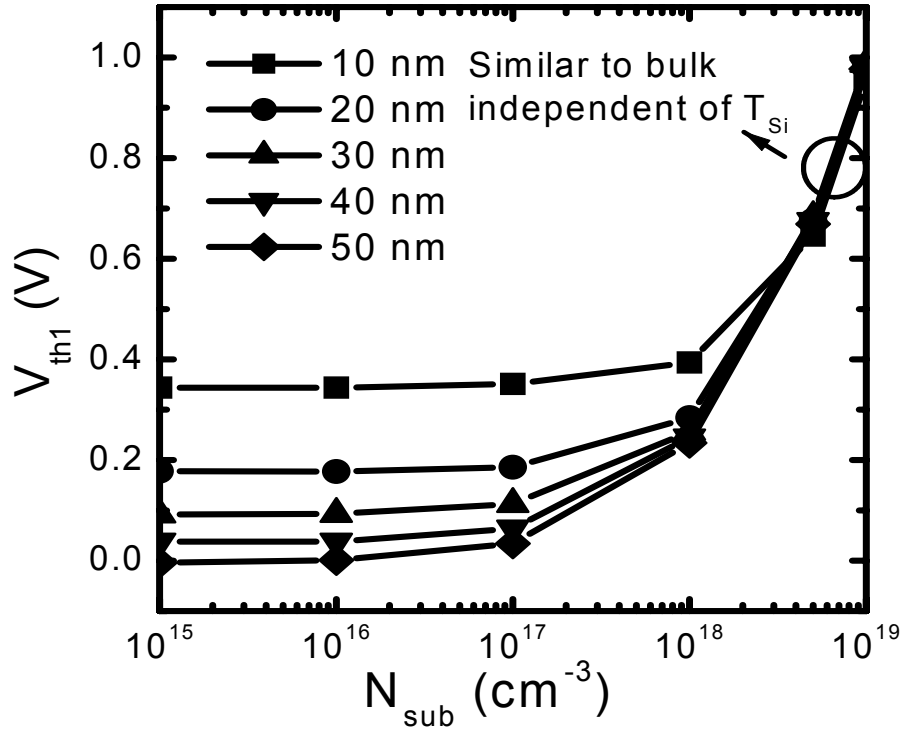
Fig. 5.12 V_{th1} vs. N_{sub} with different T_{Si} from 10 to 50nm. In low N_{sub}, V_{th1} is independent of N_{sub}, but inversely proportional to T_{Si}.

Fig. 5.13 shows the fitting results of the modified V_{th} model and DESSIS simulation for

IDG operation mode. After considering the short channel effect and drain-induced barrier

lowering, the model accurately fits the simulation output with different N_{sub} and T_{Si}.
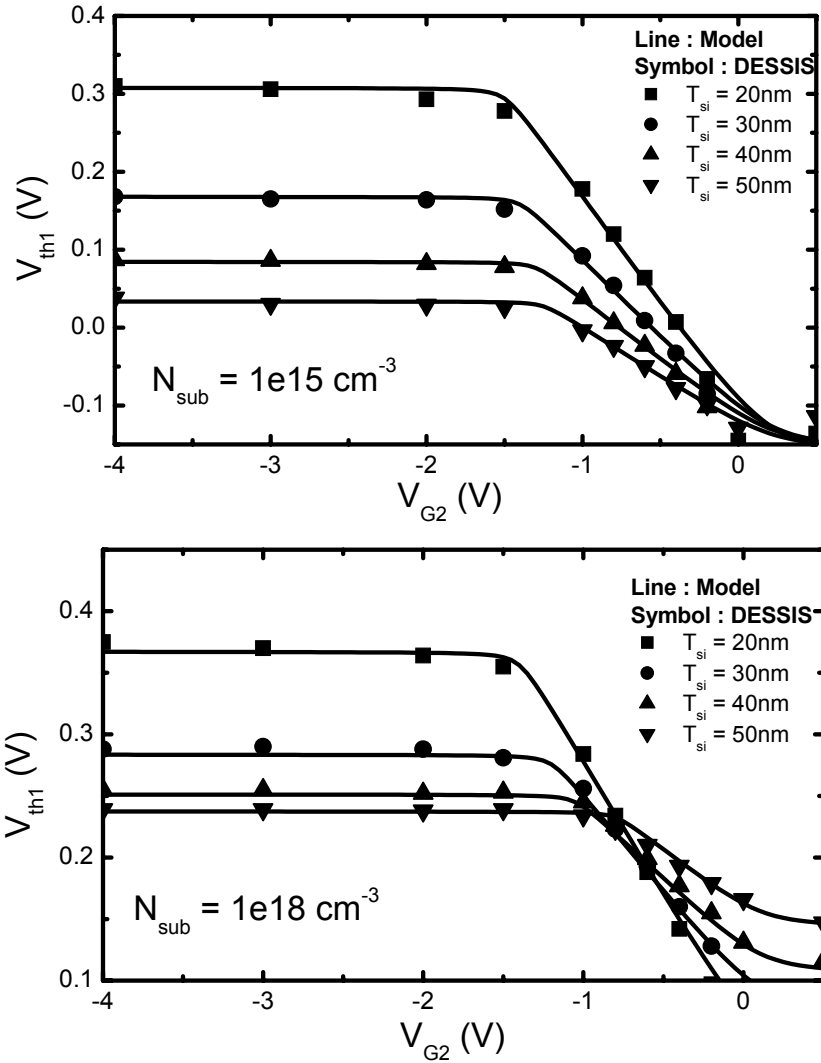
Fig. 5.13 V_{th1} vs. V_{G2} with different body thickness at $N_{sub} = 1e15 cm^{-3}$ (top) and $N_{sub} = 1e18 cm^{-3}$ (bottom). The gate length of the device is 100nm, and $T_{ox1} = T_{ox2} = 2nm$. The threshold voltage model modified from conventional FD-SOI model shows very good fitting results as compared to 2-D device simulation results

The basic charge formulation of the core FinFET compact model is based on surface-potential. The currents at the two interfaces are calculated separately and added together taking into consideration the coupling between the two gates. The I-V and intrinsic C-V models are successfully implemented into SPICE3 environment with DC, AC, and transient behavior. Fig. 5.14 shows the model calibration results of n-type CMG with multiple-dimensional device simulator. The gate length of the device is 35nm and oxide thickness is 1.7nm. The silicon fin is 20nm thick with fin height of 50nm, and the substrate doping is 10^{16}cm^{-3}.
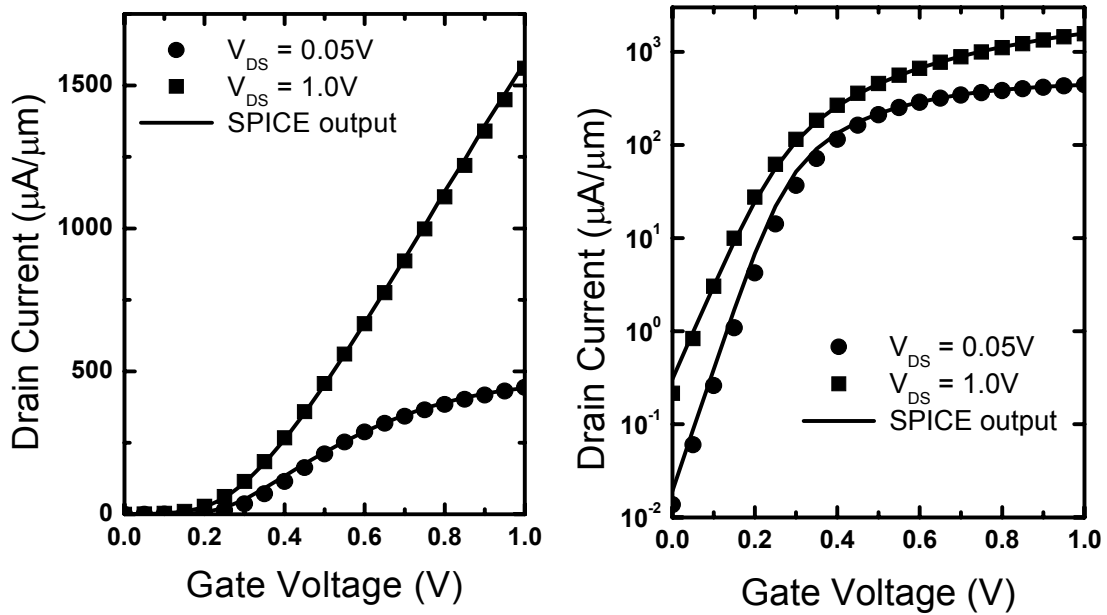
Fig. 5.14 Model verification with 3-D device simulator. The Id-Vg curves of n-type CMG at different drain voltage show good fitting results in linear (left) and log (right) scale.

The outputs of SPICE3 accurately fit the simulation results in both subthreshold and strong inversion regions with different drain biases. By using the same model card, we also obtain good fitting results for the independent-gate mode with V_{G2}=0V (ground-plane (GP), Fig. 5.15), which indicates our model can accurately predict the behaviors of FinFET under independent-gate operation mode. The 17 stage CMG and GP ring oscillator (RO) without loaded capacitance is simulated by using SPICE3. Note that for a fair comparison, the CMG and GP devices were adjusted to have similar off-state current and effective width.
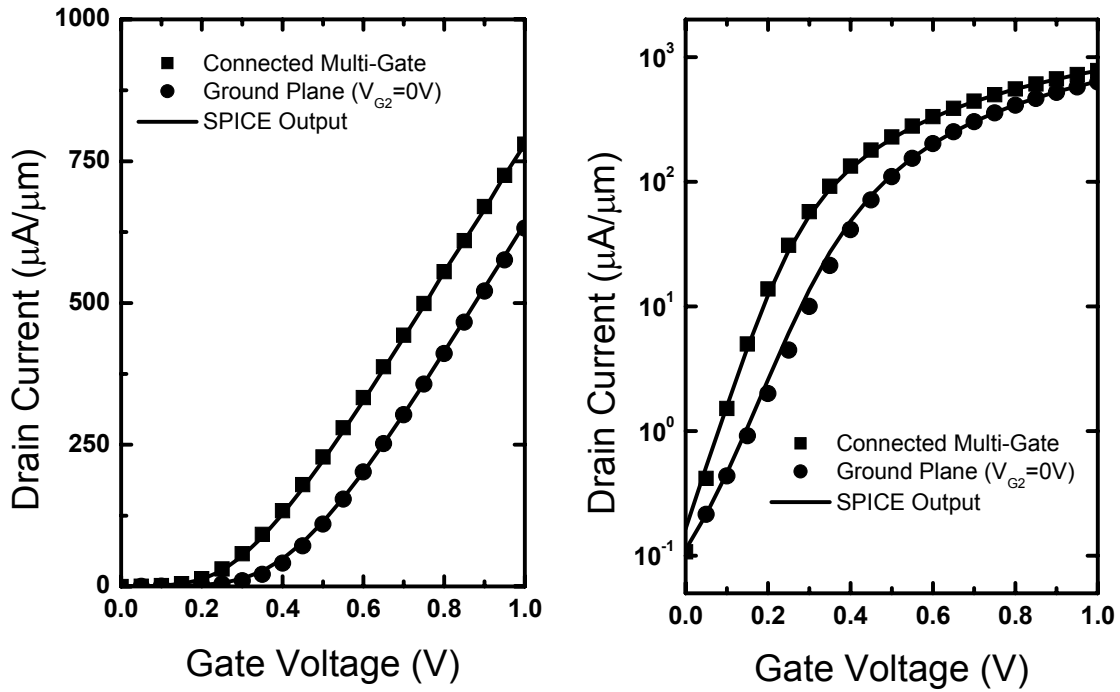
Fig. 5.15 Model verification with 3-D device simulator. The model accurately fits both CMG and independent-gate FinFET in linear (left) and log (right) scale.

Fig. 5.16 shows the stage delay and power dissipation comparison of CMG and GP as function of RO supply voltage. The CMG RO shows lower propagation delay than GP RO. For a given amount of power, the CMG achieves a lower stage delay than the GP due to the better subthreshold behavior for a fixed off-state current (Fig. 5. 17). The CMG is more efficient than the GP in the application of logic circuit. Although the FinFET under GP mode is slower than CMG, the dynamic V_{th} tuning of independent-gate FinFET is attractive to simultaneously achieve high performance during active periods and low leakage power during idle periods for ultra-low power circuit design.
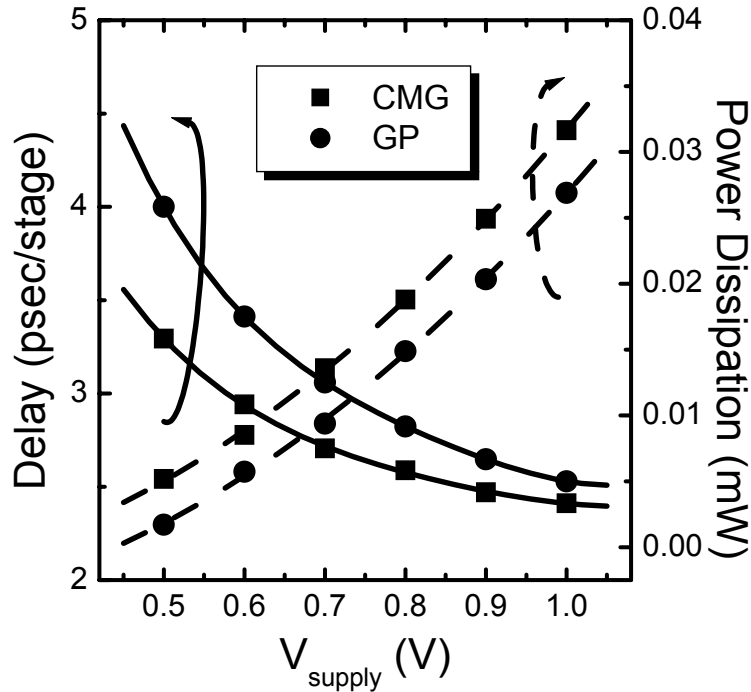
Fig. 5.16 The ring oscillator delay and power dissipation as function of the supply voltage under CMG and GP modes. The gate length is 35nm with body thickness of 20nm.
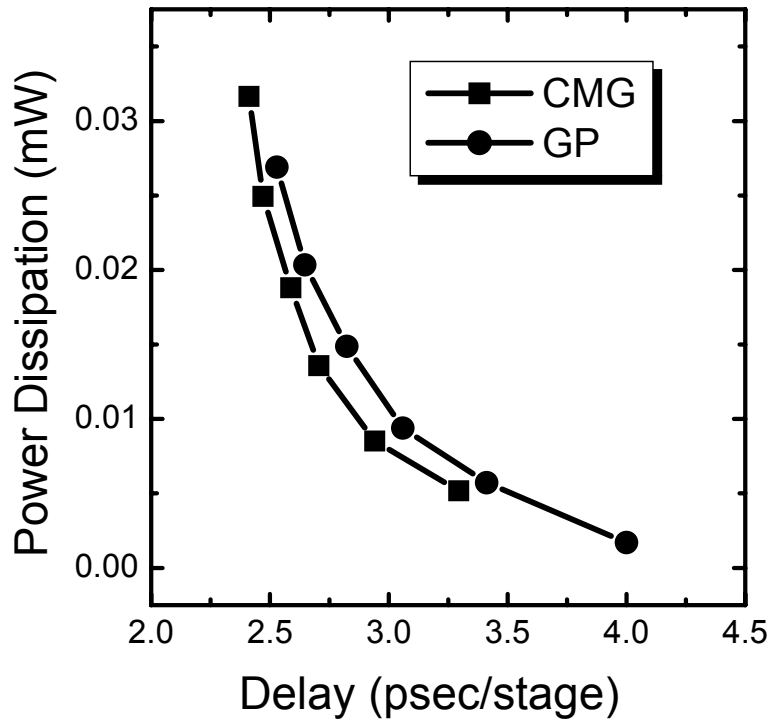
Fig. 5.17 The power dissipation vs. delay of CMG and GP. The CMG is more efficient than the GP due to better subthreshold behavior for a fixed off-state current. If C_{out} and C_{Int} are considered, the low-power advantage would be even greater.

In summary, the impact of FinFET device parameters, such as T_{Si}, and N_{sub}, on dynamic V_{th} control is discussed. The compact modeling of FinFET, which allows two gates to operate independently, has been implemented into Berkeley SPICE3 and shows good prediction of device characteristics and circuit performance.

5.4 V$_{DD}$ Scaling for FinFET Logic and Memory Circuits: Impact of Process Variations and SRAM Stability

FinFETs are considered as the most promising structure down to 22nm node [5.1, 5.8]. Both FinFET-based logic and SRAM have been demonstrated recently [5.13-5.14]. However, with scaling of the device dimensions, process-induced variations cause an increasing spread in the distribution of circuit delay and power, affecting the robustness of VLSI designs [5.15]. SRAM has become the focus of technology scaling since embedded SRAM is estimated to occupy nearly 90% of the chip area in the near future [5.5]. Due to the area-constrained limit, the device fluctuation in the SRAM cell is significant. In this section, we explore the performance of FinFET technology in digital circuit applications at 90 nm technology node under various device parameter variations. Comprehensive comparison of FinFET vis-à-vis PD-SOI has been done for logic gates as well as memory structures that are most commonly used in commercial VLSI designs. We also compare the performance of these two technologies at ultra-low voltages for future low-power applications.

The FinFET compact device models have been modified from Compact Model Council (CMC)-standard planar partially-depleted silicon-on-insulator (PD-SOI) model BSIMPD from University of California, Berkeley [5.16]. This modified model is similar to the

model used in the previous study of FinFET SRAM [5.17]. Fig. 5.18 shows the schematic

of the modified model. Digital static CMOS, transmission gate circuits, and 6-T FinFET

SRAM have been investigated. The basic digital circuit structure used for simulation is

similar to [5.18]. Appropriate fan-out (=FO) has been used for each gate type, based on

real designs. Transistor design parameters used in this study is similar to [5.17]. For

FinFET technology, the thicker gate oxide (1.5nm) is allowed for lower gate leakage

power. FinFETs have lower linear threshold voltage (V_{tlin}) and drain-induced barrier

lowering (DIBL).

Fig. 5.18 View of the schematic of FinFET model. The first order FinFET features
are included in the model.

The I_{on} of PFET is identical for the FinFET and PD-SOI. The I_{on} of N-FinFET is slightly lower.

Comparison of FinFET and PD-SOI delay under V_{DD} scaling has been shown for a FO4 inverter delay chain in Fig. 5.19. The same transistor sizes have been used for both FinFET and PD-SOI for all comparisons in this paper. It is observed that at ultra-low voltages, say 0.6 V, FinFET delay is about 38% less compared to PD-SOI. Similar trends are shown by all other logic gates; plots have been omitted for brevity. Since the transmission gate is very sensitive to (V_{DD}-V_{th}), the increased current drive of FinFET (due to lower V_{tlin}) gets reflected in additional delay improvements at ultra-low voltages. This simulation data clearly demonstrate that FinFETs are an ideal candidate for ultra-low voltage logic applications. At high voltages, the current drive of PD-SOI increases due to worse DIBL, which makes the performance of PD-SOI similar to FinFETs. Power number comparison of FinFET and PD-SOI under V_{DD} scaling is shown in Fig. 5.20. The average power of the FO4 inverter circuit is plotted versus delay under normal switching conditions. Similar trend is shown by all logic gates. For a given circuit performance, FinFETs are more power efficient than PD-SOI for ultra-low voltage application.
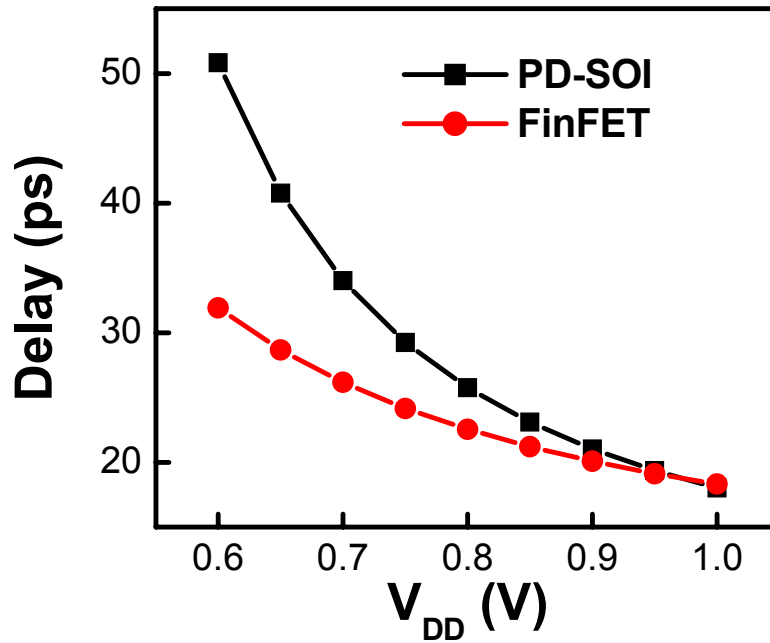
Fig. 5.18 Delay of PD-SOI and FinFET FO4 inverter chain under V_{DD} scaling. FinFET shows superior performance at ultra-low voltages.
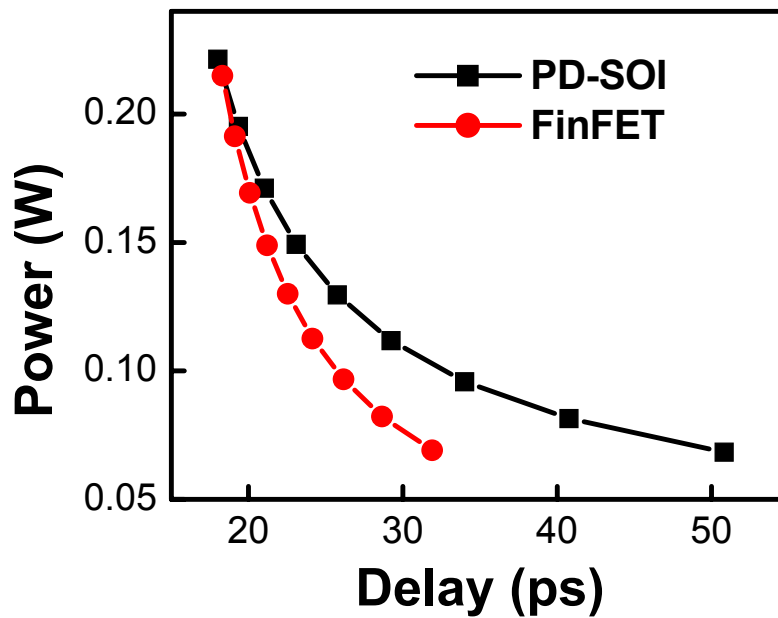

Fig. 5.19 Power-delay merit of PD-SOI and FinFET FO4 inverter chain. FinFET is more power efficient for given delay at ultra-low voltages.

Since process variation has significant impact on ultra-low voltage circuits, we study

the delay sensitivities of the various logic gates under some of the important device

parameter variations like L_{eff} (effective channel length), V_{th} and R_{ds} (source/drain

resistance). Our simulation results show that L_{eff} variation has similar impact on the

performance of all logic gates for both FinFET and PD-SOI since the device design

points are similar for both technologies. For brevity, we have included the plot for just the

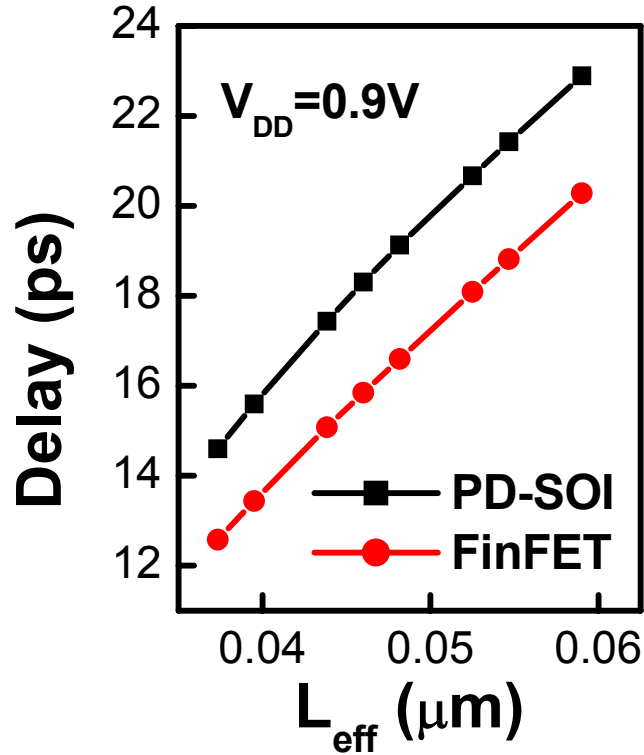transmission gate in Fig. 5.20 at V_{DD} of 0.9V.

Fig. 5.20 L_{eff} variation of transmission gate inverter chain shows similar impact for PD-SOI and FinFET due to fixed device design point.

Transmission gates are very sensitive to V_{th} variation since on-resistance is determined by (V_{DD}-V_{th}). In Fig. 5.21, the transmission gate delay sensitivity has been shown for FinFET and PD-SOI as V_{th} is varied from -80 mV to +80 mV of the nominal value. While the delay varies by 229% for PD-SOI, significantly improved stability is observed for FinFET (121%) at ultra-low voltage (V_{DD}=0.6V) due to the lower V_{tlin} (higher V_g-V_t). Similar reduction in delay sensitivity to V_{th} variation has been observed for other types of logic gates; results of NAND3 have been shown in Fig. 5.22 as an example.
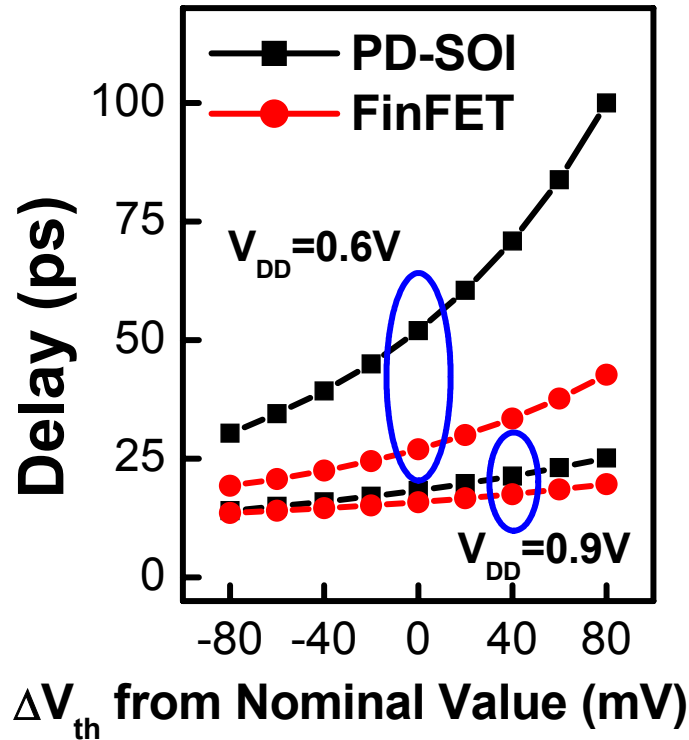
Fig. 5.21 V_{th} variation of transmission gate. Reduction of sensitivity to V_{th} variation is observed in FinFET due to lower V_{tlin}
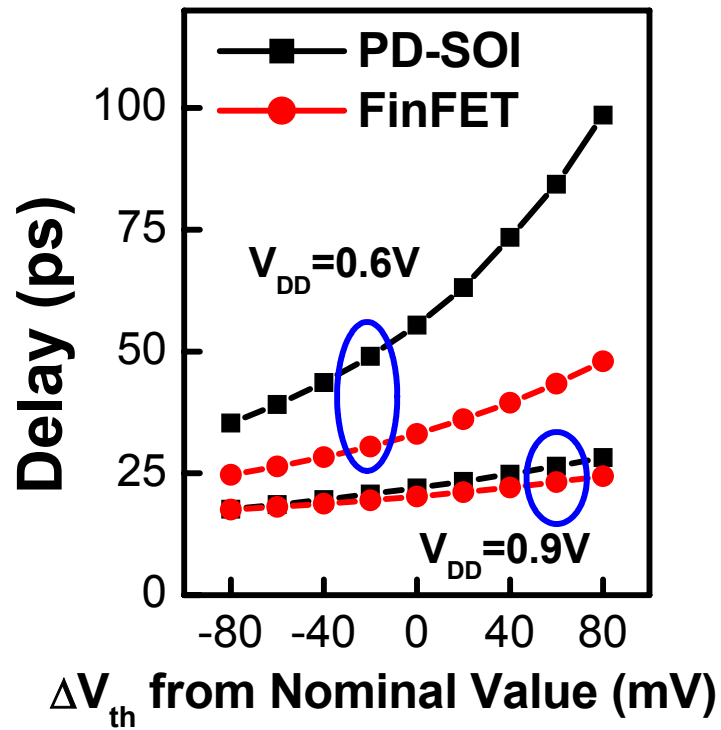
Fig. 5.22 V_{th} variation of NAND3. Reduction of sensitivity to V_{th} variation is observed in all FinFET logic gates.

Static gates with stacked NFETs/PFETs like NORs and NANDs are most sensitive to the variation of R_{ds}. In Fig. 5.23, the NAND4 gate delay is studied as R_{ds} is varied from -100 ohms to +100 ohms of the nominal value; the plots show that FinFET has a slightly reduced R_{ds} sensitivity (18%) as compared to PD-SOI (22%) due to lower gate capacitance of FinFET.
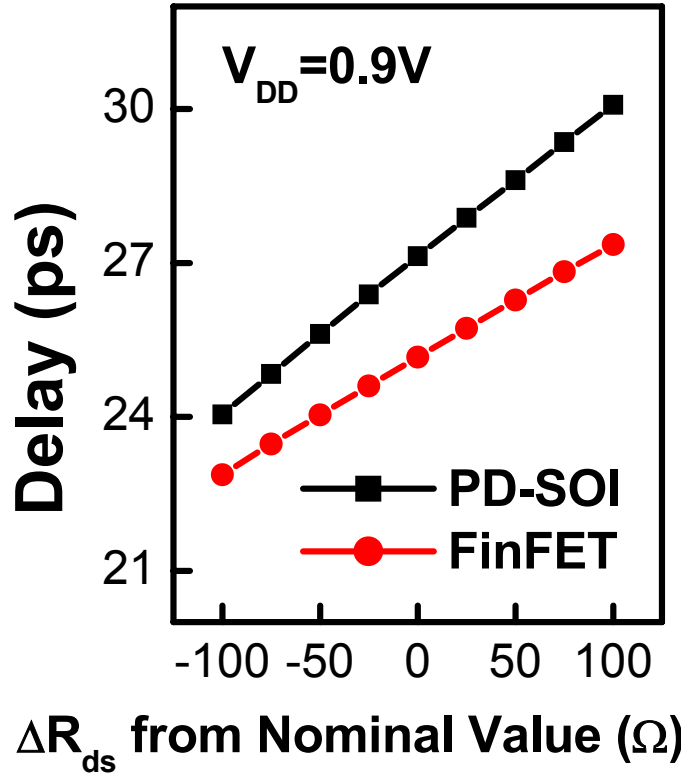
Fig. 5.23 R_{ds} variation of NAND4. Slightly reduction of sensitivity to R_{ds} variation is observed in all FinFET logic gate due to lower gate capacitance.

The fundamental stability problem in the 6T-SRAM is caused by a potential disturbance via the pass-gate to the "0" storage node in the read condition (Fig. 5.24). The static noise margin (SNM) in the read condition is investigated for transistor biasing, sizing, and variations. Fig. 5.25 shows the butterfly curve with various V_{DD} of FinFET 6T-SRAM cells. FinFET SRAM is very robust at ultra-low voltages. The beta ratio (current drive ratio between pull-down NFET and access NFET) is a quantized number.
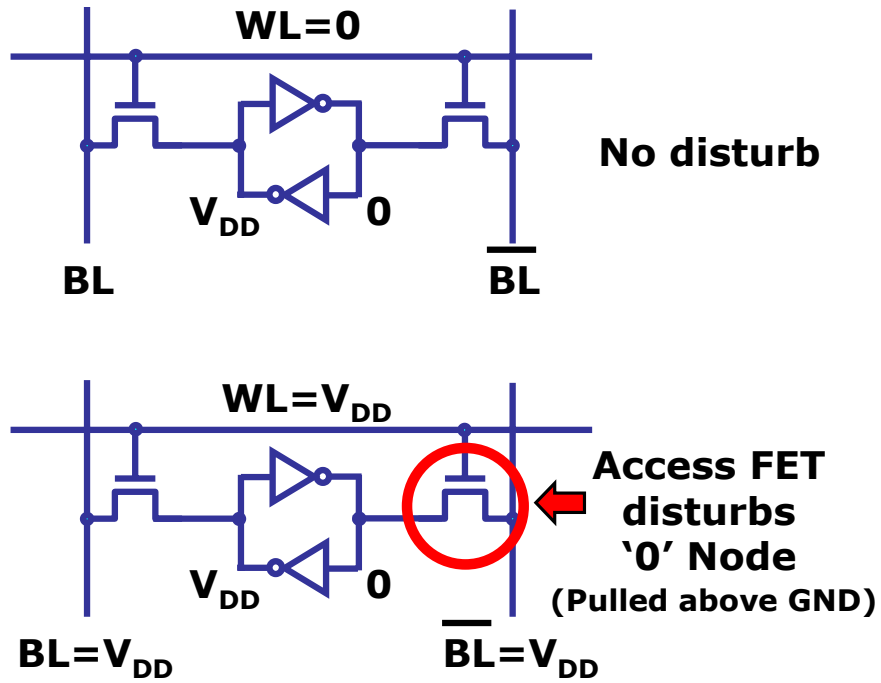
Fig. 5.24 6T-SRAM under read operation. Access transistor disturbs the internal node (passes a "1").
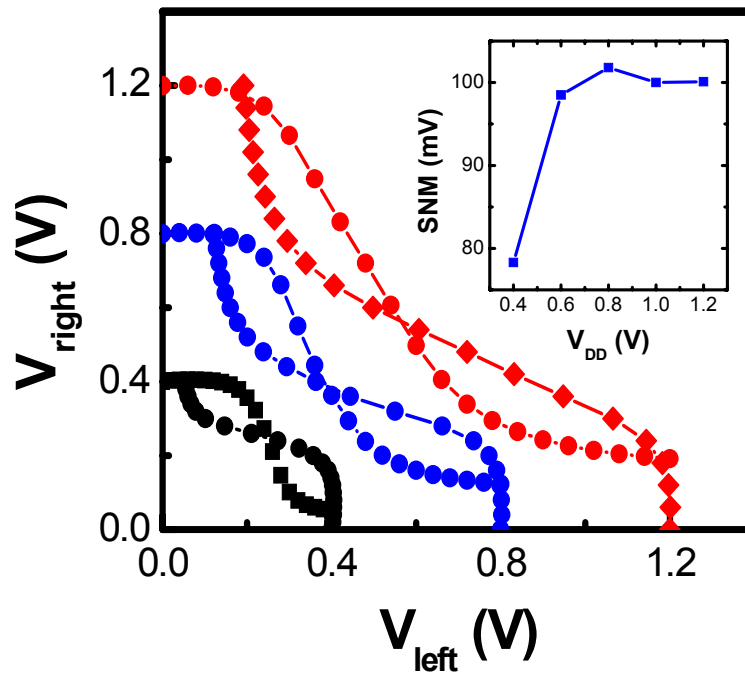
Fig. 5.25 Butterfly curves of FinFET 6T-SRAM cell under read operation at various V_{DDS}. FinFET SRAM is robust for ultra-low voltage applications.

A reliable SNM can be maintained at 0.6 V and tuned by different beta ratios (Fig. 5.26). However, there is the trade-off between the read stability and cell area. Since the process variations have significant impact on the minimum-geometry devices such as SRAM, V_{th} variation is studied by considering the worst case of variation in the paired pull-down NFETs (meaning V_{th} varies in opposite directions in the paired pull-down NFET, Fig. 5.27). The V_{th} variation due to random dopant and T_{si} fluctuations shows very significant impact on the read stability.
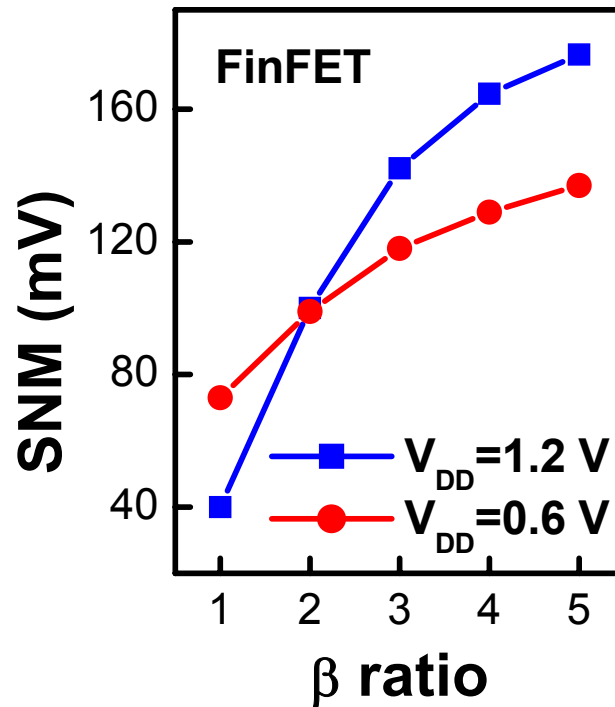
Fig. 5.26 Wider SNM FinFET SRAM cell can be achieved by increasing beta ratio w/ the trade-off of cell area. FinFET SRAM is stable for ultra-low voltage applications.
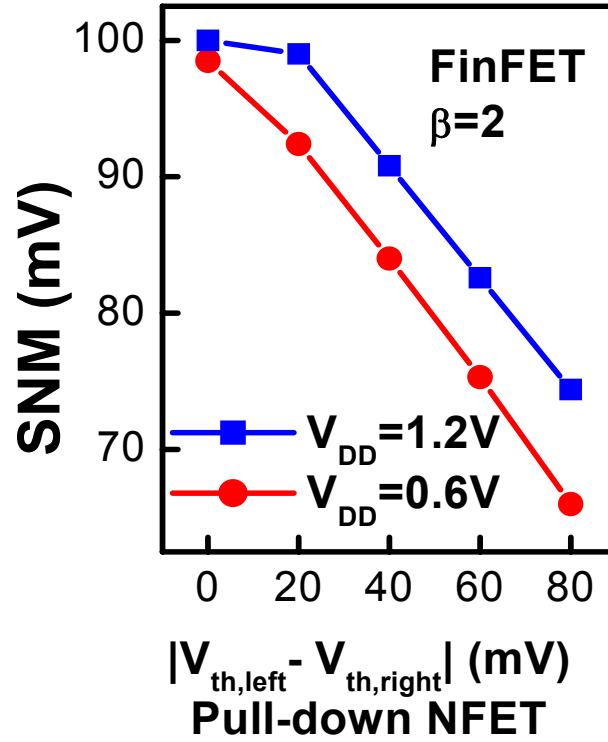
Fig. 5.27 Random dopant fluctuation- and T_{si}-induced V_{th} variation between paired pull-down NFET has significant impact on SNM.

In summary, FinFET-based logic and memory circuits for ultra-low voltage applications are analyzed in detail in this section. It clearly shows that FinFET is an ideal candidate for ultra-low voltage applications because of its robustness to various device parameter variations. It demonstrates that a reliable SNM of FinFET SRAM can be maintained at ultra-low voltages under V_{th} variation caused by random dopant and T_{si} fluctuations.

5.5 Summary

Several important concepts of designing MG-based circuits are addressed. The SDG device shows ~10-20% speed and energy efficiency superiority over the ADG device. Although the ADG structure has the advantage of threshold control flexibility, our evaluation indicates that the SDG SOI CMOS is the most promising candidate for the next generation of high performance CMOS technology.

For a given amount of power, the CMG achieves a lower stage delay than the GP due to the better subthreshold behavior for a fixed off-state current. The CMG is more efficient than the GP in the application of logic circuit. However, the dynamic V_{th} tuning of independent-gate FinFET is attractive to simultaneously achieve high performance during active periods and low leakage power during idle periods for ultra-low power circuit design. The dynamic V_{th} tuning can also be used for reducing the impact of process variations.

FinFET-based logic and memory circuits for ultra-low voltage applications are analyzed. Our results clearly show that FinFET is an ideal candidate for ultra-low voltage applications because of its robustness to various device parameter variations. A reliable SNM of FinFET SRAM can be maintained at ultra-low voltages under V_{th} variation caused by random dopant and T_{si} fluctuations.

5.6 References

[5.1]   D. Frank, S. Laux, and M. Fischetti, "Monte Carlo simulation of a 30-nm dual-gate MOSFET: How far can silicon go?," in IEDM Tech. Dig., 1992, p. 553.

[5.2]   Y. Taur, "Analytic Solution of Charge and Capacitance in Symmetric and Asymmetric Double-Gate MOSFETs," IEEE Trans. Electron Devices, vol. 48, no. 12, p. 2861, 2001.

[5.3]   Tang, L. Chang, N. Lindert, Y.-K. Choi, W.-C. Lee, X. Huang, V. Subramanian, J. Boker, T.-J. King, and C. Hu, "FinFET- A Quasi-Planar Double-Gate MOSFET," in ISSCC Dig. of Tech.l Papers, 2000, p. 118.

[5.4]   K. Kim, and J. Fossum, "Optimal Double-gate MOSFETs: Symmetrical or Asymmetrical Gates?," Proc. IEEE International SOI Conf., 1999, p. 98.

[5.5]   International Technology Roadmap for Semiconductors, 2001.

[5.6]   J. Fossum, L. Ge, and M.-H. Chiang, "Speed Superiority of Scaled Double-Gate CMOS," IEEE Trans. Electron Devices, vol. 49, no. 5, p. 808, 2002.

[5.7]   Y. Taur, and T. H. Ning, Fundamentals of Modern VLSI Devices, Cambridge University Press, 1998.

[5.8]   D. Hisamoto, Wen-Chin Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, Tsu-Jae King, J. Bokor and Chenming Hu: FinFET – A Self-Aligned

Double Gate MOSFET Scalable to 20nm. IEEE Trans. Electron Devices, vol. 47, pp. 2320-2325, Dec. 2000.

[5.9]   J. Kedzierski et al., "Metal-gate FinFET and fully-depleted SOI devices using total gate silicidation," IEDM Tech. Dig., p.247, 2002.

[5.10] I. Y. Yang, C. Vieri, A. Chandrakasan and D. A. Antoniadis, "Back-Gated CMOS on SOIAS For Dynamic Threshold Voltage Control," IEEE TED, vol. 44, no. 5, p. 822, 1997.

[5.11] Yongxun Liu, Meishoku Masahara, Kenichi Ishii, Toshihiro Sekigawa, Hidenori Takashima, Hiromi Yamauchi, and Eiichi Suzuki, "A Highly Threshold Voltage-Controllable 4T FinFET with an 8.5-nm-Thick Si-Fin Channel," IEEE EDL, vol. 25 (7), p.510, 2004.

[5.12] S. R. Banna, P.C.H Chan, P. K. Ko, C.T. Nguyen, Mansun Chan, "Threshold voltage model for deep-submicrometer fully depleted SOI MOSFET's," IEEE TED, vol. 42(11), p. 1949, 1995.

[5.13] B. A. Rainey, D. M. Fried, M. Ieong, J. Kedzierski, E. Nowak, "Demonstration of FinFET CMOS Circuits," DRC Tech. Dig., p. 47, 2002.

[5.14] E. Nowak, T. Ludwig, I. Aller, J. Kedzierski, M. Ieong, B. Rainey, M. Breitwisch, V. Gemhoefer, J. Keinert, D.M. Fried, "Scaling beyond the 65 nm node with FinFET-DGCMOS," CICC Tech. Dig., p.339, 2003.

[5.15] David Burnett, Kent Erington, Chitra Subramanian, and Kelly Baker, "Implications of Fundamental Threshold Voltage Variations for High-Density SRAM and Logic Circuits," VLSI Tech. Dig., p. 15, 1994.

[5.16] BSIMPD MOSFET Model User's Manual, University of California at Berkeley.

[5.17] R. V. Joshi, R.Q. Williams, E. Nowak, K. Kim, J. Beintner, T.Ludwig, I. Aller, C. Chuang, "FinFET SRAM for high-performance low-power application," ESSCIRC, p. 211, 2004.

[5.18] K. K. Das, S.-H. Lo, C.-T. Chuang, K. Bernstein, R. Williams, "Analysis and Design of Digital CMOS Circuits in Hybrid SOI-Epitaxial Technology with Different Crystal Orientations," International SOI conference, p. 177, 2004.

Chapter 6

Statistical Compact Modeling of Variations in Nano MOSFETs

6.1 Introduction

In the sub-45nm CMOS technology regime, the impact of device variations on circuit functionality becomes critial. The scaling of the device geometry makes device characteristics more sensitive to process-related fluctuation. This leads to greater variance of device/circuit performance around the nominal technology node of circuit design. Statistical compact modeling becomes crucial for acurately predicting the statistical variations of VLSI circuit performance such as speed, leakage power, and gain. It is also

important to predict the circuit yield given knowledge of process variations.

The major challenge of statistical circuit performance analysis is how to relate the device-level variation to the circuit-level variation in an accurate and efficient manner. The two conventional approaches for simulating the impact of device variation in a designed circuit are (i) worst- and best-case corner method, and (ii) Monte Carlo (MC) SPICE simulations [6.1]. The worst- and best-case corner method is simple and computationally effcient and thus widely adopted by circuit designers. However, the corner approach usually gives overly pessimistic or optimistic performance prediction due to insufficient attention to the correlations between variations and to electrical test (ET) variation data. The MC SPICE simulation provides more accurate prediction with statistical information but is computationally expensive and impractical for complex VLSI circuits. In addtion, the MC approach also suffers from insufficient attention to the correlation of model parameters and ET variation data [6.2].

Principal Component Analysis (PCA) was introduced to capture the complex correlations of device parameters [6.3]. PCA transforms the correlated device parameters into uncorrelated variables (principal components). Each device parameter is a linear or nonlinear combination of the principal components. There is no phyical interpretation of these principal components. However, the accuracy of PCA can be problematic in nano

scale CMOS technology due to highly nonlinear device characterisitcs and heterogeneity of device parameter patterns [6.4]. To overcome this issue, ET-based Direct Sampling Methodology (DSM) extracts device model parameters for all test sites, and preserves the existing complex nature of parameter correlations [6.4-6.5]. DSM can predict more accurate distribution of circuit performance due to stochastic process variations. However, the accuracy of the DSM relies on the number of generated device parameter sets: the more parameter sets generated, the more accurate result predicted. The efficiency of the device parameter extraction for significant number of test sites could be problematic due to more complex process introduced in advanced CMOS technology.

In this chapter, a novel methodology for generating Performance Aware (Corner/Distribution) Models (PAM) cards is presented. More accurate and application-specific (for speed, power, gain, etc) model cards can be easily generated at any distribution levels (such as $+2\sigma$, -1σ). The detail of generating PAM cards is discussed in section 6.2. In section 6.3, we demonstrate the accuracy improvement of generatd PAM cards by applying it to different scale of logic circuit. The PAM cards also improve the accuracy of Monte Carlo simulation by reconciling the physical and ET variances.

6.2 Methodology of Generating PAM Cards

Fig. 6.1 shows the proposed flow for performance aware modeling of device variation. The inputs to the Berkeley Short-Channel IGFET Model (BSIM) variation modeling process are the nominal BSIM device parameter set (model card), the ET variations (V_{th}, I_{on}, I_{off}, R_{out}, etc.) and the information of stochastic process variations (nominal value and standard deviation (sigma or σ) of L_g, T_{ox}, W, etc.). The nominal BSIM model card already includes the layout-dependent variations (strain, well proximity effect, etc.) [6.6]. To demonstrate the methodology, pseudo ET distribution data are generated using 32nm technology node Predictive Technology Model (PTM) [6.7]. The assumed $3\sigma L_g$ and $3\sigma T_{ox}$ are 10% and 5% of the nominal value, respectively. Fig. 6.2 shows the examples of ET data distribution for 32nm technology node. The mean values of threshold voltage (V_{th}) and I_{on} (I_d biased at $V_{gs}=V_{ds}=V_{dd}$) are 157.4mV and 1.48mA/μm, respectively. The ratio of sigma/mean for V_{th} and I_{on} are 28% and 6.5%, respectively. The distribution of the Ion and Ieff are shown in Fig. 6.3. The electrical variation inputs are first reconciled with the process variations and decomposed into components due to various physical variations. For example, the total V_{th} variance is decomposed into four causes in the present case: (i) L_g variation, (ii) T_{ox} variation, (iii) N_{ch} variation, and (iv) independent variations, such as random dopant fluctuation (RDF) as shown in Eq. 6.1

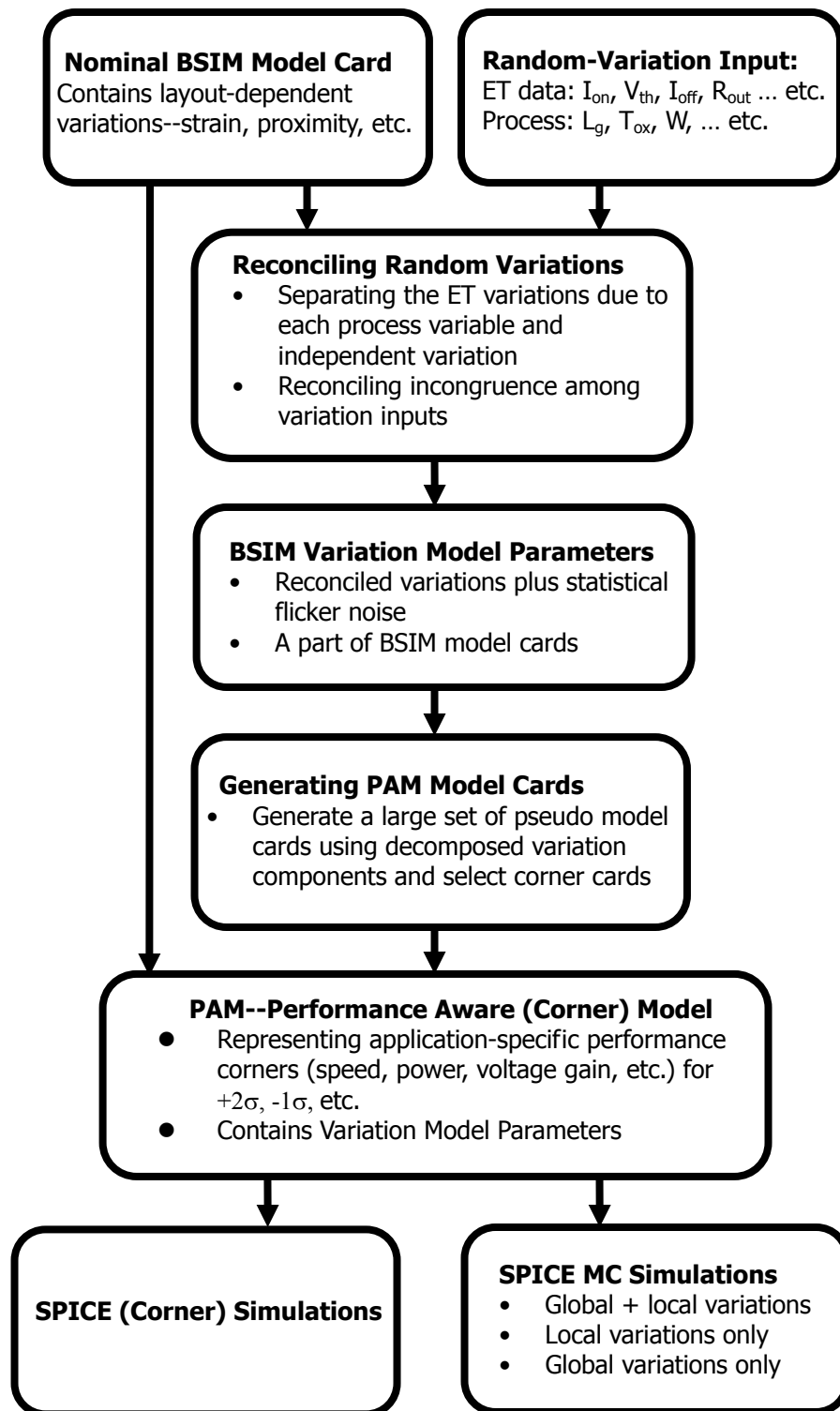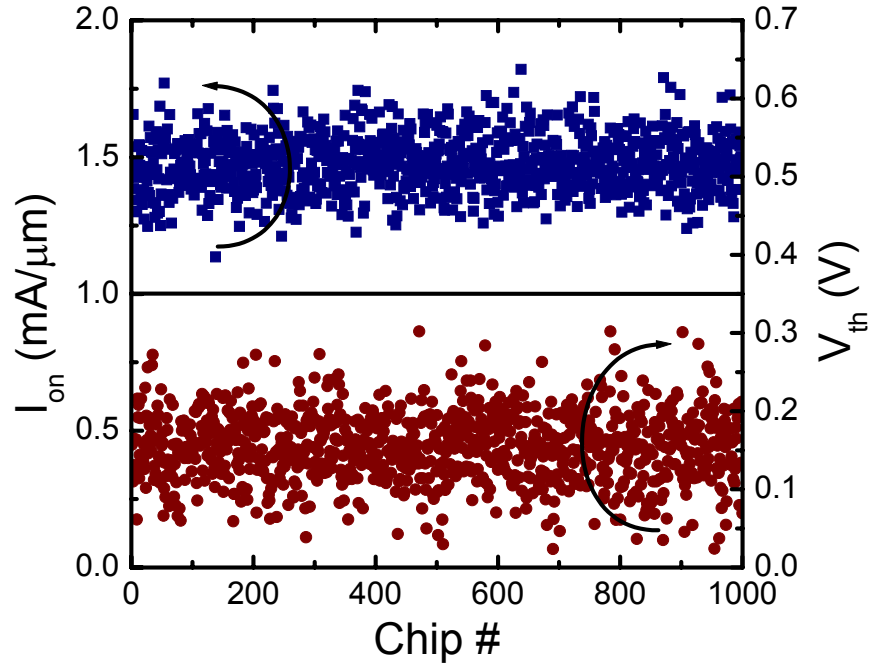Fig. 6.1 Flow chart for statistical compact modeling of device variations

Fig. 6.2 V_{th} and I_{on} variations in the 32nm technology node are generated to represent ET data for illustration. Sigma/mean is 28% for V_{th} and 6.5% for $I_{on.}$
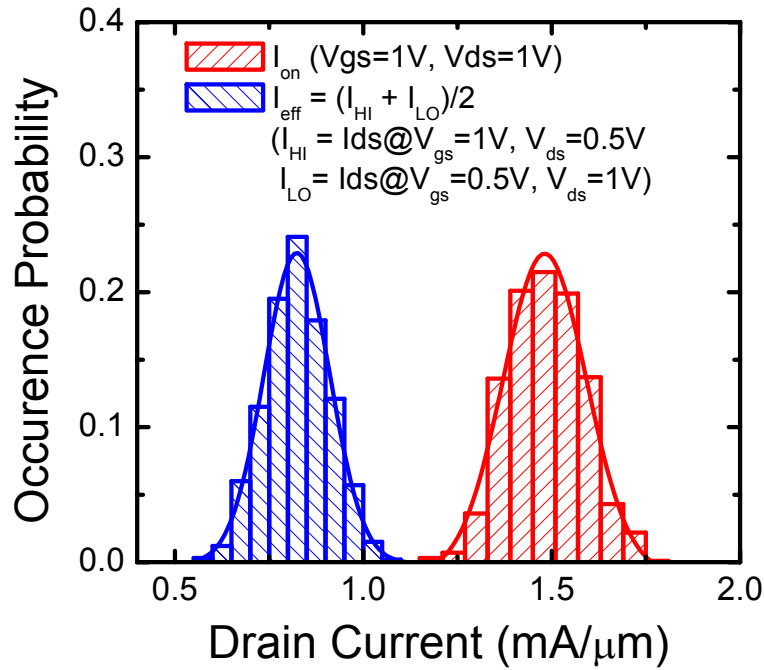

Fig. 6.3 Histogram of I_{on} and I_{eff} distribution. Mean of I_{on} and I_{eff} are 1.48mA and 0.83mA. Sigma of I_{on} and I_{eff} are 0.097mA, and 0.084mA.

$$\sigma V_{th,\,total}^2 = \sigma V_{th,\,Lg}^2 + \sigma V_{th,\,Tox}^2 + \sigma V_{th,\,Nch}^2 + \sigma V_{th,\,independent}^2 \cdot \tag{6.1}$$

The first three variations are modeled with BSIM4 equations and the nominal model card which contains the nominal model parameters. σV_{th} due to RDF is modeled with percolation theory (atomistic simulations) [6.8] with modification for extra model flexibility,

$$\sigma V_{th,\,RDF} = A \cdot \frac{T_{ox} \cdot \left(N_{DEP} \cdot \left(1 + LPE0/L_{eff}\right)\right)^B}{\sqrt{W_{eff} L_{eff}}} \; (V) \tag{6.2}$$

where LPE0 is the lateral non-uniform doping parameter, and B is the doping dependence factor of RDF. The default values of A and B are 3.18×10^{-8} and 0.4, respectively. The unknown variation category is used to reconcile the ET σV_{th} data with the variations attributed to the other causes such as etching and implantation-induced variations.

The relationship between ET variations and physical variations can be treated as nonlinear mode or linear mode. For example, the V_{th} can be expressed as

$$V_{th} = f\left(L_g + \delta L_g, T_{ox} + \delta T_{ox}, N_{ch} + \delta N_{ch}\right) \qquad \text{Nonlinear Mode} \tag{6.3}$$

$$V_{th} = V_{th,nominal} + a \cdot \delta L_g + b \cdot \delta T_{ox} + c \cdot \delta N_{ch} + V_{th,independent} \quad \text{Linear Mode} \tag{6.4}$$

The nonlinear mode is more physical and easier to be implemented in the compact model. However, it might introduce difficulty for reconciling process variations. The linear mode, on the other hand, is easy for reconciling process variations. The a, b, and c are the sensitivity of V_{th} to process variations. If users do not provide a, b, and c, compact model

like BSIM will generate the coefficients by partially differentiating the V_{th} expression in the compact model. If the correlation of the physical parameter variations (δL_g, δT_{ox}, δN_{ch}) is not given, we assume the physical variations are independent variables. By using the Eq. (6.4) combined with Eq. (6.1), we can determine the $\sigma V_{th, independent}$ by subtracting the first three terms (obtained from equations) in Eq. (6.1) from the measured $\sigma V_{th, total}$. The determined $\sigma V_{th, independent}$ is further decomposed into two components,

$$\sigma V_{th, \text{ independent}}^2 = \sigma V_{th, \text{ RDF}}^2 + \sigma V_{th, \text{ other}}^2 . \tag{6.5}$$

This decomposition is important because width and length- dependent RDF is expected to become increasingly important. Hierarchical rules govern how to reconcile cases when a negative component due to other causes is indicated. Based on the knowledge device physics and process technology, the priority of physical variations is determined. For example, if user-specified $\sigma V_{th, total}^2$ is smaller than $a^2 \cdot \sigma L_g^2$, σL_g is reduced so that $\sigma V_{th,total}^2 = a^2 \cdot \sigma L_g^2$. The same process is applied to other ET data (I_{on}, I_{off}, $G_m R_{out}$, etc.). Table 6.1 shows examples of ET σV_{th} and σI_{on} reconciled into various component contributions by this procedure.

Table 6.1 Separation of V_{th} and I_{on} variations into various causes.

ET σV_{th}	43.7mV	ET σI_{on}	96.7 ($\mu A/\mu m$)
$\sigma V_{th, Lg}$	29.6mV	$\sigma I_{on, Lg}$	69.2 ($\mu A/\mu m$)
$\sigma V_{th, Tox}$	1.4mV	$\sigma I_{on, Tox}$	6.2 ($\mu A/\mu m$)
$\sigma V_{th, RDF}$	30mV	$\sigma I_{on, RDF}$	63.5 ($\mu A/\mu m$)
$\sigma V_{th, other}$	11.5mV	$\sigma I_{on, mobility}$	22.2 ($\mu A/\mu m$)

The reconciled variations (σL_g, σT_{ox}, $\sigma V_{th, RDF}$, $\sigma V_{th, other}$, $\sigma_{mobility}$) are used to generate one thousand pseudo model cards. For easier visualization the cards are numbered according to the ascending sequence of I_{on} that each predicts as shown in Fig. 6.4. The stars indicate the median, $\pm\sigma$ and $\pm 2\sigma$ PAM cards based on the I_{on}'s these card generate. As expected, another set of PAM cards based on I_{off} are close but not identical to those based on I_{on} (Fig. 6.5). We propose to select the median, $\pm\sigma$ and $\pm 2\sigma$ PAM cards based on multiple electrical PERFORMANCE metrics (hence the name Performance Aware Models) such as I_{on}, I_{eff}, I_{off}, $G_m \times R_{out}$, and the speed of circuit fabrics.
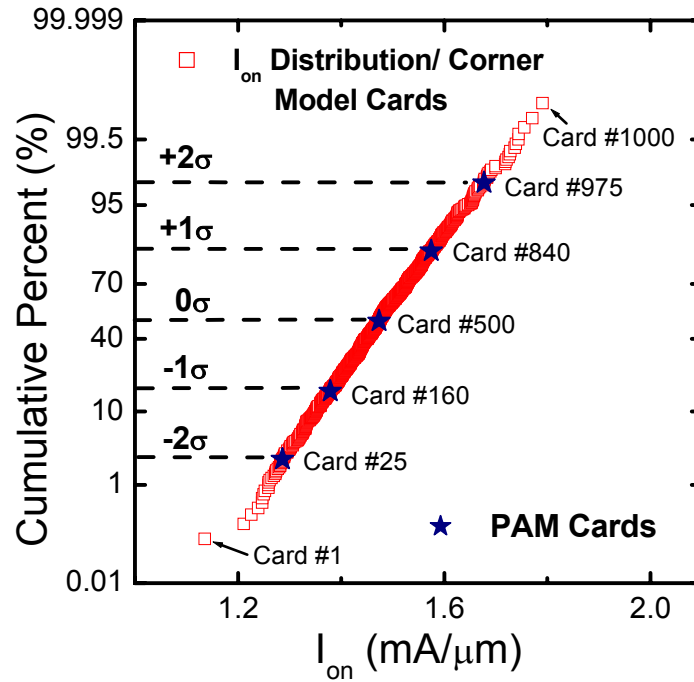
Fig. 6.4 One thousand model cards generated with reconciled variation components. The cards are numbered according to ascending sequence of I_{on} that each predicts.
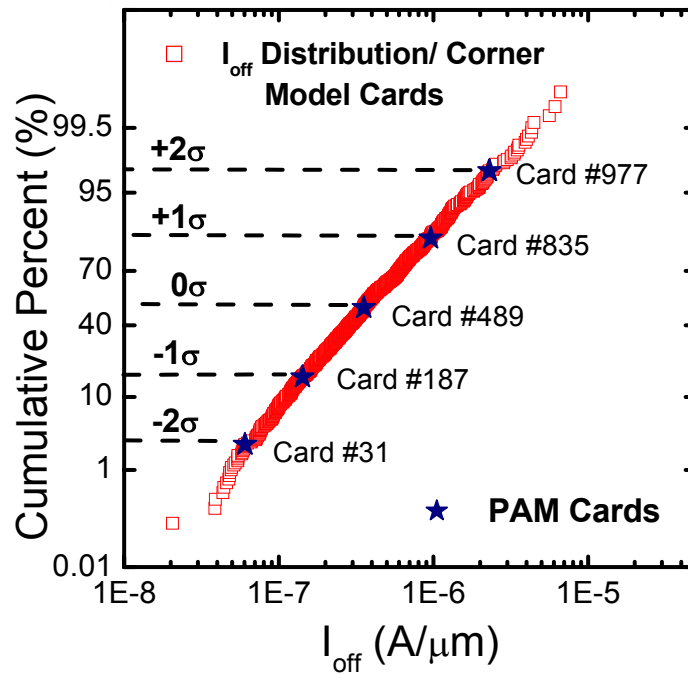

Fig. 6.5 Distribution of off-state leakage current (I_{off}). The represented cards number between $+2\sigma$ and -2σ are identified.

For example, based on the small sample of cards shown in Table 6.2, one may select #845 instead of #840 (based on I_{on} alone) as the $+1\sigma$ card since #845 predicts all electrical performance metrics more accurately than #840. One can add other electrical performance in Table 6.2, of course. One may choose to select a separate set of PAM cards for applications that are particularly sensitive to voltage gain or to leakage power. Therefore, PAM can be application specific.

Table 6.2 Based on the I_{on}, I_{eff}, I_{off}, $G_m \times R_{out}$ and ring oscillator speed, the application-specific $+1\sigma$ card (#845) is selected.

Card #	I_{on} (A/μm)	I_{eff} (A/μm)	I_{off} (A/μm)	$G_m \times R_{out}$	Ring Oscillator Speed
$+1\sigma$ target	1.58E-3	9.02E-4	9.59E-7	3.63 (-1σ)	238.7GHz
835	-0.2%	-0.4%	0%	0.12%	-3.4%
840	0%	-0.13%	-21.6%	-2.2%	-5.3%
844	0.14%	0.19%	-6.5%	-0.26%	10.3%
845	0.15%	0.57%	6.8%	2.2%	1.7%

6.3 Results and Discussion

After so selecting the 5 PAM cards, Fig. 6.6 shows the PAM prediction of 55-stage ring-oscillator (RO) speed compared to the 1000 Monte Carlo simulations. The selected 5 PAM cards predict the RO speed very well. The speed of all logic gates (Inverter, NAND, NOR, etc.) are highly correlated with each other [6.9]. One can use selected PAM cards to predict the performance of the more complex circuits. In Fig. 6.7, the same PAM cards give good prediction relative to 1000 MC simulations of the speed of a 4-bit adder, which was not considered in the PAM generation process. It indicates that the PAM cards selected based on simple circuit or transistor behave statistically similarly when applied to larger scale designed circuits. PAM is as easy to use as the traditional corner models but is more rational and accurate. The accuracy improvement arises from two fronts: i) deliberate inclusion of more electrical variation data such as I_{on}, I_{eff}, I_{off}, $G_m{\times}R_{out}$, speed of circuit fabrics and ii) a methodology to reconcile the many physical, electrical, and RDF variations.
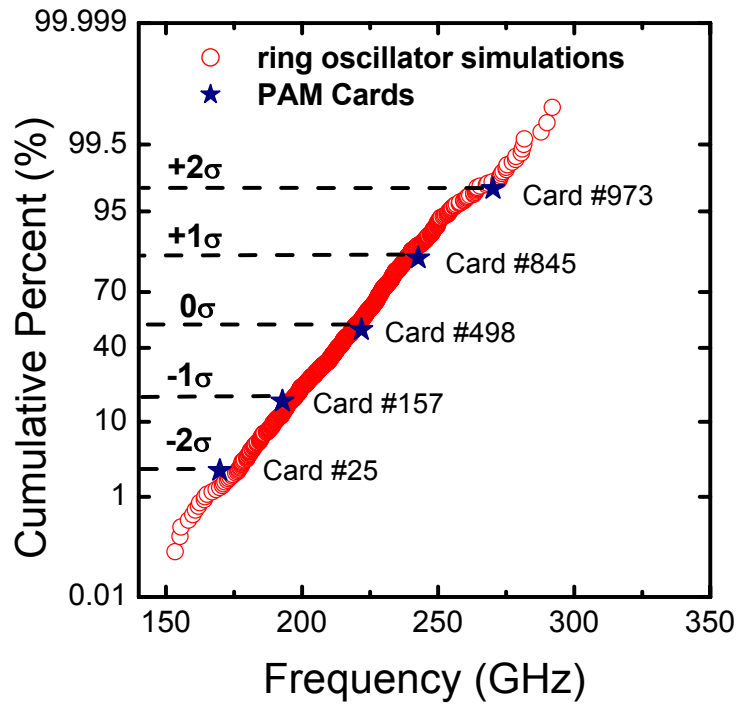
Fig. 6.6 The speed distribution of 55-stage ring-oscillator (1000 simulations). The improved corner model cards can predict the speed very well.
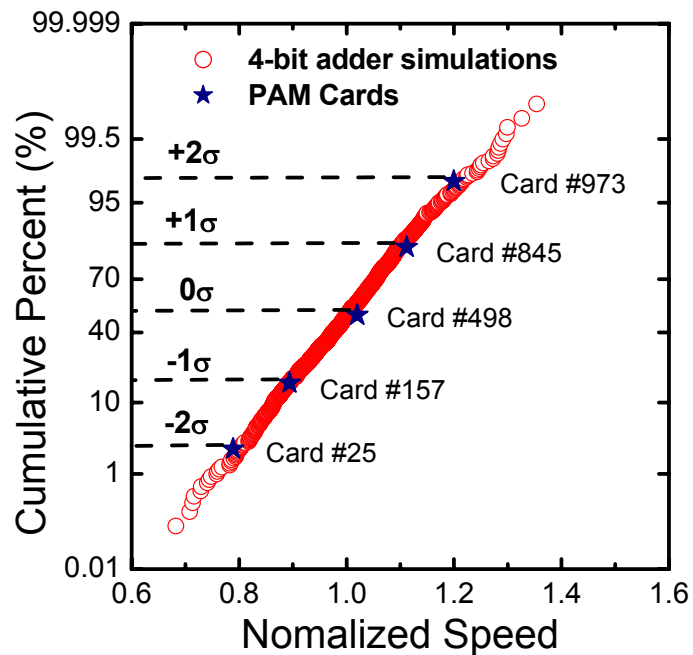

Fig. 6.7 The speed distribution of 4-bit adder (1000 simulations). The improved corner model cards calibrated from E-T data can predict the speed of larger scale circuit very well.

A reduced set of variances that are part of the PAM cards defines the technology for SPICE MC simulations. If the variation inputs are separated between global and local variations [6.10], the PAM cards will support MC simulations for global variation only, local variation only, and total variation. These three types of MC simulations are performed for RO delay study (1000 simulations). In Fig. 6.8, the local variations contribute little to the delay variation because the local variations of individual devices in a long logic chain tend to average themselves out. Fig. 6.9 shows the average power per stage vs. delay per stage of 55-stage RO. The local variation has little impact on the variation due to averaging effect in the long logic path. These type of circuits is served well by the PAM corner model. On the other hand, Fig. 6.10 shows sample butterfly curves of 6-T SRAM cell during read event using the same variation model employed in Fig. 6.8. Fig. 6.11 shows MC result of local-variation only effect on the static noise margin (SNM) of 6-T SRAM. In the SRAM case, the same local variations produce very large SNM variations while the mean of SNM is affected by the global variation. Both global and local variations are critical for the SRAM yield. New variation mechanisms such as MC flicker noise will be modeled and added to the BSIM device variation model.
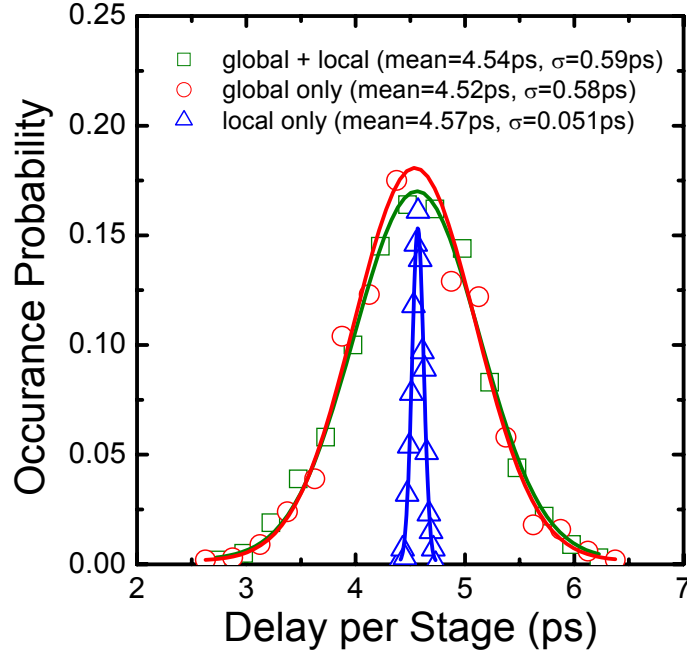
Fig. 6.8 The Monte Carlo simulation of delay distribution of 55-stage RO (1000 simulations). The separation of global and local model card is needed for improving the accuracy of Monte Carlo simulation.
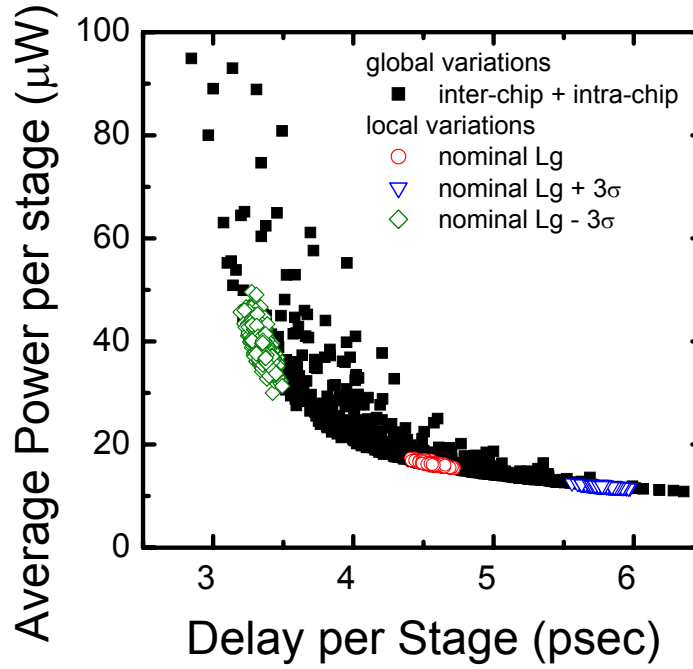

Fig. 6.9 The average power vs. delay metric of 55-stage RO (1000 simulations). The separation of global and local model card is needed for improving the accuracy of Monte Carlo simulation.
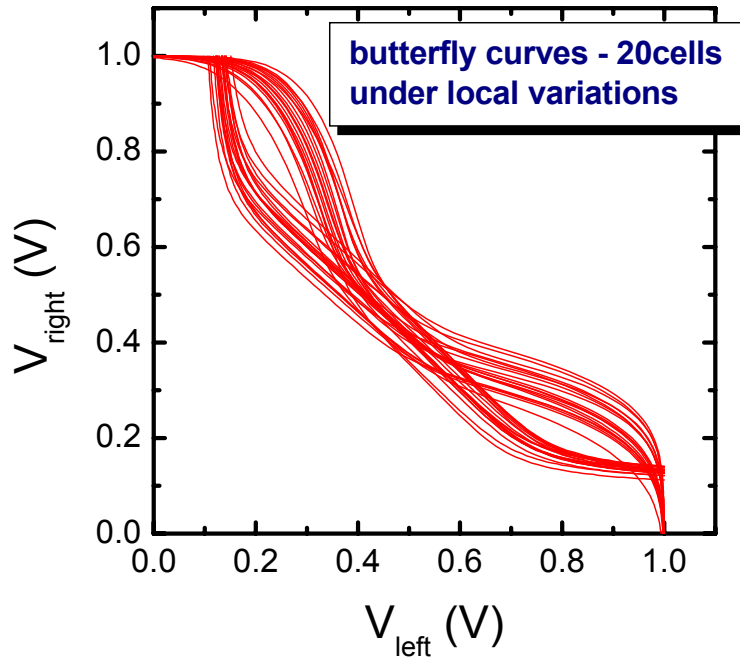
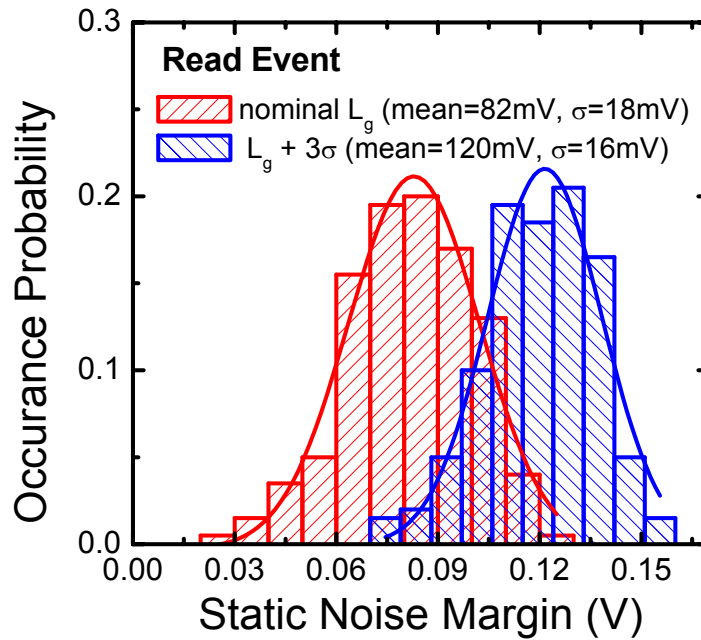Fig. 6.10 The sample butterfly curves of 6T-SRAM during the read event.

Fig. 6.11 The SNM distribution of 6T-SRAM under different global variation conditions. The sigma of SNM is only affected by the local variation while the mean of SNM is affected by the global variation.

6.4 Summary

We present a methodology to generate Performance-Aware corner Models--PAM.

PAM is easy to use as the traditional corner models but is more rational and accurate.

Accuracy is improved by emphasizing electrical variation data and reconciling the

process and electrical variation data. PAM supports corner simulation and MC simulation.

PAM can predict the performance of the more complex circuits. Furthermore, PAM

supports application-specific corner cards, for example, for gain sensitive applications.

6.5 References

[6.1]   J. Power, B. Donnellan, A. Mathewson, and W. Lane, "Relating Statistical

MOSFET Model Parameter Variabilities to IC Manufacturing Process

Fluctuations Enabling Realistic Worst Case Design," IEEE Trans. Semiconduct.

Manufact., vol. 7, no. 3, pp. 306-318, 1994.

[6.2]   J. Chen, M. Orshansky, C. Hu, and C.-P. Wan, "Statistical Circuit

Characterization for Deep-Submicron CMOS Designs," in Proc. ISSCC, 1998,

pp. 90-91.

[6.3]   J. Power et al., "An approach for relating model parameter variabilities to process fluctuations," in Proc. ICTMS, 1993, p. 63.

[6.4]   M. Orshansky, J. Chen, and C. Hu, "Direct Sampling Methodology for Statistical Analysis of Scaled CMOS Technologies," IEEE Trans. Semiconduct. Manufact., Vol. 7, no. 3, pp. 306-318, 1994.

[6.5]   J. Chen, C. Hu, D. Wan, P. Bendix, and A. Kapoor, "E-T base statistical modeling and compact statistical circuit simulation methodologies," in Proc. IEDM, 1996, pp. 635-638.

[6.6]   BSIM4 Manual, http://www-device.eecs.berkeley.edu/~bsim3/bsim4.html

[6.7]   W. Zhao, and Y. Cao, "New Generation of Predictive Technology Model for sub-45nm Design Exploration," in Proc. ISQED 2006, pp. 717-722.

[6.8]   A. Asenov, A. Brown, J. Davies, S. Kaya, and G. Slavcheva, "Simulation of Intrinsic Parameter Fluctuations in Decananometer and Nanometer-Scale MOSFETs," IEEE, Trans. Electron Devices, vol. 50, no. 9, pp. 1837-1852, 2003.

[6.9]   J. C. Chen, C. Hu, Z. Liu, and P. K. Ko, "Realistic worst-case SPICE file extraction using BSIM3," in Proc. Custom Integrated Circuits Conf., 1995, p. 375.

[6.10] M. Kanno, A. Shibuya, M. Matsumura, K. Tamura, H. Tsuno, S. Mori, Y. Fukuzaki, T. Gocho, H. Ansai, and N. Nagashima, "Empirical Characteristics and Extraction of Overall Variations for 65-nm MOSFETs and Beyond," in Proc. VLSI Tech., 2007, pp. 88-89.

Chapter 7

Conclusions

7.1 Summary of Contributions

The unique device physics in the multi-gate MOSFET are comprehensively studied. Modeling methodologies/approaches are proposed to incorporate these unique multi-gate physics in the compact model. An accurate QM correction for multi-gate MOSFET is proposed to simultaneously capture both V_{th} shift and gate capacitance degradation for the first time. The model is accurate over a wide range of device parameters. The SCE

model shows excellent agreements with 2-D TCAD simulation results without the use of any fitting parameters. Good scalability over T_{ox} and T_{si} down to 30nm channel length (L_g) is clearly visible. The SCE model is extended for modeling the triple or more gates structures by making $\lambda = f(T_{ox}, T_{si}, H_{fin})$. The SCE model implementation captures V_{th} roll-off, DIBL and subthreshold slope degradation for short channel multi-gate FETs simultaneously. A cap transistor model is introduced to model the corner portion of the multi-gate MOSFET. The height of cap transistor is half of the fin width independent of fin height, which can be explained by the charge sharing concept.

A full scale compact model for multi-gate MOSFET, BSIM-MG is developed based on surface potential. The core model agrees with TCAD simulation results very well without using any fitting parameters. It demonstrates the inherent physical predictivity and scalability of the model. BSIM-CMG model is experimentally verified against both SOI FinFET and bulk FinFET technologies. BSIM-CMG was able to describe the drain current and its derivatives for long and short channel FETs for both technologies. BSIM-IMG is verified against TCAD simulation results.

A planar double-gate SOI nMOSFET is simulated with TCAD and fitted using BSIM-IMG. The impact of back-gate length and misalignment on V_{th} is discussed. Misalignment effect is more serious when the back-gate is misplaced towards the drain

end.

Several important design concepts of MG-based circuits are addressed. For a given amount of power, the CMG achieves a lower stage delay than the GP due to the better subthreshold behavior for a fixed off-state current. The CMG is more efficient than the GP in the application of logic circuit. However, the dynamic V_{th} tuning of GP is attractive to simultaneously achieve high performance during active periods and low leakage power during idle periods for ultra-low power circuit design. FinFET-based logic and memory circuits for ultra-low voltage applications are analyzed. It clearly shows that FinFET is an ideal candidate for ultra-low voltage applications because of its robustness to various device parameter variations. A reliable SNM of FinFET SRAM can be maintained at ultra-low voltages under V_{th} variation caused by random dopant and T_{si} fluctuations.

We present a methodology to generate Performance-Aware corner Models--PAM. PAM is easy to use as the traditional corner models but is more rational and accurate. Accuracy is improved by emphasizing electrical variation data and reconciling the process and electrical variation data. PAM supports corner simulation and MC simulation. PAM can predict the performance of the more complex circuits. Furthermore, PAM supports application-specific corner cards, for example, for gain sensitive applications.

7.2 Suggestions for Future Work

Several important features of scaled multi-gate MOSFET are not included in the model yet, such as parasitic capacitance, and prasitic resistance. Currently, gate tunneling current, series resistance, substrate current and ballistic transport, are modeled using the approach in the conventional bulk MOSFET compact model. Detailed TCAD studies or data analysis are needed for identifying the physical difference of physics between multi-gate MOSFET and conventional bulk MOSFET. BSIM-MG can be improved by including these effects.

The BSIM-IMG assumes the single sided conduction which will introduce significant error when the back surface is biased to conduct current. The accuracy and yield of certain logic building blocks and memory cells using independent multi-gate structure will be improved if the model can predict the both-channel conduction well. The modeling of the impact of misalignment at the back-gate electrode is important since the back-gate tuning is a very attractive scheme for reducing the process-induced variations.

The BSIM-MG model is verified with single transistor experimental data only. It will be very important and attractive to verify the model against the characteristics of various logic gates or memory cell, such as delay, speed, and noise margins.