國立交通大學

電子工程學系電子研究所碩士班

碩士論文

奈米元件靜態隨機讀取記憶體特性之研究

Performance of SRAM with Nanoscale Transistors

研 究 生：呂建松

指導教授：黃調元 博士

林鴻志 博士

李義明 博士

中華民國九十五年七月


E
S
A
1896

奈米元件靜態隨機讀取記憶體特性之研究
Performance of SRAM with Nanoscale Transistors

研究生：呂建松　　Student：Chien-Sung Lu

指導教授：黃調元　博士　　Advisor：Dr. Tiao-Yuan Huang

林鴻志　博士　　Advisor：Dr. Horng-Chih Lin

李義明　博士　　Advisor：Dr. Yiming Li

國立交通大學
電子工程學系電子研究所碩士班
碩士論文

A Thesis
Submitted to Institute of Electronics Engineering
College of Electronical Engineering National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master
in

Electronics Engineering

July 2006

Hsinchu, Taiwan

中華民國九十五年七月

E
S
A
1896

E
S
A
1896

奈米元件靜態隨機讀取記憶體特性之研究

學生：呂建松　　　　　　　　　　指導教授：黃調元 博士
林鴻志 博士
李義明 博士

國立交通大學 電子工程 學系（研究所）碩士班

摘　　要

平面結構的金屬氧化物半導體場效電晶體目前是靜態記憶體的最常見的組成架構。不過，當設計規章向45奈米以下持續縮小時，傳統的平面的電晶體將遇到很多顯著的挑戰。立體(三維)架構的電晶體已經製造出來，並且證明有更好的電特性相較於傳統平面電晶體。對於靜態隨機讀取記憶體設計而言， 操作穩定度和記憶體所佔面積都是必須考慮的。靜態隨機讀取記憶體面積大約佔整個晶片區域的 3 分之 2，穩定度則是當製程和操作條件發生擾動時靜態隨機讀取記憶體特性的敏感性。為了增進靜態隨機讀取記憶體穩定度必須加大其面積，這兩者參數的關係是互相牽制的。

本文將探討使用三種不同結構的奈米級金屬氧化物半導體場效電晶體架構下靜態隨機讀取記憶體操作性能，三種結構分別是平面金屬氧化物半導體場效電晶體、絕緣層上矽鰭鱗式場效電晶體和全閘鰭鱗式場效電晶體，包含記憶體操作的穩定度和特性的敏感性分析將被討論。靜態雜訊邊際(SNM)參數將在考慮到量子效應下使用三維混合模式(mixed-mode)元件模擬軟體下加以萃取。首先，我們探討不同結構金屬氧化物半導體場效電

晶體的本質特性和端點特性。此外，靜態雜訊邊際對於不同供應電壓，電晶體尺寸比率(cell ratio)和操作溫度的變化將被分析和比較。

當元件縮小到 100 奈米以下時，由於電晶體濃度摻雜，晶面粗糙和通道長度變動等所引起電晶體特性的擾動將開始影響到電路的特性和功能。不同的元件架構下的靜態隨機讀取記憶體穩定度的敏感性分析將被藉由一個有系統的統計方法被發展加以分析。藉由實驗設計，混合模式(連接的元件和電路)模擬技術，以及二次反應曲面模型，我們將專注對由於通道長度變化所產生的靜態雜訊邊際擾動加以探討。

本研究提供一個方法探討不同元件結構所架構的靜態隨機讀取記憶體的特性。在尚未有元件完整模型存在情況下，我們可以使用三維混合模式元件模擬對新的元件結構所架構的靜態隨機讀取記憶體的穩定度加以研究。對於電路特性擾動的探討，可藉由實驗設計和反應曲面模型的建立達到。總之，本文提供一個系統化的統計方法去探討奈米級金屬氧化物半導體場效電晶體架構下靜態隨機讀取記憶體操作性能。在時間、成本、效率的考量上，此方法顯得很有經濟效益。

Performance of SRAM with Nanoscale Transistors

Student：Chien-Sung Lu
Advisors：Dr. Tiao-Yuan Huang
Dr. Horng-Chih Lin
Dr. Yiming Li

Department of Electronics Engineering
National Chiao Tung University

ABSTRACT

Silicon-based planar MOSFETs have been the building block for SRAM. However, as the design rule continuously shrank down beyond the 45 nm, conventional planar CMOS devices encounter significant challenges. Many three-dimensional (3D) structure transistors, such as the bulk fin field-effect transistor (Bulk FinFET), SOI fin field-effect transistor (SOI FinFET), multiple-gate FinFET, and surrounding-gate nanowire FinFET (Nanowire FinFET) have been proposed, fabricated, and demonstrated more attractive electrical characteristics than that of single-gate planar devices. Two aspects are important for SRAM cell design: the cell area and the stability of the cell. The cell area determines about two-third of the total chip area. The cell stability determines the soft-error rate and the sensitivity of the memory to process tolerance and operating conditions. The two aspects are interdependent since designing a cell for improved stability invariably requires a larger cell area.

In this thesis, we study the performance of 6-T SRAM cell with three different building 32 nm devices, planar MOSFETs, SOI FinFETs, and nanowire FinFETs. The stability and sensitivity analysis will be discussed. Static noise margin (SNM) of SRAM is computational investigated and compared by using a mixed-model three-dimensional device simulation with considering quantum mechanical effects. We firstly analyze and compare the intrinsic and terminal characteristics for the three different transistors in SRAM cells. Also, the SNM of SRAM during both hold and read modes is explored for the device with respect to different supply voltage , cell ratio, and operation temperature.

With the scaling of conventional CMOS devices to sub-100 nm and beyond, the variations

of the transistor characteristics due to local and non-local effects, such as the random dopants, the critical dimension of channel length, the interface roughness, and line edge roughness (LER) start to adversely affect the yield and functionality of the corresponding circuits. In this thesis, a systematical method for sensitivity analysis of SRAM cells with different device structures is developed. Based on a design of experiment (DOE), a mixed-mode (i.e., coupled device and circuit) simulation, and a response surface model (RSM), performances of 6T SRAM cells are explored with respect to static noise margin (SNM). Taking the channel length of different transistors in a 6-T SRAM cell as significant variables, a SNM response surface model is constructed. With the developed SNM model, the impact of channel length variation on SNM is evaluated.

Finally, the purpose of this study is to provide a systematically statistical method to analysis the performance of the SRAM cell using nano-scale device structures. The stability of SRAM cells is explored by a 3D mixed-mode simulation. The sensitivity analysis of SNM will be studied by the combination of design of experiment and second-order response surface model. We believe that he design of 6-T SRAM cell with nanowire FinFETs is a promising approach in sub-45 nm CMOS devices era.

誌謝

這份論文能夠順利完成，首先感謝我的指導老師 黃調元老師和林鴻志老師給予我最大自由度，讓我可以繼續完成感興趣的研究，並在課業以及生活上不吝惜的指導及支持。其次，我要感謝我另一共同指導老師 李義明老師，感謝老師指導我論文方向脈絡，以及研究能力激發有著深厚的影響。

論文口試期間，承蒙交大電子工程研究所陳明哲教授、以及交大電信研究所李育民教授撥冗細審，並會予寶貴意見與殷切指正使本論文更臻完備。

研究室方面，我要感謝余紹銘學長、周宏穆學長、陳璞學長、王傳盛學長、陳正凱學長、卓彥羽學長的照顧幫忙，感謝蔡孟家學弟幫我寫萃取參數的程式。感謝同窗好友的朝夕相伴，煒昕、伯賢、景嵐、宏榮及婉文的互相砥礪，在此一併致謝。

電子所方面，感謝凱翔、志誠、雨霖、振家、健銘、銑泓及行徽，不僅在課業方面的知識上會給予幫助，也在我學習的路上給予許多的建議。

最後最應該感謝的是我的媽媽，和我的弟弟以及妹妹。這期間即便我承受在大的壓力、不開心、不耐煩、及挫折，還是一直在鼓勵我、幫我調適心情。現在終於完成論文畢業了，真的很感謝你們。

感謝這段期間大家對我的包容、關懷與愛護。這篇論文，這個工作，以及我在交大的一切成長，沒有你們，是完全沒有辦法達成的，一切的功勞都歸因於全部的人。謝謝大家一直挺我鼓勵我，使我順順利利的度過這段非凡且精采的日子。在此將這篇論文獻給所有關心我以及我所愛的人，謝謝你們。

本論文感謝行政院國家科學委員會(計畫編號 NSC-93-2115-E-492-008、NSC-94-2115-E-009-084)、卓越沿續計畫(計畫編號 NSC-94-2752-E-009-003-PAE、NSC-95-2752-E-009-003-PAE)、五年五百億計畫、經濟部科專計劃(計畫編號 93-EC-17-A-07-S1-0011)以及台灣積體電路製造股份有限公司之資助。

呂建松 謹誌

中華民國九十五年七月 于風城交大

Contents

7 The Static Noise Margin of SRAM Cells 108

Appendix A

Appendix B

Appendix C

List of Figures

List of Tables

Chapter 1

Introduction

This chapter is organized as follows. First of all, the background of this work which includes the semiconductor memory and the literature review is introduced. The evolution of nanodevice structures, design Considerations of the 6-T SRAM Cell, and the circuit performance sensitivity due to device parameter fluctuations will be discussed in literature review. Then we introduce the motivation. We want to know the performance of the 6-T SRAM cell using three planar MOSFETs, SOI FinFETs, and nanowire FinFETs. In the third part, the objectives of this study will be presented. Finally, the outline in this thesis is described.

1.1 Background

In this section, we will introduce the background of this thesis. It includes the semiconductor memory classification and the literature review. Semiconductor memories can be classified on the basis of memory functionality, access patterns, ant the nature of the storage mechanism. In this thesis, we focus on the analysis of static random access memories. The overview of device structure scaling and SRAM design considerations are presented in the literature review.

1.1.1 Semiconductor Memory Classification

Semiconductor memories can be classified on the basis of memory functionality, access patterns, ant the nature of the storage mechanism. A distinction is made between read-only (ROM) and read-write (RWM) memories. The RWM structures have the advantage of offering both read and write functionality with comparable access times and are the most flexible memories. Data is stored either in flip-flops or as a charge on a capacitor. As in the classification introduced on sequential circuitry, these memory cells are called static and dynamic respectively. The former retain their data as long as the supply voltage retained, while the latter need periodic refreshing to compensate for the charge loss caused by leakage. Since RWM memories use active circuitry to store the information, they belong to the class celled volatile memories, where the data is lost when the supply voltage is turned off. Currently, DRAM and SRAM are extensively used in personal computers and work stations, mainly because of DRAM's attributes of high density and low cost, and SRAM's attribute of high speed.

Read-only memories, on the other hand, encode the information into the circuit topology - for example, by adding or removing diodes or transistors. Since this topology is hardwired, the data cannot be modified; it can only read. Furthermore, ROM structures belong to the class of the nonvolatile memories. Disconnection of the supply voltage does not in a loss of the stored data.

The most recent entry in the field are memory modules that can be classified as nonvolatile, yet offer both read and write functionality. Typically, their write operation takes substantially more time than the read. We call them nonvolatile read-write (NVRWM) memories. Members of this family are the EPROM (erasable programmable read-only memory), EEPROM (electrically erasable programmable read-only memory), and flash memories. The nonvolatile memory is used extensively in portable electronics systems such as the cellular phone, digital camera, and smart IC cards, mainly because of its attributes of low-power consumption and non-volatility.

1.1.2 Literature Review

Silicon-based conventional planar MOSFETs have been the building block for SRAM. However, as the design rule continuously shrank down beyond the 45 nm, conventional CMOS devices encounter significant challenges [1] - [2]. Double-gate (DG) undoped-channel quasi-planar CMOS technology (FinFET [3], Tri-gate [4], Omega-FET [5], Nanowire FinFET [6] etc.) has emerged as the leading candidate to replace Bulk and Partially-depleted SOI for scaling to the end of the roadmap [7]. Double-gate operation provides superior short-channel control, ideal sub-threshold slope and higher drive current compared to single-gate operation. The thin undoped (i.e. lightly-doped with $N_A \leq 10^{16}\ cm^{-3}$) body improves mobility and eliminates random dopant fluctuation effects. Quasi-planarity implies that while the device is not planar in the conventional sense, current flow is parallel to the wafer plane. It allows increased effective channel area from a given planar area by increasing height. This property is especially suitable in SRAM where load capacitances are interconnect-dominated [8].

For SRAM cell design, the cell area and the stability of the cell are the most important factors. The cell area are expected to contribute the largest fraction of chip device count (close to 70 %) in future ICs. The cell stability determines the soft-error rate and the sensitivity of the memory to process tolerance and operating conditions. These two aspects are interdependent since designing a cell for improved stability invariably requires a larger cell area. There are different methods to simulate and analyze the SRAM cell stability. One way uses the direct 3D process and device simulations [9]. The SRAM cell simulation structure is constructed with combination of 3D device simulator. It also constructs the via and wire structure. The most general way is explored the stability of SRAM cells by SPICE circuit simulation [10] - [12]. It used the device compact model to simulate the circuit performance. The mixed-mode device simulation is another way to analyze the SRAM cell stability [13] - [14]. The drift-diffuse model for carrier transport and the density-gradient model to account for quantum-mechanical effects in nanoscale MOSFETs is employed to simulate DC transfer characteristics of SRAM cells.

With the scaling of conventional CMOS devices to sub-100 nm and beyond, the variations of the transistor characteristics due to local and nonlocal effects, such as the random dopants, the critical dimension of channel length, the interface roughness, and line edge roughness (LER) [15] - [16] start to adversely affect the yield and functionality of the corresponding circuits. [17] - [18] These atomic-level intrinsic fluctuations cannot be eliminated by external control of the manufacturing process and are most pronounced in minimum-geometry transistors commonly used in area-constrained circuits such as SRAM cells [19]. Furthermore, intrinsic fluctuations are independent of transistor location on a chip. The threshold voltage mismatch between neighboring cell transistors due to intrinsic fluctuations typically contributes to larger reductions in static noise margin (SNM) than the threshold voltage mismatch due to macroscopic manufacturing-related variations in scale CMOS SRAM cells [19]. Due to the scaling limitations of the conventional MOSFETs, novel devices architectures such as double-gate MOSFETs, which are more resistant to some of sources of intrinsic parameter fluctuations, are expected to play an increasing important role beyond the 45nm technology node. There are various methods to explore the SRAM cell sensitivity induced by intrinsic parameter fluctuations. The tradition statistical way is the Monte Carlo simulation analysis [20]. It uses statistical transistor model including process and mismatch fluctuations. It can create unrealistic device behavior and may not capture all aspects of the impact of intrinsic parameter fluctuation sources on the device behavior. The better way is a robust methodology to incorporate intrinsic parameter fluctuation information into device compact model [21] - [22]. It can consider all impacts of intrinsic parameter fluctuation sources on the device performance. Another way is also a statistical method, it constructs the response surface model [23]. Using the suitable distribution, we can perform the sensitivity of SRAM cells.

1.2 Motivation

Conventional CMOS devices encounter significant challenges when the design rule continuously shrank down beyond the 45 nm. Double-gate (DG) undoped-channel quasi-planar CMOS technology has recently been of great interest and relative present better scale-down properties than than of conventional bulk mental-oxide-semiconductor field-effect transistors (MOSFETs). Besides, double-gate device is more immune to the source of intrinsic parameter fluctuations than conventional MOSFETs. However, the quasi-planar CMOS devices have attractive characteristics than conventional planar MOSFET at device level, we don't know how the performance changes at circuit level. As CMOS technology is dramatically scaled down in recent years, the operation of SRAM becomes critical issue for further scaling. Keeping enough SNM and avoiding soft errors are the two key factors in reliability. We study the performance (including the stability and sensitivity of SNM) of the 6-T SRAM cell with different device structures. Firstly, the device characteristics by using a device simulation are analyzed and compared. We verify that the the quasi-planar CMOS devices have less short channel effects than the conventional planar MOSFET. Then,we explored stability of 6-T SRAM cells with conventional planar MOSFETs and double-gate quasi-planar MOSFETS by using a mixed-mode simulation. Finally, we provide a statistical methodology to explore the sensitivity of static noise margin for 6-T SRAM cells with different device structures.

1.3 Objectives

In this thesis, a systematical method for the performance of the conventional 6-T SRAM cells with different device structures (including planar MOSFETs, SOI FinFETs, and nanowire FinFETs) is investigated. The performance consists of the static noise margin (SNM) analysis and the sensitivity analysis of SNM for SRAM cells. First, we explore intrinsic characteristics, terminal electrical characteristics and short channel effect parameters for each device structures under different technology generation nodes by using a three-dimension device simulation. The intrinsic characteristic consists of the electric potential, electrical field, and carrier density under on and off state. The terminal characteristic is curves of Id-Vd and Id-Vg. We extract short channel effect parameters which consist of drain induced barrier lowering (DIBL), subthreshold swing (S.S.), on/off current ratio, threshold voltage. Static noise margin (SNM) of 6-T SRAM is also computational investigated and compared by using a mixed-model three-dimensional device (i.e. coupled device and circuit) simulation with considering quantum mechanical effects. The SNM of the 6-T SRAM cell with 32nm different device structures during both hold and read modes is explored with respect to the supply voltage, cell ratio, and temperature. Finally, the sensitivity of the SNM for the 6-T SRAM cell with 32nm CMOS device is explored. Based on the screening design, a design of experiment (DOE), a mixed-mode simulation, and a response surface model

(RSM), the performance of 6-T SRAM cells are explored with respect to static noise margin. Taking channel length of each transistor with 6-T SRAM cells as significant variables, a SNM response surface model is constructed. The sensitivity of SNM then can be explored in a computational cost-effective way.

1.4 Outline of the Thesis

This thesis is organized as follows. In Chapter 2, nanodevices and the static random access memory cell will be introduced. The developed simulation methodology will be discussed in Chapter 3. Physical models and the mixed-mode simulation will state in Chapter 4. Detail about the computation statistical technique is in Chapter 5. The results of the device electrical characteristics of the adopted devices will be presented in Chapter 6. The static noise margin of SRAM cells will be shown in Chapter 7. A sensitivity analysis of static noise margin will be discussed in Chapter 8. Finally, we draw some conclusions and suggest future works in Chapter 9.

Chapter 2

Nanodevices and Static Random Access Memory Cell

In this chapter, we firstly introduce the evolution of nanodevice structures in section 2.1. Architectures of SRAM cells will be shown in section 2.2. Operations of 6-T SRAM cell is described in more detail in section 2.3. We introduce the standby (hold), read and write modes of memory cell. The design considerations of 6-T SRAM cell will be discussed in section 2.4. We will focus on stability of 6-T SRAM cell during holding data and reading access and sensitivity of static noise margin. In section 2.5, we will show the effect of device parameters on SNM.

2.1 Structures of Nanodevices

A single-gate planar MOSFET device built on a bulk Si substrate ha s been the workhorse for the semiconductor industry in the past four decades. A degenerately doped poly-Si is usually employed as the gate electrode, and a subsequent implant step is performed to form the source/drain in a self-aligned manner. Success of this structure is ascribed to the facts that the overall fabrication is relatively simple and the self-aligned is very suitable for device scaling. Nevertheless, the planar device will suffer from the short-channel effects(SCEs), such as the increased subthreshold and substrate leakages as the channel length is shortened. To overcome the SCE, ultra-shallow source/drain junctions as well as optimized two-dimension channel doping profile are needed [25]. This will complicate the device fabrication, especially when we enter the nanoelectronic era. Actually, the reported performance of MOSFETs [26] - [29] with channel length around 20 nm are substantially lower than the requirements defined in the ITRS 2001 [24]. Therefore, we must improve the device performance to meet the application requirements. Otherwise, alternative device structures must be considered for the technology node of 65 nm.

One alternate device structure is the single-gate MOSFET built on a silicon-on-insulator (SOI) substrate. Depending on the thickness of the silicon layer, we have the partially-depleted (PD), the fully-depleted (FD), and the ultra-thin-body (UTB) SOI structures. The

SOI devices can suppress the substrate leakage current encountered in conventional devices, because the buried oxide blocks the leakage path in the substrate. The UTB SOI structure (with Si thickness $<$ 20 nm) can further reduce the substrate leakage, thus represents a promising candidate for nano-scale MOSFET applications. According to experimental observations [28][30], the thickness of the silicon layer should be thinner than one-third of the channel length in order to effectively control the SCE. This imposes a sever constraint on the development of UTB-SOI due to the thickness fluctuation of the silicon layer and quantum confinement effects [31].

Another approach is the double-gate (DG) SOI structure [34]. This structure cab be scaled to smaller dimension than UTB SOI devices, because the electric field originating from the drain is effectively terminated by the gates. Moreover, the gate configuration relaxes the requirement of minimum channel thickness to two-third of the channel length, which is two times thicker than that for SG UTB SOI [32]. This means that the impacts of the body thickness fluctuation and quantum confinement effects on device performance are relaxed as well. In addition, the gate leakage us also reduced in the DG structure [33].

Two parallel channels perpendicular to the wafer surface has been proposed and demonstrated a decade ago [35]. It is now more generally dubbed ”FinFET” [32], since the channel island is fin-like. Multiple fins can be implemented to increase the current drive. It should be noted that, the structure is ”quasi-planar” if the height of fin channel is limited to 0.1 μm or less [36]. The FinFET has an important implication for devices scaling because modern processing equipment have the capability to fabricate such devices. The device characteristic of p- and n-channel FinFETs with a channel length of 10 nm has been reported [37]. The excellent results obtained demonstrate the feasibility of FinFET in nano-scale manufacturing.

Intel introduced a tri-gate (TG) FinFET technology [38]. Basically, the structure is similar to the DG FinFET except that there is no hard-mask on top of the fin channel. As a result the top surface of the fin channel also severs as a current conduction path during the on-state operation. Such scheme can further relax the constraint imposed on the channel thickness as compared to the DG FET, and reduce the fin height for better fabrication control.

A related structure is the surrounding gate (nanowire FinFET) scheme [39], which has the channel fully surrounded by the gate. Theoretically, it is the most appropriate configuration for nano-scale MOSFETs, it is unlikely to be implemented in practical manufacturing due to the complexity of the fabrication process.

2.2 Architectures of SRAM Cells

There are basically two types of semiconductor RAMs : static and dynamic. Static RAMs (called SRAMs for short) utilize static latches as the storage cells. Dynamic RAMs (called DRAMs), on the other hand, store the binary data on capacitors, resulting in further reduction cell area, but at the expense of more complex read and write circuitry. In particular, while static RAMs can hold their stored data indefinitely, provided the power supply remains on, dynamic RAMs require periodic refreshing to regenerate the data stored on capacitors. This is because the storage capacitors will discharge, though slowly, as a result of the leakage currents inevitably present. By virtue of their smaller cell size, dynamic memory chips are usually four times as dense as their contemporary static chips. Both static and dynamic RAMs are volatile; that is, they require the continuous presence of a power supply. By contrast, most ROMs are of the nonvolatile type.

Fig. 2.1 shows a typical static memory cell in CMOS technology. The circuit is a flip-flop comprising two cross-coupled inverters and two access transistors, M2 and M5. The flip-flop consists of two load elements (M3,M6) called pull-up(load) transistors and two storage elements (M1,M4) called pull-down(driver) transistors. Data are stored as voltage levels with the two sides of the flip-flop in opposite voltage configurations, that is, node Q is high and node QB is low in one state and node Q is low and node QB is high in the other resulting in two stable states. The access transistors are turned on when the world

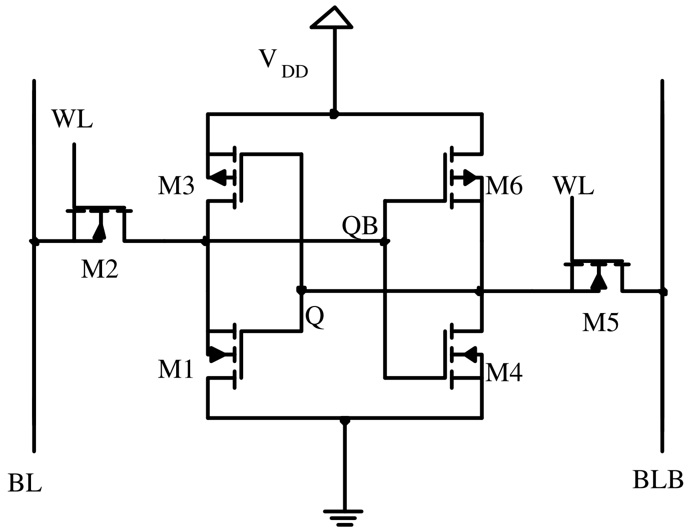


Figure 2.1: A circuit of 6-T SRAM cell used in our mixed-mode simulation. The N- and P-MOSFETs with planar MOSFET, SOI FinFET, and nanowire FinFET structures and adopted in the explored SRAM cell, respectively.

line is selected and its voltage raised to V$_{DD}$, and they connect flip-flop to the column (bit or BL) line and $\overline{\text{column}}$ ($\overline{\text{bit}}$ or BLB) line. Not that both the BL and BLB lines are utilized. The access transistors act as transmission gates allowing bidirectional current flow between the flip-flop and the BL and BLB lines. The ratio of the transconductance factors of the driver and access transistors (r $= \beta_{\text{driver}}/\beta_{\text{access}}$) is an important cell parameter called the ”cell ratio” r. It determines the cell size as well as the cell stability. The ratio of the transconductance factors of the load and access transistors referred to as q $= \beta_{\text{load}}/\beta_{\text{access}}$.

2.3 Operations of 6-T SRAM Cell

To understand the operation of a 6-T SRAM cell, we try to get some idea of how the cell works during standby, read and write operations.

2.3.1 Stand-by Operation

The standby operation are also called hold operation.The word line is not asserted, the access transistors M2 and M5 disconnect the cell from the bit lines during holding data. And the two cross coupled inverters formed by M1, M3, M4, and M6 will continue to reinforce each other as long as they are disconnected from the outside world.

2.3.2 Read Operation

Consider a first read operation, and assume that the cell is storage a ”1”. In this case, Q will be high at V$_{DD}$, and QB will be low at 0 V. Before the read operation begins, the BL and BLB lines are precharged to an high voltage, usually V$_{DD}$. (The circuit for precharging will be conjunct with the sense amplifier.) When the word line is selected and M2 and M5 are turned on, we see that current will flow from V$_{DD}$ through M6 and M5 and onto line BLB, charging the capacitance of line BLB, C$_B$. On the other side of the circuit, current will flow from the precharged BL line through M2 and M1 to ground, thus discharging C$_{\overline{B}}$.

From this description, we note that during a read ”1” operation, the voltage across C$_B$

will rise and that $C_{\overline{B}}$ will fall. Thus, a differential voltage $v_{B\overline{B}}$ develops between line BL and line BLB. Usually, only 0.2 V or so is required for the sense amplifier to detect the presence of a ”1” in the cell. Observe that the cell must be designed so that the changes in v_Q and $v_{\overline{Q}}$ are sufficiently small so that the flip-flop would not changes state during readout. The read operation in an SRAM is nondestructive.

2.3.3 Write Operation

We consider write operation. Assume that the cell is originally storage a ”1” ($v_Q = V_{DD}$ and $v_{\overline{Q}} = 0$) and that we wish to write ”0”. To do this, the BL line is lowered to 0V and the BLB line is raised to V_{DD}, and of course the cell is selected by raising the world line to V_{DD}. The node QB is being pulled up toward the threshold voltage $\mathrm{V_{DD}}/2$ and node Q is being pulled down toward $\mathrm{V_{DD}}/2$. Capacitors $\mathrm{C_Q}$ and $C_{\overline{Q}}$ are the parasitic capacitors at nodes Q and QB, respectively. An approximate analysis can be performed on either circuit to determine the time required for toggling to take place. Note that the regenerative feedback that causes the flip-flop to switch will begin when either $\mathrm{C_Q}$ or $C_{\overline{Q}}$ reaches $\mathrm{V_{DD}}/2$. Once this happens, the positive feedback takes over.

We note that this component of write delay is much smaller than the corresponding component in the read operation. This is because in the write operation, only the small capacitance $\mathrm{C_Q}$ needs to be charged (or discharged), whereas in the read operation, we have to charge (or discharge) the much larger capacitances of the BL or BLB lines. In the write operation, the BL and BLB line capacitances are charged (and discharged) relatively fast by the driver circuitry. The end result is that the delay time in the write operation is dominated by the world-line delay.

2.4 Design Considerations of the 6-T SRAM Cell

The functionality and density of a memory array are its most important properties. Functionality is guaranteed for large memory arrays by providing sufficiently large design margins which is often characterized using Static Noise Margin (SNM). The cell stability is based on the ability of the cell to resist accidental overwrites during different operating conditions in the presence of electrical noise and process variations. The factors that influence the cell stability include the device sizing (channel widths and lengths), the supply voltage, and temperature. Although upsizing the transistors increases the noise margins, it increases the cell area and thus lowers the density. We will consider static noise margin during hold and real modes in detail.

With the scaling of devices, the variations of transistor and circuit characteristics are an important issue. We also focus on the variance of SNM due to critical dimension of gate length , random discrete dopants.

2.4.1 The Stability Margin Under the Hold Mode

In standby mode, the PMOS load transistor (M3) must be strong enough to compensate for the sub-threshold and gate leakage current of all the NMOS transistors connected to the storage node QB (Fig. 2.1). The node Q is stored ”0”, and the node QB is stored ”1”. This is becoming more of concern due to dramatic increase in gate leakage and degradation in on/off current ratio (Ion/Ioff) in recent technology nodes. Coupled with the recent trend to decrease the cell supply voltage during standby to reduce static power consumption, this makes it increasingly more difficult to design robust low-power memory arrays.

Generally, hold stability is commonly quantified by the cell static noise margin (SNM) in standby mode [56]. Fig. 2.2 shows the butterfly curve for the 6-T SRAM with FinFET MOSFETs at 0.8 V operation voltage and cell ratio = 2 while holding date (un-accessed). A basic understanding of the SNM is obtained by drawing and mirroring the inverter characteristics and finding the maximum square between them. The SNM of an SRAM cell represents the minimum DC-voltage disturbance necessary to upset the cell state, and can be quantified by the length of the side of the maximum square that can fit inside the butterfly curves formed by the cross-coupled inverters [56].

2.4.2 The Stability Margin Under the Read Mode

During a read operation, Q rises above 0V, to a voltage determined by the resistive voltage divider set up by the access transistor (M5) and the pull-down transistor (M4) between BLB and node Q (Fig. 2.1). The ratio of the width/length of M4 to M5 determines how high Q will rise and is commonly referred to as the cell β-ratio. If Q exceeds the trip point of the inverter formed by M1 and M3, the cell bit will flip during the read operation, causing a read upset.

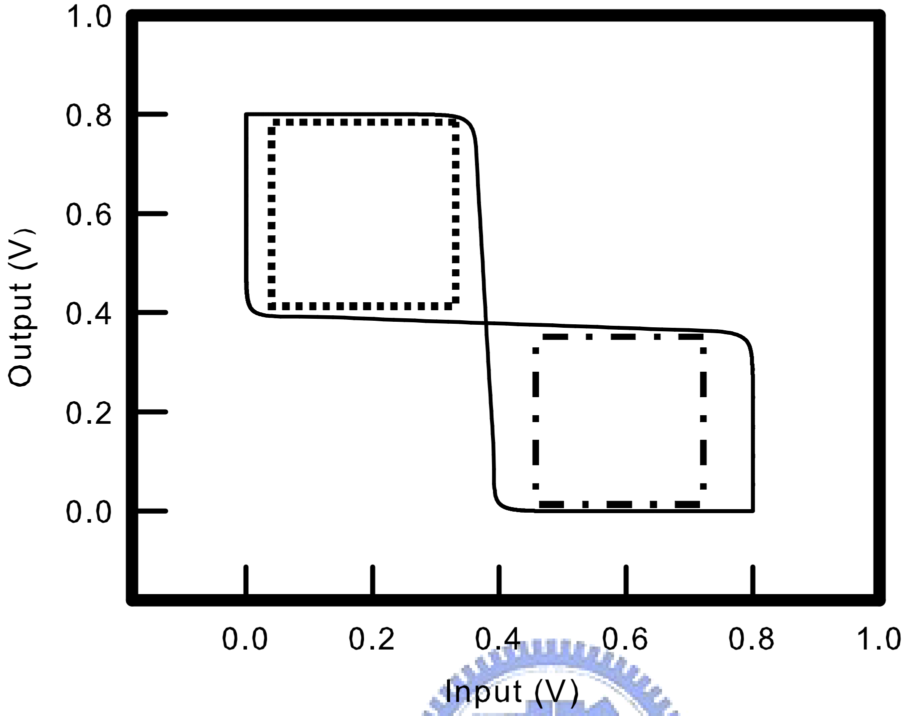


Figure 2.2: The butterfly curve for the 6-T SRAM with FinFET MOSFETs at 0.8 V operation voltage and cell ratio = 2 during hold operation.

Read stability can also be quantified by the cell SNM during a read access. Fig. 2.3 shows the butterfly curve for the 6-T SRAM with FinFET MOSFETs at 0.8 V operation voltage and cell ratio = 2 while reading operation. The cell is most vulnerable to noise during a read access since the voltage divider (M4 and M5) that exist between the BLB and ground causes the voltage at ”0” storage node to rise above ground, so even a small voltage excursion can cause a read upset. The read margin can be increased by upsizing the

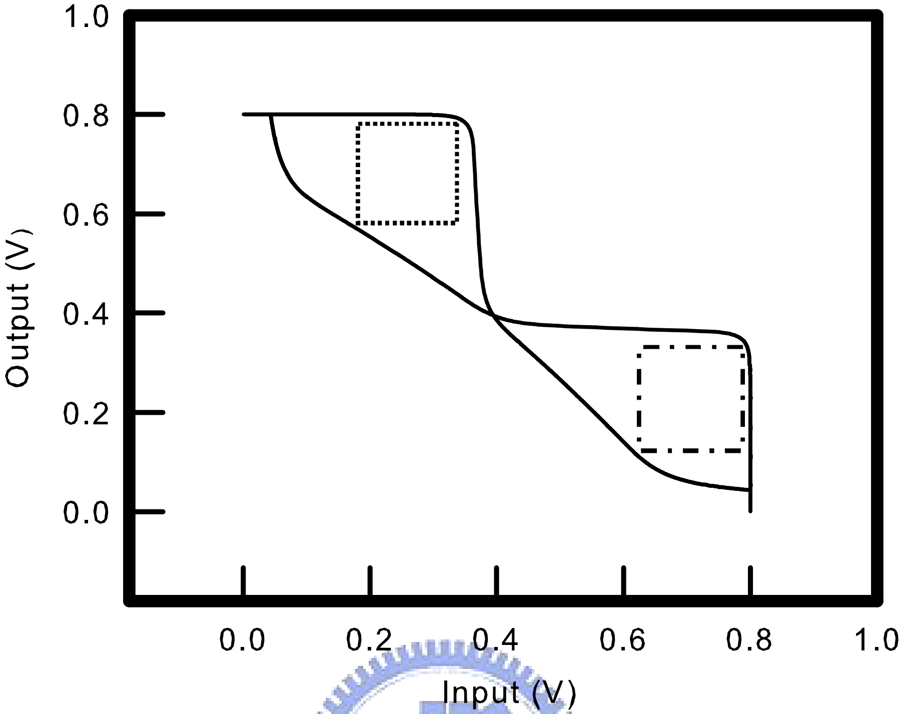


Figure 2.3: The butterfly curve for the 6-T SRAM with FinFET MOSFETs at 0.8 V operation voltage and cell ratio = 2 during read access.

pull-down transistor, which results in an area penalty and increasing the gate length of the access transistor, whice increases the world line (WL) delay and hurts the write margin.

2.4.3 Sensitivity of Static Noise Margin

As scaling continues, process induced gate length variations, random dopant fluctuations are increasingly limiting circuit performance. We analyze the effect of variations by studying their impact on V_{th}, which affects the leakage power, SNM, and performance. The SNM depends on the choice of the V_{th} for the transistors used in the SRAM cells. A high V_{th} means that drive current of these devices is small making the write operation (both accidental and intentional) more difficult, thus increasing the SNM. Then, the SNM is seen to be most sensitivity to threshold voltage fluctuations in the devices. We will study the SNM sensitivity due to the fluctuation of different device parameters. Finally, the SNM sensitivity for the SRAM cell using different device structures due to the channel length fluctuation of pull-down N-MOFETs, access transistors, and pull-up P-MOSFETs will be explored.

2.4.4 Effect of Device Parameters on SNM

The SNM can be found analytically by solving the Kirchhoff equations and appling one of the mathematically equivalent noise margin criteria [57]. The Explicit expression for the SNM of the 6T cell was obtained by using the basic MOS model equations with constant threshold voltage (equal for n- and p-channel) and neglecting second-order effects such as mobility reduction and velocity saturation. The result is given below :

$$SNM = V_{th} - (\frac{1}{k+1})\{\frac{V_{DD} - \frac{2r+1}{r+1}V_{th}}{1 + \frac{r}{k(r+1)}} - \frac{V_{DD} - 2V_{th}}{1 + k\frac{r}{q} + \sqrt{\frac{r}{a}(1 + 2k + \frac{r}{q}k^2)}}\} \tag{2.1}$$

where

$$\begin{aligned} r = \text{ratio} &= \frac{\beta_d}{\beta_a} \\ q &= \frac{\beta_p}{\beta_a} \\ k = (\frac{r}{r+1})\{\sqrt{\frac{r+1}{r+1-\frac{V_{s2}}{V_{r2}}}} &- 1\} \\ V_r = V_s - (\frac{r}{r+1})&V_{th} \end{aligned} \tag{2.2}$$

When studying the SNM expression we can drae some intersting general conclusions. The SNM for 6T cells depends only on threshold voltage, supply voltage(V_{DD}), and β ratios, and not on the abosulte value of the β's. The SNM of 6-T SRAM cells also depends on the device parameters, such ad the channel length, the gate oxide thickness, the channel doping concentration, the source/drain doping concentration, and temperature because it is depends on threshold voltage.

2.5 Summary

The evolution of nanodevice structures is presented, where there are the single-gate MOSFET, SOI MOSFET, the double-gate transistor, the quasi-planar FinFET and surrounding gates. The conventional 6-T architecture has been explored and we show the cell ratio and ”q” ratio definition. Next we talked about each operations in detail. We try to explain how the cell works during standby, read and write operations. Several issues when designing SRAM cells are discussed, and they are hold stability, read stability, and sensitivity of static noise margin. We also introduce how to extract the stability during hold or read operations. Finally, effect of device parameters on SNM will be discussed.

Chapter 3

The Developed Simulation Methodology

In this chapter we introduce the main methodology of this work in more detailed in the following sections. First the methodology the flowchart of the SRAM cell stability analysis will be shown in section 3.1, to provide a general concept about the steps of this methodology steps. A computational statistical methodology that accounts for the sensitivity of the SRAM cell stability is depicted in section 3.2. Then we state each step in following sections respectively. Three main topics of the SRAM cell stability analysis will be shown. First, preprocess, in this step we can define accurate structure and doping profile for our simulations. Second, device simulation, this step give us the device characteristic. Finally, mixed-mode simulation, it coupled device and circuit simulation to calculate SRAM circuit performance. For the the sensitivity analysis of the SRAM cell stability, we consider five main topics will be shown. First, screening designs, in this step we can select fewer but important factors from many parameters. Second, design of experiment, this step help us construct the design matrix which can be used to make experiments or circuit simulations, through this effort we can get the responses which we are interested in. Third, construct the response surface model, after we know the responses, we build the response surface model (RSM)to connect the relationship between the responses and the factors. Finally we will discuss the sensitivity analysis.

3.1 The Computational Procedure of SNM Analysis

There are four steps in this work, and the flowchart is shown in Fig 3.1. We use this systematic analysis to study the stability of SRAM cells using different device structures. By the first step, we can define different device structures, accurate doping profile, and the mesh(refinement) specification respectively.

After we define the device structure, doping profile, and the mesh, electrical characteristics of different devices are investigated by using a device simulation. Our device simulation used the drift-diffusing model for carrier transport and the density gradient model to account for quantum-mechanical effects in nanoscale MOSFETs. Taking several important electrical characteristics as evaluation criteria, planar MOSFETs, SOI FinFETs,and nanowire FinFETs, are examined and compared.

Then we can use the mixed-mode simulation, we obtain the voltage transfer characteristic curves of 6-T SRAM cells with different device structures. Mixed-mode simulation is the complex equations are solved-consistently at process, device, and circuit level. Detailed will be discussed in the corresponding section.

The last step, we extract the cell static noise margin (SNM) parameter from the 6-T-SRAM-cell DC transfer characteristics to explore stability of SRAM cells. The cell stability is based on the ability of the cell to resist accidental overwrites during different operating conditions in the presence of electrical noise and process variations.

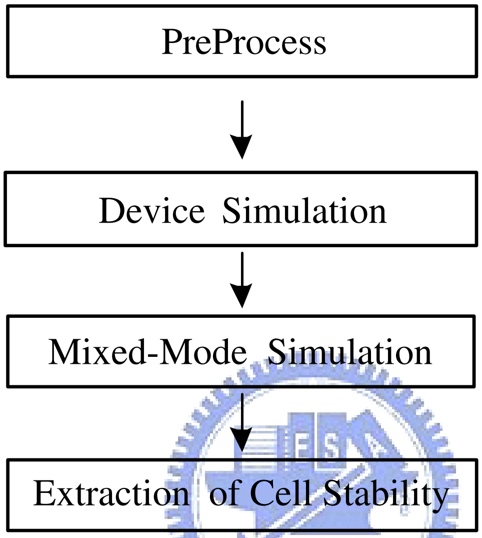


Figure 3.1: Systematic scheme for stability analysis of SRAM cells. It contains the preprocess, device simulation, mixed-mode simulation, and extraction of cell stability.

3.2 The Computational Procedure of Sensitivity Analysis

There are five steps in this work, and the flowchart will be shown in Fig 3.2. When we use this statistical analysis to explore the circuit performances sensitivity analyze, we might face many important and unimportant factors. The first step could help us to solve this problem, when there are too many factors so that we don't know how to select them, screen design is one suggested method to choose the significant factors which affect the responses most.

After we get the important factors, we can use them to build the complex response surface models which are more closed to the true nature. So we need to construct the design of experiment (DOE) before we build the response surface models.

Third step, we select the the face centered cube DOE technique, and we perform the mixed-mode simulation to extract the static noise margin parameter by solving a set of three-dimensional density-gradient based drift-diffusion equations, which is simultaneously coupled with circuit nodal equations.

Then we can use design matrix to do the experiment, from it we obtain the responses data, and next we can build the response surface model [42]. In this step the main purpose is to find the corresponding polynomial functions to describe the relationship between the responses and the factors. As we get these functions, we can use them to perform sensitivity analysis.

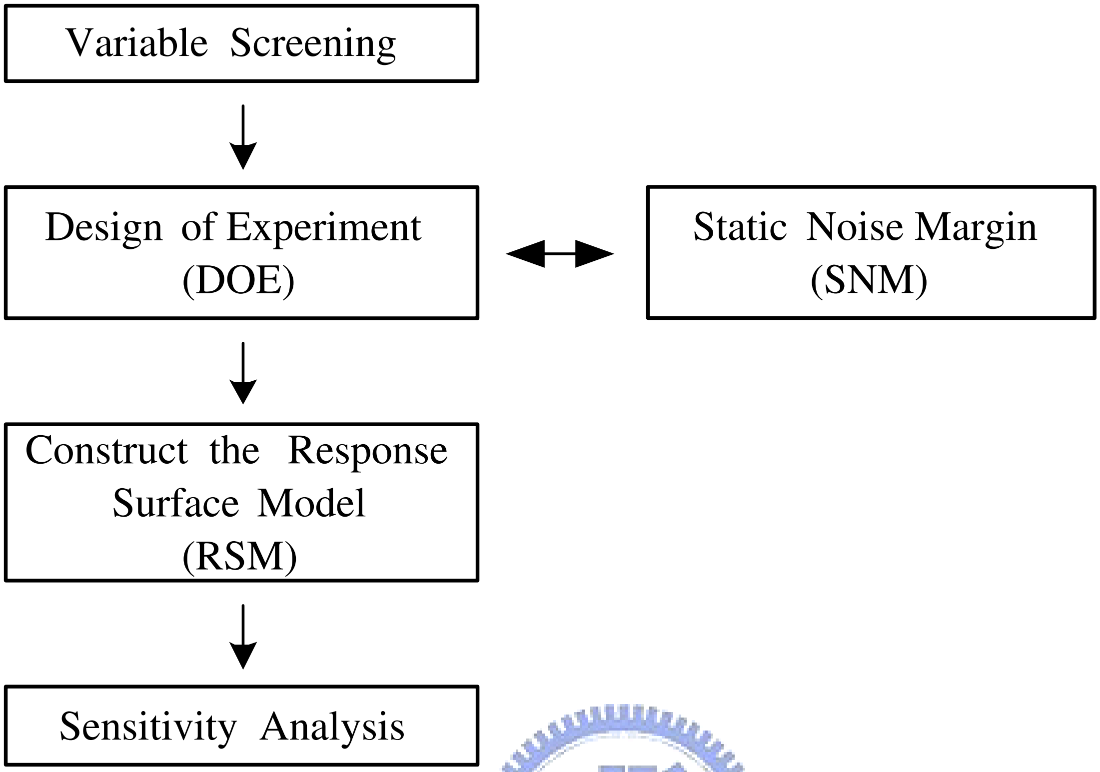


Figure 3.2: Systematic scheme for 6-T-SRAM-cell stability sensitivity analysis. It contains of the screening design, the design of experiment, mixed-mode simulation, the response surface model, and sensitivity analysis.

Finally we analyze the 6-T-SRAM-cell static noise margin (SNM) sensitivity by using the SNM's response surface model [43] - [44]. The sensitivity of SNM then can be explored in a computationally cost-effective way. We note that this analysis technique not only can be used together with device and circuit simulation programs for theoretical prediction but also can analyze experimental measurement foe realistic SRAM cell datas.

3.3 Preprocess

For avoiding time-consume simulation, we don't consider the detail process fabrication steps now. We directly define the device structure, accurate doping profile, and mesh specification by using the ISE TCAD simulator of commercial softwares.

Technology computer-aided design (TCAD) refers to using computer simulations to develop and optimize semiconductor processing technologies and devices. TCAD simulation tools solve fundamental and physical partial differential equations, such as diffusion and transport equations for discretized geometries, representing the silicon wafer or the layer system in a semiconductor device. This deep physical approach gives TCAD simulation predictive accuracy. It is therefore possible to substitute TCAD computer simulations for costly and time-consuming test wafer runs when developing and characterizing a new semiconductor device or technology. TCAD consists of two main branches: process simulation and device simulation.

In process simulation, processing steps such as etching, deposition, ion implantation, thermal annealing and oxidation are simulated based on physical equations, which govern the respective processing steps. The simulated part of the silicon wafer is discretized and represented as a finite element structure. For example, in the simulation of thermal annealing complex diffusion equations for each dopant species are solved on this mesh. For oxidation simulations the growth of the silicon oxide is simulated taking into account the oxygen diffusion, the mechanical stresses at corners, etc. In another example, Monte Carlo techniques are often used to simulate the process of ion implantation, where ion-semiconductor interactions are taken into account to compute individual ion paths.

Device simulations can be thought of as virtual measurements of the electrical behavior of a semiconductor device, such as a transistor, or a diode. Again, the device is represented as a meshed finite element structure. Each node of the device has properties associated with it, such as material type, doping concentration, etc. For each node the carrier concentration, current densities, electric field, generation and recombination rates etc., are computed. Electrodes are represented as areas on which boundary conditions, such as applied voltages. The device simulator solves the Poisson and the carrier continuity equation (and possibly other equations, such as lattice and carrier temperature equations). After solving these the resulting electrical currents at the contacts are extracted.

TCAD simulations are widely used throughout the semiconductor industry. As technologies become more and more complex the semiconductor industry relies increasingly more on TCAD to cut costs and speed up the research and development process. In addition, semiconductor-manufacturing companies use TCAD for yield analysis, i.e., monitoring, analyzing and optimizing their IC process flows, as well as to analyze the impact of IC process variation.

3.4 Device Simulation

Based on a multidimensional device simulation, the threshold voltage, ratio of the on/off current ratio (Ion/Ioff), subthreshold swing (S.S.), drain induced barrier lowering (DIBL), are extracted from the device terminal characteristics. We can also get the device intrinsic characteristics by device simulation. To focus on exploring the device characteristics for optimal nanodevice structures, phenomenological quantum correction model, a density-gradient equation together with the classical three-dimensional (3-D) drift-diffusion model is adopted and solved numerically in this study. A carefully calibrated density-gradient model has attracted more and more attention and successfully demonstrates its validity for efficient modeling of the quantum mechanical effects in a TCAD simulator using first-order quantum corrections. This simulation quantitatively predicts the main tendency of electrical and physical properties for the examined device structures. Full quantum mechanical methodologies definitely will input more accurate estimation on the characteristics, but it is believed that our simulation will not be significantly altered. The density- gradient modeling approach is computationally effective for incorporating the quantum mechanical effect in a multidimensional nanodevice TCAD simulation. The detail of the drift-diffuse model, the density-gradient model, and the numerical method are depicted in Chapter 4.

3.5 Mixed-Mode Simulation

There are three different methods to explore the 6-T SRAM cell performance. The methods are the following : the SPICE circuit simulation, the 3D direct device simulation, and the mixed-mode simulation. The SPICE circuit simulation must know the compact models of simulated devices. But for non-CMOS structures (ex. SOI FinFET or nanowire FinFET), a proper corresponding compact model is required when a circuit simulation is performed. This a reason that we do not use SPICE simulation for SRAM cells using non-CMOS structures. Alternatively, the mixed-mode simulation methodology can be employed under a full numerical way. The 3D direct device simulation is the best way among these methods, it explore circuit performance of SRAM cells by directly simulating three dimension device simulation. It can also consider the layout effect under the simulation. Fig. 3.3 shows 6T type SRAM SOI cell simulation structure. It was constructed with the combination of ISE 3D device simulator. We also constructed the via and wire structure. Although this way is the most fertile in the physical meaning, it is time-consuming. This way takes more than 50 hours of simulation time under the specific case. In this thesis, we use the mixed-mode simulation which connects the previous device simulation to analyze and compare performance of SRAM cells for saving time. The mixed-mode simulation takes only about 4 hours of simulation time under the specific case. The more detail will be discuss in section 4.5.

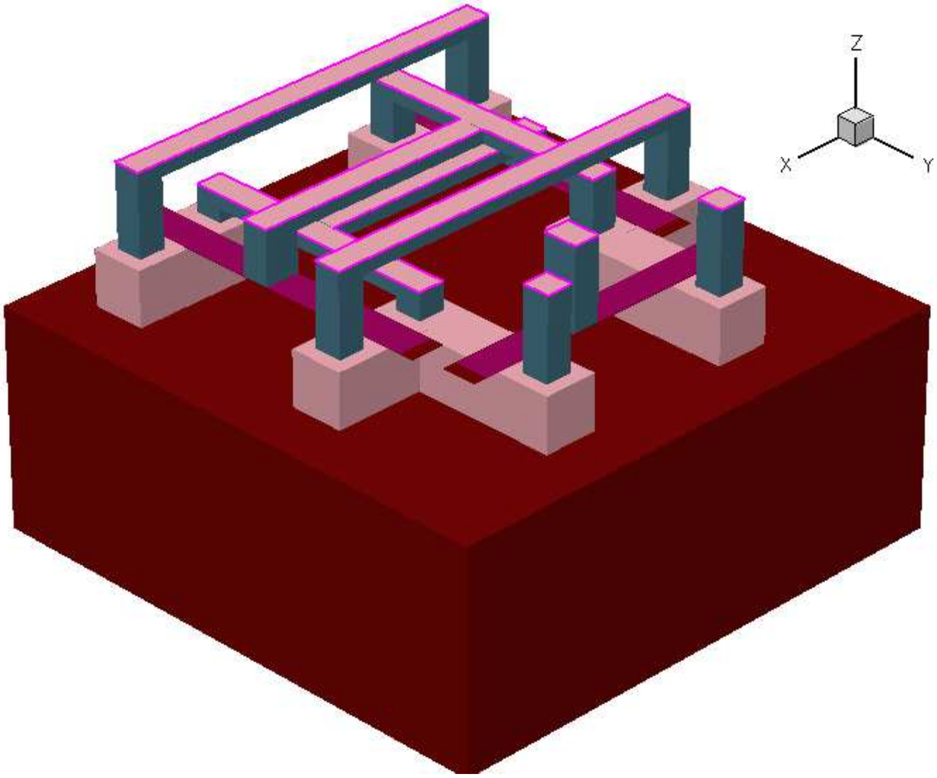


Figure 3.3: Simulated 6-T SOI SRAM structure. This structure is constructed and simulated by ISE 3D devise simulator.

3.6 Screening Designs

The term 'screening design' refers to an experimental plan that is intended to find the few significant factors from a list of many potential ones. Alternatively, we refer to a design as a screening design if its primary purpose is to identify significant main effects, rather than interaction effects, the latter being assumed an order of magnitude less important [45].

Even when the experimental goal is to eventually construct a response surface model (an RSM analysis), the first experiment should be a screening design when there are many factors to be considered.

We know there are many device factors which can affect the SNM of SRAM cells, and if we want to build a 2nd order response surface model, we need 43 completed runs for just 5 variables. This experiment takes more than 172 hours of simulation time by using a mixed-mode simulation. Thus, for the 6-T SRAM cell in which a large number of variables exist, time constraints may limit the feasibility of performing a response surface study including all possible device factors. The method of screening design may be employed during the initial stages of a response surface study to identify those processes having the greatest effect on process output. By reducing the number of factors taken into consideration, the subsequent response surface study may be streamlined, and fewer runs or tests required.

3.7 Design of Experiments

Response models are constructed relating the 6-T-SRAM-cell characteristics to the significant device factors using data generated from statistical experimentation. Mathematically, response surface models may be represented as second-order polynomials [42]:

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k} \sum_{i \neq j}^{k} \beta_{ij} x_i x_j + \varepsilon, \tag{3.1}$$

where k is the number of input factors, x_i is the ith input factor, β_i is the ith regression coefficient, and ε represents model error.

Traditional design is the central composite design (CCD) [42], and there are three special forms of it, which are the central composite circumscribed (CCC) design, the face centered cube (CCF) design, and the central composite inscribed (CCI) design [46]. If we feel the run number of the traditional central composite design is still too large, we can choose another design, which is called small composite design (SCD) [47]. Detailed will be discussed in the section 5.2. In this thesis, we use the central composite circumscribed design to create a SNM response surface model.

3.8 Response Surface Model

After we choose one kind of design, we do real experiments in the laboratory or use TCAD simulator to find the responses we are interested. If N observations are collected in an experiment, the model for them takes the form

$$y_n = \beta_0 + \sum_{i=1}^{k} \beta_i x_{ni} + \sum_{i=1}^{k} \beta_{ii} x_{ni}^2 + \sum_{i=1}^{k} \sum_{i \neq j}^{k} \beta_{ij} x_{ni} x_{nj} + \varepsilon_n, \quad \texttt{for} \quad n = 1, \ldots, N., \tag{3.2}$$

where y_n is the nth value of the response and $x_{n1}, \ldots, x_{nk}$ are the corresponding values of the k factors [42].

These N equations can be written in matrix notation as:

$$y = X\beta + \varepsilon, \tag{3.3}$$

where $y = (y_1, \ldots, y_N)^T$ is the $N \times 1$ vector of responses, $\beta = (\beta_0, \beta_i, \beta_{ii}, \beta_{ij})^T$ for $i = 1, \ldots, k$ and $1 \leq i < j \leq k$ is the $[(k+1)(k+2)/2] \times 1$ vector of regression coefficients, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_N)^T$ is the $N \times 1$ vector of errors.

The model is a univariate version. When there are m response variables to be studied together and there are dependence between each response variable, the multivariate regression model would be used to estimate the unknown β. Assuming that there are m response variables, the model has the following matrix form:

$$\begin{cases} y_{\cdot 1} = X\beta_{\cdot 1} + \varepsilon_{\cdot 1}, \\ y_{\cdot 2} = X\beta_{\cdot 2} + \varepsilon_{\cdot 2}, \\ \dots \\ y_{\cdot m} = X\beta_{\cdot m} + \varepsilon_{\cdot m}, \end{cases} \tag{3.4}$$

The unknown parameters in the model are the regression coefficients $\beta_{\cdot 1}$, ..., $\beta_{\cdot m}$ and the covariance matrix Σ. Estimates of the model coefficients in Eq.(3.4) are determined using a least-squares fit:

$$\begin{pmatrix} \hat{\beta}_{\cdot 1} \\ \hat{\beta}_{\cdot 2} \\ \dots \\ \hat{\beta}_{\cdot m} \end{pmatrix} = \begin{pmatrix} (X^TX)^{-1}X^Ty_{\cdot 1} \\ (X^TX)^{-1}X^Ty_{\cdot 2} \\ \dots \\ (X^TX)^{-1}X^Ty_{\cdot m} \end{pmatrix},$$

which is the vertical joining of the LSE's for the m-multiple linear regression model. The residuals, R^2, and MSE in each response variable will be shown in section 5.3.

3.9 Sensitivity Analysis

The objective is to investigate the sensitivity of ststic noise margin to varying a set of device parameters. Once a quadratic model for SNM response has been obtained, then it can be 'interrogated' by plotting the distribution of the response to a Gaussian or a uniform variation of the input factors [23]. In this approach, random values for each input factor are selected and the corresponding responses predicted using the response surface models. The input factors are assumed to be uniform or Gauss distributed about their mean. The standard deviation for each factor must be determined experimentally, or be set as a percentage of its mean values. Analysis of the distributions provides an estimate of the amount of response variation [44].

3.10 Summary

In this chapter, we discuss the methodology about this work. First the methodology flowchart will be shown. The methodology has two parts : the SNM and its sensitivity analysis of SRAM cells. For the SNM analysis of SRAM cells, preprocess is the first work, it can determine the device structure and doping profile. Then, we use device simulation to analyze the device electrical characteristics (including the intrinsic and terminal characteristics). After implementing device simulation, we do the mixed-mode simulation. From this step, we can get the circuit performance. Finally, we can extract the static noise margin by a graphical method. For the analysis of the SNM sensitivity, screening design is the first step in this work, it could help us to find the important factors from a lot of variables. After selecting a smaller group of factors, we do design of experiment. The traditional design matrix is based on the CCD, and there are 3 special forms (CCC, CCF, and CCI designs) of it. In this thesi, we select the face center cube (CCF) technique. After the step of design of experiment, the topic of constructing the response surface model is introduced, and through these models we can find the relationship between the responses and the selected factors. Finally sensitivity of SRAM cells analysis by using the SNM's response surface model are discussed.

Chapter 4

Physical Models and Mixed-Mode Simulation

Physical phenomena in semiconductor devices are very complicated and, depending on applications, are described by partial differential equations of different level of complexity. Coefficients and boundary conditions of equations (such as mobility, generation-recombination rate, material-dependent parameters, interface and contact boundary conditions) can be very complicated and can depend on microscopic physics, the structure of the device, and the applied bias.

In this chapter, we introduce the physical phenomena in semiconductor devices and the simulation numerical methods. First the different transport equations, the drift-diffusion model, the thermodynamic model, and the hydrodynamic model, are presented in section 3.1. We focus on the drift-diffusion model in this thesis. The different quantum correction model, the van Dort quantum correction model, the 1D Schrodinger equation, and the density gradient model are depicted in section 3.2. We focus on the density-gradient model now. The mobility model will be shown in section 3.3. It includes the doping-dependent mobility degradation, mobility degradation at interfaces, and high field saturation. Finally we briefly introduce the numerical method of our simulation. We divide the two parts : the discretization and the nonlinear solvers.

4.1 The Drift-Diffusion Model

Depending on the device under investigation and the level of modeling accuracy required, we can select four different simulation modes:

Drift-diffusion :

Isothermal simulation, described by basic semiconductor equations. Suitable for low power density devices with long active regions.

Thermodynamic :

Accounts for self-heating. Suitable for devices with low thermal exchange, particularly, high-power density devices with long active regions.

Hydrodynamic :

Accounts for energy transport of the carriers. Suitable for devices with small active regions.

Monte Carlo :

Allows for full band Monte Carlo device simulation in the selected window of the device.

The three governing equations for charge transport in semiconductor devices are the Poisson equation and the electron and hole continuity equations. The Poisson equation is:

$$\nabla^2\phi = -\frac{q}{\varepsilon_{si}}(p - n + N) \tag{4.1}$$

where ϕ is electrostatic potential, q is electric charge, ε_{si} is silicon permittivity, p is hole carrier density, n is electron carrier density, and $N = N_D - N_A$ is the electrically active net impurity concentration. The continuity equations for electrons and holes are given by

$$-\frac{1}{q}\nabla \cdot J_n - G + R + \frac{\partial n}{\partial t} = 0 \tag{4.2}$$

$$\frac{1}{q}\nabla \cdot J_p - G + R + \frac{\partial p}{\partial t} = 0 \tag{4.3}$$

where G incorporates generations phenomena, such as impact ionization or carrier generation by external radiation and R describes recombination process. J_n and J_p are the electron and hole current density given by

$$J_n = q\mu_n nE + qD_n \nabla n \tag{4.4}$$

$$J_p = q\mu_p pE - qD_p \nabla p \tag{4.5}$$

where μ_n and μ_p are the electron and hole mobility, and D_n and D_p are the corresponding diffusion coefficients. Both the mobility and diffusion coefficients depend on electric field.

The drift-diffusion model is widely used for the simulation of carrier transport in semiconductors and is defined by the basic semiconductor equations which consist of Poisson equation, electron and hole continuity equation.

4.2 The Density-Gradient Model

The scaling rules for modern sub-micron devices require a thinner oxide and higher level of channel doping. Some features of current MOSFETs (oxide thickness, channel width) have reached quantum mechanical length scales. Therefore, the wave nature of electrons and holes can no longer be neglected. The most basic quantization effects in MOSFETs are the shift of the threshold voltage and reduction of the gate capacity.

To include quantization effects in a classical device simulation, a simple approach is to introduce an additional potential-like quantity Λ in the classical density formula, which reads:

$$n = N_C exp(\frac{E_F - E_C - \Lambda}{k_B T}) \tag{4.6}$$

where n is the electron density, T is the carrier temperature, k_B is the Boltzmann constant, N_C is the conduction band density of states, E_C is the conduction band energy, and E_{F_n} is the electron Fermi energy. (For brevity, only the formulas for electrons are given; holes are handled analogously.) When using Fermi statistics, the exponential function in Eq. 4.6 is replaced by a Fermi integral of order $1/2$.

The most important effects related to the density modification (due to quantization) can be captured by proper models for Λ. Other effects (for example, single electron effects) exceed the scope of this approach.

We briefly introduce three quantization models, that is, three different models for Λ .

They differ in physical sophistication, numeric expense, and robustness:

The van Dort quantum correction model :

It is a numerically robust, fast, and proven model. It is only suited to MOSFET simulations. Structures such as quantum wells and SOI transistors with ultrathin silicon layer (below approximately 10 nm) are beyond the scope of this model. While important terminal characteristics are well described by this model, it does not give the correct density distribution in the channel.

The 1D Schrodinger equation :

It is the most physically sophisticated quantization model. It can be used for MOSFET simulation, and quantum well and ultrathin SOI simulation. Simulations with this model tend to be slow and often lead to convergence problems, which restrict its use to situations with small current flow. Therefore, the Schrodinger equation is used mainly for the validation and calibration of other quantization models.

The density gradient model :

It is numerically robust, but significantly slower than the van Dort model. It can be applied to MOSFETs, quantum wells and SOI structures, and gives a reasonable description of terminal characteristics and charge distribution inside a device. Compared to the other quantization models, it can describe 2D and 3D quantization effects.

The density-gradient model advanced by Ancona and his coworkers is an approximate approach to the QM correction of the macroscopic electron transport equation. For the density gradient model, Λ in Eq. 4.6 is given in terms of a partial differential equation :

$$\Lambda = -\frac{r\hbar^2}{12m}\nabla^2 logn + \frac{1}{2}(\nabla logn)^2 = -\frac{r\hbar^2}{6m}\frac{\nabla^2\sqrt{n}}{\sqrt{n}} \tag{4.7}$$

where $\hbar = \frac{h}{2\pi}$ is the reduced Planck constant, m is the DOS mass, and γ is a fit factor.

In this approach, an extra term is introduced in the carrier flux by making the equation of state for the electron gas density-gradient dependent (Ancona and Tiersten 1987, Ancona 1997), i.e.,

$$J_n = q\mu_n nE + qD_n\nabla n - qn\mu_n\nabla(2b_n\frac{\nabla^2\sqrt{n}}{\sqrt{n}}) \tag{4.8}$$

where $b_n = \frac{h^2}{12qm_n^*}$ is the (linear) density-gradient coefficient.

4.3 The Mobility Models

We uses a modular approach for the description of the carrier mobilities. In the simplest case, the mobility is a function of the lattice temperature. This so-called constant mobility model should only be used for undoped materials. For doped materials, the carriers scatter with the impurities. This leads to a degradation of the mobility. Section 3.3.1 introduces the models that describe this effect.

Models that describe the mobility degradation at interfaces, for example, the silicon-Voxide interface in the channel region of a MOSFET, are introduced in Section 3.3.2. These models account for the scattering with surface phonons and surface roughness. Finally, the models that describe mobility degradation in high electric fields are discussed in Section 3.3.3.

4.3.1 Doping-Dependent Mobility Degradation

In doped semiconductors, scattering of the carriers by charged impurity ions leads to degradation of the carrier mobility.

The model used to simulate doping-dependent mobility in silicon was proposed by Masetti et al.

$$\mu_{dop} = \mu_{min1} exp(-\frac{P_c}{N_i}) + \frac{\mu_{const} - \mu_{min2}}{1 + (\frac{N_i}{C_r})^{\alpha}} - \frac{\mu_1}{1 + (\frac{C_s}{N_i})^{\beta}} \tag{4.9}$$

where $N_i = N_A + N_D$ denotes the total concentration of ionized impurities.

The reference mobilities μ_{min1}, μ_{min2}, and μ_1, the reference doping concentration P_c, C_r, and C_s, and the exponents α and β accessible in the parameter file.The corresponding values for silicon are given in Table 4.1. The low-doping reference mobility μ_{const} is determined by the constant mobility model.

Table 4.1: Masetti model : Default coefficients

Symbol	**Parameter name**	**Electrons**	**Holes**	**Unit**
μ_{min1}	mumin1	52.2	44.9	cm^2/Vs
μ_{min2}	mumin2	52.2	0	cm^2/Vs
μ_1	mu1	43.4	29.0	cm^2/Vs
P_c	Pc	0	9.23×10^{16}	cm^{-3}
C_r	Cr	9.68×10^{16}	2.23×10^{17}	cm^{-3}
C_s	Cs	3.34×10^{20}	6.10×10^{20}	cm^{-3}
α	alpha	0.680	0.719	1
β	beta	2.0	2.0	1

4.3.2 Mobility Degradation at Interfaces

In the channel region of a MOSFET, the high transverse electric field forces carriers to interact strongly with the semiconductorVinsulator interface. Carriers are subjected to scattering by acoustic surface phonons and surface roughness. The models in this section describe mobility degradation caused by these effects. The model is called Lombardi model.

The surface contribution due to acoustic phonon scattering has the form:

$$\mu_{ac} = \frac{\beta}{\Gamma_{\perp}} + \frac{C(\frac{N_i}{N_0})^{\lambda}}{\Gamma_{\perp}^{\frac{1}{3}}(\frac{T}{T_0})^{k}} \tag{4.10}$$

and the contribution attributed to surface roughness scattering is given by:

$$\mu_{sr} = (\frac{(\frac{\Gamma_{\perp}}{\Gamma_{ref}})^{A^*}}{\delta} + \frac{\Gamma_{\perp}^{3}}{\eta})^{-1} \tag{4.11}$$

These surface contributions to the mobility (μ_{ac} and μ_{sr}) are then combined with the bulk mobility μ_b :

$$\frac{1}{\mu} = \frac{1}{\mu_b} + \frac{D}{\mu_{ac}} + \frac{D}{\mu_{sr}} \tag{4.12}$$

In the above formulas, $N_i = N_A + N_D$ refers to the total concentration of ionized impurities and $T_0 = 300K$. The reference field $\Gamma_{ref} = 1\frac{V}{cm}$ ensures a unitless numerator in (Eq. 4.11). $\Gamma_{\perp}$ is the transverse electric field normal to the semiconductorVinsulator interface. $D = exp(\frac{-x}{l_{crit}})$ (where x is the distance from the instance and l_{crit} a fit parameter) is a damping that switches off the inversion layer terms far away from the interface. In the

Table 4.2: Lombardi model : Default coefficients for silicon

Symbol	**Parameter name**	**Electrons**	**Holes**	**Unit**
B	B	$4.75 * 10^{7}$	$9.925 * 10^{6}$	cm/s
C	C	$5.80 * 10^{2}$	$2.947 * 10^{3}$	$cm^{5/3}/V^{2/3}s$
N_0	N0	1	1	cm^{-3}
λ	lambda	0.1250	0.0317	1
k	k	1	1	1
δ	delta	$5.82 * 10^{14}$	$2.0546 * 10^{14}$	cm^2/Vs
η	eta	$5.82 * 10^{30}$	$2.0546 * 10^{30}$	V^2/Vs
l_{crit}	lcrit	$1 * 10^{-6}$	$1 * 10^{-6}$	cm

Lombardi model, the exponent A^* in Eq. 4.11 is equal to 2. The respective parameters that are appropriate for silicon are given in Table 4.2.

Table 4.3: Canali model parameters (default values for silicon)

Symbol	Parameter name	Electrons	Holes	Unit
β_0	beta0	1.109	1.213	1
β_{exp}	betaexp	0.66	0.17	1

4.3.3 High Field Saturation

In high electric fields, the carrier drift velocity is no longer proportional to the electric field strength, instead, the velocity saturates to a finite speed v_{sat}. The Canali model is available in different versions (one for drift-diffusion and thermodynamic, and one for hydrodynamic simulations).

The Canali model originates from the CaugheyVThomas formula, but has temperature-dependent parameters, which were fitted up to 430 K by Canali et al. :

$$\mu(F) = \frac{\mu_{low}}{[1 + (\frac{\mu_{low} F}{v_{sat}})^{\beta}]^{\frac{1}{\beta}}} \tag{4.13}$$

where μ_{low} denotes the low field mobility. The exponent is temperature dependent according to:

$$\beta = \beta_0 (\frac{T}{T_0})^{\beta_{exp}} \tag{4.14}$$

where T denotes the lattice temperature and $T_0 = 300K$. The silicon default values are listed in Table 4.3.

4.3.4 The Recombination-Generation Model

The recombination and generation model will be introduced. In our simulation, we use the ShockleyVReadVHall (SRH) model to describe the behavior of carrier recombination. The avalanche generation (impact ionization) model is used to describe the behavior of carrier generation.

Recombination through deep levels in the gap is usually labeled ShockleyVReadVHall (SRH) recombination. The following form is implemented:

$$R_{net}^{SRH} = \frac{np - n_{i,eff}^2}{\tau_p(n + n_l) + \tau_n(p + p_l)} \tag{4.15}$$

with

$$n_l = n_{i,eff} e^{\frac{E_{trap}}{kT}} \tag{4.16}$$

and

$$p_l = n_{i,eff} e^{\frac{-E_{trap}}{kT}} \tag{4.17}$$

where E_{trap} is the difference between the defect level and intrinsic level. The variable E_{trap} is accessible in the parameter file. The silicon default value is $E_{trap} = 0$. The minority lifetimes τ_n and τ_p are modeled as a product of a doping-dependent, field-dependent, and temperature-dependent factor:

$$\tau_c = \tau_{dop} \frac{f(T)}{1 + g_c(F)}, c = n, p \tag{4.18}$$

where $c = n$ or $c = p$ for holes. For simulations that use Fermi statistics or quantization, Eq. 4.15 needs to be generalized. The modified equation reads:

$$R_{net}^{SRH} = \frac{np - r_n r_p n_{i,eff}^2}{\tau_p(n + r_n n_l) + \tau_n(p + r_p p_l)} \tag{4.19}$$

with

$$r_r = \frac{n}{N_C} exp(-\eta_n) \tag{4.20}$$

and

$$r_p = \frac{p}{N_V} exp(-\eta_p) \tag{4.21}$$

Electron-hole pair production due to avalanche generation (impact ionization) requires a certain threshold field strength and the possibility of acceleration, that is, wide space charge regions. If the width of a space charge region is greater than the mean free path between two ionizing impacts, charge multiplication occurs, which can cause electrical breakdown. The reciprocal of the mean free path is called the ionization coefficient α. With these coefficients for electrons and holes, the generation rate can be expressed as:

$$G^{\|} = \alpha_n n v_n + \alpha_p p v_p \tag{4.22}$$

where v_n and v_p denotes the drift velocity.

4.4 Numerical Methods foe Solving the Semiconductor Device Equation

In this section, we introduce the numerical methods in our simulation. We divide two parts : the discretization and the nonlinear solvers. First, we will show the ”box discretization” methdo. It is applied to discretize the partial differential equations (PDEs). In the next section, we introduce the solution of nonlinear systems. applied. This scheme tries to solve the nonlinear system by the Newton method.

4.4.1 Discretization

The well-known ”box discretization” is applied to discretize the partial differential equations (PDEs). This method integrates the PDEs over a test volume such as that shown in Fig. 4.1, which applies the Gaussian theorem, and discretizes the resulting terms to a first-order approximation.

In general, box discretization discretizes each PDE of the form:

$$\nabla \cdot \overrightarrow{J} + R = 0 \tag{4.23}$$

into:

$$\sum_{j \neq i} k_{ij} j_{ij} + \mu(\Omega_j) \cdot r_j = 0 \tag{4.24}$$

with the value of k_{ij} is $\frac{D_{ij}}{l_{ij}}$ for three dimension.

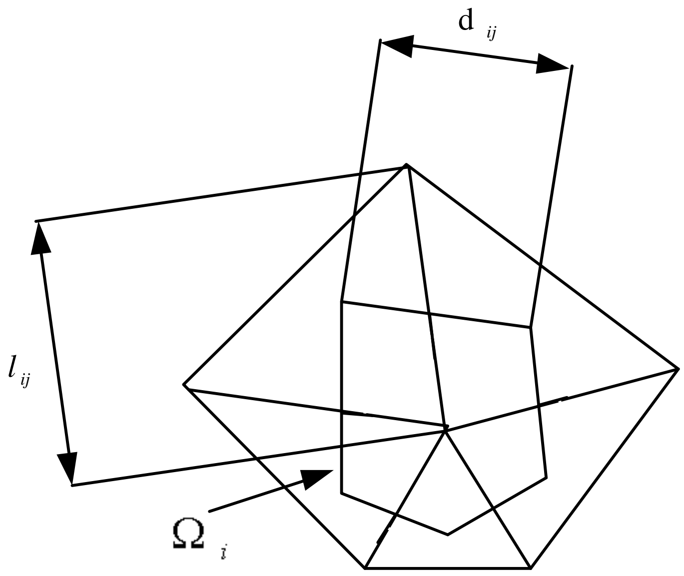


Figure 4.1: Single box for a triangular mesh in 2D.

Table 4.4: Equation

Equation	j_{ij}	r_{ij}
$Poisson$	$\varepsilon(\mu_i - \mu_j)$	$-\rho_i$
$Electron\ continuity$	$\mu^n(n_i B(\mu_i - \mu_j) - n_j B(\mu_j - \mu_i))$	$R_i - G_i + \frac{d}{dt} n_i$
$Hole\ continuity$	$\mu^p(p_j B(\mu_j - \mu_i) - p_i B(\mu_i - \mu_j))$	$R_i - G_i + \frac{d}{dt} p_i$
$Temperature$	$k(T_i - T_j)$	$H_i - \frac{d}{dt} r_i c_i$

In this case, the physical parameters j_{ij} and l_{ij} have the values listed in Table 4.4, where $B(x) = \frac{x}{e^x - 1}$ is the Bernoulli function.

One special feature is that the actual assembly of the nonlinear equations is performed elementwise, that is:

$$\sum_{\varepsilon \in Elements(i)} [\sum_{j \in Vertices(e)} k^e_{ij} j^e_{ij}] + \mu^e(\Omega_i \cdot r^e_i) = 0 \tag{4.25}$$

This expression is equivalent to (Eq. 4.24), but has the advantage that some parameters (such as ε, μ_n, μ_p) can be handled elementwise, which is useful for numeric stability and physical exactness.

4.4.2 Nonlinear Solvers

In this sections, we introduce the full coupled solution and Plugin iterations.

First, we will show the full coupled solution. For the solution of nonlinear systems, the scheme developed by Bank and Rose is applied. This scheme tries to solve the nonlinear system $g(z) = 0$ by the Newton method:

$$\overrightarrow{g} + \overrightarrow{g}'\overrightarrow{x} = 0 \tag{4.26}$$

$$\overrightarrow{z}^j - \overrightarrow{z}^{j+1} = \lambda \overrightarrow{x} \tag{4.27}$$

where λ is selected such that $\|g_{k+1}\| < \|g_k\|$, but is as close as possible to 1. It handles the error by computing an error function that can be defined by two methods.

The Newton iterations stop if the convergence criteria are fulfilled. One convergence criterion is the norm of the right-hand side, that is, $\|g\|$ in Eg. 4.26. natural criterion may be the relative error of the variables measured, such as $\|\frac{(\lambda x)}{z}\|$

We will introduce the Plugin iterations now. This is the traditional scheme, which is also known as ”Gummel iterations” in most other device simulators. Consider that there are n sets of nonlinear systems $g_j(z_1 \cdots z_n) = 0$. n can be, for example, 3 and the sets can be the Poisson equation and two continuity equations.) This method starts with values $z_{1(1)} \cdots z_{n(1)}$ and then solves each set $g_j = 0$ separately and consecutively. One loop could be:

$$
\begin{aligned}
g_1(z_1 z_2^{(i)} \cdots z_n^{(i)}) = 0 &\Longrightarrow z_1^{(i+1)} \\
&\cdots \\
g_1(z_1^{(i+1)} \cdots z_{n-1}^{(i+1)} z_n) = 0 &\Longrightarrow z_n^{(i+1)}
\end{aligned}
\tag{4.28}
$$

If an update (λx) of the solution between two successive Plugin iterations is defined as:

$$
(\lambda x) = z_j^{(i+1)} - z_j^{(i)} \tag{4.29}
$$

4.5 Mixed-Mode Simulation

To analyze and compare the 6-T SRAM cell performance with different device structures, we use the mixed-mode simulation which connects the previous device simulation to study characteristics at circuit level. The mixed-mode simulation methodology is the complex equations are solved-consistently at process, device, and circuit level in a coupled fashion in simulating 6-T SRAM circuit behaviors. In this thesis, we don't consider the detail process steps for economizing the use of time. So the mixed-model simulation include various device physics equations as discussed in preceding section, and Kirchhoff circuit equations, listed later. The voltage can get from the Poisson equation, and the current can get from integrating the current continuous equations under device simulation.

Kirchhoff circuit equations consist of two equations : Kirchhoff current equation and Kirchhoff voltage equation. Kirchhoff current equation is given by

$$\sum_{x=1}^{m} i_x = 0 \tag{4.30}$$

where m is the number of paths at ome node, and Kirchhoff voltage equation is given by

$$\sum_{y=1}^{n} v_y = 0 \tag{4.31}$$

where n is the number of branched in a closed loop.

Equations of the 6-TSRAM cell in our simulation will be described in detail. The scheme of the 6-T SRAM circuit is presented in Fig. 2.1. The world line and bit line is biased to V_{DD}, then access transistors will turn on. We assume that the node of Q is logic high ”1”, and the node of QB is logic low ”0”. The transistors of M1 and M6 will turn on, and the transistors of M3 and M4 will turn off. The Kirchhoff current equations are the following:

$$I_{DS1} = I_{DS2} \tag{4.32}$$

$$I_{DS1} = I_{DS2} \tag{4.33}$$

The Kirchhoff voltage equations are the following:

$$V_{GS1} = V_{DS4} \tag{4.34}$$

$$V_{DS6} = V_{DD} - V_{GS1} \tag{4.35}$$

$$V_{GS6} = V_{DD} - V_{DS1} \tag{4.36}$$

$$V_{GS2} = V_{DD} - V_{DS1} \tag{4.37}$$

By the Kirchhoff current and voltage equations, the connection of device simulation and circuit simulation will be achieved. We can analyze the circuit performance without compact modes by a mixed-mode simulation.

4.6 Summary

In this chapter, we first discuss the different transport models and the quantum correction models. In our simulation, we use the three-dimension drift-diffuse transport model and the density-gradient quantum correction model to account for quantum-mechanical effects in nanoscale devices. Three different effects for the carrier mobility, such as doping-dependent mobility degradation, mobility degradation at interfaces, and high-filed saturation are considered. We also state the recombination-generation model of our simulation. We use the ShockleyVReadVHall (SRH) model to describe the behavior of carrier recombination, and the avalanche generation (impact ionization) model is used to describe the behavior of carrier generation. Finally the numerical methods of our simulation is depicted in section 3.4. It consists of the discretization and the nonlinear solvers.

Chapter 5

The Computational Statistic Technique

In this chapter, we will discuss the computation statistical technique in detail. First, a method of the screening design will state in section 5.1 in detail. We take a SRAM cell using 65 nm CMOS devices to be an example, and to find find the few significant factors from a list of many potential ones. The design of experiment (DOE) will be shown in section 5.2. Many applications of response surface models involve constructing and checking the adequacy of a second-order model. The central-composite design (CCD) is perhaps the most common experimental design used to generate second-order response models. Finally, the the residuals, R^2, and MSE about the response surface model (RSM)are discussed.

5.1 Screening Design

Time constraints may limit the feasibility of performing a response surface study including all possible process conditions. Statistical factor screening methods may be employed during the initial stages of a response surface study to identify those processes having the greatest effect on process output. By reducing the number of parameters taken into consideration, the subsequent response surface study may be streamlined, and fewer runs or tests required. Due to the widespread assimilation of TCAD tools throughout the semiconductor industry, it is important to identify a factor screening method suitable for deterministic systems.

To determine factor significance, this work utilizes an approach combining statistical design of experiment methods and statistical graphical methods. To minimize the number of simulations or experimental runs, statistical designs should be used to efficiently explore the experimental design space. Two-level fractional factorial design with resolution III or Plackett-Burman design plans are ideally suited for screening design [46].

A special subset of factorial designs, Plackett-Burman designs, are published by R.L. Plackett and J.P. Burman in 1946 [47]. It provides very economical designs that permit the study of $k = N - 1$ variables in N runs, where N is a multiple of 4 (rather than a power of 2) [46]. For example, 11 device parameters may be studied in 12 experimental runs. In a Plackett-Burman design, main effects are, in general, heavily confounded with two-factor interactions. By using such design schemes, the circuit designer sacrifices the ability to clearly separate the main effects of each device parameter from certain higher-order interactions, a condition known as ’aliasing’. During factor screening, some aliasing is deemed acceptable in the face of time and or resource constraints [44].

Using a graphical method to judge effect significance is often preferred to using an extension of the t test such as the studentized maximum modulus test, because the latter depends on the estimate S^2 of σ^2, which may be unreliable, where as the former depends on seeing a deviation from linearity, which is easier to judge [47]. The S^2 is the sample variance, and the σ^2 is the true variance of the population. A related graphical method is the half-normal probability plot. Let $|\hat{\theta}|_{(1)} \leqslant \ldots \leqslant |\hat{\theta}|_{(I)}$, where $|\hat{\theta}|_{(i)}$ is the ordered statistics of the effects, for i = 1, . . ., I. Plot them against coordinates based on a half-normal distribution; the absolute value of a normal random variable is half-normal. The half-normal probability plot consists of the points

$$(\Phi^{-1}(0.5 + 0.5[i - 0.5]/I), |\hat{\theta}|_{(i)}), \tag{5.1}$$

for i = 1, . . ., I. The Φ is the cumulated density function of the standard normal distribution. The advantage of the half-normal plot is that all the large estimated effects appear in the upper right corner and fall above the line through the small estimated effects. In this work half-normal plots are used for testing effect significance. For other purposes such as outlier detection, normal plots may be preferred.

5.2 The Design of Experiment

Many applications of response surface models involve constructing and checking the adequacy of a second-order model. The central-composite design (CCD) is perhaps the most common experimental design used to generate second-order response models. These designs combine a two-level full factorial or fractional factorial design of n_f runs with $2k$ axial runs and n_c center runs, where k represents the number of control factors [42].

As seen in Fig. 5.1, central-composite designs include five input levels for each control factor $(0, \pm 1, \pm \alpha)$. Level 0, the nominal factor level, represents the base processing conditions. The cube levels (± 1) are selected to reflect the design space of interest [49].These values are typically set to a multiple of the factor's standard deviation or a percentage of its nominal value [44]. The precise value of α depends on certain properties desired for the design and on the number of factors involved.

5.2.1 Forms of the CCD

The central composite circumscribed (CCC) designs are the original form of the central composite design [46]. The axial points at some distance α from the center base on the properties desired for the design and the number of factors in the design. The axial points establish new extremes for the low and high settings for all factors. Fig 5.1 illustrates a CCC design. These designs have circular, spherical, or hyperspherical symmetry and require 5

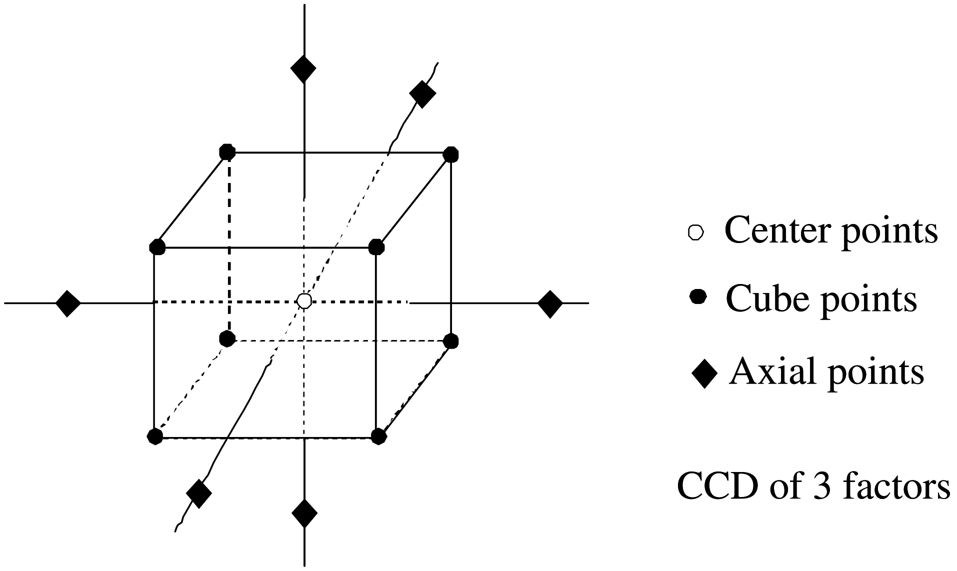


Figure 5.1: Central composite design for 3 factors.

levels for each factor. Augmenting an existing factorial or resolution V fractional factorial design with axial points can produce this design.

For those situations in which the limits specified for factor settings are truly limits, the central composite inscribed (CCI) design uses the factor settings as the axial points and creates a factorial or fractional factorial design within those limits (in other words, a CCI design is a scaled down CCC design with each factor level of the CCC design divided by α to generate the CCI design) [46]. This design also requires 5 levels of each factor.

The other special design is called the face centered cube (CCF) design. In this design the axial points are at the center of each face of the factorial space, so $\alpha = \pm 1$. For Fig.

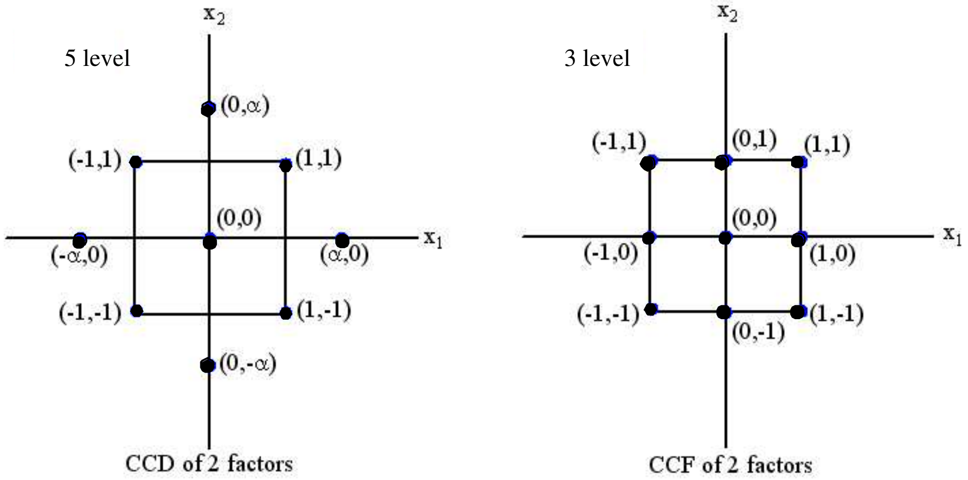


Figure 5.2: Comparison of the CCC design and the CCF design.

5.1, if the diamond points move to the face in the cube, then the design is CCF. This variety requires 3 levels of each factor. Augmenting an existing factorial or resolution V design with appropriate axial points can also produce this design. Fig. 5.2 shows the difference between the CCC and CCF designs for 2 factors design [46].

The diagrams in Fig. 5.3 illustrate the three types of central composite designs for two factors. Note that the CCC explores the largest process space and the CCI explores the smallest process space. Both the CCC and CCI are rotatable designs, but the CCF is not. In the CCC design, the design points describe a circle circumscribed about the factorial square. For three factors, the CCC design points describe a sphere around the factorial cube. To maintain rotatability, the value of α depends on the number of experimental runs

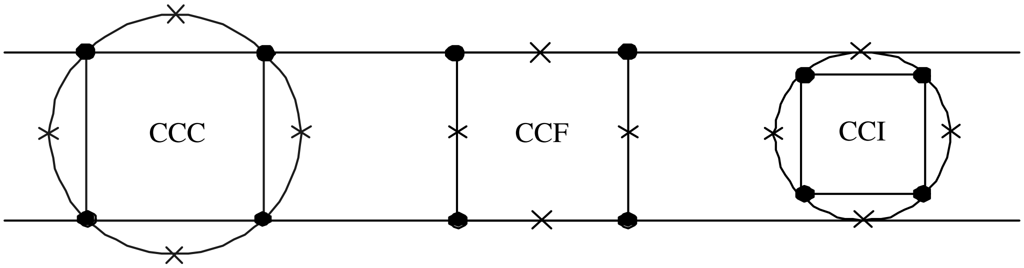


Figure 5.3: Comparison of the three types of central composite designs.

in the factorial portion of the central composite design:

$$\alpha = [n_c]^{\frac{1}{4}}, \tag{5.2}$$

where n_c is the number of experimental runs in the factorial potion of the central composite design. However, the factorial portion can also be a fractional factorial design of resolution V.

5.2.2 The Small Composite Design (SCD)

Alternative designs are small composite designs, and the term 'small' refers to an experimental design that has small runs in CCD [47]. A typical CCD requires about 9 runs for 2 factors, 15 runs for 3 factors, 25 runs for 4 factors, and 27 runs for 5 factors. It is obvious that, once you have four or more factors you wish to include in a CCD, you will need more than one lot (i.e., batch) of experimental units for your basic design. However, there is a way to cut down on the number of runs, as suggested by H.O. Hartley in his paper ’Smallest Composite Designs for Quadratic Response Surfaces’, has published in Biometrics, December 1959.

Theorem 5.2.1 *In any central composite design whose factorial portion is a* 2^{k-p} *design that does not use any main effect as a defining relation, the following parameters in Eq.(3.1) are estimable:*

1. β_0, β_i, β_{ii}, $i = 1, \ldots, k$, *and*

2. one β_{ij} *selected from each set of aliased effects for* $i < j$.

It is not possible to estimate more than one β_{ij} *from each set of aliased effects.*

This method addresses the theory that using a Resolution V design as the smallest fractional design to create a response surface design is unnecessary. The method adds axial points to designs of Resolution III and uses the axial points to clear the main effects of aliasing with the two-factor interactions. The resulting design allows estimation of the higher-order interactions [50].

Smaller central composite designs can be found by using the Plackett-Burman designs for the factorial portion and be represented by Draper and Lin. By using this method, several designs achieve to the smallest composite design [51]. Table 5.1 present the different designs for k factors.

Table 5.1: A list of the central composite designs for $2 \leq k \leq 10$ factors

Factors	**Number-of-coefficients**	**Runs**	**Runs**	**Runs**
k	$(k+1)(k+2)/2$	**CCD**	**Hartley**	**Draper-and-Lin**
2	6	9	-	7
3	10	15	11	11
4	15	25	17	17
5	21	27	-	22
6	28	45	29	29
7	36	79	47	37
8	45	81	-	47
9	55	147	83	57
10	66	149	-	67

5.3 The Response Surface Model

We will discuss the residuals, R^2, and mean-square-error (MSE) in each response variable. Following model generation, a residual analysis is conducted to evaluate the model adequacy. The residual for the nth observation in each response is defined as:

$$e_n = y_n - \hat{y}_n, \tag{5.3}$$

where y_n is the nth measured response (i.e., one physical quantity in the simulation results), and $\hat{y}_n$ is the nth predicted response (i.e., one physical quantity from the response surface model results) [42]. The normality assumption is checked by preparing a normal probability plot of the residual values. If the assumption holds, this plot will resemble a straight line. If the assumption is violated, a non-linear data transformation (e.g., $y' = log(y)$) may be applied and new RSM models generated in an attempt to improve model adequacy [42]. A second plot showing the residual values versus the predicted response values is used to verify if the variance of the original observation is constant. A random scattering of the residual values indicates that no correlation exists between the observed variance and the mean level of the response.

A relatively simple procedure is performed to check for model significance in relation to random error. This test involves calculating the test statistic:

$$F_0 = \frac{\frac{1}{p}\sum_{j=1}^{n}(\hat{y}_i - \overline{y})^2}{\frac{1}{n-p-1}\sum_{j=1}^{n}(y_i - \hat{y}_i)^2}, \tag{5.4}$$

where $\overline{y}$ is the average of measured response values, m is the number of model coefficients, and again y_i, $\hat{y}_i$, and n are the ith measured response, the ith predicted response, and the number of simulated runs, respectively [42]. If this statistic exceeds the corresponding value of the F distribution value ($F_{\alpha,p,n-p-1}$), the response model is considered significant in relation to random error. A second statistic, the coefficient of multiple determination, defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}, \tag{5.5}$$

it measures the amount of reduction in variability of the response y achieved, using the input factors $x_1, x_2, \ldots, x_k$. R^2 varies from zero to one with a value of one being ideal [42].

Because the R^2 value measures the ”proportion of total variation explained by the constructed regression model $X\hat{\beta}$”, a higher R^2 value indicates a better fit of the regression model. It can be shown that R is the correlation between $y = (y_i)_{i=1}^{N}$ and $\hat{y} = (\hat{y}_i)_{i=1}^{N}$ and thus is called the *multiple correlation coefficient* [47]. The residual mean square is commonly referred to as the *mean-square error* (MSE) and is an estimate $\hat{\sigma}^2$ for σ^2, i.e.,

$$\hat{\sigma}^2 = (y - X\hat{\beta})^T(y - X\hat{\beta})/(N - p - 1), \tag{5.6}$$

where p is equal to $(k + 1)(k + 2)/2$.

5.4 Summary

By screening design, we introduce the method of screening design to find the significant factor from a list of many potential ones. With the developed SNM model, the impact of device parameters on SNM is calculated. It is known that the variance of SNM is strongly affected by the critical dimension of device channel length. Besides, the SNM is seen to be the most sensitive to channel length fluctuations in the pull-down and access transistors and least sensitive to the channel length fluctuations in the load transistors. We introduce the classification of design of experiment (DOE). It consists of central-composite design and small composite design. The residuals, R^2, and mean-square-error (MSE) in each response variable will be depicted.

Chapter 6

Electrical Characteristics of the Explored Nanodevices

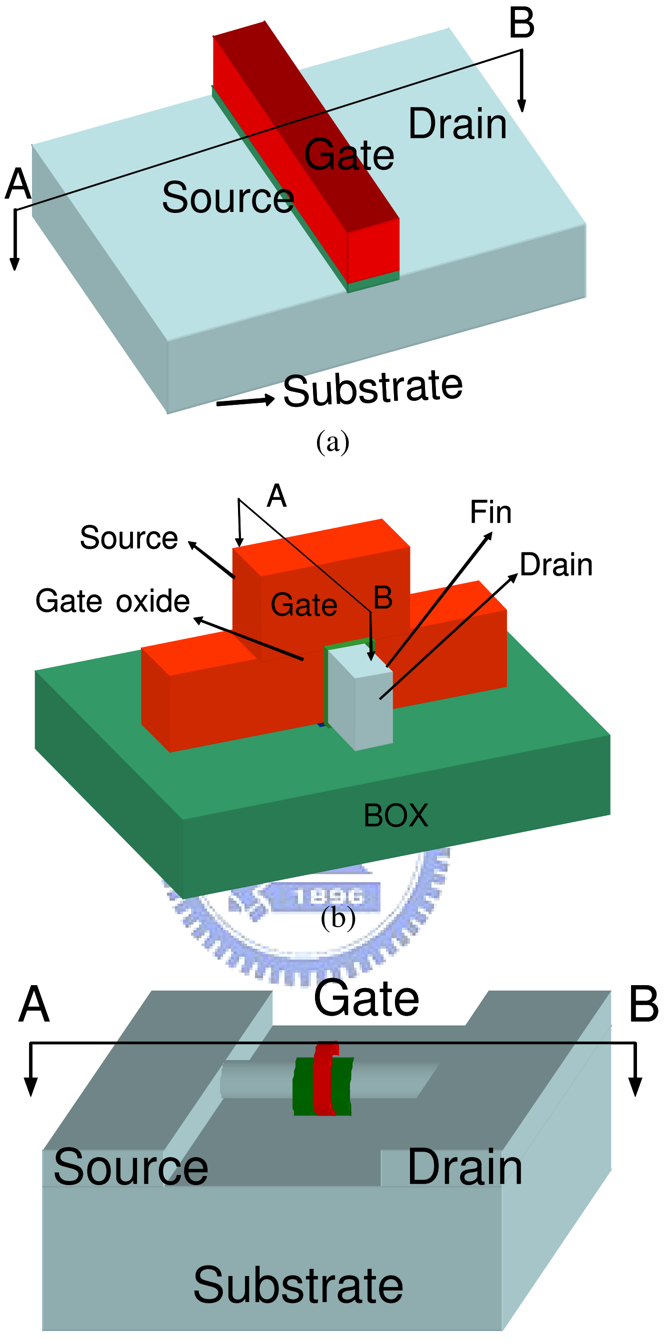
In this chapter, the results of device electrical characteristics of the adopted devices will be discussed. The simulated device structures will be presented in section 6.1. It consists of three different device structures, the planar MOSFETs the SOI FinFET, and the nanowire FinFET. The performance of the planar MOSFET, the SOI FinFET, and the nanowire FinFET will be shown in section 6.2, 6.3, and 6.4, respectively. The performance includes the intrinsic and terminal characteristics. Finally, we extract short channel effect parameters of different device structures and compare their performance in section 6.5.

6.1 Device Structures

Without lost of generality, we will analyze three different device structures, where consist of the planar MOSFET, the SOI FinFET, and the nanowire FinFET. The planar MOSFET is the planar structure, but the SOI and nanowire FinFET are non-planar structures. Fig. 6.1 is 3D illustrations of the device structures, the planar MOSFET, the SOI FinFET and the nanowire FinFET, respectively. The device dimension parameters and doping profile concentration will be shown in Table 6.1. We note non-planar structures, the SOI FinFET and the nanowire FinFET, are un-doped channel in our simulations. The use of a lightly doped or un-doped channel is desirable for immunity against dopant-fluctuation effects which give rise to threshold-voltage variation and also for reduced drain-to-body capacitance and high carrier mobility which provide for improved circuit performance. The polygate is used in planar structures (the planar MOSFET), and the concentration of polygate is 5e+20 cm^{-3}. For non-planar structures (the SOI and nanowire FinFET), we will use the mental gate, where the workfunction is set to be 4.6 ev.

Table 6.1: Device parameters used for our simulations.

	Planar MOSFET	SOI FinFET	Nanowire FinFET	Unit
Gate oxide thickness	1.5	1.5	1.5	nm
Channel width	0.05	0.01	–	μm
Fin height	–	20	–	nm
Channel radius	–	–	8	nm
Channel doping concentration	3.2e+18	3e+16	3e+16	cm^{-3}
Source/Drain doping concentration	3e+20	3e+20	3e+20	cm^{-3}



(a)

(b)

(c)

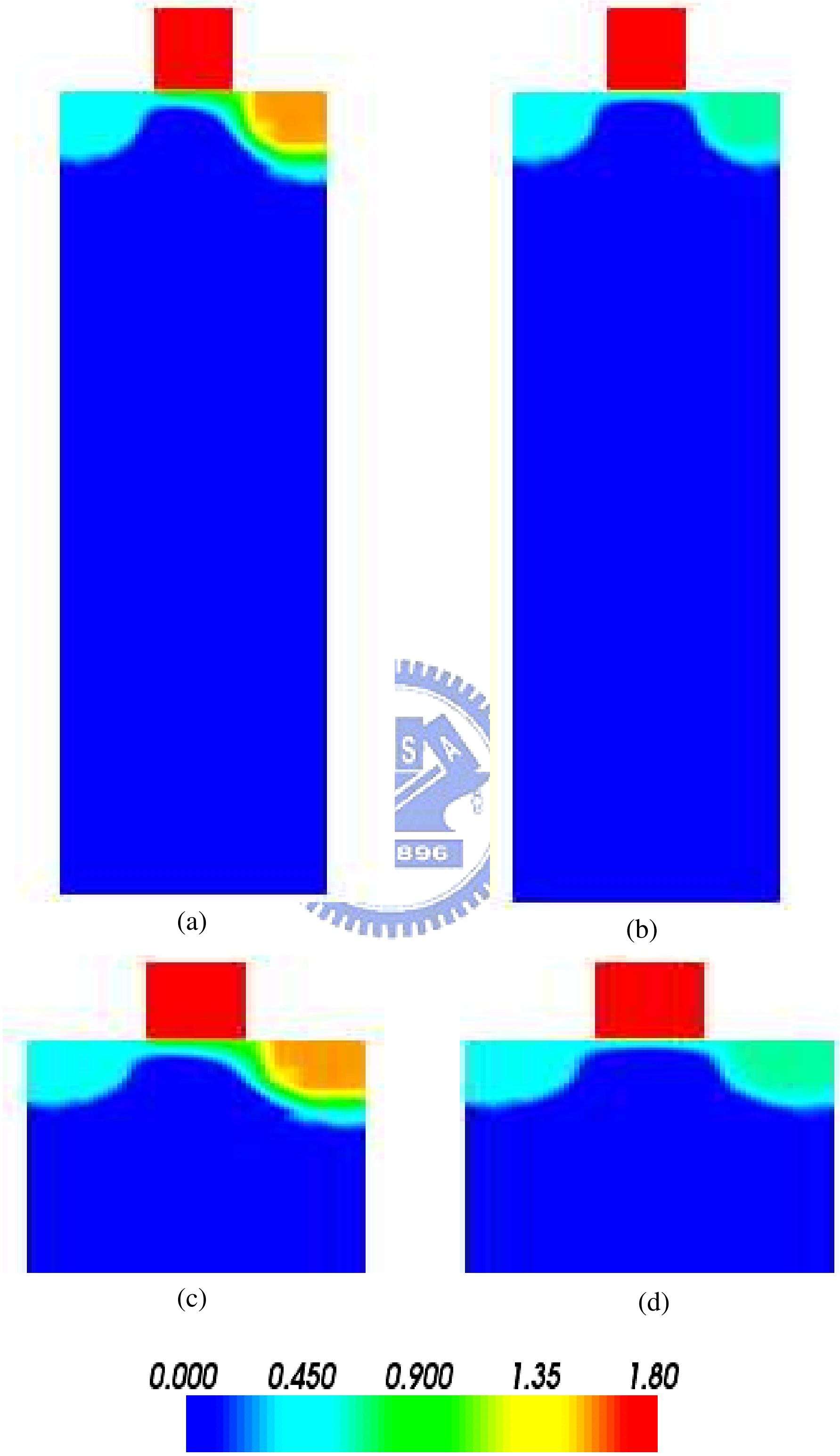
Figure 6.1: 3D illustrations of the device structure (a) the planar MOSFET, (b) the SOI FFET and (c) the nanowire FinFET. We note the SOI FinFET is fabricated on the insulator.

6.2 Performance of the Planar MOSFET

In this section, we explore the performance of the planar MOSFET. We will analyze device intrinsic and terminal characteristics under different technology generation nodes, 16 nm, 22 nm, 32 nm, 45 nm, 65 nm, 90 nm, and 130 nm. Intrinsic characteristics consist of the electrical potential, electrical field, and carrier density under both on-state and off-state. The I_D-V_G will be shown in terminal characteristics.

6.2.1 Intrinsic Characteristics

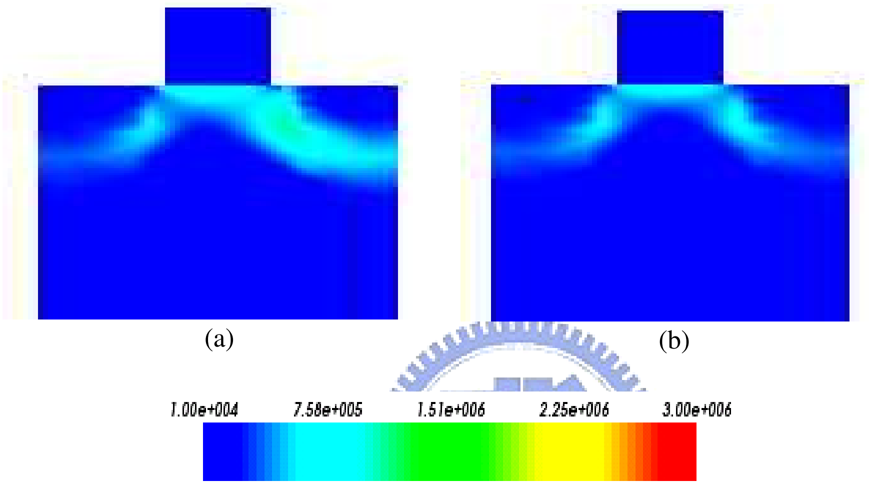
We will show the cross-section views of the simulated electrostatic potential, electric field, and carrier density for the 32 nm planar MOSFET under on-state and off-state. Plots are along the cutting planes of AB, shown in Fig. 6.1(a). Under the on-state, the device' bias is V_G = 1.2 V and V_D = 1.0 V. The device' bias is V_G = 1.2 V and V_D = 0.1 V under the off-state. The plot of electrostatic potential under both on-state and off-state will be shown in Fig. 6.2. The plot of electric field under both on-state and off-state will be shown in Fig. 6.3. The plot of electron density under both on-state and off-state will be shown in Fig. 6.4. The plot of hold density under both on-state and off-state will be shown in Fig. 6.5.



(a) (b)

(c) (d)

Figure 6.2: Cross-section views of the electrostatic potential for the 32 nm planar MOSFET under (a) on-state, (b) off-state, (c) on-state near junction regions and (d) off-state near junction regions.



(a) (b)

Figure 6.3: Cross-section views of the electric field near junction regions for the 32 nm planar MOSFET under (a) on-state and (b) off-state.

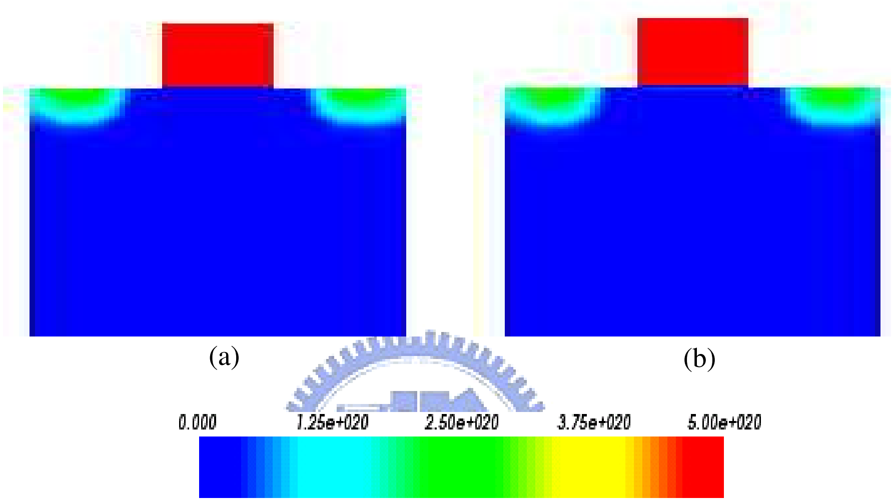


Figure 6.4: Cross-section views of the electron density near junction regions for the 32 nm planar MOSFET under (a) on-state and (b) off-state.

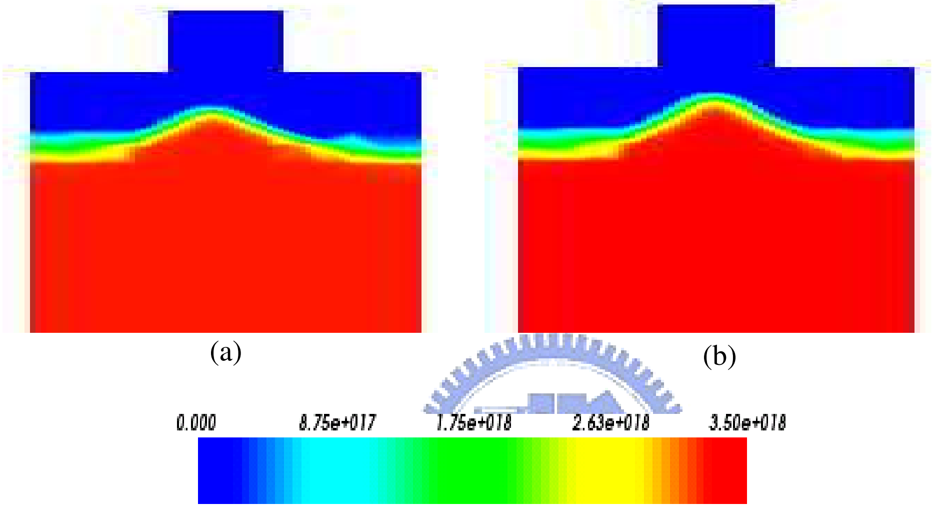

(a)
(b)

Figure 6.5: Cross-section views of the hole density near junction regions for the 32 nm planar MOSFET under (a) on-state and (b) off-state.

6.2.2 Terminal Characteristics

We simulate the device terminal characteristics at different technology generation nodes. We have seven technology generation nodes, and they (channel lengths) are 16 nm, 22 nm, 32 nm, 45 nm, 65 nm, 90 nm, and 130 nm. Fig. 6.6 will show the I_D-V_G characteristic curves for the planar MOSFET with seven different generation nodes. The drain voltage of the planar MOSFET is applied to 0.1 V and 1 V, and gate voltage is biased from 0 V to 1.2 V. The terminals of the source and substrate are biased grounded. We can extract short channel parameters, such as threshold voltage, subthreshold swing (S.S.), drain induced barrier lowering (DIBL), on/off current ratio from terminal characteristics. Then we compare and analyze the short channel effect for different device structures.

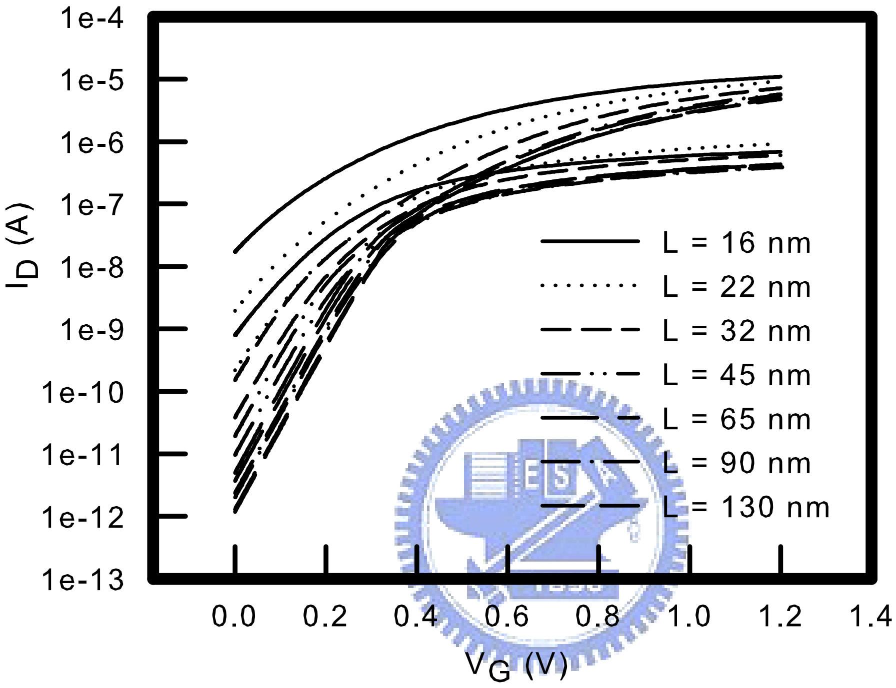


Figure 6.6: The I_D-V_G curve of the planar MOSFET with different channel lengths.

6.3 Performance of the SOI FinFET

In this section, we explore the performance of the SOI FinFET. We will analyze device intrinsic and terminal characteristics under different technology generation nodes, 16 nm, 22 nm, 32 nm, 45 nm, 65 nm, 90 nm, and 130 nm. Intrinsic characteristics consist of the electrical potential, electrical field, and carrier density under both on-state and off-state. The I_D-V_G will be shown in terminal characteristics.

6.3.1 Intrinsic Characteristics

We will show the cross-section views of the simulated electrostatic potential, electric field, and carrier density for the 32 nm SOI FinFET under on-state and off-state. Plots are along the cutting planes of AB, shown in Fig. 6.1(b). Under the on-state, the device' bias is V_G = 1.2 V and V_D = 1.0 V. The device' bias is V_G = 1.2 V and V_D = 0.1 V under the off-state. The plot of electrostatic potential under both on-state and off-state will be shown in Fig. 6.7. The plot of electric field under both on-state and off-state will be shown in Fig. 6.8. The plot of electron density under both on-state and off-state will be shown in Fig. 6.9. The plot of hold density under both on-state and off-state will be shown in Fig. 6.10.

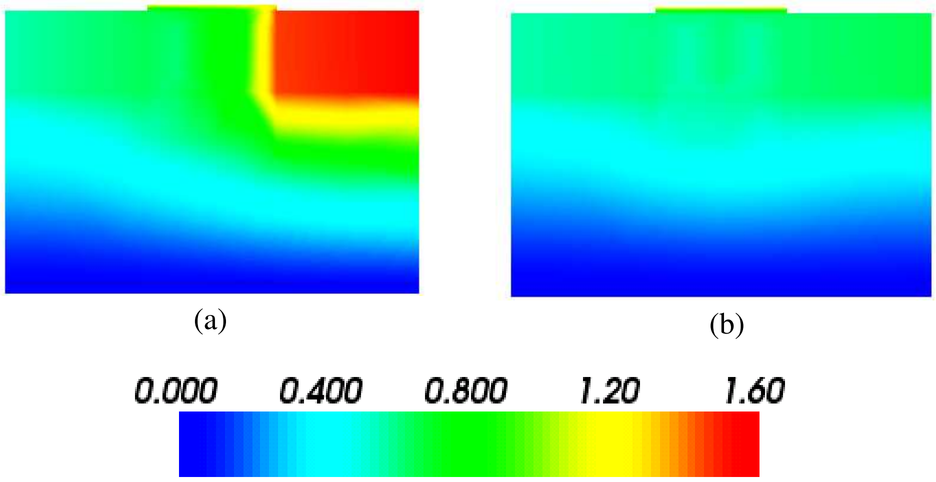

(a) (b)

Figure 6.7: Cross-section views of the electrostatic potential for the 32 nm SOI FinFET under (a) on-state and (b) off-state.

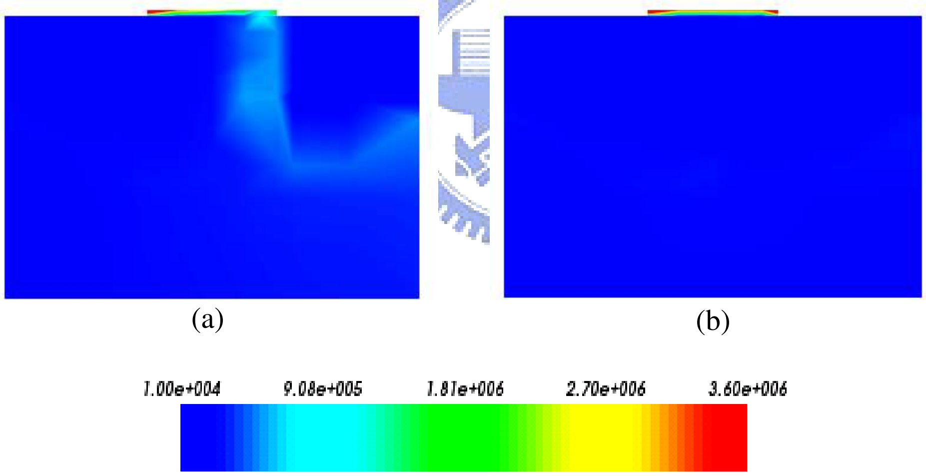
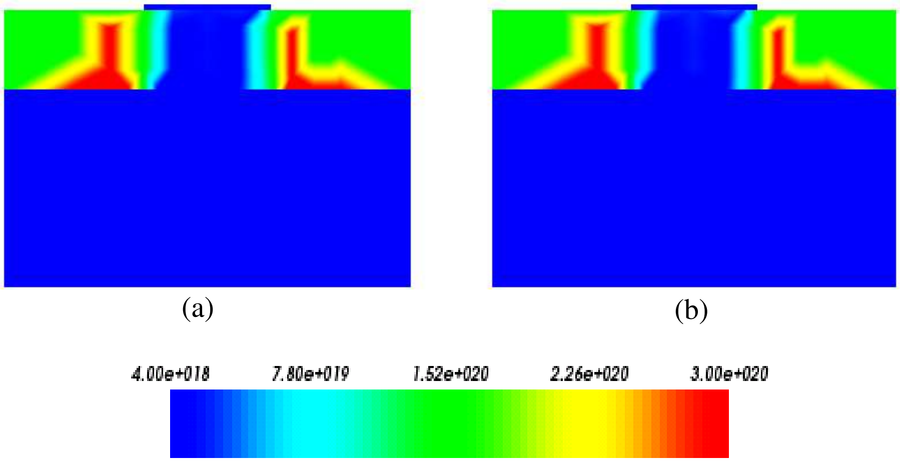
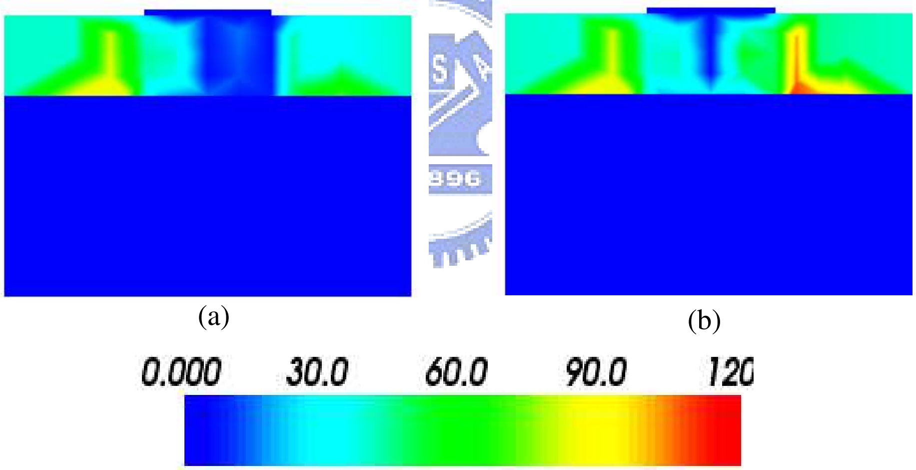

(a) (b)

Figure 6.8: Cross-section views of the electric field for the 32 nm SOI FinFET under (a) on-state and (b) off-state.



(a) (b)

Figure 6.9: Cross-section views of the electron density for the 32 nm SOI FinFET under (a) on-state and (b) off-state.



(a) (b)

Figure 6.10: Cross-section views of the hole density for the 32 nm SOI FinFET under (a) on-state and (b) off-state.

6.3.2 Terminal Characteristics

We simulate the device terminal characteristics at different technology generation nodes. We have seven technology generation nodes, and they (channel lengths) are 16 nm, 22 nm, 32 nm, 45 nm, 65 nm, 90 nm, and 130 nm. Fig. 6.6 will show the I_D-V_G characteristic curves for the SOI FinFET with seven different generation nodes. The drain voltage of the SOI FinFET is applied to 0.1 V and 1 V, and gate voltage is biased from 0 V to 1.2 V. The terminal of source is biased grounded. We can extract short channel parameters, such as threshold voltage, subthreshold swing (S.S.), drain induced barrier lowering (DIBL), on/off current ratio from terminal characteristics. Then we compare and analyze the short channel effect for different device structures.

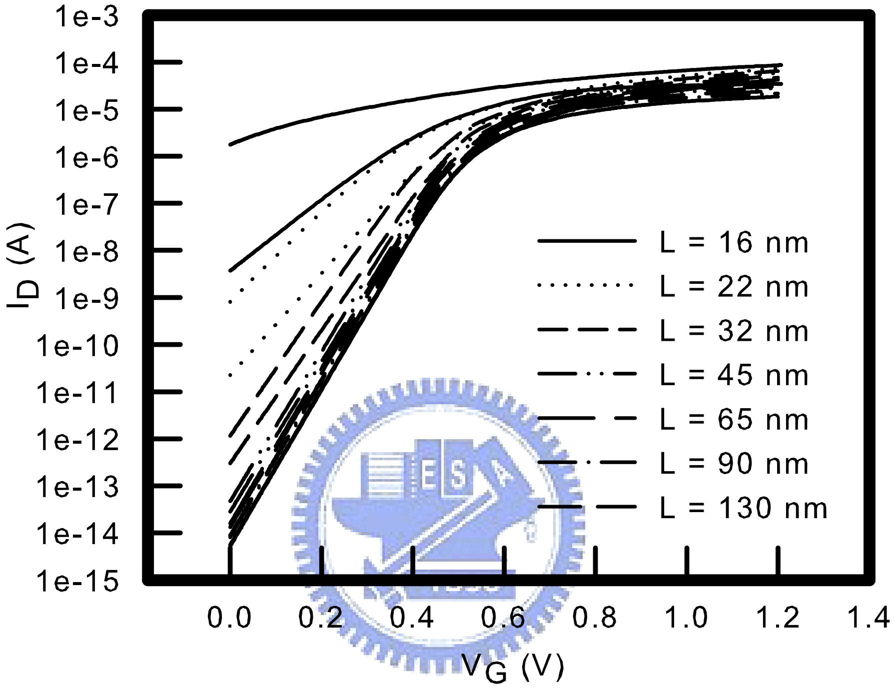


Figure 6.11: The I_D-V_G curve of the SOI FinFET with different channel lengths.

6.4 Performance of the Nanowire FinFET

In this section, we explore the performance of the nanowire FinFET. We will analyze device intrinsic and terminal characteristics under different technology generation nodes, 16 nm, 22 nm, 32 nm, 45 nm, 65 nm, 90 nm, and 130 nm. Intrinsic characteristics consist of the electrical potential, electrical field, and carrier density under both on-state and off-state. The I_D-V_G will be shown in terminal characteristics.

6.4.1 Intrinsic Characteristics

We will show the cross-section views of the simulated electrostatic potential, electric field, and carrier density for the 32 nm nanowire FinFET under on-state and off-state. Plots are along the cutting planes of AB, shown in Fig. 6.1(c). Under the on-state, the device' bias is V_G = 1.2 V and V_D = 1.0 V. The device' bias is V_G = 1.2 V and V_D = 0.1 V under the off-state. The plot of electrostatic potential under both on-state and off-state will be shown in Fig. 6.12. The plot of electric field under both on-state and off-state will be shown in Fig. 6.13. The plot of electron density under both on-state and off-state will be shown in Fig. 6.14. The plot of hold density under both on-state and off-state will be shown in Fig. 6.15.

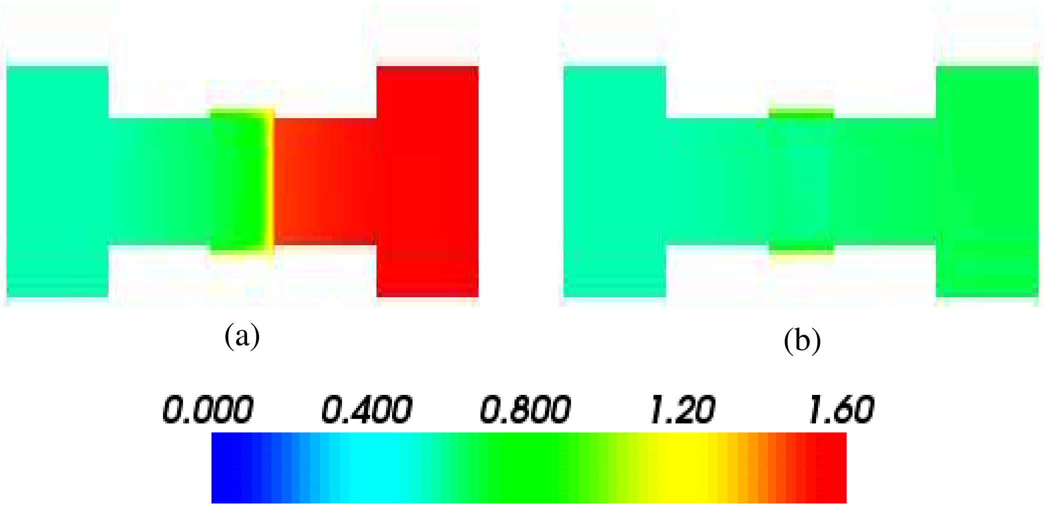

(a) (b)

Figure 6.12: Cross-section views of the electrostatic potential for the 32 nm nanowire FinFET under (a) on-state and (b) off-state.

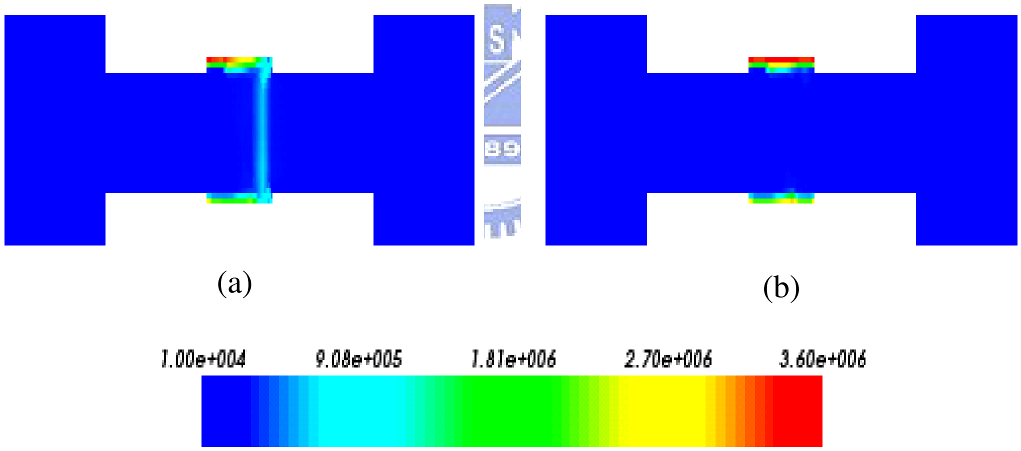
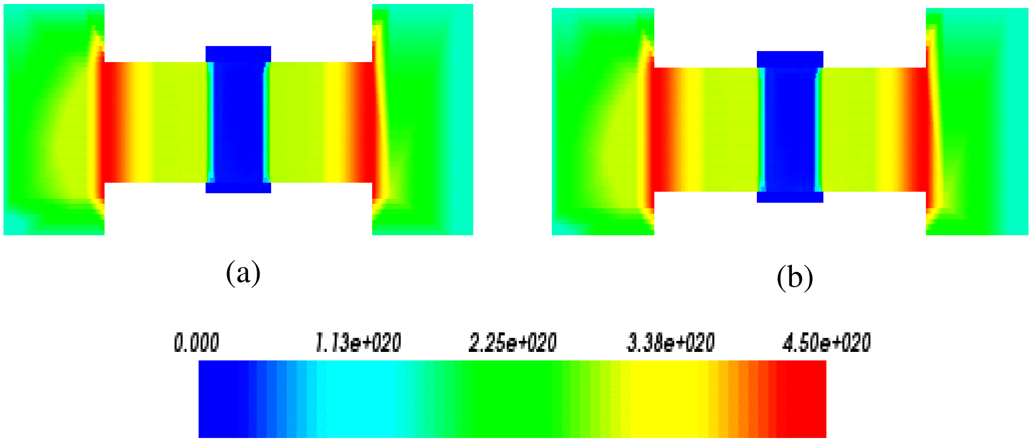
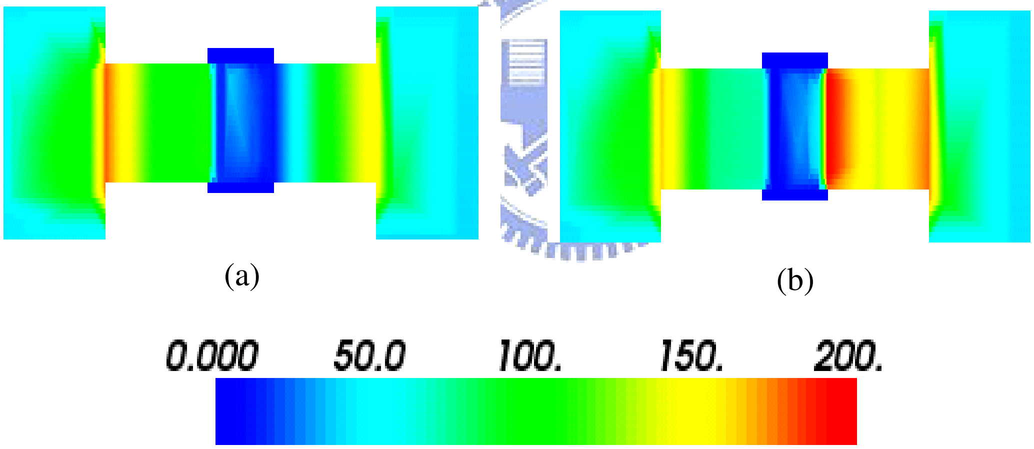

(a) (b)

Figure 6.13: Cross-section views of the electric field for the 32 nm nanowire FinFET under (a) on-state and (b) off-state.



(a) (b)

Figure 6.14: Cross-section views of the electron density for the 32 nm nanowire FinFET under (a) on-state and (b) off-state.



(a) (b)

Figure 6.15: Cross-section views of the hole density for the 32 nm nanowire FinFET under (a) on-state and (b) off-state.

6.4.2 Terminal Characteristics

We simulate the device terminal characteristics at different technology generation nodes. We have seven technology generation nodes, and they (channel lengths) are 16 nm, 22 nm, 32 nm, 45 nm, 65 nm, 90 nm, and 130 nm. Fig. 6.6 will show the I_D-V_G characteristic curves for the nanowire FinFET with seven different generation nodes. The drain voltage of the nanowire FinFET is applied to 0.1 V and 1 V, and gate voltage is biased from 0 V to 1.2 V. The terminal of source is biased grounded. We can extract short channel parameters, such as threshold voltage, subthreshold swing (S.S.), drain induced barrier lowering (DIBL), on/off current ratio from terminal characteristics. Then we compare and analyze the short channel effect for different device structures.

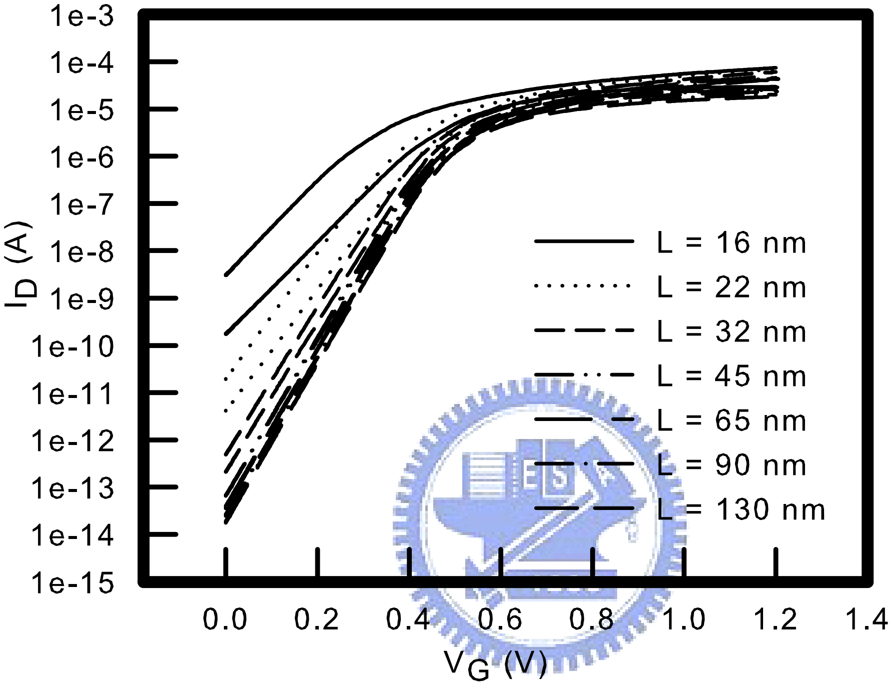


Figure 6.16: The I_D-V_G curve of the nanowire FinFET with different channel lengths.

6.5 Comparisons of the Performance

The I_D-V_G characteristics of the transistors (the planar MOSFET, SOI FinFET, and nanowire FinFET) with channel length is 32 nm are compared, shown in Fig. 6.17. The I_D-V_D characteristics, shown in Fig. 6.18, of the planar MOSFET, SOI FinFET, and nanowire FinFET with channel length is 32 nm are also compared. The gate voltage is applied to 0.2 V, 0.4 V, 0.6 V, 0.8 V, and 1 V. The nanowire FinFET, shown in Fig. 6.18, has higher output current compared to the planar MOSFET and SOI FinFET.

To explore the short channel effect with different device structures, we extract the short channel effect parameters, threshold voltage (V_{th}), subthreshold seing (S.S.), drain induced barrier lowering (DIBL), and on/off current ratio (I_{on}/I_{off}) from the I_D-V_G characteristics. The threshold voltage (V_{th}), subthreshold seing (S.S.), drain induced barrier lowering (DIBL), and on/off current ratio (I_{on}/I_{off}) versus channel length with the planar MOSFET, the SOI FinFET, and the nanowire FinFET are depicted in Fig. 6.19 - Fig. 6.22, respectively. Table 6.2 shows the value of short channel effect parameters for the planar MOSFET, the SOI FinFET, and the nanowire FinFET with channel length is 32 nm. Among these devices, the nanowire FinFET shows much less threshold voltage and drain induced barrier lowering than the planar MOSFET and SOI FinFET. Besides, the nanowire FinFET has better subthreshold swing and on/off current ratio than the planar MOSFET and SOI FinFET. Our calculation confirms the improvement of DIBL in nanowire FinFET,

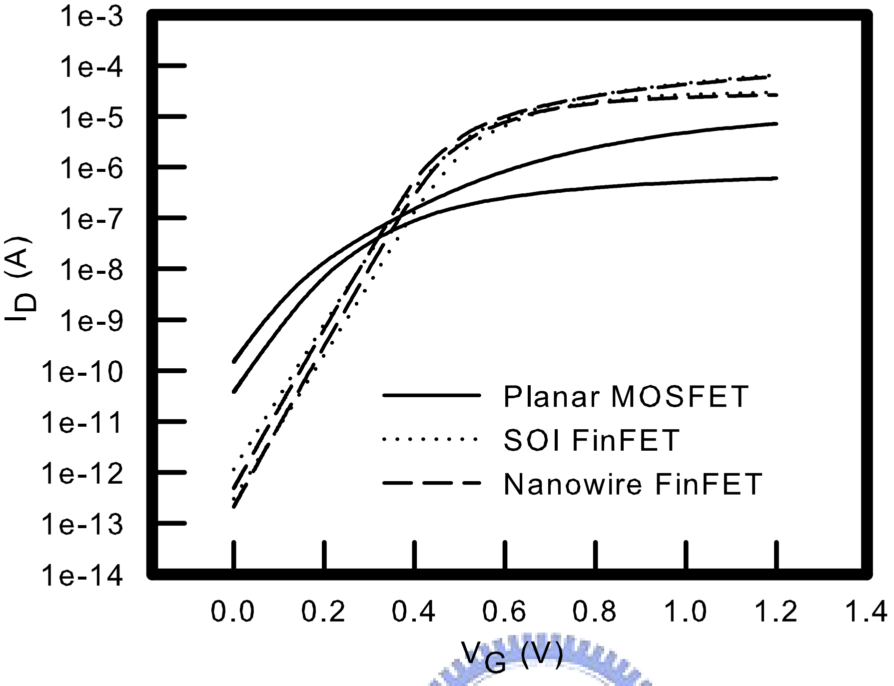


Figure 6.17: The I_D-V_G curve of the planar MOSFET, SOI FinFET, and nanowire FinFET with channel length is 32 nm. The latest device structure shows good device performance.

which provides interesting application in SRAM cells [52] - [55]. The nanowire FinFET also shows the smallest leakage current among three devices. It overcomes the limit of scaling issue and inherently have good suppression of short-channel effects (SCEs), and ideal sub-threshold swing (S.S.), large driving current, small leakage current, small DIBL, and good area efficiency.

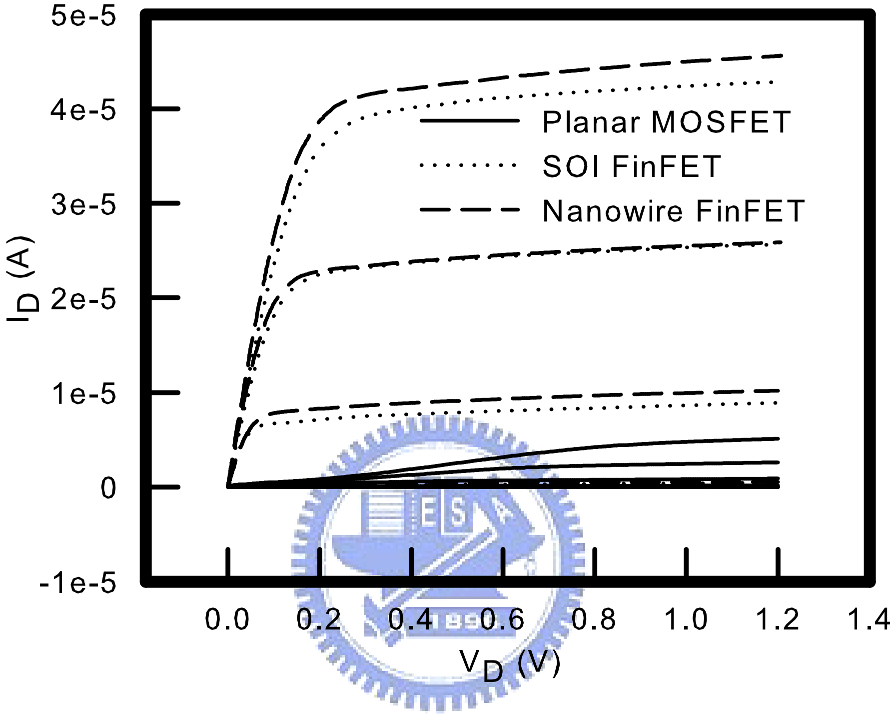


Figure 6.18: The simulated characteristics of I_D-V_D curve for the planar MOSFET, SOI FinFET, and nanowire FinFET with channel length is 32 nm. The latest device structure shows higher output current.

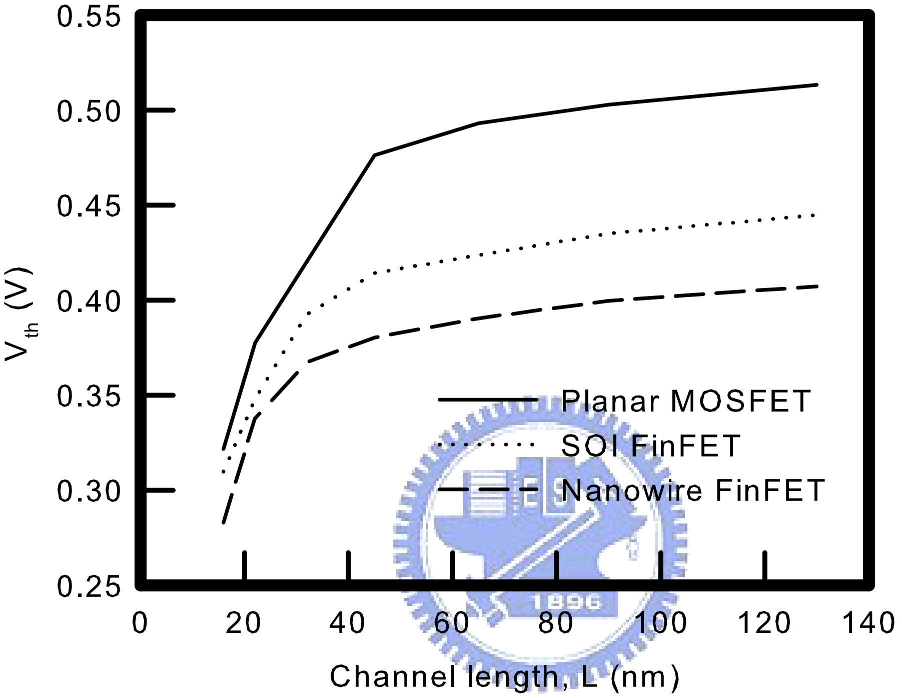


Figure 6.19: The threshold voltage (V_{th}) versus channel length with the planar MOSFET, the SOI FinFET, and the nanowire FinFET.

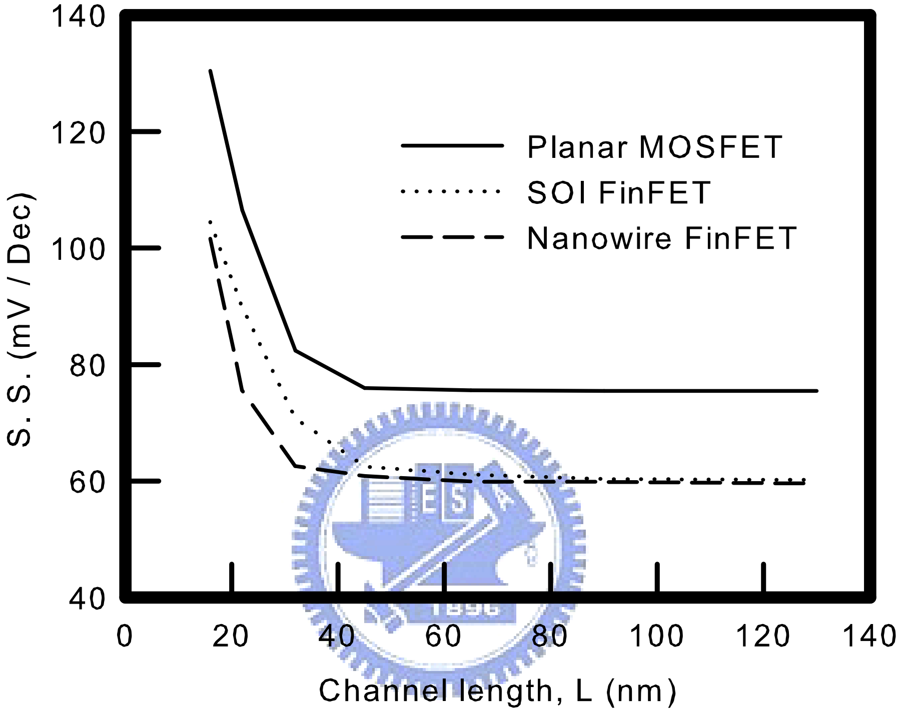


Figure 6.20: The subthreshold swing (S.S.) versus channel length with the planar MOSFET, the SOI FinFET, and the nanowire FinFET.

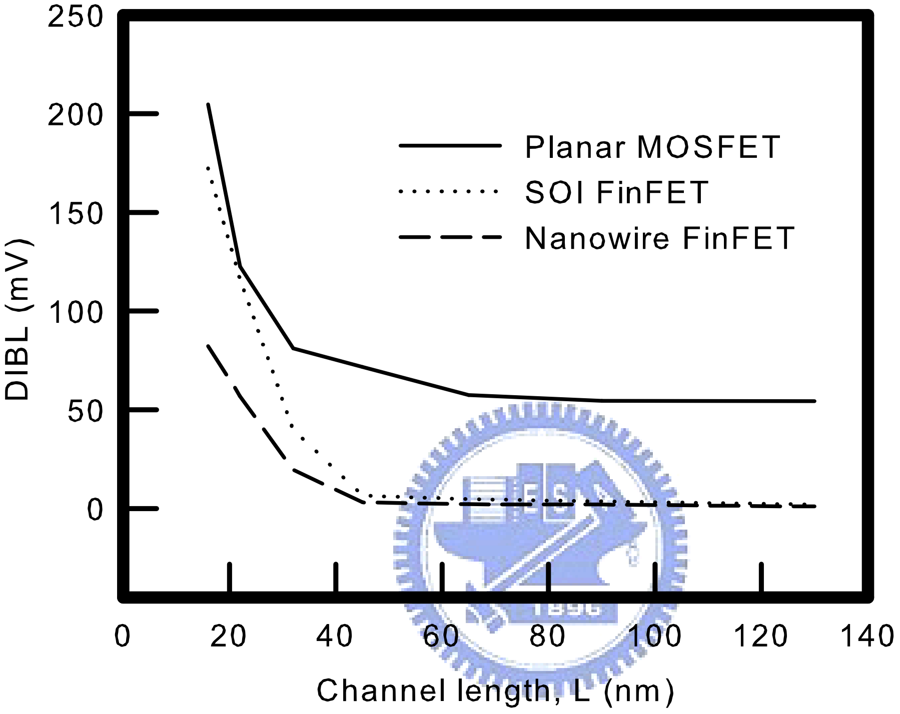


Figure 6.21: The drain induced barrier lowering (DIBL) versus channel length with the planar MOSFET, the SOI FinFET, and the nanowire FinFET.

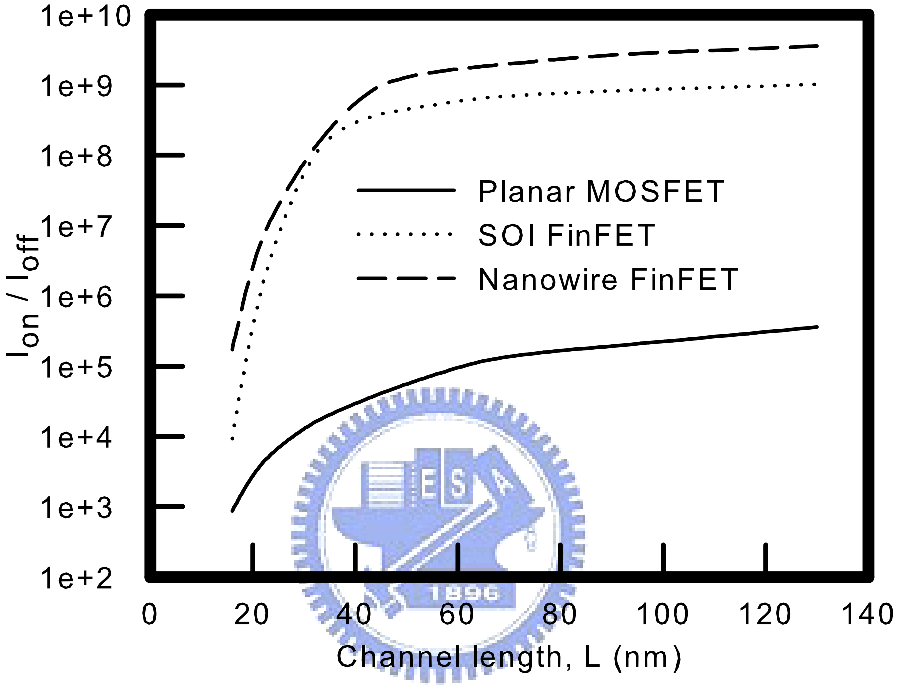


Figure 6.22: The on/off current ratio (Ion/Ioff) versus channel length with the planar MOSFET, the SOI FinFET, and the nanowire FinFET.

Table 6.2: The value of short channel parameters for the planar MOSFET, the SOI FinFET, and the nanowire FinFET with channel length is 32 nm.

	Planar MOSFET	SOI FinFET	Nanowire FinFET	Unit
Threshold voltage (Vth)	0.4203	0.3924	0.3674	V
Subthreshold swing (S.S.)	82.3731	62.4752	60.7809	mV/Dec
Drain induced barrier lowering (DIBL)	81.0123	39.8768	19.4531	mV
On/off current ratio (Ion/Ioff)	1.581e+4	9.923e+7	1.24e+8	1

6.6 Summary

We simulate the intrinsic and terminal characteristics of the planar MOSFET, the SOI FinFET, and the nanowire FinFET with respect to different technology generation nodes. For intrinsic characteristics, the electrostatic potential, electric field, electron and hole density for different device structures under on-state and off-state will be explored. We also explore and compare the device terminal. The short channel effect parameters, threshold voltage, subthreshold swing, drain induced barrier lowering, and on/off current ratio are extracted from the terminal characteristics (the I_D-V_G curve). It is found that the nanowire FinFET device has the best performance and much immunity to short channel effects (SCEs). The Nanowire FinFET will be emerged as the leading candidate to replace the planar MOSFET for scaling to the end of the roadmap. It operation provides superior short-channel control compared with conventional planar MOS devices.

Chapter 7

The Static Noise Margin of SRAM Cells

In this chapter, results of the stability for SRAM cells with planar MOSFETs, SOI FinFETs, and nanowire FinFETs will be shown. Stability, the immunity of the cell to flipping during a hold or read operation, is characterized by static noise margin (SNM). The result of SNM during both hold and read mode with respect to supply voltage is discussed in section 7.1. Section 7.2 states the result of SNM with respect to operation temperature. It is known that the SNM will decrease with increase of temperature. Finally, the result of SNM during both hold and read mode with respect to cell ratio is presented in section 7.3.

7.1 Effects of the Supply Voltage on the SNM

SNM for a bitcell with ideal voltage transfer characteristics (VTCs) is still limited to $V_{DD}/2$ because of the two sides of the butterfly curves. An upper limit on the changes in SNM with V_{DD} is thus $1/2$. The slope of the curves confirm that less than $1/2$ of V_{DD} noise will translate into SNM changes. Fig. 7.1 shows how SNM varies with Vdd for both hold and read mode, where the cell ratio is set to be 1. It is known that V_{DD} reduction induces the significant degradation of SRAM cells. As the V_{DD} increases, the SNM increase linearly for high-V_{th} SRAM cells and at high V_{DD} the noise margin levels off and may even decrease. This occurs when $V_{DD} \simeq 2V_{th}$. The 6-T-nanowire-FinFET-based SRAM cell enjoys much better stability, especially at higher supply voltage. The SNM depends on the choice of the V_{th} for the transistors used in the SRAM cells. A high V_{th} means that drive current of these devices is small making the operation (both accidental and intensional) more difficult, thus increasing the SNM. Thus, one approach to achieve a low power cell with high stability is to use high V_{th} devices at the cost of performance. FinFETs provide with a high device currebt even with larger V_{th} thereby achieving high noise margins along with good stability [57].

Table. 7.1 - 7.2 shows the SNM improvement percentage under hold and read modes of the SRAM cell using SOI FinFETs and nanowire FinFETs compared with the SRAM cell with planar MOSFETs, respectively. The definition of SNM improvement percentage is presented the following :

$$\text{SNM improvement percentage} = \frac{\text{SNM(A)} - \text{SNM(B)}}{\text{SNM(B)}} \tag{7.1}$$

where A is the SOI FinFET or the nanowire FinFET, and B is the planar MOSFET. For the hold mode, shown in Table. 7.1, SNM of the SRAM cell with SOI FinFETs for hold mode is improved more than 15 %, compared with the SRAM cell with planar MOSFETs, and about 18 % during read mode, respectively. For the hold mode, shown in Table. 7.2, SNM of the SRAM cell with nanowire FinFETs for hold mode is improved more than 23 %, compared with the SRAM cell with the planar MOSFETs, and about 30 % during read mode, respectively.

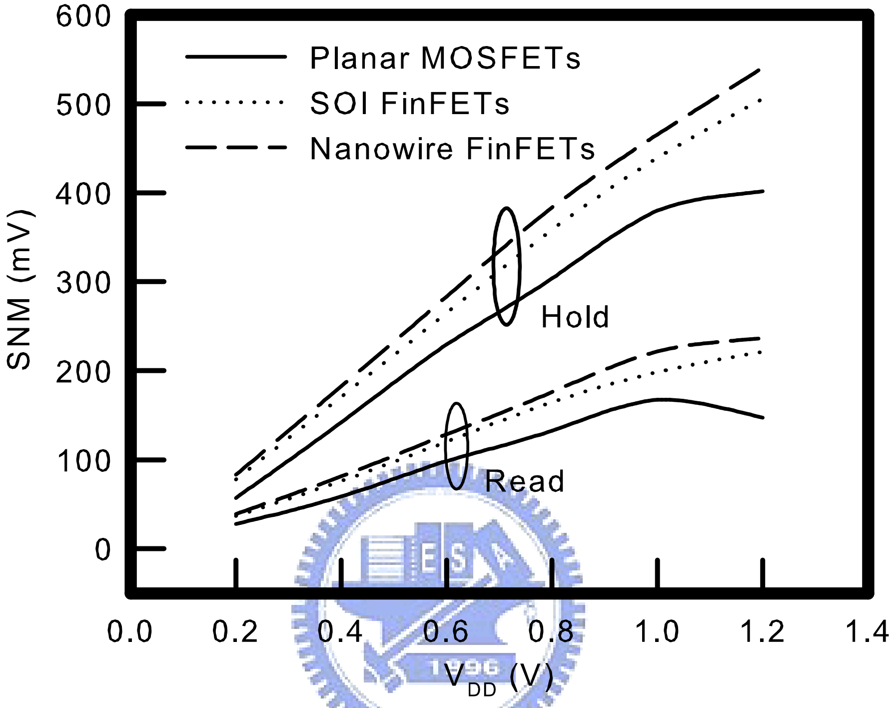


Figure 7.1: Plot of the SNM versus V_{DD} for SRAM cells using three different device structures under hold and read modes. The stability of SRAM cells with 32 nm SOI and nanowire FinFETs is better than that of planar MOSFETs. Furthermore, the 6-T-nanowire-FinFET-based SRAM demonstrates the best stability.

Table 7.1: The SNM improvement percentage under hold and read modes of the SRAM cell using SOI FinFETs compared with the SRAM cell with planar MOSFETs.

	During hold mode	During read mode
$V_{DD} = 0.2$ V	35.93 %	31.52 %
$V_{DD} = 0.4$ V	20.20 %	29.57 %
$V_{DD} = 0.6$ V	15.46 %	21.92 %
$V_{DD} = 0.8$ V	18.00 %	24.10 %
$V_{DD} = 1.0$ V	15.62 %	18.54 %
$V_{DD} = 1.2$ V	25.85 %	50.31 %

Table 7.2: The SNM improvement percentage under hold and read modes of the SRAM cell using nanowire FinFETs compared with the SRAM cell with planar MOSFETs.

	During hold mode	During read mode
$V_{DD} = 0.2$ V	45.39 %	40.68 %
$V_{DD} = 0.4$ V	28.56 %	38.59 %
$V_{DD} = 0.6$ V	23.50 %	30.41 %
$V_{DD} = 0.8$ V	26.21 %	32.73 %
$V_{DD} = 1.0$ V	22.52 %	32.49 %
$V_{DD} = 1.2$ V	34.61 %	60.77 %

7.2 Effects of the Operation Temperature on the SNM

With the technology scaling, the transistor density in a chip is approximately doubled for each new technology generation. As a result, the average operating temperature of the chip has increased. The simulation results of SNM versus different operating temperature are presented in Fig. 7.2, where the cell ratio is set to be 1 and the supply voltage is 1 V. It is known that the SNM will decrease with increase of temperature. It is attributed to a decrease of the threshold voltage when the temperature increases. Table. 7.3 shows the SNM temperature dependence under hold and read modes of the SRAM cell using planar MOSFETs, SOI FinFETs and nanowire FinFETs, respectively. The definition of SNM temperature dependence is presented the following :

$$\text{SNM temperature dependence} = \frac{\text{SNM(T = 300 K)} - \text{SNM(T = 500 K)}}{\text{200K}} \tag{7.2}$$

In the hold operation, the SNM temperature dependence with three different device structures is reduced from 0.43 to 0.22 when the temperature varies from 300 K to 500 K. A similar reduction (from 0.28 to 0.21) of the SNM temperature dependence is observed in the mode of read. According to our calculation the variation of the threshold voltage versus the temperature is almost the same among three device structures. Nevertheless, the SNM of the SRAM with the 32 nm nanowire FinFETs is still less dependent upon the temperature due to good channel controllability and improved short channel effects, compared with the SRAM using the planar MOSFETs, in particular, for the SRAM under the hold mode.

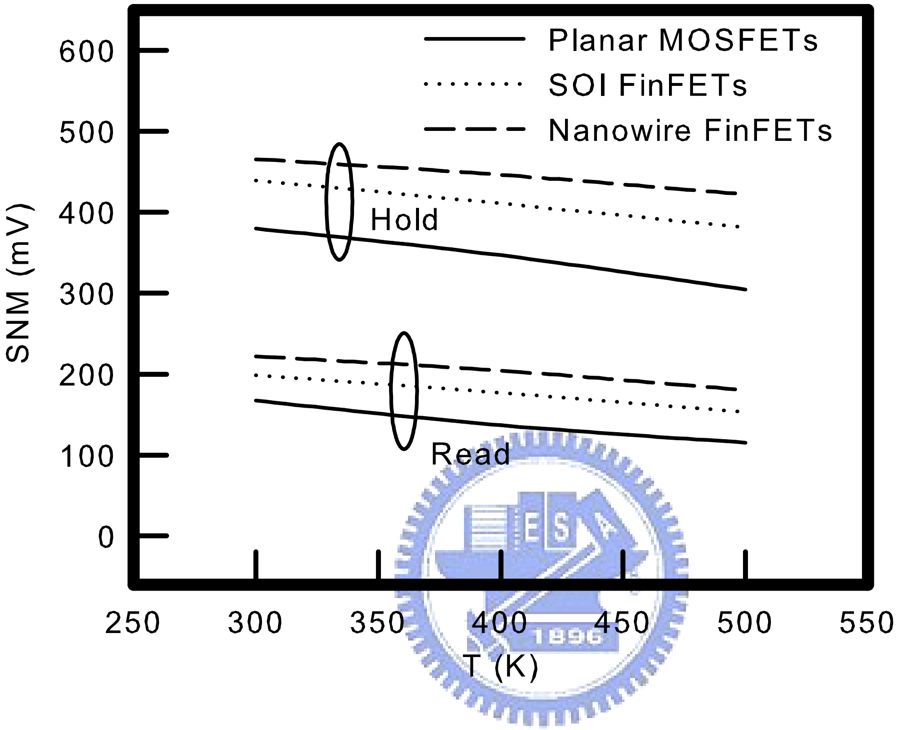


Figure 7.2: Plot of the SNM versus the operation temperature for the SRAM using three different device structures with channel length is 32 nm, where the hold and read modes are computed. The SRAM with the 32 nm nanowire FinFET demonstrates less temperature dependence.

Table 7.3: The SNM temperature dependence under hold and read modes of the SRAM cell using different device structures.

	Planar MOSFETs	SOI FinFETs	Nanowire FinFETs
Hold	0.38	0.29	0.22
Read	0.26	0.23	0.21

7.3 Effects of the Cell Ratio on the SNM

The functionality and density of a memory array are its most important properties. Functionality is guaranteed for large memory arrays by providing sufficiently large design, where are determined by device sizing (channel widths and channel lengths), the supply voltage, and temperature. Although upsizing the cell area and this increase the noise margins, it increases the cell area and thus lower the density. A careful sizing of the transistors for is required to avoid accidentally writing a logic ”1” into the cell while trying to read a stored logic ”0”, thus resulting in a read upset. The ratio of the widths of the pull-down transistor to the access transistor commonly referred to as the cell ratio (CR) determines how high the ”0” storage node rises during a read access [56].

$$CR = \frac{W_1/L_1}{W_2/L_2} = \frac{W_4/L_4}{W_5/L_5} \tag{7.3}$$

where the W_1 is the channel width of the transistor M1 (shown in Fig. 2.1), and the L_1 is the channel length of the transistor M1, respectively.

The SNM versus the V_{DD} for the SRAM using planar MOSFETs with different cell ratio during hold and read modes are presented Fig. 7.3 - Fig. 7.4, respectively. The SNM versus the V_{DD} for the SRAM using SOI FinFETs with different cell ratio during hold and read modes are presented Fig. 7.5 - Fig. 7.6, respectively. The SNM versus the V_{DD} for the SRAM using nanowire FinFETs with different cell ratio during hold and read modes are presented Fig. 7.7 - Fig. 7.8, respectively. The SNM will decrease during hold mode when increasing the cell ratio, but the SNM will increase during read mode when increasing the cell ratio.

Fig. 7.9 plots of the SNM versus the cell ratio for three different device structures, where the hold mode is compared. The supply voltage is set to be 1 V and temperature is 300 K (room temperature). It is founded that the SNM slightly reduces when the cell ratio increases under the hold mode for three different device structures. However, the change is insignificant due to an off-state operation of the transistors M2 and M5. The turned-off M2 and M5 keep the symmetry properties of the upper and lower squares in the butterfly curve and then maintain the original states at the nodes Q and QB (shown in Fig. 2.1). Therefore, an increase of the cell ratio can not significantly alter the SNM for the SRAM under the hold mode. As upsizing cell ratio, the upper and lower squares in the butterfly curve will change from symmetry to asymmetry. Then it makes the SNM decreasing when increasing cell ratio under hold mode.

Fig. 7.10 plots of the SNM versus the cell ratio for three different device structures, where the read mode is compared. The supply voltage and temperature conditions are as same as Fig. 7.9. The SNM increases when the cell ratio increases under read mode for three different device structures, especially at high voltage. It is a direct result due to high current resulting from the high cell ratio in the transistor M1, where the node Q is assumed at the logic ”1” and the node QB is at the logic ”0” (Q and QB are set to be the initial logic state). Smaller cell ratios translate into a bigger voltage drop across the pull-down transistor (M1), requiring a smaller noise voltage at the ”0” node to trip the cell. Then the SNM of SRAM cells will reduce. Because the SNM under the mode of read is the worst case, we will increase stability of SRAM cells by upsizing the cell ratio.

An operational six-transistor SRAM cell was experimentally demonstrated using Double Gate CMOS FinFET technology [58]. A cell size of 4.8 μm^2 was achieved in 180 nm node technology, with stable operation at 1.5 V using a single level of copper interconnect. The 6-T SRAM cell was experimentally demonstrated using bulk FinFET CMOS technology [59]. A cell size of 0.79 μm^2 was achieved by 90 nm node technology, with stable operation at 1.2 V using 4 levels of W and AI interconnect. Static noise margin of 280 mV was obtained at 1.2 V. Our simulation results confirm the improvement of stability for the 6-T SRAM cell using FinFETs.

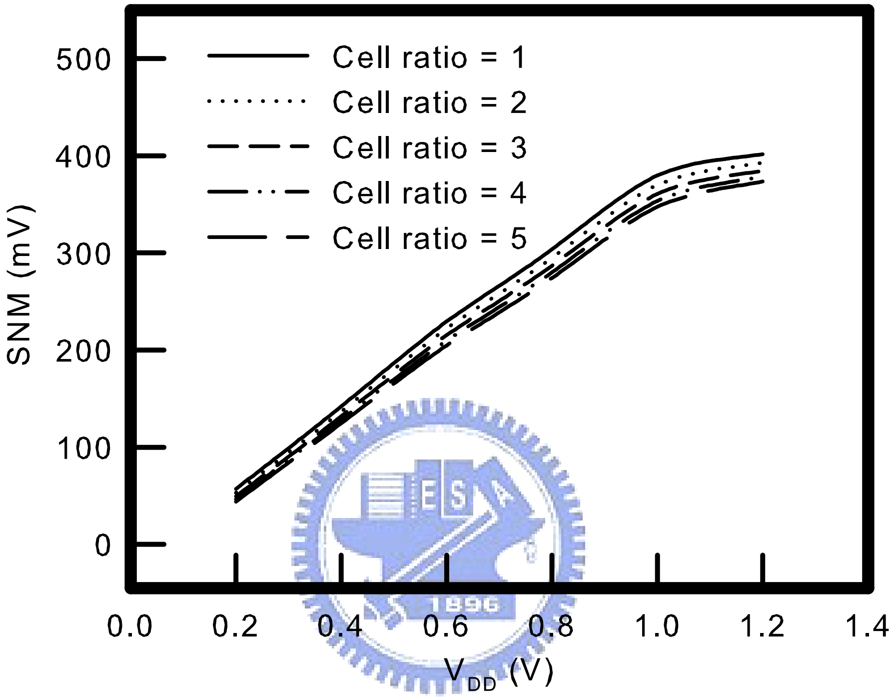


Figure 7.3: Plot of the SNM versus the V_{DD} for the SRAM cell using planar MOSFETs with different cell ratio, where the hold mode is computed.

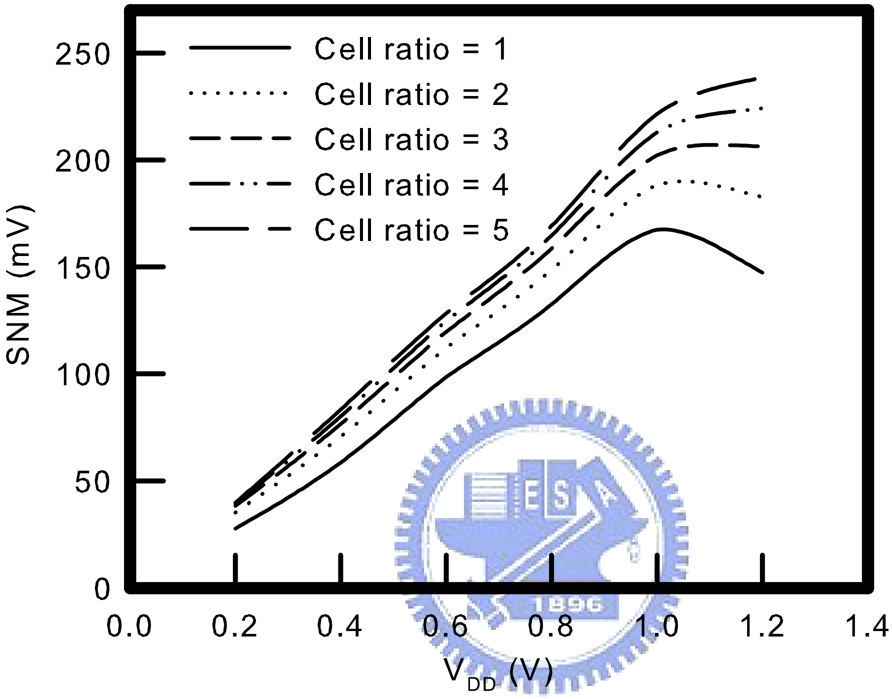


Figure 7.4: Plot of the SNM versus the V_{DD} for the SRAM cell using planar MOSFETs with different cell ratio, where the read mode is computed.

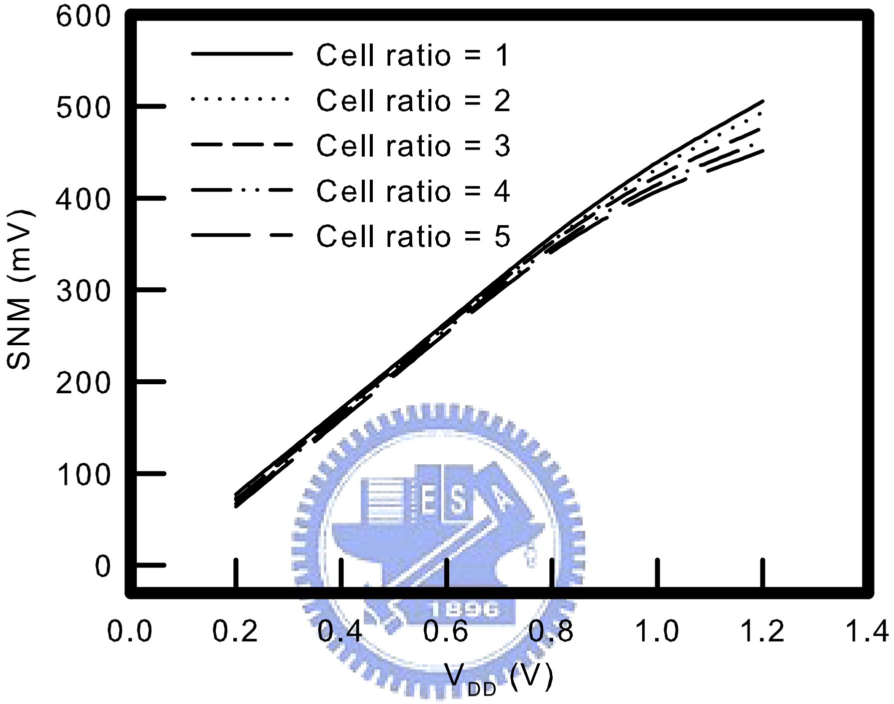


Figure 7.5: Plot of the SNM versus the V_{DD} for the SRAM cell using SOI FinFETs with different cell ratio, where the hold mode is computed.

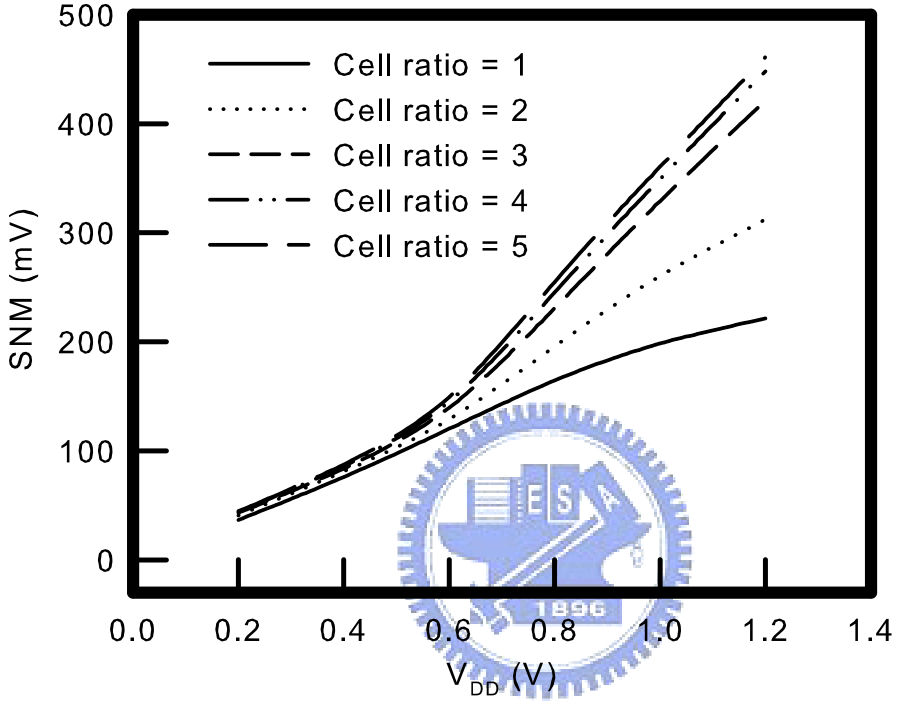


Figure 7.6: Plot of the SNM versus the V_{DD} for the SRAM cell using SOI FinFETs with different cell ratio, where the read mode is computed.

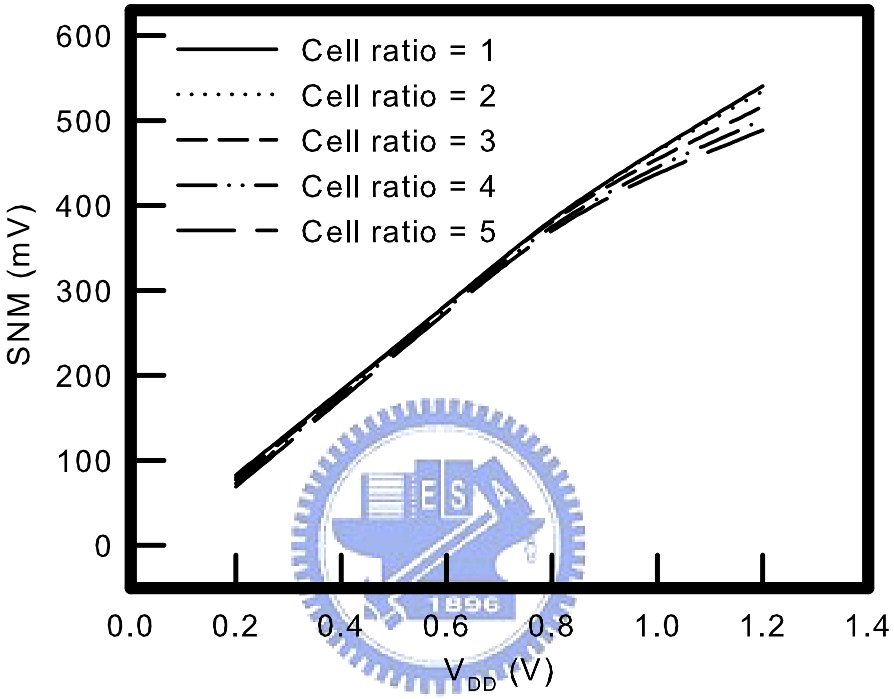


Figure 7.7: Plot of the SNM versus the V_{DD} for the SRAM cell using nanowire FinFETs with different cell ratio, where the hold mode is computed.

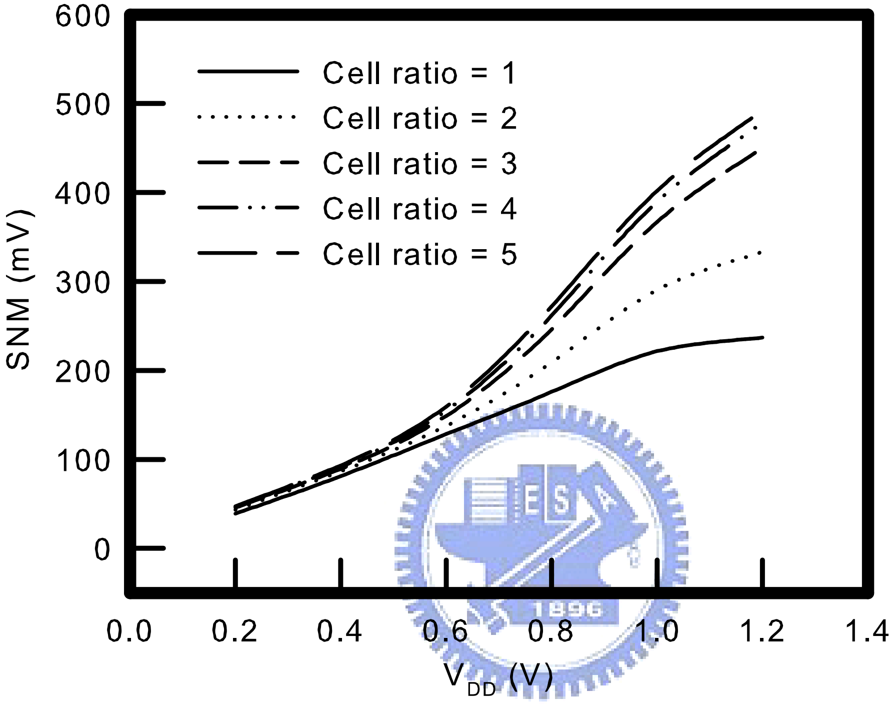


Figure 7.8: Plot of the SNM versus the V_{DD} for the SRAM cell using nanowire FinFETs with different cell ratio, where the read mode is computed.

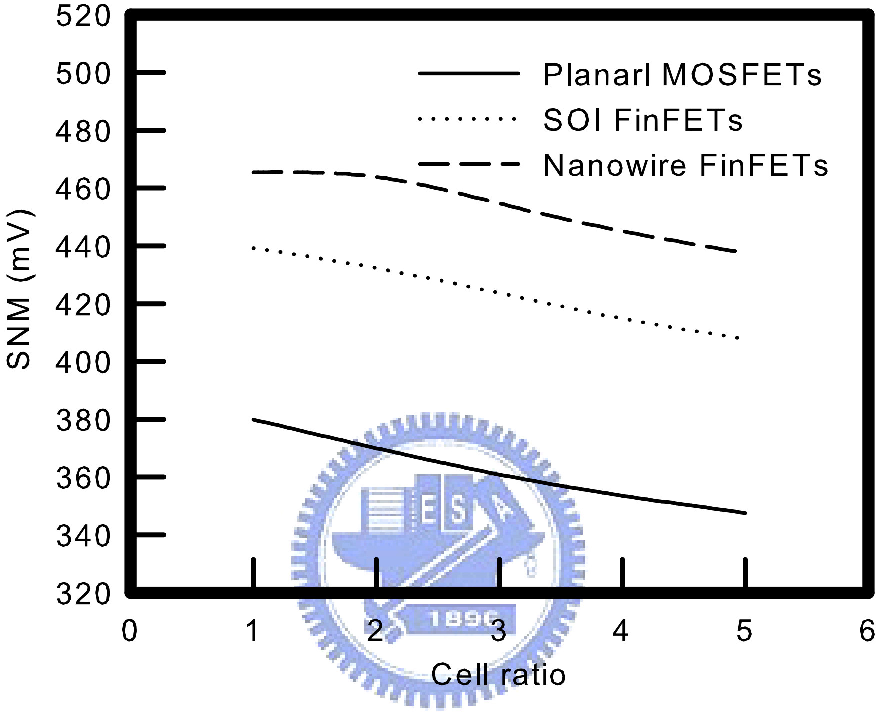


Figure 7.9: Plot of the SNM versus the cell ratio for SRAM cells with three different device structures, where the hold mode is compared. We note upsizing cell ratio will decrease the SNM.

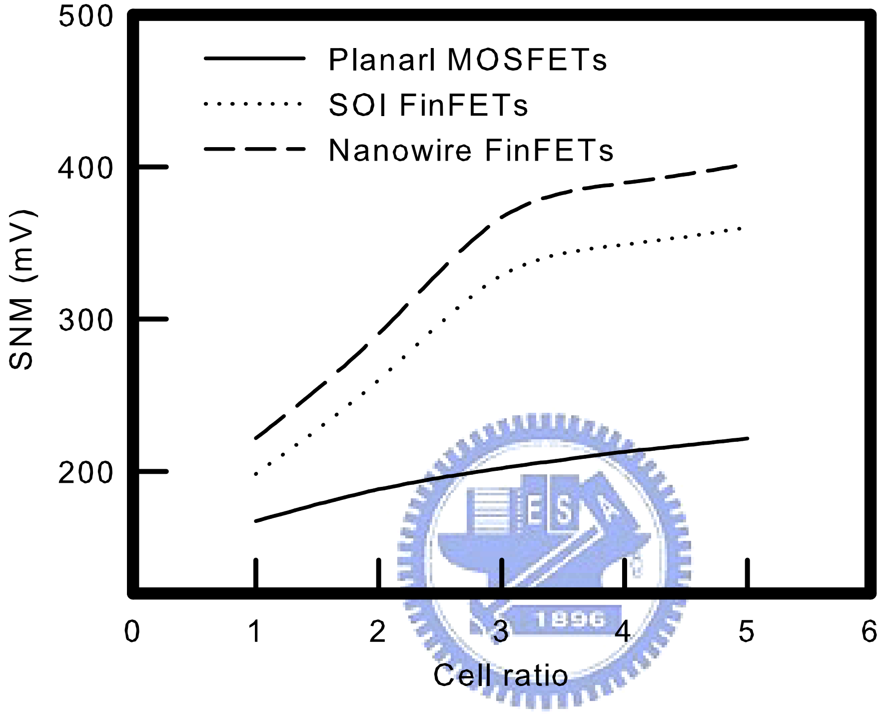


Figure 7.10: Plot of the SNM versus the cell ratio for SRAM cells three different device structures, where the read mode is compared. We note upsizing cell ratio will increase the SNM.

7.4 Summary

In this chapter, we explore the SRAM stability during both hold and read modes with respect to the supply voltage, temperature, and cell ratio. We analyze and compare SRAM cells with planar MOSFETs, SOI FinFETs, and nanowire FinFETs. The SRAM cell using nanowire FinFETs has the best stability and least temperature dependence under both hold and read mode compared with compared with the SRAM cell with planar MOSFETs and SOI FinFETs. It is due to good channel controllability and improved short channel effects. We state that V_{DD} reduction and increase of the temperature induce the significant degradation of SRAM cells. Besides, it is found that for improving the SRAM cell stability we can upsize the cell ratio of the SRAM cell.

Chapter 8

Sensitivity Analysis of the Static Noise Margin

In this chapter, we use the central-composite design (CCD) design with the full 2^{nd} order response surface model to perform the sensitivity analysis of the SNM for the SRAM cell using different device structures. Firstly, the sensitivity of SNM due to device parameter fluctuations will be discussed in section 8.1. It is found that the variance of SNM is strongly affected by the critical dimension of gate length. In section 8.2, we focus on a SNM's sensitivity due to device channel length variations. A sensitivity analysis of SNM for SRAM cells corresponding to the channel length of three different transistors will be discussed.

8.1 The SNM Variation Induced by Device Parameter Fluctuations

Without lost of generality we take a SRAM cell using 65 nm CMOS devices to be an example, and to find find the few significant device factors from a list of many potential ones.

First, we take device channel doping profile, channel length, source/drain doping profile, and thickness of gate oxide as selected variables, then a second order RSM of SNM during the read mode on the 1.2 V supply voltage at room temperature is constructed. The model is in terms of variation of the device channel length, the gate oxide thickness, the channel doping concentration, and the source/drain doping concentration

$$\begin{aligned} SNM = 218.39 - 6.46x_1 - 19.64x_2 - 6.38x_3 \\ +2.43x_4 - 1.00x_1^2 + 15.55x_2^2 + 0.73x_3^2 \\ -1.00x_4^2 + 1.87x_1x_2 + 0.19x_1x_3 \\ -0.28x_1x_4 - 0.32x_2x_3 - 1.66x_2x_4 \\ +1.66x_3x_4, \end{aligned} \tag{8.1}$$

where x_1 is the gate oxide thickness, x_2 is the channel length, x_3 is the source/drain dose, and x_4 is the channel dose. The coefficient are determined by the data obtained from the experiment. The upper and lower bounds of the process parameters are summarized in

Table 8.1. Residual normal plots are performed to verify the accuracy of the constructed model, shown in Fig. 8.1. Examination shows a high accuracy of the RSM of SNM. Table 8.1 report the factors for the studied SRAM cell with the 65 nm planar MOSFETs. The four factor levels are considered with respect to each parameter. The nominal (mean) values are also given in Table 8.1.

With the developed SNM model, the impact of device parameters on SNM is evaluated. Fig. 8.2 - 8.5 will be shown the standard deviation of the SNM due to gate oxide thickness, channel length, source/drain doping concentration, and channel doping concentration variations for the 65 nm SRAM cell based on planar MOSFETs, where the cell ratio is set to be 1. It is found that the variance of SNM is strongly affected by the critical dimension of gate length. In this thesis, we focus on the sensitive analysis of SNM due channel length variations. The SNM is seen to be the most sensitive to threshold voltage fluctuations in the pull-down and access N-type MOSFET devices and least sensitive to the fluctuations in the pull-up P-type MOSFET devices [48]. For avoiding time-consume simulation, we will choose three device parameters as significant factors, they are :

(1)The channel length of the M1 transistor

(2)The channel length of the M4 transistor

(3)The channel length of the other transistors (i.e. M2, M3, M5, M6 transistors)

where the 6-T SRAM circuit id depicted in Fig. 2.1. Now, the node Q is low level (logic

Table 8.1: The upper and lower limits of the parameters, and the nominal recipes.

Parameters	Range	Nominal values
Gate oxide thickness (nm)	$[1.3, 1.7]$	1.5
Channel length (nm)	$[60, 70]$	65
Source/drain dose (cm^{-3})	$[8e+19, 1.2e+20]$	1e+20
Channel dose (cm^{-3})	$[7e+18, 1.2e+18]$	1e+18

”0”), and the node QB is high level (logic ”1”).

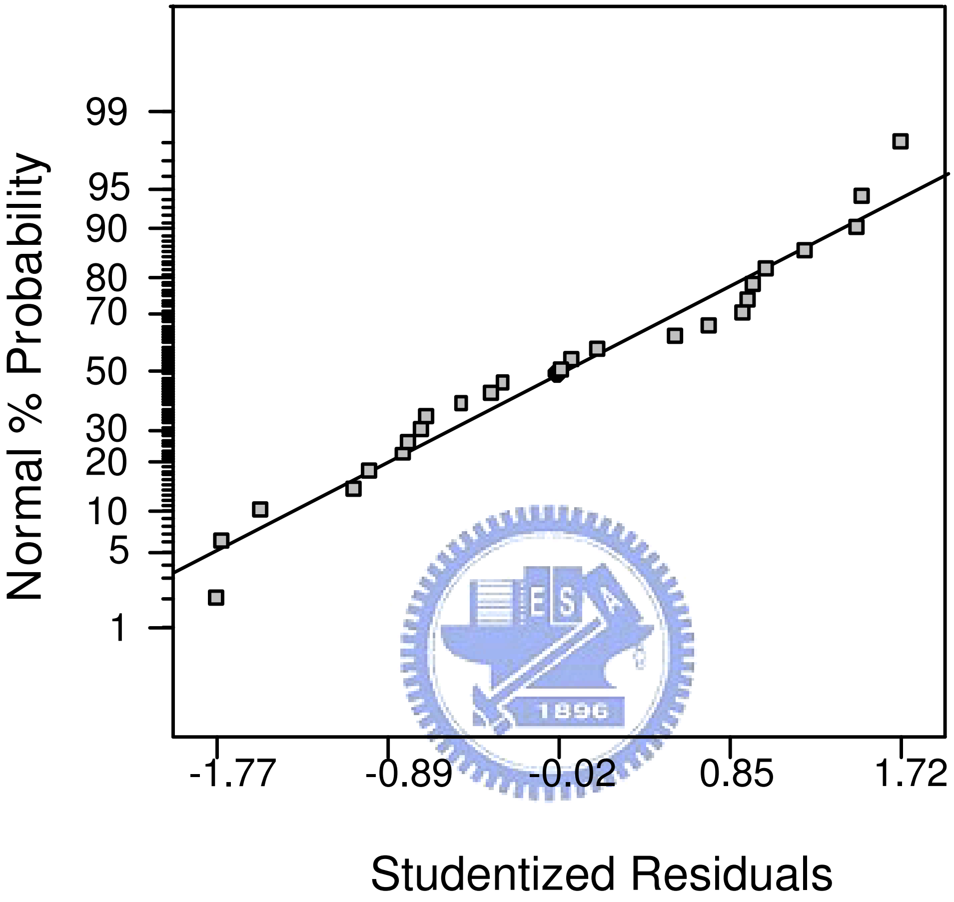


Figure 8.1: Residuals and residual normal plot of SNM.

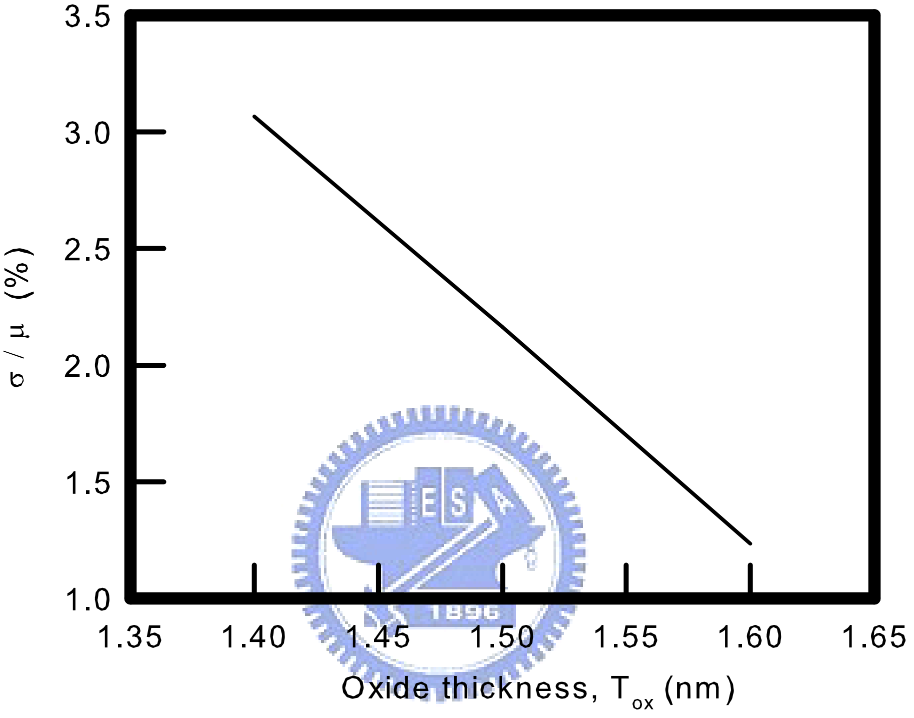


Figure 8.2: Normalized standard deviation of the SNM versus gate oxide thickness for the SRAM cell using 65 nm CMOS devices.

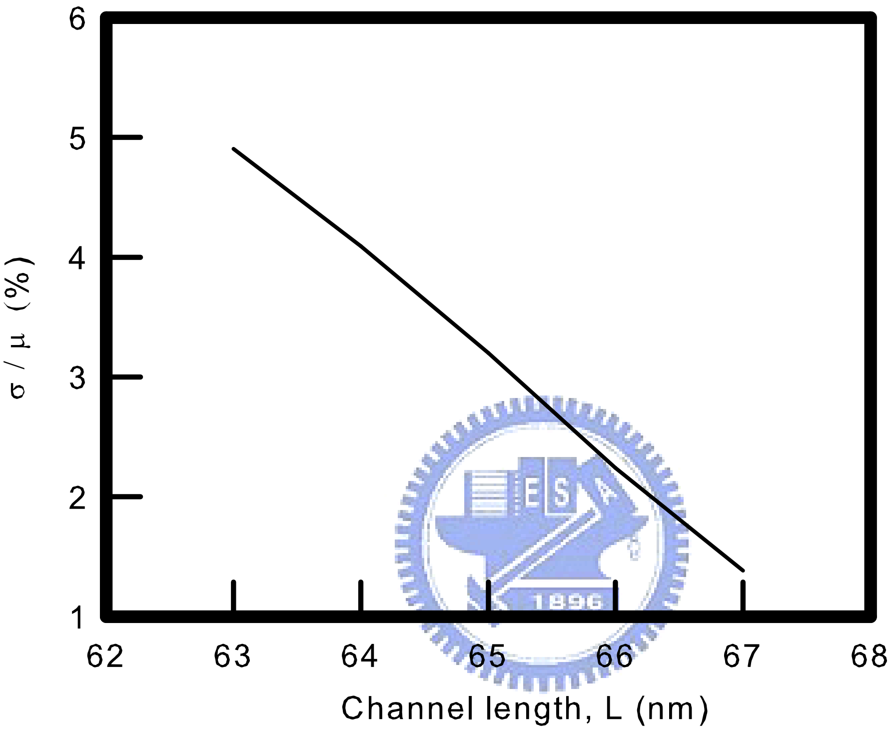


Figure 8.3: Normalized standard deviation of the SNM versus channel length for the SRAM cell using 65 nm CMOS devices.

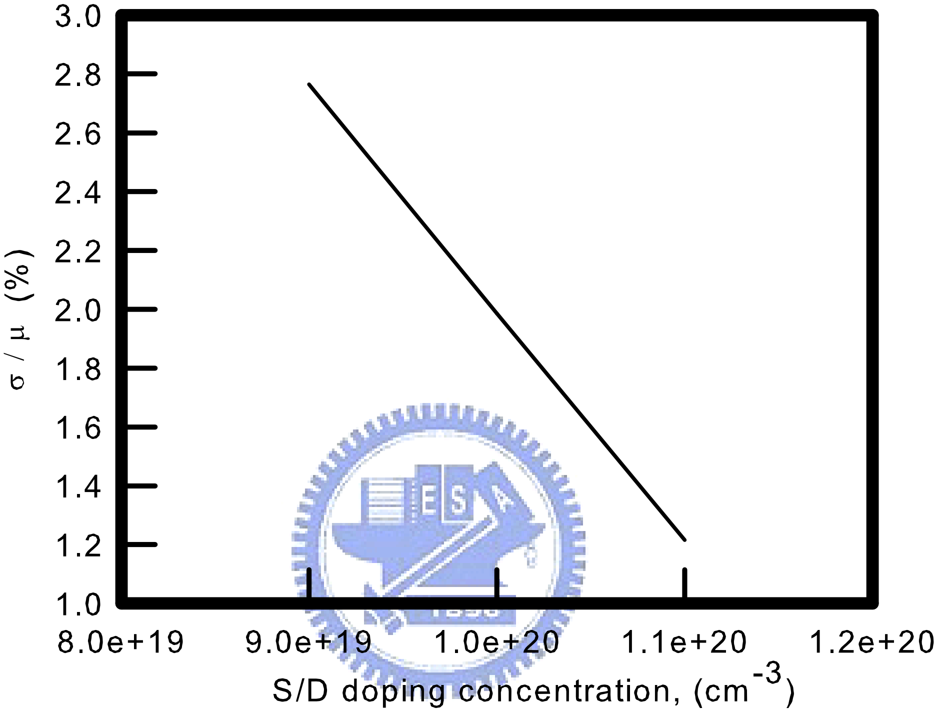


Figure 8.4: Normalized standard deviation of the SNM versus source/drain doping concentration for the SRAM cell using 65 nm CMOS devices.

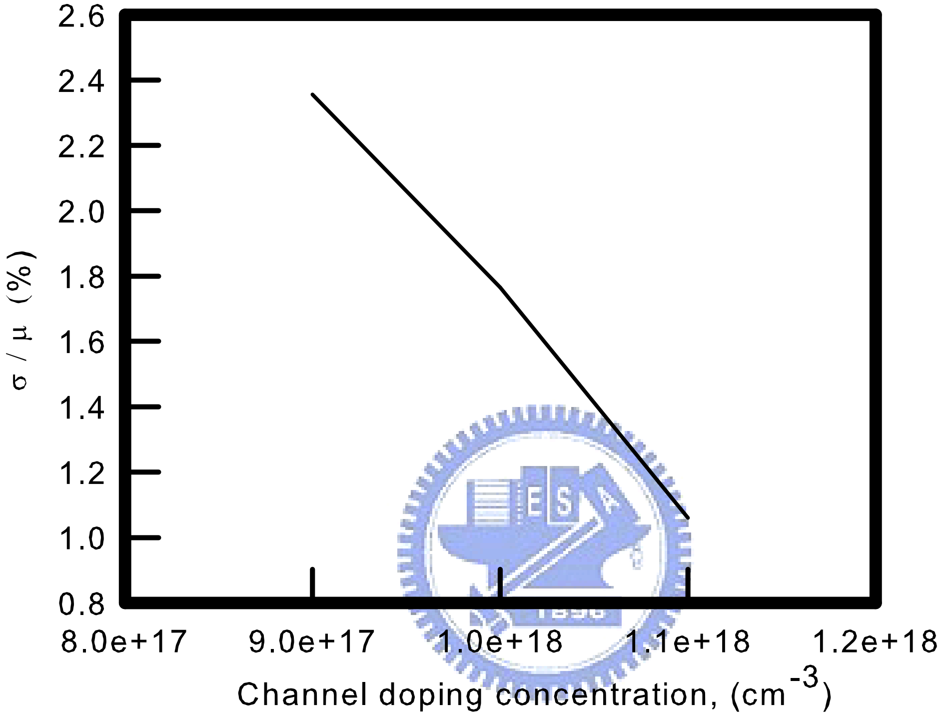


Figure 8.5: Normalized standard deviation of the SNM versus channel doping concentration for the SRAM cell using 65 nm CMOS devices.

8.2 The SNM Variation Induced by Channel Length Fluctuations

In this section, We explore the sensitivity of the SNM versus the variation of the channel length of transistors in the 6-T SRAM cell. We focus on the sensitivity of the SNM under read modes because it is the worst case. With the applied computational statistics technique, we firstly model a second-order RSM for the SNM of the 6-T SRAM cell using 32 nm device structures during the read mode, where the supply voltage is fixed at 1.0 V and the room temperature is assumed. To simply the complicity of the RSM in this examination, without loss of generality, only the six devices' channel lengths are considered and grouped. The RSM of the SNM is in terms of the channel length of the transistor M1, the channel length of the transistor M4, and the channel length of other transistors including M2, M3, M5, and M6. We note the node Q is assumed to have a logic ”0” and the node QB is with a logic ”1”, shown in Fig. 1a. The parameter levels of experiments are shown on Table 8.2.

The established RSM for the 6-T SRAM using the 32 nm planar MOSFETs, SOI FinFETs, and the nanowire FinFETs are given by :

$$\begin{aligned} SNM = 168.34 + 27.96x_1 - 18.59x_2 + 14.33x_3 - 9.04x_1^2 - 4.05x_2^2 \\ +0.23x_3^2 - 0.099x_1x_3 \end{aligned} \tag{8.2}$$

Table 8.2: Experiment levels for all factors

Factor name	**level ”0”**	**level ”-1”**	**level ”1”**
X1 : channel length of the transistor M1 (nm)	32	27	37
X2 : channel length of the transistor M4 (nm)	32	27	37
X3 : channel length of the other transistors M2, M3, M5, M6 (nm)	32	27	37

and

$$SNM = 198.64 + 14.45x_1 - 9.61x_2 + 7.41x_3 - 4.71x_1^2 - 2.14x_2^2 + 0.081x_3^2 - 0.051x_1x_3 \tag{8.3}$$

and

$$SNM = 222.60 + 9.41x_1 - 4.50x_2 + 7.76x_3 - 2.90x_1^2 - 1.47x_2^2 - 0.75x_3^2 + 0.02x_1x_2 - 0.073x_1x_3 + 1.92x_2x_3, \tag{8.4}$$

respectively, where x_1 is the channel length of M1 transistor, x_2 is the channel length of M4 transistor, and x_3 is the channel length of other transistors (M2, M3, M5, M6). The model is supported by visual analysis of residual normal probability plots of SRAM cells using three different device structures, shown in Fig. 8.6 - 8.8.

The sensitivity analysis is performed by simultaneously assuming a uniform distribution on the variables x_1, x_2, and x_3. In the distribution, the 3-sigma = 3 nm is assumed for

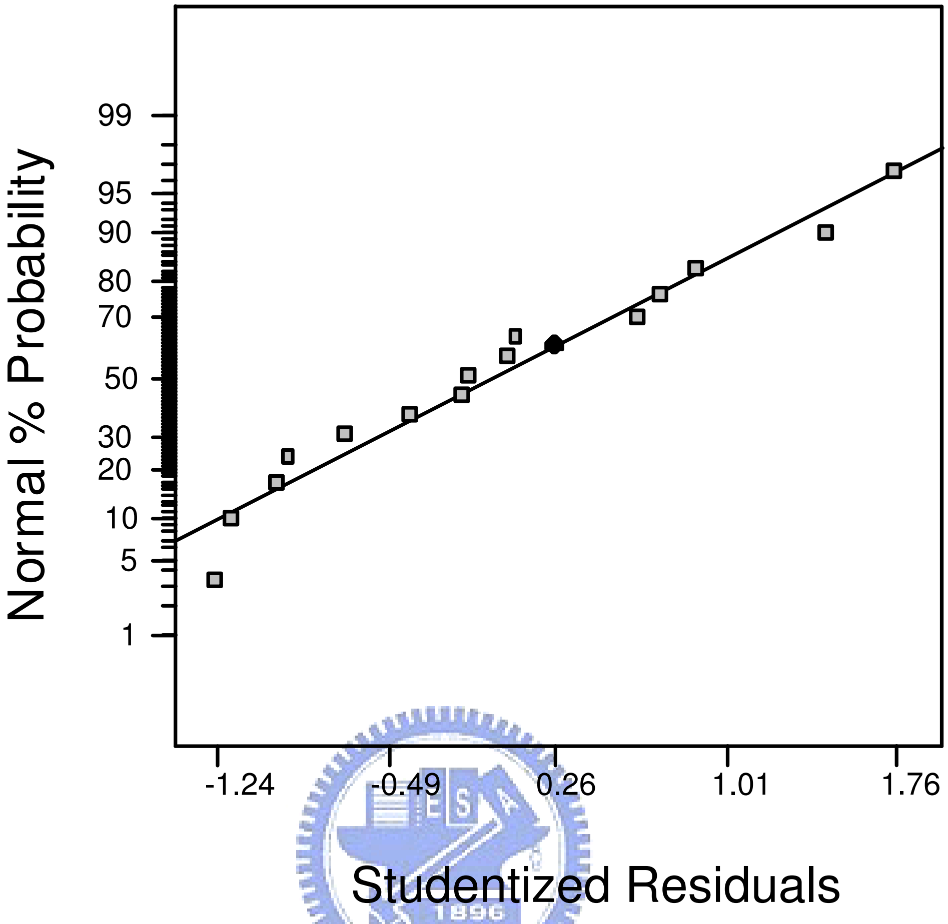


Figure 8.6: Residual normal probability plot of the SNM's response surface model using planar MOSFETs.

each nominal value. The SNM versus channel length of the M1 transistor for SRAM cells using three different device structures will be presented in Fig. 8.9, where the cell ratio is set to be 1. The standard deviation of the SNM due to the channel length variation of the M1 transistor for SRAM cells using three different device structures will be presented in Fig. 8.10. The standard deviation of the SNM for the SRAM with planar MOSFETs

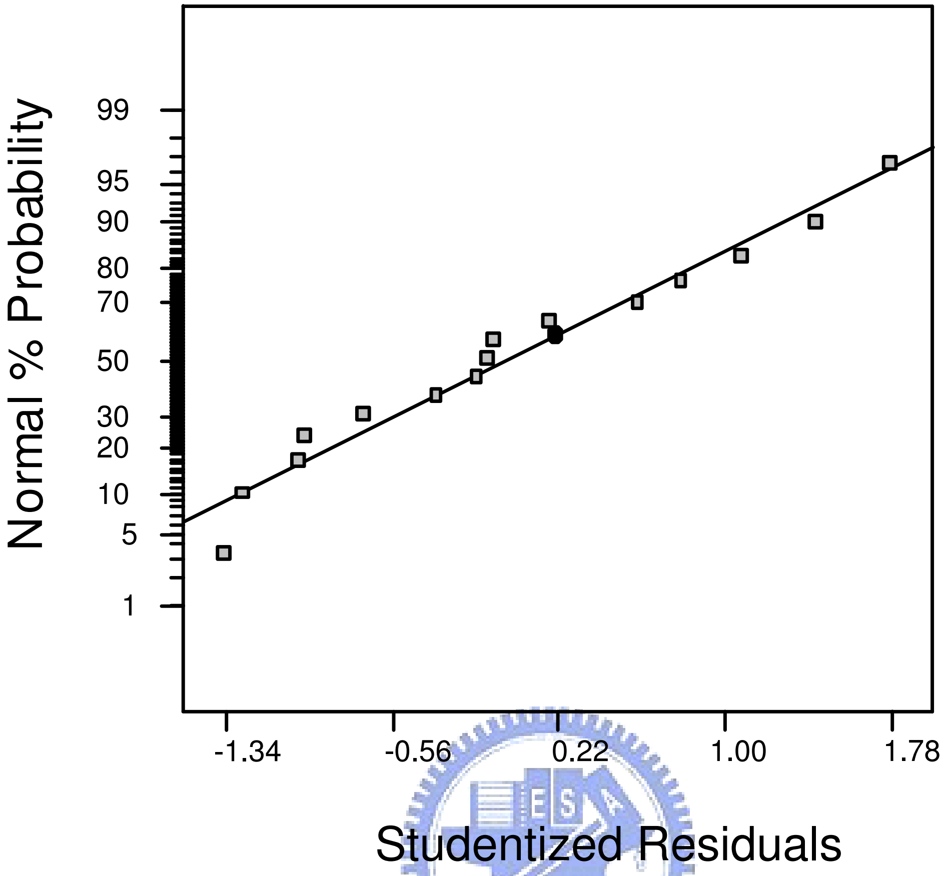


Figure 8.7: PResidual normal probability plot of the SNM's response surface model using SOI FinFETs.

is larger than that of the SOI and nanowire FinFETs due to a serious short channel effects appearing in the planar MOSFETs. For references, the normalized standard deviation of the SNM due to the channel length variation of the M1 transistor is also depicted in Fig. 8.11. Similarly, the SNM versus channel length of the M4 transistor for SRAM cells using three different device structures will be presented in Fig. 8.12, where the cell ratio is set

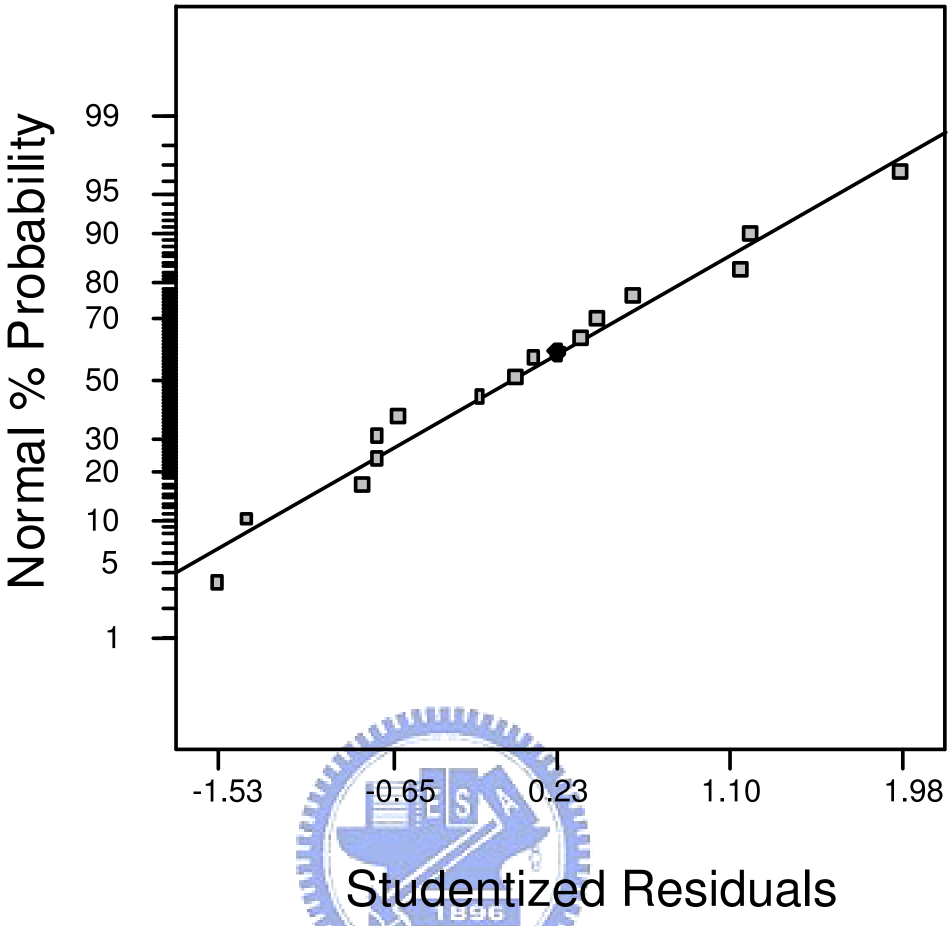


Figure 8.8: Residual normal probability plot of the SNM's response surface model using nanowire FinFETs.

to be 1. The standard deviation of the SNM due to the channel length variation of the M4 transistor for SRAM cells using three different device structures will be presented in Fig. 8.13. The normalized standard deviation of the SNM due to the channel length variation of the M4 transistor is also depicted in Fig. 8.14. The SNM versus channel length of the other transistors for SRAM cells using three different device structures will be presented in Fig. 8.15, where the cell ratio is set to be 1. The standard deviation of the SNM due to the channel length variation of other transistors for SRAM cells using three different device structures will be presented in Fig. 8.16. The normalized standard deviation of the SNM due to the channel length variation of other transistors is also depicted in Fig. 8.17. In a world, the effect of channel length variation on V_{th} is small, so the effect on the SNM is also small. Because the device of nanowire FinFETs has the least sensitivity V_{th} variations due to channel length fluctuations, so it is more resistant to SNM fluctuations.

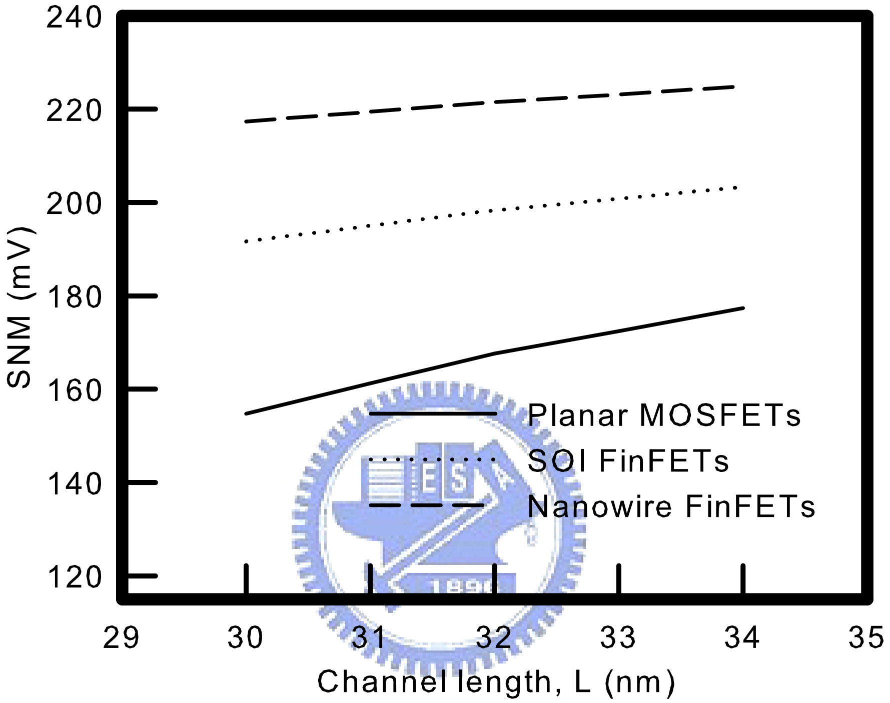


Figure 8.9: SNM versus channel length of M1 transistor for SRAM cells using three different device structures.

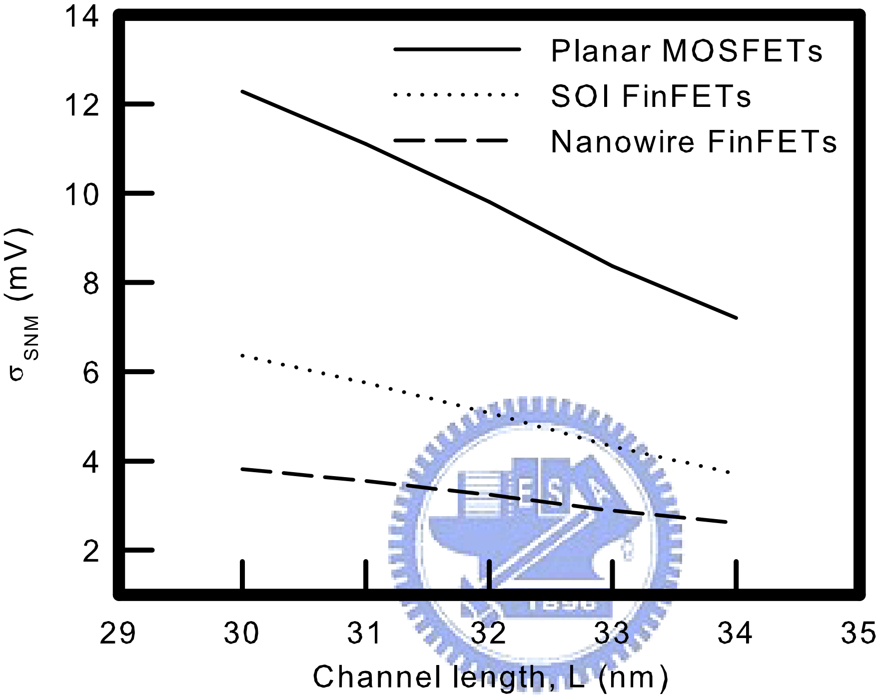


Figure 8.10: Standard deviation of the SNM versus channel length of M1 transistor for SRAM cells using three different device structures.

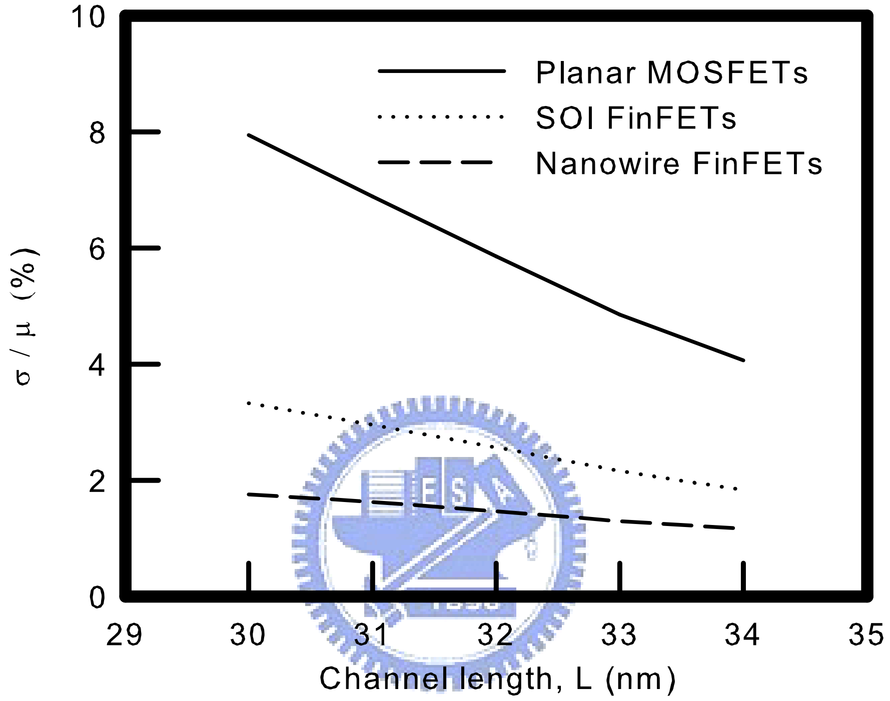


Figure 8.11: Normalized standard deviation of the SNM versus channel length of M1 transistor for SRAM cells using three different device structures.

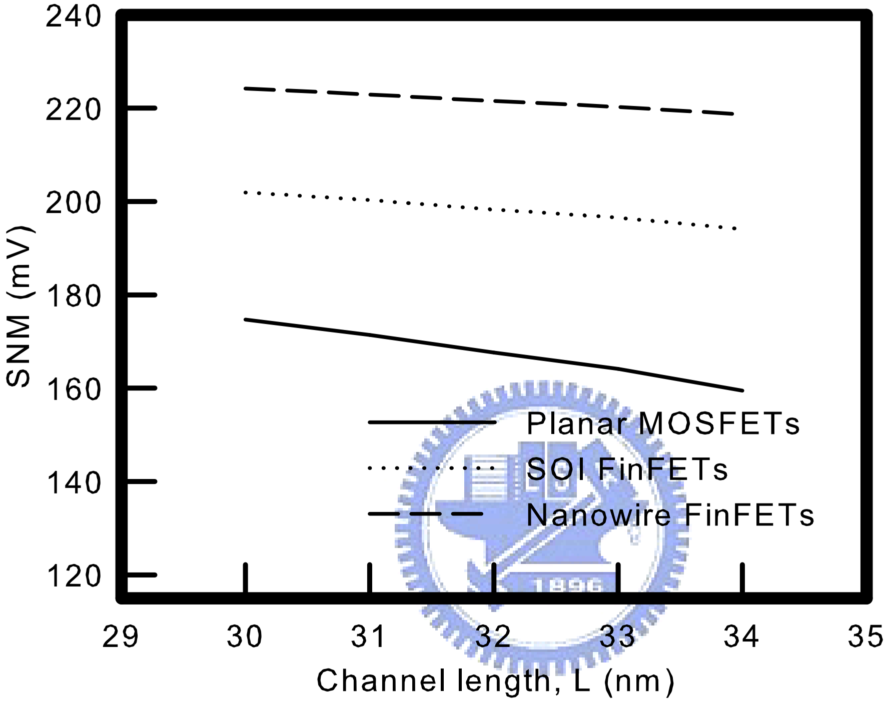


Figure 8.12: SNM versus channel length of M4 transistor for SRAM cells using three different device structures.

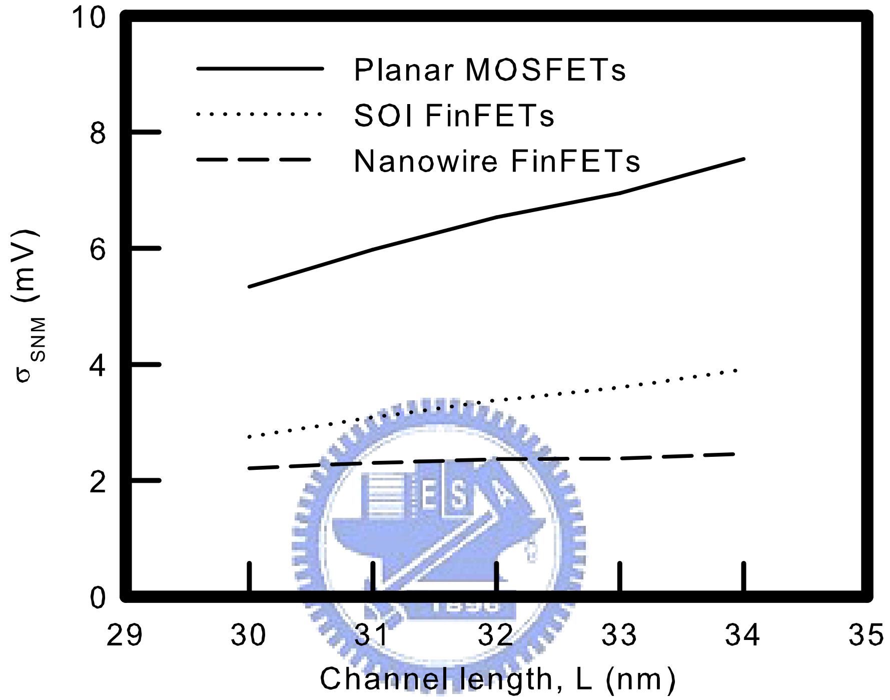


Figure 8.13: Standard deviation of the SNM versus channel length of M4 transistor for SRAM cells using three different device structures.

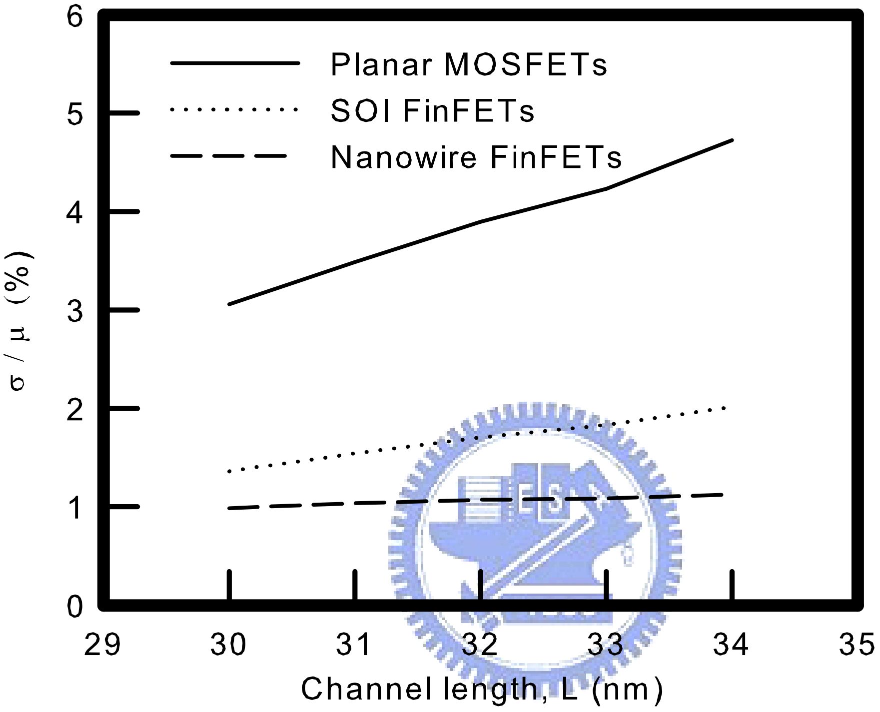


Figure 8.14: Normalized standard deviation of the SNM versus channel length of M4 transistor for SRAM cells using three different device structures.

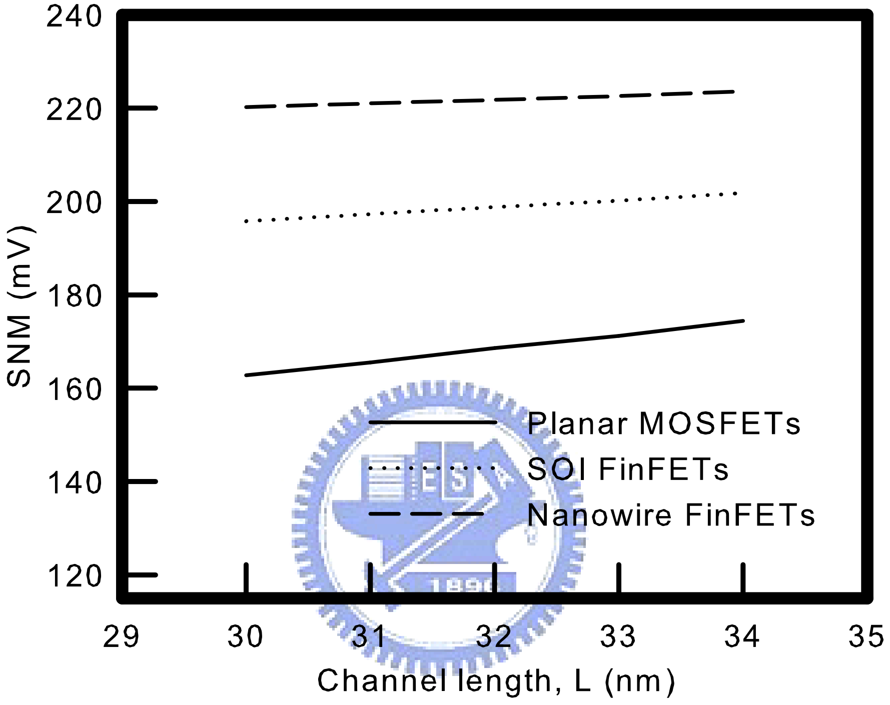


Figure 8.15: SNM versus channel length of other transistors for SRAM cells using three different device structures.

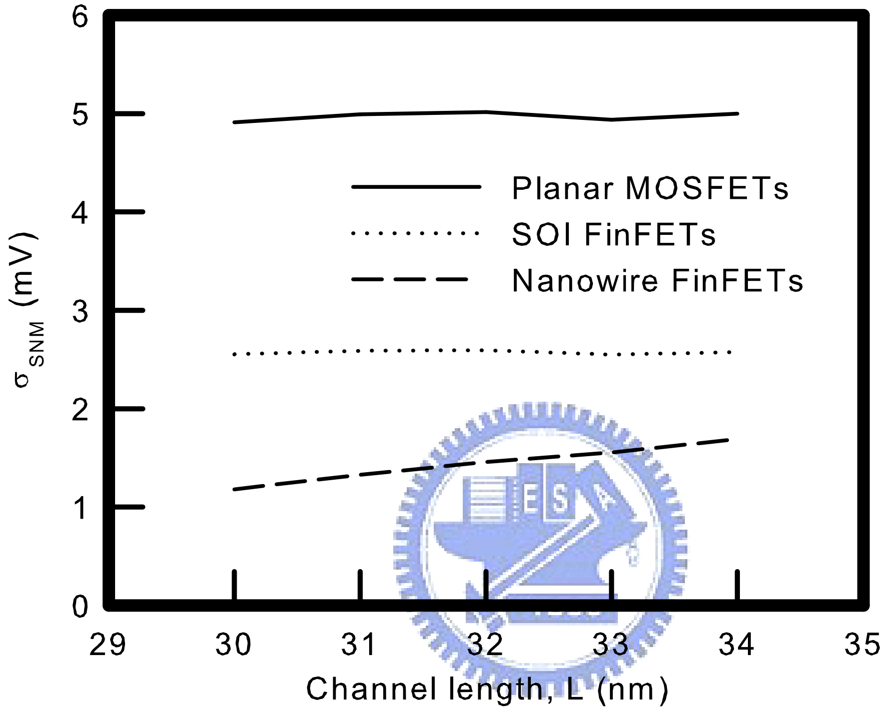


Figure 8.16: Standard deviation of the SNM versus channel length of other transistors for SRAM cells using three different device structures.

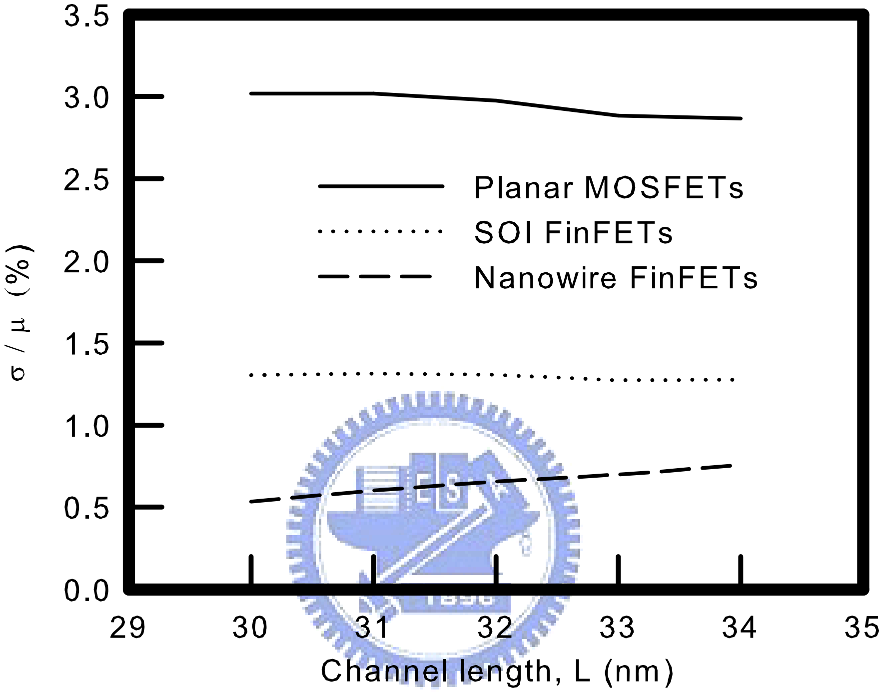


Figure 8.17: Normalized standard deviation of the SNM versus channel length of other transistors for SRAM cells using three different device structures.

8.3 Summary

The sensitivity of SNM due to device parameter fluctuations will be analyzed. It is known that the variance of SNM is strongly affected by the critical dimension of gate length for each transistor. We analyze and compare SNM's fluctuations of SRAM cells using different device structures due to channel length variation of each transistor. It is found that the nanowire FinFET is the least sensitive on SNM compare with that of the planar MOSFET and the SOI FinFET due to a better short channel effects and a smaller threshold voltage fluctuation.

Chapter 9

Conclusions and Suggestions on Future Work

In this thesis, a systematic method for the stability and sensitivity analysis of SRAM cells using different device structures have been investigated. The device characteristics of different structures are also explored. The conventional 6-T SRAM cell using planar MOSFETs, SOI FinFETs, and nanowire FinFETs with channel length is set to be 32 nm was provided to be an example. Achieved conclusions and suggested future works are listed in the following sections.

9.1 Conclusions

The device characteristics (includes the intrinsic and terminal characteristics) with different device structures by using a device simulation which includes the drift-diffusing model for carrier transport and the density gradient model to account for quantum-mechanical effects in nanoscale MOSFETs will be explored. We compare the device characteristics under different technology generation nodes. The short channel effect parameters are extracted from the terminal characteristics (the I_D-V_G curve). It is found that the nanowire FinFET device has the best performance and much immunity to short channel effects (SCEs). We also analyze the electrostatic static potential, electric field, electron density, and hole density under both on-state and off-state.

The stability of SRAM cells using different device structures will be analyze by using a mixed-model simulation. The SNM with respect to supply voltage, cell ratio, and operation temperature under both hold and read mode is discussed. It is known that V_{DD} reduction and increase of the temperature induce the significant degradation of SRAM cells. For improving the SRAM cell stability, it is a method to upsize the cell ratio of SRAM cells. The SRAM cell using nanowire FinFETs has the best stability and least temperature dependence under both hold and read mode compared with compared with the SRAM cell with planar MOSFETs and SOI FinFETS.

A computational statistical methodology that accounts for the sensitivity of the SRAM cell stability is developed. We focus on a SNM's sensitivity due to device channel length variations. It is found that the nanowire FinFET is the least sensitive on SNM compare with that of the planar MOSFET and the SOI FinFET due to a better short channel effects and a smaller threshold voltage fluctuation.

In conclusions, we have explored 6-T SRAM cells with different building CMOS devices including 32 nm planar MOSFETs, SOI FinFETs, and nanowire FinFETs by using mixed-mode 3D device simulation. The SRAM cell with nanowire FinFETs shows very interesting physical properties. The design of 6-T SRAM cell with nanowire FinFETs is promising approach in sub-45 nm CMOS devices era.

9.2 Suggestions for Future Work

In the future works, we could use the similar methodology to analyze the 4-T or other special architectures of SRAM cells. The stability and sensitivity of SNM might be discussed by using a A computational statistical methodology. When there are not device compact models, the performance of other circuits (such as DRAM, operation amplifier et al.) using non-CMOS device structures might be explored by using a mixed-mode simulation. We could write codes to simulate the circuit performance by combine the device physical equations (Poisson equation and carrier transport model) and circuit equations (Kirchhoff current and voltage equations). It is not constraint TACD simulation tools.

For the statistical method, we could use higher order RSM to explain the relationship between the responses and factors. The error between the estimated responses from the RSM and the true measurements obtained from the TCAD might be reduced, and the response surface models would be closer to the true situations. We note that this analysis technique not only can be used together with device and circuit simulation programs for theoretical prediction but also can analyze experimental measurement for realistic SRAM data. The input of DOE can be replaced with experimental measurement for a realistic diagnosis of process variation effects.

We could explore the layout effect of SRAM cells by a direct three-dimension device simulation. We could also construct the via and wire structure to consider the interconnect delay effects. SRAM characteristic might be calculated directly from Poisson equation and electron and hole continuity equations (not mixed-mode). By this method, real circuit structures can be calculated directly and it provides very powerful tool for future perspectives.

Bibliography

[1] H. Wakabayashi, S. Yamagami, N. Ikezawa, A. Ogura, M. Narihiro, K. Arai, Y. Ochiai, K. Takeuchi, T. Yamamoto, and T. Mogami, Sub-10-nm planar-bulk-CMOS devices using lateral junction control, *in Int. Electron Devices Meeting Tech. Dig.*, 2003, pp.20.7.1-20.7.4.

[2] M. V. Fischetti, Scaling MOSFET's to the limit: A physicists's perspective, *J. Comput. Electron.*, vol. 2, no. 2-4, 2003, pp. 73-79.

[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 20 nm, *IEEE Trans. Electron Devices*, vol. 47, no. 12, 2000, pp. 2320-2325.

[4] R. Chau, B. Doyle, J. Kavalieros, D. Barlage, and A. Murthy, Advanced depleted-substrate transistors: Single-gate, double-gate and tri-gate, *in Proc. Solid State Devices and Materials Conf.*, 2002, pp. 68-69.

[5] Fu-Liang Yang, Hao-Yu Chen, Fang-Cheng Chen, Cheng-Chuan Huang, Chang-Yun Chang, Hsien-Kuang Chiu, Chi-Chuang Lee, Chi-Chun Chen, Huan-Tsung Huang, Chih-Jian Chen, Hun-Jan Tao, Yee-Chia Yeo, Mong-Song Liang, and Chenming Hu, 25 nm CMOS Omega FETs, *in Int. Electron Devices Meeting Tech. Dig.*, 2002, pp. 255-258.

[6] J. P. Colinge, M. H. Gao, A. Romano-Rodriguez, H. Maes, and C. Claeys, Silicon-on-insulator ’gate-all-around device, *in Int. Electron Devices Meeting Tech. Dig.*, 1990, pp. 595-598.

[7] Semiconductor Industry Association (SIA), The International Technology Roadmap for Semiconductors, 2004.

[8] H. Ananthan, A. Bansal, and K. Roy, FinFET SRAM - device and circuit design considerations, *in Int. Symp. Quality Electronic Design*, 2004, pp. 511-516.

[9] R. Tanabe, Y. Tosaka, Y. Ashizawa, and H. Oka, , Investigation of 6T SOI SRAM Cell Stability Including Quantum and Gate Direct Tunneling Effects by Three-dimensional Device Simulation, *in Proc. Int. Simulation of Semiconductor Processes*

and Devices Conf., 2005, pp. 43-46.

[10] X. Li, J. Qin, X. Zhang, and J. B. Bernstein, SRAM Circuit Failure Modeling and Reliability Simulation with SPICE, *IEEE Trans. Device and Materials Reliability : Accepted for future publication*, vol. PP, no. 99, 2006, pp. 1.

[11] M. L. Alles, SPICE analysis of the SEU sensitivity of a fully depleted SOI CMOS SRAM cell, *IEEE Trans. Nuclear Science*, vol. 41, no. 6, 1994, pp. 2093-2097.

[12] Abu Khari bin A'Ain, Kian Sin Sim, and Cheow Kwee Siong, The effects of gate oxide short in 6-transistors SRAM cell, *in Proc. Int. Semiconductor Electronics Conf.*, 2004, pp. 5.

[13] Zheng Guo, S. Balasubramanian, R. Zlatanovici, Tsu-Jae King, and B. Nikolic, FinFET-based SRAM design, *in. Int. Symp. Low Power Electronics and Design*, 2005, pp. 2-7.

[14] Y. Kawakami,H. Hane, H. Nakamura, T. Yamada, and K. Kumagai, Investigation of soft error rate including multi-bit upsets in advanced SRAM using neutron irradiation test and 3D mixed-mode device simulation, *in Int. Electron Devices Meeting Tech. Dig.*, 2004, pp. 945-948.

[15] A. Asenov, 3D statistical simulation of instrinsic fluctuations in decanano MOSFET's induced bt discrete dopants, oxide thickness fluctuations and LER, *in Proc. Int. Simulation of Semiconductor Processes and Devices Conf.*, 2001, pp. 162-169.

[16] H. Wong, and Y. Taur, 3-Dimensional atomistic simulation of discrete effects in sub-0.1μm MOSFETs, *in Int. Electron Devices Meeting Tech. Dig.*, 1993, pp. 705-708.

[17] A. Asenov, A.R. Brown, J.H. Davies, S.Kaya, and G. Slavcheva, Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale MOSFETS, *IEEE Trans. Electron Devices*, vol. 50, no. 9, 2003, pp. 1837-1852.

[18] X. Tang, V. Die, and J. D. Meindl, Intrinsic MOSFET parameter fluctuations due to random dopant placement, *IEEE Trans. Very Large Scale Integration Systems*, vol. 5, no. 4, 1997, pp. 369-376.

[19] D. Burnett, K. Erington, C. Subramanian, and K. Baker, Implications of fundamental threshold voltage variations for high-density SRAM and logic circuits, *in VLSI Technology Tech. Synp. Dig.*, 1994, pp. 15-16.

[20] R. Heald, P. Wang, Variability in sub-100nm SRAM designs , *in Int. Computer Aided Design Conf.*, 2004, pp. 347-352.

[21] K. Samsudin, B. Cheng, A.R. Brown, S. Roy, A. Asenov, UTB SOI SRAM cell stability under the influence of intrinsic parameter fluctuation, *in Proc. Solid-State Device Research Conf.*, 2005, pp. 553-5563.

[22] K. Samsudin, B. Cheng, A.R. Brown, S. Roy, A. Asenov, Impact of random dopant induced fluctuations on sub-15nm UTB SOI 6T SRAM cells , *in Int. Silicon on Insulator Conf.*, 2005, pp. 61-62.

[23] G.J. Gaston, Process Sensitivity Analysis Using Design of Experimental and Full 2D TCAD Simulation, *Proc. IEE Colloquium on Advanced MOS and Bi-Polar Device*, 1995, pp 1-6.

[24] Semiconductor Industry Association (SIA), The International Technology Roadmap for Semiconductors, 2001.

[25] Y. Taur, C. H. Wann, and D.J. Frank, 25 nm CMOS design considerations , *in Int. Electron Devices Meeting Tech. Dig.*, 1998, pp.789-792.

[26] B. Yu; Haihong Wang; A. Joshi, Q. Xiang, E. Ibok, and Ming-Ren Lin, 15 nm gate length planar CMOS transistor, *in Int. Electron Devices Meeting Tech. Dig.*, 2001, pp.937-939.

[27] F. Boeuf, T. Skotnicki, S. Monfray, C. Julien, D. Dutartre, J. Martins, P. Mazoyer, R. Palla, B. Tavel, P. Ribot, E. Sondergard, and A. Sanquer, 16 nm planar NMOSFET

manufacturable within state-of-the-art CMOS process thanks to specific design and optimisation, *in Int. Electron Devices Meeting Tech. Dig.*, 2001, pp.637-640.

[28] B. Doyle et al., Transistor elements for 30nm hysical gate lengths and beyond, *J. Intel Tech.*, vol. 6, no. 2, 2002, pp. 42-54.

[29] 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, 14 nm gate length CMOSFETs utilizing low thermal budget process with poly-SiGe and Ni salicide, *in Int. Electron Devices Meeting Tech. Dig.*, 2002, pp.639-642.

[30] Tsu-Jae King, Chenming Hu, Yang-Kyu Choi; and D. Ha, Nanoscale ultrathin body PMOSFETs with raised selective germanium source/drain , *IEEE Electron Devices Lett.*, vol. 22, 2001, pp.447-449.

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

[32] AXuejue Huang, Wen-Chin Lee, C. Kuo, D. Hisamoto, Leland Chang, J. Kedzierski, E. Anderson, H. Takeuchi, Yang-Kyu Choi, K. Asano, V. Subramanian, Tsu-Jae King, J. Bokor, and Chenming Hu, Sub-50 nm P-channel FinFET, *IEEE Trans. Electron Devices*, vol. 48, no. 5, 2001, pp. 880-886.

[33] Leland Chang, K.J. Yang, Yee-Chia Yeo, Yang-Kyu Cho,; Tsu-Jae King, and Chenming Hu, Reduction of direct-tunneling gate leakage current in double-gate and ultra-thin body MOSFETs, *in Int. Electron Devices Meeting Tech. Dig.*, 2001, pp.99-102.

[34] H.S.P. Wong, Beyond the conventional transistors, *J. IBM Res.&Dev.*, vol. 46, 2002, pp. 133-167.

[35] D. Hisamoto, T. Kaga, E. Takeda, Impact of the vertical SOI 'DELTA' structure on planar device technology, *IEEE Trans. Electron Devices*, vol. 38, no. 6, 1991, pp. 1419-1423.

[36] N. Lindert, L. Chang, Yang-Kyu Choi, E.H. Anderson, Wen-Chin Lee, Tsu-Jae King, J. Bokor, and Chenming Hu, Sub-60-nm quasi-planar FinFETs fabricated using a simplified process, *IEEE Electron Devices Lett.*, vol. 22, 2001, pp.487-489.

[37] Bin Yu, Leland Chang, S. Ahmed, Haihong Wang, S. Bell, Chih-Yuh Yang, C. Tabery, Chau Ho, Qi Xiang, Tsu-Jae King, J. Bokor, J, Chenming Hu, Ming-Ren

Lin, and D. Kyser, FinFET scaling to 10 nm gate length, *in Int. Electron Devices Meeting Tech. Dig.*, 2002, pp.251-254.

[38] B. Doyle, B. Boyanov, S. Datta, M. Doczy, S. Hareland, B. Jin, J. Kavalieros, T. Linton, R. Rios, and R. Chau, Tri-Gate fully-depleted CMOS transistors: fabrication, design and layout, *in VLSI Technology Tech. Synp. Dig.*, 2003, pp. 133-134.

[39] J.P. Colinge, M.H. Gao, A. Romano-Rodriguez, H. Maes, and C. Claeys, Silicon-on-insulator 'gate-all-around device', *in Int. Electron Devices Meeting Tech. Dig.*, 1990, pp.595-598.

[40] E. Seevinck, F. J. List, and J. Lohstoh, Static-Noise Margin Analysis of MOS SRAM Cells, *IEEE J. Solid-state circuits*, vol. 22, no. 5, 1987, pp. 748-754.

[41] J. Lohstroh, E. Seevinck, and J. de Grooth, Worst-case static noise margin criteria for logic circuits and their mathematical equivalence. *IEEE J. Solid-state circuits*, vol. SC-18, no. 6, 1983, pp. 803-807.

[42] R. H. Myers and D. C. Montgomery, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, New York: John Wiley & Sons, Inc., 2002.

[43] H. C. Srinivasaiah and N. Bhat, Response Surface Modeling of 100nm CMOS Process Technology Using Design of Experiment, *in VLSI Technology Tech. Synp. Dig.*, 2004, pp. 285-290.

[44] S. Williams, K. Varahramyan, and W. Maszara, Statistical Optimization and Manufacturing Sensitivity Analysis of $0.18\mu m$ SOI MOSFETs, *Microelectronic Engineering*, Vol. 49, 1999, pp. 245-261.

[45] D. C. Montgomery, *Design and Analysis of Experiments*, New York: John Wiley & Sons, Inc., 1997.

[46] *Engineering Statistics Handbook* [Online]. Available: http://www.itl.nist.gov/div898/handbook/

[47] C. F. J. Wu and M. Hamada, *Experiments: Planning, Analysis, and Parameter Design Optimization*, New York: John Wiley & Sons, Inc., 2000.

[48] Y. Jing, and B. Sriram, *Design of sub-30nm FinFET based SRAMs* Available: http://www-device.eecs.erkeley.edu/ bsriram/ee241/ee241-SRAM-final-report.pdf

[49] D. C. Montgomery, *Design and Analysis of Experiments*, New York: John Wiley & Sons, Inc., 1997.

[50] H. O. Hartley, Smallest Composite Designs for Quadratic Response Surfaces, *Biometrics*, Vol. 15, 1959, pp. 611-624.

[51] N. R. Draper and D. K. J. Lin, Small Response-Surface Designs, *Technometrics*, Vol. 32, 1990, pp. 187-194.

[52] S. Xiong and J. Bokor, Sensitivity of double-gate and FinFET devices to process variations, *IEEE Trans. Electron Devices*, vol. 50, no. 11, 2003, pp. 2255-2261.

[53] J.-T. Park and J.-P. Colinge, Multiple-gate SOI MOSFET: Device design guideline *IEEE Trans. Electron Devices*, vol. 49, no. 12, 2002, pp. 2222-2229.

[54] J. P. Colinge, M. H. Gao, A. Romano-Rodriguez, H. Maes, and C. Claeys, Silicon-on-insulator 'gate-all-around device' *in Int. Electron Devices Meeting Tech. Dig.*, 1990, pp. 595-598.

[55] C. P. Auth and J. D. Plummer, Scaling theory for cylindrical, fully-depleted, surrounding-gate MOSFET's, *IEEE Lett. Electron Device*, vol. 18, no. 2, 1997, pp. 74-76.

[56] J. Robaey, A. Chandrakasan, and B. Nikolic, Digital Integrated Circuits : A Design Perspective , *2nd edition Prentice-Hall*, 2002.

[57] T. B. Hook, M. Breitwisch, J. Brown, P. Cottrell, D. Hoyniak, Lam Chung, and R. Mann, Noise margin and leakage in ultra-low leakage SRAM cell design, *IEEE Trans. Electron Devices*, vol. 49, no. 8, 2002, pp. 1499-1501.

[58] E.J. Nowak, B.A. Rainey, D.M. Fried, J. Kedzierski,M. Ieong, W. Leipold, J. Wright, M. Breitwisch, A functional FinFET-DGCMOS SRAM cell, *in Int. Electron Devices Meeting Tech. Dig.*, 2002, pp.411-414.

[59] T. Park, H.J. Cho, J.D. Choe, S.Y. Han, S.-M. Jung, J.H. Jeong, B.Y. Nam, O.I. Kwon, J.N. Han, H.S. Kang, M.C. Chae, G.S. Yeo, S.W. Lee, D.Y. Lee, D. Park, K. Kim, E. Yoon, J.H. Lee, Static noise margin of the full DG-CMOS SRAM cell using bulk FinFETs (Omega MOSFETs), *in Int. Electron Devices Meeting Tech. Dig.*, 2003, pp.2.2.1-2.2.4.

Appendix A

Commands of the ISE TCAD Tool

In this appendix we state ISE commands of our simulation. We state firstly boundary commands for the planar MOSFET, SOI FinFET, and nanowire FinFET, respectively. The mixed-mode simulation command for SRAM cells will be shown latter, where consist of both hold and read mode.

A.1 Boundary Commands

We will show boundary commands for different device structures. The device structure and doping profile will be defined by boundary commands.

A.1.1 The Planar MOSFET

```
Definitions {
# Refinement regions
Refinement ”Default Region”
{
}
Refinement ”NoName_1”
{
MaxElementSize = (0.01 0.01 )
MinElementSize = (0.01 0.01 )
}
Refinement ”NoName_2”
{
MaxElementSize = (0.005 0.005 )
MinElementSize = (0.005 0.005 )
}
Refinement ”NoName_3”
{
MaxElementSize = (5 0.05 )
```

```
MinElementSize = (0.001 0.001 )
RefineFunction = MaxTransDiff(Variable = ”DopingConcentration”, Value = 1)
}
# Profiles
Constant ”substrate”
{
Species = ”BoronActiveConcentration”
Value = 1e+18
}
AnalyticalProfile ”source”
{
Species = ”ArsenicActiveConcentration”
Function = Gauss(PeakPos = 0, PeakVal = 3e+20, ValueAtDepth = 3e+18, Depth = 0.05)
LateralFunction = Erf(Factor = 0.36)
}
AnalyticalProfile ”drain”
{
Species = ”ArsenicActiveConcentration”
Function = Gauss(PeakPos = 0, PeakVal = 3e+20, ValueAtDepth = 3e+18, Depth = 0.05)
```

```
LateralFunction = Erf(Factor = 0.36)
}
}
Placements {
# Refinement regions
Refinement ”Default Region”
{
Reference = ”Default Region”
# Default region
}
Refinement ”NoName_1”
{
Reference = ”NoName_1”
RefineWindow = rectangle [( -0.04194 -0.01944 ) , ( 0.15694 0.16944 )]
}
Refinement ”NoName_2”
{
Reference = ”NoName_2”
RefineWindow = rectangle [( 0.03 -0.001 ) , ( 0.082 0.05 )]
```

```
}
Refinement ”NoName_3”
{
Reference = ”NoName_3”
RefineWindow = rectangle [( 0.035 0 ) , ( 0.077 0.06 )]
}
# Profiles
Constant ”substrate”
{
Reference = ”substrate”
EvaluateWindow
{
Element = rectangle [( 0 0 ) , ( 0.112 0.7 )]
DecayLength = 0
}
}
AnalyticalProfile ”source”
{
Reference = ”source”
```

```
ReferenceElement
{
Element = line [( 0 0 ) , ( 0.0255 0 )]
}
}
AnalyticalProfile ”drain”
{
Reference = ”drain”
ReferenceElement
{
Element = line [( 0.087 0 ) , ( 0.112 0 )]
}
}
}
Offsetting {
#steps to perform usebox = 0
#global settings: noffset { hloc= 0 factor= 1.3 subdivide= 0 }
noffset { maxedgelength= 1e+30 terminateline= 3 maxlevel= 200 }
boundary { hglob= 1e+08 }
```

```
#isolines: }
```

A.1.2 The SOI FinFET

```
(define nm 1e-3)

(define L @L1@)

;— Bulk

(isegeo:create-cuboid (position 0 0 0) (position (+ 70 L) 80 50) ”SiO2” ”Bulk”)

;—Gateoxide

(isegeo:create-cuboid (position 35 33.5 50) (position (+ 35 L) 46.5 71.5) ”Oxide” ”Gateoxide”)

;— Channel

(isegeo:set-default-boolean ”ABA”)

(isegeo:create-cuboid (position 0 35 50) (position (+ 70 L) 45 70) ”Silicon” ”Channel”)

;— Gate contact

;—Top

(isegeo:define-contact-set ”gate” 4.0 (color:rgb 1.0 0.0 0.0 ) ”——” )

(isegeo:set-current-contact-set ”gate”)

(isegeo:set-contact-faces (find-face-id (position 45 40 71.5)))
```

```
;—Left
(isegeo:define-contact-set ”gate” 4.0 (color:rgb 1.0 0.0 0.0 ) ”——” )
(isegeo:set-current-contact-set ”gate”)
(isegeo:set-contact-faces (find-face-id (position 45 33.5 60)))
;—Right
(isegeo:define-contact-set ”gate” 4.0 (color:rgb 1.0 0.0 0.0 ) ”——” )
(isegeo:set-current-contact-set ”gate”)
(isegeo:set-contact-faces (find-face-id (position 45 46.5 60)))
;— Source contact
(isegeo:define-contact-set ”source” 4.0 (color:rgb 1.0 0.0 0.0 ) ”##” )
(isegeo:set-current-contact-set ”source”)
(isegeo:set-contact-faces (find-face-id (position 0 40 60)))
;— Drain contact
(isegeo:define-contact-set ”drain” 4.0 (color:rgb 1.0 0.0 0.0 ) ”##” )
(isegeo:set-current-contact-set ”drain”)
(isegeo:set-contact-faces (find-face-id (position (+ 70 L) 40 60)))
;— Substrate contact
(isegeo:define-contact-set ”substrate” 4.0 (color:rgb 1.0 0.0 0.0 ) ”##” )
(isegeo:set-current-contact-set ”substrate”)
```

```
(isegeo:set-contact-faces (find-face-id (position 45 40 0)))

;— Channel doping

;—base line

(isedr:define-refinement-window ”BaseLine_Channel” ”Cuboid”

(position 35 35 45)

(position (+ 35 L) 45 70) )

;—implant

(isedr:define-constant-profile ”Implant_Channel” ”BoronActiveConcentration” 3e16)

;—placement

(isedr:define-constant-profile-placement ”Placement_Channel” ”Implant_Channel” ”BaseLine_Channel”3)

;— source doping

;—base line (isedr:define-refinement-window ”BaseLine_Source” ”Cuboid”

(position 0 35 45)

(position 35 45 70)

)

;—implant

(isedr:define-constant-profile ”Implant_Source” ”ArsenicActiveConcentration” 3e20)

;—placement
```

```
(isedr:define-constant-profile-placement ”Placement_Source” ”Implant_Source” ”BaseLine_Source”
1)
;— drain doping
;—base line (isedr:define-refinement-window ”BaseLine_Drain” ”Cuboid”
(position (+ 35 L) 35 45)
(position (+ 70 L) 45 70)
)
;—implant (isedr:define-constant-profile ”Implant_Drain” ”ArsenicActiveConcentration”
3e20)
;—placement
(isedr:define-constant-profile-placement ”Placement_Drain” ”Implant_Drain” ”BaseLine_Drain”1)
;—multibox 1
(isedr:define-refinement-window ”multibox1” ”Cuboid”
(position 0 0 0)
(position (+ 70 L) 80 70) )
(isedr:define-multibox-size ”multiboxSize1”60 60 60 20 20 20 )
(isedr:define-multibox-placement ”multiboxPlacement1”
”multiboxSize1” ”multibox1”)
—multibox 2
```

```
(isedr:define-refinement-window ”multibox2” ”Cuboid”
(position 0 30 40)
(position (+ 70 L) 50 80) )
(isedr:define-multibox-size ”multiboxSize2”30 30 30 15 15 15)
(isedr:define-multibox-placement ”multiboxPlacement2”
”multiboxSize2” ”multibox2”)
;—multibox 3
(isedr:define-refinement-window ”multibox3” ”Cuboid”
(position 25 30 40)
(position (+ 45 L) 50 80) )
(isedr:define-multibox-size ”multiboxSize3”10 10 10 5 5 5)
(isedr:define-multibox-placement ”multiboxPlacement3”
”multiboxSize3” ”multibox3”)
;—multibox 4
(isedr:define-refinement-window ”multibox4” ”Cuboid”
(position 32 32 45)
(position 38 48 75) )
(isedr:define-multibox-size ”multiboxSize4”6 6 6 4 4 4)
(isedr:define-multibox-placement ”multiboxPlacement4”
```

```
”multiboxSize4” ”multibox4”)
;—multibox 5
(isedr:define-refinement-window ”multibox5” ”Cuboid”
(position (+ 32 L) 32 45)
(position (+ 38 L) 48 75) )
(isedr:define-multibox-size ”multiboxSize5”6 6 6 4 4 4)
(isedr:define-multibox-placement ”multiboxPlacement5”
”multiboxSize5” ”multibox5”)
;— Save BND and CMD and rescale to um
(ise:assign-material-and-region-names (part:entities (filter:type ”solid?”)))
(isegeo:scale (part:entities (filter:type ”solid?”)) nm nm nm )
(iseio:save-dfise-bnd (part:entities (filter:type ”solid?”))
”@boundary/o@”) (isedr:write-scaled-cmd-file ”@commands/o@” nm)
```

A.1.3 The nanowire FinFET

```
;— Units conversion from nm to the default unit on um
(define nm 1e-3)
(define L @L1@)
```

```
;———————————————————————————-;
;— Channel
(isegeo:create-cylinder (position 0 0 10) (position L 0 10) 9.5 ”SiO2” ”Gateoxide”)
”ABA”
(isegeo:create-cylinder (position 0 0 10) (position L 0 10) 8 ”Silicon” ”Channel”)
;— Source
(isegeo:create-cylinder (position -50 0 10) (position 0 0 10) 8 ”Silicon” ”Source”)
(isegeo:create-cuboid (position -100 -15 -5) (position -50 15 25) ”Silicon” ”SC”)
;— Drain
(isegeo:create-cylinder (position L 0 10) (position (+ 50 L) 0 10) 8 ”Silicon” ”Drain”)
(isegeo:create-cuboid (position (+ 50 L) -15 -5) (position (+ 100 L) 15 25) ”Silicon” ”DC”)
;— Gate contact:
(isegeo:define-contact-set ”gate” 4.0 (color:rgb 1.0 0.0 0.0 ) ”——” )
(isegeo:set-current-contact-set ”gate”)
(isegeo:set-contact-faces(find-face-id (position 10 0 19.5)))
;— Assign drain contact
(isegeo:define-contact-set ”drain” 4.0 (color:rgb 1.0 0.0 0.0 ) ”##” )
(isegeo:set-current-contact-set ”drain”)
(isegeo:set-contact-faces(find-face-id (position 88 5 25 )))
```

```
;— Assign source contact
(isegeo:define-contact-set ”source”4.0 (color:rgb 1.0 0.0 0.0 ) ”##” )
(isegeo:set-current-contact-set ”source”)
(isegeo:set-contact-faces(find-face-id (position -75 6 25 )))
;—Bulk
;—base line
(isedr:define-refinement-window ”BaseLine_Channel” ”Cuboid”
(position 0 -10 0)
(position L 10 20) )
;—implant
(isedr:define-constant-profile ”Implant_Channel” ”BoronActiveConcentration” 3e16)
;—placement
(isedr:define-constant-profile-placement
”Placement_Channel” ”Implant_Channel” ”BaseLine_Channel”0)
;—Source
;—base line
(isedr:define-refinement-window ”BaseLine_Source” ”Cuboid”
(position -100 -15 -5)
(position -50 15 25) )
```

```
;—implant
(isedr:define-constant-profile ”Implant_Source” ”ArsenicActiveConcentration” 3e20)
;—placement
(isedr:define-constant-profile-placement ”Placement_Source” ”Implant_Source” ”BaseLine_Source”2)
;—Drain
;—base line
(isedr:define-refinement-window ”BaseLine_Drain” ”Cuboid”
(position (+ 50 L) -15 -5)
(position (+ 100 L) 15 25) )
;—implant (isedr:define-constant-profile ”Implant_Drain” ”ArsenicActiveConcentration”
3e20)
;—placement
(isedr:define-constant-profile-placement ”Placement_Drain” ”Implant_Drain” ”BaseLine_Drain”2)
;—SourceLDD
;—base line
(isedr:define-refinement-window ”BaseLine_SourceLDD” ”Cuboid”
(position 0 -10 0)
(position -50 10 20) )
;—implant
```

```
(isedr:define-constant-profile ”Implant_SourceLDD” ”ArsenicActiveConcentration” 3e20)
;—placement
(isedr:define-constant-profile-placement ”Placement_SourceLDD” ”Implant_SourceLDD”
”BaseLine_SourceLDD”0)
;—DrainLDD
;—base line (isedr:define-refinement-window ”BaseLine_DrainLDD” ”Cuboid”
(position L -10 0)
(position (+ 50 L) 10 20) )
;—implant
(isedr:define-constant-profile ”Implant_DrainLDD” ”ArsenicActiveConcentration” 3e20)
;—placement
(isedr:define-constant-profile-placement ”Placement_DrainLDD” ”Implant_DrainLDD” ”Base-
Line_DrainLDD”0)
(isedr:define-refinement-window ”channel_RF” ”cylinder”
(position -70 0 10)
(position (+ 70 L) 0 10)5)
(isedr:define-refinement-size ”Cha_Mesh”40 40 40 20 20 20)
(isedr:define-refinement-material ”channel_RF” ”Cha_Mesh” ”Silicon” )
;—multibox 1
```

```
(isedr:define-refinement-window ”multibox1” ”Cuboid”
(position -100 -10 0)
(position -30 10 20) )
(isedr:define-multibox-size ”multiboxSize1”40 40 40 20 20 20 )
(isedr:define-multibox-placement ”multiboxPlacement1”
”multiboxSize1” ”multibox1”)
;—multibox 2
(isedr:define-refinement-window ”multibox2” ”Cuboid”
(position (+ 30 L) -10 0)
(position (+ 100 L) 10 20) )
(isedr:define-multibox-size ”multiboxSize2”40 40 40 20 20 20 )
(isedr:define-multibox-placement ”multiboxPlacement2”
”multiboxSize2” ”multibox2”)
;— Save BND and CMD and rescale to um
(ise:assign-material-and-region-names (part:entities (filter:type ”solid?”)))
(isegeo:scale (part:entities (filter:type ”solid?”)) nm nm nm )
(iseio:save-dfise-bnd (part:entities (filter:type ”solid?”))
”@boundary/o@”) (isedr:write-scaled-cmd-file ”@commands/o@” nm)
```

A.2 The Mixed-mode Simulation Command

We will show the mixed-mode simulation command of our simulation. It is defined the 6-T SRAM circuit architecture under the hold or read mode.

A.2.1 The Hold Mode

```
Device NMOS {
Electrode{
{ Name=”source” Voltage=0.0 }
{ Name=”drain” Voltage=0.0 }
{ Name=”gate” Voltage=0.0 Workfunction=4.6}
} File{ Grid = ”@grid—-2@”
Doping = ”@doping—-2@”
Plot = ”@dat@”
Current = ”@plot@”
Param = ”@parameter@”
}
Physics{ Mobility(DopingDependent HighFieldSaturation Enormal)
eQuantumPotential
Recombination(
```

```
SRH(DopingDependence)
eAvalanche
)
}
}
Device PMOS{
Electrode{
{ Name=”source” Voltage=0.0 }
{ Name=”drain” Voltage=0.0 }
{ Name=”gate” Voltage=0.0 Workfunction=4.6}
}
File{ Grid = ”@grid@”
Doping = ”@doping@”
Plot = ”@dat@”
Current = ”@plot@”
Param = ”@parameter@”
}
Physics{
Mobility(DopingDependent HighFieldSaturation Enormal)
```

```
eQuantumPotential
Recombination(
SRH(DopingDependence)
eAvalanche
)
}
}
File{
Output = ”@log@”
}
Plot{
eDensity hDensity eCurrent hCurrent
Potential SpaceCharge ElectricField
eMobility hMobility eVelocity hVelocity
Doping DonorConcentration AcceptorConcentration
}
Math{
Extrapolate
Derivatives
```

```
* Avalderivatives

RelErrControl

Digits=5

ErRef(electron)=1.e10

ErRef(hole)=1.e10

Notdamped=50

Iterations=40

Newdiscretization

Directcurrent

Method=ParDiSo

-VoronoiFaceBoxMethod

NaturalBoxMethod

}

System{

Vsource_pset vdd (dd 0) { dc = 0.0 }

Vsource_pset vwl (T 0) { dc = 0.0 }

Vsource_pset vb (L 0) { dc = 0.0 }

Vsource_pset vbl (R 0) { dc = 0.0 }

Vsource_pset vin (VinL 0) { dc = 0.0 }
```

```
NMOS nmos1 ( ”source”=0 ”drain”=VinR ”gate”=VinL )
NMOS nmos2 ( ”source”=0 ”drain”=VinL ”gate”=VinR )
NMOS nmos3 ( ”source”=VinR ”drain”=L ”gate”=T )
NMOS nmos4 ( ”source”=R ”drain”=VinL ”gate”=T )
PMOS pmos1 ( ”source”=dd ”drain”=VinR ”gate”=VinL )
PMOS pmos2 ( ”source”=dd ”drain”=VinL ”gate”=VinR )
Plot ”n@node@_sys_des.plt” (time() v(in) v(out) i(nmos1,out) i(nmos2,out) i(nmos3,out)
i(nmos4,out) i(pmos1,out) i(pmos2,out) i(cout,out) )
}
Solve{
NewCurrentFile=”init”
Coupled(Iterations=250){Poisson eQuantumPotential}
Coupled{Poisson Electron Hole eQuantumPotential Contact Circuit}
Quasistationary(
InitialStep=1e-3 Increment=1
MinStep=1e-14 MaxStep=0.05
Goal{Parameter=vdd.dc Voltage= @Vdd@}
){Coupled{nmos1.poisson nmos1.electron nmos1.eQuantumPotential nmos1.contact
nmos2.poisson nmos2.electron nmos2.eQuantumPotential nmos2.contact
```

```
nmos3.poisson nmos3.electron nmos3.eQuantumPotential nmos3.contact
nmos4.poisson nmos4.electron nmos4.eQuantumPotential nmos4.contact
pmos1.poisson pmos1.hole pmos1.eQuantumPotential pmos1.contact
pmos2.poisson pmos2.hole pmos2.eQuantumPotential pmos2.contact
circuit}
}
NewCurrentFile=””
Quasistationary(
InitialStep=1e-3 Increment=1
MinStep=1e-14 MaxStep=0.05
Goal{Parameter=vin.dc Voltage= @Vdd@})
){Coupled{nmos1.poisson nmos1.electron nmos1.eQuantumPotential nmos1.contact
nmos2.poisson nmos2.electron nmos2.eQuantumPotential nmos2.contact
nmos3.poisson nmos3.electron nmos3.eQuantumPotential nmos3.contact
nmos4.poisson nmos4.electron nmos4.eQuantumPotential nmos4.contact
pmos1.poisson pmos1.hole pmos1.eQuantumPotential pmos1.contact
pmos2.poisson pmos2.hole pmos2.eQuantumPotential pmos2.contact
circuit}
}}
```

A.2.2 The Read Mode

```
Device NMOS {

Electrode{

{ Name=”source” Voltage=0.0 }

{ Name=”drain” Voltage=0.0 }

{ Name=”gate” Voltage=0.0 Workfunction=4.6}

} File{ Grid = ”@grid—-2@”

Doping = ”@doping—-2@”

Plot = ”@dat@”

Current = ”@plot@”

Param = ”@parameter@”

}

Physics{ Mobility(DopingDependent HighFieldSaturation Enormal)

eQuantumPotential

Recombination(

SRH(DopingDependence)

eAvalanche

)

}
```

```
}
Device PMOS{
Electrode{
{ Name=”source” Voltage=0.0 }
{ Name=”drain” Voltage=0.0 }
{ Name=”gate” Voltage=0.0 Workfunction=4.6}
}
File{ Grid = ”@grid@”
Doping = ”@doping@”
Plot = ”@dat@”
Current = ”@plot@”
Param = ”@parameter@”
}
Physics{
Mobility(DopingDependent HighFieldSaturation Enormal)
eQuantumPotential
Recombination(
SRH(DopingDependence)
eAvalanche
```

```
)
}
}
File{
Output = ”@log@”
}
Plot{
eDensity hDensity eCurrent hCurrent
Potential SpaceCharge ElectricField
eMobility hMobility eVelocity hVelocity
Doping DonorConcentration AcceptorConcentration
}
Math{
Extrapolate
Derivatives
* Avalderivatives
RelErrControl
Digits=5
ErRef(electron)=1.e10
```

```
ErRef(hole)=1.e10
Notdamped=50
Iterations=40
Newdiscretization
Directcurrent
Method=ParDiSo
-VoronoiFaceBoxMethod
NaturalBoxMethod
}
System{
Vsource_pset vdd (dd 0) { dc = 0.0 }
Vsource_pset vwl (T 0) { dc = 0.0 }
Vsource_pset vb (L 0) { dc = 0.0 }
Vsource_pset vbl (R 0) { dc = 0.0 }
Vsource_pset vin (VinL 0) { dc = 0.0 }
NMOS nmos1 ( ”source”=0 ”drain”=VinR ”gate”=VinL )
NMOS nmos2 ( ”source”=0 ”drain”=VinL ”gate”=VinR )
NMOS nmos3 ( ”source”=VinR ”drain”=L ”gate”=T )
NMOS nmos4 ( ”source”=R ”drain”=VinL ”gate”=T )
```

```
PMOS pmos1 ( ”source”=dd ”drain”=VinR ”gate”=VinL )
PMOS pmos2 ( ”source”=dd ”drain”=VinL ”gate”=VinR )
Plot ”n@node@_sys_des.plt” (time() v(in) v(out) i(nmos1,out) i(nmos2,out) i(nmos3,out)
i(nmos4,out) i(pmos1,out) i(pmos2,out) i(cout,out) )
}
Solve{
NewCurrentFile=”init”
Coupled(Iterations=250){Poisson eQuantumPotential}
Coupled{Poisson Electron Hole eQuantumPotential Contact Circuit}
Quasistationary(
InitialStep=1e-3 Increment=1
MinStep=1e-14 MaxStep=0.05
Goal{Parameter=vdd.dc Voltage= @Vdd@}
){Coupled{nmos1.poisson nmos1.electron nmos1.eQuantumPotential nmos1.contact
nmos2.poisson nmos2.electron nmos2.eQuantumPotential nmos2.contact
nmos3.poisson nmos3.electron nmos3.eQuantumPotential nmos3.contact
nmos4.poisson nmos4.electron nmos4.eQuantumPotential nmos4.contact
pmos1.poisson pmos1.hole pmos1.eQuantumPotential pmos1.contact
pmos2.poisson pmos2.hole pmos2.eQuantumPotential pmos2.contact
```

```
circuit}
}
Quasistationary(
InitialStep=1e-3 Increment=1
MinStep=1e-14 MaxStep=0.05
Goal{Parameter=vwl.dc Voltage= @Vdd@}
){Coupled{nmos1.poisson nmos1.electron nmos1.eQuantumPotential nmos1.contact
nmos2.poisson nmos2.electron nmos2.eQuantumPotential nmos2.contact
nmos3.poisson nmos3.electron nmos3.eQuantumPotential nmos3.contact
nmos4.poisson nmos4.electron nmos4.eQuantumPotential nmos4.contact
pmos1.poisson pmos1.hole pmos1.eQuantumPotential pmos1.contact
pmos2.poisson pmos2.hole pmos2.eQuantumPotential pmos2.contact
circuit}
}
Quasistationary(
InitialStep=1e-3 Increment=1
MinStep=1e-14 MaxStep=0.05
Goal{ Parameter=vb.dc Voltage= @Vdd@ }
){Coupled{nmos1.poisson nmos1.electron nmos1.eQuantumPotential nmos1.contact
```

```
nmos2.poisson nmos2.electron nmos2.eQuantumPotential nmos2.contact
nmos3.poisson nmos3.electron nmos3.eQuantumPotential nmos3.contact
nmos4.poisson nmos4.electron nmos4.eQuantumPotential nmos4.contact
pmos1.poisson pmos1.hole pmos1.eQuantumPotential pmos1.contact
pmos2.poisson pmos2.hole pmos2.eQuantumPotential pmos2.contact
circuit}
}
Quasistationary(
InitialStep=1e-3 Increment=1
MinStep=1e-14 MaxStep=0.05
Goal{Parameter=vbl.dc Voltage= @Vdd@}
){Coupled{nmos1.poisson nmos1.electron nmos1.eQuantumPotential nmos1.contact
nmos2.poisson nmos2.electron nmos2.eQuantumPotential nmos2.contact
nmos3.poisson nmos3.electron nmos3.eQuantumPotential nmos3.contact
nmos4.poisson nmos4.electron nmos4.eQuantumPotential nmos4.contact
pmos1.poisson pmos1.hole pmos1.eQuantumPotential pmos1.contact
pmos2.poisson pmos2.hole pmos2.eQuantumPotential pmos2.contact
circuit}
}
```

```
NewCurrentFile=””
Quasistationary(
InitialStep=1e-3 Increment=1
MinStep=1e-14 MaxStep=0.05
Goal{Parameter=vin.dc Voltage= @Vdd@})
){Coupled{nmos1.poisson nmos1.electron nmos1.eQuantumPotential nmos1.contact
nmos2.poisson nmos2.electron nmos2.eQuantumPotential nmos2.contact
nmos3.poisson nmos3.electron nmos3.eQuantumPotential nmos3.contact
nmos4.poisson nmos4.electron nmos4.eQuantumPotential nmos4.contact
pmos1.poisson pmos1.hole pmos1.eQuantumPotential pmos1.contact
pmos2.poisson pmos2.hole pmos2.eQuantumPotential pmos2.contact
circuit}
}}
```

Appendix B

The Code of SNM’s Extraction

Here we describe the code for extraction of SNM for SRAM cells. After we will get the DC transfer characteristics by three-dimension mixed-mode simulation, this code will be performed. We can extract the static noise margin from the DC transfer characteristics. The code is the following :

```
#!/usr/bin/perl

use IO::Handle;

#{{{

$file=shift @ARGV;

open FH, ”<$file”;

while(<FH>){
```

```
chomp;
split /\s+/;
push @x1, $_[0];
push @y1, $_[1];
}
close FH;
$file=shift @ARGV;
open FH, ”<$file”;
while(<FH>){
chomp;
split /\s+/;
push @x2, $_[0];
push @y2, $_[1];
}
close FH;
for ($i=0; $i<= $#x1 ; $i++ ) {
$tmp=”$x1[$i] $y1[$i]\n”;
push @command, $tmp;
}
```

```
$tmp=”e\n”;

push @command, $tmp;

for ($i=0; $i<= $#x1 ; $i++ ) {

$tmp=”$x2[$i] $y2[$i]\n”;

push @command, $tmp;

}

$tmp=”e\n”;

push @command, $tmp;

#}}}

#print ”@x1\n @x2\n”;

print 1／(10**(length($#x1)-1)).”\n”;

$r=10**-2;

for($i=1; $i<$#x1; $i++){

$j=$#x2;

# print ”$x2[$j] $x1[$i] $y2[$j] $y1[$i]\n”;

$flag=0;

while( $x2[$j]<=$x1[$i]){

if($y1[$i]<$y2[$i]){ $flag=$j;}

# norm 2
```

```
$slope=($y1[$i]-$y2[$j])/($x1[$i]-$x2[$j]) if( ($x2[$j]-$x1[$i])!=0);

# print ”$i $j ($x1[$j], $y1[$j]),($x2[$i],$y2[$i])...$slope\n”;

if( abs($slope-1)< $r ){

# print ”$i $j ($x1[$j], $y1[$j]),($x2[$i], $y2[$i])...$slope\n”;

$tmp=sqrt(($y2[$j]-$y1[$i])**2+($x2[$j]-$x1[$i])**2);

push @norm, $tmp;

push @recx1, $x1[$i];

push @recy1, $y1[$i];

push @recx2, $x2[$j];

push @recy2, $y2[$j];

}

$j–;

}

if($flag){ last;}

# if( abs($y1[$i]-$y2[$j])>($r) ){ last; }

}

#print ”@norm\n”;

$max=0;

for($j=0; $j<$#norm+1; $j++){
```

```
if($norm[$j]>=$norm[$max]){

$max=$j;

}

}

$haha=$norm[$max]/sqrt(2);

print ”max $haha [$recx1[$max], $recy1[$max]],[$recx2[$max],$recy2[$max]]\n”;

push @command, ”$recx1[$max] $recy1[$max]\n”;

push @command, ”$recx1[$max] $recy2[$max]\n”;

push @command, ”$recx2[$max] $recy2[$max]\n”;

push @command, ”$recx2[$max] $recy1[$max]\n”;

push @command, ”$recx1[$max] $recy1[$max]\n”;

push @command, ”e\n”;

$#norm=$#recx1=$#recx2=$#recy1=$#recy2=-1;

for($j=0; $j<$#x2; $j++){

if(abs($x1[$i]-$x2[$j])<$r){ $flag=$j; last }

}

#print ”$i......$#x1......$#norm.....$x1[$i] $x2[$flag]\n”;

while($i<=$#x1){

$j=$flag;
```

```
while( $j>=0 ){
# norm 2
$slope=($y2[$j]-$y1[$i])/($x2[$j]-$x1[$i]) if( ($x2[$j]-$x1[$i])!=0);
# print ”$i $j ($x1[$i], $y1[$i]),($x2[$j], $y2[$j])...$slope\n”;
if( abs($slope-1)<$r ){
$tmp=sqrt(($y2[$j]-$y1[$i])**2+($x2[$j]-$x1[$i])**2);
# print ”$i $j ($x1[$i], $y1[$i]),($x2[$j], $y2[$j])...$slope\n”;
push @norm, $tmp;
push @recx1, $x1[$i];
push @recy1, $y1[$i];
push @recx2, $x2[$j];
push @recy2, $y2[$j];
}
$j–;
}
$i++; }
$max=0;
for($j=0; $j<$#norm+1; $j++){
if($norm[$j]>=$norm[$max]){ $max=$j;}
```

```
}
$haha=$norm[$max]/sqrt(2); print ”max $haha [$recx1[$max],
$recy1[$max]],[$recx2[$max],$recy2[$max]]\n”; push @command, ”$recx1[$max] $recy1[$max]\n”;
push @command, ”$recx1[$max] $recy2[$max]\n”;
push @command, ”$recx2[$max] $recy2[$max]\n”;
push @command, ”$recx2[$max] $recy1[$max]\n”;
push @command, ”$recx1[$max] $recy1[$max]\n”;
push @command, ”e\n”;
#### gnuplot ####
my $PIPE=new IO::Handle;
open $PIPE, ”—gnuplot -persist” or die ”\n$0 : failed to open pipe to \”gnuplot\” :
$!\n”;
$PIPE->autoflush(1);
print $PIPE <<END;
set terminal x11;
set xlabel \”V gate\”;
set ylabel \”V drain\”;
set xrange [];
set yrange [];
```

```
set size square;
END
#print $PIPE ”set terminal x12;\n”;
#print $PIPE ”set xlabel \”TE (ms)\”;\n”;
#print $PIPE ”set ylabel \”Signal Intensity\”;\n”;
#print $PIPE ”set xrange [];\n”;
#print $PIPE ”set yrange [];\n”;
print $PIPE ”plot ’-’ using 1:2 with lines, ’-’ using 1:2 with lines, ’-’ using 1:2 with lines, ’-’ using 1:2 with lines;\n”;
#push @command, ”plot ’-’ using 1:2 with lines\n”;
print $PIPE @command;
#print $PIPE ”pause 3;\n”;
close $PIPE; ##### end #######
exit;
my $file, %h2, %h3;
$file=shift @ARGV;
open FH, ”<$file”;
$buf=<FH>;
$len=0;
```

```
while($buf=<FH>){
$len++;
split ／\s+／, $buf;
push @v12, $_[2];
push @v13, $_[3];
}
print ”..@v12\n”;
print ”...@v13\n”;
close(FH);
$file=shift @ARGV;
print ”$file\n”;
open FH, ”<$file”;
$buf=<FH>;
$len=0;
while($buf=<FH>){
$len++;
split ／\s+／, $buf;
push @v22, $_[2];
push @v23, $_[3];
```

```
}
print ”..@v22\n”;
print ”...@v23\n”;
close(FH);
for($i=0; $i<$len; $i++){
$x2=$v12[$i];
$y2=$v13[$i];
print ”$i \n”;
for($j=$len-1; $j>=0; $j–){
$x3=$v22[$j];
$y3=$v23[$j];
$slope=($y3-$y2)/($x3-$x2) if( ($x3-$x2)!=0);
if( abs($slope-1)<0.01){
print ”...$slope $x2,$y2 $x3,$y3”;
push @recx2, $x2;
push @recy2, $y2;
push @recx3, $x3;
push @recy3, $y3;
$tmp=sqrt(($y3-$y2)**2+($x3-$x2)**2);
```

```
print ” $tmp\n”;

push @norm, $tmp;

}

}

} $j=0;

while($j<2){

$max=0;

for($i=0; $i<$#norm+1; $i++){

if($norm[$i]>$norm[$max]){ $max=$i;}

}

print ”max $norm[$max] [$recx2[$max],

$recy2[$max]],[$recx3[$max],$recy3[$max]]\n”;

splice @norm, $max, 1;

splice @recx2, $max, 1;

splice @recy2, $max, 1;

splice @recx3, $max, 1;

splice @recy3, $max, 1;

$j++;

}
```

VITA

Name: Chien-Sung Lu, 呂建松

Permanent address: No. 1-2, Sanyue Rd., Taichung, Taiwan

Degree and date to be conferred: Master of Electrical & Computer Engineering, July, 2006

Date of birth: July 31st , 1982

Place of birth: Taichung, Taiwan

Collegiate institutions attended	Degree	Date of graduate
Department of Mathematics National Cheng Kung University, Tainan, Taiwan	BS	June, 2004
Department of Electronics Engineering National Chaio Tung University, Hsinchu, Taiwan	MS	July, 2006

Master thesis title: Performance of SRAM with Nanoscale Transistors