Statistical Modelling of Nano CMOS Transistors with Surface Potential Compact Model PSP

Daryoosh Dideban

Submitted in fulfilment of the requirements for

the degree of Doctor of Philosophy in Electronics and Electrical Engineering

School of Engineering

College of Science and Engineering

University of Glasgow

December 2011

Abstract

The development of a statistical compact model strategy for nano-scale CMOS transistors is presented in this thesis. Statistical variability which arises from the discreteness of charge and granularity of matter plays an important role in scaling of nano CMOS transistors especially in sub 50nm technology nodes. In order to achieve reasonable performance and yield in contemporary CMOS designs, the statistical variability that affects the circuit/system performance and yield must be accurately represented by the industry standard compact models. As a starting point, predictive 3D simulation of an ensemble of 1000 microscopically different 35nm gate length transistors is carried out to characterize the impact of statistical variability on the device characteristics. PSP, an advanced surface potential compact model that is selected as the next generation industry standard compact model, is targeted in this study. There are two challenges in development of a statistical compact model strategy. The first challenge is related to the selection of a small subset of statistical compact model parameters from the large number of compact model parameters. We propose a strategy to select 7 parameters from PSP to capture the impact of statistical variability on current-voltage characteristics. These 7 parameters are used in statistical parameter extraction with an average RMS error of less than 2.5% crossing the whole operation region of the simulated transistors. Moreover, the accuracy of statistical compact model extraction strategy in reproducing the MOSFET electrical figures of merit is studied in detail. The results of the statistical compact model extraction are used for statistical circuit simulation of a CMOS inverter under different input-output conditions and different number of statistical parameters. The second challenge in the development of statistical compact model strategy is associated with statistical generation of parameters preserving the distribution and correlation of the directly extracted parameters. By using advanced statistical methods such as principal component analysis and nonlinear power method, the accuracy of parameter generation is evaluated and compared to directly extracted parameter sets. Finally, an extension of the PSP statistical compact model strategy to different channel width/length devices is presented. The statistical trends of parameters and figures of merit versus channel width/length are characterized.

Acknowledgements

I would like to thank, first and foremost, my supervisors, Prof. Asen Asenov and Dr. Binjie Cheng for their guidance, encouragement and support throughout my research. Professor Asenov helped me to achieve more knowledge in the wonderful world of device physics. He was an excellent leader and a kind father for the members of the Device Modelling Group. Valuable technical discussions with Dr. Cheng learned me a lot in the field of compact modelling. I had the chance to meet him at any time and get the advice for my research challenges. I would like also to acknowledge the other members of Device Modelling Group for their cooperation and support. And a special thank to my beloved wife, Negin, who was my colleague and without her patience and support this work would not have been possible.

Publications

1- **D. Dideban**, B. Cheng, N. Moezi, N. A. Kamsani, C. Millar, S. Roy and A. Asenov, "Impact of Input Slew Rate on Statistical Timing and Power Dissipation Variability in nano CMOS" in *Proceedings of Ultimate Integration on Silicon (ULIS)*, Glasgow, UK, Mar. 17-19, 2010, pp. 45–48.

2- **D. Dideban**, B. Cheng, N. Moezi, X. Wang, A. Asenov, "Evaluation of 35nm MOSFET Capacitance Components in PSP Compact Model", *Proceedings of ICEE 2010*, Isfahan, Iran, 11-13 May 2010.

3- B. Cheng, **D. Dideban**, N. Moezi, C. Millar, G. Roy, X. Wang, S. Roy and A. Asenov, "Capturing Intrinsic Parameter Fluctuations using the PSP Compact Model" in *Proceedings of Design, Automation and Test in Europe (DATE)*, Dresden, Germany, Mar. 8-12, 2010, pp. 650–653.

4- B. Cheng, **D. Dideban**, N. Moezi, C. Millar, G. Roy, X. Wang, S. Roy and A. Asenov, "Statistical Variability Compact Modeling Strategies for BSIM4 and PSP", *IEEE Journal of Design and Test of Computers*, Vol. 27, issue 2, pp. 26–35, Mar. /Apr. 2010.

5- U. Kovac, **D. Dideban**, B. Cheng, N. Moezi, G. Roy and A. Asenov, "A Novel Approach to the Statistical Generation of Non-normal Distributed PSP Compact Model Parameters using a Nonlinear Power Method" in *Proceedings of Simulation of Semiconductor Processes and Devices (SISPAD)*, Bologna, Italy, Sept. 6-8, 2010, pp. 125–128.

6- B. Cheng, N. Moezi, **D. Dideban**, G. Roy, S. Roy and A. Asenov, "Benchmarking the Accuracy of PCA Generated Statistical Compact Model Parameters Against Physical Device Simulation and Directly Extracted Statistical Parameters" in *Proceedings of Simulation of Semiconductor Processes and Devices (SISPAD)*, Sept. 9-11, 2009, pp. 143–146.

7- A. Asenov, B. Cheng, **D. Dideban**, U. Kovac, N. Moezi, C. Millar, G. Roy, A. Brown and S. Roy, "Modeling and Simulation of Transistor and Circuit Variability and Reliability"**,** *Custom Integrated Circuit Conference (CICC)*, San Jose, California, Sept. 19-22, 2010.

8- N.Moezi, **D.Dideban**, B.Cheng, S.Roy, A.Asenov, "Impact of Statistical Parameter Set Selection on the Statistical Compact Model Accuracy: BSIM4 and PSP Case Study", *Microelectronics Journal*, in press (corrected proof).

* Publications numbered 2,3,4,5,6,7 are indexed by IEEE and can be accessed from: http://www.ieeexplore.ieee.org

* Publication number 8 is indexed by Elsevier and can be accessed from: http://www.journals.elsevier.com/microelectronics-journal/

Contents

List of Figures

List of Tables

Nomenclature

AFM	Atomic Force Microscope
BPV	Backward Propagation of Variance
BTE	Boltzman Transport Equation
CESL	Contact Etch Stop Layer
CMP	Chemical Mechanical Polishing
CM	Compact Model
CLM	Channel Length Modulation
DD	Drift Diffusion
DG	Density Gradient
DIBL	Drain Induced Barrier Lowering
EDA	Electronic Design Automation
GA	Genetic Algorithm
GCA	Gradual Channel Approximation
GIDL	Gate Induced Drain Lowering
HKMG	High K Metal Gate
IPF	Intrinsic Parameter Fluctuations

ITRS	International Technology Roadmap for Semiconductors
KS	Kolmogorov-Smirnov (statistical test)
LM	Levenberg-Marquardt
LER	Line Edge Roughness
OTF	Oxide Thickness Fluctuations
OPC	Optical Proximity Effects
PCA	Principal Component Analysis
PSP	Name of a surface potential MOSFET compact model
PGG	Polysilicon Gate Granularity
PSO	Particle Swarm Optimization
MC	Monte Carlo
MOS	Metal-Oxide-Semiconductor
MGG	Metal Gate Granularity
NPM	Nonlinear Power Method
NBTI	Negative Bias Temperature Instability
Q-Q	Quantile-Quantile (statistical plot)
RDF	Random Dopant Fluctuations
RDD	Random Discrete Dopants
RMS	Root Mean Square
RSM	Response Surface Methodology
RTA	Rapid Thermal Annealing
SV	Statistical Variability
SD	Standard Deviation
SS	Subthreshold Slope
SCM	Statistical Compact Model

SEM	Scanning Electron Microscope
SLM	Symmetric Linearization Method
SRAM	Static Random Access Memory
STI	Shallow Trench Isolation
STA	Statistical Timing Analysis
TCAD	Technology Computer Aided Design
WPE	Well Proximity Effects

Chapter 1

Introduction

The aim of this PhD research is to develop a statistical compact modelling framework capable of capturing the impact of statistical variability on deca-nanometer scale bulk MOSFETs. The focus is on the adoption of an advanced surface potential compact model PSP, which is the next generation industry standard compact model. A brief overview of the challenges in the context of statistical compact modelling in ultra-scaled MOSFETs reflects our motivation in this study. Following this, we present the aims and objectives of this study, in line with the demands of a state of the art statistical compact model. The introduction is completed with an outline of subsequent chapters.

1.1 Motivation

Compact Models (CMs) are explicit description of device physics in terms of a set of parameters. They act as key components of the interface between technology and design. Although the initial impetus behind CM development was the requirement to accurately model circuit components in analog circuit design, CMs are also extensively used in transistor-level digital circuit design and verification, especially in standard cell characterization procedures. The importance of device matching properties in the analog domain has driven initially the transistor mismatch modelling efforts. The first systematic mismatch models were reported in the early 1980s for MOS capacitors and MOSFETs [1], and still today this remains an active research area [2,3].

Most MOSFET mismatch models are based on simple drain current analytical or

numerical models [4]. Although such models could provide critical information regarding the trends of transistor mismatch as a function of the device size, they cannot be integrated into design tools to directly support design activities.

A natural way to incorporate the mismatch into the design flow is to employ statistical compact modelling techniques. Most CM-based mismatch approaches assume normal, uncorrelated distribution of CM parameters [5]. With the scaling of the CMOS transistors to deca-nanometer dimensions, the statistical variability (SV) which was exclusive to the analog domain has now entered the digital domain, affecting the performance and yield of digital circuits and systems [6,7,8]. Furthermore, SV in nano-CMOS devices which arises from discreteness of charge and granularity of matter does not follow normal distribution [9]. Thus statistical compact model approaches based on the assumption of uncorrelated normal distributions of the statistical compact model parameters could introduce considerable errors in statistical circuit simulations.

Achieving reasonable performance and yield in contemporary CMOS design necessitates the use of transistor compact models that can accurately reproduce the impact of SV on circuit performance and yield. Thus, the investigation of statistical compact modelling strategies that are flexible and accurate, yet economical is of great importance for variability aware design. This is due to the fact that the overall accuracy of circuit/system simulation is determined by the accuracy of circuit component models in the presence of SV. Moreover, development of tools to predict yield loss caused by SV needs to be carried out using these statistical compact models. Finally, to forecast the SV for technology generations ahead, it is essential to develop variability aware design strategies based on statistical compact models.

In order to facilitate the process of developing statistical compact models in line with the aforementioned demands, we have used an advanced surface potential based compact model PSP [10], as a test-bed compact model. This new compact model has many advantages compared to the traditional compact models like BSIM [11], including physics-based expressions, symmetry and reciprocity of MOSFET trans-capacitance components and smooth transition from sub-threshold to strong inversion [12,13].

1.2 Aims and Objectives

The main goal of this project is to develop a statistical compact model strategy which will be based on the surface potential compact model, PSP. The proposed statistical compact model strategy should capture accurately and efficiently the effects of statistical variability in contemporary and next generation bulk deca-nanometer MOSFETs. This will be achieved by:

1- Development of strategy and methodology for extraction of an accurate nominal PSP model parameter set based on the simulation of a template bulk MOSFET.

2- Calibration of the nominal PSP model parameter set to reproduce accurately the MOSFET trans-capacitance components and evaluate the accuracy of the transient time SPICE simulations in the case of deca-nanometer bulk MOSFETs.

3- Identification of an optimal set of statistical compact model parameters based on the physical simulation of the statistical variability and sensitivity analysis of the compact model parameters.

4- Research on development of efficient and accurate techniques for the generation of statistical compact model parameters using principal component analysis (PCA) and nonlinear power method (NPM).

5- Investigation of the impact of statistical compact model parameter selection and statistical parameter generation techniques on the accuracy of reproducing MOSFET electrical figures of merit and also on the accuracy of statistical circuit simulations.

6- Extension of proposed statistical compact model strategy to characterize the statistical variability of transistors with different channel width/lengths.

1.3 Outline

The main chapters of this thesis are organized as follows:

In chapter 2 we first review existing challenges of CMOS scaling into deca-nano meter regime with emphasis on the statistical variability. We classify the variability into different axes and then we will enumerate the important sources of systematic and statistical variability. We introduce the 'atomistic' simulation of statistical variability using the drift-diffusion technique with density gradient quantum corrections. This chapter is completed with a section on the statistical modelling concepts including mismatch models, corner models, numerical and analytical models.

Chapter 3 is devoted to the PSP parameter extraction and optimization. It introduces the advanced surface potential compact model PSP, its structure, and important expressions. The design of a uniform 35nm gate length MOSFET is introduced as a test bed device in this study. The procedure of DC parameter extraction to obtain an accurate compact model parameter set is then discussed. The important parameters in the AC part of the compact model are extracted to match trans-capacitance components in respect to physical TCAD simulations. The accuracy of these components in transient time SPICE simulations is carefully evaluated.

Chapter 4 is an extensive study of statistical atomistic simulation and parameter extraction for the 35nm template MOSFET. A method of identification of the most responsible compact model parameters using sensitivity analysis is presented followed by the method and strategy of statistical parameter extraction, with an emphasis on the impact of initial conditions on the accuracy of statistical parameter extraction. MOSFET electrical figures of merit are simulated using statistical compact model parameter to evaluate the accuracy of direct parameter extraction approach in respect to atomistic simulation results. Statistical circuit simulation for a CMOS inverter is carried out to investigate the impact of different number of parameters in statistical parameter sets and the input-output conditions on statistical distribution of the delay and the energy of inverter.

Chapter 5 introduces efficient statistical parameter generation techniques. We first investigate the statistical properties of directly extracted parameters. The statistical

distribution and correlation between pairs of parameters are used as the input information of different parameter generation techniques. Gaussian parameter generation ignores the correlation between parameters but is simple. Principal Component Analysis (PCA) maintains the correlation between parameters assuming Gaussian distributions. The Nonlinear Power Method (NPM) is introduced to improve the accuracy of parameter distributions by considering four moments of each parameter distribution. The accuracy of each parameter generation technique in reproducing the MOSFET figures of merit is examined. Statistical circuit simulation of a CMOS inverter is used to evaluate the impact of each parameter generation technique on the accuracy in reproducing the statistical distribution of the inverter delay and energy.

Chapter 6 extends the statistical compact model strategy to devices with different width or length. Two methods are used to investigate the impact of the transistor width on figures of merit. Atomistic simulations of different width/length devices have been carried out to obtain the most accurate results and the slicing method is introduced to facilitate production of statistical characteristics of wider devices. The statistical behaviour of statistical compact model parameter versus width/length is plotted and the accuracy of the slicing method in circuit simulation is investigated. Finally, a non-integer width ratio device parameter generation strategy is developed by using the interpolation of the statistical properties of parameters.

Chapter 7 concludes the findings in this research and proposes possible future work in this area.

Chapter 2

Background

The evolution of electronics has been mainly enabled by progressive scaling of Complementary Metal Oxide Semiconductor (CMOS) transistors to achieve higher circuit density, lower power and better performance [14]. With continued scaling of MOSFET devices to sub-30nm dimensions, the historical growth captured by Moore's famous law [15,16], doubling number of transistors per unit area and increased performance by about 40% in each new generation, becomes difficult to sustain. The semiconductor industry is facing fundamental challenges at technology and device level which impedes the design of the next generations of integrated circuits and systems by implementing the conventional scaling theory. Increasing sub-threshold and gate leakage currents, carrier mobility degradation, hot carrier effects, direct drain to source tunnelling, gate depletion, parasitic resistance and capacitance, leakage junction currents and gate-induced drain leakage (GIDL) are among the challenges after the end of the 'happy scaling' period [17].Various innovations in materials and device structures have been introduced to extend lifespan of conventional bulk MOSFETs.

Figure 2.1 shows the cross section of a typical advanced CMOS transistor, along with the major technology enhancements which have been used to overcome scaling limitations of these devices. High-k dielectrics have been introduced in place of traditional SiO_2 to reduce gate leakage current which significantly impacts on the static power dissipation and proper device operation [18]. The higher permittivity material allows a thicker dielectric which is more resistant to tunnelling while maintains the high gate capacitance. Recent

high-k materials are HfO_2 ($\varepsilon_k \approx 22$) and $HfSiO_x$ ($\varepsilon_k \approx 12-16$) [19].

The poly silicon gate depletion effect dramatically impacts on the gate capacitance and drive current. This was initially improved by increasing the poly-silicon doping but this solution was temporary as the doping concentrations were close to the equilibrium solid solubility in silicon. Hence, metal gates were introduced to remove the gate depletion effects [20]. To mitigate short channel effects (SCE), advanced MOSFETs use ultra-shallow source/drain extensions with non-uniform retrograde doping profile in the channel and pocket halo implant adjacent to source/drain regions [21]. Moreover, strain engineering is introduced to enhance the state-of-the art MOSFET's drivability [22].

However, the complexities associated with advanced processing technology and new materials along with atomistic and quantum mechanical limitations have introduced an increasing amount of statistical variability which in turn leads to device performance variability. This type of variability can no longer be modelled with conventional worst-case circuit design techniques [23,24]. Increasing performance variability has become a critical issue in scaling and integration for the present and next generation of nano CMOS transistors and circuits [25,26].

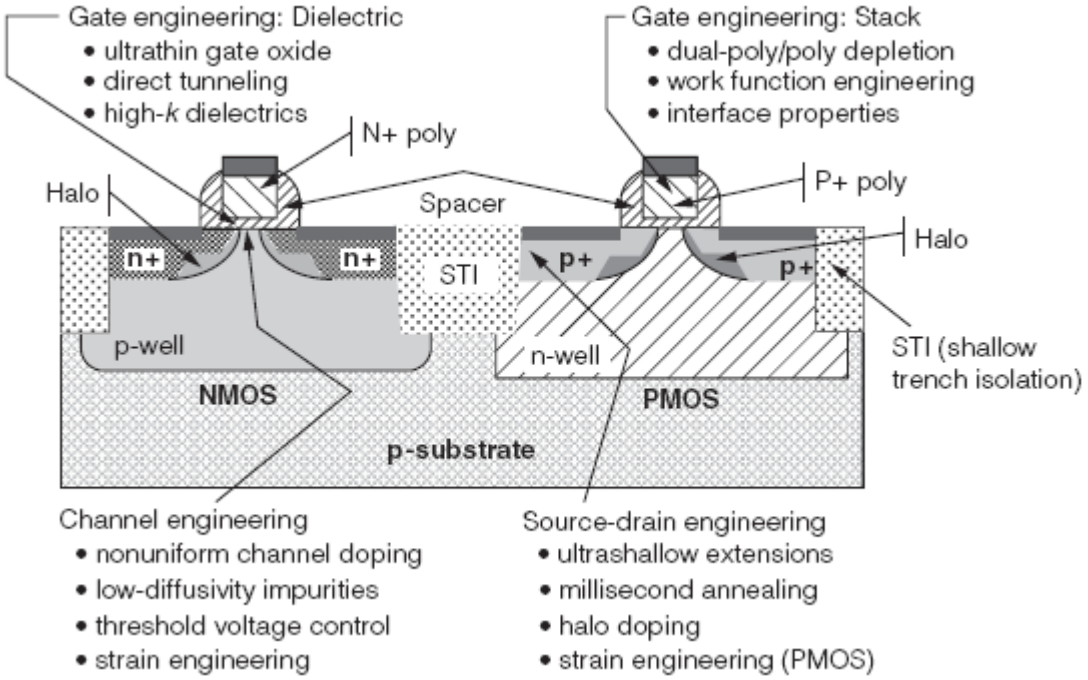


Figure 2.1: Cross section of a typical advanced CMOS device with associated complexities, [23].

While 35nm gate length MOSFETs are in mass production in the 45nm technology generation, the impact of variability keeps increasing as CMOS technologies continue to scale down. The effects of variability was a concern previously in analogue design [27], however it progressively introduces leakage and timing uncertainties in digital circuits [28] and critically affects SRAM cell scaling [29].

In this chapter, we will focus on statistical variability in CMOS technology which becomes the dominant source of variability in 45-nm technology generation and beyond [30,31]. In the first section, we provide a classification of variability in contemporary CMOS devices and circuits, and then enumerate some important sources of systematic and statistical variability. In the second section, we review the important sources of statistical variability in nano CMOS transistors. In the third section, incorporating of statistical variability into Glasgow University atomistic simulator will be discussed. This chapter ends with a review on the existing statistical modelling approaches to account the impact of device variability into circuit operation.

2.1 Variability Classification

The general word of 'variability' often confuses the device/circuit designers because it is not accurate unless it is clearly defined. It is generally agreed that the variability effects can be divided along three different axes [32]:

- Time independent versus time dependent effects.
- Global Variations versus local variations.
- Deterministic (systematic) versus random (statistical) effects.

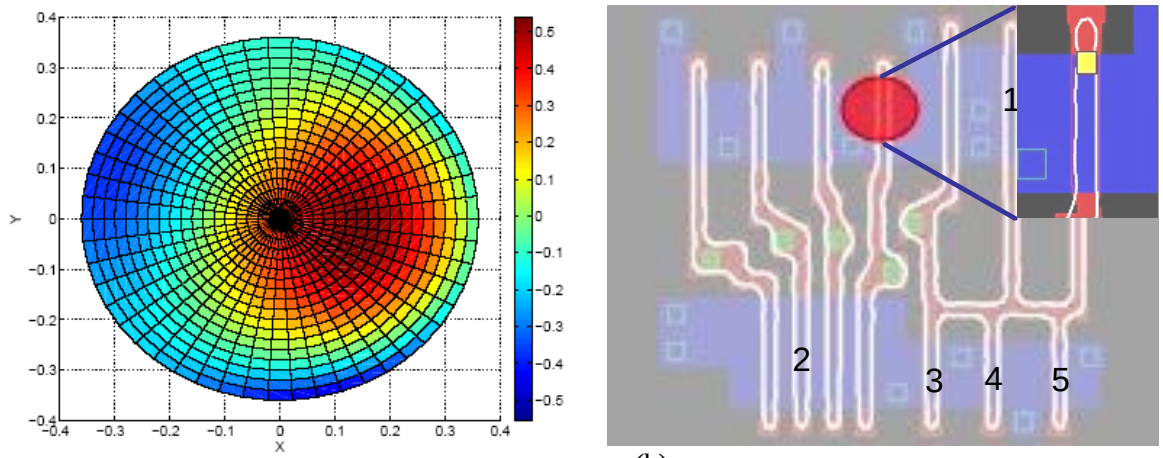
To mitigate the impact of the variability in design, a successful model has to be developed and the designer must choose appropriate design solutions. Therefore, specific properties of each of the variability axis must be considered prior to any modelling or design solution effort. The first axis defines the variability effects in respect to their time domain behaviour. The variability introduced by hot carriers, NBTI, noise, jitter, temperature gradients, wiring IR drop and soft breakdown varies with time, while Random Dopant Fluctuations (RDF), Well Proximity Effect (WPE) and STI stress are static effects.

The second axis is global versus local variations. While local variability refers to the variability of identical MOSFETS within a short distance (inside a chip), the global variability refers to the changes for identical transistors from wafer to wafer or chip to chip (across a wafer). The global variability is caused by inaccuracy in the process parameters control and non-uniformity of equipment and results in slow variation of device dimensions, layer thicknesses and doping concentrations and the corresponding electrical parameters [23]. It introduces change of the mean values of transistor geometry parameters like L (channel length), W (channel width), doping concentrations and oxide thickness which in turn leads to device parameter variations across the chip, from chip to chip on a single wafer and from wafer to wafer.

The third axis which is limited to time-independent variations, refers to the variability from statistics point of view. Effects such as STI stress and WPE are deterministic (systematic) while others like RDF, Line Edge Roughness (LER) and Polysilicon Gate Granularity (PGG) are random (statistical). The systematic variability can be reduced by layout design compensation rules or tightening fabrication process control while such effects have no impact on the statistical variability. Since the random variations arise from discreteness of charge and granularity of matter, they are intrinsic to transistor and are commonly referred to as intrinsic parameter fluctuations [33].

2.1.1 Sources of Systematic Variability

Figure 2.2 illustrates two significant types of systematic variability induced by process and layout and at global and local levels. Figure 2(a) shows the frequencies of ring oscillators in a 300mm wafer [34]. Since the ring oscillators are very sensitive to gate length variation, the color bar which represents the frequency variations (in MHz), indirectly shows the gate length variation across the wafer. These long range effects are different from local variations and are often called global variations. In design stage, they can be dealt with by the definition of process corners [32]. Figure 2(b) shows the local layout induced variability inside a chip, caused by the lithography process [30]. The dimensional deformations will exist even when using Optical Proximity Correction (OPC) techniques [35] and as a result, the shapes of supposedly identical transistors will be different after fabrication. OPC partially compensates the photolithography errors due to diffraction and the remaining errors are predictable when using simulation techniques.



(a) (b)

Figure 2.2: Examples of systematic variability induced by: (a)-process [34], (b)-layout [35].

The introduction of strain to improve the carrier mobility and device drive current at 90nm technology node caused another important source of systematic variability [36]. Impact of stress related to the etch-stop layer or Shallow Trench Isolation (STI) has been subject of various studies [37,38]. An analytical formulation for the variability of device threshold voltage and carrier mobility induced by STI is given in [39]. A threshold voltage variation of 10mV and maximum current deviation of 12% in a specific technology with STI are reported in [38]. Variation associated with embedded silicon germanium (eSiGe) is discussed in [40]. Metal layers are another source of variation in the stress pattern in MOSFETs because they cause mobility reduction due to incomplete annealing of interface states [41].

The Well Proximity Effect (WPE) is caused by the interaction of implanted ions in the formation of the wells and the photoresist boundary [32]. It results in lateral non-uniform doping in the well region and causes MOSFET electrical characteristic variations [42].

It should be noted that the interconnect variability can result in systematic circuit performance variability [43,44]. For a real interconnect structure in 0.18 micrometer technology, an overall delay variation of 18% is reported in [45], from which 48% is due to device variation and 52% is related to interconnect variation. An important source of interconnect systematic variability is Chemical-Mechanical Polishing (CMP) which is used in different steps of the fabrication process to provide smooth and planar surfaces from which subsequent layers are fabricated. The primary effect of CMP is that the metal lines can be dished and eroded when polished and this results in the variation of the copper line

thicknesses [46]. An important manifestation of interconnect variability in digital circuits is clock skew [43,47] which is the difference between propagation times of a single clock to two similar destination points. A list of mitigation strategies to reduce systematic variability in the device and circuit level can be found in [46,32].

2.1.2 Sources of Statistical Variability

The concept of statistical variability is relatively new compared with the systematic variability which has been researched from early stages of the semiconductor industry. For example, a comprehensive research in the impact of process variations on MOSFET threshold voltage has been published in 1974 [48]. An attempt to model process variation with a TCAD simulator was published in 1984 [49]. The subject of statistical variability became important since the introduction of the deca-nanometer CMOS technologies. This is due to the fact that scaling forced the device feature size to become comparable to countable multiples of inter-atomic distances [50]. Although using high-k dielectric and metal gate material had significant impact on the reduction of statistical variation in 45-nm generation technology node compared with 65-nm technology, the magnitude of statistical variability can still be 40% of the systematic counterpart, as reported in [51]. Statistical variability will be of great importance with further shrinking of the device dimensions in near future because discreteness of charge and matter naturally have stronger effect in reduced channel length technologies, and hence its simulation and modeling is becoming one of the hot topics in both industry and academia. Three main sources of statistical variability will be discussed in this section. They are Random Dopant Fluctuations (RDF), Line Edge Roughness (LER) and Polysilicon or Metal Gate Granularity (PGG/MGG).

RDF is major source of statistical variability in conventional bulk MOSFETs and is caused by the variations in the position and number of dopant atoms along the channel and in the S/D regions [52]. It alters the conductivity of the channel as the device turns on, and also affects the leakage current of the device when it turns off. In modern MOSFETS there are many doping stages. Creation of well regions for CMOS technology, halo/pocket doping used for reduction of short channel effects and definition of source/drain regions are some examples that all needs to be introduced by ion implantation techniques. The impurity atoms are introduced in the different device regions with this technique and then annealing is performed to activate the ions. The exact location of each ion after

implantation process is not deterministic because there will be a lot of scattering events when the ion penetrates through the silicon surface. Apart from that, annealing is a high temperature process which results in the diffusion of the implanted ions inside the silicon. The overall result of this implantation/annealing process will be a random dopant distribution for each device among an ensemble of macroscopically identical devices. Figure 2.3 shows a typical 35nm gate length transistor generated by the Glasgow University 'atomistic' simulator with the effect of RDD shown in the bulk, source and drain regions [53]. While the number of dopant atoms in the channel region of a 1-micron transistor was about 5000, this number is reduced to about 100 in 35nm gate length transistor. Therefore, the accurate position of individual dopants will affect the device behavior. This in turn leads to the fact that for two transistors in identical dimension and process conditions, the electrical characteristics will be different.

Line Edge Roughness (LER) is another important source of statistical variability which causes fluctuations in the local gate length which is a critical device parameter. LER arises from the granular nature of the photoresist material and subwavelength lithography used in the fabrication process. Since in modern CMOS processes, source/drain regions are self aligned with the gate edge, LER affects both the gate length and source/drain junctions.

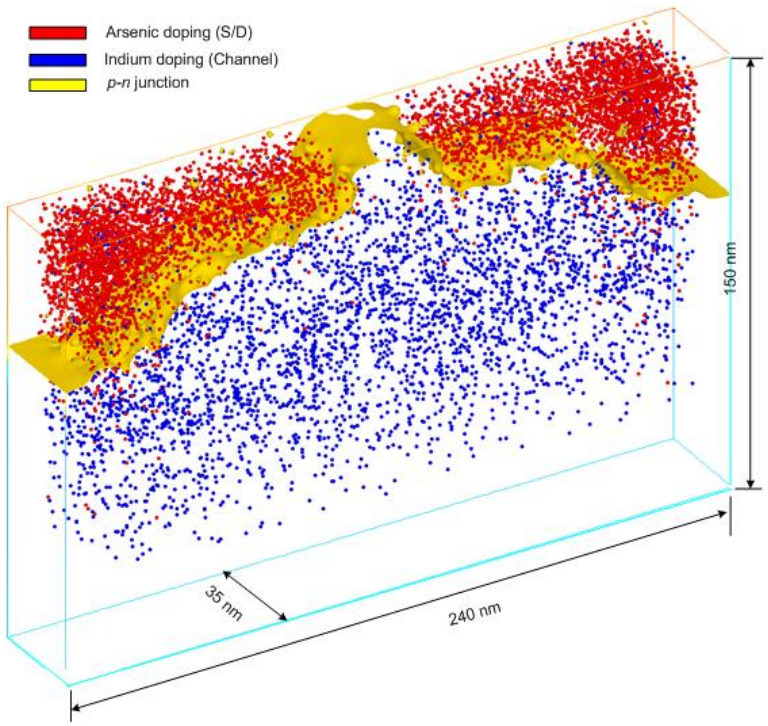
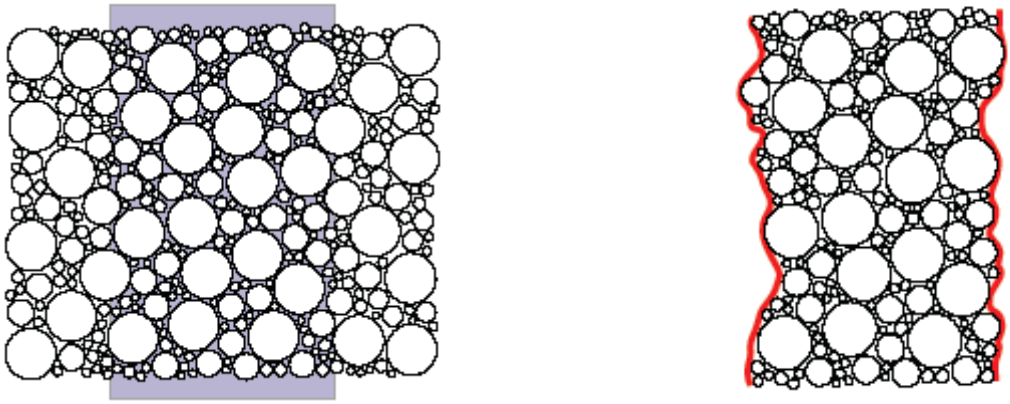


Figure 2.3: The position of random discrete dopants in a typical 35nm gate length N-MOSFET. Blue circles are acceptors in bulk/channel and red circles are donors in source/drain [53].

The polymer chemistry of the photoresist used for subwavelength lithography with 193nm light source imposes LER on the order of 5nm in gate edges [54]. Figure 2.4 illustrates the LER effects in the gate edges caused by photolithography [55]. The other important source of the variability is the Polysilicon Gate Granularity (PGG). An atomic force microscope (AFM) image of PGG is shown in Figure 2.5 [56]. Since polysilicon has a high density of defects at the grain boundaries, a physical phenomena called 'Fermi level pinning' occurs between grains which in turn can cause threshold voltage fluctuations [57]. In new process technologies and device architectures, polysilicon has been replaced by metal gates. However, depending on the actual fabrication procedures, it can introduce a new variability source: Metal Gate Granularity (MGG) [58]. Moreover, Oxide Thickness Fluctuations (OTF) is also reviewed as a source of statistical variability in [59]. By evaluating the impact of individual and combined sources of intrinsic parameter fluctuation (IPF) sources on a 45-nm technology node NMOSFET, it has been shown that RDD, LER and PGG are statistically independent [60].



(a) (b)

Figure 2.4: (a)-A negative photoresist is laid down over active region to be used for exposure using a mask, (b)-pattern of gate edges after removing unexposed region [55].

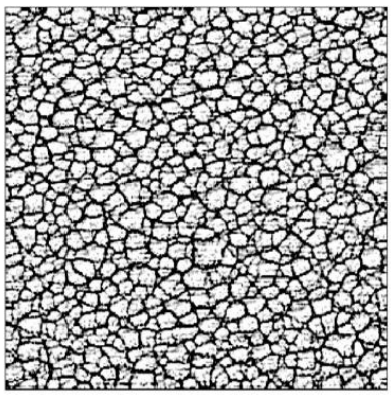


Figure 2.5: An AFM image of ploy silicon gate granularity [56].

2.2 Atomistic Simulation of IPF

In order to study the impact of intrinsic parameter fluctuations (IPF) on the device behaviour, it is important to use predictive simulations to capture 'atomistic' effects of different sources of statistical variability. Moreover, such predictive simulations enable the investigation of the statistical properties of IPF and highlight the role of individual sources of variability on the device electrical parameters. In the first subsection, we review some aspects of the simulation techniques used in the variability analysis of sub-100nm MOSFETs with emphasis on the drift diffusion approach. Using density gradient quantum correction is a key component of this approach and will be discussed in more detail. In the second subsection, we discuss the main equations embedded in the heart of Glasgow University atomistic simulator which will be used in Chapter 4 for the statistical simulation of intrinsic parameter fluctuations. In the third subsection, the introduction of different statistical variability sources into atomistic simulator will be reviewed.

2.2.1 Simulation Technique

Various techniques have been used for the simulation of decananometer MOSFETs [61] including Drift-Diffusion (DD) simulations, Monte Carlo (MC) and Quantum Mechanical (QM) approaches. From a computational efficiency point of view, 3D DD simulations of a MOSFET needs a couple of hours to run on a CPU cluster and as a result is the most efficient and feasible method for simulation of thousands of identical devices with microscopically different sources of variability [55]. DD simulations have been used in combination with density gradient quantum corrections for 3D simulations of sub-100nm MOSFETs in presence of statistical variability [52,54,62]. The DD method provides accurate results in the sub-threshold region of MOSFET characteristics while MC simulations are more accurate in the above-threshold region [55]. MC techniques need a lot of computational resources and thus they are expensive [63].

The DD equations are obtained from the Boltzmann Transport Equation (BTE) following some approximations [55] and as a result the electron and hole currents are approximated using two components: a drift component caused by electric field and a diffusion component as a result of carrier density gradient. Sum of these two components

give the total current through the device. For example in NMOSFET device, drift and diffusion components are given by:

$$J_{n,drift} = qn\mu_n E = -qn\mu_n \nabla \psi \tag{2.1}$$

$$J_{n,diff} = qD_n \nabla n \tag{2.2}$$

where q is absolute charge of electron and n is the electron concentration. E is the electric field which equals to the gradient of electrostatic potential, ψ. Coefficients μ and D denote to electron mobility and electron diffusion constant, respectively. In Boltzmann statistics approximation, they are related via the Einstein's relation:

$$\frac{D_n}{\mu_n} = \frac{kT}{q} \tag{2.3}$$

where k is Boltzmann constant and T is the absolute temperature. The assumptions employed in deriving DD equations from BTE, limit the validity of the DD model [64]. However, empirical mobility models and quantum corrections will improve the validity of DD model to take into account effects such as high field carrier velocity saturation and quantum confinement. The mobility model published by Caughey and Thomas [65] can be used to represent smooth mobility behavior in transition from low to high field:

$$\mu(E) = \mu_o [1 + (\frac{\mu_o E}{v_{sat}})^{\beta}]^{-1/\beta} \tag{2.4}$$

where E is the parallel electric field and μ_o is the low field mobility. Also, v_{sat} is the saturation velocity and β is a constant equal to 2 for electrons and 1 for holes. In order to take into account the concentration dependency of the mobility, the zero field mobility is described by:

$$\mu_o = \mu_{\min} + \frac{\mu_{\max} - \mu_{\min}}{1 + (\frac{N_{total}}{N_{ref}})^{\alpha_n}} \tag{2.5}$$

In Equation (2.5), $\mu_{\min}$ and $\mu_{\max}$ are maximum and minimum mobilities, N_{ref} is a reference concentration and N_{total} is total doping concentration at a particular position in the simulated device and α_n is a fitting parameter. Further improvement in mobility model to take into account the vertical field dependence of the mobility and the value of parameters in Equation (2.5) can be found in [55].

Modeling quantum effects becomes important with devices scaled down. The quantum correction allows quantum confinement effects and some aspects of tunneling phenomena to be taken into account in DD simulations. Two well known methods for quantum corrections in classical DD simulations are the Density Gradient (DG) approach and Effective Potential (EP) approach [66]. The Glasgow University 'atomistic' simulator incorporates DG approach. The quantum corrections in DG approach are introduced by addition of an extra term in the DD current expression [66]:

$$J_n = qD_n \nabla n - q\mu_n n \nabla \psi - 2qn\mu_n \nabla (b_n \frac{\nabla^2 \sqrt{n}}{\sqrt{n}}) \tag{2.6}$$

where the first two terms in Equation (2.6) describe classical diffusion and drift components of the current respectively, as already introduced in Equations (2.1) and (2.2). The third term is called 'quantum driving force' because inclusion of it results in pushing carriers away from the Si/SiO_2 interface yielding carrier distribution which is consistent with the solution of Poisson-Schrödinger equation [63]. The shift in the peak of carrier concentration away from the interface is called quantum confinement which in turn increases the effective gate oxide thickness and the MOSFET threshold voltage [67]. The component b_n determines the magnitude of density gradient corrections and is given by:

$$b_n = \frac{\hbar^2}{4m_n^* qr} \tag{2.7}$$

where $\hbar$ is the reduced Planck constant, m_n^* is the effective mass of electrons and *r* is a dimensionless parameter which depends on the temperature and physical properties of the energy band diagram. It has been shown that *r* approaches to 3 for high temperatures (above 77K in Silicon) [68]. By incorporating the driving force term of Equation (2.6) into drift term, an effective potential can be introduced in the DD current density as:

$$J_n = qD_n\nabla n - q\mu_n n\nabla\psi_{eff} \quad ; \quad \psi_{eff} = \psi + 2b_n\frac{\nabla^2\sqrt{n}}{\sqrt{n}} \tag{2.8}$$

Using the quasi-Fermi potential as a link between effective potential and electron density, results in [66]:

$$\phi_n = \psi_{eff} - \frac{kT}{q}\ln(\frac{n}{n_i}) = \psi + 2b_n\frac{\nabla^2\sqrt{n}}{\sqrt{n}} - \frac{kT}{q}\ln(\frac{n}{n_i}) \tag{2.9}$$

where ϕ_n is the quasi-Fermi potential and n_i is the intrinsic carrier density and other symbols have already been defined.

2.2.2 The Atomistic Simulator

The Glasgow University 'atomistic' simulator has been developed within the Glasgow Device Modeling Group and it is the basic simulation tool for the study of intrinsic parameter fluctuations in sub-100nm MOSFETs. The main equations in the heart of simulator are Poisson's equation, Density gradient equation and Current continuity equation which are numerically solved together in an iterative approach to obtain convergence. The Poisson's equation which relates the charge and the potential is given by:

$$\varepsilon\nabla^2\psi = q(n - p + N_A^- - N_D^+) \tag{2.10}$$

where ψ is the potential, ε is the silicon permittivity, *q* is the unit charge, *n* is the electron concentration, *p* is hole concentration, N_A^- is the concentration of ionized acceptors and

N_D^+ is the concentration of ionized donors. The density gradient equation was discussed in Equation (2.9) and the current continuity equation is given by:

$$\nabla . J_n = 0 \tag{2.11}$$

where J_n is the current density flowing through the device and is given by Equation (2.8) to account for the effect of quantum corrections. The set of equations consisting (2.9), (2.10) and (2.11) are discretised onto non-uniform 3D Cartesian mesh using a finite difference scheme with mesh size between 0.5nm and 1nm [55]. The flowchart of the solver part of the atomistic simulator is shown in Figure 2.6 [69]. First, Density Gradient equations are solved self-consistently with Poisson's equation. The results are then used in obtaining of current in self-consistent iterative method using current continuity equation.

2.2.3 Implementing Sources of Parameter Fluctuations

It is essential to review the techniques and models used to introduce some of the most important sources of statistical variability in the atomistic simulator. RDD is introduced into atomistic simulator using the methodology described in [55]. In this method, all sites of silicon lattice covering the simulated transistor are scanned one by one and dopants are introduced randomly in the sites with a probability equal to corresponding dopant to silicon

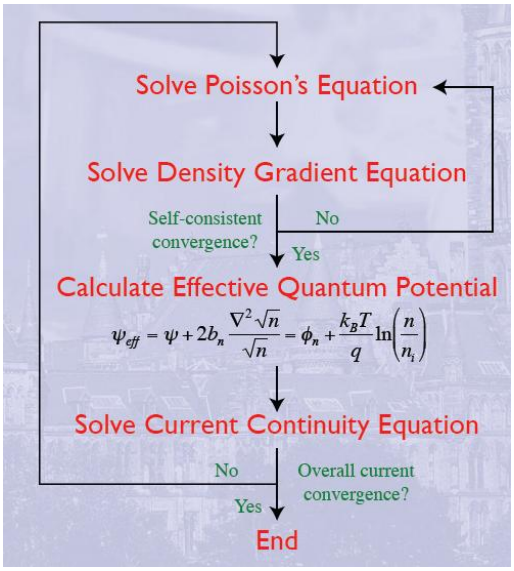


Figure 2.6: Flowchart of implemented solver in Glasgow University atomistic simulator [69].

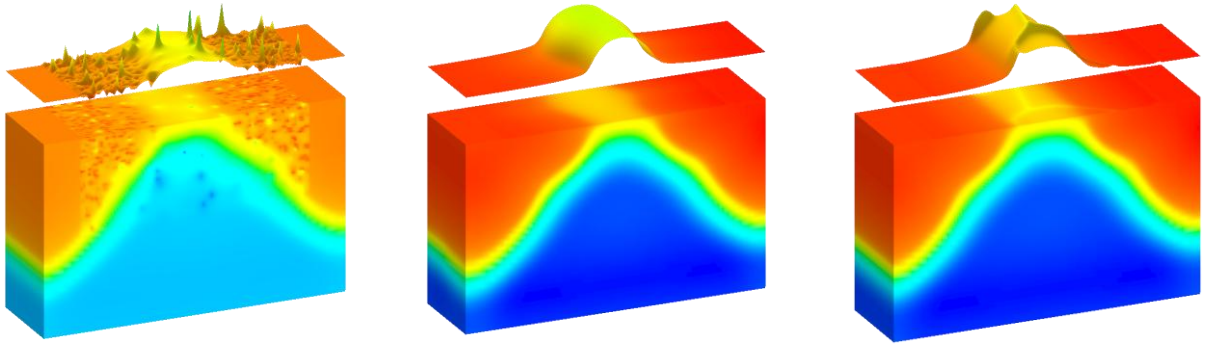
concentration ratio. One typical problem in introducing random dopants into simulator is the charge trapping [70]. Discrete dopants create deep Coulomb potential wells in classical simulations and these deep wells are able to artificially trap carriers. This problem can be reduced by using Cloud-in-Cell charge assignment technique which is addressed in [55,63]. Also, it has been shown that the inclusion of Density Gradient quantum correction in atomistic simulations leads to significant decrease in amount of trapped charge around the impurity atoms [55,63].

LER is introduced into the simulator using 1-D Fourier analysis which generates the gate edge lines from a Gaussian power spectrum. The Gaussian power spectrum used in the simulator is given by [71]:

$$S_G(k) = \sqrt{\pi}\Delta^2 \exp(-\frac{k^2\Lambda^2}{4}) \quad ; \quad k = i(\frac{2\pi}{Ndx}) \quad ; \quad 0 \le i \le \frac{N}{2} \tag{2.12}$$

where N is the number of mesh points in the Fourier space and dx is the mesh spacing. The corresponding autocorrelation function of the power spectrum is characterized using two parameters, the RMS amplitude Δ and the correlation length Λ.These characteristic values are extracted from fitting the data obtained from SEM micrographs and electron beam lithography of real devices. The results indicate that Λ is in the range of 20-30nm and 3Δ is in the range of 3-5nm. More details of generating LER profile based on Equation (2.12) can be found in [63].

The introduction of PGG into the atomistic simulator has been carried out using AFM image of polysilicon grains, as shown in Figure 2.5. This type of variability is important for n-channel MOSFETs. This is due to the presence of acceptor type interface states in the upper half of bandgap in n-channel MOSFET which pins the Fermi level and the absence of donor type interface states in lower half of bandgap in p-channel MOSFETs which leaves the Fermi level unpinned [71]. The grain boundaries are then traced in black color using graphical software, leaving the grains white. The image is then scaled and saved in a template format readable by the simulator. The simulator imports a random section of scaled template and pins the Fermi level along the defined black grain boundaries.



(a) (b) (c)

Figure 2.7: Potential distribution resulted from 3D simulation of a typical 35 nm gate length n-channel MOSFET subject to (a)-RDD, (b)-LER, (c)-PGG [24].

The individual impact of RDD, LER and PGG on the potential distribution in a typical 35nm gate length bulk MOSFET resulted from 3D simulation of IPF with Glasgow University atomistic simulator is illustrated in Figure 2.7 [24].

The simulation results for the standard deviation of threshold voltage (σV_T) from individual and combined sources of statistical variability has been compared with measured data from 45nm technology node in [60]. The simulation results are in very good agreement with measured data, i.e. less than 2% difference in σV_T of both N- and P-channel MOSFETs biased at low or high drain bias has been reported which verifies high accuracy of simulation techniques.

2.3 Statistical Modeling

Statistical models are an essential part of manufacturability-aware design. They provide the insight for design margins and pave the road for the modeling of statistical properties of contemporary and next generation CMOS technology in presence of statistical variability. First statistical models were developed in 1980 to model matching properties of passive elements such MOS capacitors on integrated circuits [1,72]. A few years later, characterization and modeling of mismatch in MOSFET transistors were reported with focus on analog design [73]. In this section, we first review existing mismatch models which are able to predict the variations in electrical parameters of the device. In the second subsection, the traditional corner models as well as advanced statistical methods such as PCA and BPV will be addressed based on literature review.

2.3.1 Mismatch Models

Mismatch is defined by time-independent differential performance of identical transistors in a single integrated circuit [74]. Threshold voltage mismatch caused by random dopant variations is one of the key performance factors in circuit design because it is a significant contributor in determining minimum operating voltage of SRAM cells and register file arrays in a chip [75]. An analytical expression for σ_{V_T} in bulk MOSFETs caused by RDD was proposed by Mizuno *et al.* [76], and is given by:

$$\sigma_{V_T} = \frac{(4q^3\varepsilon_{si}\phi_B N)^{1/4}}{2(W_{eff}L_{eff})^{1/2}}\frac{t_{ox}}{\varepsilon_{ox}} \tag{2.13}$$

where N is bulk doping concentration, ϕ_B is the bulk Fermi potential and other symbols have their conventional meanings. A similar expression with slightly different coefficients was proposed by Stolk *et al.* in [77], and is given by:

$$\sigma_{V_T} = \frac{(4q^3\varepsilon_{si}\phi_B N)^{1/4}}{(3W_{eff}L_{eff})^{1/2}}\frac{t_{ox}}{\varepsilon_{ox}} \tag{2.14}$$

The reduction in the variability for larger area devices is justified by the physical fact that the number of random dopants is averaged for larger devices. It is usual to measure the random variation of the threshold voltage for closely spaced pairs of MOSFETs on a chip. The measured $\sigma_{\Delta V_T}$ is related to σ_{V_T} by:

$$\sigma_{\Delta V_T} = \sqrt{2}\sigma_{V_T} \tag{2.15}$$

Either $\sigma_{\Delta V_T}$ or σ_{V_T} are usually plotted versus $1/\sqrt{area}$ to compare the threshold voltage variability between different geometries and technology structures. Based on Equation (2.13) or (2.14), these types of plots are almost linear and the line slope is termed A_{VT}. The deviations from the linear behavior are results of the impact of non-RDD

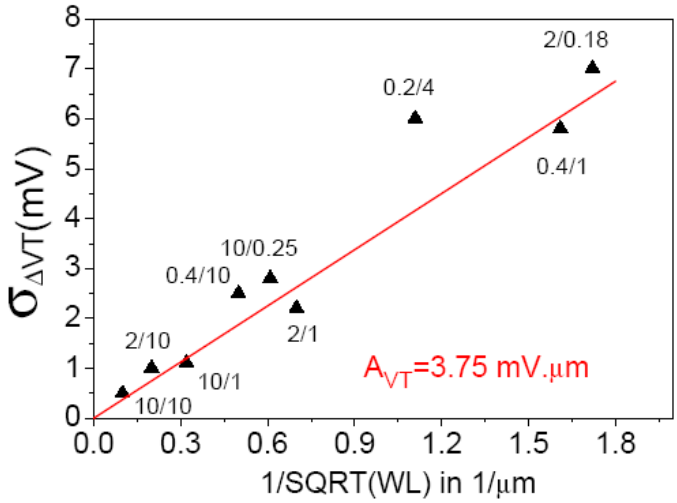


Figure 2.8: Variance of the difference in threshold voltage of transistor pairs in a 180nm CMOS process, the aspect ratio (W/L) for each device is written on symbols in μm. After [32].

components on the σ_{V_T} [75]. Figure 2.8 illustrates plot of $\sigma_{\Delta V_T}$ versus the inverse square root of the device area for a set of different geometry n-channel MOSFETS in a typical 180nm CMOS process [32]. For the smaller size transistors, the effective gate area is smaller due to relative larger overlap between gate and source/drain area which in turn reduces the effective channel length by ΔL . ($L_{eff} = L - \Delta L$)

The current mismatch is the other important performance factor of MOSFETs which has been extensively researched in the literature [74,78,79]. The propagation of variance method can be used to derive analytical expressions for current mismatch [80]:

$$\sigma_{\Delta I}^2 = (\frac{\partial I}{\partial V_T})^2 \sigma_{\Delta V_T}^2 + (\frac{\partial I}{\partial \beta})^2 \sigma_{\Delta \beta}^2 \tag{2.16}$$

Using simple square law of $I = 0.5\beta(V_{gs} - V_T)^2$ in combination with propagation of variance, Equation (2.16), results in:

$$(\frac{\sigma_{\Delta I}}{I})^2 = \frac{4\sigma_{\Delta V_T}^2}{(V_{gs} - V_T)^2} + \frac{\sigma_{\Delta \beta}^2}{\beta^2} \tag{2.17}$$

The simple square law which has been used to derive Eq. (2.17) limits the application of current mismatch formula to saturation regime and long channel devices. Nevertheless, it gives an insight for the important contributors of the MOSFET current mismatch. The mismatch in the current factor can be further simplified using the current factor definition of $\beta = \mu C_{ox} W / L$ and assuming independent components by:

$$\frac{\sigma_{\Delta\beta}^2}{\beta^2} = \frac{\sigma_W^2}{W^2} + \frac{\sigma_L^2}{L^2} + \frac{\sigma^2 C_{ox}}{C_{ox}^2} + \frac{\sigma_\mu^2}{\mu^2} \tag{2.18}$$

The variation in mobility originates from different sources of statistical variability. The variation in oxide capacitance originates from OTF and the variations in W and L arise from LER. One of disadvantages of Eq. (2.17) occurs when the underlying cause of current mismatch is OTF. In this case, the oxide thickness variations will be accounted for both $\sigma_{\Delta V_T}$ and $\sigma_{\Delta\beta}$. Therefore, $\sigma_{\Delta I}$ will be overestimated and the correlations between two component need to be taken into account to avoid this problem [74,80].

A more accurate mismatch modeling approach which employs BSIM compact model and SPICE simulator to derive partial derivatives in Eq. (2.16) and is applicable in weak and strong inversion, linear and saturation regions across different bias conditions proposed in [74,80]. It has been shown that the impact of local dopant fluctuations on current mismatch of MOSFETs operating in saturation region is significant [81,82]. Local dopant fluctuations cause a meandering boundary of depletion region across the channel and as a result the mobile charge in the inversion layer will be modulated which in turn leads to current fluctuations. A compact model of current mismatch which takes into account the local dopant fluctuation is presented in [81,82].

The impact of halo and well design on the MOSFET mismatch in 32nm HKMG technology shows potential need for the extension of conventional approaches to take these effects into account [83]. Halo or pocket implants can introduce significant variations in local doping profile of random dopants which in turn causes mismatch.

2.3.2 Corner Models

Corner or worst-case models have been used traditionally to characterize the impact of process variability in circuit design [23,84]. They are generated based on a Gaussian distribution of process-sensitive compact model parameters. The mean value of the parameters are selected from a nominal transistor and then the means are shifted by $n\sigma$ (where n is integer) and σ is the assumed value of the standard deviation for a target parameter. Therefore $\mu \pm n\sigma$ sets the lower and upper limit of a parameter which will be used to produce worst-case corners of transistor or circuit electrical performance.

The important corners for analog design are SS (slow-n, slow-p) which models worst-case speed and FF (fast-n, fast-p) which takes the worst-case power consumption into account. Significant corners for digital design are FS (fast-n, slow-p) which creates worst-case 1 and SF (slow-n, fast-p) which corresponds to worst-case zero [23]. Figure 2.9 illustrates simulation examples of corner models for device and circuit applications. Figure 2.9(a) shows the cloud of saturation current for NMOS versus saturation current of the PMOS in a typical process with simulated process corners [23]. Figure 2.9(b) shows the dissipated energy versus maximum delay of a 1-bit adder circuit with their corners at different levels of the variability [85]. As is evident from both figures, that corners models are either pessimistic or cannot cover full cloud of real data and they may cause over-design. Moreover, considering statistical variability in addition to process variability leads to global or statistical corners [85].

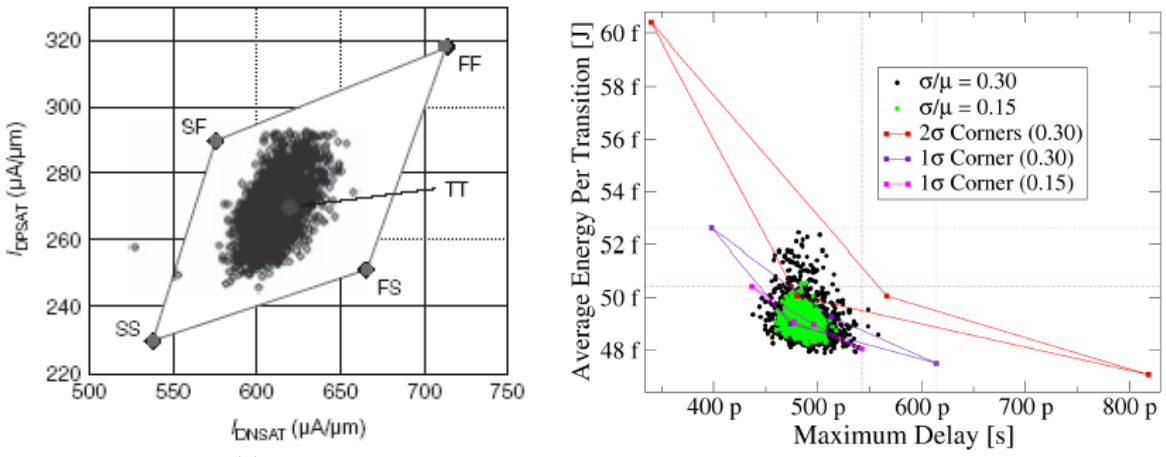


Figure 2.9: (a)- Drive current of PMOS versus NMOS with definition of 4 corners [23], (b)- Average energy versus maximum delay of a 1-bit adder at different levels of the variability [85].

2.3.3 Numerical Models

Several numerical models have been used for the purpose of MOSFET statistical modeling. One famous method is Principal Component Analysis (PCA) [2,86,87,88]. We briefly review this method. Other numerical techniques such as Statistical Timing Analysis (STA) and Response Surface Methodology (RSM) have been introduced with the aim of statistical circuit simulations [87,89].

Principal Component Analysis (PCA) is a valuable technique of transforming statistically correlated variables into uncorrelated variables called "Principal Components". Assuming that X is a matrix of measured or simulated data with m rows and n columns, m is the number of parameters associated with one particular device and n is the number of sample devices. The covariance matrix of X will be a symmetrical m by m matrix with diagonal terms equal to variance of each parameter and off-diagonal terms equal to covariance between measurements. The covariance is directly related to correlation of parameters. The goal of PCA is to find a transformation matrix P for the original matrix X such that the resulting matrix $Y=PX$ should have a diagonal covariance matrix whose columns are called principal components of X. Examining the covariance matrix of transformed matrix gives:

$$C_Y = \frac{1}{n-1} YY^T = \frac{1}{n-1}(PX)(PX)^T = \frac{1}{n-1} P(XX^T)P^T \tag{2.19}$$

If we define a new matrix $A = XX^T$ in the middle of Eq. (2.19), this matrix is symmetrical. Any symmetrical matrix can be diagonalized using an orthogonal matrix of its eigenvectors [90]. Hence:

$$A = EDE^T \tag{2.20}$$

where E is the matrix of eigenvectors of A arranged as columns and D is a diagonal matrix. By choosing $P=E^T$, the covariance matrix C_Y will be diagonal. This can be verified by substituting (2.20) into (2.19) which gives $C_Y=D/(n-1)$.Choosing this P, the principal components are calculated from:

$$Y=E^{T}X \tag{2.21}$$

In practice, calculating principal components of original data is performed in two stages: first, the data are normalized by subtracting the mean values of each row and dividing the results by original standard deviations. This will convert all of data to normal distributions with zero mean and standard deviation of one. In the next step, the eigenvectors of the covariance matrix will be calculated. Based on Eq. (2.21), reconstructing original data from its principal component, Y, can be performed by:

$$X=EY \tag{2.22}$$

It should be pointed out here that two assumptions in PCA are normal distribution and linearity of the original data [90]. Any deviation from these assumptions leads to degradation of the accuracy of PCA in real applications.

2.3.4 Analytical Models

The Backward Propagation of Variance (BPV) has been introduced recently and it lies in the category of analytic models [91,92]. It is based on formulating statistical models for device electrical performance parameters as a function of independent normally distributed process parameters encapsulated in SPICE models for device and circuit simulations. It is essential for this method to find enough process parameters in SPICE models to capture observed fluctuations in electrical behavior of device and circuit. Developers of the BPV have suggested using process parameters such as lateral and vertical geometry variations, corresponding to parameters ΔL and t_{ox} in different SPICE models, and material properties such as sheet resistance, doping density and flat-band voltage in SPICE models as the basis of modeling. BPV tries to model correlation between electrical performance parameters and the nonlinearities associated with their distributions up to third moment, skewness. Considering N_p independent normally distributed process variables as a vector p and N_e electrical performance parameter of interest as a vector e, the problem is to characterize the $e(p)$ mapping. Taylor expansion approximation for a particular component of electrical performance of interest, e_m, gives:

$$e_m(p) \cong e_m(\bar{p}) + \sum_{i=1}^{N_p} s_{m,i}\delta p_i + \sum_{i,j=1}^{N_p} s_{m,ij}\delta p_i \delta p_j \tag{2.23}$$

where $\bar{p} = (\bar{p}_1, \bar{p}_2, ..., \bar{p}_{N_p})$ is vector of the mean or median value of independent, normally distributed process parameters and $\delta p_i = p_i - \bar{p}_i$ is the deviation of a particular process parameter from its mean value. The first and second order sensitivities introduced in Eq. (2.23) are given by:

$$s_{m,i} = \left.\frac{\partial e_m}{\partial p_i}\right|_{p=\bar{p}} \quad , \quad s_{m,ij} = \left.\frac{1}{2}\frac{\partial^2 e_m}{\partial p_i \partial p_j}\right|_{p=\bar{p}} \tag{2.24}$$

Note that the normal distribution of process parameters, *p*, does not imply that the electrical performance parameter e_m is normally distributed and its distribution can be skewed. The variance and skewness of e_m are defined by:

$$\sigma_{e_m} = < (e_m - < e_m >)^2 > \quad , \quad \gamma_{e_m} = \frac{< (e_m - < e_m >)^3 >}{\sigma_{e_m}^3} \tag{2.25}$$

And the covariance between two electrical performance parameters will be given by:

$$\sigma_{e_m,e_n} = < (e_m - < e_m >)(e_n - < e_n >) > \tag{2.26}$$

If the independent process parameters have the mean value of $\bar{p}_i$ and variance σ_i^2, then the mean, variance and skewness of electrical performance parameter of interest, e_m, can be found by using Eq. (2.23) and the definitions of Eq. (2.25). More simplifications results in [91,92]:

$$\mu_{e_m} = e_m(\bar{p}) + \sum_{i=1}^{N_p} s_{m,ii}\sigma_i^2 \tag{2.27}$$

$$\sigma_{e_m}^2 = \sum_{i=1}^{N_p} s_{m,i}^2 \sigma_i^2 + 2 \sum_{i,j=1}^{N_p} s_{m,ij}^2 \sigma_i^2 \sigma_j^2 \tag{2.28}$$

$$\gamma_{e_m} = \frac{1}{\sigma_{e_m}^3} \left(\begin{array}{c} 6 \sum_{i,j=1}^{N_p} s_{m,i} s_{m,j} s_{m,ij} \sigma_i^2 \sigma_j^2 \\ + 8 \sum_{i,j,k=1}^{N_p} s_{m,ij} s_{m,jk} s_{m,ki} \sigma_i^2 \sigma_j^2 \sigma_k^2 \end{array} \right) \tag{2.29}$$

and the covariance between two electrical performance parameters of interest, e_m and e_n, will be given by:

$$\sigma_{e_m,e_n} = \sum_{i=1}^{N_p} s_{m,i} s_{n,i} \sigma_i^2 + 2 \sum_{i,j=1}^{N_p} s_{m,ij} s_{n,ij} \sigma_i^2 \sigma_j^2 \tag{2.30}$$

If the variance of some process parameters such as oxide thickness can be specified directly, they are treated as Forward Propagation of Variance (FPV) and this should be considered by modifying Equations (2.27) to (2.30). The details of modifications can be found in [91,92].

2.4 Summary

The significance of statistical variability for the design of contemporary and next generation of CMOS technology was discussed in this chapter. Statistical variability arises from discreetness of charge and granularity of matter and plays as a barrier among other challenges of CMOS technology. We classified different axis of the variability and enumerated various sources of systematic and statistical variability. A review on the simulation techniques used in Glasgow University atomistic simulator as well as the implementation of different sources of statistical variability into atomistic simulator was presented. Existing statistical methods and approaches used in the literature for MOSFET modeling such as mismatch models and numerical and analytical approaches were discussed.

Chapter 3

Uniform Device PSP Parameter Extraction and Optimization

The ITRS 2010 forecasts that the scaling of bulk MOSFETs will be extended until 2016 [93]. This introduces a significant demand for better modeling of new technology generation of bulk MOSFETs with reduced gate length, especially at sub-45nm technology nodes. Compact models are concise mathematical description of the complex device physics in the transistor. A compact model maintains a balance between the amount of detailed physics embedded for model accuracy and model compactness (computational efficiency). The simplifications in the device physics enable very fast analysis of device/circuit behavior when compared to the much slower numerical based TCAD simulations. Compact models have been at the heart of EDA tools for more than several decades, and are playing an increasingly important role in the deca-nanometer system-on-chip era [94,95].

In recent years, surface potential based compact models like PSP [10] have attracted significant attention because of their better physical description of device characteristics compared to industry standard compact models like BSIM4 which is a threshold voltage based model. Threshold voltage based models like BSIM4 [11] use fitting parameters to achieve continuity between weak and strong inversion and this is not based on physical considerations, and as a result the accuracy of fitting varies from device to device. The surface potential based compact models have shown better continuity and smoothness in device behavior especially in the transition from weak inversion to strong inversion. This

smooth behavior is very important in circuits which operate with scaled supply voltages around one volt. With such low supply voltage, moderate inversion regime will play an important role in device operation. Accurate description of the device behavior in this region cannot be easily obtained with threshold voltage based compact models.

This chapter is organized into four sections. In the first section, after a concise introduction to device model classification, the principles and mathematical background of PSP compact model alongside with the hierarchical construction of the model from local to global level are reviewed. Then in the second section, physics and design of a 35nm MOSFET device are investigated. This particular device is considered as a basic continuous doping device for subsequent statistical compact modeling studies in this thesis. In the final section, the context of PSP parameter extraction and optimization for the basic uniform device will be elaborated in two DC and AC parts. The accuracy of corresponding HSPICE model card will be evaluated in both device and circuit levels compared with TCAD simulation results.

3.1 PSP Compact Model

3.1.1 Device Model Classification

The aim of widely used advanced circuit simulators like HSPICE is to verify the circuit operation before being manufactured in real silicon [94,95]. This verification is a requirement in the industry to reduce both the production cost and time to market for new IC designs. By simulating the complicated circuits, designers are able to detect any deviation from desired operation of the circuit and modify it in the design stage to avoid the costly and time consuming prototyping.

It is essential to use accurate device models in the circuit simulators to achieve a reliable prediction in the design stage. With no doubt the accuracy of device models affects the reliability of the circuit simulation results. Device models provide a bridge between circuit designers and chip manufacturers.

The device models can be classified into three categories. They are numerical models, look-up tables and compact models. The first category which is the most accurate between

others is used in this thesis to create a set of accurate device current-voltage characteristics under the influence of statistical variability sources. Numerical models solve self consistently the Poisson's equation and the current continuity equation using different degrees of approximation for the carrier transport equations. They consider quantum corrections and other physical phenomena and finally they are able to represent the device characteristics under different bias conditions, temperatures and device geometries.

Although the numerical models are the best to produce accurate device characteristics under different operating conditions, they are difficult to use in circuit simulators with large number of transistors because they need a lot of CPU time which slows down the circuit simulation process and restricts the number of transistors that can be simulated. The circuit analysis using these models with normal workstation is extremely slow, therefore two benefits of using circuit simulators in the design process, the cost reduction and the fast time to market, are lost. In the second approach a table of transistor characteristics is constructed either from measurement data or from comprehensive numerical simulations. Then in a circuit simulator employing table look-up models, the value of device characteristics can be read or interpolated from the table very fast. The only limitation to this type of device model is that it needs an extensive data measurement or simulations to cover different bias and temperature conditions. The most feasible device models in circuit simulators are the so called 'compact models'. Compact models are mathematical description of device physics.

The main requirement for a state of the art compact model is that it should provide a reasonable trade-off between model accuracy and computational efficiency [96]. The interface between a compact model and a circuit simulator is through a set of parameters which is called a 'model card'. A complete model card consists of a few hundred parameters [10,11]. Each parameter tries to capture a physical aspect of the device physics. However, due to the requirement of high computational efficiency in compact models, there is a trade-off between the accuracy and physical meaning of each parameter in compact model related to the feasibility of its implementation in circuit simulation. This needs a balance between the accuracy and compactness with both of them remaining in their tolerated margins. An important result of this discussion is that the accuracy of a model card should be evaluated for each particular device through a process which is called 'Parameter Extraction and Optimization', the main content of this chapter.

With scaling the device dimensions below 100nm which has delivered billions of devices on a chip and with the emergence of new physical phenomena for these deca-nano MOSFETs, the conflict between the model accuracy and computational efficiency becomes more severe. Hence, more concerns should be considered and addressed by the compact model developers to introduce new parameters in the model while keeping the model computational efficiency at a reasonable level to simulate large circuits in less than a few seconds with desktop computers. Quantum mechanical corrections, polysilicon gate depletion effects, gate tunneling current, gate induced drain/source leakage (GIDL and GISL) currents, drain induced barrier lowering (DIBL), extensive carrier mobility model, HALO doping, retrograde channel profile, STI-induced stress model and enhanced channel modulation effects are some examples of these important phenomena which should be incorporated in the compact model. A general purpose state-of-the-art compact model should also satisfy another main requirement which is generality to remain independent of particularities associated with fabrication process and device design [95].

3.1.2 PSP Structure

PSP is a new surface potential based compact MOSFET model which has been developed by merging features of other two surface potential based models, named SP and MM11 [10]. SP has been developed in Pensylvania state University and MM11 has been developed at Phillips [97,98].

Three approaches have been used so far in the design of compact models for MOSFET transistors [95,99]. They are the threshold voltage-based approach, the inversion charge-based approach and the surface potential-based approach. The most important example of the first category is BSIM (Berekley Short Channel IGFET Model) series compact models which have served industry for over 20 years [11]. The main products of this series are BSIM3 and BSIM4. There are other formats of V_{th}-based compact models like MM9 [99]. Although the V_{th}-based compact models are formulated in a way which is easier to use for circuit designers, there is a widely shared consensus that the models of this type cannot represent the full features of sub 100nm devices [97,100,101].

The compact models utilizing the inversion charge-based approach are powerful alternatives of V_{th}-based models [95,97]. There are some limitations for this approach to

model device characteristics especially in accumulation region where inversion charge is not available which is important for some device applications [95,97]. The compact model EKV is a famous member of q_i-based compact model family [102].

It has been shown that surface potential-based approach (ψ_S -based) includes both V_{th}-based and q_i-based methods as special cases when additional assumptions are to be considered [95]. The origin of surface potential compact modeling is from the Pao-Sah formula [103] which was published at 1966. Although it was able to describe all regions of device operation with a single expression, it was practically useless in compact modeling area until the last decade because of the complexities associated with numerical solution of the surface potential equation [95].

PSP has a hierarchical approach from local to global parameters. Global parameters include geometry dependencies and before evaluating a particular device current – voltage characteristics, they are converted to local parameters that correspond to the geometry and the dimension of a particular device. Although the use of local parameters can facilitate the parameter extraction process, the use of global parameters has the advantage of batch extraction and description for a set of different channel length devices. The simplified PSP structure is shown in figure 3.1.

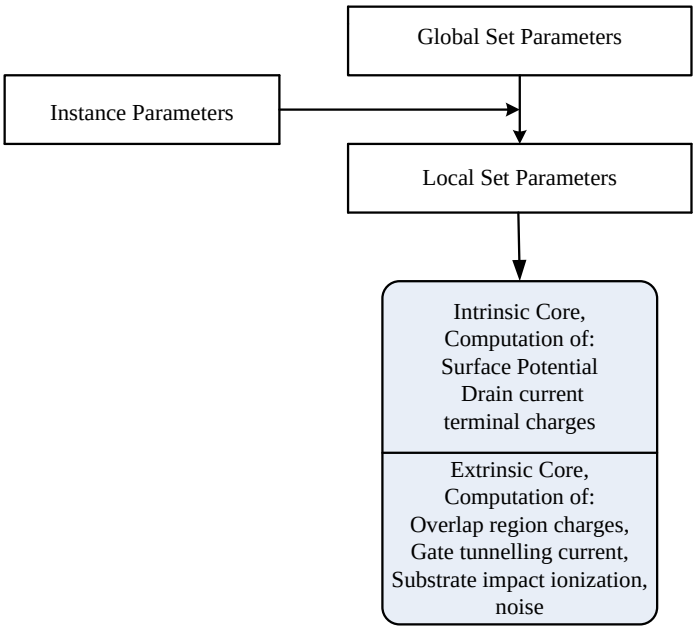


Figure 3.1: Simplified structure of PSP compact model.

In Figure 3.1 instance parameters refer to the geometry and the size of the device including channel length (*L*), channel width (*W*), Source/Drain areas and perimeters. The instance parameters do not enter a 'model card' but are supplied by the circuit designer into circuit simulator 'netlist' to create necessary local parameter set. For example the description for flat band voltage (*VFB*) is given by:

$$VFB = VFBO.(1 + VFBL.\frac{L_{EN}}{L}).(1 + VFBLW.\frac{W_{EN}}{W}).(1 + VFBLW.\frac{W_{EN}}{W}.\frac{L_{EN}}{L}) \tag{3.1}$$

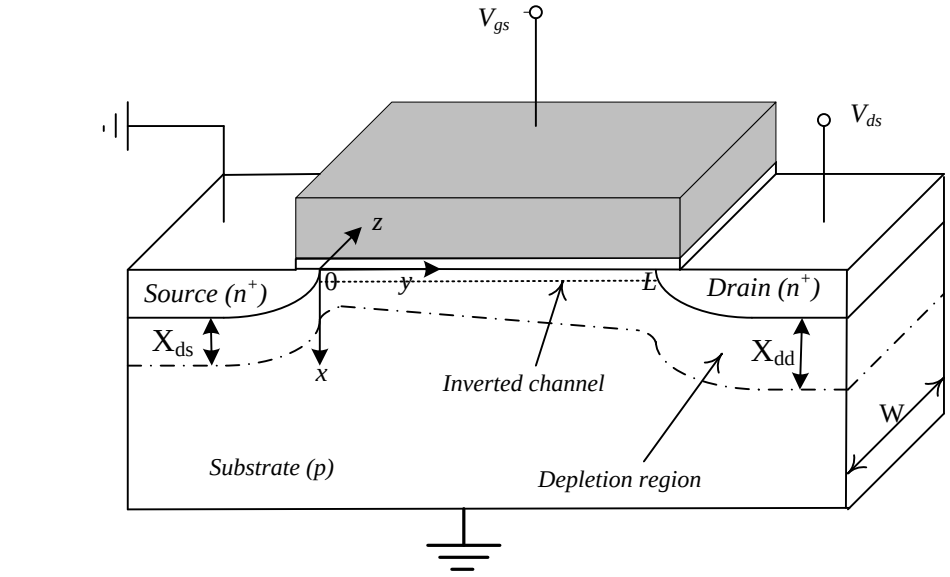
In the Equation 3.1, *VFBO* is the general global parameter (geometry independent part) of MOSFET flat band voltage, *VFBL* is the length dependent counterpart of it, *VFBW* is the width dependent counterpart of it, and *VFBLW* is the area dependent parameter of flat band voltage in global mode. L_{EN} and W_{EN} are standard reference constants equal to one micro meter. *L* and *W* are device length and width respectively.

3.1.3 Surface Potential Equation

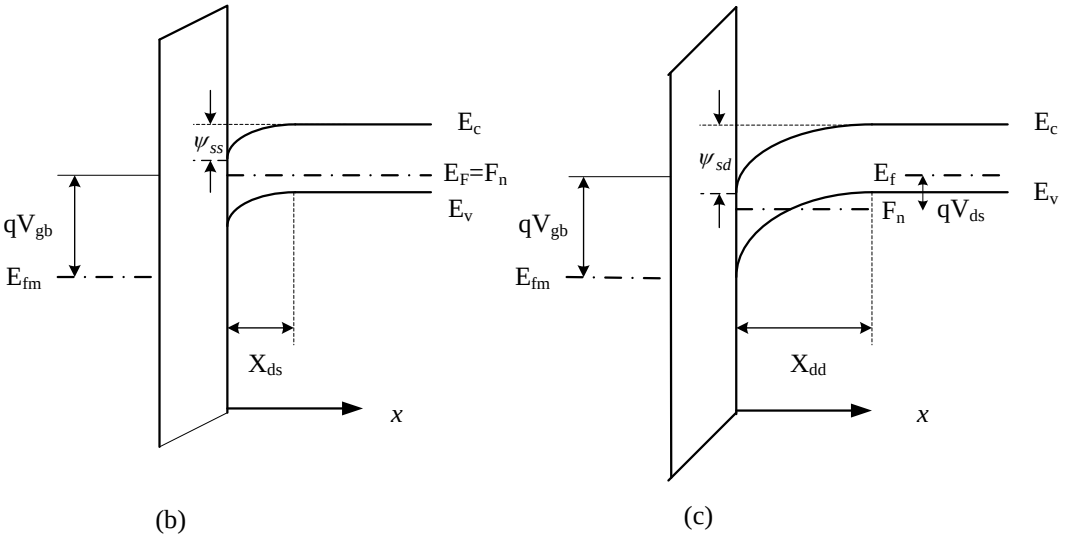
The surface potential, ψ_s, which is the potential at the Si/SiO_2 interface is the basis of PSP and other surface potential based compact models [95,97]. It is derived by solving Poisson's equation under some simplified assumptions and then it is used in the core of compact model for further calculations such as current and charges of terminals. The physical schematic of a conventional bulk MOSFET is shown in Figure 3.2(a) with associated coordinate system. Figures 3.2(b) and 3.2(c) depict its energy band diagram at source and drain ends respectively. In Figure 3.2(a), the *x* direction is the distance into depth of silicon measured from Si/SiO_2 interface. The *y* direction shows the distance along the length of the channel measured from source junction and increasing toward the drain junction, and *z* direction denotes distance along the channel width. In general, the Poisson's equation for this 3-D structure can be written as:

$$\nabla^2\psi = -\frac{\rho}{\varepsilon_s} \tag{3.2}$$

where ψ is the electrostatic potential, ρ denotes the charge density and ε_s is the silicon permittivity.



(a)



(b) (c)

Figure 3.2: (a) Schematic diagram of N-channel MOSFET. Its energy band diagram at (b) the source end, and (c) the drain end.

As the MOSFET structure is symmetric along z direction, it can be treated as a 2-D system. At this point an important simplifying assumption which is called gradual channel approximation (GCA), turns the Poisson Equation (3.2) to an one-dimensional problem. GCA states that the variation of the electric field along the channel is much less than the corresponding variation in the x-direction. Although it is valid for most parts of the channel region, it fails in the pinch-off region of the MOSFET channel [104]. In PSP, this problem is solved by introduction of an "effective drain-source voltage" which will be discussed later. In a mathematical description, GCA can be stated as:

$$\frac{\partial E_y}{\partial y} << \frac{\partial E_x}{\partial x} \Rightarrow \left|\frac{\partial^2 \psi_y}{\partial y^2}\right| << \left|\frac{\partial^2 \psi_x}{\partial x^2}\right| \tag{3.3}$$

With GCA assumption, equation (3.2) becomes one dimensional and the simplified version of it becomes:

$$\frac{\partial^2 \psi}{\partial x^2} = -\frac{\rho}{\varepsilon_s} \tag{3.4}$$

Denoting substrate doping with N_A and hole and electron densities as p and n respectively and assuming complete ionization of acceptors, charge density is given by:

$$\rho = q(p - n - N_A) \tag{3.5}$$

Since the electron current is the major contribution of total current in an N-channel MOSFET, the hole current is negligible. Hence, the hole semi-Fermi potential (occasionally called hole imref, F_p) and the hole density in the channel can be related to electrostatic potential by:

$$p = p_b.\exp(-\psi / V_t) \tag{3.6}$$

where $p_b = N_A$ is the hole concentration in the bulk and ψ denotes the electrostatic potential and $V_t = kT/q$ is the thermal potential, roughly 26mV at room temperature. For electrons, the semi-Fermi potential is dependent on the channel voltage at each point along the channel. It makes more band bending near drain as shown in Figure 3.2(b) and 3.2(c) compared to the source. The density of electrons can be given by:

$$n = n_b.\exp[(\psi - \phi_n)/V_t] \tag{3.7}$$

where ϕ_n is the channel voltage which results in more split in semi-Fermi potential when moving from the source towards the drain. ϕ_n is zero at the source and it is equal to V_{ds} at the drain. It can be expressed in terms of semi-Fermi potential split between electrons and holes as:

$$\phi_n = \frac{F_p - F_n}{q} \tag{3.8}$$

Substituting equations (3.6) and (3.7) into (3.5) and considering the charge neutrality in the bulk ($N_a = p_b - n_b$) results in:

$$\rho = q[p_b(e^{-\beta\psi} - 1) - n_b(ke^{\beta\psi} - 1)] \tag{3.9}$$

where $\beta = 1/V_t$ and $k = \exp(-\beta\phi_n)$. On the other hand, equation (3.4) can be easily converted into another form to show surface electric field by integrating it once and using the boundary condition $\partial\psi/\partial x = 0$ for $\psi = 0$. The surface electric field is defined by $E_s = -(d\psi/dx)_{x=0}$ .Hence:

$$E_s{}^2 = -\frac{2}{\varepsilon_s}\int_0^{\psi_s} \rho d\psi \tag{3.10}$$

Considering Gauss's law in a box from Si/SiO_2 interface into bulk, the total charge in

the semiconductor can be found from $Q_s = -\varepsilon_s E_s$. Therefore the MOSFET general input equation can be shown either in terms of surface electric field or Si charge as:

$$V_g = V_{fb} + \psi_s - \frac{Q_s}{C_{ox}} = V_{fb} + \psi_s + \frac{\varepsilon_s E_s}{C_{ox}} \tag{3.11}$$

where V_{fb} is the flat band voltage and $C_{ox} = \varepsilon_{ox} / t_{ox}$ is the unit area oxide capacitance [104]. Substituting Equation (3.10) into Equation (3.11), the Surface Potential Equation (SPE) can be derived as:

$$(V_g - V_{fb} - \psi_s)^2 = \gamma^2 V_t h \tag{3.12}$$

where γ is the body factor which is defined by $\gamma = \sqrt{2q\varepsilon_s p_b} / C_{ox}$ and dimensionless parameter *h* is the normalized square of surface electric field:

$$h = \frac{\varepsilon_s {E_s}^2}{2qV_t p_b} = -\frac{1}{qV_t p_b}\int_0^{\psi_s} \rho d\psi \tag{3.13}$$

The integral in Equation (3.13) with the charge density given by (3.9) can be obtained in a closed form for MOS capacitor where F_n=F_p. For MOSFET compact modeling, using reasonable approximations is inevitable. Ignoring position dependence of electron quasi-Fermi potential along the *x* direction is an approximation which is used in classical text books [104] and results in the following form of *h*:

$$h = e^{-u} + u - 1 + \frac{n_b}{p_b} k_0 (e^u - m(u)) \tag{3.14}$$

where $k_0 = \exp(-\beta\phi_{n0})$ and $m(u) = 1 + u / k_0$ while $u = \beta\psi_s$ is the normalized surface potential. It was later realized that this form of *h* is problematic for a narrow region of gate voltage around the flat-band voltage where *h* becomes negative which is inconsistent with equation (3.12) [105]. Other forms of $m(u)$ have been proposed to rectify this problem are

addressed in [95,106]. Apart from the selection of $m(u)$, the SPE equation (3.12) should be solved either using an iterative procedure [107,108] or analytical approximations for the surface potential [109,110]. PSP uses analytical approximations which are preferable in compact modeling era due to higher computational speed. To achieve these goals, PSP uses a form of $m(u)$ which is given by:

$$m(u) = 1 + u + \frac{u^2}{u^2 + 1} \tag{3.15}$$

This selection of $m(u)$ ensures existence of an analytical solution for SPE and is particularly accurate with an absolute error under 1nV relative to exact numerical solution of SPE [10]. Furthermore, it ensures the right hand side of SPE to remain always positive avoiding occasional circuit simulator crash. It is also shown that in contrast to other possible selections of $m(u)$ reported in [95], the derivatives of surface potential are continuous near flat-band voltage and the surface potential is valid for very negative values of the gate voltage with feasibility of assigning the surface potential difference between drain and source to be zero in accumulation region when needed.

3.1.4 Drain Current

All surface potential based compact models use charge sheet approximation [111] to simplify Pao-Sah double integral formula [103] for the drain current. The resulting expressions for the current and terminal charges are still complicated [112,113] and they need to be simplified more but yet accurate to be implemented in compact models. PSP uses a method which is called symmetric linearization (SLM) developed in [113]. To introduce SLM, an average surface potential (or surface potential midpoint) is introduced:

$$\psi_m = \frac{1}{2}(\psi_{ss} + \psi_{sd}) \tag{3.16}$$

where ψ_{ss} is the surface potential in the source side and ψ_{sd} is the surface potential in the drain side. The next step is to find inversion charge density for the midpoint, q_{im}, which can be found in [95,105]. Applying the inversion charge density in the standard expression

of charge sheet model [111,114] gives the drain current in the form:

$$I_d = \mu \frac{W}{L} C_{ox}(-q_{im} + \alpha V_t)\Delta\psi \tag{3.17}$$

where W and L are the width and the length of device, respectively. $\Delta\psi$ is the surface potential difference between drain and source and $V_t = kT/q$ is the thermal potential. The coefficient α is defined by this expression:

$$\alpha = 1 + \frac{\gamma(1 - e^{-u_m})}{2\sqrt{V_t(e^{-u_m} + u_m - 1)}} \tag{3.18}$$

where $u_m = \psi_m / V_t$ is the normalized midpoint surface potential and γ is the body factor already introduced after equation (3.12). The parameter μ refers to MOSFET channel effective mobility and PSP uses a mobility model as follows:

$$\mu = \frac{kUo}{1 + (MUE.E_{eff})^{THEMU} + CS[q_{bm}/(q_{bm} + q_{im})^2] + G_R} \tag{3.19}$$

where *k* is a proportionality constant and *Uo* is a model parameter roughly corresponding to vertical low-field mobility. *MUE* and *THEMUE* are two model parameters responsible for mobility degradation due to vertical field E_{eff}. Coulomb scattering is taken into account by model parameter *CS* while q_{im} and q_{bm} are midpoint inversion charge density and bulk midpoint charge density, respectively [105]. G_R accounts for the effect of source/drian series resistance (R_s) on the mobility and is equal to:

$$G_R = Uo.(W/L).q_{im}.R_s \tag{3.20}$$

PSP is capable of modeling short channel effects which are inevitable is sub-100nm MOSFETs. Velocity saturation effect is modeled in PSP as follows:

$$I_d = \frac{\mu \frac{W}{L} C_{ox}(-q_{im} + \alpha V_t)\Delta\psi}{G_{vsat}} \tag{3.21}$$

In equation (3.21), the numerator term is the same as drain current equation for long channel MOSFET which was mentioned in (3.17), but G_{vsat} in the denominator models drain current degradation due to velocity saturation effect. G_{vsat} is calculated based on a model parameter which is called *THESAT* (or θ_{sat}):

$$G_{vsat} = \frac{1}{2}[1 + \sqrt{1 + 2(\theta_{sat}\Delta\psi)^2}] \tag{3.22}$$

Equation (3.22) intends to account for velocity-electric field dependence for electrons in nMOSFET and holes in pMOSFET. Direct incorporation of velocity-field relation into compact models produces unphysical negative output conductance. The problem arises from the fact that GCA approximation (which was assumed in expression (3.3)) is not valid in the pinch-off region of the channel [105]. In PSP, this problem is solved by introducing an "effective drain-source voltage" according to:

$$V_{dse} = \frac{V_{ds}}{[1 + (V_{ds}/V_{dsat})^{a_x}]^{1/a_x}} \tag{3.23}$$

where a_x is found from $a_x = AXO/(1 + AXL/L)$ using two PSP model parameters, *AXO* and *AXL*. The minimum allowed value of a_x is 2. Finally in the drain current formulation, PSP includes semi-empirical expressions to model channel length modulation effects [105]. The multiplication factor which models the drain current increase caused by the channel modulation effect is expressed by:

$$\varsigma = T_1.[ALP + \frac{ALP1.q_{im}}{(q_{im} + \alpha.V_t)^2}] + T_2.ALP2.q_{bm}.(\frac{\alpha.V_t}{q_{im} + \alpha.V_t})^2 \tag{3.24}$$

where *ALP*, *ALP1* and *ALP2* are PSP model parameters and T_1 and T_2 are functions of terminal voltages and surface potential across the channel [115]. Other terms have been introduced previously.

3.2 The Uniform 35nm MOSFET

The test bed device in this thesis is a 35nm poly silicon gate bulk MOSFET designed to match the performance of state-of-the-art 45nm technology devices. This particular device is considered as a basic continuous doping device for subsequent statistical compact modeling studies in the next chapters. The device has gas annealed oxynitride gate oxide and retrograde and super-halo doping profile to suppress punch-through and short channel effects.

The design of the device was initially based on the 35nm gate length transistor published by Toshiba in 2002 [116] but its structure was updated to incorporate latest technology features embedded in 45nm CMOS technologies which were reported by Intel in 2007 and published in [117,118,119]. A device design includes two major steps: process simulation and device simulation. The device channel doping profile design is based on the channel doping profile from Toshiba 35nm gate length MOSFET. The gate oxide material is silicon oxynitride (SiON) with a thickness of 1.4nm and relative permittivity of 5.45, which results in 'Equivalent Oxide Thickness (EOT)' equal to 1nm. Both source and drain have ultra-shallow junction extensions which has a depth of 20nm for NMOS and 28nm for PMOS. For NMOS, a primary Indium implantation with high dose/high energy has been used followed by two additional implantations with lower dose to achieve a doping profile as close as possible to experimental data. For source/drain junction extensions, arsenic has been used as dopant with low energy followed by a 4 seconds RTA process. Figure 3.3 shows final retrograde doping profile in the channel and the doping profile in the source/drain shallow junction extensions resulted from TCAD process simulations compared with the measurement data [120]. The halo doping has been introduced by Boron implantation via source/drain side with 30^o tilt which enables the dopants to penetrate into the substrate in a region in front of source/drain under the gate. Introduction of halo doping in this region effectively avoids lateral source/drain field penetration into the substrate which degrades the performance of small dimension MOSFETs.

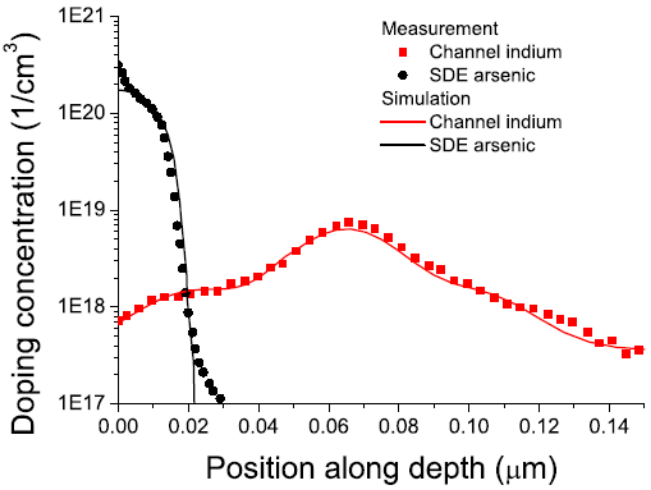
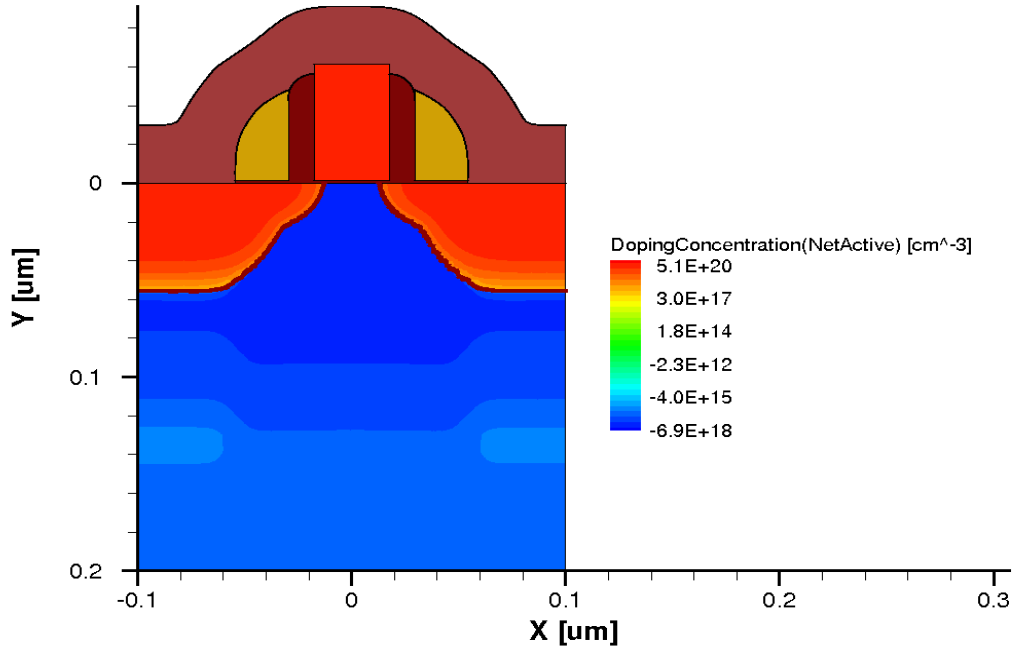


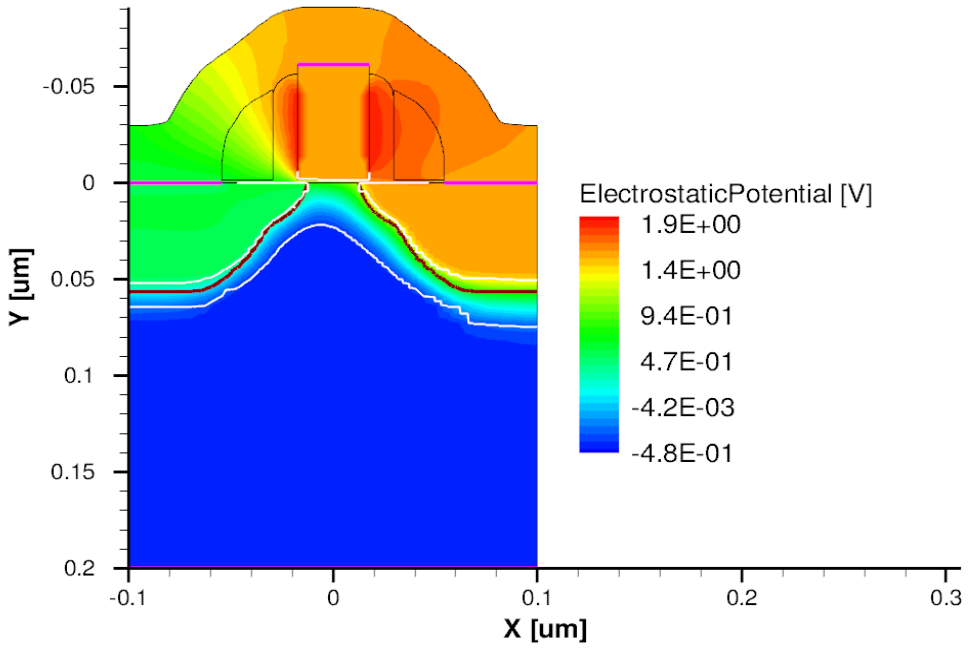
Figure 3.3: Comparison of channel and source/drain extension doping profiles of nMOSFET between TCAD simulation and measurement data, from [120].

Strain engineering has been introduced in the last step of process simulation by using tensile 'Contact Etch Stop Layer (CESL)' to increase the electron mobility in the channel for n-channel MOSFET. Figure 3.4(a) shows the final doping profile of n-channel device based on a realistic process flow simulations using the TCAD tool Sentaurus from Synopsys. Figure 3.4(b) demonstrates the electrostatic potential for the same transistor under high gate, high drain bias conditions (V_d=V_g=1v). The depletion region is obviously wider on the drain side compared to the source side.

Similar process simulation steps have been carried out for the complementary 35nm PMOS device with one additional step to introduce compressive strain. This additional step involves Silicon recess etching in source/drain regions and then formation of SiGe epitaxial layer in those regions. Introducing 30% of Ge content in embedded SiGe results in enhancing of mechanical stress into channel which in turn will increase the hole mobility in p-channel MOSFET as reported by Intel [119,121]. SiGe regions are created in Σ shape and have a close proximity of around 12nm to the channel. After source/drain formation, a compressive CESL layer is deposited over PMOS transistor as a part of stress enhancement techniques. The final doping profile of the designed P-channel device is shown in Figure 3.5(a). The potential profile at high gate, high drain bias also demonstrated in Figure 3.5(b) which clearly depicts Σ shaped SiGe source/drain regions.

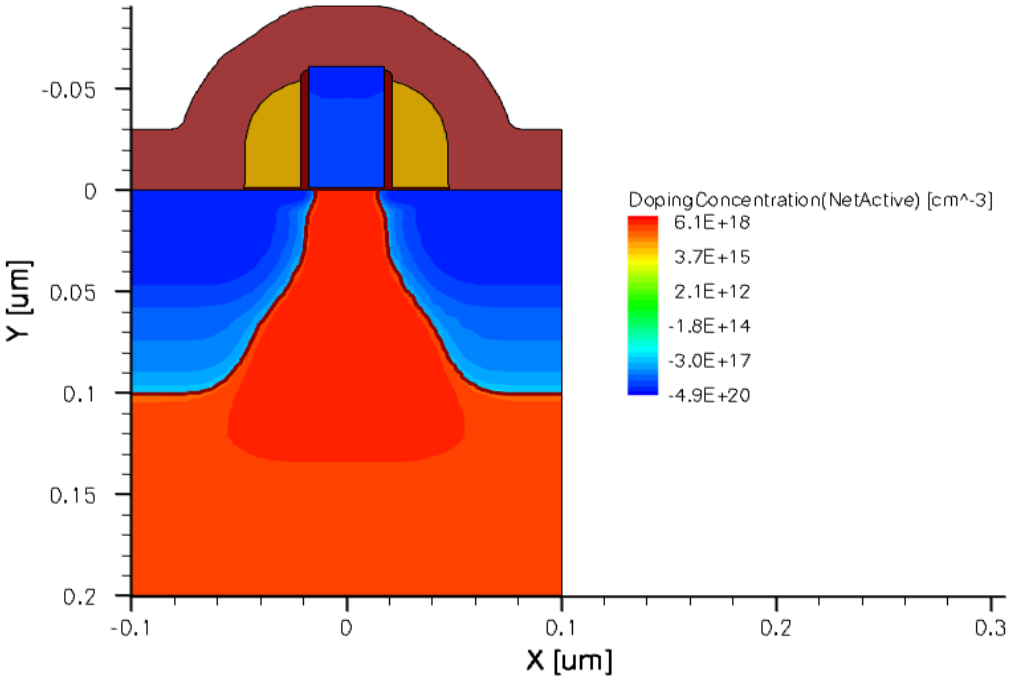


(a)

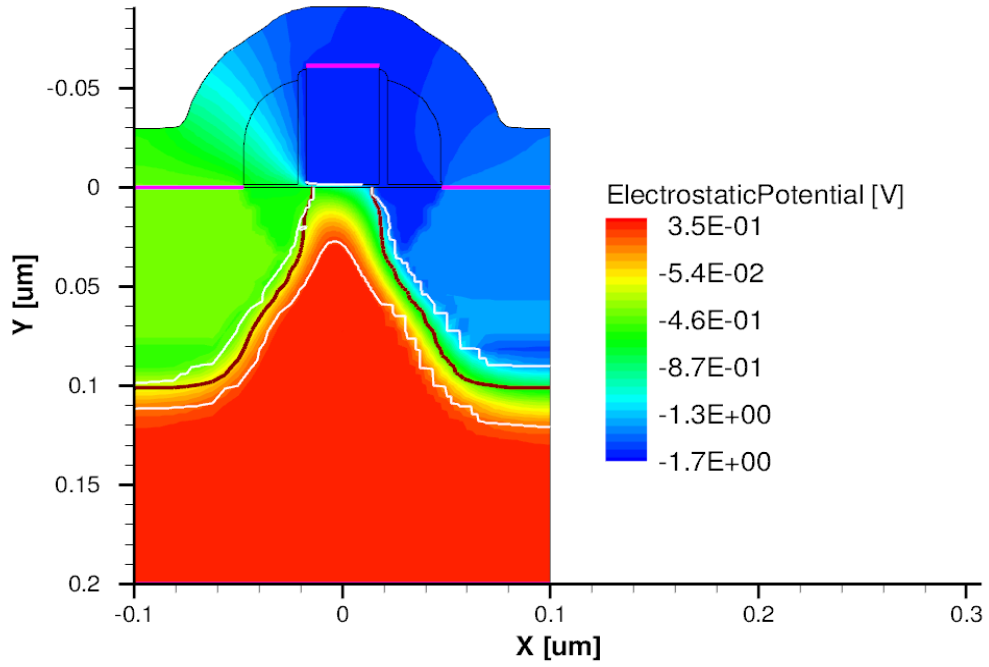


(b)

Figure 3.4: N-channel 35nm MOSFET used as a test bed device, (a) doping profile; (b) electrostatic potential profile for high drain and gate bias conditions (V_d=V_g=1v).



(a)



(b)

Figure 3.5: P-channel 35nm MOSFET used as a test bed device, (a) Doping profile,(b) electrostatic potential profile for high drain and high gate bias conditions (V_d=V_g= -1v).

The next step is the device simulation. The primary targets of device simulation steps are the electrical I_d-V_g and I_d-V_d characteristics. The calibration starts by adjusting the device electrostatics property to match the subthreshold slope (SS) and DIBL parameters in respect to corresponding values from fabricated transistor by Intel [117,118,119]. Then the mobility models available in TCAD tool should be adjusted to match I_d-V_g characteristics in 'above threshold' region. Also, I_d-V_d curves will be checked in TCAD simulation results to make sure that they are in reasonable errors in respect to measurement data [120]. Table 3.1 includes the value of device 'figures of merit' at high voltage drain bias conditions (V_d=1v) from physical TCAD simulations for one micrometer width devices.

Table 3.1: Important electrical 'figures of merit' at high voltage drain bias (V_d=1v)

Device	**I_{on} (mA)**	**I_{off} (μA)**	**V_{th} (V)**	**DIBL**	**SS (mv/dec)**
NMOS	1.26	0.1	0.18	0.105	87.6
PMOS	0.55	0.115	0.19	0.136	88.5

In Table 3.1, the drive current (I_{on}) is defined as the device current at high gate and high drain bias (i.e. V_g=V_d=1v for NMOS), the leakage current (I_{off}) is defined as the device current at low gate and high drain bias (i.e. V_g=0v, V_d=1v for NMOS). The threshold voltage (V_{th}) is extracted from I_d-V_g curves at a constant threshold current of 10 micro Ampers per micrometer. Drain Induced Barrier Lowering (DIBL) parameter is calculated from the absolute difference in the values of the threshold voltage at high and low drain bias voltages divided by the corresponding difference in drain bias voltages. Sub-threshold slope is calculated from I_d-V_g characteristics at logarithmic scale and is defined as the inverse of the arithmetic slope of I_d-V_g between first and second current data points.

Significant reduction of the drive current in PMOS compared to NMOS is due to lower mobility of holes in comparison with electrons. The leakage current in PMOS is 15% more than NMOS while the DIBL is increased about 30% in PMOS compared with NMOS device. The sub-threshold slope is a few percent worse. This results from stronger short channel effects (SCE) in the P-channel devices.

3.3 Parameter Extraction

Any compact model in terms of device characteristics can be separated into two major parts: static (DC) part and dynamic or capacitance (AC) part. In this thesis, we focus on the capability of PSP to capture the basic DC and AC behavior of MOSFETs at 35nm gate length technology node. Some second order effects such as substrate junction leakage current, GIDL, gate leakage and noise are not included in this study. A DC characteristics extraction strategy is the process of parameter extraction and optimization to match the I_d-V_g and I_d-V_d characteristics between the compact model results and TCAD or measurement data. The capacitance extraction strategy can be implemented based on DC extraction results to match the trans-capacitance (between terminal) components.

3.3.1 DC Parameter Extraction

DC parameter extraction process consists of a series of optimization steps. This task is carried out by assigning initial values to compact model parameters and then changing the values of target parameters in extraction steps in such a way that the device DC characteristics can replicate the reference curves as closely as possible. In other words, a global optimization should be performed to find model parameters that will fit reference data with minimum RMS error. The reference data is either physical device simulation or practical measurement data. The relative RMS error can be used to measure the accuracy of generated compact model set by:

$$E_{RMS} = \sqrt{\frac{1}{N}\left(\frac{y_{i,SIM} - y_i}{y_i}\right)^2} \tag{3.25}$$

where N denotes the number of data points, y_i is the data point from device physical simulation and $y_{i,SIM}$ is the corresponding data point from compact model simulation using conventional circuit simulators like SPICE. Different steps in parameter extraction must be followed to prevent sub-optimization or unphysical value assignment to compact model parameters [122]. Such a procedure which determines the smaller set of parameters which are targeted for extraction in each step and the order of different steps is called “parameter

extraction strategy". The device DC characteristics which should be used as a fitting target in an appropriate parameter extraction strategy are I_d-V_g (drain current versus gate voltage) at two different drain biases and I_d-V_d (drain current versus drain voltage) at different gate biases.

The PSP parameter extraction was carried out using BSIMProPLUS [123] software based on the latest supported version of PSP (102.2) available in the circuit simulator HSPICE [124]. Prior to parameter extraction procedure, some process related parameters should be provided in the input data file or assigned as a locked parameter in the extractor software. These most important parameters are listed in Table 3.2.

Table 3.2: Process related parameters in PSP extraction procedure

Parameter	**Physical Meaning**
L	Drawn channel length
W	Drawn channel width
LAP	Effective gate overlap with source/drain
TOXO	Gate oxide thickness
EPSROXO	Relative permittivity of gate dielectric
NPO	Gate polysilicon doping
NOVO	Effective doping of overlap region
TR	Reference temperature

The proposed parameter extraction strategy consists of six modules; each module tries to extract parameters from corresponding physical phenomena in the simulated device. It includes the following steps:

1. Extraction of "threshold voltage (V_{th})" related parameters: Although threshold voltage is not a parameter for surface potential compact models, it is one of the basic figures of merit that determines the device characteristics. Therefore, it's necessary to capture the threshold voltage behavior in the PSP at the first stage of parameter extraction. At the first order, threshold voltage is defined by the gate voltage at which the surface potential becomes twice of bulk potential ($\psi_s = 2\Phi_b$), a manipulation of equation (3.11) gives the threshold voltage by:

$$V_{th} = V_{fb} + 2\Phi_b + \frac{\sqrt{2\varepsilon_s q N_a (2\Phi_b - V_{bs})}}{C_{ox}} \tag{3.26}$$

where N_a is the substrate doping and other parameters mentioned previously in subsection 3.1.3. Equation (3.26) justifies extraction of two PSP model parameters in the first sub-step of this parameter extraction stage. These parameters are *VFBO* and *NSUBO* which model the flat-band voltage and substrate doping parameter in the compact model, respectively. These two parameters are extracted from I_d-V_g characteristics at low drain bias (V_{ds}=50mv) and different bulk bias (V_{bs}) voltages. In order to achieve this goal, V_{th} vs V_{bs} is present in Figure 3.6 to examine accuracy of PSP extraction compared to data from device TCAD simulations, with the RMS error of less than 1%. Different ways of extracting V_{th} from device I_d-V_g characteristics have been introduced in articles [125], among them two most popular methods are available in the extractor which are constant current criteria and maximum transconductance method. The former defines the threshold voltage as a gate voltage corresponding to a constant level of drain current but the latter extracts the gate voltage by extrapolating the derivative of I_d-V_g characteristic to zero or the maximum transconductance point. The maximum transconductance method has been used in the first step of parameter extraction.

2. In the second step, the parameters which affect the sub-threshold behavior of device characteristics will be added to those two basic V_{th} determining parameters to fit sub-threshold part of the I_d-V_g characteristics at low drain bias (V_{ds}=50mv) and different bulk bias (V_{bs}) voltages. A full set of parameters for this sub-step consist of *NSUBO, VFBO, CTO, DPHIBO, DNSUBO* and *VNSUBO*. The first two parameters were already introduced and *CTO* is called "interface state factor" and it is used in PSP to directly control sub-threshold slope. *DPHIBO* models "Offset for Φ_b" while *DNSUBO* and *VNSUBO* are two "effective doping bias-dependence parameters" [122]. The last three parameters are not extracted normally and will be set to their default value of zero. They will be extracted if the accuracy of fitting by using *NSUBO, VFBO* and *CTO* is not acceptable, i.e. if the RMS error for the reproduced sub-threshold part of the I_d-V_g characteristics with PSP model is more than 10% in comparison with physical TCAD data points.

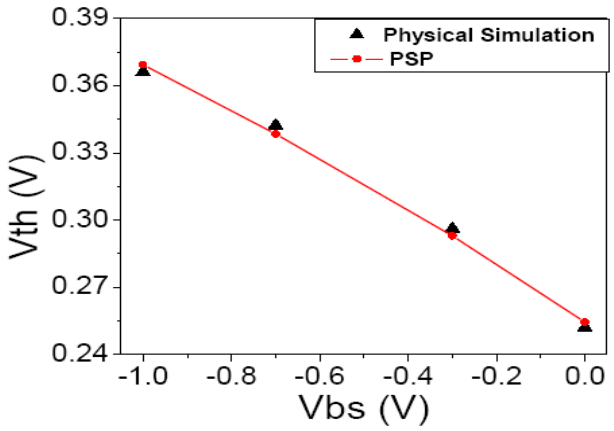


Figure 3.6: A comparison of V_{th} extraction between PSP model and device physical simulation at different substrate bias voltages.

3. Extraction of "Drain Induced Barrier Lowering (DIBL)" related parameters:

DIBL explains the threshold voltage reduction at high drain bias [104,125], therefore two DIBL related PSP parameters are extracted from I_d-V_g subthreshold characteristics at high drain bias (V_d=1v) and different bulk bias (V_{bs}) voltages. These parameters are *CFL* and *CFBO* and their role in DIBL description is given by:

$$\Delta V_{th} = CF.V_{ds}(1+CFBO.V_{sb}) \quad ; \; CF = K.CFL \tag{3.27}$$

where K is a device geometry dependant parameter and *CF*, *CFL* and *CFBO* are called “DIBL parameter”, "length dependence DIBL parameter" and "back-bias dependence DIBL parameter", respectively.

4. Extraction of Mobility and Source/Drain resistance related parameters:

Since the carrier mobility affects the shape and slope of device I_d-V_g characteristics in the above-threshold region, extraction of mobility related parameters will be carried out with several sub-steps. The target of parameter extraction is I_d-V_g at low drain bias (V_d=50mV). Most of mobility related parameters and the equation governing mobility in PSP were introduced in equation (3.19) and a list of parameters which are aimed for extraction are *UO*, *MUEO*, *THEMUO*, *CSO*, *XCORO*, *RSW1*, *RSBO* and *RSGO*. They are called "zero field mobility", "mobility reduction coefficient", "mobility reduction exponents", "coulomb

scattering parameter", "non universality parameter", "source/drain series resistance", "back bias dependence of series resistance" and "gate bias dependence of series resistance", respectively. The last three parameters are responsible for further reduction of drain current at high gate/high bulk bias voltages.

5. Extraction of "Velocity Saturation" related parameters:

As velocity saturation occurs for short channel devices, it is essential to extract parameter models from device I_d-V_d characteristics at different gate bias voltages. *THESATO, AXO* and *THESATGO* are the velocity saturation parameters which are the aim of the extraction at this stage and the equations governing this phenomenon were discussed in equations (3.21) to (3.23). *THESATO* is the "velocity saturation parameter" in PSP, *AXO* is "linear/saturation transition factor" and *THESATGO* is the "gate bias dependence of *THESATO*". Following the extraction of saturation related parameters, a sub-step for the re-extraction of mobility related parameters will be necessary since it will affect on the drain current. At the end, a final re-extraction of velocity saturation parameters will be carried out from I_d-V_d characteristics.

6. Extraction of "Channel length modulation (CLM)" related parameters:

Channel length modulation models non-zero slope of drain current in I_d-V_d characteristics and therefore in the final stage of parameter extraction, all of channel length modulation related parameters should be extracted from R_{out}-V_d at different gate bias voltages. The output resistance (R_{out}) is defined by the derivative of drain current versus drain voltage (dI_d/dV_d) to show the slopes of drain current more clearly in the saturation region of any I_d-V_d set. Equation (3.24) models the channel length modulation in PSP where two of model parameters are sufficient to be extracted at this stage; *ALP* and *ALP1* which are called "CLM pre-factor" and "CLM enhancement factor above threshold". The other parameter, *ALP2* which is called "CLM enhancement factor below threshold", will be extracted if at least one set of I_d-V_d for gate voltages below threshold voltage is available.

The above proposed six stage strategy has been used to extract PSP model parameters of 35nm uniform test bed MOSFETs which were discussed in section 3.2. Figures 3.7 and 3.8 illustrate the different device characteristics from compact model extraction compared with physical simulations for NMOS and PMOS devices, respectively. Both linear and

logarithmic drain current scale have been used to clearly demonstrate the accuracy of fitting for above-threshold and sub-threshold regions. Table 3.3 shows the RMS error of PSP extraction for each device characteristics.

Table 3.3: RMS error of PSP parameter extraction for uniform test bed devices

Device Characteristics	**I_d-V_g @V_d=50mv**	**I_d-V_g @V_d=1v**	**I_d-V_d**
NMOS	**2.93%**	**2.48%**	**2.16%**
PMOS	**2.70%**	**2.05%**	**2.41%**

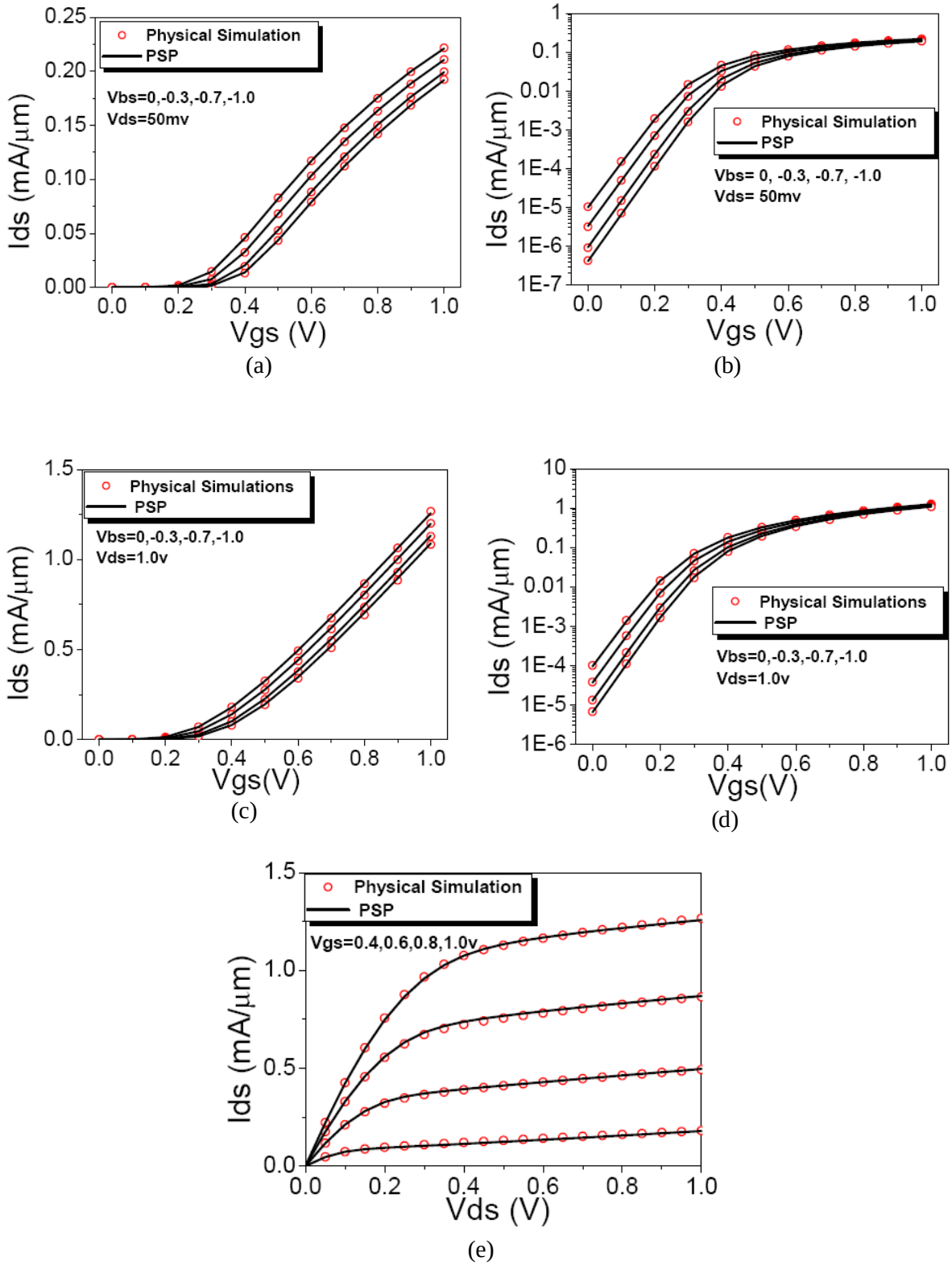


Figure 3.7: Uniform 35nm NMOS characteristics from PSP parameter extraction (solid lines) and physical simulations (symbols); (a,b): I_d-V_g at low drain bias (V_{ds}=50mv) for different substrate bias voltages in linear and logarithmic demonstration; (c,d): I_d-V_g at high drain bias (V_{ds}=1.0v) for different substrate bias voltages in linear and logarithmic demonstration; (e): I_d-V_d at different gate bias voltages.

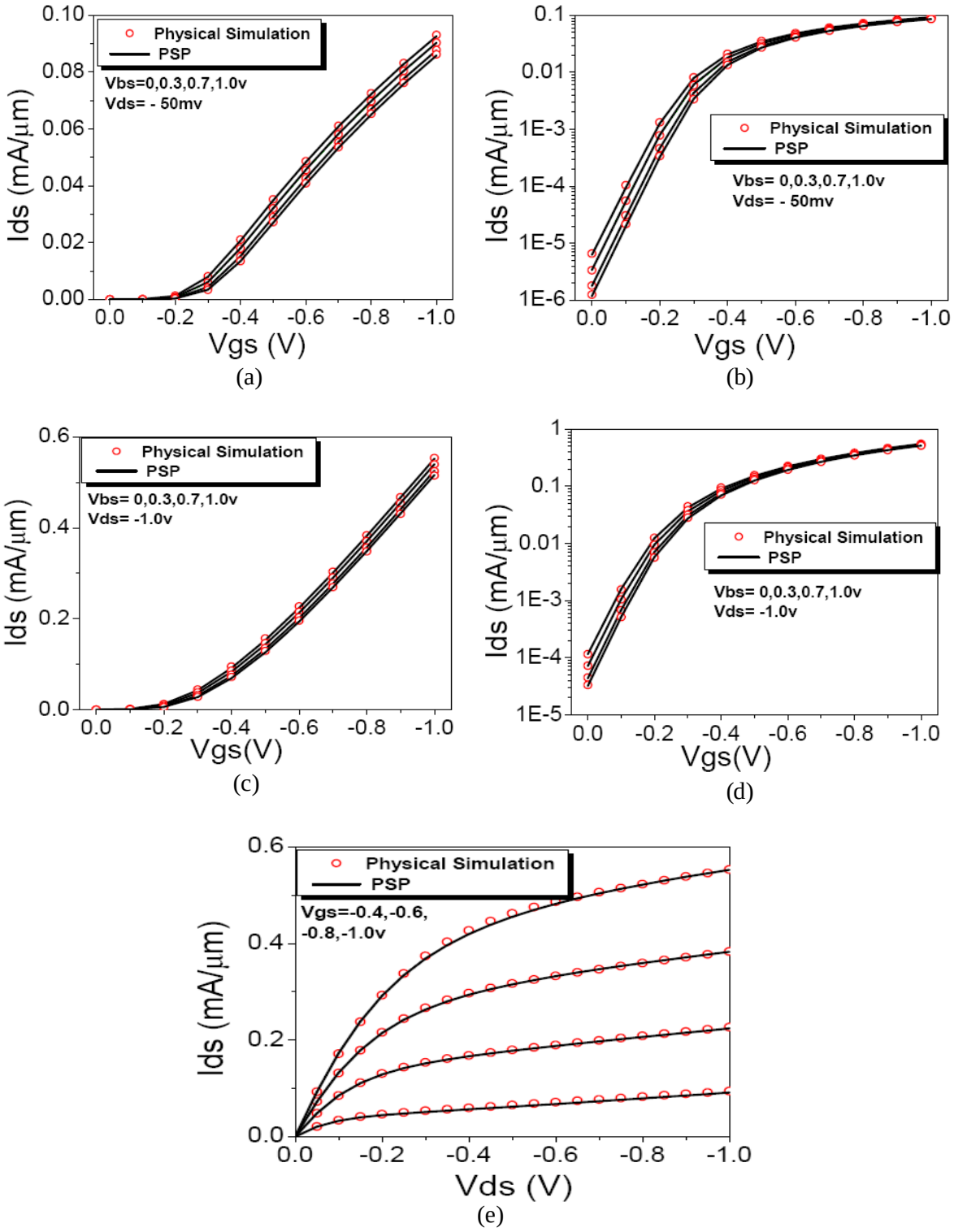


Figure 3.8: Uniform 35nm PMOS electrical characteristics from PSP parameter extraction (solid lines) and physical simulations (symbols); (a,b): I_d-V_g at low drain bias (V_{ds}= -50mv) for different substrate bias voltages in linear and logarithmic demonstration; (c,d): I_d-V_g at high drain bias (V_{ds}= -1v) for different substrate bias voltages in linear and logarithmic demonstration; (e): I_d-V_d at different gate bias voltages.

Apart from the above single device parameter extraction, the proposed extraction strategy can be employed to extract parameters of a group of devices with different channel length/width as well, and this is called “Batch Extraction”.

For the batch extraction, one additional stage will be necessary following extraction of parameters for each single device. This extra stage aims to extrapolate different parameters of batch devices in such a way that the trend of parameters can be described either versus length or width with corresponding “geometry independent part”, “length dependence part”, “width dependence part” and “area dependence part” as was discussed in the hierarchy of PSP model from local to global level with an example Equation (3.1).

A graphical representation of additional stage in ‘Batch Extraction’ is shown in Figure 3.9 for two typical PSP parameters *THESAT* and *CF*; both of them have already been defined. The devices under batch parameter extraction consist of the previous 35nm test bed NMOS along with two other NMOS devices in the same process and the same width of one micrometer but different channel lengths of 30nm and 40nm. The benefit of this batch extraction is that a circuit designer can use any channel length in the interval between 30nm to 40nm for different parts of a circuit due to the fact that the parameters in the compact model resulted from batch extraction have a continuous trend made with batch extraction strategy, i.e. red lines in Figure 3.9. The drawback of batch extraction strategy is that it is less accurate in reproducing physical simulation results due to the additional error at the parameter length or width dependent trend. The errors in Table 3.3 will be increased to about 8% for 35nm NMOS in the batch extraction compared with 2% RMS error in single device extraction.

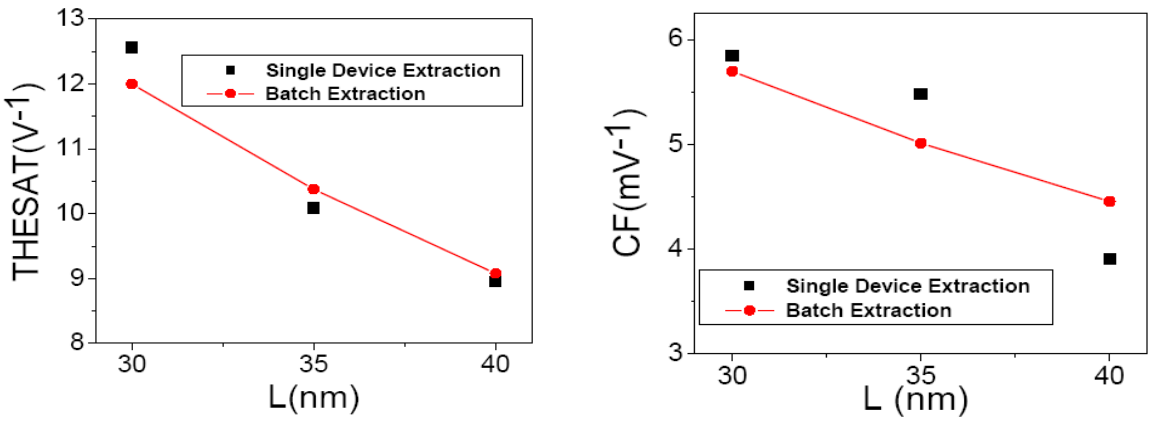


Figure 3.9: Batch extraction of two typical parameters from single device extraction results.

3.3.2 AC Parameter Extraction

The description of dynamic charge distributions in both intrinsic and extrinsic parts of a transistor is an important part of any device compact model. This is due to the fact that in real circuit, most of the devices work in transient or AC conditions rather than DC. The good accuracy of this charge descriptions are proven in rather long channel devices [112,125], however for the deca-nanometer scaled devices some degradations in the accuracy of these charge expressions are inevitable.

Under quasi static approximation, MOSFET terminal charges follow terminal voltages. It is clear that sum of these terminal charges should be zero to preserve charge conservation principle. Modern compact models describe terminal charges with some explicit expressions and change of these terminal charges in response to terminal voltages defines the intrinsic capacitances of a MOSFET. The extrinsic charge or parasitic capacitance of MOSFET is caused by source/drain overlap region with the gate and fringing effect which comes from the non-uniform electric field in the gate edges. These two capacitance components (overlap and fringing) make a significant part of the total gate capacitance for devices at the deca-nanometer regime. Both intrinsic and extrinsic charges add together to make total charge of each terminal. The capacitances between terminals are given by [101,105]:

$$C_{ij} = (2\delta_{ij} - 1)\frac{\partial Q_i}{\partial V_j} \quad ; \quad i, j \in \{g, s, d, b\} \tag{3.28}$$

where Q_i and V_j denote terminal charge and terminal voltage in reference to ground, respectively and δ_{ij} is the "Kronecker" delta function which is given by:

$$\delta_{ij} = \begin{cases} 1 & i \neq j \\ 0 & i = j \end{cases} \tag{3.29}$$

A complete strategy of parameter extraction in order to match the capacitances with

original TCAD simulations is as follows: It starts from DC parameter extraction to match I_d-V_g and I_d-V_d characteristics with TCAD simulations; this part also makes the internal charge or intrinsic capacitance available. The second stage is to extract extrinsic part which is made by right selection of two PSP model parameters: *LOV* and *CFRW*. These parameters are called "overlap length for gate/drain and gate/source overlap capacitance" and "outer fringe capacitance for a channel width of one micrometer", respectively. The overlap capacitance is then calculated internally in the model from:

$$CGOV = W.LOV.C_{ox} \tag{3.30}$$

where W is the device width and C_{ox} is the oxide capacitance per unit area. In PSP, it is possible to assign different values to *LOV* and *LAP*, the latter is an important parameter in DC part which was introduced in Table 3.1 with the name of "Effective gate overlap with source/drain". This gives PSP more capability to simultaneously match both C-V and I-V characteristics independently. The exact value of *LAP* is a bit different from metallurgical overlap length and details of its extraction from gate-bulk capacitance component are given in [126,127].

The third stage of the PSP capacitance extraction strategy needs to extract source/bulk and drain/bulk junction parameter values. The important junction parameters in PSP are *CJORBOT* and *CJORGAT* which are called "zero bias junction capacitances per unit area of the bottom component" and "zero bias junction capacitance per unit length of the gate edge component". If the device has STI, *CJORSTI* which is a parameter which defines the "zero bias capacitance per unit length of STI edge" will be important and has to be extracted. The important PSP flag parameter *SWJUNCAP* makes the junction part of the PSP available for the simulations. This parameter is used to determine the meaning and usage of the junction instance parameters such as junction area and junction perimeter for Source/Bulk and Drain/Bulk junctions [122]. Figure 3.10 illustrates the graphical representation of five capacitance component which are essential components of designing appropriate AC parameter extraction strategy.

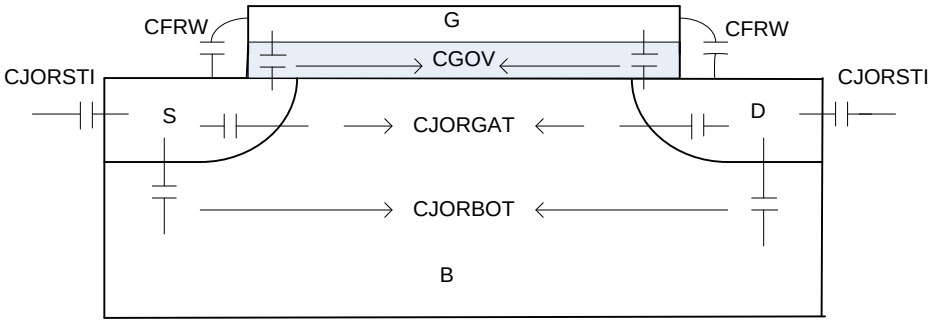


Figure 3.10: Five capacitance components in extrinsic and junction part of PSP.

Based on Equation (3.28), a so called transcapacitance matrix is given by:

$$C_T = \begin{bmatrix} C_{gg} & C_{gs} & C_{gd} & C_{gb} \\ C_{sg} & C_{ss} & C_{sd} & C_{sb} \\ C_{dg} & C_{ds} & C_{dd} & C_{db} \\ C_{bg} & C_{bs} & C_{bd} & C_{bb} \end{bmatrix} \tag{3.31}$$

In the transcapacitance matrix, the sum of elements on each row and on each column should be equal to zero [125], thus from these 16 elements, only 9 elements are independent. We have chosen C_{gg}, C_{gs}, C_{gd} from gate related components, C_{dd}, C_{ds}, C_{dg} from drain related components and C_{bb}, C_{bs}, C_{bd} from bulk related components as independent components for this study. These transcapacitances can be simulated with TCAD software to obtain physical simulation results and then with PSP compact model which is extracted by proposed AC extraction strategy to examine the accuracy of compact model to capture each transcapacitance component. Figures 3.11, 3.12 and 3.13 illustrate transcapacitance simulation results of test bed NMOS 35nm device from both TCAD physical simulations and PSP compact model results at two drain bias voltage of V_d=0V and V_d=1V. Figures 3.14, 3.15 and 3.16 show the same components for PMOS test bed 35nm device at two drain bias voltages of V_d=0V and V_d=-1V. All of transcapacitance components are in the order of femto Farads with device width of 1 micrometer. Tables 3.4 and 3.5 represent RMS error of PSP simulated transcapacitance components in respect to physical simulations at above mentioned drain bias conditions. The criterion for the RMS error calculation was given in Equation (3.25).

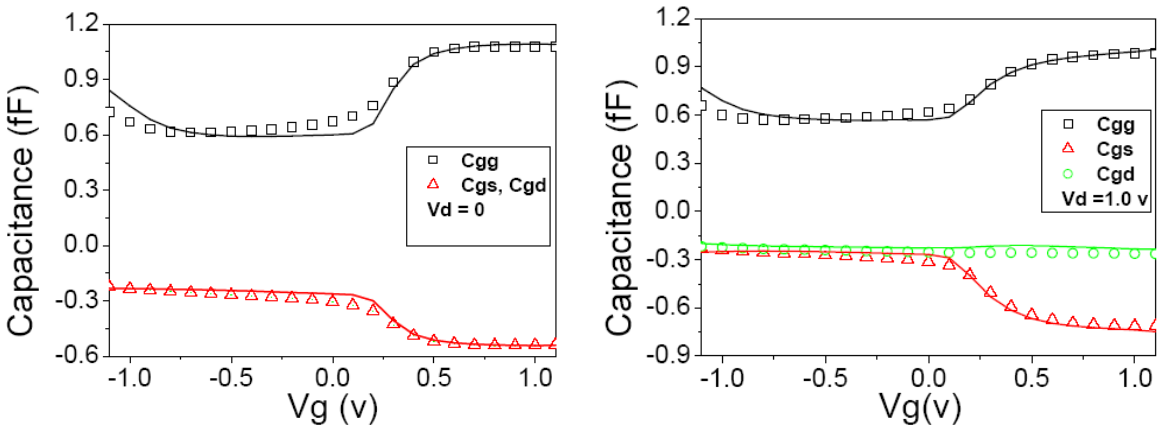


Figure 3.11: NMOS gate capacitance components versus gate voltage for low drain bias (left) and high drain bias (right); solid lines from PSP and symbols from TCAD.

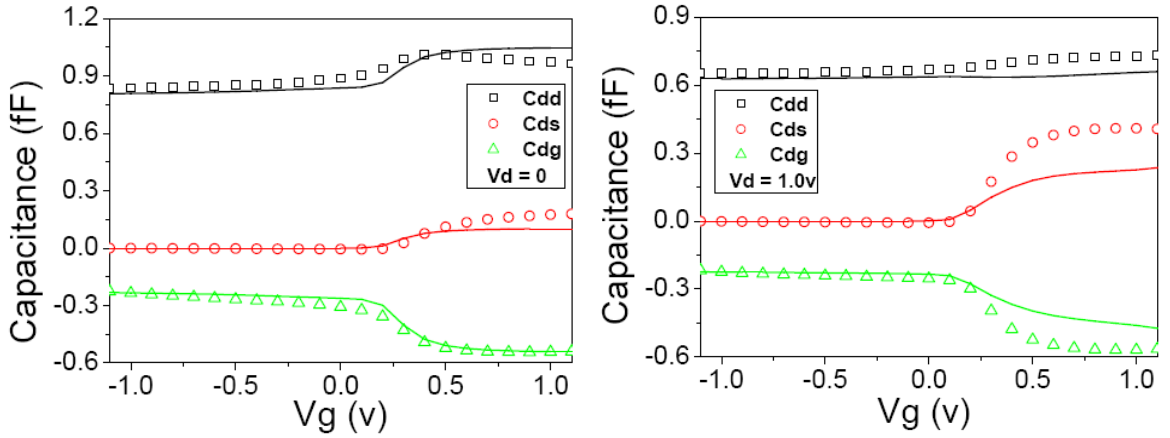


Figure 3.12: NMOS drain capacitance components versus gate voltage for low drain bias (left) and high drain bias (right); solid lines from PSP and symbols from TCAD.

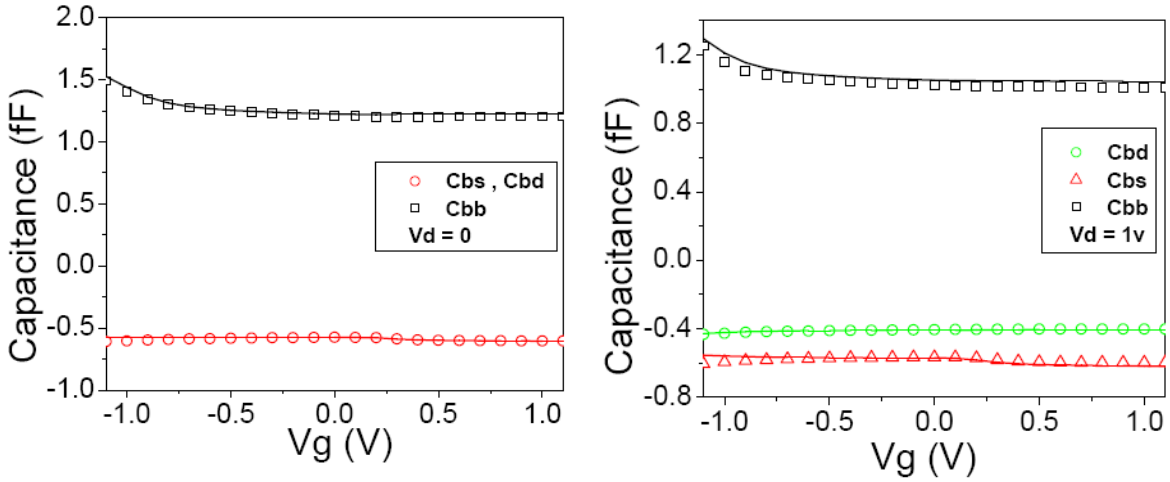


Figure 3.13: NMOS bulk capacitance components versus gate voltage for low drain bias (left) and high drain bias (right); solid lines from PSP and symbols from TCAD.

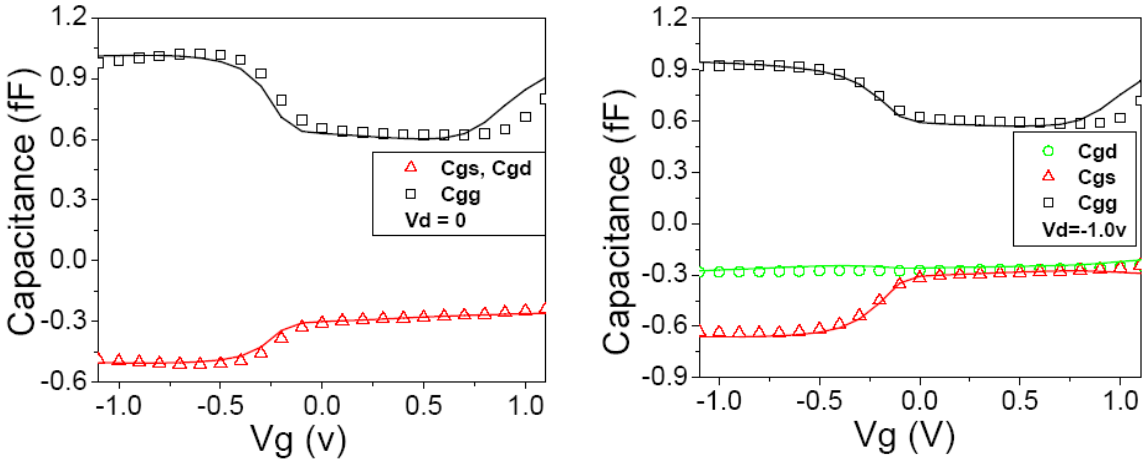


Figure 3.14: PMOS gate capacitance components versus gate voltage for low drain bias (left) and high drain bias (right); solid lines from PSP and symbols from TCAD.

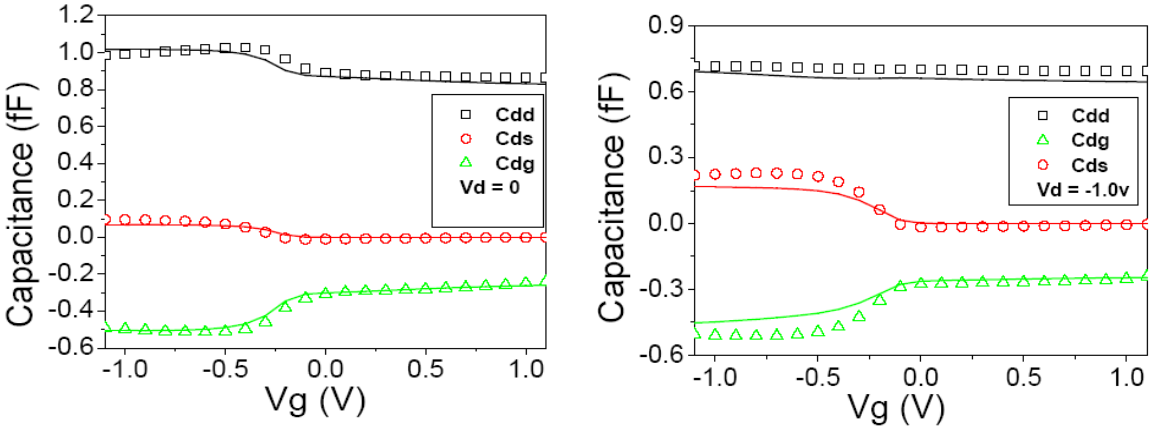


Figure 3.15: PMOS drain capacitance components versus gate voltage for low drain bias (left) and high drain bias (right); solid lines from PSP and symbols from TCAD.

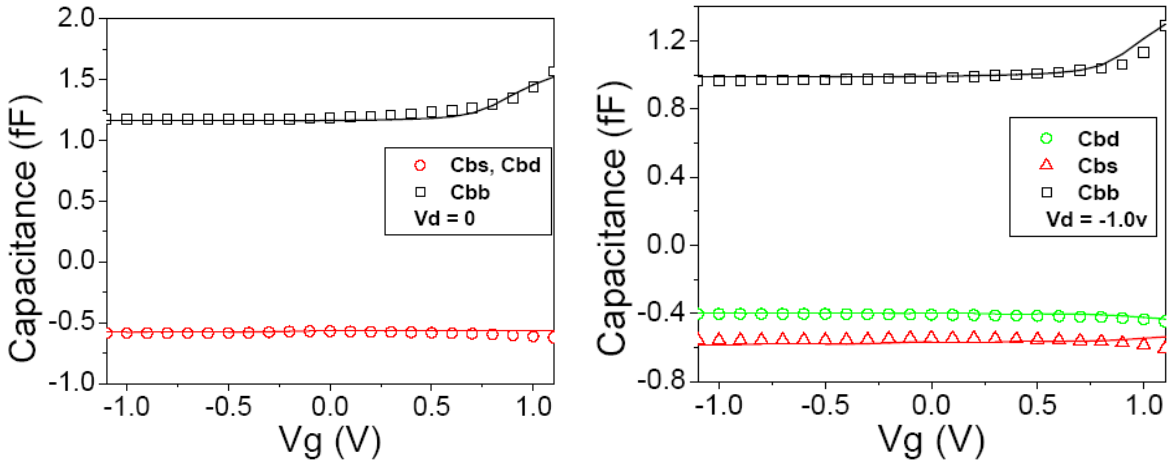


Figure 3.16: PMOS bulk capacitance components versus gate voltage for low drain bias (left) and high drain bias (right); solid lines from PSP and symbols from TCAD.

Table 3.4: RMS error of PSP transcapacitance components for NMOS 35nm test bed device

	Gate capacitance components			**Drain capacitances components**			**Bulk capacitances components**		
Drain bias	$\mathbf{C_{gg}}$	$\mathbf{C_{gs}}$	$\mathbf{C_{gd}}$	$\mathbf{C_{dd}}$	$\mathbf{C_{ds}}$	$\mathbf{C_{dg}}$	$\mathbf{C_{bb}}$	$\mathbf{C_{bs}}$	$\mathbf{C_{bd}}$
V_d=0v	7.25%	7.88%	7.88%	5.05%	20.05%	7.88%	1.42%	1.86%	1.86%
V_d=1v	6.14%	7.33%	11.76%	7.13%	28.58%	13.95%	3.31%	3.07%	0.82%

Table 3.5: RMS error of PSP transcapacitance components for PMOS 35nm test bed device

	Gate capacitance components			**Drain capacitances components**			**Bulk capacitances components**		
Drain bias	$\mathbf{C_{gg}}$	$\mathbf{C_{gs}}$	$\mathbf{C_{gd}}$	$\mathbf{C_{dd}}$	$\mathbf{C_{ds}}$	$\mathbf{C_{dg}}$	$\mathbf{C_{bb}}$	$\mathbf{C_{bs}}$	$\mathbf{C_{bd}}$
V_d=0v	7.51%	3.97%	3.97%	3.18%	17.84%	4.08%	2.23%	3.14%	3.14%
V_d= -1v	6.93%	5.30%	8.04%	5.96%	17.58%	10.21%	2.37%	4.36%	2.19%

By looking into Figures 3.11 to 3.16 and Tables 3.4 and 3.5, we conclude that for both n- and p-channel devices, the most accurate transcapacitances belong to bulk components though they are less dependent to gate voltage variations. The gate components are less accurate due to implemented channel charge partitioning scheme between source and drain. PSP uses Ward-Dutton charge partitioning scheme [128] which is a result of 1-D current continuity equation. This charge partitioning is accurate for uniformly doped channels but numerical studies have shown that in halo-doped channels, some errors are introduced by this charge partition method [105]. The least accurate transcapacitances observed in Tables 3.4 and 3.5 are in the drain components, especially $C_{ds,}$ which is almost zero for accumulation and weak inversion and PSP is able to reproduce it in these regions but the error is noticeable in strong inversion region.

The other conclusion from Tables 3.4 and 3.5 is that the impact of drain bias on the accuracy of transcapacitance components is not strong although with a few exceptions, the error in high drain bias components is a few percent larger than the error in low drain bias conditions. And the final conclusion is that PSP simulation results preserve device

symmetry between source and drain at zero drain bias, i.e. C_{gs}=C_{gd} and C_{bs}=C_{bd} at V_d=0, as expected from physical simulation results and it also preserves reciprocity at zero drain bias, i.e. C_{gd}=C_{dg} at V_d=0, which is seen from Figures 3.11, 3.12, 3.14 and 3.15. In other drain bias conditions in general, the transcapacitance components are not reciprocal and the assumption of reciprocity is inconsistent with charge conservation law [125].

3.3.3 Accuracy of Transient Time Simulations

When the MOSFET terminals are connected to voltage sources subject to AC or transient time variations, a displacement current will be added to each terminal current. It causes a non-zero current for the gate terminal due to capacitive coupling between gate and other terminals. Other terminal currents will be changed from their DC bias values and in general each terminal current will be expressed by [124,125]:

$$I_i(t) = I_{i,DC} + \sum_j \frac{\partial Q_j}{\partial V_j}\frac{\partial V_j}{\partial t} \quad ; \quad i, j \in \{g, s, d, b\} \tag{3.32}$$

where the first term of equation (3.32) accounts for the DC current of each terminal and the second term represents the effect of transcapacitance components on the current. In order to investigate the impact of the transcapacitances accuracy of the PSP compact model on circuit simulation, a CMOS digital inverter which is made from two complementary test bed 35nm devices has been considered for further transient time simulations.

In order to match the output rise and fall delay times, both transistors should have equal drive currents. Therefore, a wider PMOS has been chosen with a gate width of 2.3 micro meters while the NMOS has basic width of one micrometer. The applied input voltage is a pulse, linearly rising and falling between 0 and 1 with rise/fall time of 0.05 ps. The output is connected to a variable load capacitance with value $C_L = nC_0$ fF where n is an integer with the maximum value of the inverter fan-out and C_0 is the unit load of the inverter which is equal to total input capacitance of another similar CMOS inverter. Thus, the unit load capacitance will be equal to sum of oxide capacitances of two NMOS and PMOS devices which is 1.08 femto Farads. The supply voltage is 1.0V. Figure 3.17 shows the CMOS inverter with its input voltage.

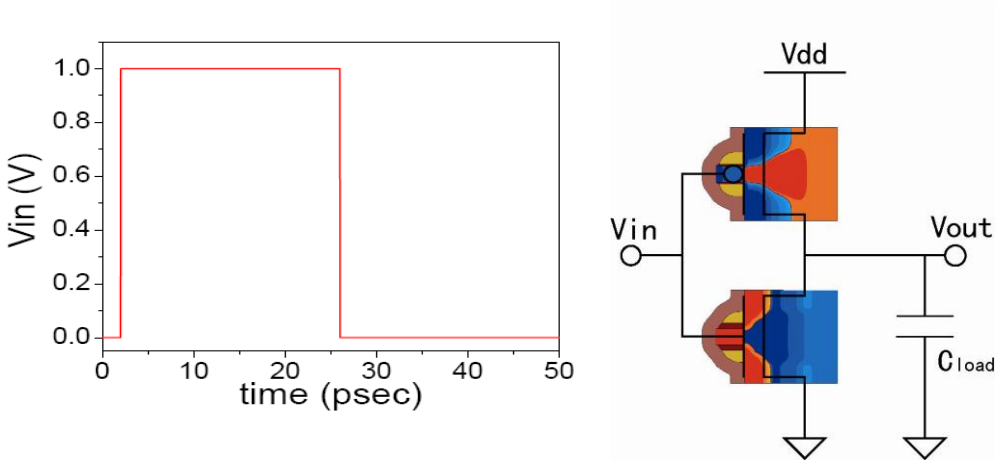


Figure 3.17: CMOS inverter consisting of two complementary 35nm test bed devices used for the purpose of transient time simulations (right) with its pulsed input voltage (left).

The inverter is simulated by both mixed mode TCAD (physical device simulator) and HSPICE (circuit simulator) with different load capacitances and the simulation results from TCAD are served as physical reference, while HSPICE simulation is carried out with PSP model cards to evaluate the accuracy of PSP built-in capacitance expressions in transient time simulations for our test bed 35nm devices. Figure 3.18 illustrates the output voltage from both simulations for two typical load capacitances which are 1.08 fF, the unit load capacitance of the inverter, and 10.8 fF corresponding to load condition with ten times of unit load, respectively. Although two methods produce close outputs, the transition delay times of HSPICE simulations using PSP model are a few percent more than TCAD physical simulations.

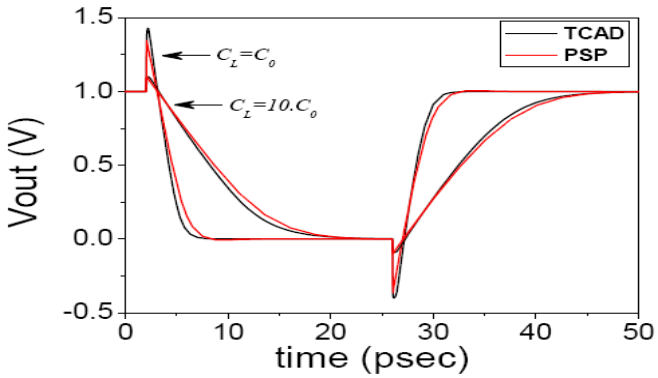


Figure 3.18: Simulated output of inverter shown in Figure 3.17 with two different load capacitances.

The rising output transition delay time (T_{DLH}) is determined by PMOS characteristics while the falling output transition delay time (T_{DHL}) will be controlled by NMOS. In this study, the inverter propagation delay times are defined as the time difference between the 50% transition points of the input and output signals.

Due to better fitting of drain transcapacitance components for the PMOS, it can be seen from Figure 3.18 that the T_{DLH} error is less than corresponding T_{DHL} error. The origin of both errors can be explained by observed transcapacitance accuracies. Starting with expansion of Equation (3.32) and using Equation (3.28) for one of devices (i.e. NMOS), the drain current is given by:

$$I_D(t) = I_{DC} + C_{ds}\frac{dV_S}{dt} + C_{db}\frac{dV_b}{dt} + C_{dg}\frac{dV_g}{dt} - C_{dd}\frac{dV_d}{dt} \tag{3.33}$$

Since the source and bulk terminals of the device are connected to ground, no time varying signal is present at these terminals and by ignoring the transient time of the gate voltage due to fast input slew rate, Equation (3.33) can be further simplified for a first order analysis when the gate voltage stays at high state (V_g=1v):

$$I_D(t) \cong I_{DC} - C_{dd}\frac{dV_d}{dt} \tag{3.34}$$

Considering the fact that PMOS is off and the NMOS current should be passed from load capacitance (i.e. $I_D(t) = -C_L(dV_{out}/dt)$ and $V_{out} = V_d$), Equation (3.34) can be used to evaluate rate of output voltage decay in terms of total drain capacitance and load capacitance:

$$\frac{dV_{out}}{dt} = -\frac{I_{DC}}{C_L - C_{dd}} \tag{3.35}$$

Since I_{DC} in Equation (3.35) is almost equal between PSP model and TCAD physical simulations, the main source of difference in the decay rate of the output voltage is caused by drain capacitance discrepancy. By Referring to Figures 3.12 at high drain bias conditions, it is understood that C_{dd} is underestimated with PSP model compared with TCAD simulations. Hence, the rate of output voltage decay from PSP is less than the corresponding rate from TCAD simulations.

With an empirical effort it was realized that this error can be substantially reduced by just adding a small negative capacitance in parallel with the load capacitance. The best value of this compensation capacitance was obtained around -0.88fF. However, due to different characteristics of transcapacitance components between NMOS and PMOS devices, a single compensation capacitance cannot result in perfect matching for both rise and fall propagation delay times.

Figure 3.19 shows the transition time error trend after using a compensation capacitance on both T_{DLH} and T_{DLH} versus different load capacitance values and compares it with original errors before using any compensation techniques. With exception of T_{DLH} error at unit load capacitance which has been increased, all other errors are improved after introduction of compensation capacitance. The error for the $T_{DHL,}$ in the inverter simulated with original compact models, varies from 6% to 16% while it is decreased to an interval between 0 and 5% for different load capacitance values after using compensation capacitance.

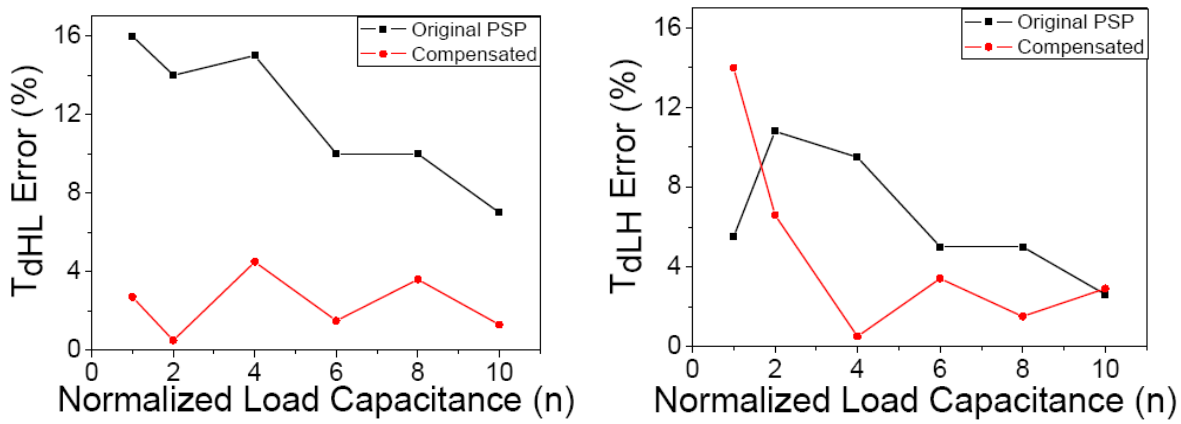


Figure 3.19: Trend of rising and falling output transition delay times for the test bed CMOS inverter with emphasis on the effect of compensation capacitance.

3.4 Summary

In this chapter we reviewed device model classifications, and then particularly focused on a surface potential based compact model, PSP. The mathematical background of PSP in terms of surface potential equation and drain current equation was investigated and important parameters of PSP in each part were reviewed. A continuous doping profile 35nm gate length MOSFET in both n-channel and p-channel types was introduced. These devices will be used as a test bed in this thesis. A PSP parameter extraction strategy was introduced to match the DC device electrical characteristics from compact model simulation with TCAD physical simulations while the RMS error remains less than 3%. An AC parameter extraction procedure then introduced and applied on test bed devices. The transcapacitance components were simulated with extracted PSP parameter set and compared with TCAD simulation results.

For the first time the accuracy of the SPICE transient simulations using deca-nanometer size transistors was evaluated in respect of mixed mode TCAD simulations. A compensation scheme was introduced to reduce propagation delay time errors of the inverter.

Chapter 4

Statistical Atomistic Simulation and Parameter Extraction

Statistical Variability (SV) which arises from discreteness of charge and granularity of matter is one of the fundamentally limiting factors of CMOS scaling and integration in the nanometer regime [24,30]. Due to its purely statistical nature, SV introduces increasing challenges for accurate compact modeling and statistical circuit simulation [2,29,129]. In order to achieve reasonable performance and yield in contemporary CMOS design, the SV has to be accurately simulated with an 'atomistic' simulator and then the resulting fluctuations in device characteristics should be translated into appropriate set of compact model parameters in order to assess the impact of these variations on circuit operation.

In the first section of this chapter, the impacts of SV on 35nm MOSFET characteristics are presented. This device will serve as the test-bed device in the statistical compact modeling study. In the second section, the context of statistical parameter set selection will be discussed based on sensitivity and error analysis of PSP compact model parameters. The methodology of statistical parameter extraction and optimization, and the accuracy of statistical parameter extraction are investigated in section 3. Section 4 outlines statistical circuit simulation using created SCM libraries. Statistical timing and power simulation of an inverter with emphasis on the impact of different input-output specifications and also the impact of the number of parameters chosen to capture the statistical behavior of delay and energy variability will be discussed in this section.

4.1 Simulation of Statistical Variability

The test bed device in this study is the 35nm gate length bulk MOSFET which was studied in the previous chapter but in presence of different sources of statistical variability. We used the Glasgow University 'atomistic' simulator GARAND which is essentially a 3D drift-diffusion simulator which employs density-gradient quantum corrections for electrons and holes to resolve the impact of individual impurities [66]. We simultaneously simulated the combined impacts of random discrete dopants (RDD), line-edge roughness (LER) and polysilicon grain granularity (PGG) in square devices (*W/L=1*) to reduce the computational complexity. The RDDs were generated based on the continuous doping profile from the Sentaurus process simulation by randomly placing dopant atoms on silicon lattice sites, with the probability determined by the local ratio between the dopant and silicon atom concentrations [55]. LER was introduced using 1D Fourier synthesis, with a power spectrum corresponding to a Gaussian autocorrelation function with a correlation length of 30nm and RMS amplitude of 1.3nm [71]. PGG was introduced by importing a random section of a large template polycrystalline silicon grain image for the entire gate region, with the average grain size of 65nm obtained through X-ray diffraction measurements. Because of the presence of acceptor type interface states along the grain boundaries, the Fermi level remained pinned at a certain position in the silicon band gap, 200 mV about the middle gap [71]. The impact of PGG has proved to be insignificant for PMOS atomistic simulations [130]. Figure 4.1 shows the electrostatic potential profile for a 35nm NMOSFET under influence of RDD, LER and PGG effects. The bias conditions for the simulation results shown in Figure 4.1 are V_g=0v and V_d=50mv. The surface potential is also shown in the same figures with the peaks at the point of random dopants. The RDD effect is clearly demonstrated in the bulk, source and drain regions as distinguished bright spots with higher potentials inside these regions. LER impact on the potential can be seen at the edges of source/drain regions and PGG effect is manifested by potential ridges in the channel region.

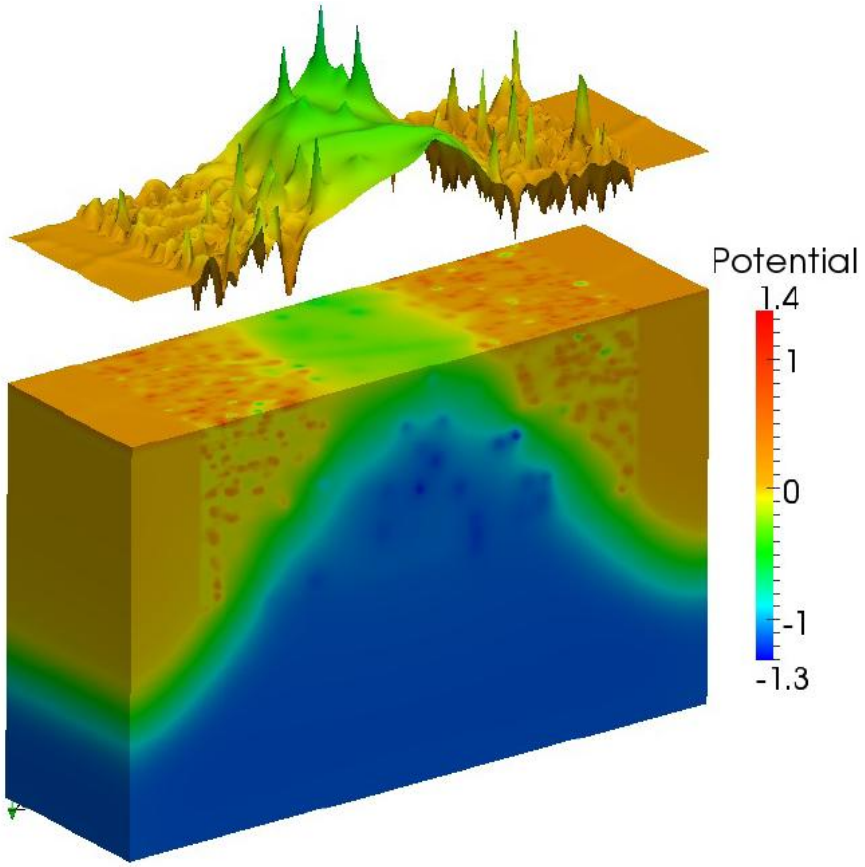


Figure 4.1: Typical potential profile in bulk, source and drain regions of a 35nm gate length device subject to RDD, LER and PGG effects. The surface potential is shown as a plot above the device.

Atomistic simulations were carried out for 1000 microscopically different devices at 35nm gete length for both NMOS and PMOS devices. The grid size for each device is 120 grid nodes along *x*-axis, 35 grids along *y*-axis and 75 grids along *z*-axis. This gives a total number of 315,000 mesh nodes. Three main equations which should be solved for each mesh node are Poisson's equation, current continuity equation and density gradient quantum correction equation which already stated in chapter 1. In order to facilitate numerical computations on this scale, it was necessary to employ Glasgow University Device Modeling Group's cluster. This cluster resource consist of 98 AMD Opteron CPUs, 90 Quad Core Intel Xeon E5530 CPUs and 31 Quad Core Intel Xeon E5462 CPUs at the time of running these simulations (February 2010).

Due to large scale of parallel simulations on the cluster, significant technical challenges associated with data convergence and tracking had to be overcome. Each device

simulation task on the cluster is called a job and the method which was used to enable the simulation of 1000 devices on the shared cluster is called "batch job submission". To achieve the aim of atomistic simulation of statistical variability, at least 2 batch job submissions were necessary to account for different drain bias conditions. While the simulation time of each job was expected to be about two days, several issues were found in job monitoring which resulted in the fact that we had to resubmit up to 30% of jobs in each batch at least once. These issues were observed as the active problems during simulations. Some of them failed to be completed in medium queues within 3 days necessitating job re-submission on long queues. Numerically unstable and non-converged jobs were resubmitted with new set of built-in iteration-convergence parameters. Other intermittent failures occurred during simulations due to issues with either hardware or software on the cluster resulted in complete loss of a set of running jobs. All of these problems increased the simulation time of each submitted batch to more than one month.

Figure 4.2 illustrates simulated I_d-V_g characteristics of 1000 samples of 35nm gate length N-MOSFET square devices (W=L) under the influence of combined sources of statistical variability. The current is normalized for the width of 1 micro meter. Red lines represent the current in a logarithmic scale to emphasize the impact of statistical variability on the sub-threshold region while blue lines show the current in linear scale to illustrate this impact on above-threshold region. Solid black line in the middle of each graph corresponds to 35nm uniform device characteristics without associated source of variability as we discussed it in previous chapter. The distribution of three important electrical parameter of the device obtained from physical device simulations are shown in Figure 4.3. They are: (a)- Drive current, I_{on}, which is the drain current at the bias point of V_g=V_d=1v, (b)- Leakage current, I_{off}, which is the drain current at the bias point of V_g=0v, V_d=1v, (c)- Threshold voltage, V_{th}, which is based on the constant current criterion of 1e-5A/um.

The leakage current spreads more than three orders of magnitude, indicating that SV strongly impacts the device electrostatic-dominated sub-threshold behavior. It is well known that drift-diffusion based simulations can underestimate I_{on} variability [71], but the difference between the maximum and minimum drive currents in these simulations is still 80% of the current's mean. Although digital designs commonly use multi-width devices, this variation level will still significantly impact the yield and performance of the circuit

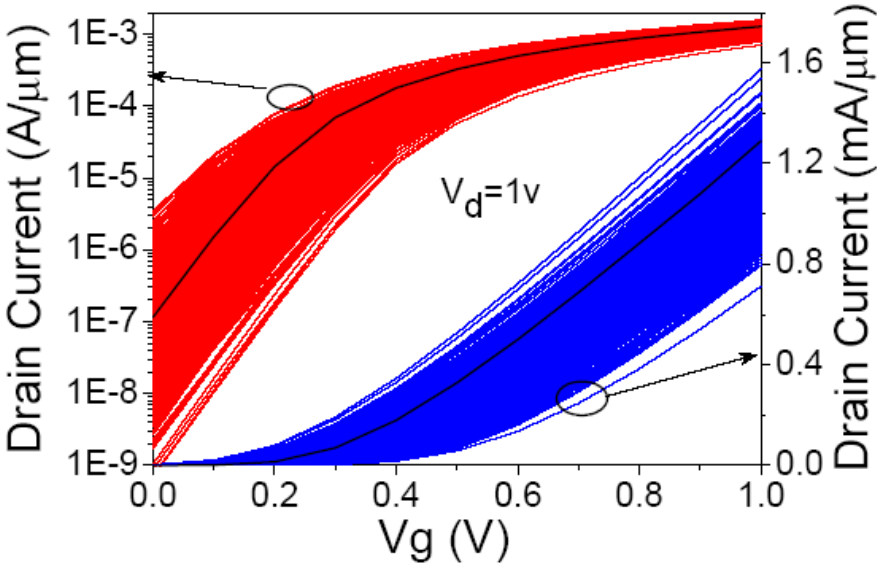


Figure 4.2: Simulated variability in I_d-V_g characteristics of a statistical sample of 1000 microscopically different 35nm gate length N-MOSFETs at V_d=1v. Black solid lines in the middle represent the I_d-V_g of uniform device without source of variability.

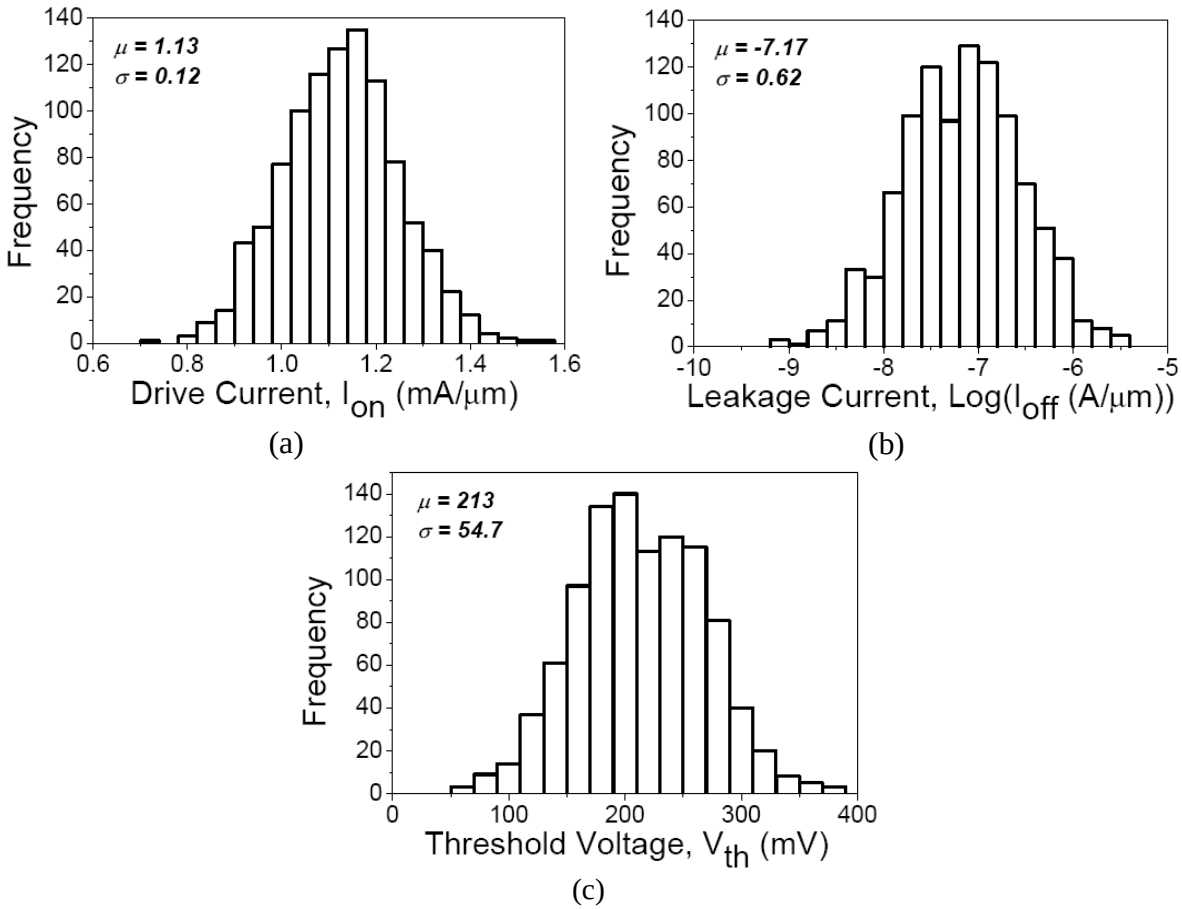


Figure 4.3: Distribution of (a)-I_{on}, (b)-I_{off}, (c)-V_{th} resulted from 'atomistic' simulation of statistical variability for 1000 microscopically different 35nm gate length N-MOSFETs at V_d=1v.

and system. Using the methodology described in [129], we can employ compact models extracted for the simulated devices in circuit simulations involving realistic devices with larger channel width. The histograms of device figure of merits distributions are presented in figure 4.3 and the normalized standard deviations (σ/μ) of I_{on}, $\log(I_{off})$ and V_{th} are at 10%, 8.7% and 25%, respectively.

4.2 Statistical Parameter Set Selection

The impact of statistical variability (SV) in the device characteristics, as discussed and simulated in the previous section of this chapter, must be accurately represented by a set of carefully chosen parameters in a compact model. This is due to the fact that the current industrial standard compact models do not have natural parameters designed to incorporate seamlessly the truly statistical variability associated with RDD, LER, PGG and other relevant variability sources. Once we have an accurate nominal PSP model parameters of a 35nm uniform test bed device with no source of variability inside it, as described in chapter 3, selection of an optimal set of parameters will be necessary to translate SV induced drain current deviations in the compact model. Statistical compact modelling is the bridge between device technology and circuit design and is essential in variability-aware contemporary CMOS circuit design.

We have carried out a first order sensitivity analysis to identify most important statistical parameters that should be extracted at the second stage of statistical extraction. The small signal, first order sensitivity response of drain current in respected to variations of each PSP model parameter is defined by:

$$S = \left| \frac{\Delta I_D / I_D}{\Delta P / P} \right| \tag{4.1}$$

where I_D is the drain current of the nominal compact model and P is the compact model parameter under investigation. ΔI_D is the increment in the drain current resulting from an increment in parameter value by ΔP.

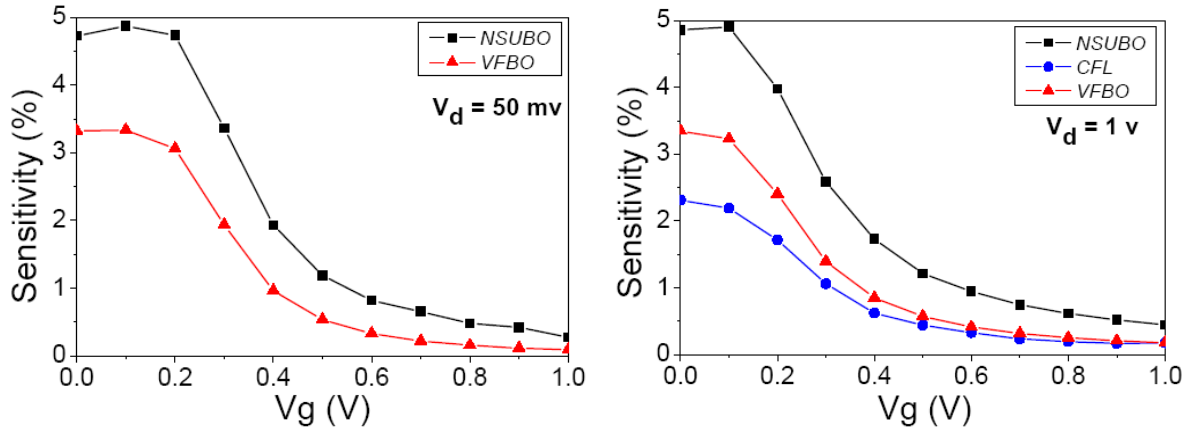


Figure 4.4: Sensitivity of drain current versus gate voltage for 3 model parameters at two drain bias voltages; left-V_d=50mv and right-V_d=1v.

Therefore, we ran the device simulation with circuit simulator HSPICE in two situations: one for nominal compact model parameter set and another for a modified compact model parameter set with targeted parameter value of $P + \Delta P$. Since the sensitivity is dependent on the bias conditions of the device, the sensitivity analysis was carried out for a range of gate voltages between 0 and 1 volts and two particular drain bias points of interest, 50mv and 1v.

Figure 4.4 illustrates the results of drain current sensitivity analysis for those parameters that lead to maximum sensitivities of more than 2%. Two parameters *NSUBO* and *VFBO* produce sensitivities almost independent of drain bias voltages and decreasing versus gate voltage while parameter *CFL* which models DIBL effects in PSP, causes drain current variations just at high drain bias as expected.

Figure 4.5 illustrates the drain current sensitivity for all other candidate parameters if the analysis leads to less than 2% but more than 0.1% sensitivity. In other words, the parameters which produce drain current sensitivity less than 0.1% are withdrawn from the following discussion. In Figure 4.5, six parameters are producing current sensitivities more than 0.1% at low drain bias (V_d=50mv) while a larger set of seven parameters are considered for high drain sensitivity analysis (V_d=1v).

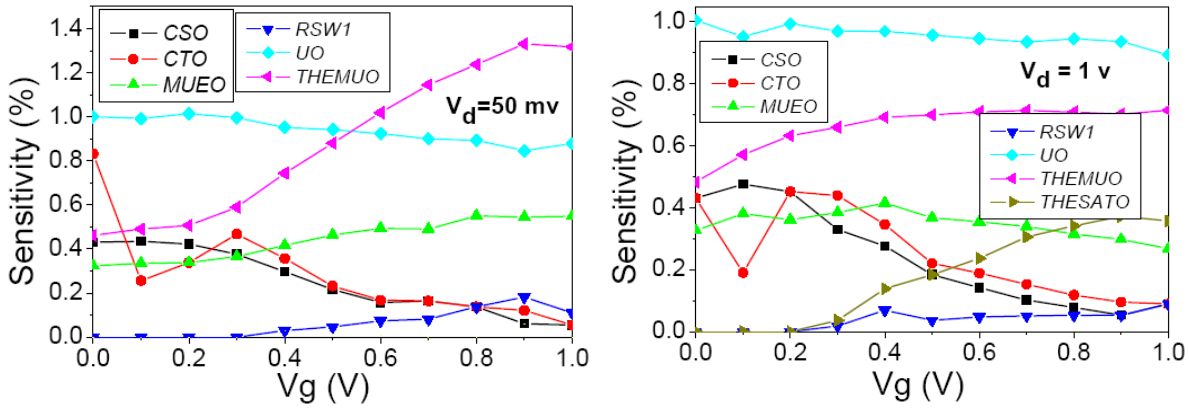


Figure 4.5: Sensitivity of drain current versus gate voltage for 7 model parameters at two drain bias voltages; left-V_d=50mv and right- V_d=1v.

The least sensitive parameter is *RSW1* which describes the source/drain series resistance in the model. It has a maximum sensitivity of 0.2% at above-threshold region and weak sensitivity value at sub-threshold region as expected. The parameter *THESATO* accounts for velocity saturation and is important at high drain bias. Parameter *UO* which models the mobility has almost flat trend versus different gate voltages. Three parameters *CSO*, *THEMUO* and *MUEO* model mobility degradation and the parameter *CTO* affects on the sub-threshold swing of the device characteristics. For parameter sensitivity study, a set of 10 PSP model parameters is identified as potential candidates to be used for further adjustment in a statistical parameter extractions procedure.

In order to capture the impact of SV with optimum set of parameters in PSP compact model, an average RMS error of less than 3% will be desirable. This is due to the fact that the corresponding error in full parameter extraction of a uniform device was around or less than 3%, with referring to Table 3.3 in the previous chapter. We have used the RMS error as defined in Eq. 3.25 to demonstrate quality of fitting for a target parameter set in statistical extractions. Based on the parameter sensitivity strength and their physical content, an optimum set of 7-parameter set has been selected. The following discussion clarifies the process of selecting 7-parameters from existing set of 10 parameters.

NSUBO and *VFBO* can be used to adjust the threshold voltage (V_{th}) of each atomistic device and *CTO* is the interface states factor parameter and it can be used to adjust the sub-threshold swing accordingly. However, we found that by use of flat-band voltage

parameter (*VFBO*) in statistical extractions, it is not possible to effectively reduce the RMS fitting error. This is due to the fact that the dominant source of SV in 35nm device which is Random Dopant Fluctuations (RDD), is captured effectively by using the substrate doping parameter (*NSUBO*) alone. A recent study on another surface-potential based model, HiSIM2, presented similar observations [131]. The parameter *CFL* is needed in statistical parameter extraction to capture DIBL variation. The parameter *UO* models the low field mobility but 3 parameters *THEMUO*, *MUEO* and *CSO* model mobility degradations caused by different physical process. Although *THEMUO* gives the higher first order sensitivity in the drain current compared with *MUEO* and *CSO*, the effectiveness of each parameter in total RMS error reduction in combination with aforementioned parameters is different. The average fitting error for each set was evaluated with our statistical extractor. It was 3.03%, 4.65% and 4.36% for inclusion of *CSO*, *MUEO* and *THEMUO*, respectively. This means that use of *CSO* is more efficient in error reduction for statistical parameter extraction and also makes it possible to reduce the statistical parameter set size. Moreover, *CSO* has introduced in PSP to model the mobility degradation due to coulomb scattering phenomena and selection of it in presence of RDD makes more physical insight in parameter selection.

A final 7 parameter set is selected to capture impact of SV on device operation based on the sensitivity analysis and the role of each parameter in reproducing device characteristics. Table 4.1 represents the final statistical parameter set with the physical meaning of each parameter. The parameters are ordered based on their statistical significance in error reduction. If the parameters are selected based on the order shown in Table 4.1, SCMs with different number of parameters can be made. For instance a 4-parameter set means that we use first 4 parameters from Table 4.1 which are *NSUBO*, *CFL*, *CTO* and *UO*.

Table 4.1: PSP parameter set to capture statistical variability

Parameter	Physical Meaning
NSUBO	Substrate doping
CFL	DIBL parameter
CTO	Interface states factor
UO	Zero field mobility
CSO	Coulomb scattering parameter
THESATO	Velocity saturation parameter
RSW1	Source/drain series resistance

4.3 Statistical Parameter Extraction

It is very important to be able to capture the simulated or measured statistical variability in statistical compact models since this is the only way to communicate this information to designers. Based on the discussion in the previous section on the sensitivity of model parameters, we identified the important parameters which are used to describe the impact of SV in compact model level. After identification of a statistical parameter set, we need to select an appropriate optimization method and parameter extraction strategy to commence statistical parameter extraction procedure.

4.3.1 Method and Strategy

In this work, a widely used optimization method - Levenberg-Marquardt (LM), which is essentially a gradient-based search algorithm, is used for PSP statistical parameter extractions. The LM method is implemented in standard non linear least-square routines [125], and many commercial MOSFET optimization and parameter extraction packages use this method for device parameter extraction [123,132]. Although some global optimization techniques like Genetic Algorithm (GA) [115] and Particle Swarm Optimization (PSO) [133,134] have been reported recently for PSP parameter extraction of a single device, they are less useful for the statistical parameter extraction task due to difficulties associated with algorithm implementation, low computational efficiency and non-convergence problems. Furthermore, no optimization algorithm can guarantee a real

global minimum in a multi-dimensional space trajectory and in addition to this fact, a good supply of initial conditions in gradient-based search algorithms can lead to the result close to a minimum error [125]. We will examine our statistical extractor with different initial conditions to see the trend of fitting error and its stability when selecting different statistical parameter sets.

The parameter extraction has been implemented with a Python [135] script. A nominal HSPICE 'model card' acts as the initial value of the statistical extraction procedure. The script introduces LM nonlinear least square optimization technique from the Python libraries [135] to fit selected parameters in the compact model by minimizing the RMS error calculated from HSPICE simulator. The target parameter list will be supplied into the script and the script will change just these parameters on each step starting from their nominal values. The iteration of calling simulator and changing the selected parameters with the least square optimizer will be terminated when the minimum RMS value is obtained.

The statistical parameter set contains both high sensitivity parameters (*NSUBO, CFL*) and low sensitivity parameters (*UO, CSO, CTO, THESATO, RSW1*). Hence, we have developed a two step parameter extraction strategy. In the first step, high sensitivity parameters will be extracted and in the next step, all other parameters will be extracted. On the other hand, parameters can be categorized as high field and low field parameters. High field parameters (*CFL, THESATO*) need to be extracted from I_d-V_g characteristics at high drain bias (V_d=1v) while other parameters will be target to be extracted from I_d-V_g characteristics at low drain bias (V_d=50mv).

4.3.2 Impact of Initial Conditions on the Accuracy

As it was pointed out in previous section, the LM method has been used in our parameter extractions due to its simplicity and computational efficiency. This method converges towards a minimum RMS error but cannot distinguish between a local and global minimum. In a multidimensional optimization application such as our statistical parameter extraction, the convergence trajectory depends on the initial value of the parameters. The uniform device full PSP parameter set is a natural initial guess for statistical parameter extraction but it is not necessarily the best initial guess. In practice, the

best initial guess is not known, therefore, we decided to use another base parameter set and then repeat the statistical parameter extraction for different number of parameters to ensure the stability of optimization method in terms of initial conditions.

The first step to create a reasonable initial set of parameter values for the statistical extraction is to select one of suitable 'atomistic' devices for this purpose. The devices which have I_{on} and I_{off} in the middle of I_d-V_g characteristics of Figure 4.2 are good choices because they provide the minimum current deviations for the other devices in the statistical set. To select a device which has drain current close to the median of I_{on} and I_{off} distributions, first we find the median value of each distribution and then use the I_{on}-I_{off} scatter plots to identify devices around the median point. Three devices were detected with I_{on} and I_{off} in a range between 0.95 and 1.05 times of the median values (i.e. as close as 5% to median values) for both high and low drain bias conditions. We call each one of these three devices a 'median' device. A 'median' device is an 'atomistic' simulated device with particular sources of variability while for the uniform device, no variability was assumed. Figure 4.6 illustrates the method used to identify the median devices. The uniform device lies on the edge of I_{on}-I_{off} scatter plot pattern. Figure 4.7 illustrates one of the median device characteristics among 1000 sample of atomistic simulations and compares it with uniform device characteristics.

The next step is the simulation of selected median devices to extend device characteristics to different bulk bias points. This is necessary to have an accurate device characteristic set before proceeding to full PSP parameter extraction. We used our atomistic simulator for this purpose to create two sets of I_d-V_g at low drain and high drain bias points for different bulk bias voltages equal to 0, -0.3, -0.7 and -1V and one set of I_d-V_d points for different gate voltages equal to 0.4, 0.6, 0.8 and 1V. The simulated data is used for single device DC parameter extraction as was discussed in detail in chapter 3.

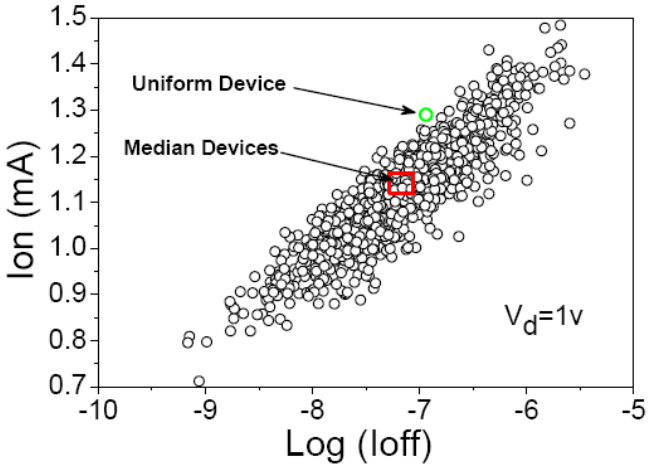


Figure 4.6: I_{on}-I_{off} scatter plot to identify devices with median drive and leakage current.

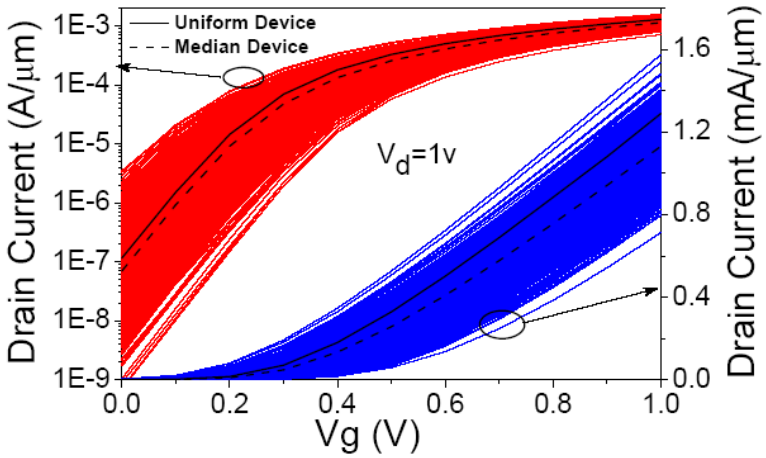


Figure 4.7: Comparison of median and uniform devices in 1000 sample of I_d-V_g characteristics.

In the last step we applied the proposed statistical parameter extraction strategy as discussed in section 4.3.1 to obtain the statistics of RMS error. Figure 4.8 compares the mean value and standard deviations of RMS error achieved in the statistical parameter extractions between two initial conditions; one based on a median device and the other one based on the uniform device. Although starting from a median device can produce an average error of 16.6% with one parameter which is considerably less than 21.3% when the starting point is a uniform device, the difference between error mean values of two methods is negligible when the number of parameters involved in statistical extraction is 4 or more. The difference in the standard deviation of the error between two approaches remains less than 1.5% for all of parameter sets.

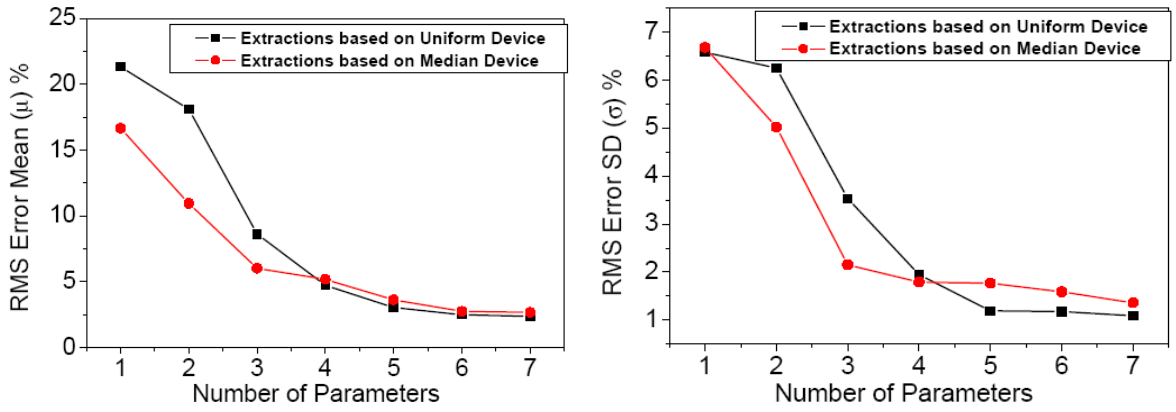


Figure 4.8: Comparison of RMS error mean (left) and standard deviation (right) versus different number of parameters in statistical parameter extraction based on uniform and median device.

The mean and standard deviation of RMS error for full 7- parameter set extractions based on the uniform device are equal to 2.36% and 1.39%, respectively. These values are slightly lower than the corresponding values for the extractions based on a median device which are equal to 2.66% and 1.36%. For this reason we will use the uniform device to set the initial conditions prior to statistical extractions. Figure 4.9 presents the distribution of RMS fitting error for different number of parameters for each statistical parameter set. The mean and standard deviation of each error distribution corresponds to the values shown in Figure 4.8 based on the uniform device initial conditions. The present discussion also proves the stability of our optimization method against using different initial conditions for statistical sets involving more than 4 parameters.

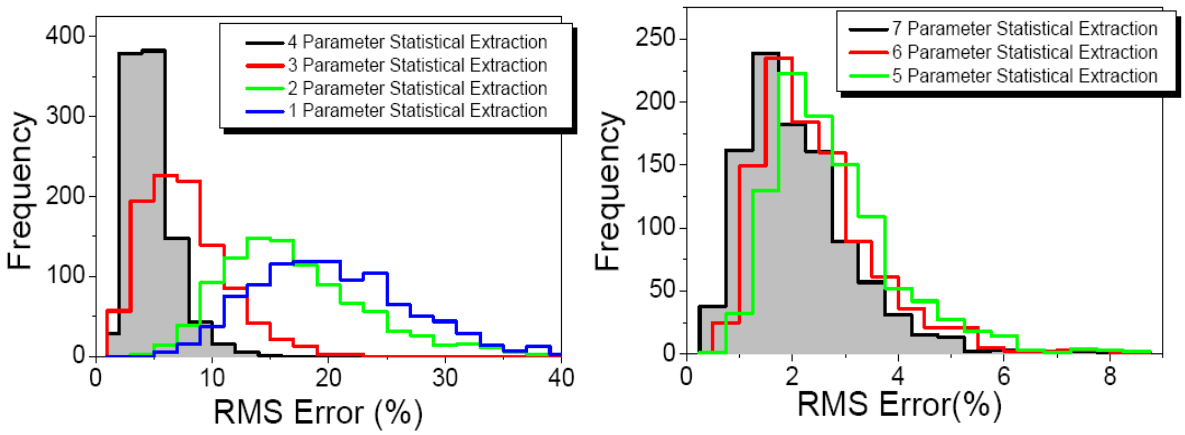


Figure 4.9: The impact of parameter set size on the RMS error of statistical parameter extraction.

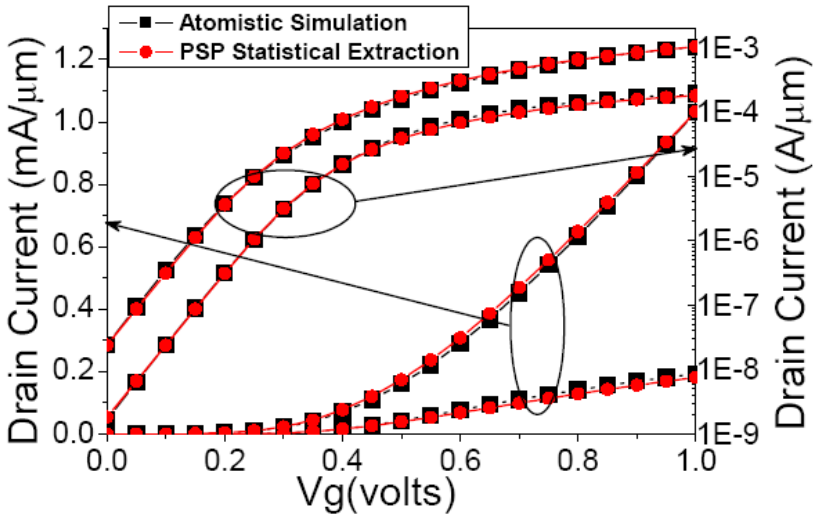


Figure 4.10: Statistical PSP parameter extraction for a typical device at two drain bias voltages with RMS error of 3%.

Figure 4.10 illustrates the result of statistical parameter extraction for a typical device chosen from 1000 devices in Fig. 4.2, with RMS error of 3%. Both linear and logarithmic scales at two drain bias voltages of V_d=50mv and V_d=1.0v have been used to demonstrate accuracy of fitting in sub-threshold and above-threshold regions.

4.3.3 Correlation between SCM and Electrical Parameters

It is important to investigate the correlation between CM parameters and physical properties of the device. A strong correlation indicates that the physical meaning of the compact model parameter is maintained during statistical extraction. Moreover, such correlation can provide guidelines for the different techniques to generate statistical CM sets based on the distribution of the device characteristics figures of merit.

Figure 4.11 illustrates the correlation between three key electrical and statistical compact model parameters. The extracted parameter *NSUBO* is statistically correlated with device threshold voltage (V_{th}) extracted from atomistic simulations. The parameter *CFL* is correlated with calculated DIBL parameter based on atomistic simulations and parameter CTO shows a direct correlation with sub-threshold slope (SS) of atomistic simulation results. The measured correlation coefficient between each parameter and electrical 'figures of merit' are shown on each graph.

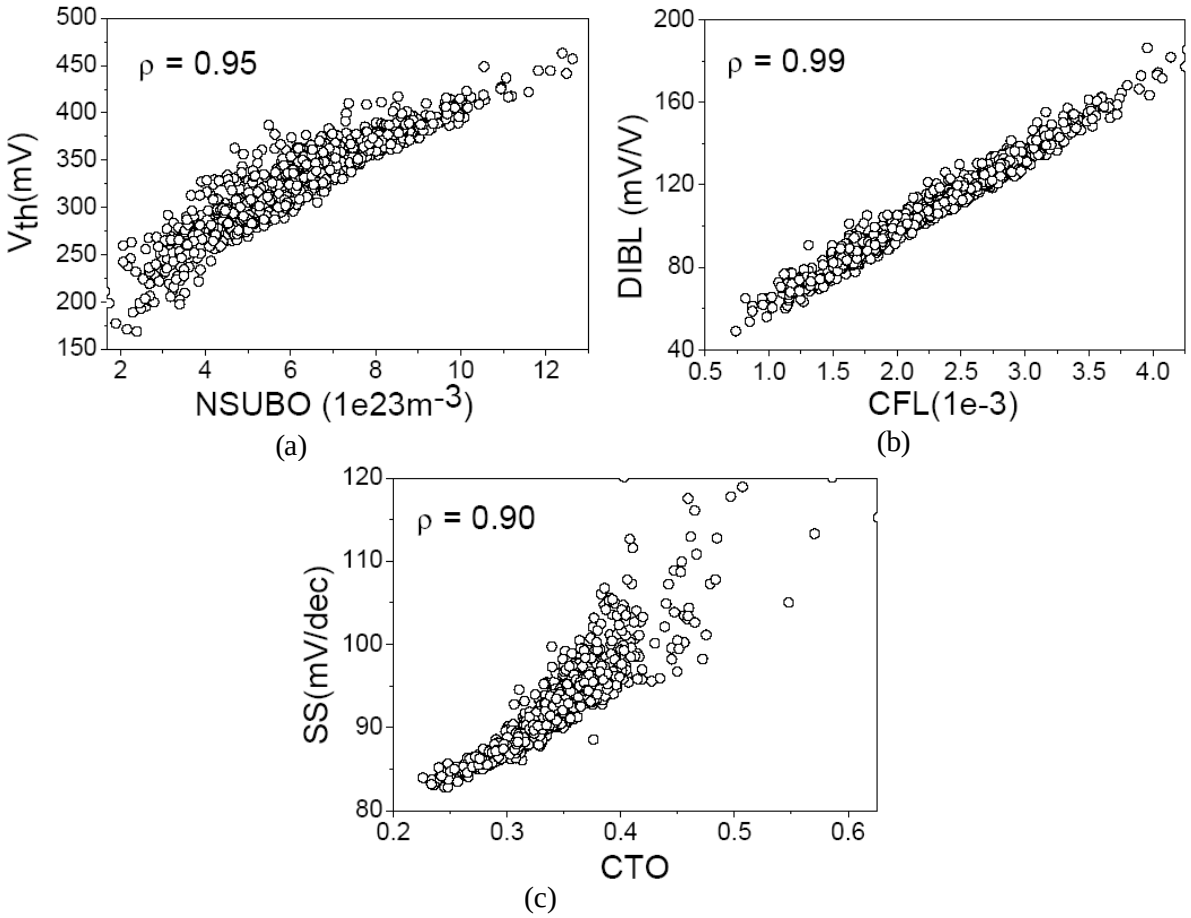


Figure 4.11: The correlation between typical electrical parameters and PSP statistical parameters: (a)- V_{th} and *NSUBO*, (b)- DIBL and *CFL*, (c)- Sub-threshold Slope and *CTO*.

4.3.4 Playback of MOSFET 'Figures of Merit'

Five key electrical parameters of a MOSFET have been considered to evaluate the accuracy of statistical compact model set in respect to 'atomistic' simulation results. These 'figures of merit' are the threshold voltage (V_{th}), drive current (I_{on}), leakage current (I_{off}), DIBL parameter and sub-threshold slope (SS). The definitions of these figures of merit were presented in chapter 3 for uniform MOSFET and we used the same definitions to calculate corresponding values from simulated I_d-V_g characteristics of microscopically different MOSFETs selected from statistical compact model library set. This 'Statistical Compact Model' (SCM) library has 1000 samples of HSPICE model cards, each one different in just 7-parameters from the others. Figure 4.12 shows the histogram plots of I_{on}, I_{off} and V_{th} as three typical electrical 'figures of merit', extracted from SCM simulations and compares them with atomistic simulation results which were presented in Figure 4.3. Normal distribution plots are fitted on each distribution. It clearly demonstrates

reproduction of 'figures of merit' distributions using SCM with high accuracy. The absolute values and relative errors associated with the mean values (μ) and the standard deviations (σ) of device 'figures of merit' are calculated and shown in Table 4.2 for both low drain (V_d=50mv) and high drain (V_d=1.0v) bias conditions. The produced error in the mean values of SCM simulations are less than 3% in respect to atomistic results. The standard deviation errors remain also less than 4% for all of electrical parameters except sub-threshold slope parameter (SS) which exhibits an upper margin of around 20%. This is due to calculation method of SS which creates a cumulative error caused by errors in both first and second point of I_d-V_g simulated curves.

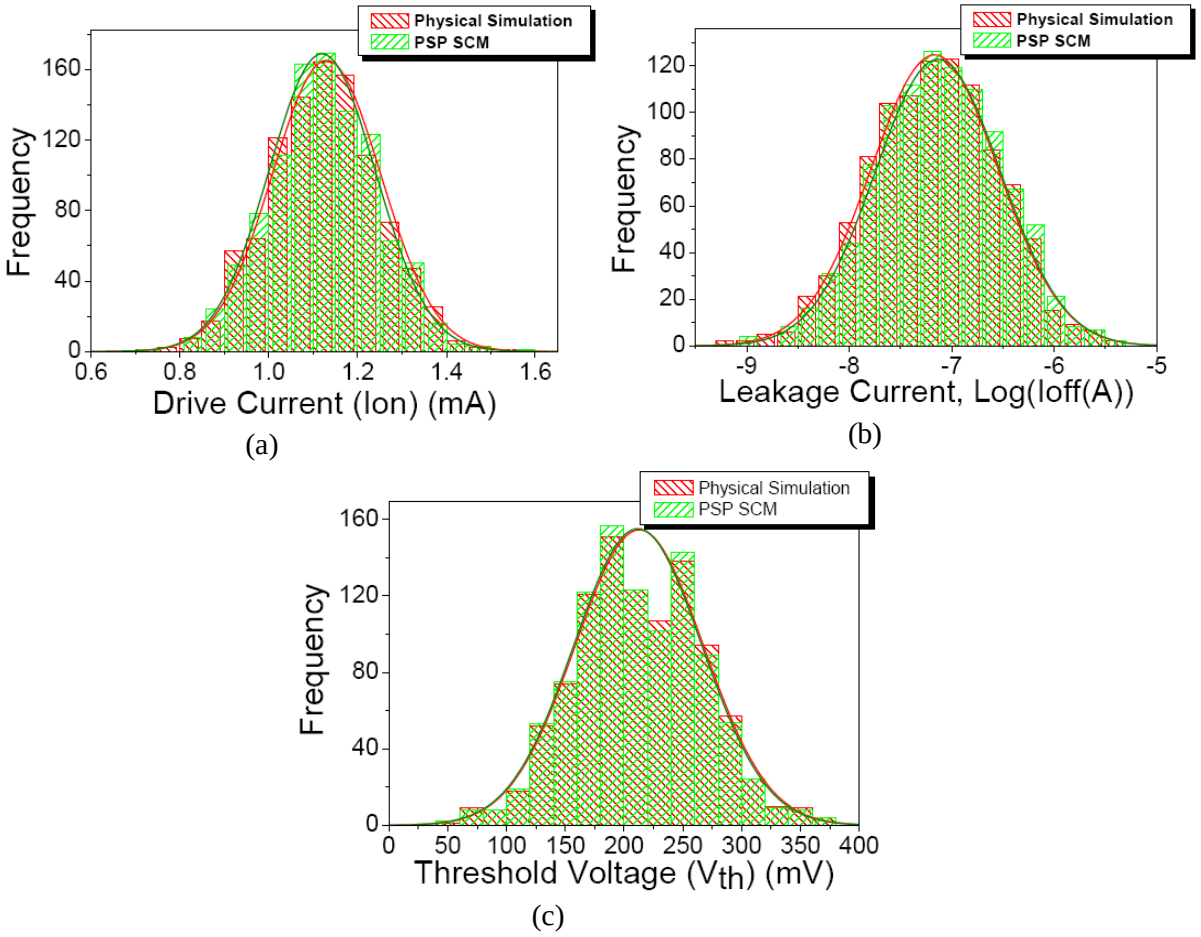


Figure 4.12: Distribution of (a)-I_{on} , (b)- I_{off} and (c)- V_{th} compared between physical simulations and statistical PSP compact model with fitted normal distributions on them, V_d=1.0v.

Table 4.2: Comparison of statistical properties of device electrical 'figures of merit' between SCM results and atomistic simulations

Figures of merit	**Drain Bias (V)**	**Atomistic Simulations**		**SCM Simulations**		**Absolute Relative Error (%)**	
		μ	σ	μ	σ	μ	σ
Log (I_{off}(A))	1.0	-7.17	0.62	-7.13	0.61	0.5	1.6
	0.05	-8.17	0.55	-8.13	0.55	0.5	0
I_{on} (mA)	1.0	1.125	0.1225	1.119	0.1197	0.5	2.3
	0.05	0.192	0.0155	0.196	0.0161	2.1	3.9
V_{th} (mV)	1.0	213.0	54.7	211.7	54.4	0.6	0.5
	0.05	320.4	49.1	317.5	48.2	0.9	1.8
DIBL (mV/V)	-	97.6	23.1	96.5	22.4	1.1	3.0
SS (mV/dec)	1.0	91.99	5.85	93.62	5.38	1.8	8.0
	0.05	89.78	2.55	91.82	3.10	2.3	21.5

Since the MOSFET 'figures of merit' are inter-correlated and they are nonlinear functions of the statistical parameter sets, it will be important to investigate the playback of figures of merit correlation obtained from SCM simulation and compare it with physical 'atomistic' simulations. Figure 4.13 illustrates this playback with scatter plots between SCM simulations results and atomistic figures of merit. It clearly shows that direct extracted compact models based on the 7-parameter set can regenerate the correlations of device 'figures of merit' with high accuracy which is consistent with small errors observed in Table 4.2. High degree of correlation between V_{th} and I_{off} is caused by log-linear relationship in sub-threshold region.

However, although increasing the size of statistical parameter set can produce better accuracy results for particular device geometry, it is usually difficult to develop corresponding scaling strategy due to the empirical nature associated with the large set size. In order to evaluate the impact of size of parameter selection set on the accuracy of reproduction of device 'figures of merit' statistical behaviors, we use statistical compact models based on different parameter sets to generate I_d-V_g characteristics and calculate the corresponding statistical trend of device electrical parameters.

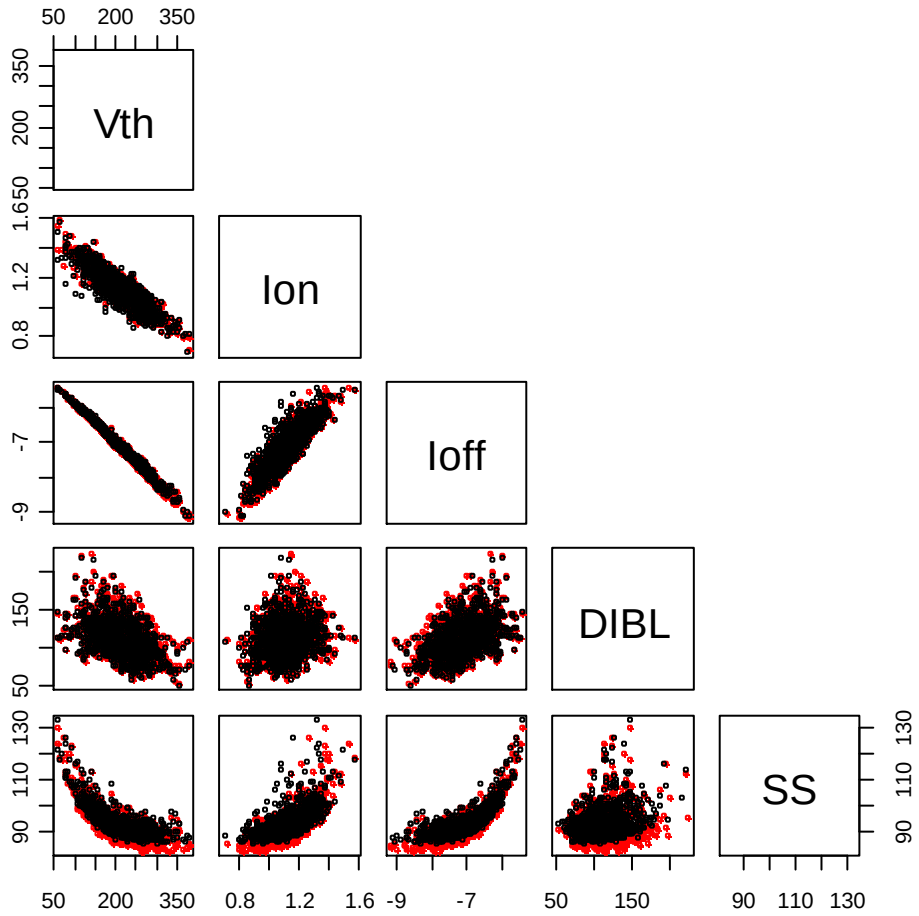


Figure 4.13: Playback of MOSFET 'figures of merit'. Black circles: physical atomistic device simulation, Red circles: direct statistical PSP compact model simulation (V_d=1.0v), Dimensions: V_{th}(mV), I_{on}(mA), Log(I_{off}(A)), DIBL(mV/V), SS(mV/dec).

Figure 4.14 shows the mean and standard deviations of the leakage current (I_{off}) versus different parameter sets in PSP statistical compact model with the absolute relative errors in respect to the original 'atomistic' results. Selection of 4-parameter set results in the settlement of the error trends. This 4-parameter set gives approximately 0.6% error in the statistical mean value of Log(I_{off}) and 1.5% error in the standard deviations (SD) of Log(I_{off}). Increasing the number of parameters in statistical set to more than 4 parameters does not lead to further error reduction in the mean or SD of Log(I_{off}). This result is consistent with role of parameters added to increase the SCMs to more than 4 parameters. The parameters *CSO*, *THESATO* and *RSW1* have weak influence on subthreshold region of the device characteristics and hence have little impact on I_{off}.

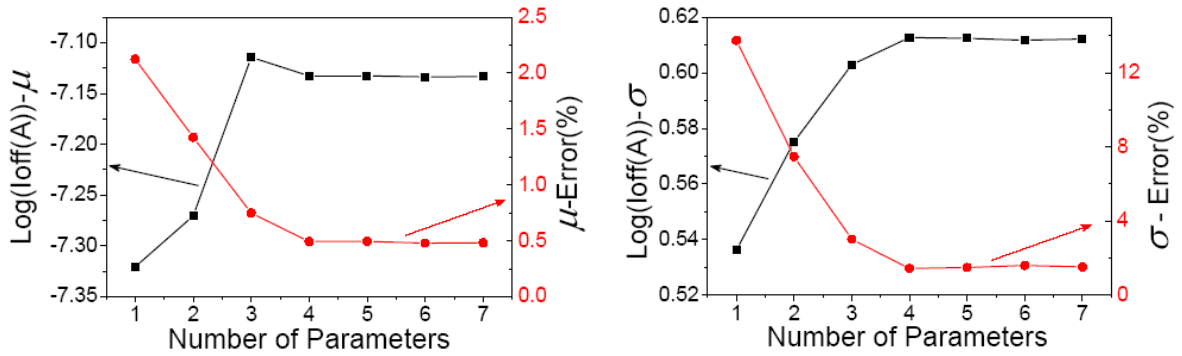


Figure 4.14: Impact of parameter set selection on the statistical mean (left) and standard deviations (right) of MOSFET leakage current with their absolute relative errors in respect to corresponding values in atomistic simulations (V_d=1.0 v).

Figure 4.15 shows the similar statistical trend for the transistor drive current (I_{on}). It shows a non monotonic trend but selection of 6 parameters settles the trend of mean value of I_{on}. A five parameter set slightly underestimates the mean value of I_{on} and leads to 1% increase in the mean error obtained by either 6 or 7 parameter set. However, in the standard deviation of I_{on}, selection of 5, 6 or 7 parameters gives close errors, equal to 2.6%, 2.45% and 2.3%, respectively. Since the direct extraction targets are the full set of I_d-V_g characteristics, the overall monotonic reduction of the total RMS error with the increasing of the parameter-set size, demonstrated in Figures 4.8 and 4.9 does not guarantee for the monotonic error reduction for a particular device figure of merit.

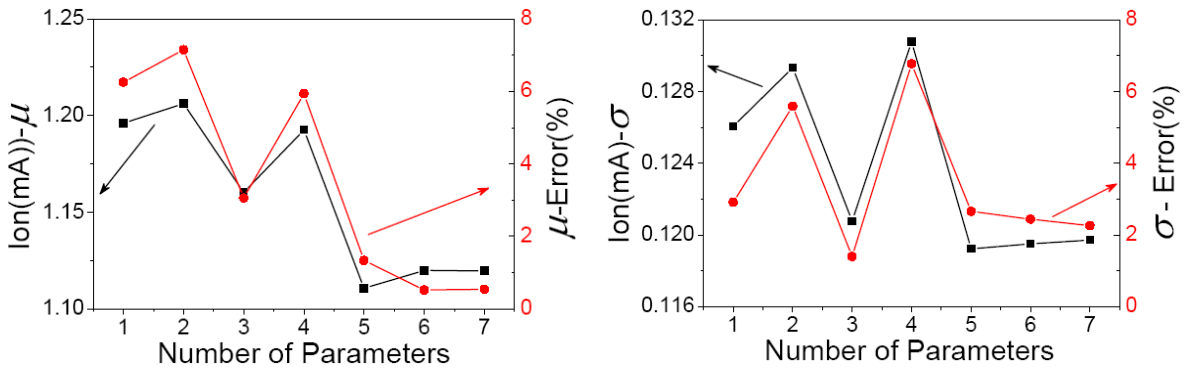


Figure 4.15: Impact of parameter set selection on the statistical mean (left) and standard deviations (right) of MOSFET drive current with their absolute relative errors in respect to corresponding values in atomistic simulations (V_d=1.0 v).

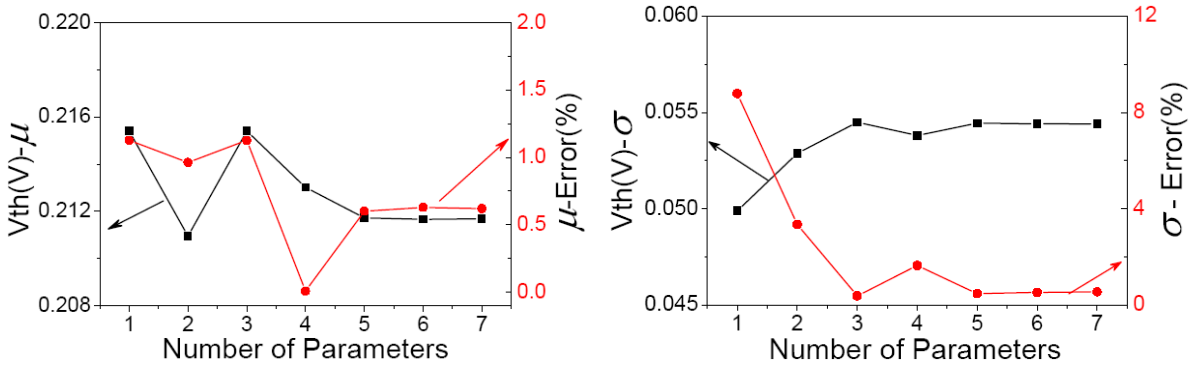


Figure 4.16: Impact of parameter set selection on the statistical mean (left) and standard deviations (right) of MOSFET threshold voltage with their absolute relative errors in respect to corresponding values in atomistic simulations (V_d=1.0 v).

Figure 4.16 shows the impact of selecting different number of parameters in SCM on the threshold voltage, V_{th}. The trend for the mean has always an error less than 1.2% which is negligible but for the standard deviation, it is settled for 5-parameter set with 0.5% error in respect to physical simulation results. Increasing the number of parameters to more than five does not affect on the accuracy of mean value or standard deviations of the threshold voltage. The flat behavior of mean and standard deviation is better understood if we consider the parameters added to the associated SCMs. The parameter *THESATO* has been added to make 6-parameter SCM and *RSW1* has been added to make a 7-parameter SCM. Both parameters have influence on above-threshold region of device characteristics with little impact on the threshold voltage.

4.4 Statistical Circuit Simulation

Statistical Variability (SV) becomes the driving force behind the fundamental shift of circuit design paradigm from deterministic to parametric approach. Thus, it is important to study the impact of SV in the circuit level. We have selected the basic building block of digital circuits, a CMOS inverter, as a test bed circuit to carry out a comprehensive study of SV using our statistical compact model (SCM) library set. Because of random nature of statistical variability, HSPICE circuit simulation is carried out by using Monte Carlo approach in order to accurately assess the circuit performance. In this approach, model cards are selected randomly for both n and p-channel MOSFETs from previously built

SCM library sets using a random number generator. The CMOS inverter which has been used as a test bed in this work has minimum device size of *W/L* equal to 35nm/35nm for NMOSFET and *W/L* of 70nm/35nm for PMOSFET in order to highlight the variability trends. The supply voltage is 1 volt. Wider PMOSFET is necessary to balance the output transition delay times of the inverter. Since SCM library for PMOS is limited to the width of 35nm, we have implemented wider PMOS in CMOS inverter with parallel combination of two randomly selected basic width devices. Figure 4.17 illustrates the inverter schematic diagram and its implementation for statistical simulation. The indices i, j and k have been generated with a uniform random number generator in the range 1 to 1000. Figure 4.18(a) shows the distribution for one of typical indices while Figure 4.18(b) proves no correlation between different indices. However, there is small fluctuation in the distribution of indices and non-zero correlation between different indices due to finite sample size.

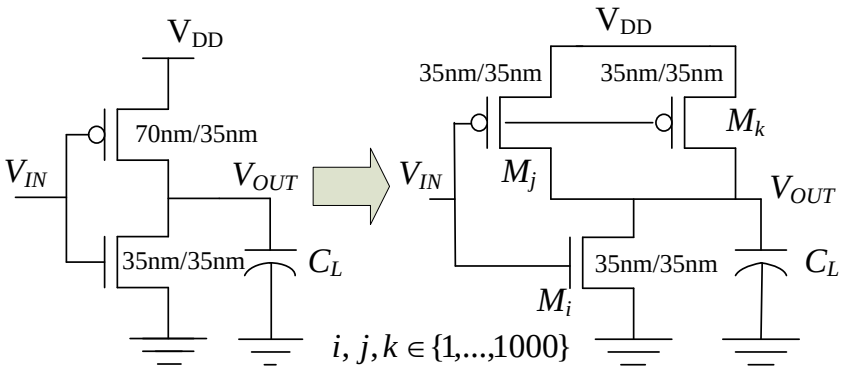


Figure 4.17: Schematic of 35nm CMOS inverter used as a test bed for statistical circuit simulations.

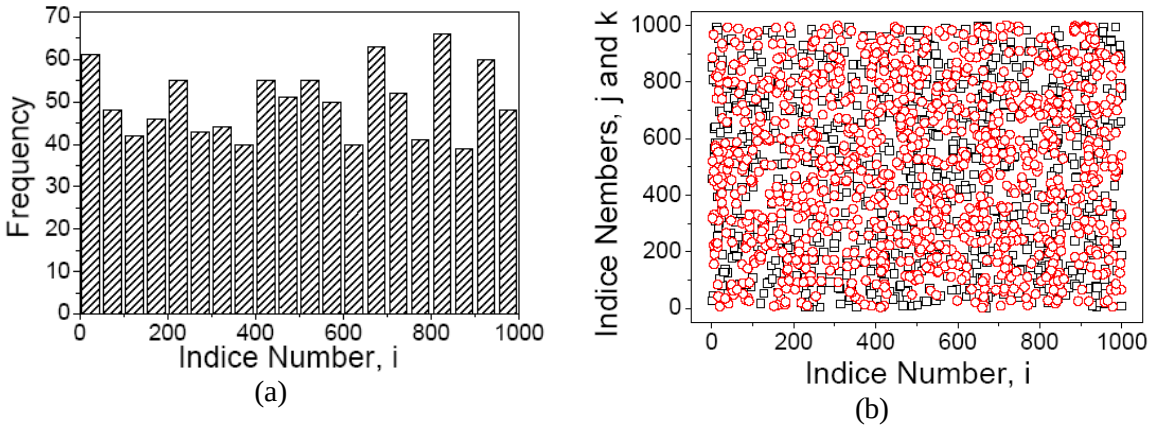


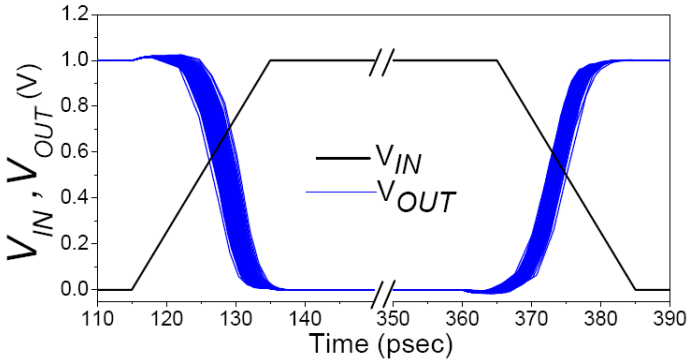
Figure 4.18: (a)-Distribution of one typical indice number, (b)-Scatter plots between different indice numbers; Red symbols represent j versus i while black symbols are k versus i plot.

The output voltage and the supply current of the inverter with no-load conditions and input rise/fall-time of 20ps in a 2GHz pulse simulated by using the Monte Carlo approach are illustrated in Figure 4.19. It demonstrates that the relatively larger size PMOS reduces the variation in the output voltage rise time.

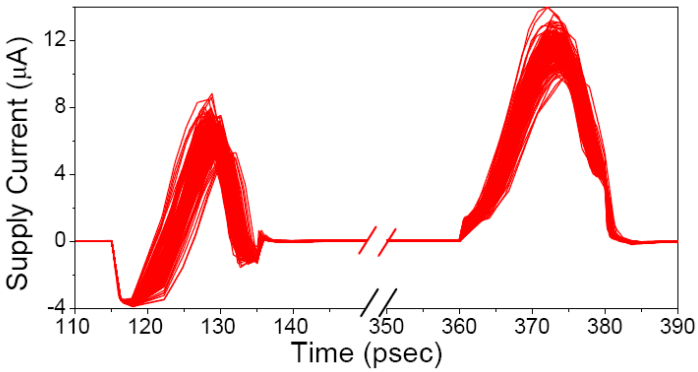
The 'figures of merit' for timing and power analysis of an inverter are rise time propagation delay (t_{dLH}), fall time propagation delay (t_{dHL}) and dissipated energy (E_{diss}). The definitions of t_{dLH} and t_{dHL} are illustrated in Figure 4.20. The dissipated energy is given by:

$$E_{diss} = \int_0^T V_{DD} I_{DD} dt \tag{4.2}$$

where V_{DD} is the power supply voltage and I_{DD} is the power supply current.



(a)



(b)

Figure 4.19: Transient time statistical simulation of 35nm CMOS inverter for an ensemble of 1000 samples in no-load conditions and 20ps input rise/fall time; (a) Output voltage, (b) Supply current.

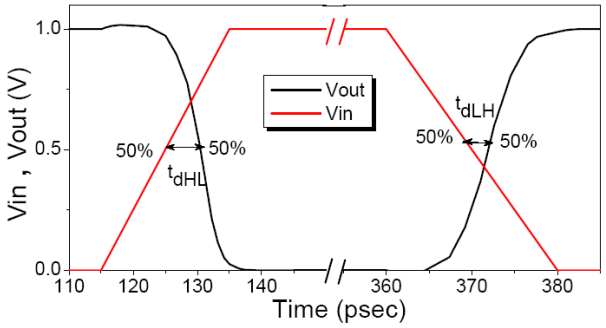


Figure 4.20: Definitions of t_{dLH} and t_{dHL}.

In this section, we will use statistical circuit simulations to evaluate impact of number of parameters in SCMs on the inverter 'figures of merit' and then as a case study, we will study impact of input-output specifications on the variability trends of inverter 'figures of merit' using the most accurate 7-parameter SCM.

4.4.1 Impact of Parameter Set Size

To investigate the impact of the number of parameters in each SCM set on the accuracy of statistical circuit simulation, we have carried out a series of statistical timing and power simulations on test bed CMOS inverter. No load output capacitance and input rise/fall time of 50ps has been considered for all of simulations to amplify the impact of statistical variability on the inverter delay and dissipated energy. Monte Carlo HSPICE simulations for 1000 samples have been carried out for each statistical parameter set and the average and standard deviations were calculated to obtain the statistical trend of delay and dissipated energy against the number of parameters in SCM.

Figure 4.21 illustrates the trend of mean and standard deviation of dissipated energy versus number of parameters in each SCM set. The errors have been obtained in reference to most accurate 7-parameter case. It is concluded that the trend has been settled for 5-parameter set.

Figure 4.22 shows the same trends for rise time propagation delay (t_{dLH}) and fall time propagation delay (t_{dHL}) of the inverter. The reduced variability of t_{dLH} compared to t_{dHL} which is seen from standard deviation plot, is due to larger width of PMOS compared to

NMOS. The errors for the mean value and standard deviation in reference to most accurate 7-parameter set are shown in Figure 4.23. It is seen that the mean and standard deviation of t_{dLH} has been settled before t_{dHL}. The mean errors are almost 1% for 5-parameter set while for standard deviation of 5-parameter set it is about 0.5% which is negligible.

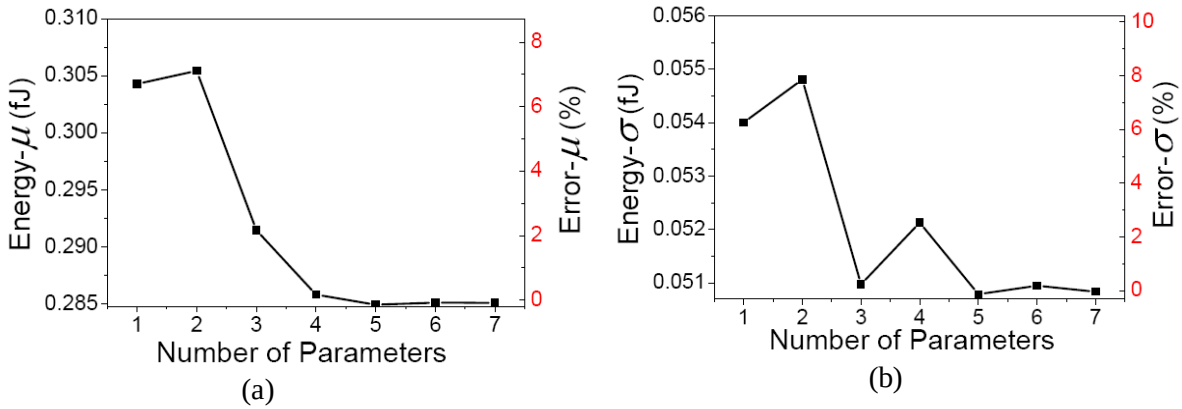


Figure 4.21: Impact of parameter set selection on the statistical mean (left) and standard deviations (right) of dissipated energy in inverter with their relative errors in reference to most accurate 7-parameter set.

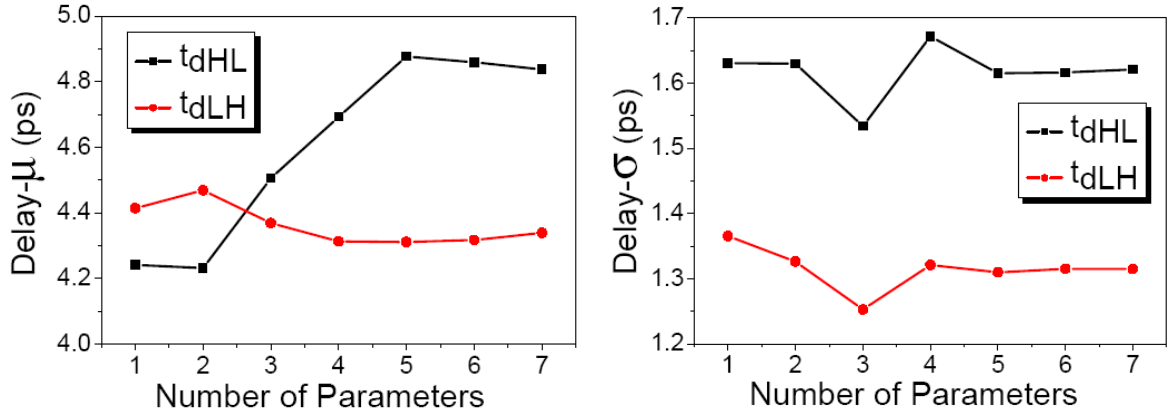


Figure 4.22: Impact of parameter set selection on the statistical mean (left) and standard deviation (right) of inverter delays.

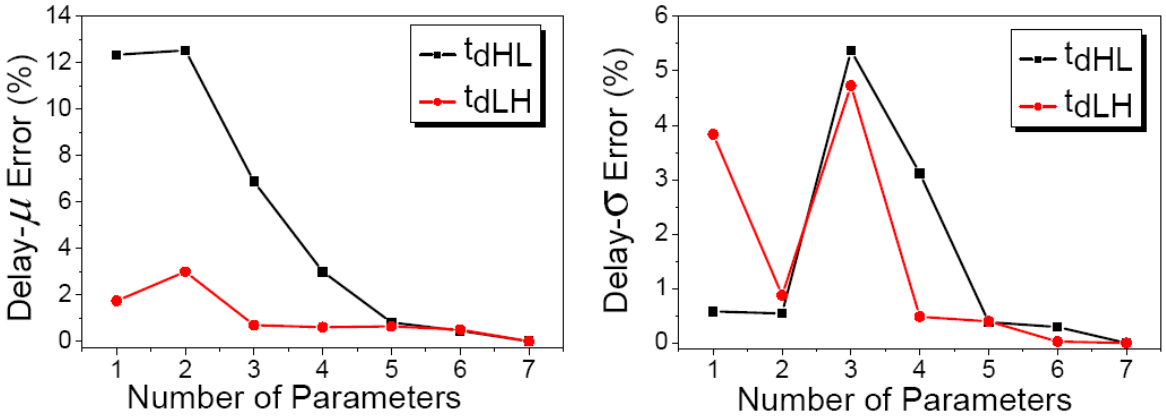


Figure 4.23: Absolute relative errors for inverter delay mean (left) and standard deviation (right) versus number of parameters in reference to most accurate 7-parameter set.

4.4.2 Impact of Input-Output Specifications

In modern CMOS digital circuit, not only the drive current but also the accurate shape of the full transistor I_d-V_d characteristic plays role during switching, and the actual current switch trajectory is affected by both input slew rate and the output load capacitance [136]. In this study, based on the accurate direct statistical compact modeling approach, the effects of the input waveform on variation of delay and power dissipation of an inverter at 35nm gate length MOSFETs are investigated in detail under various load conditions. The results can provide guidelines for reliable statistical standard cell characterization that is still a research hotspot.

The same inverter as demonstrated in Figure 4.17 has been considered in this study. The unit capacitive load of 0.105fF (equivalent to fan-out of 1 under minimum size configuration) is assumed and various load conditions equal to 0, 1, 2, 4 and 6 times of unit load are considered in this study. In order to explicitly demonstrate the effect of input slew rate on circuit performance variation, instead of generating input signal through an inverter chain, a 2 GHz ideal symmetrical clock pulse with various rise/fall times (2, 10, 20 and 50ps) is considered, as shown in Figure 4.24.

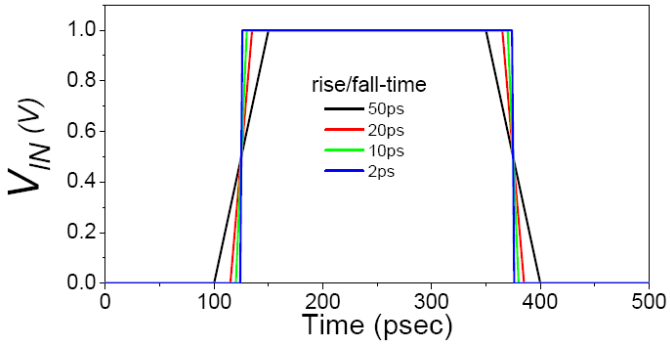


Figure 4.24: Inverter test input signals with various rise/fall times.

NMOS current trajectories under different input rise time conditions with the unit load and 6 times of unit load are illustrated in Figures 4.25(a). For rise time of 50ps, NMOS current never reaches the drive current for bias condition $V_g=V_{dd}$. It clearly demonstrates that the slope of input waveform has a direct impact on device bias conditions during the MOSFET switching. This in turn can strongly affect the circuit variability behavior in the presence of SV that is bias condition dependent. Increasing the load capacitance will improve the situation by pushing the switch trajectory up to the high gate bias regime, as demonstrated in Figure 4.25(b) for NMOS current trajectories for six times of the unit load at the output node. As a result, we expect that the circuit variability will be less affected by different input rise times when the load capacitance is increased.

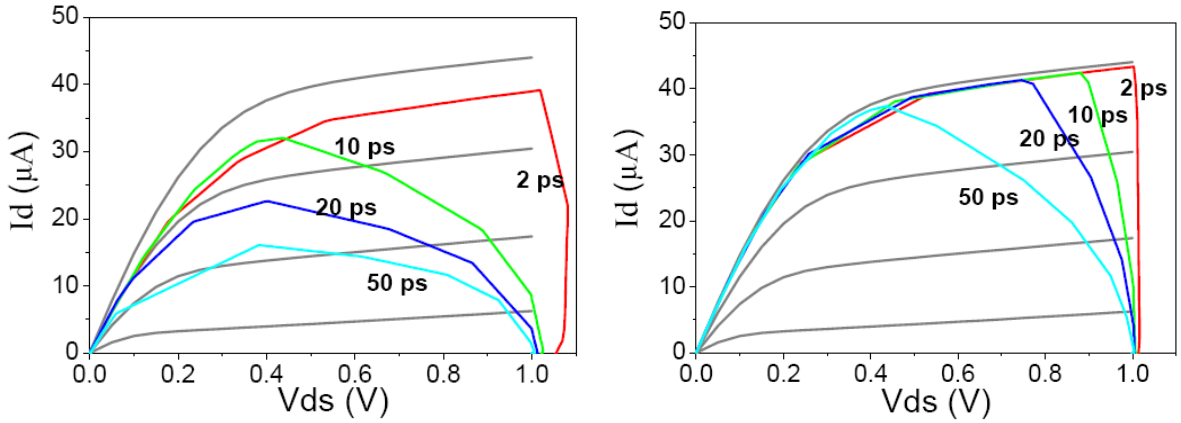
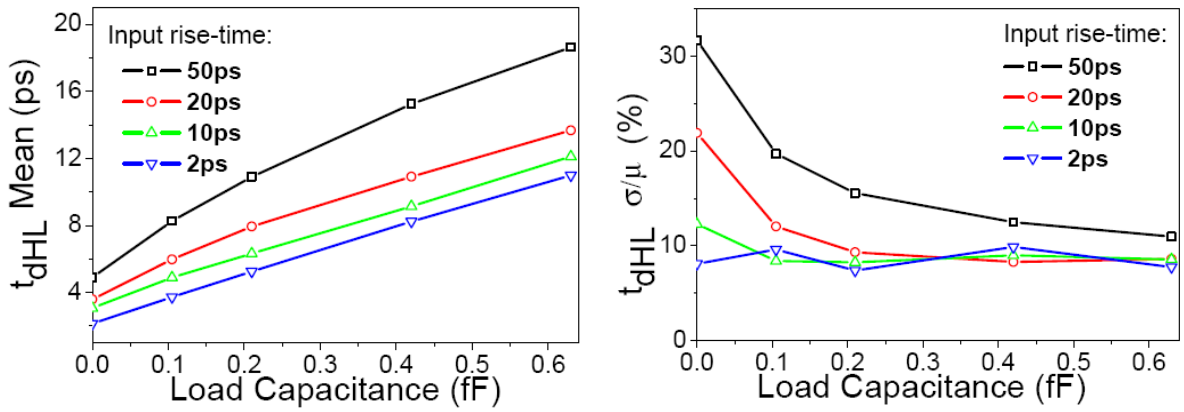


Figure 4.25: NMOS current trajectories under different input slope conditions; (a) for unit load, (b) for six times of unit load.

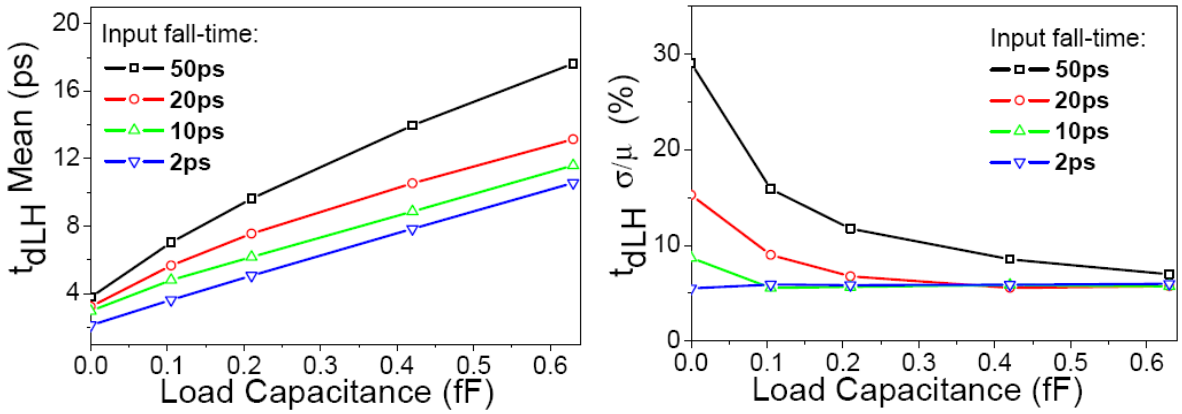
The mean and normalized standard deviations of t_{dHL} versus different loads under different input rise time slopes are shown in Figure 4.26. As expected, the mean values increase almost linearly with the increase in load capacitance. Also, the longer the rise time, the longer the delay since the transistor spends less time in strong inversion region during a switching event. Using load of 2 units as an example, the mean value of t_{dHL} increases from 5.3ps at input rise time of 2ps to 11ps at input rise time of 50ps. The normalized standard deviation (or coefficient of variation, σ/μ) is strongly modulated by input signal slope due to changes of device switching trajectory. For load of 2 units, σ/μ increases from 7.4% at input rise time of 2ps to 15.6% at input rise time of 50ps. Increasing the load capacitance will reduce the impact of input rise time on σ/μ and this is consistent to the previous discussion on Figure 4.25(b).

The statistical timing simulation results for t_{dLH} are presented in Figure 4.27. As expected, its behavior is similar to t_{dHL} counterpart but with reduced variability due to relatively large size of PMOSFET. Using load of 2 units as an example, the maximum normalized standard deviation of delay (σ/μ) is reduced from 15.6% in t_{dHL} to 11.8% in t_{dLH}. Figure 4.28 illustrates the level of impact that slope of input signal can impose on delay variation with histograms. The statistical inverter simulation has been carried out for the unit load of 0.105fF. The mean and standard deviations for each histogram can be read from Figure 4.27 for unit load.



(a) (b)

Figure 4.26: Results of statistical simulation for t_{dHL} of 35nm CMOS inverter; (a): mean value, (b): coefficient of variation.



(a) (b)

Figure 4.27: Results of statistical simulation for t_{dLH} of 35nm CMOS inverter; (a): mean value, (b): coefficient of variation.

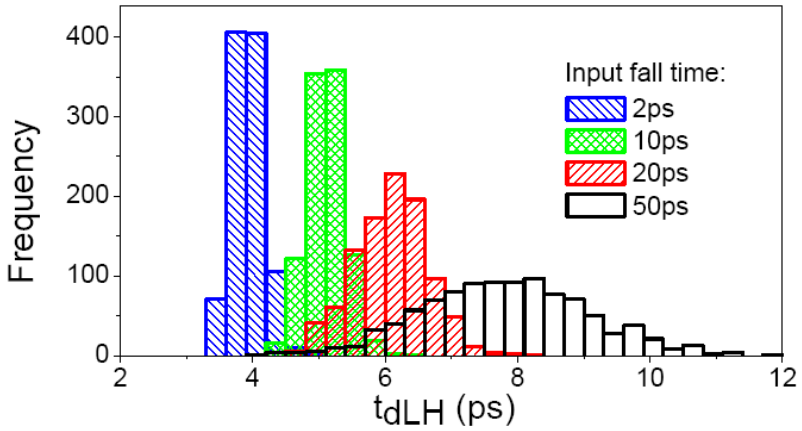
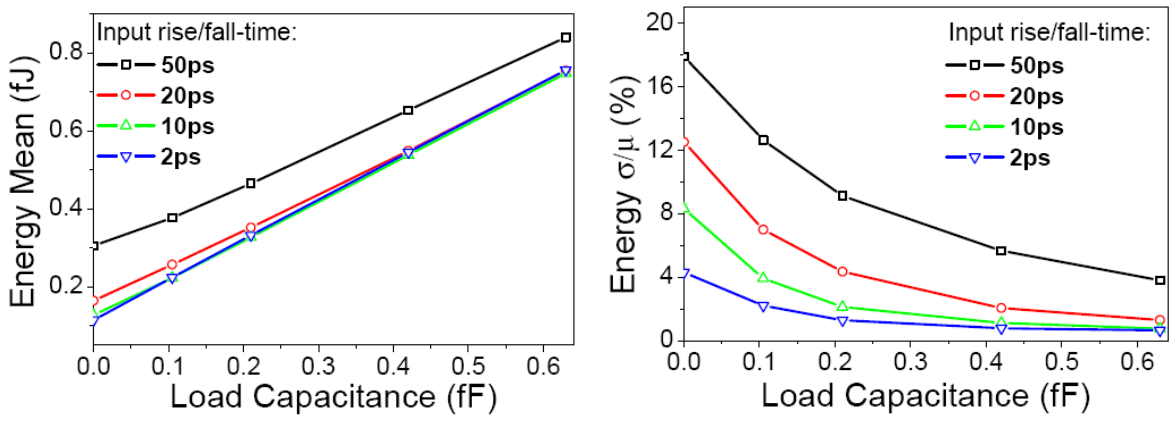


Figure 4.28: t_{dLH} distribution versus different input fall times for unit capacitive load.

The power dissipation in digital circuits can be separated into static and dynamic parts. Static dissipation is due to sub-threshold and gate leakage current flowing through the supply when devices are off. Dynamic dissipation depends on the size of the capacitive component of the load [104]. There is a third part in the power dissipation that is short circuit or crowbar power dissipation [137]. This component is caused by the existence of a DC path for the current flowing from supply to ground during the switching. Since input signal rise/fall time will strongly determine the amount of time that inverter can stay at short circuit status, it'll have a dramatic impact on power dissipation variation. Figure 4.29 illustrates the results of average and normalized standard deviation of dissipated energy (E_{diss}) during a full input cycle. Variation in the energy dissipation at this particular case is introduced by short-circuit component since leakage power dissipation is negligible. Apart

from input rise/fall time of 50ps, the average energy dissipation μ is practically the same when load of unit is larger than 2, which indicates that for most load conditions, at least up to input rise/fall time of 20ps, dynamic power dissipation dominates overwhelmingly. As expected, the maximum σ/μ of energy dissipation happens at intrinsic load (no-load) case. For a typical load of 2 units case, the σ/μ increases from 1.3% at input rise/fall time of 2ps to 9% at input rise/fall time of 50ps.

Based on the definition of inverter delay which is given by $t_d = \max\{t_{dHL}, t_{dLH}\}$, there is a correlation between delay and energy dissipation. This correlation depends on load condition and input slew rate. Figure 4.30 shows the scatter plots between delay and energy dissipation at input slew rate of 1V/2ps and 1V/50ps respectively for load of 6 units, which clearly demonstrates input slew rate will increase the correlation between delay and energy for this case. Considering the standard definition of correlation coefficient, we can calculate this coefficient under various load conditions and input slew rates. Figure 4.31 summarizes the dependence of delay-energy correlation on input slew rate and output load, which indicates in standard cell characterization, we cannot treat timing and power dissipation as statistically independent 'figures of merit'.



(a) (b)

Figure 4.29: Results of statistical simulation for dissipated energy of 35nm CMOS inverter; (a): mean value, (b): coefficient of variation.

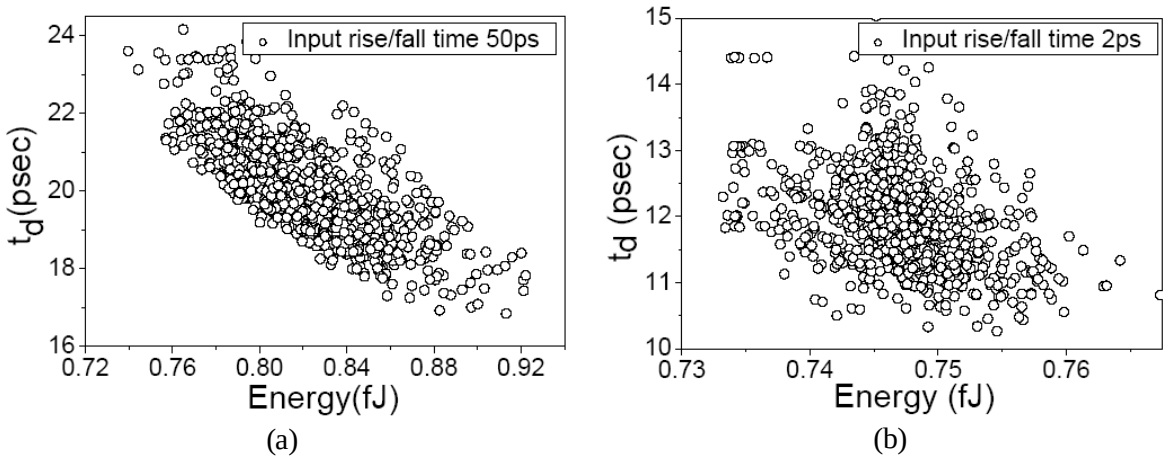


Figure 4.30: Scatter plots between delay and energy dissipation for load of 6 units; (a): input slew rate 1V/50ps, (b): input slew rate 1V/2ps.

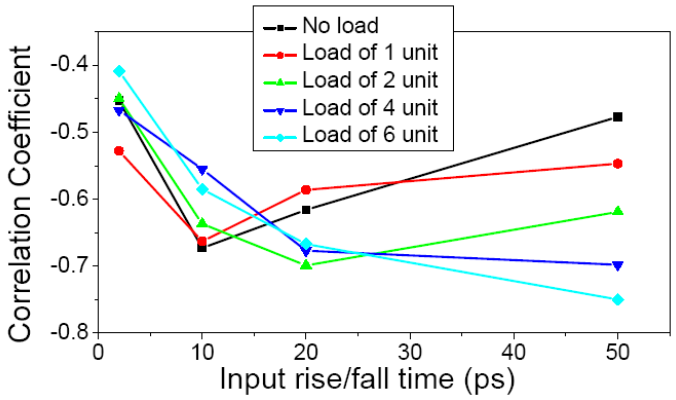


Figure 4.31: Correlation coefficient between delay and energy dissipation against input slew rates for various loads.

4.5 Summary

In this chapter, we carried out 'atomistic' simulation of statistical variability to account for the combined effects of RDD, LER and PGG in 35nm physical gate length MOSFETs. Then, based on the sensitivity analysis of PSP compact model parameters, we identified 7 parameters which are able to translate variability from device level to circuit level.

Since the reliable variability-aware circuit and system design requires the support of statistical compact model (SCM), we proposed a method and strategy to directly extract target parameters in PSP compact model from 'atomistic' simulation current-voltage characteristics. Extracted parameters were used to make SCM library set. The main advantage of this direct approach is that is does not need any pre-assumption about the parameter distributions. We have also evaluated the accuracy of the parameter extraction method by changing the initial conditions from uniform to a median device and observed stability of method in settlement of mean and standard deviation of RMS error particularly for SCMs containing 5 parameters or more.

The accuracy in the reproduction of the MOSFET 'figures of merit' was assessed in comparison with 'atomistic' simulation results. Both histograms and scatter plots were used to compare I_{on}, I_{off} and V_{th} extracted from the SCM with their counterparts obtained from gold standard 'atomistic' simulation results. Moreover, we investigated the accuracy in the distributions of these 'figures of merit' extracted for different number of statistical compact model parameters. The results show that the minimum required numbers of parameters which are needed to reproduce mean and standard deviations of I_{on}, I_{off} and V_{th} are 7, 4 and 5 parameters, respectively.

In the last section of this chapter, based on an accurate direct statistical compact modeling approach, we simulated a CMOS inverter as a digital circuit building block to obtain the distribution of delay and dissipated energy in different input-output conditions and also using different number of statistical compact model parameters. Our statistical simulations for a most accurate 7 parameter set showed that the normalized delay variability σ/μ can increase by two times, and σ/μ on energy dissipation can increase by more than 7 times, for a fan-out of 2 when input rise/fall time increases from 2ps to 50ps. Furthermore, the degree of correlations between delay and energy depends on both output load and input slew rate.

Chapter 5

Statistical Parameter Generation Techniques

In the previous chapter, we described our Statistical Compact Model (SCM) extraction methodology. Using the ensemble of models generated by the direct method, a SCM library can be constructed and devices in circuits can be selected from the library for the purpose of statistical circuit simulation. Although this is the most rigorous method of performing statistical circuit simulation, the available statistical sample size is pre-determined by the size of the SCM library. Common practice in Monte Carlo circuit simulation is to generate statistical parameter values on the fly. Since the accuracy of circuit simulation results is determined by the accuracy of the transistor compact models, it is critically important for the designer to understand the limitations of different statistical compact model parameter generation techniques.

In this chapter, we first review the statistical properties of the parameters obtained from direct extraction method. The mean and standard deviation of each parameter distribution and the correlation between parameters will be investigated. In the second and third sections, two statistical parameter generation techniques based on independent Gaussian distributions of the individual parameters and employing principal component analysis (PCA), are introduced. The accuracy of each parameter generation technique in reconstructing MOSFET electrical performance distribution will be examined. In the last section, the context of Non-linear Power Method (NPM) will be elaborated as an advanced parameter generation technique that is capable of taking high moments of statistical distribution into account. The accuracy of this method will be compared with independent Gaussian and PCA methods.

5.1 Statistical Analysis of SCM Parameters

Following a direct parameter extraction strategy discussed in previous chapter, we have obtained a 7-parameter statistical set for an ensemble of 1000 microscopically different n-channel MOSFETs of 35nm gate length, with an average RMS error of 2.3%. The impact of sample size on the variance of parameters is negligible with the choice of more than 500 samples in nano-CMOS technology [75]. In this section we study statistical properties of directly extracted parameter set including individual parameter distribution and correlations between pairs of parameters. The results form a basis to different proposed parameter generation techniques.

5.1.1 Parameter Distributions

Figure 5.1 illustrates the statistical distribution of each parameter with its measured mean (μ) and standard deviation (σ) in the full 7-parameter set. Every normal distribution can be characterized using its mean and standard deviation. Since the normal distribution of parameters is a basic assumption in the existing SCM approaches discussed in background chapter, it is important to examine this assumption for real parameter distributions at nanometer regime. Shapiro-Wilk [138], Kolmogorov-Smirnov [139] and chi-square [140] tests are examples of standard methods for the statistical 'hypothesis test' which has been introduced by statisticians to compare a distribution with any hypothesized distribution like Gaussian distribution. In addition, normal probability plots or Q-Q plots can be used to assess normal distribution of parameters by visual inspection [141]. Here, we used Kolmogorov-Smirnov (KS) test in addition to Q-Q plots to examine the normality of parameters. KS test is a non-parametric and robust test proposed in [139] for characterizing variability in semiconductors, especially when there are more than 50 samples in the sample space [125]. It is also reported that KS test is more powerful than chi-square test for any sample size [142]. The procedure of KS test calculates a *D* statistics given by:

$$D = \max\left|F^*(x) - S_N(x)\right| \tag{5.1}$$

where *N* is the number of samples and $S_N(x)$ is the sample cumulative distribution function and $F^*(x)$ is the cumulative normal distribution function with the mean and standard

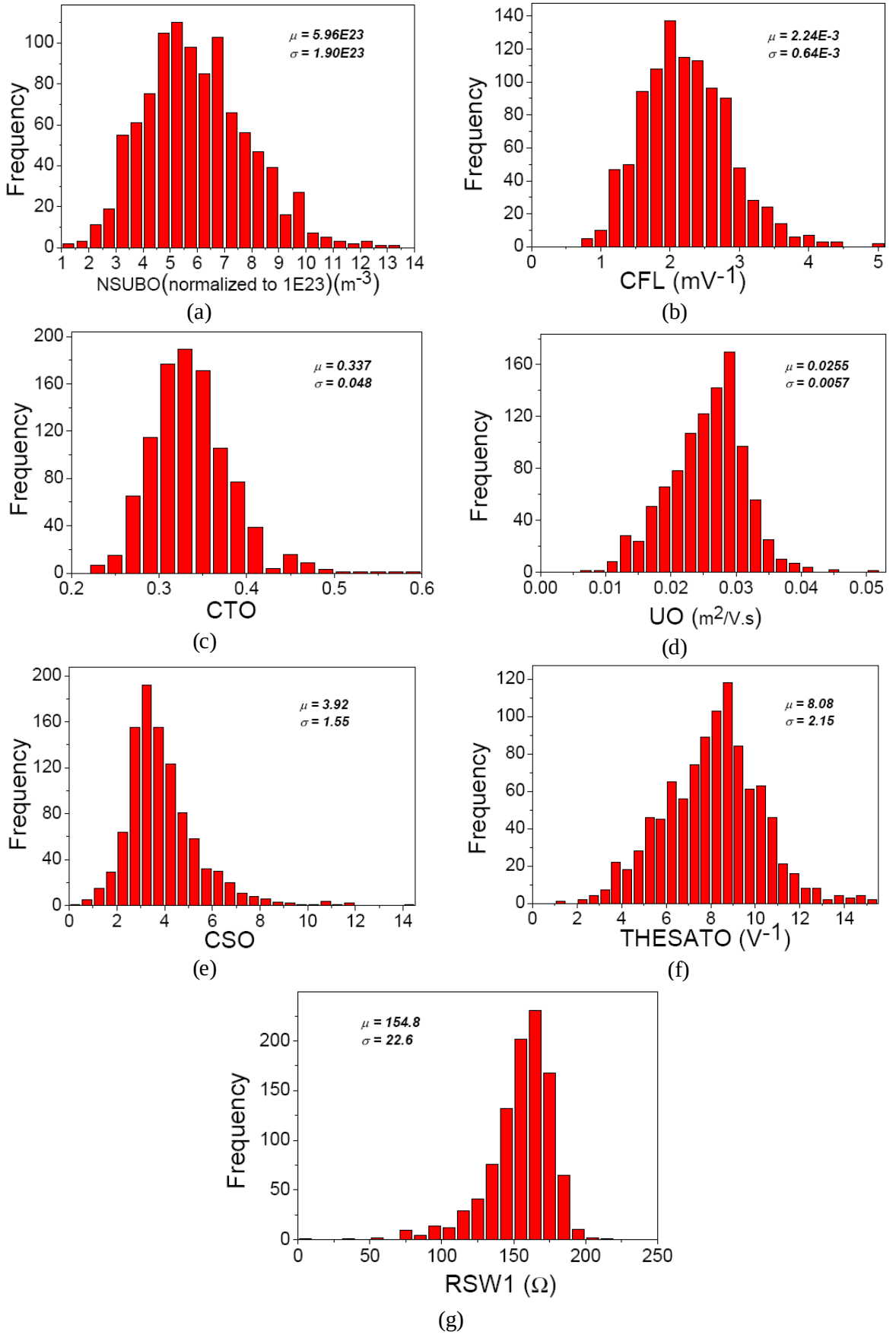


Figure 5.1: Distribution of parameters in 7-parameter statistical extraction.

deviations estimated from samples. If the calculated value of *D* is less than the tabulated critical values, then the null hypothesis will be accepted and the data are assumed to be drawn from a normal distribution. In other conditions, where *D* is larger than critical value, the null hypothesis will be rejected indicating that there is not enough evidence that the data follow a normal distribution. The critical value of *D* is given by $1.63/\sqrt{N}$ at 1% significance level [143], which is equal to 0.0515 for 1000 samples. Due to randomness of data, a significance level is associated with every hypothesis test which is defined by the probability of making error of the null hypothesis rejection whilst it is true. Table 5.1 represents the results of KS test in terms of *D* statistics for each parameter. The immediate result of comparing the *D* statistic with the critical value of 0.0515 is that the null hypothesis for 4 parameters will be accepted and 3 parameters will reject the test hypothesis. In other words, the parameter distributions for *NSUBO, CFL, UO* and *THESATO* follow a normal distribution while for other parameters including *CTO, CSO* and *RSW1* there is not enough evidence for normality assumption.

The Q-Q or normality plots depicted in Figure 5.2 help to verify the KS test results. The linearity of the data shows it is close to normal distribution. Significant deviation from linearity observed in *CTO, CSO* and *RSW1* verifies the fact that they do not follow a normal distribution. For non-normal parameters, the normality plots indicate that *CTO* and *CSO* have right skewness while RSW1 has left skewness. This is consistent with the histograms of those parameters, as presented in Figure 5.1. Defining the mean and variance as the first and second order statistical moments of a distribution, the skewness is the third moment of the distribution. We discuss higher moments of the parameter distributions and their impact on parameter generation accuracy in the section 5.4.

Table 5.1: *D*-statistic results of KS normality test for PSP directly extracted parameters.

Parameter	*NSUBO*	*CFL*	*CTO*	*UO*	*CSO*	*THESATO*	*RSW1*
***D*-statistic**	0.042	0.038	0.066	0.057	0.101	0.046	0.100
Normality	Approved	Approved	Rejected	Approved	Rejected	Approved	Rejected

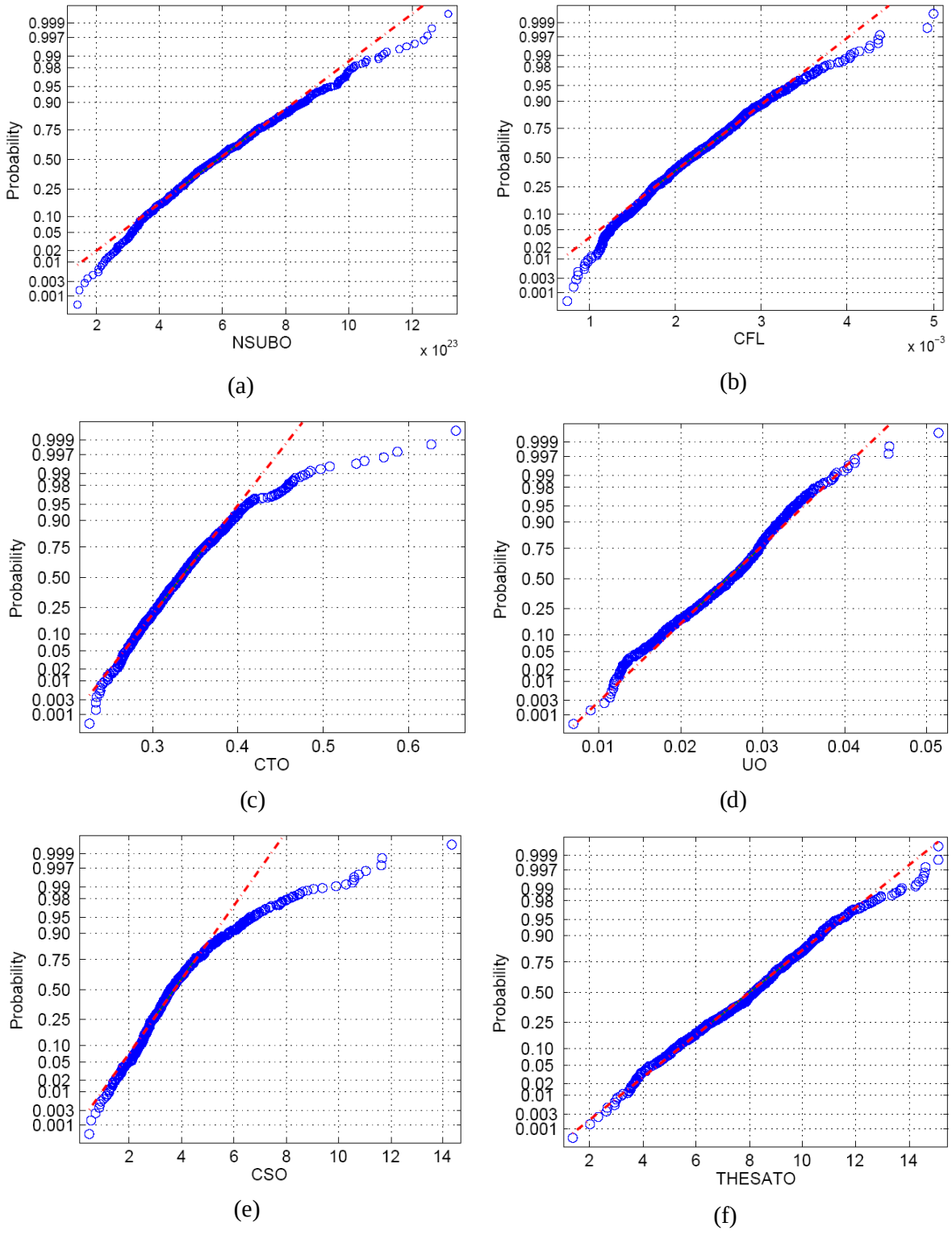
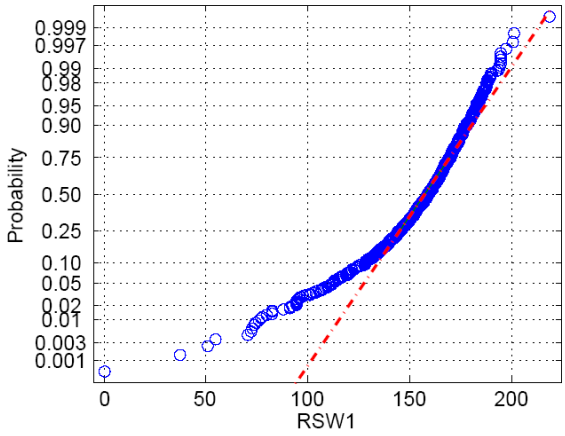


Figure 5.2: Normality plots for directly extracted statistical parameters; (a)-NSUBO, (b)-CFL, (c)-CTO, (d)-UO, (e)-CSO, (f)-THESATO (continue on the next page).



(g)

Figure 5.2 (continue): Normality plots for directly extracted statistical parameters; (g)-RSW1

5.1.2 Parameter Correlations

An intuitive analysis of parameter correlations is illustrated by pairs of scatter plots as shown in Figure 5.3. A pattern which is scattered randomly around the plane indicates no correlation between parameters while a directed pattern shows a correlation between them. In statistical language, a correlation coefficient is a measure of dependency between two parameters *X* and *Y*, and is given by [125]:

$$\rho = \frac{COV(X,Y)}{\sqrt{VAR(X)VAR(Y)}} = \frac{COV(X,Y)}{\sigma_X \sigma_Y} \tag{5.2}$$

where *COV(X,Y)* is the covariance between *X* and *Y* and *VAR* in the denominator is the variance of each parameter which is square of its standard deviation. With the definition given in Eq. 5.2, the correlation coefficient always lies between -1 and +1. Table 5.2 presents the correlation coefficients of the parameters involved in direct statistical parameter extraction. The table is a symmetric matrix and strong correlations can be observed between 5 pairs of parameters: (*NSUBO, CTO*), (*UO, THESATO*), (*UO, RSW1*), (*RSW1, THESATO*) and (*CSO, THESATO*). The parameter *CFL* is almost uncorrelated with all other parameters and the parameter *THESATO* has the most significant correlations with other parameters.

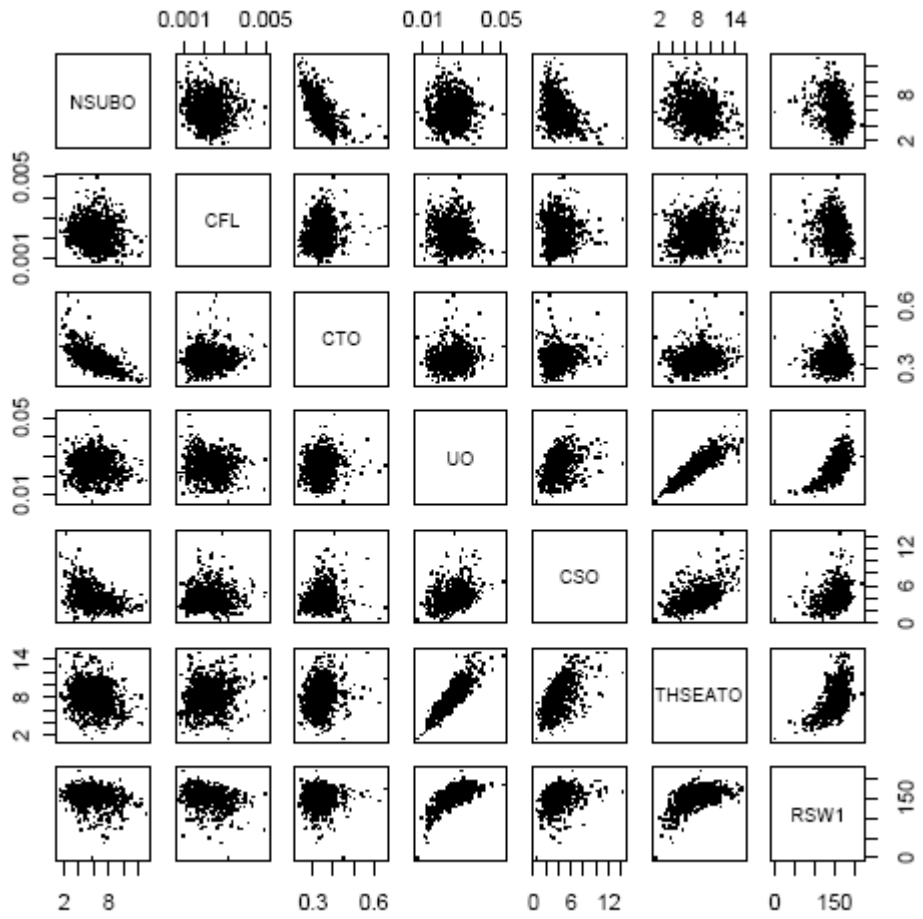


Figure 5.3: Scatter plots for pairs of parameters in 7-parameter statistical set, NSUBO is normalized to 1E23 and other scales correspond to real range of parameters.

Table 5.2: Correlation coefficients of parameters in directly extracted statistical set

	NSUBO	CFL	CTO	UO	CSO	THESATO	RSW1
NSUBO	1.00	-0.09	-0.67	-0.07	-0.39	-0.22	-0.22
CFL	-0.09	1.00	0.11	-0.12	0.00	0.17	-0.30
CTO	-0.67	0.11	1.00	0.08	0.14	0.14	-0.03
UO	-0.07	-0.12	0.08	1.00	0.30	0.87	0.71
CSO	-0.39	0.00	0.14	0.30	1.00	0.52	0.27
THESATO	-0.22	0.17	0.14	0.87	0.52	1.00	0.58
RSW1	-0.22	-0.30	-0.03	0.71	0.27	0.58	1.00

5.2 Gaussian Parameter Generation

Based on the statistical properties of the directly extracted parameters including the normality of four parameters, a simple way to generate statistical compact models is by using independent Gaussian distributions for each parameter with mean and standard deviation obtained from the directly extracted parameter distributions. In other words, two simplifying assumptions have been made in the Gaussian parameter generation technique:

1- All parameters are normally distributed. In reality, the distributions of three parameters are non-normal as was discussed in section 5.1.1. Hence, this is an approximation. However, this assumption could be held after introduction of a non-linear transformation such as Box-Cox transformation which brings non-normal parameters to normal distribution as we will discuss in this section.

2- The parameters are statistically independent. This assumption is not accurate and there are five significant correlations between 21 pairs of parameters as was discussed in section 5.1.2. However, the error caused by using this assumption on the accuracy of MOSFET figures of merit will be evaluated and compared with other parameter generation techniques.

In this section, we implement two types of Gaussian parameter generation. In the type I method, the parameters are generated independently based on the mean and standard deviation of the original directly extracted parameters. In the method type II, the Box-Cox transformation is used to provide normality for non-normal distributed parameters and then the mean and standard deviation of the converted parameters will be used for the independent Gaussian generation of parameters. In the last step of type II method, the generated parameters will be inversely transformed to their original range using an inverse Box-Cox transformation. The accuracy of methods type I and type II will be compared together using histograms and scatter plots for the parameters and Q-Q plots for the MOSFET figures of merit.

5.2.1 Box-Cox Transformation

The Box-Cox transformation was suggested by statisticians George Box and David Cox in 1964 [144]. It has generated a great deal of interest both in theoretical work and practical applications because it is used to convert non-normal distributions to normal. It has been shown in certain cases that Box-Cox transformation cannot bring the distribution to exactly normal; it can lead to a symmetrical distribution with certain restrictions on the first four moments [145]. The transformed samples denoted by y', are related to original samples, *y*, by:

$$y' = \begin{cases} \frac{1}{\lambda}(y^{\lambda} - 1) & \lambda \neq 0 \\ \ln y & \lambda = 0 \end{cases} \tag{5.3}$$

where the parameter λ can be determined using the Maximum Likelihood Estimation (MLE) approach [146]. The transformation has been implemented for non-normal parameters using MATLAB software. Figure 5.4 shows the normality plots for the transformed parameters. A further KS normality test on the transformed parameters has been carried out to investigate effectiveness of this particular transformation. Table 5.3 presents the estimated λ parameter for the Box-Cox transformation of each parameter, the mean and standard deviation of the transformed parameters and the *D*-statistic results obtained from KS test on the transformed parameters. Comparing the *D*-statistic with the critical value of 0.0515 indicates that the transformed parameters follow a normal distribution.

Table 5.3: Statistical analysis of the application of transformation on non-normal parameters

Parameter	*CTO*	*CSO*	*RSW1*
Estimated λ for Box-Cox	-0.8751	0.1824	2.9768
Mean of transformed parameter	-1.860	1.476	1.184E6
SD of transformed parameter	0.349	0.481	0.42E6
***D*-statistic of KS test**	0.0263	0.0475	0.0346
Normality of transformed parameter	Approved	Approved	Approved

The transformed parameters will not be used directly in SPICE modelcards, but instead, their first two moments (mean and SD) will be used for the Gaussian generation of parameters followed by an inverse of Box-Cox transformation to recover the original data range which will be used in SPICE simulations. Having the estimated λ for each parameter as given in Table 5.3, the inverse Box-Cox transformation is:

$$y = \exp\left(\frac{1}{\lambda}\ln(y'\lambda + 1)\right) \tag{5.4}$$

The mentioned procedure in this sub-section is so called Gaussian parameter generation-Type II, as discussed in section 5.2.

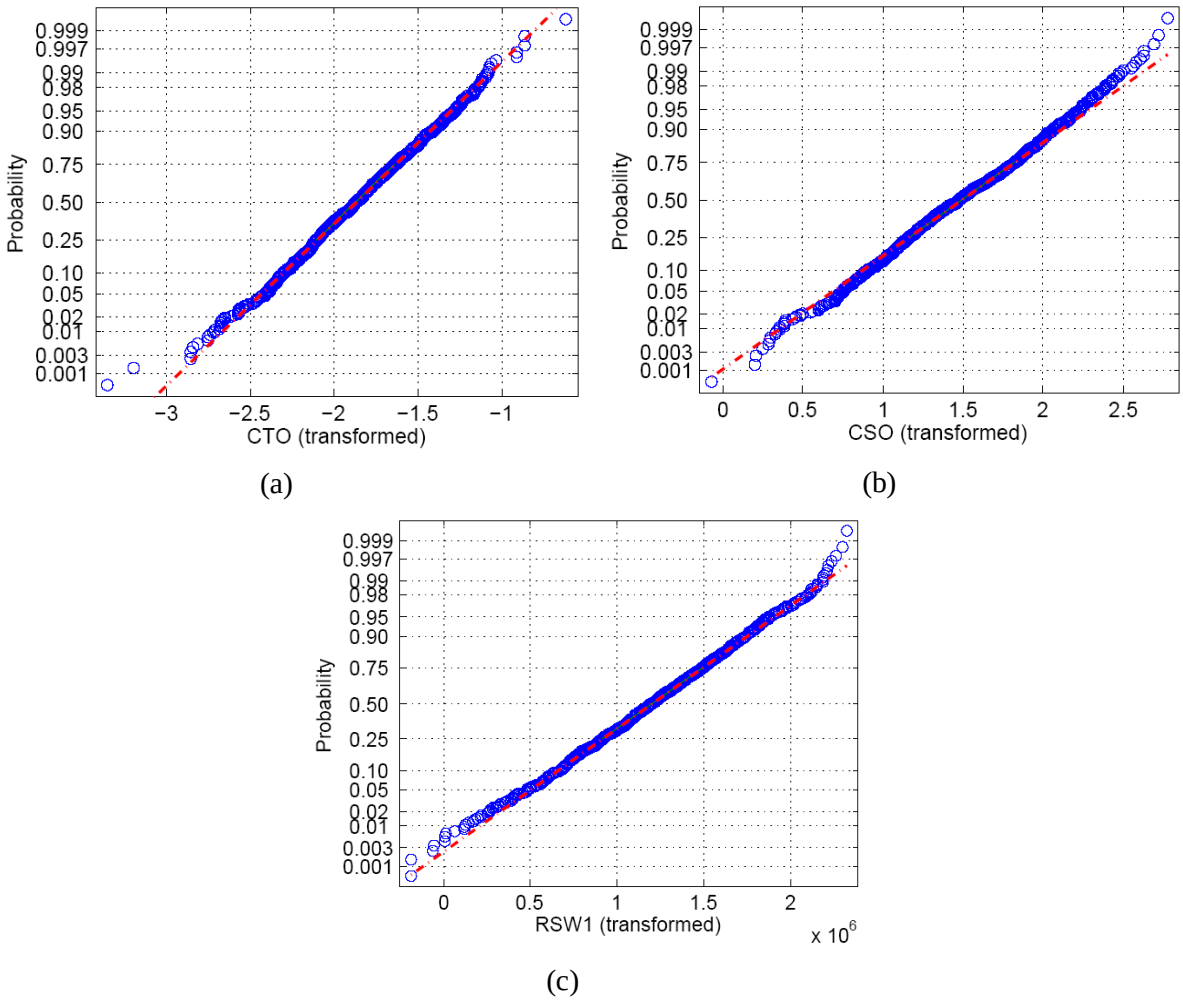
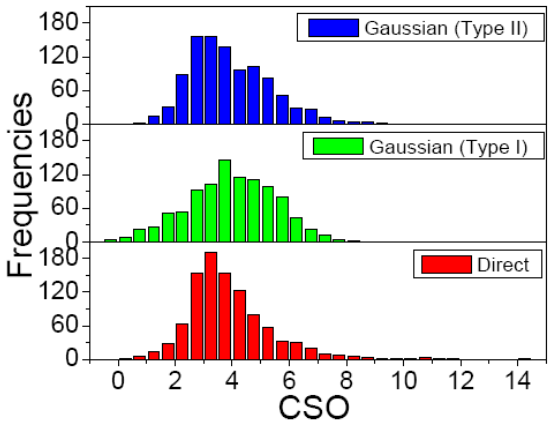


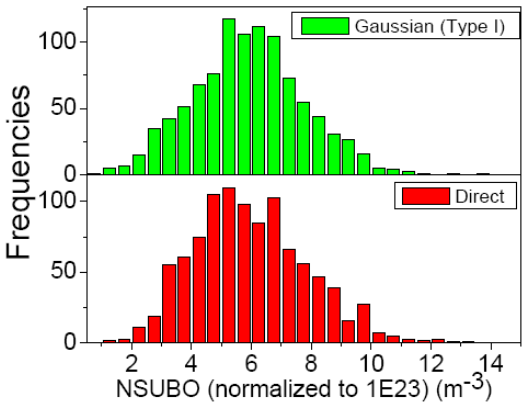
Figure 5.4: Normality plots for the Box-Cox transformed version of originally non-normal statistical parameters; (a)- CTO, (b)-CSO and (c)-RSW1.

5.2.2 Parameter Distributions and Correlations

The first step in the evaluation of the accuracy for a parameter generation technique is to investigate the parameter distributions and correlations in respect to directly extracted original parameters. A typical non-normally distributed parameter is chosen to highlight the difference between two types of Gaussian parameter generation techniques. Figure 5.5(a) compares the distributions of the parameter *CSO* generated using two proposed types of Gaussian parameter generations and the direct approach. It is understood that the parameter generated with Gaussian type II technique is in a closer match in respect to direct approach, owing to the benefits of nonlinear Box-Cox transformation. Particularly, the parameter generated with type II technique has a skewness to right, like directly extracted parameter and it does not have the negative values in the lower tail of the distribution which is the case for type I generation. Figure 5.5(b) shows the distributions for the parameter *NSUBO*, which was essentially normal and parameter generation using Gaussian type I method does not make significant change in the distribution.



(a)



(b)

Figure 5.5: Distribution of two typical parameters in Gaussian technique; (a)-CSO, (b)-NSUBO.

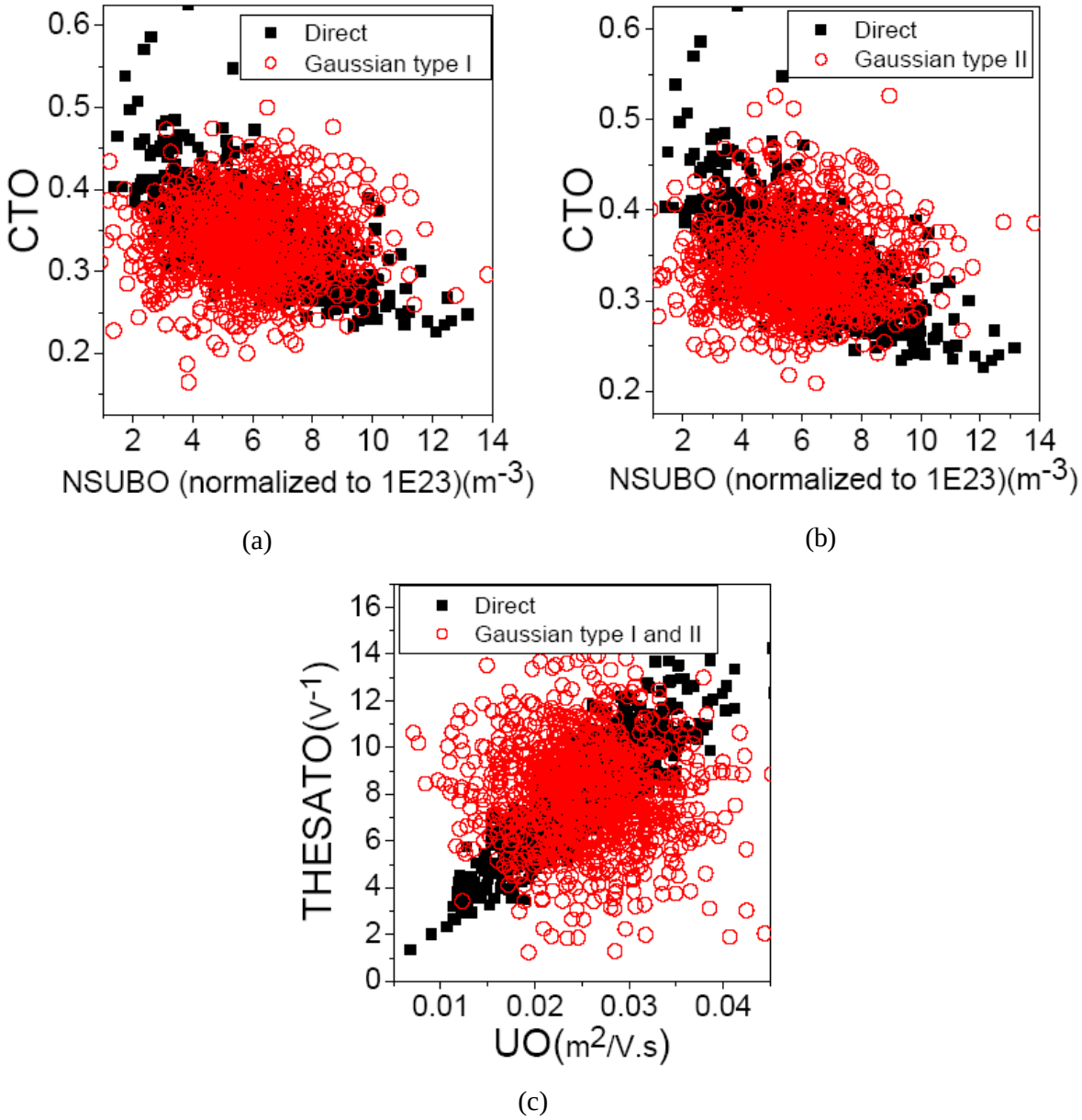


Figure 5.6: Scatter plots between typical correlated PSP parameters; (a)-NSUBO and CTO in direct and Gaussian type I, (b)-NSUBO and CTO in direct and Gaussian type II, (c)-THESATO and UO in direct and Gaussian type I and II.

Figure 5.6 shows the playback of correlation between two significantly correlated PSP parameter pairs, (*NSUBO-CTO*) and (*THESATO-UO*). It is clear that neither Gaussian type I nor type II generation can preserve the correlations between parameters. This is due to the fact that the parameters have been generated independently so the correlations will be lost after parameter generation. Based on the correlation coefficient definition given in Equation (5.2), the correlation coefficients are equal to 0.03, 0.02 and 0.01 for Figures 5.6(a), (b) and (c) respectively. In theory, uncorrelated parameters should have zero correlation coefficient but in the reality, the impact of finite sample size provides such

small correlations. However, these small numbers are negligible compared with the correlation coefficient of 0.67 for (*NSUBO-CTO*) directly extracted pair and 0.87 for (*THESATO-UO*) directly extracted pair.

5.2.3 Impact on MOSFET Figures of Merit

Based on the proposed Gaussian parameter generation techniques discussed in this section and its capabilities to reproduce the parameter distributions missing the correlations, it is important to evaluate the reproduction accuracy of MOSFET figures of merits. To achieve this aim, we have built statistical compact model libraries for each generation technique, and then have simulated with HSPICE a whole range of I_d-V_g characteristics. A drain bias voltage of 1V has been used and the gate voltage has been swept between 0 and 1. Three MOSFET figures of merit were extracted from simulations: I_{off}, I_{on} and V_{th} as defined in the previous chapter.

Figure 5.7 illustrates Q-Q plots for the simulated MOSFET figures of merits which are selected from their statistical compact model libraries, accordingly. The two Gaussian generation techniques are indistinguishable in all plots. Both Gaussian generation techniques can reproduce V_{th} and I_{off} plots compared with the direct approach with a good accuracy, except the upper tail for I_{off} and the lower tail for V_{th}. The main deficiency of Gaussian technique becomes clear by looking into I_{on} plot. There is a significant deviation in the slope of the trend generated with the Gaussian techniques in respect to direct approach. This deviation is translated into an error in the standard deviation (σ) of the I_{on} generated by the Gaussian techniques in respect to the direct approach. While σI_{on} value for the direct approach is 0.12mA, it is equal to 0.3mA from the simulations using Gaussian statistical compact model. The main reason of this error introduced when using Gaussian technique is loss of parameter correlations, particularly *UO* and *THESATO* which both of them affect the drive current of the MOSFET in saturation regime. As a result, using either Gaussian type I or II methods will be acceptable in variability-aware design of very low power circuit if the transistors are operating in sub-threshold regime. Using Gaussian techniques in high power applications is not advisable due to high error in predicting the variability of MOSFET drive current. The main message of this section is that the correlation of parameters is a more important factor than the normality. This fact has been used to develop the PCA parameter generation technique.

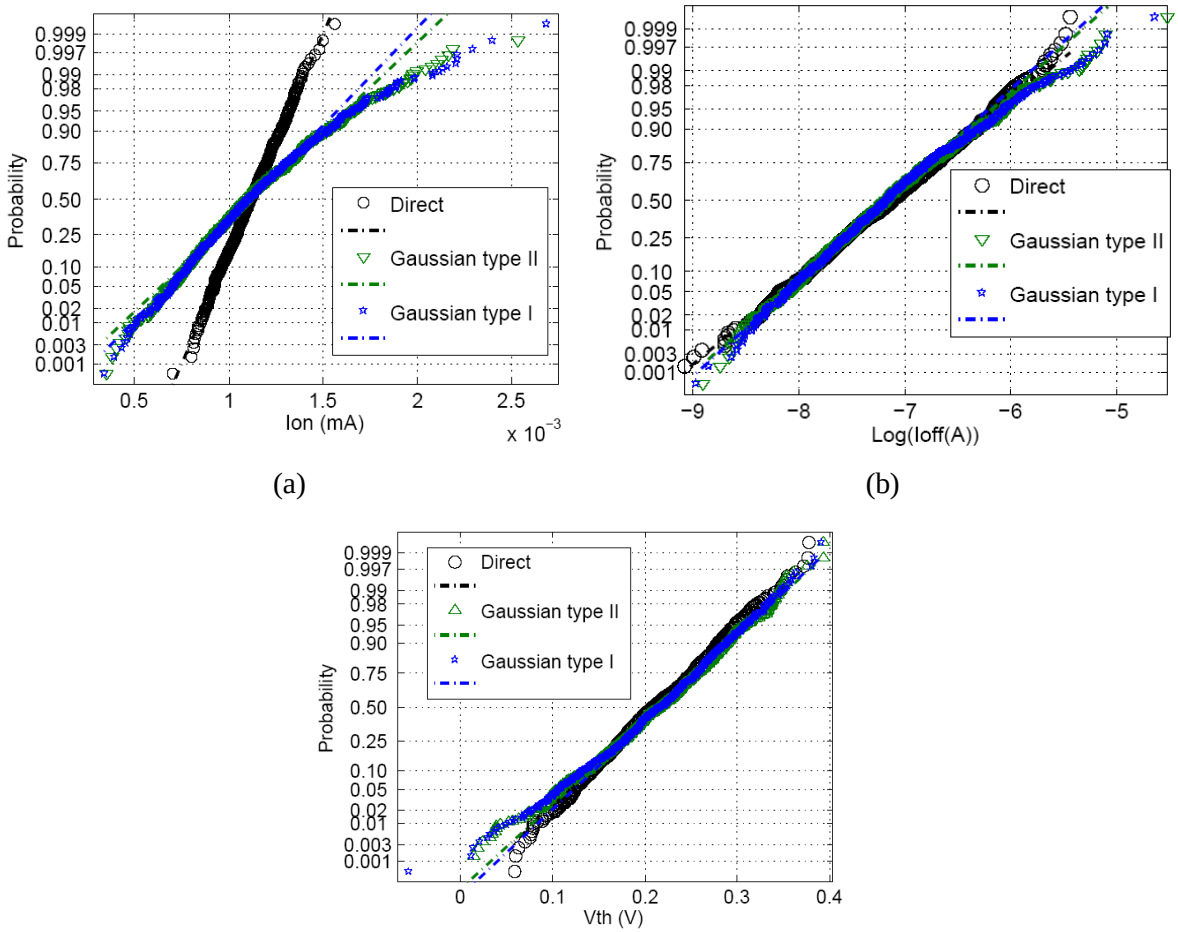


Figure 5.7: Probability plots of MOSFET figures of merit; a comparison between two Gaussian parameter generation techniques and direct approach for (a)-Ion, (b)-Log(Ioff), (c)-Vth.

5.3 Parameter Generation Based on PCA

In this section we employ the Principal Component Analysis (PCA) technique [147] to generate statistical parameters. The main benefit of PCA approach is that it preserves the correlation between parameters. In the PCA approach, the covariance matrix *S* is generated on the basis of the normalized direct parameter extraction results. The normalization process involves subtracting the mean value of a parameter set from each parameter and dividing the result by their standard deviations. Therefore, each parameter set will have μ of zero and σ of one. The key step of PCA is to find eigenvectors and eigenvalues of *S* from:

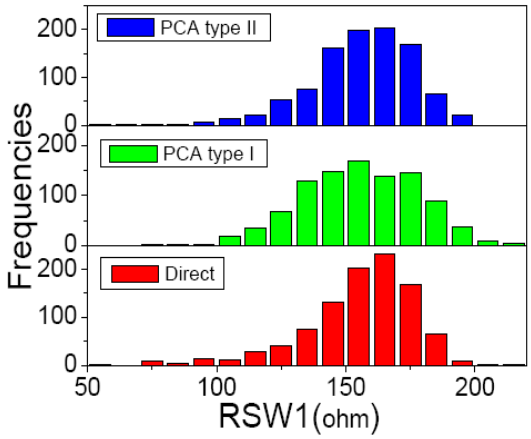
$$U'SU = L \tag{5.5}$$

where U represents the eigenvectors and L represents the eigenvalues of the covariance matrix S. The transformed variables $z = U'x$ are the principal components, where x represents the original matrix of data. In our application, the matrix of data is the statistical 7-parameter set for 1000 sample of directly extracted parameters. Thus, it has 7 rows and 1000 columns. PCA itself does not require that the original multi-dimensional data follow a particular distribution. However, if the original data closely approximates Gaussian distributions, we can reconstruct the data from statistical independent principal components using $x=Uz$ where the corresponding principal components follow a Gaussian distribution with a mean of zero and variances of eigenvalues L.

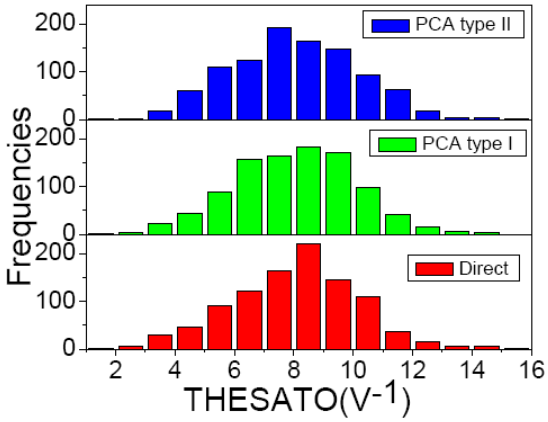
In this section we implement two types of parameter generation using PCA approach. In method type I, we follow the procedure outlined above and Equation (5.5) to construct parameter set based on the principal components. In method type II, we use Box-Cox transformation to convert non-normal distributions to normal as discussed in sub-section 5.2.1, and then we generate parameters based on principal components of the new set. The idea behind this method is that reconstructing data from their principal components will be more accurate based on normal distributions. The last stage of type II method is to perform inverse Box-Cox transformation to recover the range of the original data. However, we will assess the accuracy of this two PCA approaches by simulating MOSFET figures of merit from generated statistical compact model libraries.

5.3.1 Parameter Distributions and Correlations

As discussed in section 5.2.1, the distribution and correlation of parameters are important factors which need to be investigated for each parameter generation technique. Figure 5.8 illustrates the parameter distributions for two typical parameters. One normal and one non-normal parameter has been chosen to demonstrate the impact of different PCA generation techniques on the distribution of parameters. The parameter RSW1 in Figure 5.8(a) produces a closer distribution when generated by PCA type II. This is due to the fact that PCA type II employs Box-Cox transformation before doing PCA procedure. An inverse transformation which will be applied to the generated data will account for the skewness of the original distribution. For an originally normal parameter, both PCA techniques produce close distributions to direct approach as presented in Figure 5.8(b). However, even in this case the parameter distributions resulting from two PCA approaches



(a)



(b)

Figure 5.8: Distribution of two typical parameters in PCA technique; (a)-RSW1, (b)-THESATO.

are not identical because the data sets over which the PCA procedure has been carried out, are different. PCA type I applies PCA on original variables but PCA type II applies PCA on transformed variables.

Figure 5.9 illustrates statistical correlation between typical parameters generated with the proposed PCA techniques. The correlation between (NSUBO, CTO) pair is maintained by PCA type I while the correlation between the pair (THESATO, UO) is maintained by both PCA methods type I and type II. PCA type I always preserves the correlations which exist between directly extracted parameters. This is expected from theory of principal components. PCA type II preserves the correlation between normally distributed parameters which are the parameters which have not been used in the Box-Cox transformation. This is because of the fact that nonlinear power transformation of parameters which is carried out prior to PCA procedure in PCA type II method, distorts the

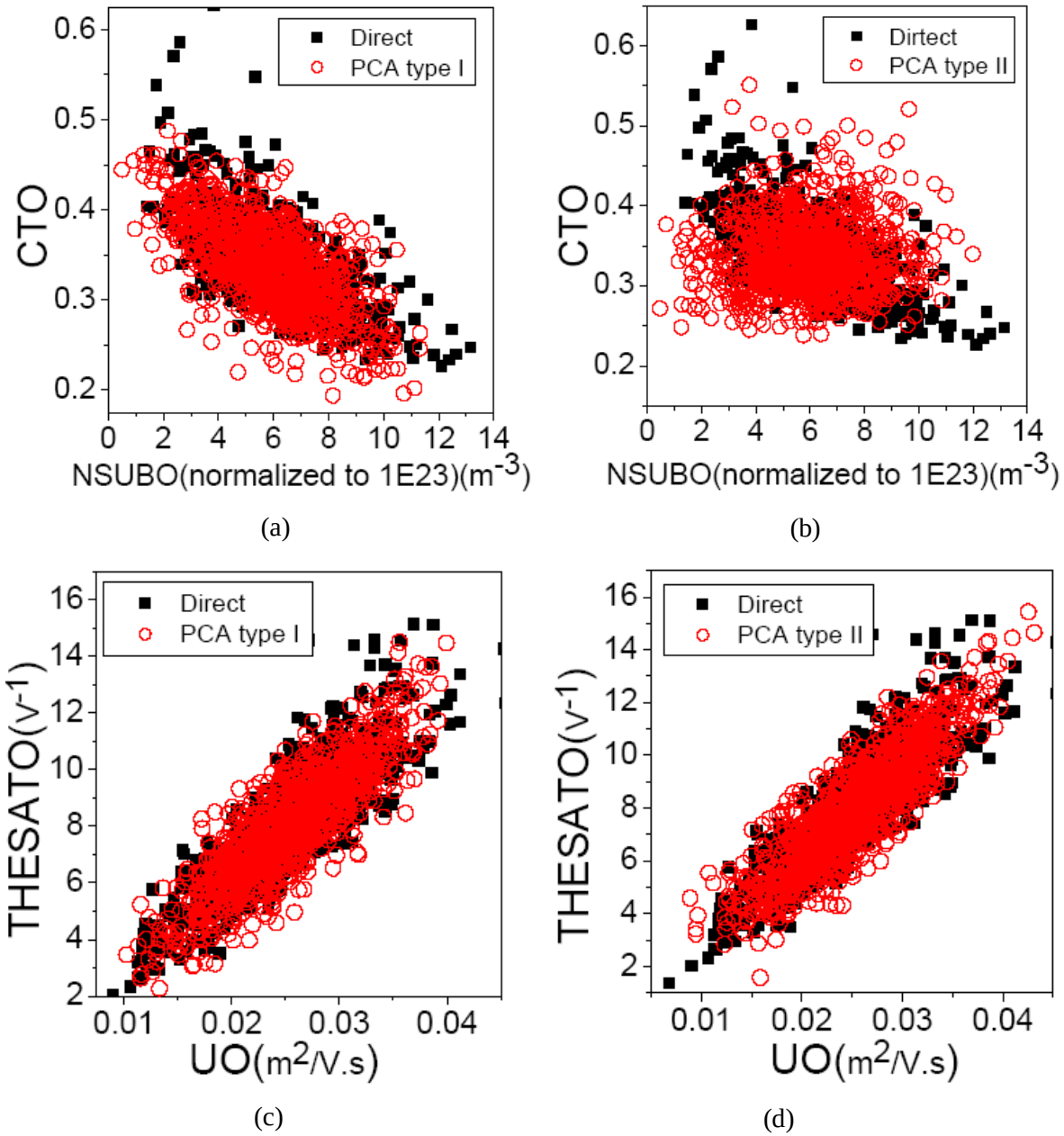


Figure 5.9: Scatter plots between generated PSP parameters; (a)-NSUBO and CTO in PCA type I compared with direct, (b)-NSUBO and CTO in PCA type II compared with direct, (c)-THESATO and UO in PCA type I compared with direct, (d)-THESATO and UO in PCA type II compared with direct approach.

correlations. This can be investigated further by intuitive comparison of the correlations between the typical parameters after and before transformation without doing any PCA, subject of the data which has been shown in Figures 5.2 and 5.4. The correlation distortion is the main deficiency of Box-Cox transformation regardless of its main benefit which is converting non-normal to normal distributions.

5.3.2 Impact on MOSFET Figures of Merit

The statistical compact models generated with two PCA techniques are used to build appropriate statistical compact model libraries. These SCM libraries have been used to simulate MOSFET figures of merits. Figure 5.10 illustrates the Q-Q plots of I_{on}, I_{off} and V_{th} distributions. For I_{on}, there are errors in the standard deviations of both PCA approaches in respect to direct approach. These errors are evident from the different slopes of the lines in Figure 5.10(a). The error of σI_{on} for PCA type I and type II methods are 10% and 44%, respectively. The higher error of PCA type II method is due to loss of some of parameter correlations as discussed in previous section.

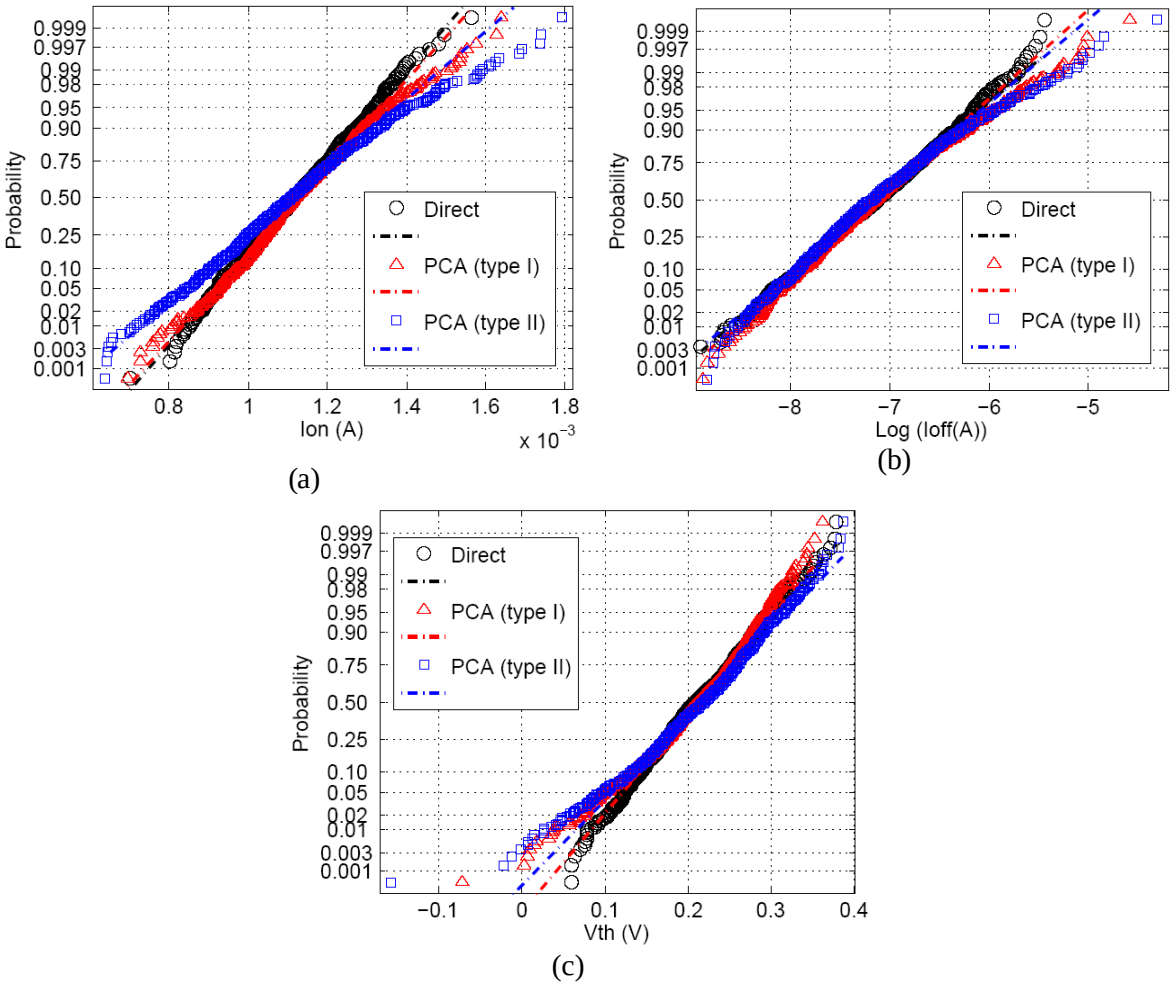


Figure 5.10: Probability plots of MOSFET figures of merit; a comparison between two PCA parameter generation techniques and direct approach for (a)-Ion, (b)-Log(Ioff), (c)-Vth.

For I_{off} in Figure 5.10(b), both PCA approaches are matched together and provide good agreement with respect to the direct approach, except in the upper tail. The error of the mean and SD of I_{off} in both PCA approaches in respect to the mean and SD of the I_{off} in direct approach is negligible (less than 1%).

The investigation on V_{th} which is subject of Figure 5.10(c) shows that both PCA approaches produce a few samples with negative threshold voltage. This is a result of the fact that the I_{off} value for those devices are above the current criteria used to extract the threshold voltage. A current criteria of $10\,\mu A$ has been used to extract V_{th} which gives positive value for all of directly extracted devices. As a result of these negative samples and the trend produced by both PCA approaches in V_{th} plot, there is a significant deviation in the lower tail of both PCA approaches in respect to direct approach. However, there is a slightly better match in the upper tail of PCA type II in respect of direct approach. The error in the mean value of the threshold voltage for PCA approaches is negligible but the σV_{th} error is 7.5% and 22% for PCA types I and II, respectively.

Overall, PCA type I produces less error in all of the simulated MOSFET figures of merit compared with PCA type II. Moreover, the error produced by PCA type I is almost comparable with Gaussian method in V_{th} and I_{off}. In particular, there is identical deviation in upper tail of I_{off} and lower tail of V_{th} for both PCA and Gaussian techniques but the I_{on} distribution is more accurate in PCA compared to the Gaussian approach due to the preservation of the parameter correlations in the PCA approach.

5.4 Nonlinear Power Method

The use of conventional PCA approach in generating SCM parameters results in normally distributed parameters preserving the correlation between parameters and the first two moments of the parameter distributions. However, it does not reconstruct the accurate shape of the distribution of the extracted SCM parameters resulting in significant error especially in I_{on} standard deviation, I_{off} upper tail and V_{th} lower tail. The accurate transfer of the statistical information contained in the extracted SCM parameter distributions into a representative set of randomly generated parameters requires the preservation of the higher moments of the parameter distributions during the generation process. This can be

achieved by applying the Nonlinear Power Method (NPM), an advanced statistical simulation strategy capable of preserving the correlations and reproducing the higher moments of the SCM parameter distributions. The key advantage of NPM method stems from the capability to generate univariate or multivariate non-normal distributions with an arbitrary covariance matrix from a set of analytical equations [148,149].

The NPM is based on a moment-matching technique, and an accurate approximation to the distribution can be quickly obtained with modern computational resources, even when the calculation involves a large number of moments. The number of moments that are required in NPM approach depends on the degree of irregularity of the target distribution function. In this section, we first introduce the NPM method. Then, the distribution and correlation of parameters generated with the NPM approach will be investigated. In the last sub-section the accuracy of the reproduction of the device figures of merit generated using NPM approach will be evaluated.

5.4.1 Formulation

NPM generates the non-normal random variable Y_i using the polynomial transformation of the standard normal variable $Z_i \sim N(0,1)$. Considering the mean and variance as the first and second moments of a distribution, the skewness and kurtosis are defined as third and fourth normalized moments given by [150]:

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \quad ; \quad \gamma_2 = \frac{\mu_4}{\sigma^4} - 3 \tag{5.6}$$

where μ_i represents the *i*'th central moments of any distribution and σ is its standard deviation. The central moments can be related to the moments about the mean which are more convenient to use in applications. The relationships which are important to use in our application are up to 4^{th} order and are given by [150]:

$$\begin{aligned}\mu_2 &= \mu_2' - \mu_1'^2 \\ \mu_3 &= \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3 \\ \mu_4 &= \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4\end{aligned} \tag{5.7}$$

where μ_i' represents the i'th moment about the mean. Since it is desirable to control the value of skewness and kurtosis of any distribution, a polynomial transformation of normal variable Z_i to non-normal variable Y_i is of the form:

$$Y_i = a + bZ_i + cZ_i^2 + dZ_i^3 \tag{5.8}$$

In order to find the coefficients of the polynomial transformation (5.8), it is necessary to find the various moments of Y_i. By substituting Equation (5.7) into Equation (5.6) and algebraic simplification of the moments of Y_i given by Equation (5.8) and knowing the fact that Y_i should have mean of zero, variance of 1, skewness of γ_1 and kurtosis of γ_2, a set of 4 equations with 4 variables should be solved [148]:

$$\begin{aligned}0 &= a + c \\ 1 &= b^2 + 6bd + 2c^2 + 15d^2 \\ \gamma_1 &= 2c(b^2 + 24bd + 105d^2 + 2) \\ \gamma_2 &= 24(bd + c^2[1 + b^2 + 28bd] + d^2[12 + 48bd + 141c^2 + 225d^2])\end{aligned} \tag{5.9}$$

Although solving Equation set (5.9) guarantees the right selection of coefficients *a*,*b*,*c* and *d* in order to produce a distribution for each parameter which is identical to its original distribution up to 4^{th} moment, it does not reproduce the correlations between parameters. In other words, for a 7-parameter SCM set, Equation set (5.9) needs to be solved 7 times to obtain different coefficients of transformation for each parameter but the next important question is how to choose normalized random number vector Z_i to account for the correlation between parameters? As an answer to this question, a method for simulating multivariate non-normal distributions can be used [149]. To describe this method, Equation (5.8) will be rewritten in matrix form as:

$$Y_i = w^T z_i \;\; ; \;\; w^T = [a,b,c,d] \;\; ; \;\; z_i^T = [1, Z_i, Z_i^2, Z_i^3] \tag{5.10}$$

If $r_{Y_1Y_2}$ denotes the correlation between two typical non-normal variables Y_1 and Y_2 corresponding to the normal variables Z_1 and Z_2, and considering the fact that the variables are normalized to have zero mean and variance of one, hence the correlation between Y_1 and Y_2 will be equal to their cross product:

$$r_{Y_1Y_2} = E(Y_1Y_2^T) = E(w_1^T z_1 z_2^T w_2) = w_1^T R w_2 \tag{5.11}$$

where R is the expected matrix product of z_1 and z_2^T which is given by:

$$R = E(z_1 z_2^T) = E\begin{bmatrix} 1 & Z_2 & Z_2^2 & Z_2^3 \\ Z_1 & Z_1Z_2 & Z_1Z_2^2 & Z_1Z_2^3 \\ Z_1^2 & Z_1^2Z_2 & Z_1^2Z_2^2 & Z_1^2Z_2^3 \\ Z_1^3 & Z_1^3Z_2 & Z_1^3Z_2^2 & Z_1^3Z_2^3 \end{bmatrix} \tag{5.12}$$

Equation (5.12) can be further simplified using the method proposed by L.Isserlis [151], where it is proved that the expected value of the odd powers of normalized variables is zero, the expected value of the even powers is one and the expected value of other combinations can be related to the correlations between these two variables. Therefore, Equation (5.12) is simplified to:

$$R = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & \rho_{Z_1Z_2} & 0 & 3\rho_{Z_1Z_2} \\ 1 & 0 & 2\rho_{Z_1Z_2}^2 + 1 & 0 \\ 0 & 3\rho_{Z_1Z_2} & 0 & 6\rho_{Z_1Z_2}^3 + 9\rho_{Z_1Z_2} \end{bmatrix} \tag{5.13}$$

where $\rho_{Z_1Z_2}$ is the correlation between normal variables Z_1 and Z_2. Substituting Eq. (5.13) into Eq. (5.11), results in:

$$r_{Y_1Y_2} = \rho_{Z_1Z_2}(b_1b_2 + 3b_1d_2 + 3d_1b_2 + 9d_1d_2) + \rho^2_{Z_1Z_2}(2c_1c_2) + \rho^3_{Z_1Z_2}(6d_1d_2) \quad (5.14)$$

Having desired values of correlation between parameters for $r_{Y_1Y_2}$ and solving Eq. (5.14), gives the intermediate correlation between two random normal variables Z_1 and Z_2. In the general case of N parameters, the total of N(N-1)/2 polynomial third order equations needs to be solved in order to obtain a complete intermediate correlation matrix. In case of our application with 7 parameters, the roots of 21 cubic polynomials need to be calculated. Since the full intermediate correlation matrix (E) is symmetrical, it can be decomposed into a form:

$$E = AA^T = U^T DU \quad (5.15)$$

where A is a lower triangle matrix and A^T is its transpose. This decomposition can be carried out using matrix of eigenvectors (U) and diagonal matrix of eigenvalues (D). Choosing $A=(D^{1/2}U)^T$ satisfies Eq. (5.15). After obtaining matrix A from decomposition of intermediate correlation matrix E, random variables Z_i are given by:

$$Z = A'V \;;\; A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} \\ 0 & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} \\ 0 & 0 & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} \\ 0 & 0 & 0 & a_{44} & a_{45} & a_{46} & a_{47} \\ 0 & 0 & 0 & 0 & a_{55} & a_{56} & a_{57} \\ 0 & 0 & 0 & 0 & 0 & a_{66} & a_{67} \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{77} \end{pmatrix} \;;\; V = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \\ V_6 \\ V_7 \end{bmatrix} \quad (5.16)$$

where V_1, V_2,…,V_7 are independent normalized Gaussian random variables (~N(0,1)).

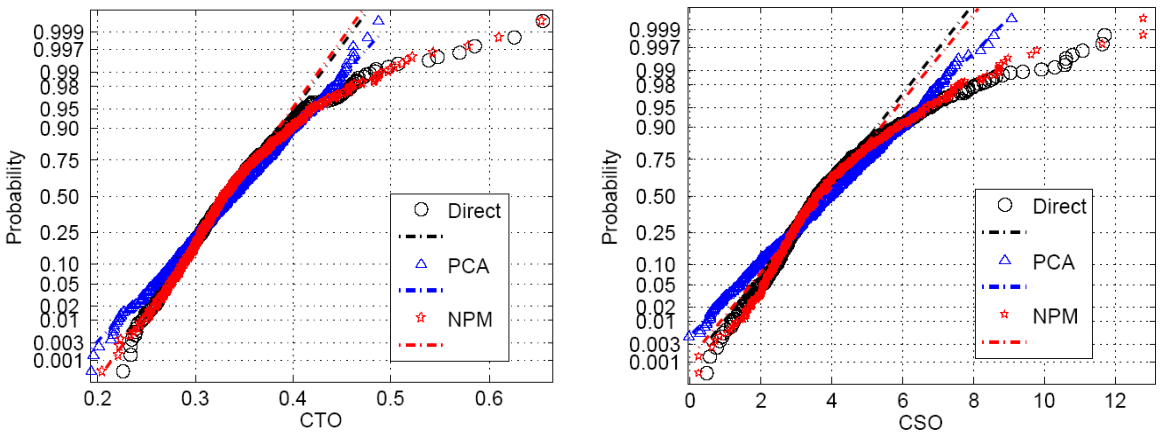


Figure 5.11: Distribution of two non-normal parameters in generated with NPM approach and compared with PCA and direct results; (a)-CTO, (b)-CSO.

5.4.2 Parameter Distributions and Correlations

By applying the NPM approach and considering the preservation of the first 4 moments, statistical parameters have been generated based on directly extracted parameters. Figure (5.11) compares generated parameter distributions using NPM and PCA (type I) with the original directly extracted parameter distributions. Clearly, by maintaining the higher moments of the distributions through the NPM approach, the non-normal shape of SCM parameter distributions are better recovered compared with PCA. Figure 5.12 shows the correlation between chosen generated parameters through NPM approach and compares the correlations with directly extracted parameters. It is clearly seen that the correlations have been well preserved.

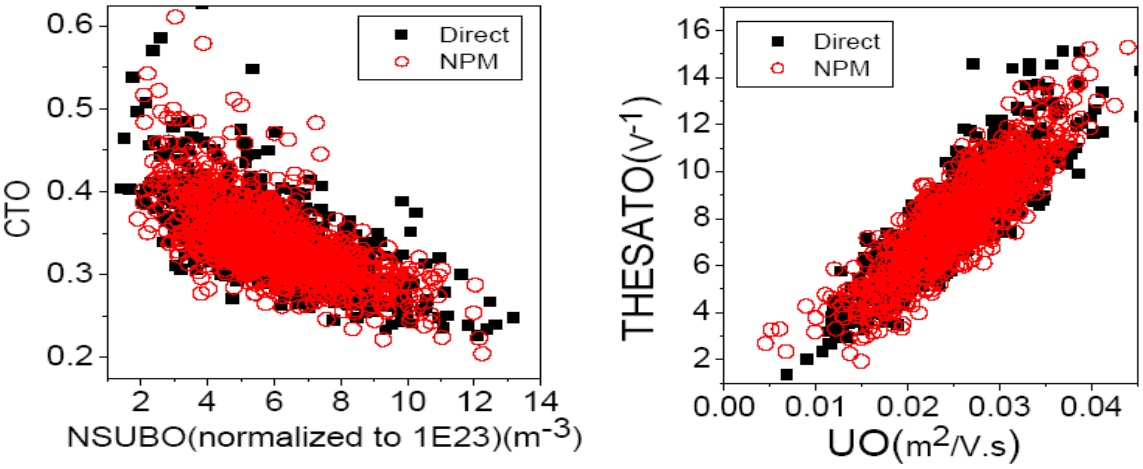


Figure 5.12: Scatter plots between SCM parameters generated with NPM approach and compared with direct approach, left is CTO versus NSUBO, right is THESATO versus UO.

5.4.3 Impact on MOSFET Figures of Merit

Based on SCM library including the parameters generated with NPM approach, a simulation of MOSFET figures of merit has been carried out to extract the exact distributions for I_{on}, I_{off} and V_{th}. Figure 5.13 illustrates the Q-Q plots of MOSFET figures of merit simulated with NPM approach and compares them to PCA (type I) distribution and the direct approach. For I_{on} displayed in Fig. 5.13(a), the upper tail of NPM is more close to the direct distribution compared with PCA approach but for the lower tail, the PCA approach provides a better match. On the other hand, the error in σI_{on} of NPM in respect to direct is less than 3% compared with 10% for PCA. Hence, NPM provides better prediction compared with PCA. For I_{off} which is subject of Fig. 5.13(b), The NPM is in a very close match to the direct approach while PCA produces a significant deviation in the upper tail. The error in σI_{off} is negligible for both NPM and PCA.

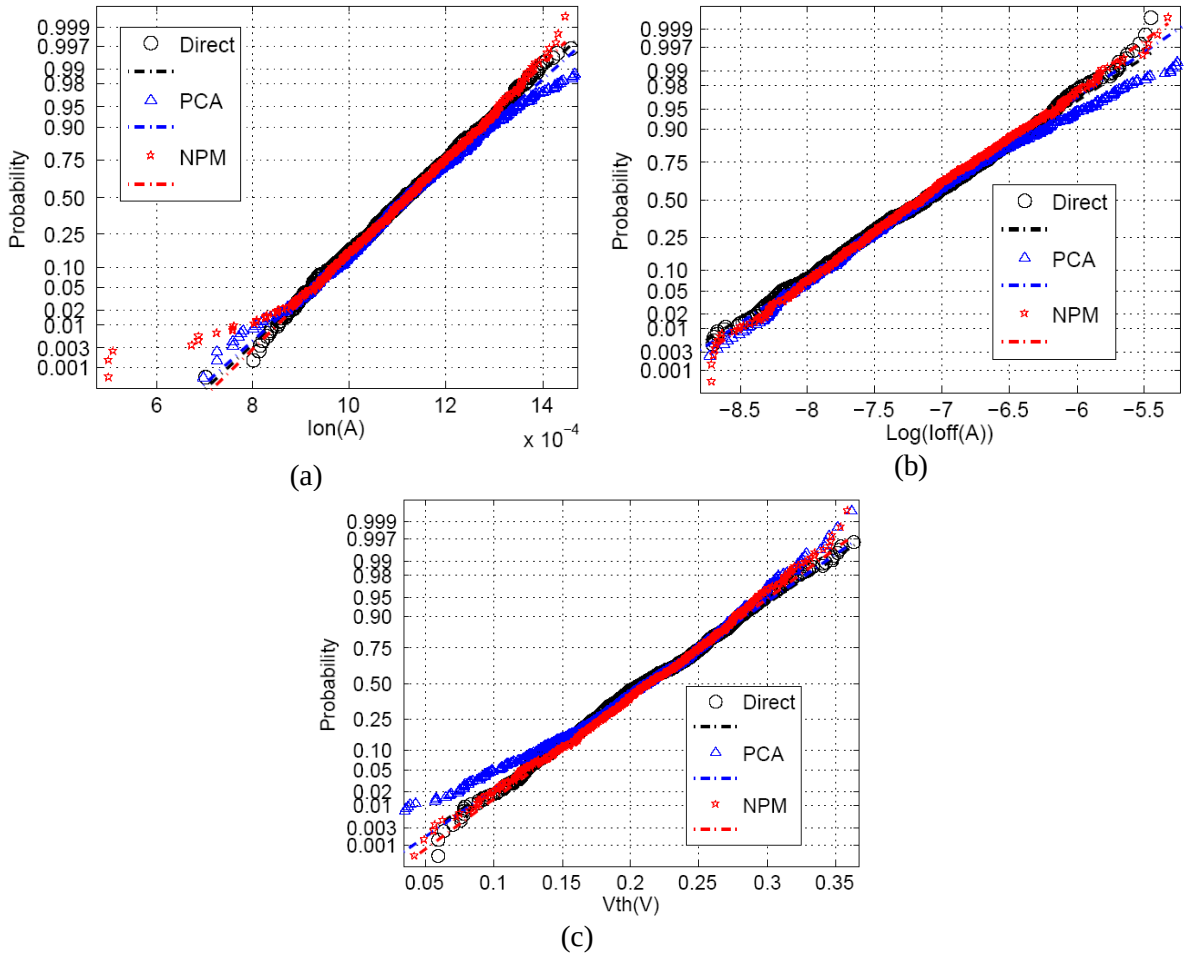


Figure 5.13: Probability plots of MOSFET figures of merit; a comparison between NPM and PCA generation techniques and direct approach for (a)-Ion, (b)-Log(Ioff), (c)-Vth.

Fig. 5.13(c) illustrates the distributions of V_{th} using NPM, PCA and direct approaches. NPM has a very close match to the direct approach compared with PCA, particularly in the lower tail of V_{th} where PCA significantly deviates from the trend produced by the direct approach. The σV_{th} error produced by NPM is 3% which is slightly lower than 7% error produced by PCA.

5.5 Statistical Circuit Simulation

In order to assess the accuracy of different statistical compact model parameter generation strategies on circuit simulations, we simulate figures of merits for a CMOS inverter. The CMOS inverter schematic and specifications are identical to those defined in Figure 4.17, in the previous chapter. An input rise/fall time of 10psec has been considered for the input and the inverter has no external capacitive load to highlight the impact of statistical variability in the inverter figures of merit. The delay and the dissipated energy of the inverter have been simulated as two inverter figures of merit using four SCM libraries: Direct, Gaussian, PCA and NPM. For Gaussian and PCA techniques, type I methods have been considered as for the Gaussian parameter generation it gives the same accuracy of type II and for PCA it leads to more accurate MOSFET figures of merits compared with type II method.

Figure 5.14 illustrates Q-Q plots of the dynamic energy dissipation for the CMOS inverter using four different parameter generation techniques. It is clear that the Gaussian

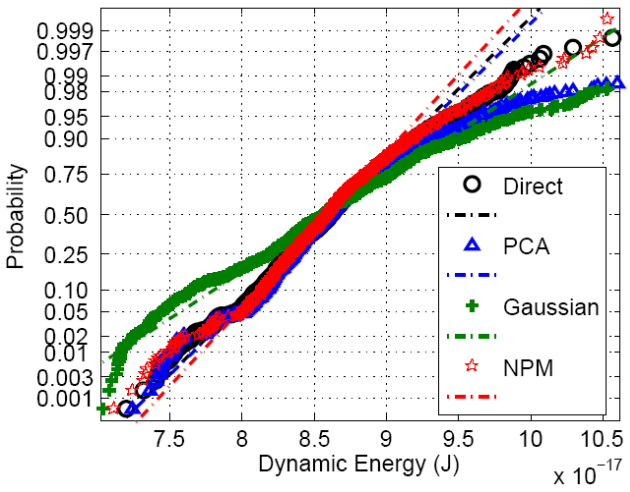


Figure 5.14: Q-Q plots of the distribution of dynamic energy in CMOS inverter using different statistical compact model generation techniques.

technique has the worst accuracy as it is producing a large error (61%) on the standard deviation of the distribution. The most accurate technique is the NPM which matches accurately the upper and the lower tails in comparison to the direct distribution. The PCA approach shows significant deviations in the upper tail but matches in the lower tail. Figure 5.15 presents the same comparison for the inverter delay. NPM provides the best match in respect to the direct method. PCA has deviations in the lower tail. Gaussian has a different slope which provides considerable error (75%) in delay standard deviation in respect to direct approach.

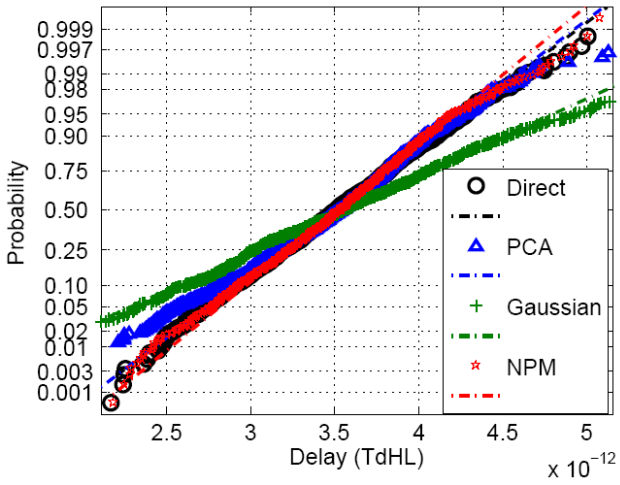


Figure 5.15: Q-Q plots of the distribution of inverter delay using different statistical compact model generation techniques.

5.6 Summary

In this chapter we proposed different statistical parameter generation techniques. Having the results of the most accurate SCM from the previous chapter, we studied the statistical properties of parameters including their distributions and correlations. Using normal probability plots clearly showed the deviations of the tails of statistical parameters from the normal distribution. This problem was further exacerbated by the fact that the extracted parameters were not statistically independent, owing to the complex physical mechanism involved in the device operation at the deca-nanometer scale and some unavoidable aspects of the empirical nature of compact model. These statistical properties were used in creating other techniques which have the capability of replicating the

statistical properties of directly extracted parameters.

Three parameter generation techniques were introduced: Gaussian, PCA, and NPM. Gaussian generation technique relies on generation of parameters independently based on the first two moments of their original distributions. PCA relies on principal components of the parameter set, hence, it considers the parameter correlations. Two flavors of Gaussian and PCA techniques were discussed in this chapter: type I and type II. While type I refers to the normal implementation of each method, type II method involves the Box-Cox transformation of non-normally distributed parameters to preserve the assumption of normality. By evaluating the MOSFET figures of merits, it was clear that both types produce the same accuracies in Gaussian technique while type I produces more accurate results in PCA. This was due to loss of correlation between parameters introduced by the application of Box-Cox transformation.

The NPM approach relies on higher moments matching technique for the parameters. NPM produces the most accurate results in reproduction of MOSFET figures of merits compared with other methods. The accuracy of parameter generation techniques was assessed in statistical circuit simulation using a CMOS inverter as discussed in the previous chapter. It is clearly seen that the NPM provides very accurate match in both inverter delay and dissipated energy compared with other techniques.

Chapter 6

Statistical Modelling of Different Width/Length MOSFETs

Although the evolution of CMOS technology is motivated by scaling of the device dimensions, the circuit designers tend to use a variety of different width/length MOSFETs in their designs rather than using minimum dimension square devices. Wider MOSFETs are preferable in the bias stages of analog integrated circuit design because they provide larger drive current. Moreover, different width devices are being used in digital integrated circuit blocks like SRAM cells and CMOS logic gates. Larger length MOSFETs are employed to decrease the leakage currents in sensitive parts of a mixed-mode circuits in extremely low power CMOS design where devices are biased in subthreshold region. The longer/wider devices exhibit less statistical variability because the fluctuations introduced by RDD, LER and PGG will be averaged in larger areas.

This chapter is devoted to statistical modeling of different width/length MOSFETs based on the extension of the methodologies, simulation techniques and statistical compact models which we have developed in the previous chapters. In the first section, we discuss statistical modeling of different width devices. We will explore the trend for both device figures of merits and SCM parameters as a function of the channel width adopting two methods. In the second section, statistical compact models will be developed for variable length but fixed width devices to investigate the trends as a function of length.

6.1 Statistical Modeling of Width Dependence

It is important to model width dependencies of both MOSFET figures of merit and statistical compact model parameters. In order to achieve these aims, the first step is to simulate the impact of different sources of statistical variability in multi-width devices. Having the I_d-V_g characteristics of different width devices will enable us to extract MOSFET figures of merit and their distribution behavior versus transistor width. In the next step, using the methodology and strategy of the direct statistical parameter extraction which has been developed in chapter 4 for a basic L=W transistor, a set of SCM parameters will be extracted for the devices with multiple widths. Finally, the trend of SCM parameters versus width will be investigated.

6.1.1 Simulation Methods

Two simulation methods are used for generation of different width devices. The first method relies on 'atomistic' simulation of devices in presence of statistical variability sources using Glasgow University atomistic simulator, GARAND. This is the most accurate method but takes a lot of time to complete the simulations. The second method is based on the simulation results of a transistor with L=W and hence is computationally less demanding compared to the first method but perhaps less accurate. It employs the gate slicing of wider devices [152]. In this method, a wider device can be sliced into several basic width devices. Therefore it is possible to provide statistical characteristics of a wider device in terms of the simulated statistical variations of a basic width device. For instance, the I_d-V_g characteristics of transistors with width W=2L will be calculated by:

$$I_{k,W_2} = I_{i,W_1} + I_{j,W_1} \tag{6.1}$$

where *i, j* are random and independently selected integer indices between 1 and N_s where N_s is the size of the simulated statistical sample of W=L transistors. $I_{i,W1}$ and $I_{j,W1}$ represent the current of particular device numbers *i* and *j* from basic width device library of N_s devices. $I_{k,W2}$ is the resulting current of a particular W=2L width transistor composed of two basic width devices. In practice, two steps were followed to produce the I_d-V_g statistical characteristics of devices which are N times wider than the basic width

transistor: first, independent random selection of N devices from I_d-V_g library of basic width transistors, and second, adding currents of selected devices at each gate bias voltage. These two steps were implemented in a Python script.

The simulation of different width devices with the Glasgow atomistic simulator was carried out in presence of different sources of statistical variability including RDD, LER and PGG. The implementation of these sources followed the procedure discussed in chapter 4 for basic width device. The only difference for the simulation of wider devices was that the grid size along *y*-axis was increased accordingly leading to a linear increase in the number of mesh points which in turn increased the run time significantly. Total of 1000 microscopically different devices were simulated for widths of 2, 4, 6 and 8 times of the basic width device at two drain bias voltages of 50mv and 1v. The basic width device has a width of 35nm as discussed in chapter 4. Figure 6.1 illustrates the statistical I_d-V_g characteristics of different width devices simulated with the atomistic simulator. It clearly indicates that the variability is reduced for wider devices. For instance, the leakage current spread which is more than three orders of magnitude in the basic width transistors as shown in Fig. 6.1(a), is reduced to less than two orders of magnitude for the transistors with widths W=8L as shown in Fig. 6.1(d). The reduction in the drive current variability and threshold voltage variability with the increase of the channel width can also be seen in Fig. 6.1.

6.1.2 Impact of Width on MOSFET Figures of Merit

Three MOSFET figures of merit have been extracted from I_d-V_g simulation results of different width transistors: I_{on}, I_{off} and V_{th} as defined in chapter 4. I_{on} and I_{off} are considered per one micrometer width of different width devices. Figures of merit were extracted twice: First from most accurate atomistic simulation results and second from the slicing of wider devices in terms of basic width device as discussed in the last section.

Figure 6.2 shows the statistical trend of the leakage current versus width. The drain bias is 1V. It is clear that the results of slicing method closely follow the atomistic simulation results with an error of maximum 2%. The mean value of $\text{Log}(I_{off})$ is slightly increased with the increase of width while the standard deviation of $\text{Log}(I_{off})$ reduces significantly, from 0.62 for 35nm width devices to 0.25 for 240nm channel width devices.

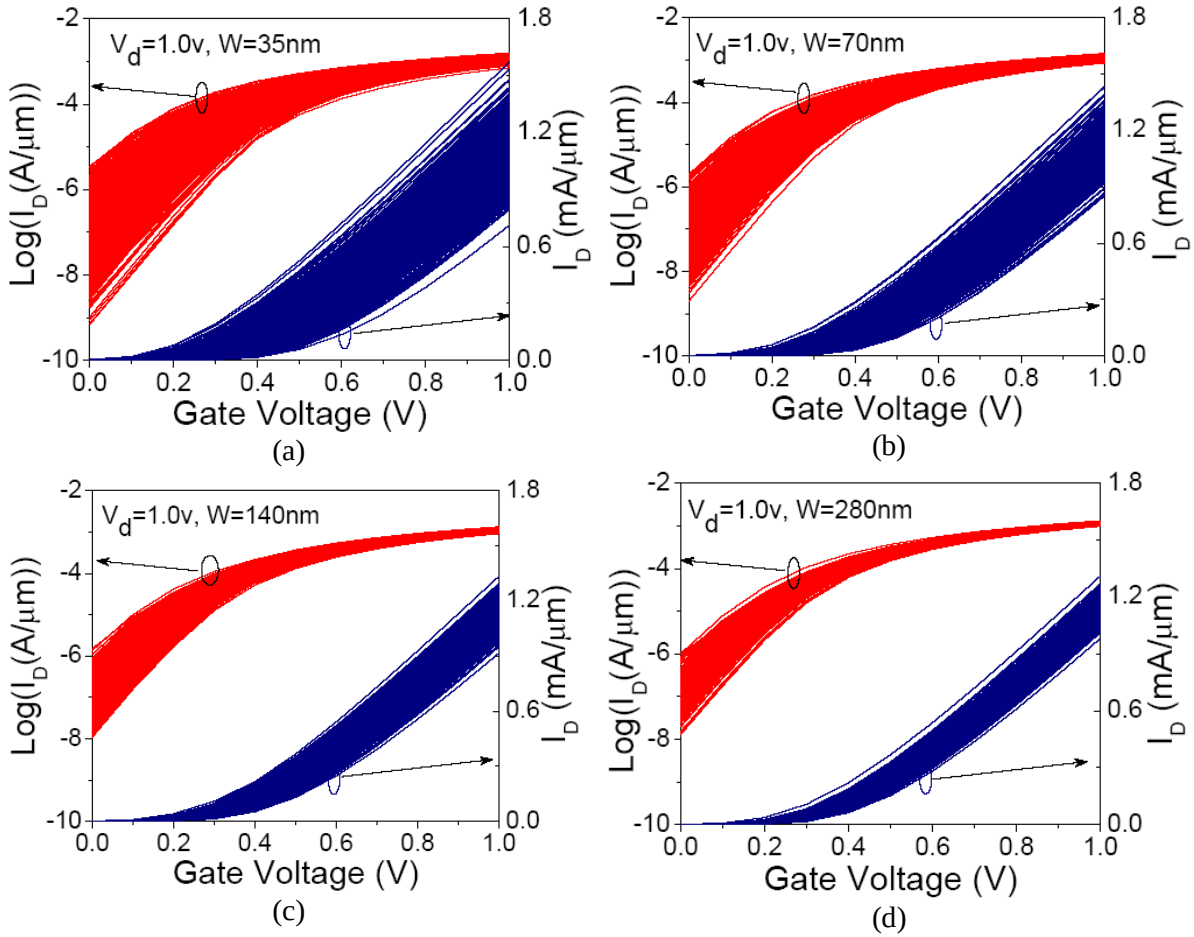


Figure 6.1: Atomistic simulation of SV for devices having different widths of, (a)-35nm, (b)-70nm, (c)-140nm, (d)-280nm.

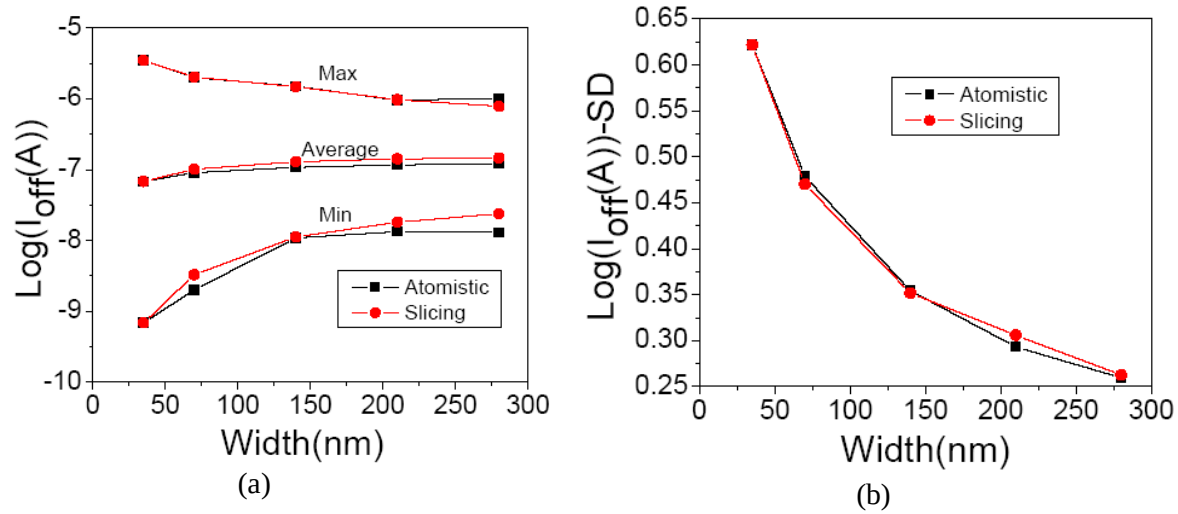


Figure 6.2: Statistical analysis of I_{off} versus width resulted from atomistic simulations and slicing method; (a)-min, max and average of Log(I_{off}), (b)-SD of Log(I_{off}).

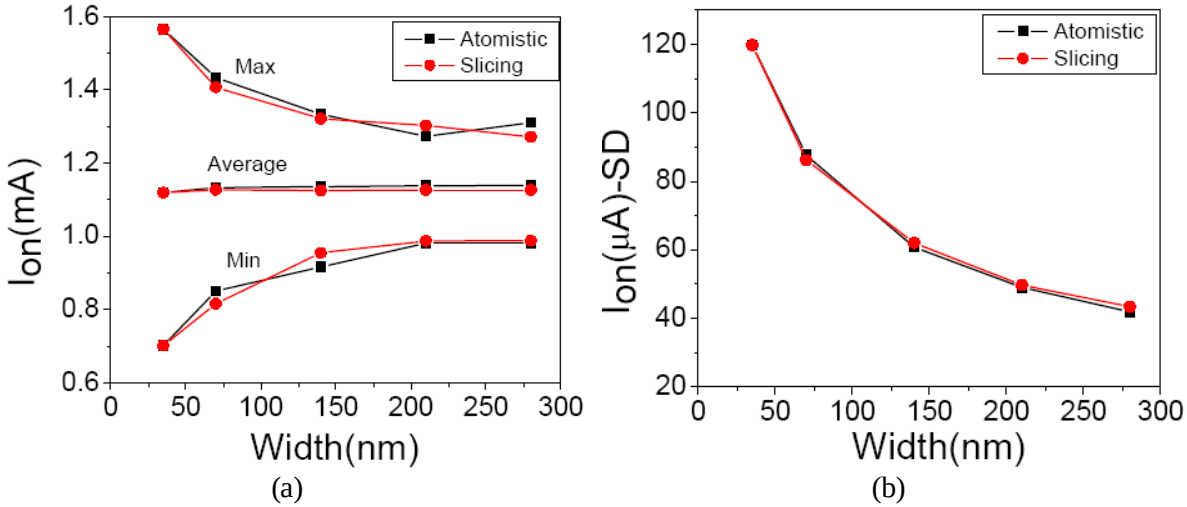


Figure 6.3: Statistical analysis of I_{on} versus width resulted from atomistic simulations and slicing method; (a)-min, max and average of I_{on}, (b)-SD of I_{on}.

Figure 6.3 illustrates the impact of width on the statistical trend of MOSFET drive current (I_{on}). The drain bias is 1V. It clearly shows that the impact of statistical variability is significantly reduced in wider devices. The SD of I_{on} decreases from 120 μA for 35nm width transistors to 40 μA for 280nm width devices. The mean value of the drive current remains constant with the increase of the channel width. Similarly to I_{off}, the error of the slicing method for mean and SD of I_{on} are less than 2% compared to the physical simulations of transistors with different channel width.

Figure 6.4 presents the trend of threshold voltage (V_{th}) versus width. A constant current criteria of 10 µA/µm has been used to extract V_{th} for every I_d-V_g characteristic obtained from atomistic simulator or the slicing method at a drain bias of 1V. It can be seen that the mean values are almost constant with less than 5% fluctuations in respect to the mean of V_{th} in basic width devices. Simultaneously, σV_{th} is decreasing from 55mV to 20mV when the width increases from 35nm to 280nm. The pattern of decay for σV_{th} is very similar to that for I_{on} and I_{off}. The results for the mean and SD of the slicing method are very close to the atomistic simulation results with error less than 2%.

It is also important to plot σV_{th} versus $1/\sqrt{WL}$ as this is a common way to benchmark variations in threshold voltage as discussed in chapter 2. The line slope termed A_{VT} is an indication of threshold voltage variability for a particular technology and can be used to

compare the variability between different technologies. Figure 6.5 is a plot of σV_{th} versus $1/\sqrt{WL}$ which results in the value of A_{VT} equal to 2.0(mV.um). If the plot represents $\sigma\Delta V_{th}$, the resulting A_{VT} will be multiplied by $\sqrt{2}$. The small deviation of the points from the straight line indicates that RDD is a dominant factor of statistical variability in 35nm gate length devices compared to LER and PGG. The zero intercept of A_{VT} line in the plot verifies the correct selection of effective length and width in the calculation of σV_{th} and is consistent with the expectation of zero variability for very-large-area devices.

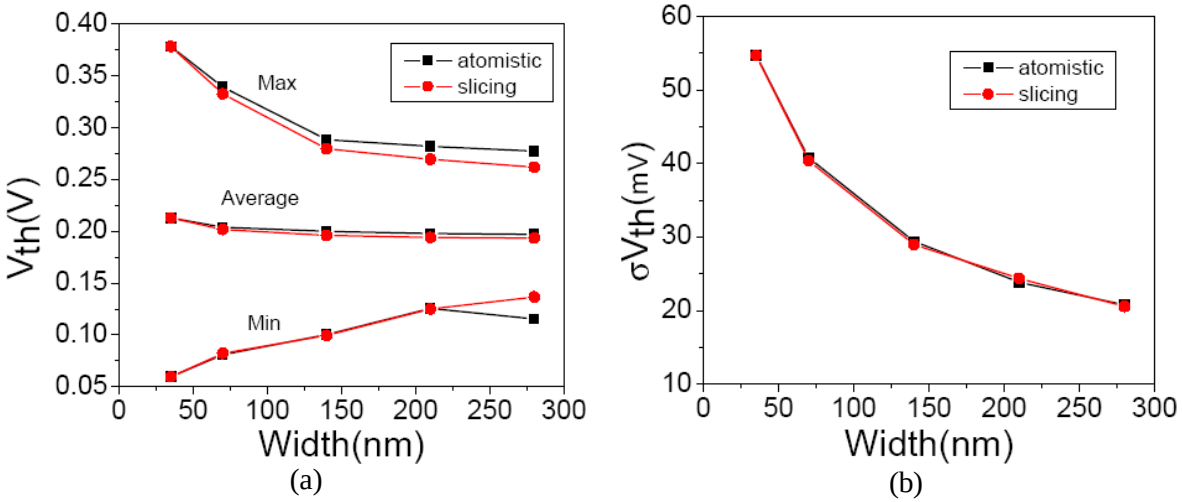


Figure 6.4: Statistical analysis of V_{th} versus width resulted from atomistic simulations and slicing method; (a)-min, max and average of V_{th}, (b)-SD of V_{th}.

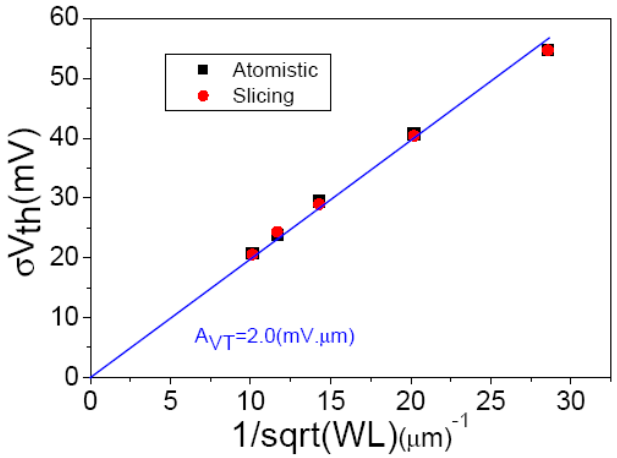
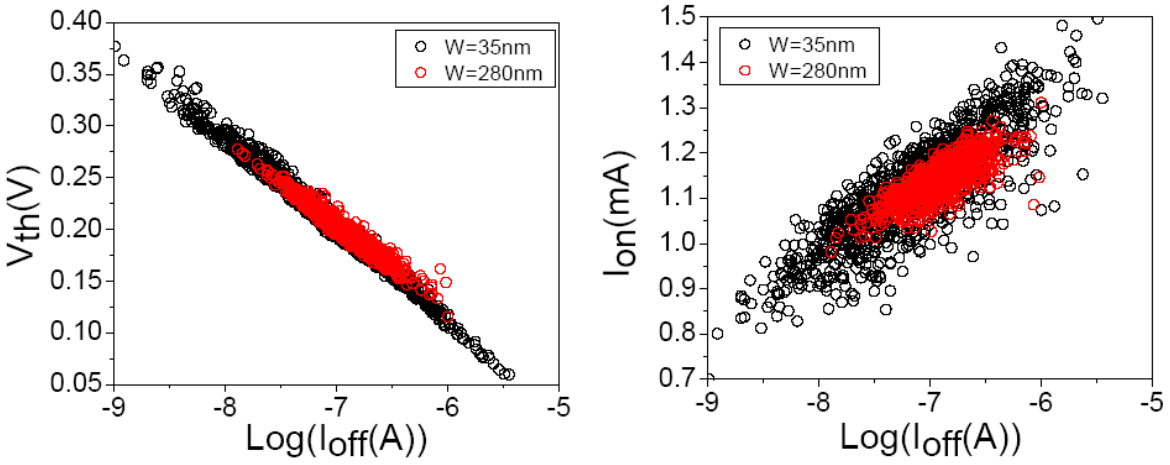


Figure 6.5: SD of V_{th} versus $1/\sqrt{area}$ obtained from simulation of different width devices with slicing method and atomistic simulator.



(a) (b)

Figure 6.6: Correlation between typical MOSFET figures of merit in different width devices, (a)-Scatter plots of V_{th} versus Log(I_{off}), (b)-Scatter plots of I_{on} versus Log(I_{off}).

The main cause of the small error (less than 2%) observed in mean and SD of MOSFET figures of merit using slicing method is the discontinuity of LER pattern in the interface between parallel basic width devices which have been used to make a wider device using the slicing method. Since RDD is dominant source of variability in 35nm gate length devices as discussed previously, this discontinuity of LER pattern plays a secondary role and results in a very small error in generating wider devices using the slicing method.

Figure 6.6 illustrates the correlation of typical MOSFET figures of merit by means of scatter plots. It is clear that the correlations are maintained with a reduced amount of variability. For instance, total spread of leakage current is reduced from more than 3 orders of magnitude for 35nm width devices to two orders of magnitude for 280nm width devices. This reduction in the spread of I_{off} corresponds to reduction in the spread of V_{th} and I_{on} as shown in Fig. 6.6.

6.1.3 Impact of Width on SCM parameters

In order to investigate the impact of the transistor width on the statistical compact model parameters, the statistical compact model parameter extraction strategy developed in chapter 4 has been applied to the data obtained from atomistic simulation or slicing of different width devices. The mean and SD of four most sensitive parameters selected from the statistical parameter set were plotted as a function of the channel width in Figure 6.7.

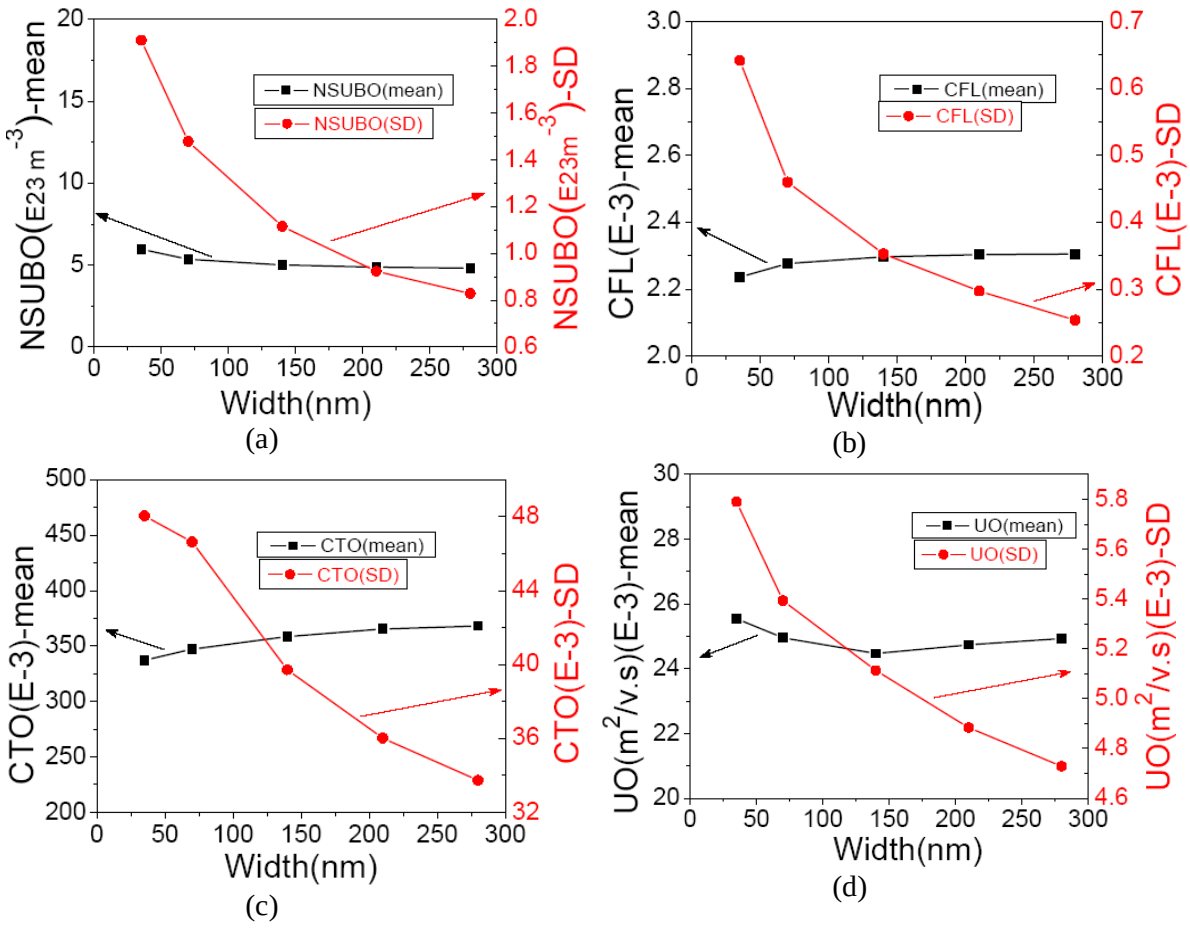


Figure 6.7: Mean and SD of typical SCM parameters versus width; (a)-NSUBO, (b)-CFL, (c)-CTO, (d)-UO.

As expected, the standard deviations decrease with the increase of the channel width but the mean values remain almost constant. This is consistent with the trends observed in respect of the MOSFET figures of merit and discussed in previous sub-section.

In Figure 6.7(a), the mean and SD of parameter NSUBO is plotted versus the channel width. The mean of this parameter decreases by a factor of 20% in devices with 280nm channel width compared to the basic width transistors while a reduction of 60% is observed for the SD of this parameter in the corresponding devices. Figure 6.7(b) plots the same trends for the parameter CFL. There is an increase of less than 3% for the mean value, and a decrease of 60% for SD of this parameter in the 280nm channel width devices compared to the basic width devices. Figure 6.7(c) illustrates the statistical trends of parameter CTO. There is an increase of 10% for the mean value, and a decrease of 30% for SD of this parameter in the 280nm width devices compared to the basic width devices.

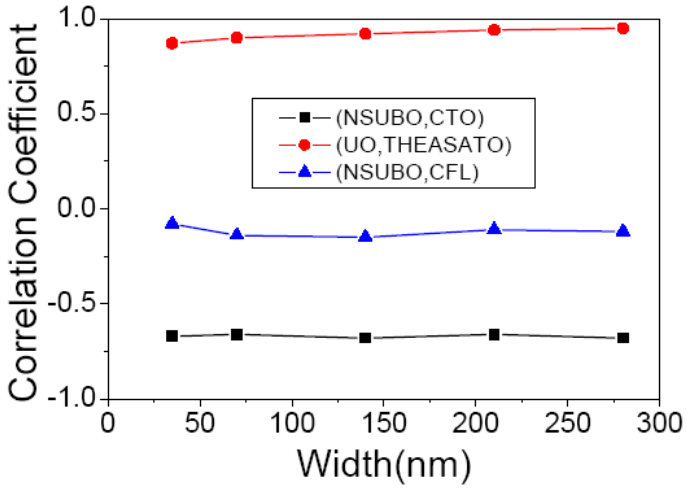


Figure 6.8: Correlation coefficient of three typical pairs of statistical parameter versus width.

The dependence of the mean and SD of parameter UO versus the channel width are shown in Figure 6.7(d). Maximum fluctuation of the mean of this parameter is 5% for different width devices compared with basic width. There is also 18% reduction in the SD of this parameter in 280nm width devices in respect to basic width devices.

It is important to investigate the correlation between the extracted parameters as a function of the channel width. Figure 6.8 illustrates the correlation coefficient of three typical pairs of parameters. Two pairs (NSUBO,CTO) and (UO,THESATO) were significantly correlated as discussed in chapter 5 and the plot shows that the correlations are maintained for different width devices. The pair (NSUBO,CFL) is an example of uncorrelated parameters and the lack of correlation is maintained with the change of width.

The RMS error of the statistical parameter extraction is another important factor which has statistical trend as a function of the channel width. The errors are calculated for each device width in respect to atomistic simulation results and based on the definition given in chapter 4. Figure 6.9 shows the trend for the mean and SD of the RMS error versus the channel width. It is clear that the mean values remain constant, equal to 2.3%, but the SDs slightly decrease with the increase of the channel width. The SD of the error is reduced by 13% in devices with 280nm channel width compared to the 35nm basic width devices. This is a result of reduced statistical variability for wider devices, as shown in Figure 6.1.

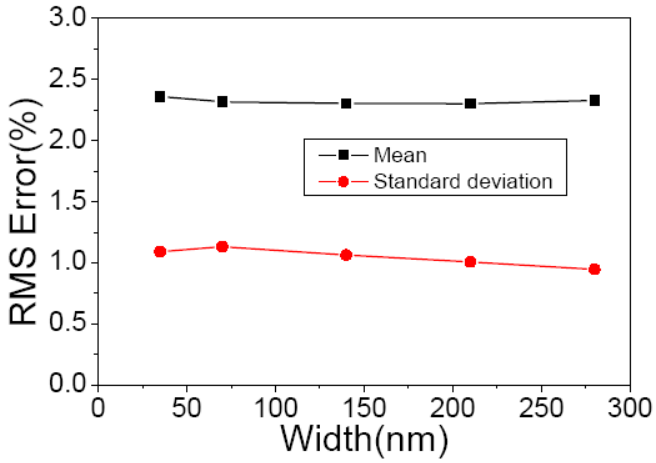
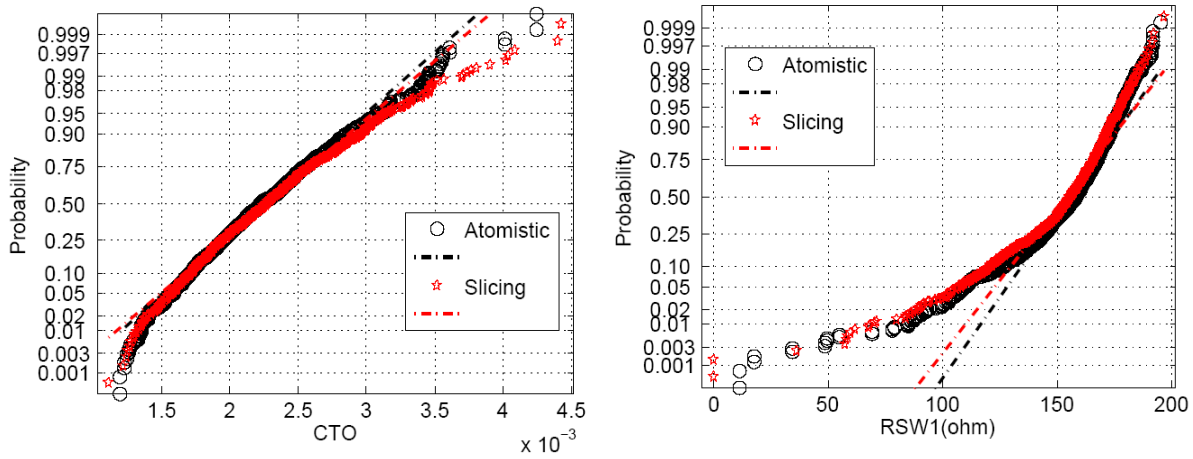


Figure 6.9: Mean and SD of RMS error for statistical parameter extraction of different width devices.

6.1.4 Accuracy of SCM Parameters Using Slicing Method

In the previous section it was shown that using the slicing method to produce larger devices leads to a very small error (less than 2%) in the mean and SD of MOSFET figures of merit. However, it is important to evaluate the impact of this error on SCM parameters extracted using the slicing method. The statistical parameter extraction has been carried out twice: first based on the data obtained from atomistic simulations as discussed in the previous sub-section and second, based on the data obtained from slicing of wider devices into basic width devices. The same method and strategy which was discussed in chapter 4 has been followed and the uniform 35nm gate length model parameters were used as initial values for both statistical parameter extractions. Figure 6.10 shows the Q-Q plots of two typical parameters extracted from atomistic data and slicing method for 70nm channel width devices. There is a close match between mean and SD of these parameters. Moreover, the tail of parameters extracted from the slicing method follows the tail of the parameters extracted from atomistic simulation results with a discrepancy for few devices in the tails. There are more atomistic devices than sliced devices on the lower tail of *RSW1*. For *CTO*, the number of sliced devices is larger than number of atomistic devices in the upper tail. For parameter *CTO*, the relative errors of the mean and SD of the slicing method in respect to atomistic results are 3% and 6%. The mean and SD errors associated with slicing method for the parameter *RSW1* are 2% and 7% respectively.



(a) (b)

Figure 6.10: Q-Q plots of two typical parameters extracted from atomistic simulation results and slicing method of 70nm width devices; (a)-CTO, (b)-RSW1.

To evaluate the impact of the small discrepancy in the parameters tails and SDs on circuit simulation accuracy, we have simulated a CMOS inverter with 70nm width n-channel MOSFET and 140nm width p-channel MOSFET. Apart from the devices width, the other specifications of the inverter are identical to the inverter simulated in chapter 4. Two SCM libraries has been used for the simulations: one extracted from atomistic simulation results of 70nm width devices and the other one extracted from slicing of 70nm devices into two parallel basic width devices. No-load has been included in the simulations to highlight the impact of SV on the inverter delay and the input pulse with rise/fall time of 20psec has been assumed. Figure 6.11 shows the inverter delay distribution comparing between simulations obtained using atomistic SCM library and slicing method SCM library. There is a close match in the tails of distributions and the mean and SD errors caused by using of slicing SCM in respect to atomistic SCM are 1.4% and 2.6%, respectively.

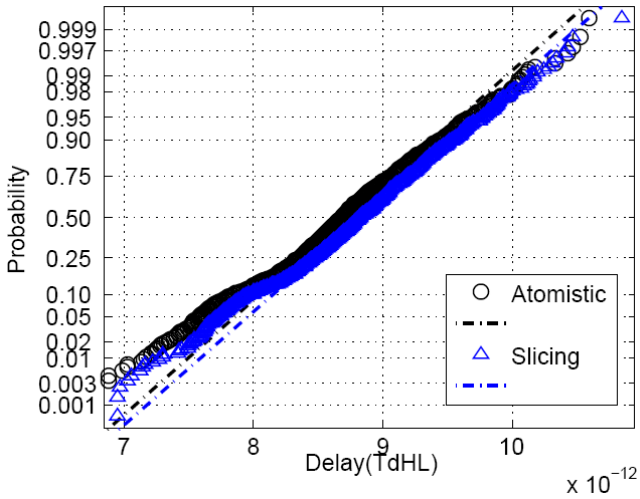


Figure 6.11: Q-Q plot of the inverter delay (T_{dLH}) using SCM libraries from atomistic simulations and slicing method of 70nm width devices.

6.2 Statistical Modeling of Length Dependence

The study on different length devices in relation to statistical compact model extraction is limited to a set of close length MOSFETs because it is not possible to extract a universal statistical SPICE modelcard for devices with any length. In an ideal MOSFET without narrow width effects, a change in the device width does not affect the internal electric fields and therefore the device current per unit width will not change. However, any change in the device length will result in the modification of the device electrostatic property and therefore the device characteristics will be changed. In order to study the impact of the channel length on the statistical compact model parameters, devices with the lengths of 30nm, 35nm and 40nm have been chosen while all of them have a constant width of 35nm. The design of uniform 30nm and 40nm gate length devices has been carried out using the same TCAD simulator and process simulation steps of the template 35nm gate length device as mentioned in chapter 3. The only difference is in the gate patterning step which uses a larger mask for 40nm gate length device and a smaller one for 30nm gate length device, compared to 35nm gate length template device. Figure 6.12 illustrates the doping profile of uniform 30nm and 40nm devices based on the process simulation steps of the uniform 35nm template device.

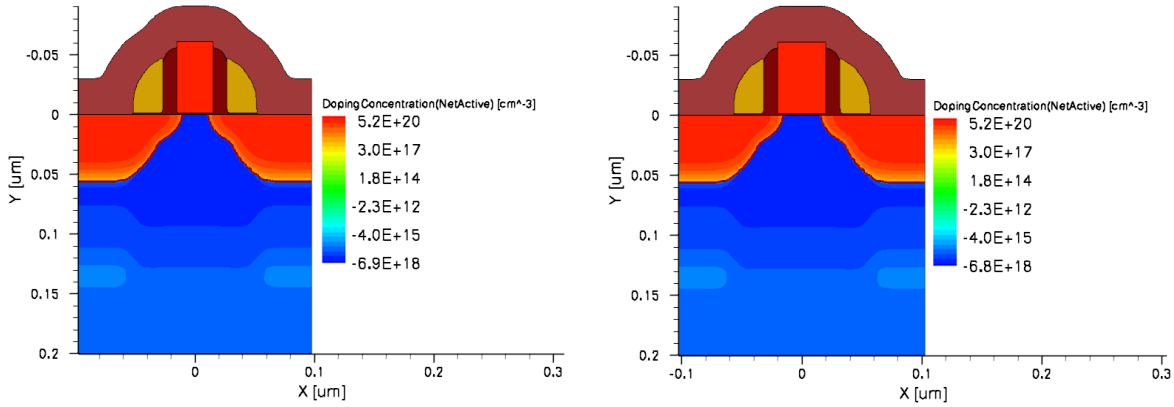


Figure 6.12: Doping profile of different gate length uniform n-channel MOSFETs designed based on the process simulation steps of template 35nm gate length device, (a)-30nm gate length device, (b)-40nm gate length device.

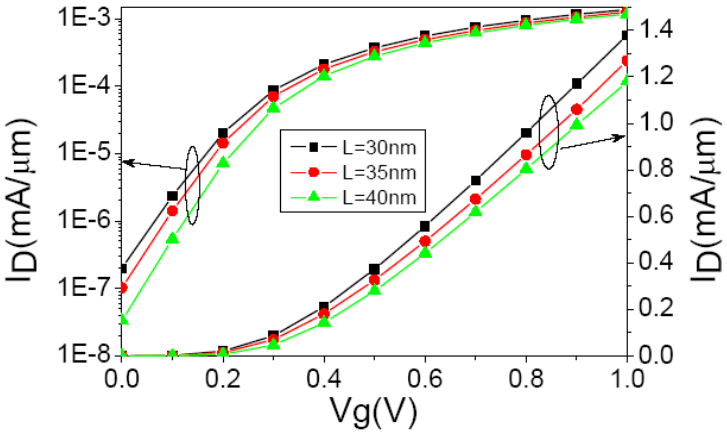
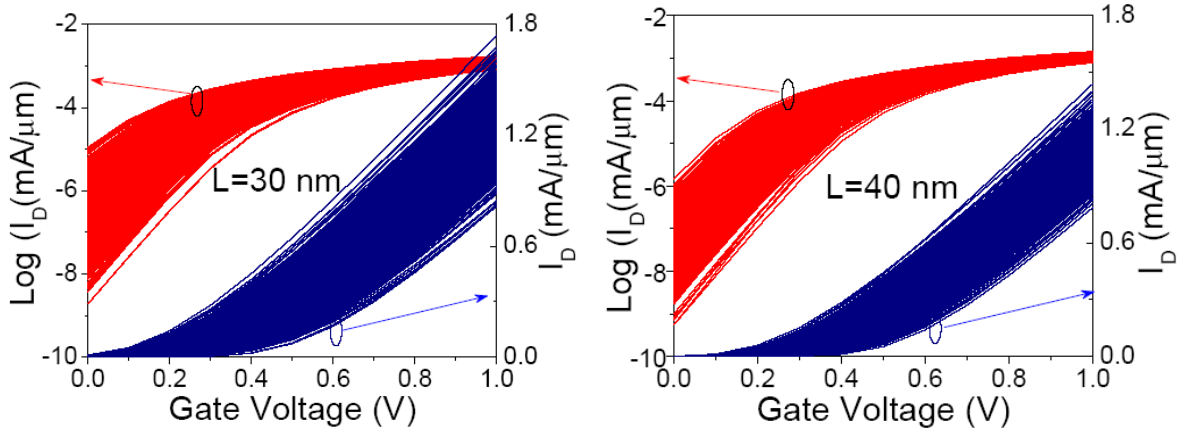


Figure 6.13: I_d-V_g characteristics of uniform 30nm, 35nm and 40nm gate length devices.

Figure 6.13 compares the I_d-V_g characteristics of 30nm, 35nm and 40nm gate length devices in the same plot, the reduction of gate length improves the drive current, but in the expense of increasing the leakage current.

Figure 6.14 shows I_d-V_g characteristics obtained from atomistic simulation of 30nm and 40nm gate length MOSFETs subject to different sources of statistical variability. We omitted 35nm gate length device from Figure 6.14 as it is the base of different width/length study and it was already shown in Figure 6.1. The impact of statistical variability on the device characteristics is reduced for larger gate length devices. The aim is to characterize this impact on the statistical compact model parameters which is discussed in the next section.



(a) (b)

Figure 6.14: Atomistic simulation of SV for devices having different widths of, (a)-30nm, (b)-40nm.

6.2.1 Impact of Length on SCM parameters

Prior to statistical parameter extraction of different length devices, it is necessary to provide a SPICE modelcard capable of describing different length device characteristics with just one parameter set. This task is accomplished with the ability of 'batch or global extraction' of PSP as described in chapter 3. For different width study in previous section, the uniform 35nm modelcard was used to set initial conditions of statistical parameter extraction for all devices of different width but for different length study, using 35nm modelcard is not appropriate and a global modelcard arising from the batch extraction of different length devices of 30nm, 35nm and 40nm will be used for setting the initial conditions of statistical parameter extraction. Since the resulting global modelcard has non-zero components of length-dependent factors for compact model parameters, it is more appropriate to use length-dependent components of seven important parameters in the statistical parameter extraction procedure. Therefore, we have used *CTL*, *CSL* and *THESATL* instead of *CTO*, *CSO* and *THESATO*. The former are the length dependent component of the parameter while the latter are global parameters as used in chapter 4. The other parameters of 7-parameter set include *NSUBO*, *CFL*, *UO* and *RSW1* do not have separate length dependent components. Figure 6.15 illustrates the trend of typical parameter standard deviations versus transistor length. The standard deviations are decreasing for longer devices and roughly follow linear trend.

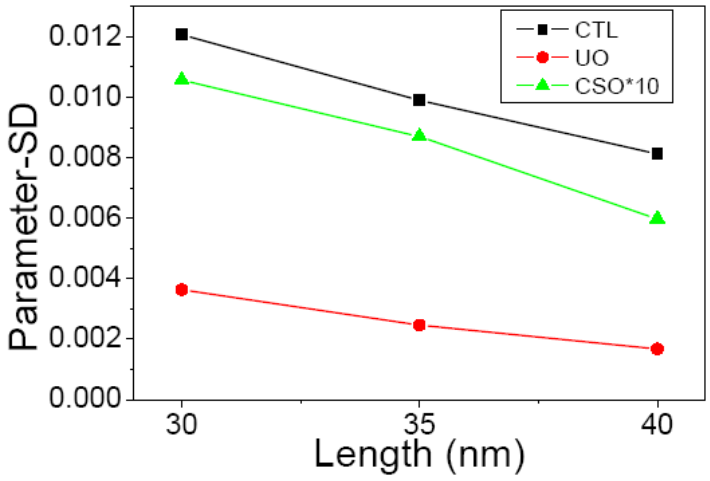


Figure 6.15: Standard deviation of typical parameters versus length obtained from statistical parameter extraction of different length devices.

6.3 Parameter Generation for Fractional Width

It is common in circuit design to use non integer multiplies of basic width devices, which are called non-integer width ratio devices. For instance, using a width of 49nm in our study of different width devices in a 35nm gate length technology is equal to a width of 1.4 times of square devices having the basic width of the 35nm transistors. Proposing slicing method for non-integer width ratio devices would be impossible due to lack of statistical information about devices with less than basic width in the same technology. One valid way to obtain statistical data of these devices is to use atomistic simulation. However, the atomistic simulations are width specific and running the atomistic code on the cluster for many fractional width devices with large number of samples is computationally expensive. Thus, we are looking for easier ways to predict the variability behavior of non-integer width ratio devices using the accurate parameter generation strategies proposed in the previous chapter.

Linear interpolation of statistical properties has been used to generate parameters of fractional width devices. The main statistical properties of parameters are the four moments of the statistical distribution (mean, variance, skewness and kurtosis) for each parameter and the correlation matrix of parameters. For a set of arbitrary non-integer width ratio devices, the linear interpolation of the statistical properties of parameters can be accomplished based on the linear interpolation of statistical properties of two integer width

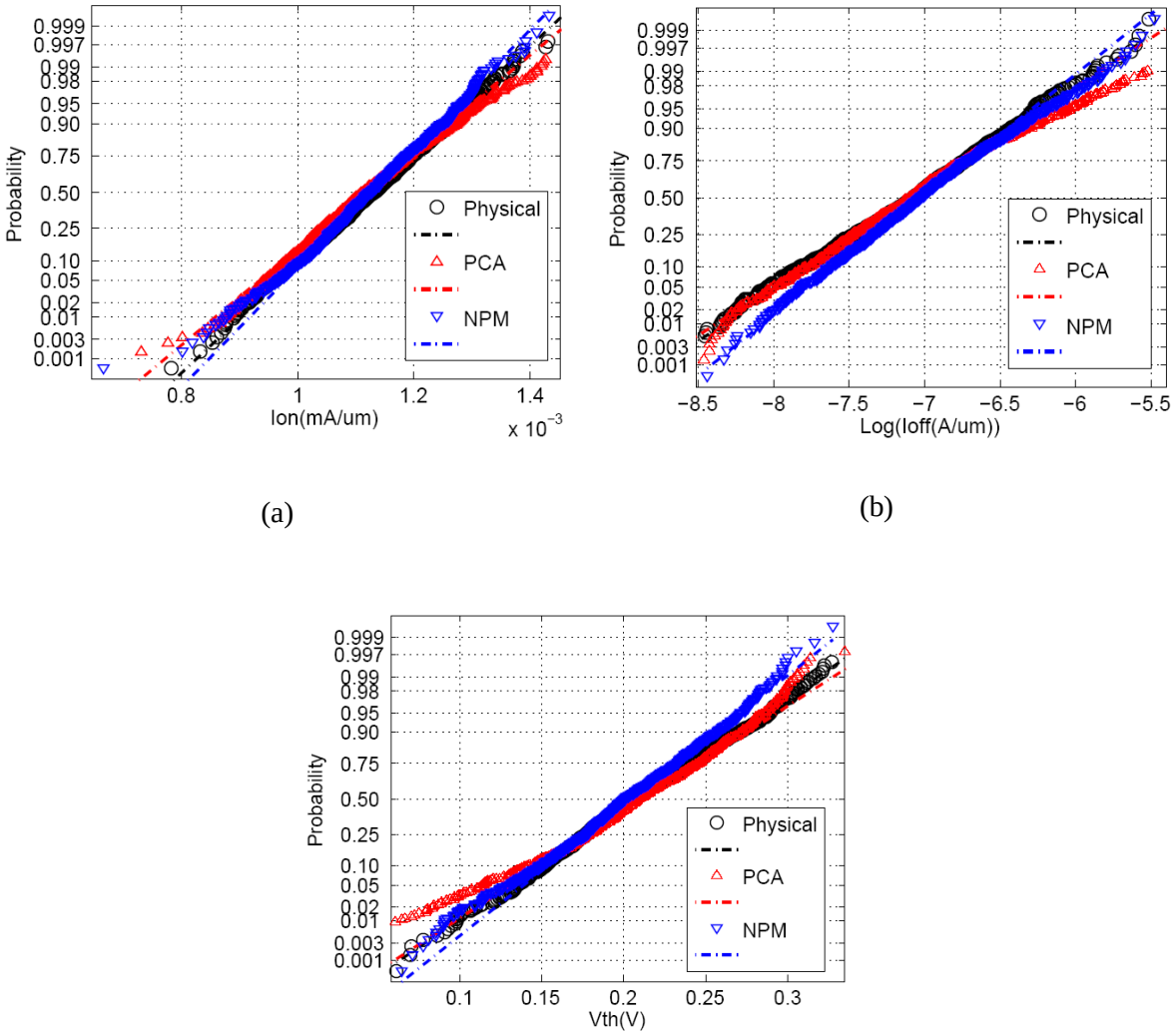
neighboring devices. For instance, the interpolation for devices with width of 1.5 times of basic width device (W1.5) is based on averaging between statistical properties of 'basic width devices' (W1) and 'two times of basic width' (W2) devices. This linear interpolation scheme can be easily extended for any fractional width devices with arbitrary width W between W1 and W2, as follows:

$$p_{i,W} = \frac{p_{i,W2} - p_{i,W1}}{W2 - W1} W + p_{i,W1} \quad ; i = 1,2,3,4 \tag{6.2}$$

where $p_{i,W1}$ and $p_{i,W2}$ represent one of 4 moments of a typical parameter p for devices with width of W1 (35nm) and width of W2 (70nm), respectively. Equation (6.3) can be also used to interpolate correlation matrix of fractional width devices based on the known correlation matrix of W1 and W2 devices.

Here, we have assessed the accuracy of this method using NPM and PCA parameter generation strategies. For PCA, only first two moments of parameter distribution have been used in the interpolation while for NPM all of 4 moments have been used. The eigenvalues and eigenvectors of the interpolated covariance matrix have been used to generate PCA parameters. Since the generated PCA parameters have the mean of zero and variance of 1, the de-normalization of each parameter has been carried out using the interpolated mean and standard deviation of that parameter.

The parameter generation techniques for non-integer width ratio devices need to be assessed in playback of MOSFET figures of merit. We have carried out the atomistic simulation of W1.5 devices to evaluate the accuracy of simulated MOSFET figures of merit using the linear interpolation of statistical properties between W1 and W2. Figure 6.16 illustrates the Q-Q plots of the simulated I_{on}, I_{off} and V_{th} of devices using generated parameters with linear interpolation strategies.



(a) (b)

(c)

Figure 6.16: Q-Q plots of I_{on}, I_{off} and V_{th} of fractional width devices (W1.5) compared between PCA, NPM and Physical approaches. (V_d=1.0v)

For I_{on} plot in Figure 6.16(a), both PCA and NPM provide a good match in respect to physical simulations. NPM provides slightly closer match in upper and lower tails of I_{on}. For I_{off} and V_{th}in Figures 6.16(b,c), there is a tradeoff between PCA and NPM. A better match between NPM and physical results is considered in upper tail of I_{off} while PCA matches closer to physical results in the lower tail of I_{off}. The opposite tradeoff can be seen in the V_{th} plot. NPM matches closer the physical results in lower tail while PCA is closer to physical results in upper tail.

6.4 Summary

In this chapter the statistical compact model strategies were extended to different width/length transistors. As expected, the impact of statistical variability becomes less important in wider/longer transistors. The characterization of this impact in statistical compact models was the main focus of research in this chapter.

Two approaches were used to simulate the impact of SV on wider devices: Atomistic simulations and slicing method. Atomistic simulations of different width devices were carried out to obtain the most accurate results. The slicing method was introduced using parallel basic width devices to make a wider device. There is an error in using this method which is caused by discontinuity of LER pattern in the interface between basic width devices. However, after simulating MOSFET figures of merit for wide devices in both approaches, it was clear that the error of mean and SD is less than 2% in respect to atomistic simulation results. A study of the mean and SD of MOSFET figures of merit showed that the mean values stay almost constant with the channel width while their SDs monotonically decrease with the increase of the channel width. It was shown that the correlation between figures of merit remains independent of the channel width.

Statistical compact model extraction was carried out to study the impact of the channel width on parameter distribution trends. It was observed that the parameters SDs decrease with the increase of the channel width. The RMS error of the statistical parameter extraction also was characterized as a function of the channel width. The accuracy of the statistical compact model parameters extracted using the slicing method was assessed using simulation of a CMOS inverter. It shows 1.4% and 2.6% errors in the mean and the SD of rise time delay.

Investigation of the impact of the channel length on the parameters statistical trend was carried out on a range of different length MOSFETs. A parameter generation strategy for non-integer width ratio devices was also developed. Linear interpolations of statistical properties of neighboring devices with integer width ratio were used to generate distribution of parameters for non-integer width ratio devices. The approach requires the interpolation of the first 4 moments of the parameter distributions for NPM and the first 2 moments of parameter distribution for PCA as well as the interpolation of the correlation

coefficient matrix for both NPM and PCA approaches. Assessment of the figures of merit for 1.5 width ratio device showed that I_{on} is better approximated with NPM while there is a tradeoff between NPM and PCA in predicting I_{off} and V_{th}.

Chapter 7

Conclusion

The aim of research presented in this thesis was to develop statistical compact model strategy for nano-scaled CMOS devices for the surface potential compact model, PSP. Statistical compact models are essential part of variability aware design in contemporary and next generation ultra scaled CMOS technologies and hence, we have developed tools and strategies to account for the impact of statistical variability on circuit performance and yield.

RDD, LER and PGG are important sources of statistical variability which result in statistical fluctuations in the transistor characteristics. These fluctuations arise from discreteness of charge and granularity of matter at microscopic level that in turn leads to variability in operational characteristics of devices in macroscopic level. For instance, exact location of dopant atoms in the channel region is not identical in two fabricated devices with the same geometry in a given technology and this fact results in different values of threshold voltage, leakage current and drive current for two macroscopically identical devices. Predictive simulation of statistical variability in device level has been carried out using an ‘atomistic’ drift-diffusion simulator which takes into account quantum corrections based on the density-gradient formulism.

PSP, an advanced surface potential compact model has been used as a basis to develop statistical compact model strategy. There are several advantages in using PSP compared to traditional compact models like BSIM. For instance, it has physical expressions for the model parameters, presents symmetrical trans-capacitance components and smooth transition between weak and strong inversion regions of the transistor operation.

Chapter 2 begins with a survey of the challenges in scaling of nano CMOS devices. The classification of variability into systematic and statistical types is then reviewed followed by reviewing the different sources of systematic and statistical variability. Simulation techniques for deep sub-micrometer MOSFETs with emphasis on drift diffusion techniques with quantum corrections are given. This is due to the fact that the drift diffusion simulations exhibit higher computational efficiency compared with Monte Carlo or quantum mechanical approaches. Incorporation of RDD, LER and PGG into the GSS 'atomistic' simulator GARAND which allows predictive simulation of statistical variability is then discussed. This chapter is completed with a review on existing statistical compact modelling approaches including mismatch models, corner models, numerical and analytical models. Mismatch models are used for predicting variance of electrical performance parameters. Corner models are worst-case or best-case models used in circuit simulation. Numerical models like PCA have been used in other chapters of this thesis to produce ensemble of statistical parameters.

In Chapter 3, after a description of the transistor models, the basic elements of surface potential compact model PSP are reviewed. The corresponding surface potential equation and drain current expression have important parameters which are used in statistical parameter extraction procedure. Parameter extraction of a template 35nm gate length MOSFET has been carried out in DC and AC modes to provide a complete set of uniform PSP model. Accuracy of both parts of the model has been evaluated in comparison with TCAD simulations. For the DC part, the RMS error of PSP model remains less than 3% in respect to TCAD simulations. For AC part, nine independent trans-capacitance components were simulated and it was concluded that the most accurately fitted components are associated with the bulk terminal. The gate components were less accurate due to use of Ward-Dutton charge partitioning scheme which is perfect only for uniformly doped devices. The error in the drain capacitance components was noticeable particularly in the strong inversion region. Finally, the accuracy of transient time SPICE simulations was evaluated using extracted PSP modelcard in a CMOS inverter and the result was compared with TCAD simulation of the same inverter. Different capacitive load conditions were employed for the inverter and the discrepancies between SPICE and TCAD transient time simulations were found to be in line with the error in trans-capacitance components. Finally, a compensation scheme was introduced to reduce the propagation delay time errors of the inverter.

Chapter 4 described the procedures needed to develop an accurate statistical compact model based on PSP. Atomistic simulations were performed to obtain the device characteristics under the influence of combined sources of variability including RDD, LER and PGG for an ensemble of 1000 microscopically different 35nm gate length MOSFETs. First order sensitivity analysis of model parameters is carried out to identify important parameters in the statistical parameter extraction procedure. A method and strategy is then proposed to extract 7-parameter statistical set. Excellent statistical RMS error mean of 2.36% and RMS error standard deviation of 1.09% were achieved. Strong correlation between the SCM parameters and the electrical figures of merit indicates that the physical meaning of compact model parameters is maintained during the statistical extraction. The extracted SCM parameters were used to simulate devices under different drain bias conditions. The results of the simulations were used to extract MOSFET figures of merit (I_{on}, I_{off}, V_{th}, *SS and DIBL*) and compare them with the values obtained from the atomistic simulations. The accuracy of these figures of merits was also evaluated for SCM sets using different number of parameters. As a result, the minimum required numbers of parameters needed to reproduce the mean and the standard deviations of I_{on}, I_{off} and V_{th} were estimated to be 7, 4 and 5 parameters, respectively. In order to assess the impact of the statistical variability on circuit operation, we have studied timing and power variability of a CMOS inverter. A Monte Carlo circuit simulation scheme was introduced to simulate CMOS inverter using NMOS and PMOS statistical compact model libraries. The statistical simulations for the most accurate 7 parameter set showed that the normalized delay variability (σ/μ) can increase two times, and (σ/μ) of energy dissipation variability can increase by more than 7 times, for a fan-out of 2 when the input rise/fall time increased from 2ps to 50ps. Furthermore, the degree of correlations between delay and energy was dependant on both output load and input slew rate.

Chapter 5 proposed novel parameter generation techniques. The distribution of each parameter and the correlation between pairs of parameters were investigated from directly extracted statistical parameters to produce an input for subsequent parameter generation techniques. A Box-Cox transformation was also introduced to convert non-normally distributed parameters to normal distributions. In the Gaussian parameter generation method, correlations between parameters were ignored and hence, it introduced a significant error in the variance of I_{on} distributions. Generation of parameters based on their principal component analysis (PCA) considering the correlation between the

parameters greatly improves the quality of results, although it still introduces error into the variance of I_{on} distribution since the parameter distributions are considered to be normal. Scatter plots indicated that the application of Box-Cox transformation distorts the parameter correlations and hence, more error was observed in the variance of MOSFET figures of merits when Box-Cox is followed by PCA analysis compared to the PCA analysis alone. Both the Gaussian and the PCA methods present a good match in distribution of $\mathrm{Log}(I_{off})$ and V_{th} except for the upper tail of $\mathrm{Log}(I_{off})$ and the lower tail of V_{th}. The nonlinear power method (NPM) was introduced as a moment matching technique which replicates both parameter distributions and correlations. The evaluation of the device electrical figures of merit showed that this method is capable of preserving tails of the distributions with a high accuracy. It also maintained a high degree of matching in statistical timing and power simulation of CMOS inverter and can be considered as the most accurate parameter generation technique among the proposed techniques.

Chapter 6 presents an extension of the PSP statistical compact model strategy to different width/length devices. Statistical atomistic simulation of different width/length devices was carried out as the most accurate way of obtaining I_d-V_g characteristics for different geometry devices. However, with the gate slicing method of generating wider devices as parallel combination of basic width devices, we were able to reproduce very accurate MOSFET figures of merit for multi-width ratio devices with a very high computational efficiency. Evaluating the mean and SD of the figures of merit showed that the slicing method gives maximum error of 2% in respect to the atomistic simulations. The accuracy of SCM parameters using slicing method was assessed by the simulation of a CMOS inverter. The errors are 1.4% and 2.6%, respectively, for the mean and SD of rise time delay. The trends of statistical parameters were investigated versus width/length. There was small variation for the mean of each parameter while the behavior for SD showed a monotonic decay. This was due to reduced effect of statistical variability for wider/longer devices. A parameter generation strategy for non-integer width ratio devices was developed. Assessment of figures of merit for a device of width ratio of 1.5 showed that I_{on} is better approximated with NPM while there is a tradeoff between NPM and PCA in predicting I_{off} and V_{th}.

7.1 Future Work

The research presented in this work can be extended in several directions. The areas to future research comprise statistical compact models for new device structures, efficient parameter generation techniques and their incorporation in statistical modeling of multi-width and multi-length devices.

Although bulk MOSFETs will be in use in the next few years and bulk scaling will continue toward 16nm CMOS technology generation at 2016, as predicted by ITRS, the focus for extreme scaling has moved to new device architectures like thin body Silicon-on-Insulator (SOI) transistors, Multi-Gate transistors and others, which exhibit better scaling properties [153,154,155,156]. This will provide a high demand for development of statistical compact model strategies for these new transistor architectures.

To find other efficient parameter generation techniques capable of reproducing distribution of directly extracted parameters and their correlation is an interesting and important research area. These methods can be compared with the nonlinear power method in terms of computational efficiency and their accuracy in statistical circuit simulations. To develop accurate but computationally efficient statistical parameter generation technologies for arbitrary width and length devices is another interesting but challenging research area.

Bibliography

[1] J.B. Shyu, G.C. Temes and F.Krummenacher, "Random Error Effects in Matched MOS Capacitors and Current Sources", *IEEE Journal of Solid State Circuits*, Vol. 19, No. 6, pp. 948-955, 1984.

[2] K. Takeuchi and M. Hane, "Statistical Compact Model Parameter Extraction by Direct Fitting to Variations", *IEEE Transactions on Electron Devices*, Vol. 55, No. 6, pp. 1487-1493, 2008.

[3] M. Miura-Mattausch, "Compact Modeling of the MOSFET Performance Distribution for Statistical Circuit Simulation", *Workshop on Simulation and Characterization of Statistical CMOS Variability and Reliability, SISPAD 2010*, 6-8 Sep, Bologna, Italy.

[4] M.J.M. Pelgrom, A.C.J, Duinmaijer and A.P.G. Welbers, "Matching Properties of MOS Transistors", *IEEE Journal of Solid State Circuits*, Vol. 24, No. 5, pp. 1433-1439, 1989.

[5] L. Chung-Hsun, M.V. Dunga, L. Darsen, A.M. Niknejad and C. Hu, "Statistical Compact Modeling of Variations in Nano MOSFETs", *Proceedings of International Symposium on VLSI Technology, Systems and Applications (VLSI_TSA 08)*, IEEE Press, pp. 165-166, 2008.

[6] Y. Cao, C.C. McAndrew, "MOSFET Modeling for 45nm and Beyond", *Proceedings of IEEE/ACM International Conference on Computer Aided Design* (ICCAD 07), IEEE Press, pp. 638-643, 2007.

[7] E. Malavasi, S. Zanella, J. Uschersohn, M. Misheloff and C. Guardiani, "Impact Analysis of Process Variability on Digital Circuits with Performance Limited Yield", *IEEE International Workshop on Statistical Methodology*, pp. 60-63, 2001.

[8] M. Merrett, P. Asenov, Y. Wang, M. Zwolinski, D. Reid, C. Millar, S. Roy, Z.

Liu, S. Furber and A. Asenov, "Modeling Circuit Performance Variations due to Statistical Variability: Monte Carlo Static Timing Analysis", *Proceedings of the Conference on Design, Automation and Test in Europe* (DATE 11), pp. 1-4, March 2011.

[9] C. Millar, D. Reid, G. Roy, S. Roy, A. Asenov, "Accurate Statistical Description of Random Dopant-Induced Threshold Voltage Variability", *IEEE Electron Device Letters*, Vol. 29, No. 8, Aug. 2008.

[10] G. Gildenblat, X. Li, W. Wu, H. Wang, A. Jha, R. Langevelde, G.D.J. Smit, A.J.Scholten and D.B.M. Klaassen, "PSP: An Advanced Surface-Potential-Based MOSFET Model for Circuit Simulation", *IEEE Transactions on Electron Devices*, Vol. 53, No. 9, pp. 1979-1993, Sep. 2006.

[11] BSIM model manual available from: http://www-device.eecs.berkeley.edu/bsim3/

[12] A.J. Scholten, G.D.J. Smit, B.A. De Vries, L.F. Tiemeijer, J.A. Croon, D.B.M. Klaassen, R. Langevelde, X. Li, W. Wu and G. Gildenblat, "The new CMC standard compact MOS model PSP: advantages for RF application", in *Radio Frequency Integrated Circuits Symposium* (RFIC 08), pp. 247-250, 2008.

[13] A.J. Scholten, G.D.J. Smit, DBM. Klaassen, R. Langevelde, G. Gildenblat, W. Wu and X. Li, "The PSP Compact MOSFET Model: an update", in *MOS_AK workshop*, Eindhoven, 2008. (available from http://www.mos-ak.org/eindhoven/)

[14] H.S.P. Wong, D.J. Frank, P. Solomon, C. Wann and J. Welser, "Nanoscale CMOS", *Proceedings of the IEEE*, Vol. 87, No. 4, pp. 537-570, 1999.

[15] G.E. Moore, "Cramming More Components onto Integrated Circuits", *Electronics*, 38:114-117, 1965.

[16] G.E. Moore, "Progress in Digital Integrated Electronics", *International Electron Device Meeting Technical Digest*, pages 11-13, Washington DC 1975, Dec 1-3.

[17] Y.B. Kim, "Challenges for Nanoscale MOSFETs and Emerging Nanoelectronics", *Transactions on Electrical and Electronic Materials*, Vol. 11, No. 3, pp. 93-105, June 2010.

[18] P.J. Wright, K.C. Saraswat, "Thickness limitations of SiO_2 gate dielectrics for MOS ULSI", *IEEE Transactions on Electron Devices*, Vol. 37, No. 8, pp. 1884-1892, 1990.

[19] P.E. Gusev, V. Narayanan, M.M. Frank, "advanced high-k dielectric stacks with polySi and metal gates: recent progress and current challenges", *IBM Journal of Research and development*, Vol. 50, Issue 4/5, pp. 387-410, July 2006.

[20] G.M.T. Wong, "An Investigation of the Work Function of Metal Gate Electrodes for Advanced CMOS Applications", *Ph.D. dissertation*, Stanford University, 2008.

[21] M. Stockinger, "Optimization of Ultra-Low-Power CMOS Transistors", *Ph.D. dissertation*, Institute for Microelectronics, Vienna, Austria, 2000.

[22] T. Krishnamohan, "Physics and Technology of High Mobility, Strained Germanium Channel, Heterostructure MOSFETs", *Ph.D. dissertation*, Stanford University, 2006.

[23] S.K. Saha, "Modeling Process Variability in Scaled CMOS Technology", *IEEE Design and Test of Computers*, Vol. 27, No. 2, pp. 8-16, March/April 2010.

[24] A. Asenov, B. Cheng, "Modeling and Simulation of Statistical Variability in Nanometer CMOS Technologies", In: H. Casier, M. Steyaert, A.H.M. van Roermund, (eds.) "*Analog Circuit Design: Robust Design, Sigma Delta Converters, RFID*", Springer, pp. 17-33, 2011.

[25] K. Bernstein, D.J. Frank, A.E. Gattiker, W. Haensch, B. J. Li, S. R. Nassif, E. J. Nowak, D.J. Pearson, N. J. Rohrer , "High Performance CMOS Variability in the 65-nm regime and beyond", *IBM Journal of Research and development* , Vol. 50, Issue 4/5, pp. 433-449, July 2006.

[26] A.R. Brown, G. Roy, A. Asenov, "Poly-Si Gate Related Variability in decananometer MOSFETs with Conventional Architecture", *IEEE Transactions on Electron Devices*, Vol. 54, No. 11, pp. 3056-3063, Nov. 2007.

[27] K.R. Lakshmikumar, R.A. Hadaway, M.A. Copeland, "Characterization and Modelling of Mismatch for Analog MOS Integrated Circuits", *IEEE Journal of Solid State Circuits*, pp. 1057-1066, 1986.

[28] D. Sylvester, K. Agarwal, S. Shah, "Variability in nanometer CMOS: Impact, analysis and minimization", *INTEGRATION, the VLSI journal*, No. 41, pp. 319-339, 2008.

[29] B. Cheng, S. Roy, A. Asenov, "The Impact of Random Dopant Effects on SRAM Cells", *Proc. 30th European Solid State Circuit Conference (ESSCIRC)*, pp. 219-222, Lueven, 2004.

[30] A. Asenov, "Statistical Nano CMOS Variability and Its Impact on SRAM", In: A. Singhee and R.A. Rutenbar (eds.) "*Extreme Statistics in Nanoscale Memory Design*", Springer, pp. 17-50, 2010.

[31] H. Aikawa, E. Morifuji, T. Sanuki, T. Sawada, S. Kyoh, A. Sakata, M. Ohta, H.

Yoshimura, T. Nakayama, M. Iwai and F. Matsuoka, "Variability aware modeling and characterization in standard cell in 45 nm CMOS with stress enhancement technique", *Symposium on VLSI Technology*, pp. 90-91, Honolulu, 2008.

[32] M. Pelgrom, H. Tuinhout, M. Vertregt, "Modeling of MOS Matching", In: G.Gildenblat (ed.) "*Compact Modeling: Principles, Techniques and Applications*", Springer, pp. 453-490, 2010.

[33] A. Asenov, A.R. Brown, J.H. Davies, S. Kaya, G. Slavcheva, "Simulation of Intrinsic Parameter Fluctuations in Decananaometer and Nanometer-Scale MOSFETs", *IEEE Transactions on Electron Devices*, Vol. 50, No. 9, pp. 1837-1852, September 2003.

[34] K. Patel, "Intrinsic and Systematic Variability in Nanometer CMOS Technologies", *Ph.D. dissertation*, University of California at Berkeley, 2010.

[35] N. Cobb, "Fast Optical and Process Proximity Correction Algorithms for Integrated Circuit Manufacturing", *Ph.D. dissertation*, University of California at Berkeley, 1998.

[36] X. Wang, B. Cheng, S. Roy and A. Asenov, "Simulation of strain induced variability in nMOSFETs", *Proceedings of the 9th International Conference on Ultimate Integration on Silicon (ULIS)*, pp. 89-92, Udine, Italy, March 2008.

[37] L. Ge, V. Adams, K. Loiko, D. Tekleab, X.Z. Bo, M. Foisy, V. Kolagunta, S. Veeraraghavan, "Modeling and simulation of poly-space effects in uniaxially-strained etch stop layer stressors", *IEEE International SOI Conference*, pp. 25-26, 2007.

[38] N. Wils, H.P. Tuinhout, M. Meijer, "Characterisation of STI-edge effects on CMOS Variability", *IEEE Transactions on Semiconductor Manufacturing*, Vol. 22, No. 1, pp.59-65, Feb. 2009.

[39] R.A. Bianchi, G. Bouche, O. Roux-dit-Buisson, "Accurate modeling of trench isolation induced mechanical stress effects on MOSFET electrical performance", In: *Digest. International Electron Devices Meeting*, pp.117-120, 2002.

[40] L.T. Pang, K. Qian, C.J. Spanos, B. Nikolic, "Measurement and analysis of variability in 45-nm strained-Si CMOS technology", *IEEE Journal of Solid State Circuits*, Vol. 44, No. 8, pp. 3056-3063, Nov. 2007.

[41] H.P. Tuinhout, M.J.M. Pelgrom, R. Penning de Vries, M. Vertregt, "Effects of metal coverage on MOSFET Matching", *Technical Digest International Electron Devices Meeting*, pp. 735-739, 1996.

[42] P.G. Drennan, M.L. Kniffin, D.R. Locascio, "Implication of proximity effects for analog design", *IEEE Custom Integrated Circuit Conference*, pp.169-176, 2006.

[43] V. Mehrotra, S.L. Sam, D. Boning, A. Chandrakasan, R. Vallishayee, S. Nassif, "A Methodology for Modeling the Effects of Systematic Within-Die Interconnect and Device Variation on Circuit Performance", *Proceedings of 37th Design Automation Conference*, pp. 172-175, 2000.

[44] D. Boning, S. Nassif, "Models of Process Variations in Device and Interconnect", In: A. Chandrakasan, W.J. Bowhill and F. Fox (eds.) "*Design of High Performance Microprocessor Circuits*", John Wiley, pp. 98-115, 2000.

[45] M. Orshansky, C. Spanos, C. Hu, "Circuit Performance Variability Decomposition", *4th International Workshop on Statistical Metrology*, pp. 10-13, June 1999.

[46] N.A. Drego, "Characterization and Mitigation of Process Variation in Digital Circuits", *Ph.D. dissertation*, Massachusetts Institute of Technology, 2009.

[47] P. Zarkesh-ha, T. Mule, J.D. Meindl, "Characterization and Modeling of Clock Skew with Process Variations", *IEEE Custom Integrated Circuit Conference*, pp. 441-444, 1999.

[48] W. Schemmert, G. Zimmer, "Threshold Voltage Sensitivity of Ion Implanted MOS Transistors due to Process Variations", *Electronic Letters*, Vol. 10, Issue 9, pp. 151-152, May 1974.

[49] Y. Aoki, T. Toyabe, S. Asai, T. Hagiwara, "CASTCAM: A Process Variation Analysis Simulator for MOS LSI's", *IEEE Transactions on Electron Devices*, Vol. 31, No. 10, pp. 1462-1467, Oct. 1984.

[50] K. Kuhn, "Reducing Variations in Advanced Logic Technologies: Approaches to Process and Design for Manufacturability of Nanoscale CMOS", *IEEE International Electron Device Meeting*, pp. 471-474, 2007.

[51] K. Kuhn, C. Kenyon, A. Kornfeld, M. Liu, A. Maheshwari, W.K. Shih, S. Sivakumar, G. Taylor, P. VanDerVoorn, K. Zawadzki, "Managing process variation in Intel's 45-nm CMOS technology", *Intel Technology Journal*, Vol. 12, No. 2, pp. 93-110, Jun. 2008.

[52] A. Asenov, "Random Dopant Induced Threshold Voltage Lowering and Fluctuations in Sub-0.1um MOSFETs: A 3D Atomistic Simulation Study", *IEEE Transactions on Electron Devices*, Vol. 45, No. 12, pp. 2505-2513, Dec. 1998.

[53] F. Adamu-Lema, "Scaling and Intrinsic Parameter Fluctuations in Nano-CMOS

Devices", *Ph.D. dissertation*, University of Glasgow, 2005.

[54] A. Asenov, S. Kaya, A.R. Brown, "Intrinsic Parameter Fluctuations in Decananometer MOSFETs Introduced by Gate Line Edge Roughness", *IEEE Transactions on Electron Devices*, Vol. 50, No. 5, pp. 1254-1260, May 2003.

[55] G. Roy, "Simulation of Intrinsic Parameter Fluctuations in Nano-CMOS Devices", *Ph.D. dissertation*, University of Glasgow, 2005.

[56] S. Uma, A.D. McConnell, M. Asheghi, K. Kurayabashi, K.E. Goodson, "Temperature dependent thermal conductivity of undoped polycrystalline silicon layers", *International Journal of Thermophysics*, Vol. 22, No. 2, pp. 605-616, Mar. 2001.

[57] A.R. Brown, G. Roy and A. Asenov, "Poly-Si-Gate-Related Variability in Decananometer MOSFETs With Conventional Architecture", *IEEE Transactions on Electron Devices*, Vol. 54, No. 11, pp. 3056-3063, Nov. 2007.

[58] Y. Zhang, J. Li, M. Grubbs, M. Deal, B. Magyari-Kope, B.M. Clemens and Y. Nishi, "Physical model of impact of metal grain work function variability on emerging dual metal gate MOSFETs and its implication for SRAM reliability", *IEDM Technical Digest*, pp. 57-60, Dec. 2009.

[59] A. Asenov, S. Kaya and J.H. Davies, "Intrinsic Threshold Voltage Fluctuations in Decanano MOSFETs Due to Local Oxide Thickness Variations", *IEEE Transactions on Electron Devices*, Vol. 49, No. 1, pp. 112-119, 2002.

[60] A. Cathignol, B. Cheng, D. Chanemougame, A.R. Brown, K. Rochereau, G.Ghibaudo and A. Asenov, "Quantitative Evaluation of Statistical Variability Sources in a 45-nm Technological Node LP N-MOSFET", *IEEE Electron Device Letters*, Vol. 29, No. 6, June 2008.

[61] U. Ravaioli, "Hierarchy of simulation approaches for hot carrier transport in deep submicron devices", *Semiconductor Science and Technology*, Vol. 13, No. 1, pp. 1-10, 1998.

[62] A. Asenov, R. Balasabrumaniam, A.R. Brown and J.H. Davies, "RTS Amplitudes in Decananometer MOSFETs: 3-D Simulation Study", *IEEE Transactions on Electron Devices*, Vol. 50, No. 3, pp. 839-845, Mar. 2003.

[63] D.T. Reid, "Large Scale Simulations of Intrinsic Parameter Fluctuations in Nano-Scale MOSFETs", *Ph.D. dissertation*, University of Glasgow, June 2010.

[64] T. Grasser, T.W. Tang, H. Kosina and S. Selberher, "A review of hydrodynamic and energy transport models for semiconductor device simulation", *Proceedings*

of the IEEE, Vol. 91, No. 2, pp. 251-274, 2003.

[65] D.M. Caughy and R.E. Thomas, "Carrier Mobilities in Silicon Empirically Related to Doping and Field", *Proceedings of the IEEE*, Vol. 55, pp. 2192-2193, 1967.

[66] A.R. Brown, J.R. Watling, G. Roy, C. Riddet, C.L. Alexander, U. Kovac, A. Martinez and A. Asenov, "Use of density gradient quantum corrections in the simulation of statistical variability in MOSFETs", *Journal of Computational Electronics*, Vol. 9, pp. 187-196, 2010.

[67] S. Jallepalli, J. Bude, W.K. Shih, M.R. Pinto, C.M. Maziar and A.F. Tasch, "Electron and Hole Quantization and Their Impact on Deep Submicron Silicon p- and n- MOSFET Characteristics", *IEEE transactions on Electron Devices*, Vol. 44, No. 2, pp. 297-303, Feb. 1997.

[68] M.G. Ancona, "Finite temperature density gradient theory", *Proceedings of IWCE*, pp. 151-154, 1992.

[69] A.R. Brown, G. Roy, "Everything you always wanted to know about drift-diffusion but were too lazy to ask", in DMG (Device Modelling Group) seminars, University of Glasgow, 2010. (Available on request from Andrew Brown, email: Andrew.Brown@glasgow.ac.uk)

[70] G. Roy, A.R. Brown, A. Asenov, S.Roy, "Quantum Aspects of Resolving Discrete Charges in Atomistic Device Simulations", *Journal of Computational Electronics*, Vol. 2, pp. 323-327, 2003.

[71] A. Asenov, A.R. Brown, G. Roy, B. Cheng, C.L. Alexander, C. Riddet, U. Kovac, A. Martinez, N. Seoane and S. Roy, "Simulation of statistical variability in nano-CMOS transistors using drift-diffusion, Monte Carlo and non-equilibrium Green's function techniques," *Journal of Computational Electronics*, Vol. 8, No. 3-4, pp. 349–373, 2009.

[72] J.L. McCreary, "Matching properties, and voltage and temperature dependence of MOS capacitors", *IEEE Journal of* Solid *State Circuits*, pp. 608-616, 1981.

[73] K.R. Lakshmikumar, R.A. Hadaway, M.A. Copeland, "Characterization and Modelling of Mismatch in MOS Transistors for Precision Analog Design", *IEEE Journal of Solid State Circuits*, pp. 1057-1066, 1986.

[74] P.G. Drennan, C.C. McAndrew, "Understanding MOSFET Mismatch for Analog Design", *IEEE Journal of Solid State Circuits*, Vol. 38, No. 3, pp. 450-456, 2003.

[75] K.J. Kuhn, M.D. Giles, D. Becher, P. Kolar, A. Kornfeld, R. Kotlyar, S.T. Ma, A.

Maheshwari, S. Mudanai, “Process Technology Variation”, *IEEE Transactions on Electron Devices*, Vol. 58, No. 8, pp. 2197-2208, Aug. 2011.

[76] T. Mizuno, J.I. Okamura, A. Toriumi, “Experimental Study of Threshold Voltage Fluctuations due to Statistical Variation of Channel Dopant Number in MOSFETs”, *IEEE Transactions on Electron Devices*, Vol. 41, No. 11, pp. 2216-2221, Nov. 1994.

[77] P. Stolk, F. Widdershoven, D. Klaassen, “Modeling Statistical Dopant Fluctuations in MOS Transistors”, *IEEE Transactions on Electron Devices*, Vol. 45, No. 9, pp. 1960-1971, Sep. 1998.

[78] M. Quarantelli, S. Saxena, N. Dragone, J.A. Babcock, C. Hess, S. Minehane, S. Winters, J. Chen, H. Karbasi, C. Guardiani, “Characterization and Modeling of MOSFET Mismatch of a Deep Submicron Technology”, *IEEE International Conference on Microelectronic Test Structures*, Vol. 16, March 2003.

[79] J.A. Croon, M. Rosmeulen, S. Decoutere, W. Sansen, H.E. Maes, “An Easy-to-Use Mismatch Model for the MOS Transistor”, *IEEE Journal of Solid State Circuits*, Vol. 37, No. 8, pp. 1056-1064, Aug. 2002.

[80] P.G. Drennan, CC. McAndrew, “Understanding MOSFET Mismatch for Analog Design”, *IEEE Custom Integrated Circuit Conference*, pp. 449-452, 2002.

[81] C. Galup-Montoro, M.C. Schneider, H. Klimach, A. Arnaud, “A Compact Model of MOSFET Mismatch for Circuit Design”, *IEEE Journal of Solid State Circuits*, Vol. 40, No. 8, pp. 1649-1657, Aug. 2005.

[82] H. Klimach, A. Arnaud, C. Galup-Montoro and M.C. Schneider, “MOSFET Mismatch Modeling: A New Approach”, *IEEE Design and Test of Computers*, pp. 20-29, Jan./Feb. 2006.

[83] T.B. Hook, J.B. Johnson, J.P. Han, A. Pond, T. Shimizu, G. Tsutsui, “Channel Length and Threshold Voltage Dependence of Transistor Mismatch in a 32-nm HKMG Technology”, *IEEE Transactions on Electron Devices*, Vol. 57, No. 10, pp. 2440-2447, Oct. 2010.

[84] B.P. Wong, A. Mittal, Y. Cao, G. Starr, “Nano CMOS Circuit and Physical Design”, John Wiley and Sons, 2005.

[85] P. Asenov, N.A. Kamsani, D. Reid, C. Millar, S. Roy and A. Asenov, "Combining Process and Statistical Variability in the Evaluation of the Effectiveness of Corners in Digital Circuit Parametric Yield Analysis," *European Solid-State Circuits Conference*, Sept. 13-17, 2010.

[86] V. Wang, K. Agarwal, S.R. Nassif, K.J. Nowka, D. Markovic, "A Simplified Design Model for Random Process Variability", *IEEE Transactions on Semiconductor Manufacturing*, Vol. 22, No. 1, Feb. 2009.

[87] A. Srivastava, D. Sylvester, D. Blaauw, "Statistical Analysis and Optimization for VLSI: Timing and Power", Springer, 2005.

[88] C.C. McAndrew, "Efficient Statistical Modeling for Circuit Simulation" In: R. Reis, J.AG. Jess (eds.), "*Design of System on a Chip: Devices and Components*", Kluwer Academics Publishers, 2004.

[89] C. Forzan, D. Pandini, "Statistical static timing analysis: A survey", *INTEGRATION, the VLSI journal*, Vol. 42, pp. 409-435, 2009.

[90] J. Shlens, "A Tutorial on Principal Component Analysis", 2009, available from http://www.snl.salk.edu/~shlens

[91] C.C. McAndrew, I. Stevanovic, X. Li, G. Gildenblat, "Extensions to Backward Propagation of Variance for Statistical Modeling", *IEEE Design and Test of Computers*, pp. 36-43, March/April 2010.

[92] C.C. McAndrew, "Statistical Modeling Using Backward Propagation of Variance", In: G.Gildenblat (ed.) "*Compact Modeling: Principles, Techniques and Applications*", Springer, pp. 491-520, 2010.

[93] International Technology Roadmap for Semiconductors (ITRS) website: http://www.itrs.net

[94] L. Schaffer, L. Lavagno, G. Martin (eds.), "EDA for IC Implementation, Circuit Design and Process Technology", Published by CRC Press, Taylor and Francis, 2006.

[95] W.Grabinski, B.Nauwelaers, D.Schreurs (eds.), "Transistor Level Modeling for Analog/RF IC Design", Springer, 2006.

[96] M.V. Dunga, "Nanoscale CMOS Modeling", *PhD thesis*, University of California at Berkeley, 2008.

[97] G. Gildenblat, H. Wang, T.L. Chen, X. Gu, X. Cai, "SP: An Advanced Surface-Potential-Based Compact MOSFET Model", *IEEE Journal of Solid State Circuits*, Vol. 30, No. 9, pp. 1394-1406, Sep. 2004.

[98] MOS Model 11 (MM11) manual, 2010 update is available from: http://www.nxp.com/wcm_documents/models/mos-models/model-11/m1102.pdf

[99] J. Watts, C.C. McAndrew, C. Enz, C. Glaup-Montoro, G. Gildenblat, C. Hu, R.v. Langevelde, M. Miura-Mattausch, R. Rios, C.T. Sah, "Advanced Compact

Models for MOSFETs", *Proceedings of NSTI Bio Nano Conference*, WCM, pp. 3-12, 2005.

[100] G. Gildenblat, X. Cai, T.L. Chen, X. Gu, H. Wang, "Reemergence of the surface potential based compact models", *Proceedings of IEDM Technical Digest*, pp. 863-866, 2003.

[101] A.J. Scholten, G.D.J. Smit, B.A. De vries, L.F. Tiemeijer, J.A. Croon, D.B.M. Klaassen, R. van Langevelde, X. Li, W. Wu, G. Gildenblat, "The New CMC Standard Compact MOS Model PSP: Advantages for RF Applications", *IEEE Journal of Solid State Circuits*, Vol. 44, No. 5, pp. 1415-1424, May 2009.

[102] The manual and other material related to EKV MOSFET compact model are available at: http://ekv.epfl.ch/

[103] H.C. Pao, C.T. Sah, "Effects of Diffusion Current on Characteristics of Metal-Oxide-Semiconductor Transistors", *Journal of Solid State Electronics*, No.9, pp. 927-937, 1966.

[104] Y. Taur, T.H.Ning, "Fundamentals of Modern VLSI Devices", Cambridge University Press, 2nd Edition, 2009.

[105] G. Gildenblat, "Compact Modeling: Principles, Techniques and Applications", Springer, 2010.

[106] W. Wu, T.L. Chen, G. Gildenblat, C.C. McAndrew, "Physics Based Mathematical Conditioning of the MOSFET Surface Potential Equation", *IEEE Transactions on Electron Devices*, Vol.51, No. 7, pp.1196-1200, July 2004.

[107] R. Rois, S. Mudanai, W.K. Shih, P. Packan, "An efficient surface potential solution algorithm for compact MOSFET models", *Proceedings of IEDM Technical Digest*, pp. 755-758, 2004.

[108] M. Miura-Mattausch, N. Sadachika, D. Navarro, G. Suzuki, Y. Takeda, M. Miyake, T. Warabino, Y. Mizukane, R. Inagaki, T. Ezaki, H.J. Mattausch, T. Ohguro, T. Iizuka, M. Taguchi, S. Kumashiro, S. Miyamoto, "Hisim2: Advanced MOSFET Model Valid for RF Circuit Simulation", *IEEE Transactions on Electron Devices*, Vol. 53, No. 9, pp. 1994-2007, Sep. 2006.

[109] T.L. Chen, G. Gildenblat, "Analytical approximation for the MOSFET surface potential", *Solid State Electronics*, No. 45, pp. 335-339, 2001.

[110] R. van Langevelde, F.M. Klaassen, "An explicit-surface potential based MOSFET model for circuit simulation", *Solid State Electronics*, No. 44, pp. 409-418, 2000.

[111] Y. Tsividis, "Operation and Modeling of the MOS Transistor", McGraw-Hill, 2nd

edition, 1999.

[112] C.C. McAndrew, J.J. Victory, "Accuracy of Approximations in MOSFET Charge Models", *IEEE Transactions on Electron Devices*, Vol. 49, No. 1, pp. 72-81, January 2002.

[113] H. Wang, T.L. Chen, G. Gildenblat, "Quasi Static and Non Quasi Static Compact MOSFET Model based on Symmetrically Linearization of the Bulk and Inversion Charges", *IEEE Transactions on Electron Devices*, Vol. 50, No. 11, pp. 2262-2272, Nov. 2003.

[114] J.R. Brews, "A charge sheet model of the MOSFET", *Solid State Electronics*, Vol. 21, Issue 2, pp. 345-355, Feb. 1978.

[115] Q. Zhou, W. Yao, W. Wu, X. Li, Z. Zhu, G. Gildenblat, "Parameter extraction for the PSP MOSFET model by the combination of genetic and Levenberg-Marquardt algorithm*", International Conference on Microelectronic Test Structures (ICMTS 2009),* pp. 137-142, 2009.

[116] S. Inaba, K. Okana, S. Matsuda, M. Fujiwara, et al, "High Performance 35nm Gate Length CMOS with NO Oxynitride Gate Dielectric and Ni Salicide", *IEEE Transactions on Electron Devices*, Vol. 49, No. 12, pp. 2263-2270, Dec. 2002.

[117] Z. Luo, N. Rovedo, S. Ong, B. Phoong, et al, "High performance transistors featured in an aggressively scaled 45nm bulk CMOS technology", *Symp. VLSI Tech. Dig.* pp. 16-17, 2007.

[118] K.L. Cheng, C.C. Wu, Y.P. Wang, D.W. Lin, et al, "A highly scaled high performance 45nm bilk logic CMOS technology with 0.242um^2 SRAM cell", *IEDM Technical Digest*, pp. 243-246, 2007.

[119] T. Miyashita, K. Ikeda, Y.S. Kim, T. Yamamoto, et al, "High performance and low power bulk logic platform utilizing FET specific multiple stressors with highly enhanced strain and full porous low-k interconnects for 45nm CMOS technology", *IEDM Technical Digest*, pp.251-254, 2007.

[120] X. Wang, "Simulation study of scaling design, performance characterization, statistical variability and reliability of deca-nanometer MOSFETs", *PhD dissertation*, University of Glasgow, 2010.

[121] K. Mistry, C. Allen, C. Auth, B. Beattie, et al, "A 45nm logic technology with high k +metal gate transistors, strained silicon, 9 Cu interconnect layers, 193nm dry patterning and 100% free Pb-packaging", *IEDM Technical Digest*, pp. 247-250, 2007.

[122] PSP 102.2 manual, developed by Arizona State University and NXP Semiconductor Research, October 2007, available online from: http://pspmodel.asu.edu/downloads/psp1022_summary.pdf

[123] BSIMProPLUS user's guide and manual available on request from: http://www.proplussolutions.com/

[124] HSPICE user's guide and manual available on request from: http://synopsys.com.

[125] N. Arora, "MOSFET Modeling for VLSI Simulation: Theory and Practice", World Scientific Publishing, 2007.

[126] T.S. Hsieh, Y.W. Chang, W.J. Tsai, T.C. Lu, "A New Leff Extraction Approach for Devices with Pocket Implants", *Proceedings of IEEE 2001 International Conference on Microelectronic Test Structures*, Vol. 14, pp. 15-17, 2001.

[127] A.J. Scholten, R. Duffy, R. Van Langevelde, D.B.M. Klaassen, "Compact modelling of pocket-implanted MOSFETs", *Proceedings of 31st European Solid State Device Research Conference*, pp. 311-314, 2001.

[128] D.E. Ward, R.W. Dutton, "A Charge-Oriented Model for MOS Transistor Capacitances", *IEEE Journal of Solid State Circuits*, Vol. 13, No. 5, pp. 703-708, 1978.

[129] B. Cheng, S. Roy, G. Roy, F. Adamu-Lema, A. Asenov, "Impact of Intrinsic Fluctuations in decanano MOSFETs on yield and functionality of SRAM cells", *Solid State Electronics*, Vol. 49, pp. 740-746, 2005.

[130] A. Asenov, A. Cathignol, B. Cheng, K. P. McKenna, A. R. Brown, A. L. Shluger, D. Chanemougame, K. Rochereau and G. Ghibaudo, "Origin of the Asymmetry in the Magnitude of the Statistical Variability of n- and p-Channel Poly-Si Gate Bulk MOSFETs," *IEEE Electron Device Letters*, Vol. 29, No. 8, pp. 913–915, 2008.

[131] H.J. Mattausch, N. Sadachika, A. Yumisaki, A. Kaya, W. Imafuku, K. Johguchi, T. Koide, M. Miura-Mattausch, "Correlating Microscopic and Macroscopic Variation With Surface-Potential Compact Model", *IEEE Electron Device Letters*, Vol. 3, No. 8, pp. 873-875, Aug. 2009.

[132] Aurora manual, Aurora is a parameter extraction and optimization software, available on request from http://synopsys.com.

[133] R.A. Thakker, N. Gandhi, M.B. Patil, K.G. Anil, "Parameter Extraction for PSP MOSFET Model using Particle Swarm Optimization", *International Worshop on Physics of Semiconductor Devices*, pp. 130-133, 2007.

[134] R.A. Thakker, M.B. Patil, K.G. Anil, "Parameter extraction for PSP MOSFET

model using hierarchical particle swarm optimization", *Elsevier Journal of Engineering Applications of Artificial Intelligence*, Vol. 22, pp. 317-328, 2009.

[135] Python programming language manuals are available from http://python.org/.

[136] M.H. Na, E.J. Nowak, W. Haensch, J. Cai, "The effective drive current in CMOS inverters", *IEDM Technical Digest*, pp. 121-124, 2002.

[137] V. Adler, E.G. Friedman, "Delay and Power Expressions for Short Channel CMOS Inverter Driving Resistive Interconnect", *Analog Integrated Circuits and Signal Processing*, Vol. 14, No. 1-2, Sep. 1997.

[138] S.S. Shapiro, M.B. Wilk, "An analysis of variance test for normality (complete samples)", *Biometrika*, Vol. 52, No. 3/4, pp. 591-611, Dec. 1965.

[139] H.W. Lilliefors, "On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown", *Journal of the American Statistical Association*, Vol. 62, No. 318, pp. 392-402, Jun. 1967.

[140] P.E. Greenwood, M.S. Nikulin, "A Guide to Chi-Squared Testing", John Wiley and Sons, 1996.

[141] M.B. Wilk, R. Gnanadesikan, "Probability plotting methods for the analysis of data", *Biometrika*, Vol. 55, No. 1, pp. 1-17, 1968.

[142] H.W. Lilliefors, "On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown", *Journal of the American Statistical Association*, Vol. 64, No. 325, pp. 387-389, Mar. 1969.

[143] D statistics critical values are available online from:http://www.mathematik.uni-kl.de/~schwaar/Exercises/Tabellen/table_kolmogorov.pdf

[144] G.E.P. Box, D.R. Cox, "An Analysis of Transformations", *Journal of the Royal Statistical Society*, Series B, 26, pp. 211-252, 1964.

[145] N.R. Draper, D.R. Cox, "On Distributions and Their Transformation to Normality", *Journal of the Royal Statistical Society*, Series B, pp. 472-476, 1969.

[146] I.J. Myung, "Tutorial on maximum likelihood estimation", *Journal of Mathematical Psychology*, No. 47, pp. 90-100, 2003.

[147] J.E. Jackson, "A User's Guide to Principal Components", John Wiley & Sons, 2003.

[148] A.I. Fleishman, "A Method for Simulating Non-Normal Distributions", *Psychometrika*, Vol. 43, No. 4, pp. 521-532, Dec. 1978.

[149] C.D. Vale, V.A. Maurelli, "Simulating Multivariate Non-Normal Distributions",

Psychometrika, Vol. 48, No. 3, pp. 465-471, Sep. 1983.

[150] G. Opton, I. Cook, "Oxford Dictionary of Statistics", 2nd edition, Oxford University Press, 2008.

[151] L. Isserlis, "On A Formula for the Product-Moment Coefficient of any Order of A Normal Frequency Distribution in any Number of Variables", *Biometrika*, Vol. 12, No. 1/2, pp. 134-139, Nov. 1918.

[152] Y. Ye, F. Liu, M. Chen, S. Nassif, Y. Cao, "Statistical Modeling and Simulation of Threshold Variation Under Random Dopant Fluctuations and Line-Edge Roughness", *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 19, No. 6, June 2011.

[153] T. Skotnicki, "Materials and device structures for sub-32nm CMOS nodes", *Microelectronic Engineering*, Vol. 84, No. 9-10, pp. 1845-1852, 2007.

[154] Y. Jiang, T. Liow, N. Singh, L. Tan, G. Lo, D. Chan, D. Kwong, "Performance breakthrough in 8nm gate length gate-all-around nanowire transistors using metallic nanowire contacts", *VLSI Technology Symposium Technical Digest*, pp. 34-35, 2008.

[155] T.Y. Liow, K.M. Tan, et al, "5nm gate length nanowire FETs and planar UTB-FETs with pure germanium source/drain stressors and laser-free melt-enhanced dopant diffusion and activation technique", *VLSI Technology Symposium Technical Digest*, pp. 36-37, 2008.

[156] Xin Sun, "Nanoscale Bulk MOSFET Design and Process Technology for Reduced Variability", *PhD thesis*, University of California at Berkeley, 2010.