Statistical Compact Model Strategies for Nano CMOS Transistors Subject of Atomic Scale Variability

Negin Moezi

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

February 2012

Abstract

One of the major limiting factors of the CMOS device, circuit and system simulation in sub 100nm regimes is the statistical variability introduced by the discreteness of charge and granularity of matter. The statistical variability cannot be eliminated by tuning the layout or by tightening fabrication process control. Since the compact models are the key bridge between technology and design, it is necessary to transfer reliably the MOSFET statistical variability information into compact models to facilitate variability aware design practice.

The aim of this project is the development of a statistical extraction methodology essential to capture statistical variability with optimum set of parameters particularly in industry standard compact model BSIM. This task is accomplished by using a detailed study on the sensitivity analysis of the transistor current in respect to key parameters in compact model in combination with error analysis of the fitted I_d-V_g characteristics. The key point in the developed direct statistical compact model strategy is that the impacts of statistical variability can be captured in device characteristics by tuning a limited number of parameters and keeping the values for remaining major set equal to their default values obtained from the "uniform" MOSFET compact model extraction. However, the statistical compact model extraction strategies will accurately represent the distribution and correlation of the electrical MOSFET figures of merit. Statistical compact model parameters are generated using statistical parameter generation techniques such as uncorrelated parameter distributions, principal component analysis and nonlinear power method. The accuracy of these methods is evaluated in comparison with the results obtained from 'atomistic' simulations. The impact of the correlations in the compact model parameters has been analyzed along with the corresponding transistor figures of merit. The accuracy of the circuit simulations with different statistical compact model libraries has been studied. Moreover, the impact of the MOSFET width/length on the statistical trend of the optimum set of statistical compact model parameters and electrical figures of merit has been analyzed with two methods to capture geometry dependencies in proposed statistical models.

Acknowledgment

I would like to take the opportunity to express my gratitude towards all the people who have assisted me in this work. Truly, finishing this PhD has been a milestone in my life and it has not been possible without their support. First and foremost, I would like to thank my supervisor Prof. Asen Asenov for his great ideas, countless suggestions, valuable comments and encouragements throughout this study. My second supervisor, Dr. Scott Roy devoted a lot of time and patience to the reading and correction of this thesis. My sincere thanks go to other members of Device Modelling Group at Glasgow University for their fruitful suggestions, among them I owe too much to Dr.Binjie Cheng for day to day discussions and progress of my research. Last but not least, I thank my parents and my beloved spouse for their endless support.

Publications:

1. 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 Proc. Simulation of Semiconductor Processes and Devices,* Sept. 9-11, 2009, pp. 143–146.

2. 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 Proc. Simulation of Semiconductor Processes and Devices (SISPAD),* Bologna, Italy, Sept. 6-8, 2010, pp. 125–128.

3. N. A. Kamsani, B. Cheng, C. Millar, **N. Moezi**, X. Wang, S. Roy and A. Asenov, “Impact of Slew Rate Definition on the Accuracy of nanoCMOS Inverter Timing Simulations”, *in Proc. Ultimate Integration on Silicon,* Glasgow, Scotland, UK, Mar. 17-19, 2010.

4. 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 Proc. Ultimate Integration on Silicon*, Glasgow, Scotland, UK, pp. 45–48, Mar. 17-19, 2010.

5. D. Dideban, B. Cheng, **N. Moezi**, X. Wang and A. Asenov, “Evaluation of 35nm MOSFET Capacitance Components in PSP Compact Model”, *ICEE2010,* Isfahan, Iran, May 11-13, 2010.

6. B. Cheng, **N. Moezi**, D. Dideban, C. Millar, S. Roy and A. Asenov, “Impact of Statistical Parameter Set Selection on Accuracy of Statistical Compact Modeling”, *MOS-AK Workshop*, Sapienza Università di Roma, Apr. 8-9, 2010.

7. 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 Proc. Design, Automation and Test in Europe*, Dresden, Germany, pp. 650–653, Mar. 8-12, 2010.

8. B. Cheng, D. Dideban, **N. Moezi**, C. Millar, G. Roy, X. Wang, S. Roy and A. Asenov, “Statistical Variability Compact Modelling Strategies for BSIM4 and PSP”, *IEEE Design and Test of Computers*, Vol. 27, No. 2, pp. 26–35, Mar./Apr. 2010.

9. A. Asenov, B. Cheng, D. Dideban, U. Kovac, **N. Moezi**, C. Millar, G. Roy, A. Brown and S. Roy, “Modelling and Simulation of Transistor and Circuit Variability and Reliability”, *Custom Integrated Circuit Conference*, San Jose, California, Sept. 19-22, 2010.

10. S. Markov, X. Wang, **N. Moezi** and A. Asenov, “Drain Current Collapse in Nano-Scaled Bulk MOSFETs Due to Random Dopant Compensation in the Source/Drain Extensions”, *IEEE transaction on Electron Device*, VOL. 58, NO. 8, August 2011.

Accepted:

11. **N. Moezi**, D. Dideban, B. Cheng, S. Roy and A. Asenov, “Impact of Statistical Parameter Set Selection on the Statistical Compact Model Accuracy: BSIM4 and PSP Case Study”, *Accepted for publication in Microelectronics Journal*.

Contents

List of Figures

List of Tables

Nomenclature

Physical Constants

ε_0	Permittivity of free space ($8.85418782\times10^{-14}\,F/cm$)
ε_{si}	Silicon permittivity (11.68 ε_0)
Q	Elementary (unit) charge ($1.60\times10^{-19}C$)
n_i	Intrinsic carrier concentration ($8.72\times10^{9}cm^{-3}$)
k_B	Boltzmann's constant ($1.38\times10^{-23}J/K$)

Other Symbols

σ	standard deviation
θ_{th}	Basic form of physical expression for SCE
μ	Mean
μ_0	Low-field mobility
μ_{eff}	Effective mobility
λ	Introduced to model the non-saturation effects which are found for P-MOSFETs

Φ_S	Surface potential
Y	A constant which depends on device type and technology
A_{bulk}	Bulk charge effects
C_0	Constant
C_{cen}	Centroid Channel Charge Capacitance
C_{oxe}	Equivalent Gate Oxide Capacitance
C_{oxeff}	Effective Gate Oxide Capacitance
E_0	Critical electric field
E_{eff}	Effective field
I_d	MOSFET drain current
I_{off}	Channel leakage current
I_{on}	Drive current
G_x	Conductance
l_t	Characteristic length
L	Device length
L_{active}	Effective length when the gate to source/drain regions is under flat-band condition.
L_{eff}	Effective channel length
L_g	Gate length
N	Multiple of the base device gate width
N	Sub-threshold swing parameter
Q_b	Charge density in the bulk
Q_{INV}	Charge density in the channel
Q_{ch}	Channel charge density
t_{dLH}	Rise time delay

t_{dHL}	Fall time delays
T	Temperature
V_A	Early voltage
V_{ADITS}	Effect and drain-induced threshold shift
V_{ASCBE}	Early voltage due to substrate current
V_{bi}	Source/drain junction built-in voltage
V_{bs}	Bulk/source bias
V_{bseff}	Effective substrate bias
V_{ds}	Drain/source bias
V_{dsat}	Saturation voltage
V_F	Quasi-Fermi potential
V_{fbSD}	Flat-band voltage between the gate and source/drain diffusion regions
V_g	MOSFET gate voltage, relative to the source contact
V_{gse}	Effective gate voltage including the poly-silicon gate depletion effect
V_{gsteff}	Effective gate bias
V_{in}	Input voltage (V)
V_t	Thermal voltage
V_{th}	Threshold voltage
V'_{off}	Potential offset parameter
W	Device width
W_{active}	Effective width when the gate to source/drain regions is under flat-band condition.
W_{eff}	Effective channel width
W_{effcj}	Effective source/drain diffusion width

X_{DC}	Equivalent DC centroid of the channel charge layer
X_{dep}	Depletion layer width in the channel with the influence of V_{bs}
X_{dep0}	Depletion layer width in the channel when $V_{BS} = 0$

Acronyms

1D	One Dimension
2D	Two Dimensions
3D	Three Dimensions
AC	Alternating Current
BPV	Backward Propagation of Variance
BSIM	Berkeley Short-channel IGFET Model
BSIM-MG	BSIM Multi Gate
BSIM CMG	BSIM Common and Multi Gate
BSIMSOI	BSIM Silicon On Insulator
BTE	Boltzmann Transport Equation
CESL	Contact Etch Stop Layer
CLM	Channel Length Modulation
CMC	Compact Model Council
CMOS	Complementary Metal Oxide Semiconductor
DIBL	Drain Induced Barrier Lowering
DITVS	Drain-Induced Threshold Voltage Shift
DC	Direct Current
DD	Drift Diffusion
DG	Density Gradient
EIA	Electronic Industries Alliance

EOT	Equivalent Oxide Thickness
FCLT	Finite Charge Layer Thickness
FOM	Figures of Merit
GIDL	Gate Induced Drain Lowering
IC	Integrated Circuit
LDD	Low-Doped Drain
LER	Line Edge Roughness
MC	Monte Carlo
MOSFET	Metal Oxide Semiconductor Field effect Transistor
NBTI	Negative Bias Temperature Instability
NULD	Non-Uniform Lateral Doping
NPM	Nonlinear Power Method
NWE	Narrow Width Effect
OTF	Oxide Thickness Fluctuations
OTV	Oxide Thickness Variations
PBTI	Positive Bias Temperature Instability
PCA	Principal Component Analysis
PGG	Poly-silicon Gate Granularity
QT	Quantum Transport
RDD	Random Discrete Dopants
RDF	Random Dopant Fluctuations
RMS	Root Mean Square
SCE	Short Channel Effect
SEM	Scanning Electron Microscope
SS	Subthreshold Slope

SPICE	Simulation Program with Integrated Circuit Emphasis
SRAM	Static Random Access Memory
STI	Shallow Trench Isolation
TCAD	Technology Computer Aided Design
ULSI	Ultra Large Scale Integration
VHDL	Very High Design Language

Chapter 1

Introduction

Computer-aided design (CAD) tools are an essential part of the integrated-circuit (IC) design flow. Among them, the Simulation Program with Integrated Circuit Emphasis (SPICE) is widely used for analysis and verification of analogue, mixed mode and digital circuits employing billions of transistors [1]. Transistor compact models are key to the utility of SPICE, and acting as an interface between technology and design. Although the initial driving force of compact model development was the requirement of accurate modelling of circuit components in analogue IC design domain, compact models are now extensively used in transistor-level digital circuit design and verification, especially in the characterization of standard cells and in the SRAM (Static Random Access Memory) design process. Therefore, compact model accuracy is key to analogue, mixed mode and digital IC design.

BSIM4 (Berkeley Short-channel IGFET Model) is a threshold voltage based compact model for integrated circuit design purposes. This model is constructed on top of the BSIM3 framework. It shares the same basic equations with that bulk model so that the physical nature and smoothness of BSIM3 are preserved. BSIM series compact models have served the electronic industry for more than 20 years with focus on the simulation of planar bulk MOSFETs [2,3]. Recently, other flavours of BSIM compact models such as

BSIMSOI and BSIM-CMG have been developed to simulate SOI and multi gate MOSFETs [4,5]. Since planar bulk MOSFETs are still the workhorse of the semiconductor industry and their scaling is expected to continue until 2015, BSIM4 is still being used by semiconductor companies such as IBM and AMD and foundries such as TSMC. BSIM4 was selected by Electronic Industries Alliance (EIA) and Compact Model Council (CMC) as the standard MOSFET compact model in March 2000 [3]. Various compact models like PSP, HiSIM and EKV have been introduced later and are being used by semiconductor industry [6]. BSIM4 is used as the target compact model in this thesis.

Although the design of compact models, and the extraction of compact model parameters is a mature field (especially for bulk devices), this long standing interface between technology and design is now being seriously affected by a rapidly growing problem in device technology – atomic scale statistical MOSFET variability. As devices scale further into the sub-100nm regime, intrinsic parameter fluctuations between devices which result from the discreteness of charge and granularity of matter are now one of the major obstacles which limit scaling and integration [7]. In general, device variability can be broadly classified as global or local variability. Global variability refers to the changes in the behavior of nominally identical transistors from wafer to wafer or die to die [8,9], typically related to variations in processing resulting in change of transistor structural dimensions and doping profiles. Local variability can be subdivided into two types: systematic variability and statistical variability. Systematic variability is introduced by layout dependent optical proximity effects and strain variations. It can be reduced by the adoption of restricted design rules. Moreover, it is locally predictable using sufficient computing power. However, statistical variability, introduced by the granularity of matter and discreteness of charge (for example due to the precise configuration of discrete dopant atoms in a nanoscale device channel) cannot be eliminated by tuning the layout or by tightening fabrication process control and cannot be deterministically predicted using simulations. It is a significant and growing problem – it accounts for more than 50% of device variability in 45nm CMOS technology [10] – and is predicted to become the major source of transistor variations for future technology generations. In addition, because it effects every nominally identical transistor, even those placed side by side in a circuit, it is

an uncontrollable source of circuit mismatch. Simulation and modelling of intrinsic parameter fluctuations is crucial for understanding on predicting the statistical variability. Simultaneously circuit designers would much prefer information on intrinsic parameter fluctuations to be supplied in the familiar framework of compact models. This becomes crucial for high yield nanometer CMOS design.

Mismatch has historically been the subject of specific modelling, and the importance of device matching in the analogue domain drove early transistor mismatch modelling efforts. The first systematic mismatch models were reported in early '80s for MOS capacitors and MOSFETs [11,12], and this still remains an active research area today [13,14]. Mismatch studies target random, uncorrelated variation that cannot be improved by matching techniques [1,12]. However, most MOSFET mismatch models are based on simple MOSFET drain current formulae in the linear or saturation region [15,16]. Although they can provide important information regarding mismatch trends versus transistor design and dimensions, they cannot be integrated into design tools to directly support design activities. A natural way to incorporate mismatch into design flow is to employ *statistical compact modelling* techniques, and investigating such statistical compact modelling will be the aim of this work.

Most previous compact model based mismatch approaches rest upon the assumption of normal, uncorrelated distributions of compact model parameters [17,18,19]. Although this assumption was primarily made to ease theoretical prediction, it also allows the storage and transfer of relatively small compact model data files containing only the mean and standard deviation of parameters, rather than large statistical 'card indexes' capturing the full distributions and correlations of compact model parameters. If necessary, approximate statistical BSIM parameter ensembles could then be generated from these mean (μ) and standard deviation (σ) values. A key aspect of our study will be to investigate the efficacy of this simplifying assumption, and study to what extent BSIM parameters generated from μ and σ values, or from additional moments and correlations can provide accurate distributions and correlations of transistor characteristics and figures of merit.

1.1 Aims and objectives

The aim of this work is to investigate accurate and efficient techniques for incorporation of statistical variability into industry standard compact models such as BSIM4. The main objectives include:

1. To develop a BSIM parameter extraction and optimization methodology using as a test bed a template 35nm and 18nm bulk MOSFETs.

2. To determine a subset of the BSIM compact model parameters that can accurately capture the effects of statistical variability over a statistical device ensemble.

3. To evaluate the accuracy of the methodology as a function of statistical parameter subset size, comparing with benchmark physics based 'atomistic' simulation results, and evaluating with respect to the distributions of typical MOSFET electrical figures of merit.

4. To develop statistical BSIM parameter generation techniques and to evaluate their accuracy and their ability to reproduce typical MOSFET electrical figures of merit in comparison with benchmark physics based simulations and directly extracted compact models.

5. To perform statistical circuit simulation of simple CMOS circuits using a statistical compact model library as a practical test of the accuracy of different parameter extraction and generation techniques.

6. To study the impact of statistical variability on different channel width/length MOSFETs and their electrical figures of merit.

1.2 Thesis Outline

The rest of this thesis is organized as follows. The second chapter presents background information about variability and the classification of variability in nanoscale MOSFETs. Statistical variability and its sources including Random Discrete Dopants (RDD), Line Edge Roughness (LER), Oxide Thickness Fluctuations (OTF), high-κ granularity and Poly-silicon Gate Granularity (PGG) are discussed. Physical simulation of statistical variability and the impact of variability on devices and circuits are explained. Existing techniques for statistical variability simulation are explored in details with emphasis on drift-diffusion simulation of statistical variability using the Glasgow University 'atomistic' simulator. Finally a literature review of the concept of statistical compact modeling is presented.

The third chapter focuses on the BSIM4 compact model parameter extraction and optimization with reference to the template/exemplar MOSFETs. Elements of the key BSIM4 compact model expressions for threshold voltage and drain current and the relationship between BSIM4 compact model parameters are reviewed. The physics and design of the template MOSFETs are then discussed, and the methodology of compact model parameter extraction and optimization for such devices is explored. The accuracy of results calculating the static behaviour of n- and p-MOSFET transistors is reviewed. The accuracy criteria are based on the RMS error for drain current between compact model results and the original TCAD physical simulations data from which the compact models were extracted.

In chapter 4, the statistical compact model extraction strategy is developed. First, atomistic simulations are carried out to obtain an ensemble of MOSFET I_d-V_g characteristics which capture the main sources of variability in 35nm channel length transistors. A subset of compact model parameters which can accurately capture the impact of statistical variability on these MOSFET characteristics is determined based on first order sensitivity analysis of I_d. The accuracy of the fitting between the compact models and original I_d-V_g curves is then discussed. The impact of the compact model parameter subset size on the statistical properties of MOSFET figures of merit is studied, and the accuracy

of figures of merit obtained from statistical parameter extraction is evaluated in comparison with the original atomistic simulations (including evaluation of the correlations between figures of merit). A simple CMOS inverter is simulated to study the impact of statistical parameter set size on the accuracy of practical circuit figures of merit – in this case the propagation delay and power dissipation of the inverter.

Chapter 5 is devoted to statistical compact model parameter generation techniques. The statistical properties of the directly extracted BSIM compact model parameters are first reviewed, including their distributions and the correlations between each pair of parameters. Naïve and PCA approaches are then introduced in an attempt to generate compact model parameter ensembles with the same statistical properties as the directly extracted compact model parameters, and the accuracy of these approaches is evaluated by comparing the parameter distributions, and how well they reproduce MOSFET and simple circuit figures of merit. In the last section, the Nonlinear Power Method, a statistical moment matching technique, is introduced and its accuracy is evaluated.

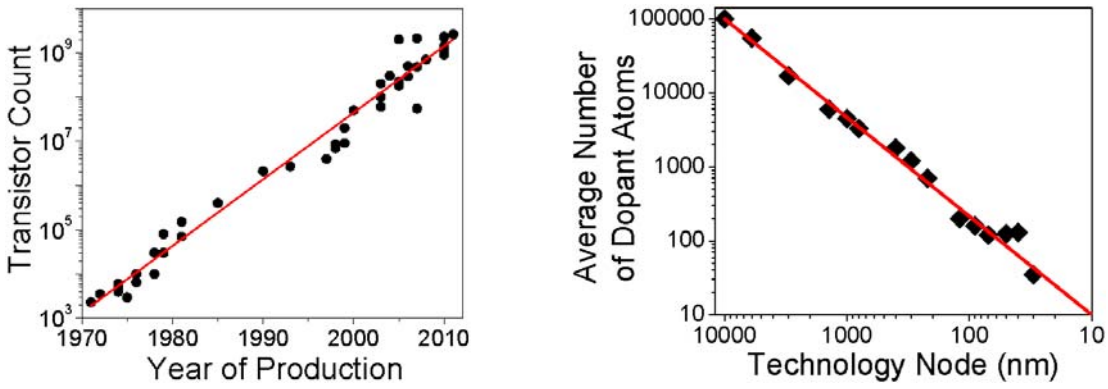
Chapter 6 investigates statistical compact modelling for different channel length/width MOSFETs. Full statistical atomistic simulation of different width/length devices is carried out to provide benchmark results. A computationally efficient approximation to obtain statistical results for different width/length devices – the parallel component approach – is described and its accuracy is investigated by considering trends in extracted BSIM parameters and device figures of merit.

Chapter 7 concludes this work. A summary of the results obtained in the bulk of the thesis is presented and future directions of research are given.

Chapter 2

Background

The semiconductor industry has always been driven by the scaling of CMOS transistors aiming to produce high performance devices (higher speed and lower power) and increased circuit density, following Moore's famous law [20]. Moore's law predicts the long-term trend in the number of transistors which can be located on an integrated circuit (IC), indicating that the number of transistors in a chip will double every two years [20]. Scaling has two important aims, first is to shrink device dimensions in order to obtain higher device density and therefore increased system functionality, and second to achieve higher devices/circuits performance. Figure 2.1(a) shows transistor count versus date of introduction for a number of noted microprocessors [21]. Due to scaling, the numbers of dopant atoms are reduced in scaled devices as shown in Figure 2.1(b) [22]. This reduction results in the impact of Random Dopant Fluctuations (RDF) on device characteristics. As a result, RDF has emerged as a significant source of statistical variability in contemporary and next generation CMOS devices. Moreover, device variability is the main factor restricting the scaling of the supply voltage, which for the last four technology generations has remained constant [22].



(a) (b)

Figure 2.1: Two manifests of MOSFET scaling: (a) Microprocessor transistor counts from 1971 to 2011, after [21], (b) Reduction in the average number of dopant atoms per technology generation, after [22].

Reduction of channel length in bulk MOSFETs requires increasing the channel doping to control short channel effects such as threshold voltage roll-off and punch-through. High channel doping causes an increase in ionized impurity scattering which in turn degrades carrier mobility and reduces the drain current [23,24]. Moreover, devices with high level of channel doping are more susceptible to degradation in device performance caused by direct band-to-band leakage current in the drain region and gate-induced-drain-leakage (GIDL) effects [25,26]. Shallow source and drain extensions and lateral non-uniform doping such as pocket implants are needed to compensate for threshold voltage roll-off and punch-through [27]. Threshold voltage scaling leads to an increased gate overdrive (difference between gate voltage and threshold voltage) and hence increased drive current and switching speed. Unfortunately, the source-drain leakage current increases exponentially with the reduction in threshold voltage and this increase results in high static power consumption. New materials and process technologies may need to be introduced to improve performance and continue scaling. For instance, strain has been introduced to compensate for the performance loss of scaled transistors by boosting the mobility and drive current [28,29,30].

To ensure adequate control of the channel by the gate, MOSFET oxide thickness should scale proportionately to channel length. However, decreasing oxide thickness increases the

gate tunneling current exponentially, which critically affects low power applications [31]. High-κ hafnium-based dielectrics were introduced at the 45nm technology node to prevent such tunneling currents in the gate [32].

Intrinsic parameter fluctuations associated with discreteness of charge and granularity of matter are now one of the major obstacles which limit scaling [7], integration and the reduction of supply voltage and power consumption in ULSI applications. The accurate modelling and simulation of such effects is very important for the development of present and future generation semiconductor devices and their integration into giga-transistor count chips [33,34,35]. In this chapter we focus on the effects of statistical variability which have become dominant at the 45nm technology generation and cannot be further reduced by tightening process control on the device and circuit fabrication steps.

2.1 Classification of Variability

Several axes have been introduced in the literature for classification of device variability [36]. One axis classifies the variability into global and local variations [8,37]. Global variations refer to variation across or between fabricated wafers. This kind of variability is caused by lack of control on uniformity in the fabrication and is the difference in oxide layer thickness, physical gate length/width and doping concentration of two devices in different wafers or two devices in the same wafer but usually at a distance apart. By using better manufacturing equipment and process technologies this kind of variability can be controlled and significantly reduced.

Local variations refer to variation between adjacent devices in a chip and falls in two categories: systematic and statistical variations. Systematic variation is the component of the physically varying parameters that follow a well understood behavior and can be predicted and modelled. For instance, optical proximity effects [38], layout mediated strain [39] and well proximity effects [40] are some sources of systematic variability. Statistical variability sources will be considered in next section.

Another classification splits the variability into intrinsic variations and extrinsic variations [37]. Intrinsic variability is caused by the discreteness of charge and granularity of matter in a device which does not depend on fabrication accuracy. Extrinsic variability is associated with the operating dynamics, layout effects and uncontrolled changes in the fabrication process conditions [41].

In this thesis, the focus will be upon intrinsic or statistical variability. At the 45nm technology generation, statistical variability accounts for more than 50% of total variability [10,42] and in smaller devices in new technologies, the manifestation of statistical variability will be increased.

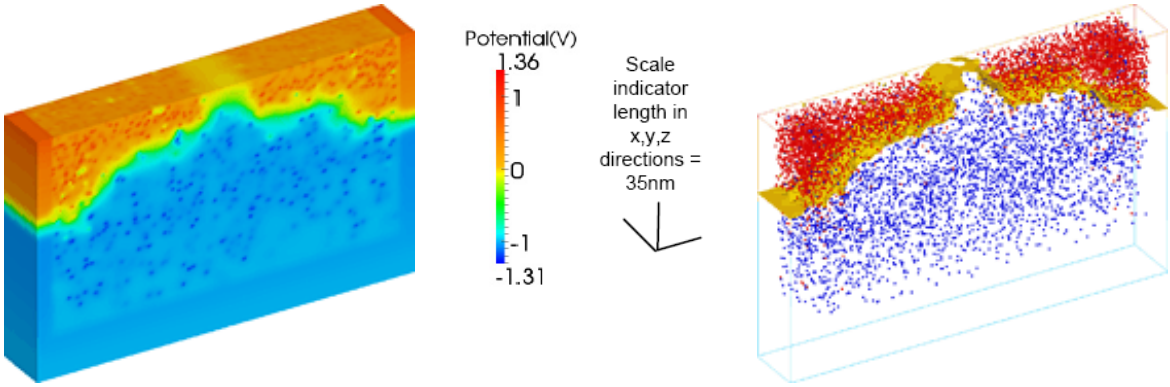
2.2 Statistical Variability in Nano-CMOS

Statistical variability which arises from discreteness of charge and granularity of matter is one of the fundamental limiting factors of CMOS scaling and integration in the nanometer regime [43]. Due to its purely random nature, statistical variability introduces increasing challenges for accurate compact modelling and statistical circuit simulation [13,17,19]. The major sources of statistical variability are introduced below.

2.2.1 Random Discrete Dopants

Random discrete dopants (RDD) are the most significant source of statistical variability in bulk MOSFETs. Threshold voltage variation due to random variations in the number and position of donor and acceptor atoms in the channel, source and drain regions is an increasingly important problem as device dimensions shrink and has received increasing attention [44,45]. The ionized dopant atoms are typically introduced into the silicon lattice by ion implantation. After annealing, they replace Si atoms in the lattice. These dopants have many collisions with atoms of lattice before coming to rest and consequently replacing Si atoms; therefore the final position of impurities will be determined by the implantation and annealing conditions and will inevitably result in a random dopant

distribution for each transistor. This effect leads to a random current-voltage characteristic change from device to device [46]. Figure 2.2 shows an example of random dopant distribution for a 35nm gate length MOSFET. Each impurity atom creates a discrete Coulomb potential peak in the channel of MOSFET which in turn will have an impact on electrical characteristics and current flow in the device.



(a) (b)

Figure 2.2: Typical atomistic discrete dopant distribution inside a 35nm gate length MOSFET: (a) Potential distribution (potential bar indicates higher positive potentials around donors located in source and drain and and negative potentials around acceptors in the bulk); (b) Dopant position map; Red bubbles are donors in source/drain; Blue bubbles are acceptors in bulk; Yellow color indicates the P-N junction, used with permission from [47].

Several research articles have been published which investigate the effects of discrete random dopants using analytical models [15,48,49,50,51]. Initial simulation studies of the effects of the discrete dopant distribution were accomplished using 2D simulations [51,52,45]. Since the random dopant fluctuations are essentially three dimensional in nature, more advanced and accurate simulators employ 3D numerical techniques [53,54,55].

2.2.2 Line Edge Roughness

Line edge roughness (LER) is another source of intrinsic parameter fluctuations. It is caused by limitations inherent in the materials and tools used in the lithographic process [56,57]. LER is manifest as local variations in the active channel length along the channel width. Where the local channel length is langer, the localized threshold voltage is increased and both leakage and drive current decrease. Where the local channel length is smaller, the localized threshold voltage is reduced and leakage current increases exponentially. This is due to the fact that in those regions where local channel length decreases, short channel effects become strong. Figure 2.3 shows LER in a device with 50nm gate length.

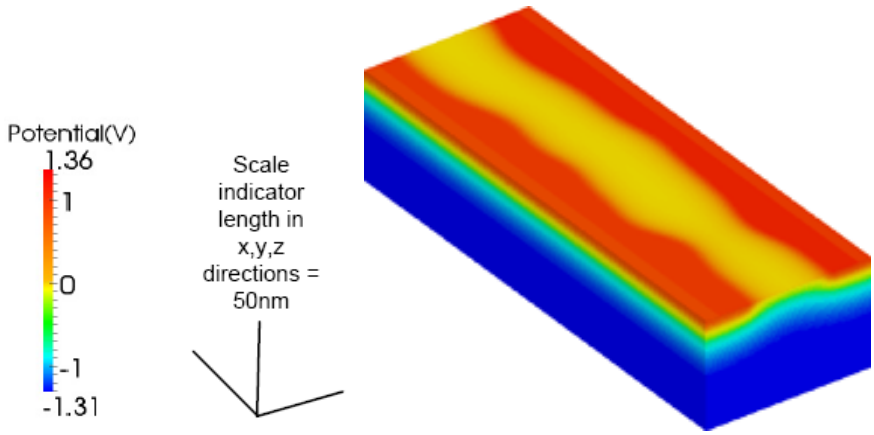


Figure 2.3: LER pattern in a 50nm×200nm MOSFET. Red color shows the high potential associated with source/drain area the blue color indicates the low potential in the bulk. Yellow color illustrates the potential area in the channel and depletion regions, [58].

LER is characterized by two parameters: The RMS magnitude Δ and the correlation length Λ. The usual definition of LER magnitude is 3Δ. The typical value for 3Δ is 5nm and Λ varies between 10nm to 50nm [59]. Simulations have shown that if the magnitude of the line edge roughness is not reduced below current levels, then the impact of LER will overtake that of RDD at devices scaled to approximately 18nm channel length [60]. LER effects in interconnects also result in resistance and capacitance variability which leads to variability in delay and power consumption of systems and circuits [61].

2.2.3 Oxide Thickness Variations

The atomic scale roughness of the Si/SiO_2 and $gate/SiO_2$ interfaces will introduce significant intrinsic parameter fluctuations [62]. Indeed, when the oxide thickness is equivalent to only a few silicon atomic layers, the atomic scale interface roughness will result in significant relative oxide thickness variation within the gate region of an individual MOSFET. Figure 2.4 shows the variations in the oxide thickness.

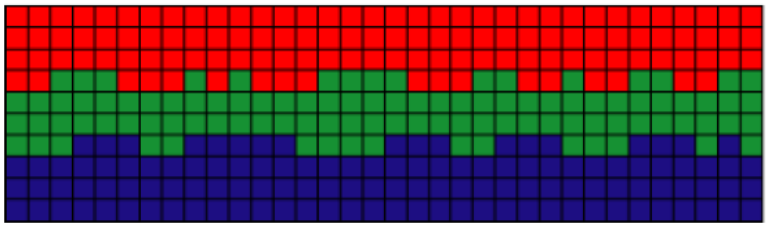


Figure 2.4: The interface between gate, oxide and silicon in an atomic layer abstraction. The red, green and blue colors indicate atomic layers in poly-silicon gate, oxide and silicon, respectively. The fluctuations in the oxide thickness are clearly shown, used with permission from [46].

The random pattern of the gate oxide thickness and interface landscape makes a unique characteristic for each nano-scale MOSFET because the correlation length of the interface becomes comparable to the channel length, different from its counterparts and leads to variations in the surface roughness limited mobility and threshold voltage from device to device [63,64].

2.2.4 High-κ Granularity

The reduction in oxide thickness of MOSFET devices leads to increased gate leakage. High-κ dielectrics such as Al_2O_3, HfO_2 and $ZrSiO_4$ have therefore superseded SiO_2 or $SiON$ in new technologies because they have higher permittivity and therefore, using them with larger thickness will provide the same equivalent oxide thickness of SiO_2 while reducing the gate leakage current [46].

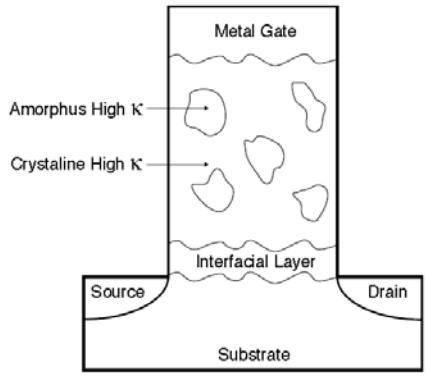
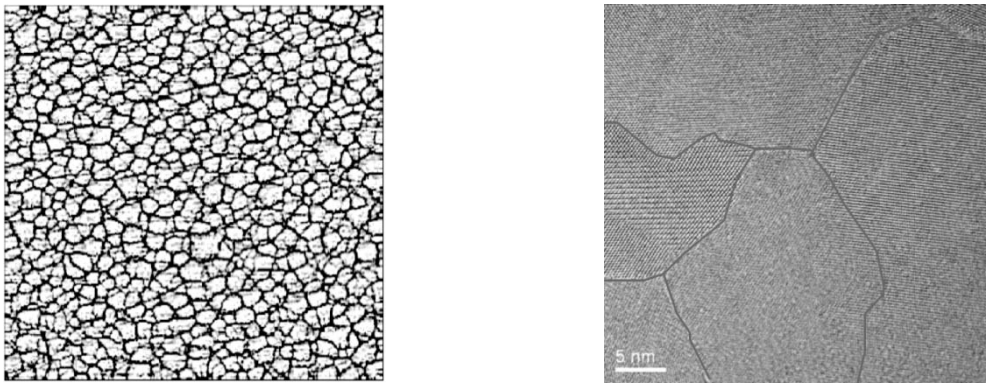


Figure 2.5: The gate stack of a nano-CMOS using high-κ dielectrics and polycrystalline of high-κ which leads to vary dielectric thickness in each point of gate, used with permission from [46].

The use of high-κ material in the gate stack of nano-CMOS devices leads to parameter fluctuations between devices. The polycrystalline nature of high-κ materials illustrated in Figure 2.5 may lead to non-uniformity of the dielectric properties across the oxide film, which in turn results in fluctuations in important device parameters such as threshold voltage [65,66,67].

2.2.5 Poly/Metal Gate Granularity

The polycrystalline granular structure of the poly-silicon gate has also been identified as an important source of variability. Fermi level pinning at the boundaries of grains due to a high density of defect states introduces surface potential fluctuation within the MOSFET channel and leads to variation in threshold voltage and current characteristics [68,69]. The use of metal gates eliminates poly gate granularity (PGG) induced variability but there is still gate work-function variability associated with metal gate granularity (MGG). Figure 2.6 shows SEM micrographs of two typical poly-silicon and metal gates.



(a) (b)

Figure 2.6: SEM micrograph of (a) poly-silicon and (b) metal gates, with 5nm scale for both figures as indicated in (b).Used with permission from [70].

2.3 Physical Simulation of Statistical Variability

There are many simulation techniques used in the simulation of sub 0.1 μm MOSFET devices to forecast the statistical variability in future technology generations. All of the simulation approaches are differentiated by the level of approximations used in the solution of the Boltzmann transport equation and the handling of quantum effects [71]. Recently, advanced techniques have been reported in predictive physical simulation of statistical variability using Drift Diffusion (DD), Monte Carlo (MC) and Quantum Transport (QT) techniques [43]. The simulations must be carried out in full 3D due to the 3D nature of intrinsic parameter fluctuations. DD is the most computationally effective technique to simulate statistical variability and the Glasgow atomistic simulator, which has been used for physical atomistic simulations of statistical variability in this work, is a drift diffusion simulator.

A Monte Carlo technique can be used to solve the Boltzmann Transport Equation (BTE) [72]. In this technique an ensemble of particles and the corresponding spatial energy distributions evolve through real space acceleration and randomly chosen scattering events. This approximates the solution of the BTE in a computationally tractable way. This approach is the best method to simulate high current flow in small MOSFETs but requires

long simulation times to allow reliable statistical averaging, therefore this method is not suitable to study the effects of intrinsic parameter fluctuations due to the large number of devices which need to be simulated to obtain statistical information.

Quantum mechanical approaches like Non Equilibrium Green Functions (NEGF) couple the Poisson and time-independent Schrödinger equations, and can give great predictive power, but are extremely computationally expensive.

2.3.1 Drift Diffusion Simulations

For many years 3D drift diffusion simulations have been the workhorse of the statistical variability simulations [43]. The DD technique uses the first two moments of the Boltzmann Transport Equation and Poisson's equation [73]. This model includes a local relationship between the velocity and the electric field and cannot represent properly non-equilibrium transport effects and scattering variations and hence under-estimates drive current and its variations [46]. However, it has been shown that the accuracy of this approach is very good in the sub-threshold region of device operation [46].

In this technique, electron and hole current density have two components: a drift component derived from the electric field and a diffusion part derived from the carrier density gradient. The total current density of electrons and holes is given by:

$$J_n = qD_n \nabla n - q\mu_n n \nabla \Phi \tag{2.1}$$

$$J_p = -qD_p \nabla p - q\mu_p p \nabla \Phi \tag{2.2}$$

where μ is the mobility, D is the diffusion coefficient and Φ is the electrostatic potential. In the Boltzmann approximation, the mobility and diffusion coefficient for both electrons and holes, is related via the Einstein relation:

$$D = \frac{k_B T}{q} \mu \tag{2.3}$$

where k_B is Boltzmann's constant and T is the temperature. Since the drift diffusion approach has low computational cost, it is perfect for the simulation of large scale statistical devices and investigating the impact of various sources of intrinsic parameter fluctuations. The bulk of our simulations have been carried out with the 3D Glasgow 'atomistic' drift diffusion simulator with quantum corrections. By shrinking MOSFETs into the nanometer scale, the influences of quantum effects become increasingly important. Since pure quantum mechanical simulations for a statistical set of devices requires very high level of computational resources, quantum corrections are used instead in combination with DD simulations. Two frequently used methods to incorporate quantum correction in classical DD simulation are the Density Gradient approach and the Effective Potential approach [74,72]. The Glasgow atomistic simulator uses the first method and further details can be found in [75].

2.3.2 Incorporation of Statistical Variability into Atomistic Simulator

Three important sources of statistical variability have been introduced in Glasgow atomistic simulator: RDD, LER and PGG. RDD is introduced by random placement of dopant atoms in the source, drain and channel regions of the MOSFET. The probability of a dopants being placed in each region is determined by local ratio between dopant and silicon atom concentration. Since the basis of the silicon lattice is 0.543nm, a fine mesh of 0.5nm or 1nm is used to ensure a high resolution of dopant atoms in simulations [46].

LER is introduced through 1D Fourier synthesis which generates random gate edges from a power spectrum corresponding to a Gaussian or exponential autocorrelation function [56]. The correlation length Λ and the RMS (root mean square) amplitude Δ describe this random gate edge. Current lithography systems have $\Lambda = 30\,nm$ and

$\Delta = 1.3\,nm$ [76]. In most cases the quoted values of LER in the literature are equal to 3Δ [77].

For simulating PGG, random generation of poly-grains is used for the whole gate region [69]. A polycrystalline silicon grains image has been used as a template and it is scaled according to the average grain diameter in real MOSFETs. Then the simulator imports a random (in both location and orientation) section of the grain template image that corresponds to the gate dimension of the simulated device. The applied gate potential in the poly-silicon is modified in such a way that the Fermi level remains pinned at a certain position in the silicon band gap. The impact of poly-silicon grain boundary variation on device characteristics is simulated through the pinning of the potential in the poly-silicon gate along the grain boundaries [77].

2.4 Effects of Variability in Device/Circuit

Figure 2.7 shows the potential variation due to RDD, LER and PGG in channel, source and drain of a 35nm gate length device simulated using the Glasgow 'atomistic' simulator. Since the structure does not use high-κ material, this source of variability is not included in the simulations.

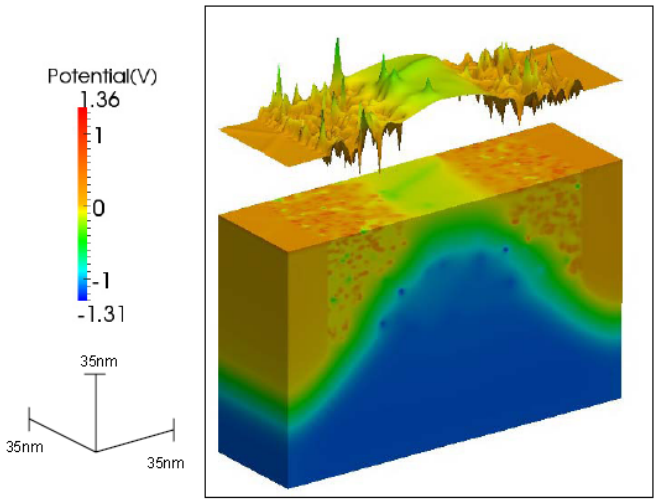


Figure 2.7: Potential profile in a 35nm physical gate length device with RDD, LER and PGG, the numbers shown on the color bar indicate the electric potential in Volts.

The combined effects of intrinsic parameter fluctuations will have significant impact on devices, circuits and systems. Figure 2.8 shows the spread in I_d-V_g characteristics of a 35nm gate length minimum width MOSFET subject to RDD, LER and PGG using Glasgow drift diffusion 'atomistic' simulator. The physical design of the template device will be reviewed in chapter 3.

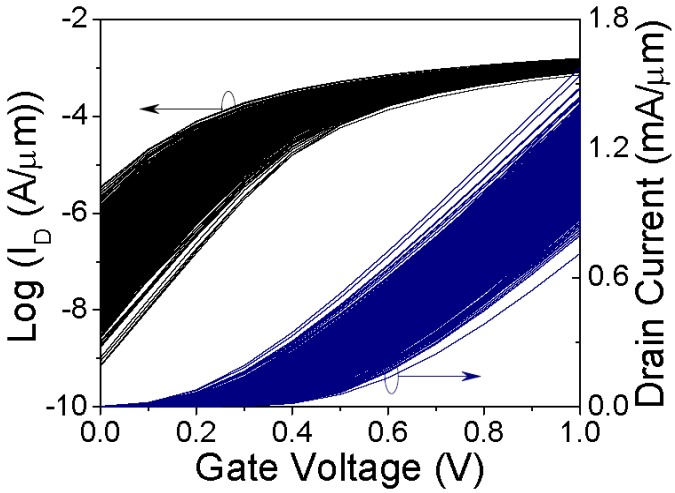


Figure 2.8: The I_D-V_G characteristics of 1000 MOSFETs with 35nm gate lengths, obtained by Glasgow university atomistic simulator.

The leakage current variation has a span of almost 3 orders of magnitude and the drive current variation, although underestimated by drift diffusion simulation, still has a spread of almost 50% of the mean drive current. Even in a practical digital circuit design where minimum width devices are rare, and many devices are subject to averaging effects due to increased device width, this level of statistical variation cannot be ignored. Due to limitation in the possible supply voltage scaling, circuit power density has begun to become prohibitive and intrinsic parameter fluctuations play an important role in the power crisis [30]. Variability of threshold voltage and leakage current are directly responsible for increased margins in the power, speed and yield design trade off. Intrinsic parameter fluctuations have already started to affect the performance of digital systems [78].

Since compact models act as the bridge between the designer and the foundry, statistical compact modelling is the obvious way to effectively transfer device statistical variability information to designers. The investigation and development of flexible and accurate, yet

economical strategies for capturing statistical variability in industrial standard compact models is of great importance for variability aware design. Integrating the effect of these fluctuations into compact models can result in the correct physical prediction of circuit characteristics variation in the dc, ac and transient regimes. As an example, Figure 2.9 shows a CMOS inverter transfer characteristic variation induced by RDD, LER and PGG based on a 35nm physical gate length technology. The n-MOSFET is a square device with 35nm gate width and length and the p-MOSFET has 35nm gate length and 70nm gate width. The choice of 2:1 for p-MOSFET to n-MOSFET width has been chosen to balance the drive currents due to reduced mobility of holes compared with electrons. The supply voltage is 1Volt. The simulation results can provide valuable yield information for inverter designers. Two inverter figures of merit are extracted and their distributions are shown in Figure 2.9. The switching threshold voltage is where the output voltage equals the input voltage and the gain is measured as the maximum slope of output transfer characteristics.

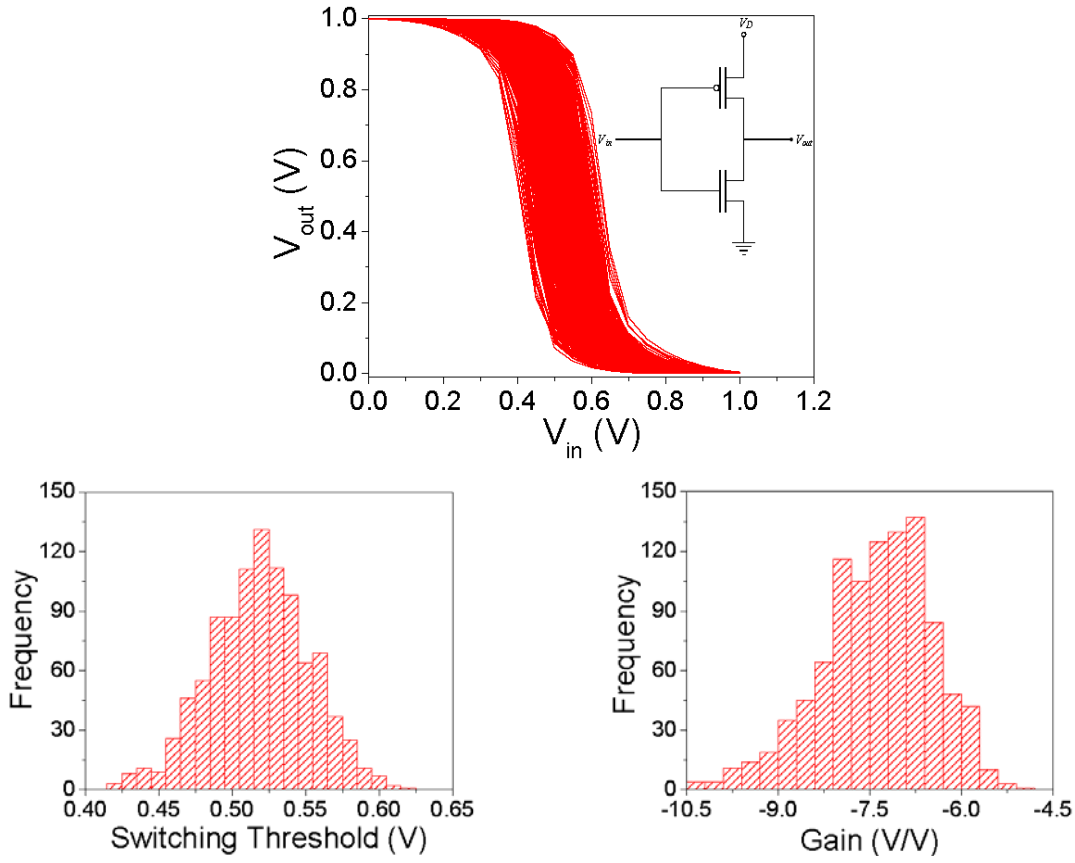


Figure 2.9: The CMOS inverter transfer characteristic variations and distributions of the switching threshold and gain induced by RDD, LER and PGG.

2.5 Review of Statistical MOSFET Models

The following is a brief review of various statistical modelling approaches published in the literature. Statistical MOSFET models are based on various statistical methods. Some of them are applicable in analytical or numerical solutions and some others have a particular target compact model. It is aimed to highlight the basic idea behind each method and their main assumptions and limitations.

Corner models have been in use from early stages of the semiconductor industry. Although statistical variability was not a major concern at that time, process induced variability was the major obstacle to determining exact values of parameters in the design stage. Effects such as lateral diffusion of source/drain, mask alignment tolerances, sheet resistance variations and imperfect etching were the main sources of variability due to manufacturing process inaccuracies. Hence, MOSFET physical parameters such as length and width and doping profiles were subject to process induced variations and reducing the cost and increasing the yield of fabrication procedure was not possible without taking these variations into account. Statistical corner models were developed to represent the worst and best case device performance [79,80,81,82,83,84,85]. The factor determining the performance depends to the application for which the circuit is designed. For instance, in digital CMOS design, I_{on} or the drive current, is a convenient performance factor because it is inversely proportional to the delay. There are four types of corners describing the relative performance of n- and p-channel MOSFETs subject to process variability [8, 101]: SS (slow p-channel, slow n-channel), FF (fast p-channel, fast n-channel), FS (fast p-channel, slow n-channel) and SF (slow p-channel, fast n-channel). The main drawbacks of these corner models are:

1- It gives an idea about the limits of the device/circuit performance but it cannot be used to estimate the number of devices in the tails of the performance distribution and the corresponding yield.
2- It assumes that the device performance factors are statistically independent. Hence, the correlation between performance factors will be ignored.

3- Although corner models guarantee good yield, they are pessimistic [8]. In other words, they over-estimate the actual performance spread. This can lead to large chip area and power consumption and hence, increased cost [86].
4- They cannot easily be used in deca-nano meter CMOS technologies where the effects of statistical variability are added to the effects of global or process induced variability. Figure 2.10 illustrates these corners in a 90nm technology node where the impacts of both local and global variability are taken into account [36]. It indicates that the variation due to local variability is as equal importance to the global or process induced variability.

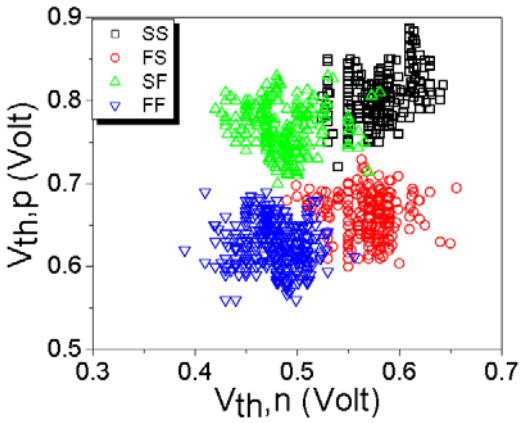


Figure 2.10: Simulation of 200 p- and n-MOS transistors in 90nm process corners, after [36].

A performance aware corner model approach was presented in [17] with both global and local variations taken into account. A more reliable corner modelling approach is presented in [86]. These improved corner models are called "statistical corner models" against "fixed corner models" [8]. They use a database of electrical test (ET) data or electrical current-voltage (*I-V*) characteristics. The critical stage of this modelling approach is to determine SPICE parameters to map variations in ET data or *I-V* data into model. Then the corners are obtained by adding a realistic standard deviation to the typical value of the parameters. However, modelling parameter correlations will still be a challenge.

Principal Component Analysis (PCA) is a statistical technique that is used to identify patterns in multi-dimensional data. It takes into account the correlation between statistical

parameters [87]. It is used for the purpose of statistical modelling of MOSFETs in different applications [88,89,90]. Using principal components of measured data as the fitting target of a statistical compact model is proposed in [90]. PCA based approaches rely on the assumption that variations follow normal distributions and this is one of the significant limitations of this method. However, if the performance parameters are close to normal distribution, PCA provides an accurate statistical model as it considers all statistical correlations.

Statistical modelling with surface potential based compact models is presented in [18,91]. Statistical modelling with the PSP (Pennsylvania Surface Potential-based) MOSFET model is proposed in [18], which is based on backward propagation of variance (BPV). The main benefit of BPV method is that it can model higher moments of MOSFET electrical performances and their correlations with an analytical formulation. However, in order to do this task, one needs to find out a set of normally distributed and independent process parameters with their corresponding parameters in SPICE, and this is a challenging requirement. The mathematical details of BPV statistical modelling method can be found in [92]. Statistical modelling of different MOSFET electrical parameters with another surface potential compact model, HiSIM (Hiroshima University STARC IGFET Model), is presented in [91]. Statistical compact modelling based on a data sampling strategy was proposed in [93].

Negative or positive bias temperature instability (NBTI or PBTI) is also becoming a hot research topic in CMOS statistical compact modelling [94,95,96]. Compact modelling of NBTI considering process variations is reviewed in [97] and modelling of statistical variability in presence of NBTI is presented in [98]. A well accepted simulation tool that can investigate the impact of reliability (which was presented in the context of hot-electron degradation) on NMOS circuit was reported in [99], and it had been further extended to p-MOSFET in [100]. An analytical model for NBTI was proposed in [94]. A compact modelling study of NBTI was reported in [95], and an analytic solution of NBTI-induced circuit aging under the influence of process variation was presented in [96], however, it does not consider the statistical nature of NBTI. The involvement of atomic-scale defects

in NBTI was reported in [97]. A composite model which can be extended to include NBTI statistical variability was presented in [98].

2.6 Summary

In this chapter, the challenges associated with device scaling were discussed with an emphasis on statistical variability. A classification of variability and important sources of intrinsic parameter fluctuations was presented. Simulation techniques used to study variability were outlined and the drift diffusion technique was described in more detail. The impact of intrinsic parameter fluctuation on device/circuit characteristics was outlined. Finally, the existing statistical models which have been reported in the literature were reviewed and their main benefits and drawbacks were highlighted. The aim of this background chapter was to study the major sources of statistical variability and their impact on the transistor electrical characteristics. This is particularly important in order to establish a bridge between the data obtained from 'atomistic' simulation of statistical variability in device level to circuit level using a statistical compact model.

Chapter 3

Uniform Device BSIM4 Parameter Extraction and Optimization

In modern circuit design, circuit simulation is essential due to the complexity and component count of circuits, and the requirements to predict circuit behaviour, optimize circuit performance, and allow for manufacturing tolerances at the design stage [101]. Accurate device models are at the heart of any circuit simulator and it has been found that in circuits based on metal-oxide-semiconductor field-effect transistors (MOSFETs), more than 70% of total circuit analysis time can be taken in evaluating currents at the terminals of each transistor [102]. Thus, the simplicity and accuracy of the MOSFET models used in a circuit simulator will directly affect the corresponding accuracy and speed of circuit simulation.

In general, transistor device modelling falls in two categories: Technology Computer Aided Design (TCAD) physical models and equivalent circuit models. TCAD physical models are more accurate because they solve semiconductor equations for each terminal bias condition, but due to the 3D nature of physical effects in small transistors, the analysis time is considerable, and TCAD is mainly useful for device design purposes. Equivalent circuit models simplify the device characteristics by using circuit elements which are

derived either as analytical expressions or via a table look up approach. The circuit simulator SPICE uses these equivalent circuit models as they have much higher computational efficiency than TCAD physical device models [1].

This chapter starts with an overview of the BSIM4 compact model and its analytical equations, which are embedded in the heart of any circuit simulator like HSPICE. In the next section the design of the 35nm gate length MOSFET used as an exemplar in this thesis will be reviewed. Finally, using this 35nm device as an example, we describe in detail our BSIM4 parameter extraction strategy for uniform devices.

3.1 BSIM4 Compact Models

BSIM (Berkeley Short-channel Insulated-gate field-effect-transistor Model) is a physical-based MOSFET model for circuit simulation which is developed by Berkeley University. This model has been widely used for the simulation of planar bulk MOSFETs and has served the industry for 20 years [103].

The early generations of BSIM models used separate model expressions for different device operating regimes such as sub-threshold and strong inversion [104]. The expressions accurately described device behaviour within their respective regime of operation. However, matching problems occurred in the transition between the sub-threshold and strong inversion regimes. With the arrival of the first generation of industry standard compact model BSIM3v3 and the new BSIM4 generations, the models have been designed to not only preserve region-specific device physics but also to ensure the continuities of current (I_{ds}), conductance (G_x) and their derivatives with respect to all the terminal voltages (V_{gs}, V_{ds} and V_{bs}) to prevent nonphysical results in circuit simulation.

BSIM3v3 describes a device's current-voltage characteristics from sub-threshold to strong inversion as well as from the linear to the saturation operating regimes with a single expression and guarantees the continuity of current and conductance at all voltage bias conditions. Furthermore, the model accounts for all the major physical effects in state-of-

the art MOSFET devices such as threshold voltage roll-off, non-uniform doping effect, mobility reduction due to the vertical field, carrier velocity saturation, channel length modulation (CLM), drain induced barrier lowering (DIBL), sub-threshold conduction, the parasitic resistance effect, and so on [2]. BSIM3v3 has extensive built-in dependencies on important dimensional and processing parameters such as channel length, width, gate oxide thickness, junction depth, substrate doping concentration, etc. This allows users to accurately model the MOSFET over a wide range of channel lengths and widths for various technologies and makes the model flexible enough to be used for statistical modelling, as discussed in the next chapter. BSIM4 is an extension of the BSIM3v3 model and addresses the physical effects which become important for sub-100nm gate length devices.

In this thesis we focus on the capability of BSIM in capturing the basic MOSFET drain current behaviour at advanced technology nodes. Hence, some second order effects such as junction and gate leakage currents and noise are not included in this study.

3.1.1 Gate Dielectric Model

With the shrinking of equivalent gate oxide thickness in modern MOSFETs to less than 1.2nm, the effect of channel quantization, which arises from the finite charge layer thickness (FCLT) in the channel, becomes significant. BSIM4 accounts for this effect in both static (DC) and dynamic (AC) models [105]. To activate the effect of FCLT in the simulation, the equivalent electrical gate oxide thickness $TOXE$[1], the physical gate oxide thickness $TOXP$, or their difference $DTOX = TOXE - TOXP$ can be used as input parameters. The appendix lists the model parameters. Based on these parameters, BSIM4 can model the FCLT effect by introducing an effective gate oxide capacitance C_{oxeff} in both *I-V* and *C-V* models:

[1] Capital and italic alphanumeric variable names in this thesis designate model parameters.

$$C_{oxeff} = \frac{C_{oxe} C_{cen}}{C_{oxe} + C_{cen}} \tag{3.1}$$

where C_{oxe} and C_{cen} are called equivalent gate oxide capacitance and centroid channel charge capacitance, respectively, and are given as follows:

$$C_{oxe} = \frac{\varepsilon_0 \times EPSROX}{TOXE} \tag{3.2}$$

$$C_{cen} = \frac{\varepsilon_{si}}{X_{DC}} \tag{3.3}$$

where *EPSROX* is a model parameter which describes the gate dielectric constant relative to vacuum; ε_0 and ε_{si} are the permittivity of free space and Silicon, respectively; X_{DC} is the equivalent DC centroid of the channel charge layer which is given by:

$$X_{DC} = \frac{1.9 \times 10^{-11}}{1 + \left(\frac{V_{gsteff} + 4(VTH0 - VFB - \Phi_S)}{2TOXP} \right)} \tag{3.4}$$

where *VTH*0 and *VFB* are model parameters for the long-channel threshold voltage at zero V_{BS} and the flat-band voltage, respectively; V_{gsteff} is effective gate voltage that will be discussed in detail in section 3.1.4; and Φ_S is the surface potential.

3.1.2 Effective DC and AC Channel Length and Width

The effective channel length and width take into account the source/drain sub-diffusion and the impact of shallow trench isolation (STI). Parameters *XL* and *XW* are introduced in BSIM4 to account for the channel length and width offset due to the processing factors such as mask and etch effects and processing non-uniformity [106]. The effective channel length and width for DC calculations (current-voltage characteristics) are given by:

$$L_{eff} = L_{drawn} + XL - 2dL \tag{3.5}$$

$$dL = LINT + \frac{LL}{L_{eff}^{LLN}} + \frac{LW}{W_{eff}^{LWN}} + \frac{LWL}{L_{eff}^{LLN} W_{eff}^{LWN}} \tag{3.6}$$

$$W_{eff} = W_{drawn} + XW - 2dW \tag{3.7}$$

$$\begin{aligned} dW = {} & WINT + \frac{WL}{L_{eff}^{WLN}} + \frac{WW}{W_{eff}^{WWN}} + \frac{WWL}{L_{eff}^{WLN} W_{eff}^{WWN}} + V_{gsteff}\, DWG \\ & + DWB\left(\sqrt{\Phi_S - V_{bseff}} - \sqrt{\Phi_S}\right) \end{aligned} \tag{3.8}$$

where *LINT*, *LL*, *LW*, *LWL*, *LLN* and *LWN* are model parameters to describe the dependence of *dL* on device geometry; *WINT*, *WL*, *WW*, *WWL*, *WLN* and *WWN* are also additional model parameters to describe the geometry dependence of *dW*. The *DWG* and *DWB* parameters are used to account for both the gate and substrate bias effects while V_{gsteff} and V_{bseff} are effective gate and substrate biases. Figure 3.1 illustrates the definitions of *XL*, *XW*, *dW*, *dL*, W_{eff} and L_{eff}.

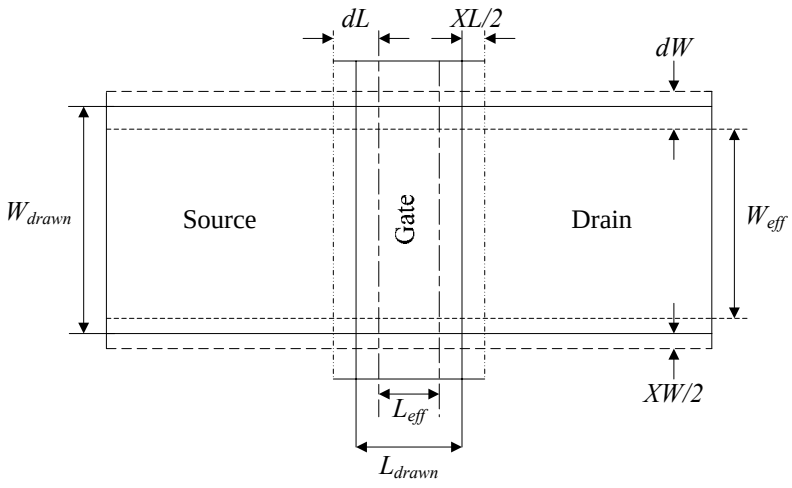


Figure 3.1: Definition of *XL*, *XW*, *dW*, *dL*, W_{eff} and L_{eff}.

For the capacitance (AC) or transient time simulations, the effective channel length and width are represented by L_{active} and W_{active}, and are defined by:

$$L_{active} = L_{drawn} + XL - 2dLC \tag{3.9}$$

$$dLC = DLC + \frac{LLC}{L_{eff}^{LLN}} + \frac{LWC}{W_{eff}^{LWN}} + \frac{LWLC}{L_{eff}^{LLN} W_{eff}^{LWN}} \tag{3.10}$$

$$W_{active} = W_{drawn} + XW - 2dWC \tag{3.11}$$

$$dWC = DWC + \frac{WLC}{L_{eff}^{WLN}} + \frac{WWC}{W_{eff}^{WWN}} + \frac{WWLC}{L_{eff}^{WLN} W_{eff}^{WWN}} \tag{3.12}$$

where *DLC*, *LLC*, *LWC* and *LWLC* are model parameters to describe the relation between *dLC* and device geometry parameters, and *DWC*, WLC, *WWC* and *WWLC* are model parameters to describe the geometry dependence of *dWC*. At default values, $DLC = LINT$ $DWC = WINT$ and *DWC*, *DLC*, *LLC*, *LWC*, *LWLC*, WLC, *WWC* and *WWLC* will be equal to the values of their DC counterparts. As shown in Figure 3.2, BSIM4 introduced *DWJ* to calculate the effective source/drain diffusion width, W_{effcj} which is used to calculate source/drain series resistance, gate resistance and gate-induced drain leakage (GIDL).

$$\begin{aligned} W_{effcj} &= W_{drawn} - 2dWJ \\ &= W_{drawn} - 2\left(DWJ + \frac{WLC}{L_{eff}^{WLN}} + \frac{WWC}{W_{eff}^{WWN}} + \frac{WWLC}{L_{eff}^{WLN} W_{eff}^{WWN}} \right) \end{aligned} \tag{3.13}$$

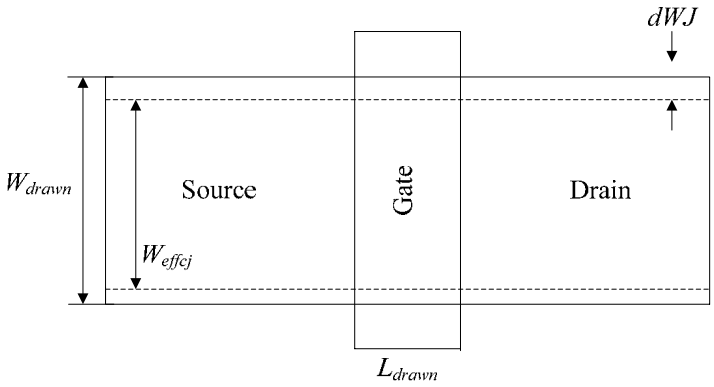


Figure 3.2: Definitions of *DWJ* and W_{effcj}

3.1.3 Threshold Voltage Model

Several important phenomena are considered in the modelling of threshold voltage (V_{th}) for sub 100nm MOSFETs. They are halo (pocket) implantation or lateral channel engineering, short channel effects (SCE) and narrow width effects. The main reason for using halo implantation is to reduce the V_{th} roll-off effect when the channel length is decreased [107]. The usual way of creating these halo regions is to implant extra dopants near the source/drain junctions compared with the middle of the channel. Figure 3.3 illustrates the doping profile of modern MOSFETs with emphasis on the 35nm gate length device which will be used as a test bed in this thesis. For the 35nm device, additional halo dopant distributions around the source and drain overlap each other and hence results in an increase in the net doping along the channel.

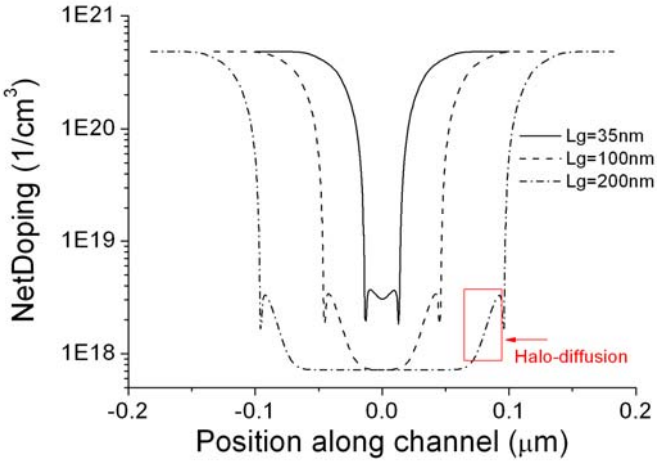


Figure 3.3: Doping profile along the channel in a 35nm, 100nm and 200nm n-MOSFET with halo implantation from [108].

To model the V_{th} roll-off under the influence of non-uniform lateral doping (NULD) caused by halo dopants, BSIM4 uses the following V_{th} model.

$$V_{th} = VTH0 + K1\left(\sqrt{\Phi_S - V_{BS}} - \sqrt{\Phi_S}\right)\sqrt{1+\frac{LPEB}{L_{eff}}} - V_{BS}K2 + K1\left(\sqrt{1+\frac{LPE0}{L_{eff}}} - 1\right)\sqrt{\Phi_S} \tag{3.14}$$

where *VTH*0, *K*1, *K*2, *LPE*0 and *LPEB* are BSIM4 model parameters. *K*1 and *K*2 are first and second body bias coefficient, respectively. They have analytical expressions based on approximated doping profile but treating them as parameters will give better fitting results; *LPE*0 is the lateral non-uniform doping parameter at Φ_S is the surface potential which is defined as $V_{BS}=0$ and is introduced to consider the increase of effective doping concentration for short channel devices due to the halo implant. *LPEB* is related to the lateral non-uniform doping effect on *K*1 which is added to the model to account for the impact of halo implant on the substrate bias effect. Furthermore halo implants cause a drain-induced threshold voltage shift (DITVS) in long channel devices, which is described by the following equation:

$$\Delta V_{th,DITVS} = -n \times V_t \times \ln\left(\frac{L_{eff}}{L_{eff} + DVTP0\left(1+e^{-DVTP1\cdot V_{DS}}\right)}\right) \tag{3.15}$$

where *n* is subthreshold swing, V_t is the thermal voltage which equals $\frac{kT}{q}$, *DVTP*0 and *DVTP*1 are the coefficients of drain-induced V_{th} shift for long-channel pocket implanted devices; L_{eff} is the effective channel length and V_{DS} is the drain/source bias.

In short channel devices, the value of the threshold voltage not only depend on channel length but is also affected by the drain voltage. This can be stated by:

$$\Delta V_{th,SCE} = \Delta V_{th}\left(L_{eff}\right) + \Delta V_{th}\left(V_{DS}\right) \tag{3.16}$$

where $\Delta V_{th}(L_{eff})$ is the threshold voltage change caused by the SCE without the impact of drain/source bias and $\Delta V_{th}(V_{DS})$ is the change in threshold voltage due to non-zero V_{DS}. The second term in equation (3.16) arises from the fact that the depletion layer thickness will be modulated by drain bias voltage as shown in Figure 3.4. Increasing the drain bias will increase the depletion layer width near the drain which in turn will reduce the gate controlled charge in the channel thereby lowering V_{th} compared to its value when $V_{DS} = 0$. This effect is called Drain Induced Barrier Lowering (DIBL) [101].

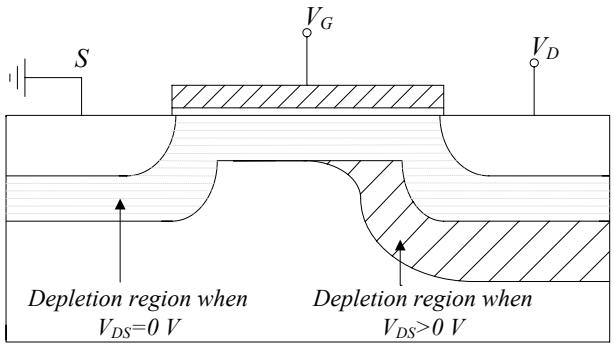


Figure 3.4: Depletion layer thickness modulation caused by applied V_{DS}.

By solving the quasi 2D Poisson equations along the channel of an ideal device, the basic physical behaviour of the short channel effect can be described by following equation [109]:

$$\Delta V_{th,SCE} = -\theta_{th}(L_{eff})\left[2(V_{bi} - \Phi_S) + V_{DS}\right] \tag{3.17}$$

where V_{bi} is the source/drain junction built-in voltage and $\theta_{th}(L_{eff})$ is used as a basic form of physical expression for SCE (width parameters are introduced into it to consider the complexity of model in reality [2]) and is given by:

$$\theta_{th}\left(L_{eff}\right)=\frac{1}{2\cosh\left(\frac{L_{eff}}{l_t}\right)-1} \tag{3.18}$$

where l_t is called the characteristic length. The effect of SCE in BSIM4 is described by following equation:

$$\Delta V_{th,SCE}=-\frac{DVT0}{2\cosh\left(DVT1\frac{L_{eff}}{l_t}\right)-1}\left(V_{bi}-\Phi_S\right)-\frac{\left(ETA0+ETAB\times V_{BS}\right)V_{DS}}{2\cosh\left(DSUB\frac{L_{eff}}{l_{t0}}\right)} \tag{3.19}$$

where

$$V_{bi}=\frac{k_BT}{q}\ln\left(\frac{NDEP\cdot NSD}{n_i^2}\right) \tag{3.20}$$

$$l_t=\sqrt{\frac{\varepsilon_{si}X_{dep}TOXE}{EPSROX}}\left(1+V_{BS}DVT2\right) \tag{3.21}$$

$$l_{t0}=\sqrt{\frac{\varepsilon_{si}X_{dep0}TOXE}{EPSROX}} \tag{3.22}$$

where *DVT*0 and *DVT*1 are model parameters which are called the first and second coefficients of SCE on V_{th} respectively; *ETA*0 is a model parameter which is named DIBL coefficient in sub-threshold region; *ETAB* is another model parameter to account for body bias coefficient for the sub-threshold DIBL effect; *DSUB* is model parameter to describe length dependent DIBL behaviour; *NSD* is the doping concentration in the source/drain regions; *NDEP* is the doping concentration at the edge of the channel depletion layer at $V_{BS}=0$; *EPSROX* was introduced in equation (3.2); *TOXE* is the equivalent electrical gate oxide thickness; *DVT*2 is a model parameter to account for body-bias coefficients of SCE

on V_{th}; X_{dep} is depletion layer width in the channel with the influence of V_{BS}; X_{dep0} is depletion layer width in the channel when $V_{BS} = 0$:

$$X_{dep} = \sqrt{\frac{2\varepsilon_{si}\left(\Phi_S - V_{BS}\right)}{q \cdot NDEP}} \tag{3.23}$$

$$X_{dep0} = \sqrt{\frac{2\varepsilon_{si}\Phi_S}{q \cdot NDEP}} \tag{3.24}$$

The narrow width effect (NWE) is very process sensitive and no universally accurate physical model is available. In BSIM4, an empirical approach is employed to take into account the overall NWE. By introducing several fitting parameters, *K*3, *K*3*B* and *W*0, the narrow width effect is modelled in BSIM4 by:

$$\begin{aligned}\Delta V_{th,W} = &\left(K3 + V_{BS}\, K3B\right)\frac{TOXE}{W_{eff} + W0}\Phi_S \\ &+ \frac{DVT0W}{2\cosh\left(DVT1W\,\frac{L_{eff}\,W_{eff}}{l_{t0}}\right) - 1}\left(V_{bi} - \Phi_S\right)\end{aligned} \tag{3.25}$$

where the second term is introduced into equation (3.25) to take into accounts the dependence of NWE in short channel length devices. *K*3 is narrow width coefficient; *K*3*B* is body effect coefficient of *K*3; *W*0 is the narrow width parameter and *DVT*0*W* and *DVT*1*W* are model parameters which are called first and second coefficient of NWE on V_{th} for small channel length, respectively.

3.1.4 Channel Charge Model

V_{gsteff} is the effective gate voltage which is the core of the BSIM model. It is given by:

$$V_{gsteff} = \frac{nV_t \ln\left\{1+\exp\left[\frac{m^*\left(V_{gse}-V_{th}\right)}{nV_t}\right]\right\}}{m^* + nC_{oxe}\sqrt{\frac{2\Phi_S}{q\cdot NDEP\varepsilon_{si}}}\exp\left[-\frac{\left(1-m^*\right)\left(V_{gse}-V_{th}\right)-V'_{off}}{nV_t}\right]} \tag{3.26}$$

where $m^* = 0.5 + \arctan\left(MINV\right)/\pi$, and the parameter *MINV* is a model parameter introduced to improve the model accuracy in the moderate inversion region; *n* is the sub-threshold swing parameter, and is modeled in BSIM4 by:

$$n = 1 + NFACTOR\frac{C_{dep}}{C_{oxe}} + \frac{CIT}{C_{oxe}} + \frac{CDSC + V_{DS}\,CDSCD + V_{bseff}\,CDSCB}{2C_{oxe}\cosh\left(DVT1\frac{L_{eff}}{l_t}\right)-1} \tag{3.27}$$

where *NFACTOR*, *CIT*, *CDSC*, *CDSCD*, *CDSCB* describe the sub-threshold swing factor, interface trap capacitance, coupling capacitance between source/drain and channel, drain bias sensitivity of *CDSC* and body-bias sensitivity of *CDSC*, respectively. The second term of equation (3.27) models the sub-threshold swing factor for long channel devices by using the model parameter *NFACTOR*, the third term introduces the effect of interface states with model parameter *CIT* and the last term models the coupling between drain/source and channel [105].

V_{gse} is the effective gate voltage including the poly-silicon gate depletion effect; V_{th} is the threshold voltage; V'_{off} is a potential offset parameter which equals to $VOFF + VOFFL / L_{eff}$ and describes the channel-length dependence of V'_{off} in devices with non-uniform doping profile; L_{eff} is the effective channel length; V_t is the thermal voltage and is equal to $k_B T/q$; Φ_S is the surface potential which is defined as:

$$\Phi_S = 0.4 + \frac{k_B T}{q} \ln\left(\frac{NDEP}{n_i}\right) + PHIN \tag{3.28}$$

where *NDEP* was introduced in equation (3.20); n_i is the intrinsic carrier concentration in the channel region, and *PHIN* is a model parameter to describe the non-uniform vertical doping effect on surface potential.

The aim of the introduction of V_{gsteff} in BSIM is to describe channel charge density from sub-threshold to strong inversion regions in a unified way [110]. With V_{gsteff}, the channel charge density at a location *y* along the channel can be described by:

$$Q_{ch}(y) = C_{oxeff} V_{gsteff}\left[1 - \frac{V_F(y)}{V_b}\right] \tag{3.29}$$

where $V_F(y)$ is the quasi-Fermi potential at specified point *y* along the channel with respect to the source and V_b is given by:

$$V_b = \frac{V_{gsteff} + 2V_t}{A_{bulk}} \tag{3.30}$$

where A_{bulk} is a factor describing the bulk charge effects and is given by:

$$\begin{aligned} A_{bulk} = &\left\{1 + F_{NUD}\left[\frac{A_0 L_{eff}}{L_{eff} + 2\sqrt{X_{dep} XJ}}(1 - AGS)V_{gsteff}\right]\left(\frac{1}{1 + V_{bseff} KETA}\right)\right\} \\ &+ \left\{F_{NUD}\left[\left(\frac{L_{eff}}{L_{eff} + 2\sqrt{X_{dep} XJ}}\right)^2 + \frac{B0}{W_{eff} + B1}\right]\left(\frac{1}{1 + V_{bseff} KETA}\right)\right\} \end{aligned} \tag{3.31}$$

where *A*0, *AGS*, *B*0, *B*1 and *KETA* are model parameters to describe coefficient of channel length dependence of bulk charge effect, coefficient of V_{gs} dependence of bulk charge effect, bulk charge effect coefficient for channel width, bulk charge effect width offset and body bias coefficient of bulk charge effect, respectively. *XJ* is the source/drain junction depth and F_{NUD} is used to model the non-uniform doping effects by:

$$F_{NUD} = \frac{\sqrt{1+\frac{LPEB}{L_{eff}}}K1}{2\sqrt{\Phi_S - V_{bseff}}} + K2 - K3B\frac{TOXE}{W_{eff}+W0}\Phi_S \tag{3.32}$$

3.1.5 Mobility Model

There are three scattering mechanisms that explain the carrier transport behavior. They are phonon scattering, Coulomb scattering and surface roughness scattering. Under specific conditions of voltage bias, doping concentration and temperature one of these mechanisms may be dominant. A universal mobility model is defined by [111]:

$$\mu_{eff} = \frac{\mu_0}{1+\left(E_{eff}/E_0\right)^{\upsilon}} \tag{3.33}$$

where μ_0 is the low-field mobility; E_0 is called the critical electric field; υ is a constant which depends on device type and technology and E_{eff} is an effective field defined experimentally by:

$$E_{eff} = \frac{Q_B + Q_{INV}/2}{\varepsilon_{si}} \tag{3.34}$$

where Q_B and Q_{INV} are the charge density in the bulk and in the channel, respectively. BSIM4 provides three different models of the effective mobility by using a flag which is named *mobMod*. If *mobMod = 0* or *1*, models are selected based on built-in BSIM4

expressions and if mobMod = *2*, a universal mobility model based on equation (3.33) will be considered which is more accurate and suitable for predictive modelling [2]. Effective mobility expressions based on possible selection of mobility flag follows:

when *mobMod = 0*

$$\mu_{eff} = \frac{U_0(T)}{1+\left(U_A(T)+U_C(T)V_{bseff}\right)\left(\frac{V_{gsteff}+2V_{th}}{TOXE}\right)+U_B(T)\left(\frac{V_{gsteff}+2V_{th}}{TOXE}\right)^2} \tag{3.35}$$

when *mobMod = 1*

$$\mu_{eff} = \frac{U_0(T)}{1+\left(1+U_C(T)V_{bseff}\right)\left[U_A(T)\left(\frac{V_{gsteff}+2V_{th}}{TOXE}\right)+U_B(T)\left(\frac{V_{gsteff}+2V_{th}}{TOXE}\right)^2\right]} \tag{3.36}$$

when *mobMod = 2*

$$\mu_{eff} = \frac{U_0(T)}{1+\left(U_A(T)+U_C(T)V_{bseff}\right)\left(\frac{V_{gsteff}+C_0\left(VTH0-VFB-\Phi_S\right)}{TOXE}\right)^{EU}} \tag{3.37}$$

where C_0 is a constant which equals to 2 for n-MOS (n-channel Metal Oxide Semiconductor) and 2.5 for p-MOS (p-channel Metal Oxide Semiconductor) devices; *U*0, *UA*, *UB*, *UC* and *EU* are model parameters to describe the coefficient of first-order mobility degradation due to the vertical field; the coefficient of second-order mobility degradation due to the vertical field; and the coefficient of mobility degradation due to the body-bias effect and exponent for mobility degradation of *mobMod = 2*, respectively. Figure 3.5 shows the trend of mobility versus V_{GS} for different *mobMod*.

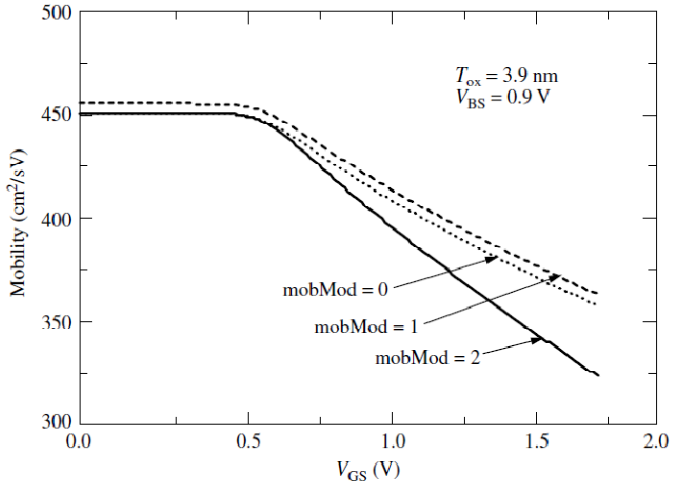


Figure 3.5: V_{GS} dependence of mobility behavior to different *mobMod* options from [105].

3.1.6 Source/Drain Resistance Model

There is a flag called *rdsMod* which was introduced in BSIM4 to select different source/drain resistance models. If *rdsMod = 0*, a symmetric source/drain resistance model is assumed; *rdsMod = 1* the external source/drain resistance model will be asymmetric. When *rdsMod = 0,* the drain/source resistance is expressed by:

$$R_{DS} = \frac{R_{DSWMIN}(T) + R_{DSW}(T)\left[PRWB\left(\sqrt{\Phi_S - V_{bseff}} - \sqrt{\Phi_S}\right) + \frac{1}{1 + PRWGV_{gsteff}}\right]}{\left(10^6 W_{effcj}\right)^{WR}} \tag{3.38}$$

where *PRWB*, *PRWG* and *WR* are model parameters describing the body-bias dependence of the low-doped drain resistance; the gate-bias dependence of the low-doped drain (LDD) resistance and the channel width dependence parameter of the LDD resistance, respectively; W_{effcj} was described in equation (3.13); $R_{DSWMIN}(T)$ and $R_{DSW}(T)$ are two temperature varying parameters described by:

$$R_{DSW}(T) = RDSW + PRT\left(\frac{T}{TNOM} - 1\right) \tag{3.39}$$

$$R_{DSWMIN}(T) = RDSWMIN + PRT\left(\frac{T}{TNOM} - 1\right) \tag{3.40}$$

where *RDSWMIN* and *RDSW* are model parameters to describe the source/drain resistance at *TNOM* (nominal temperature) and *PRT* is a temperature coefficient for *RDSW*.

When *rdsMod = 1,* the source/drain resistance model will be given by two separate equations:

$$R_D = \frac{R_{DWMIN}(T) + R_{DW}(T)\left[-V_{BD}\,PRWB + \frac{1}{1 + PRWG\left(V_{GD} - V_{fbSD}\right)}\right]}{N_F\left(10^6 W_{effcj}\right)^{WR}} \tag{3.41}$$

$$R_S = \frac{R_{SWMIN}(T) + R_{SW}(T)\left[-V_{BS}\,PRWB + \frac{1}{1 + PRWG\left(V_{GD} - V_{fbSD}\right)}\right]}{N_F\left(10^6 W_{effcj}\right)^{WR}} \tag{3.42}$$

where V_{fbSD} is the calculated flat-band voltage between the gate and source/drain diffusion regions.

$$V_{fbSD} = \frac{K_B T}{q} \ln \frac{NGATE}{NSD} \tag{3.43}$$

where *NGATE* and NSD are model parameters for the doping concentration in the gate and source/drain regions.

3.1.7 Drain Current Model

In MOSFETs, the following expression is used to take account of both drift and diffusion current in the linear region [112]:

$$I_d(y) = W_{eff}\, Q_{ch}(y)\, \mu_{eff}(y) \frac{dV_F(y)}{dy} \tag{3.44}$$

where W_{eff} was described in equation (3.7); $Q_{ch}(y)$ is given in equation (3.29); $\mu_{eff}(y)$ is the effective mobility and described previously in equation (3.33) and $V_F(y)$ is the carrier quasi-Fermi potential at a point y along the channel.

By integrating equation (3.44) from source to drain and substituting equation (3.29) into it, the expression of linear drain current without including the source/drain resistance is given by:

$$I_{ds0} = \frac{W_{eff}\, \mu_{eff}\, C_{oxeff}\, V_{gsteff}\, V_{dseff} \left(1 - \frac{V_{dseff}}{2V_b}\right)}{L_{eff} + \frac{\mu_{eff}\, V_{dseff}}{2V_{SAT}(T)}} \tag{3.45}$$

where L_{eff} is given by equation (3.5) and V_{dseff} is introduced to ensure a smooth transition from triode to saturation region and is expressed as:

$$V_{dseff} = V_{dsat} - V_{DS} - DELTA + \sqrt{(V_{dsat} - V_{DS} - DELTA)^2 + 4V_{dsat}\, DELTA} \tag{3.46}$$

where $DELTA$ is the smoothing parameter; V_{dsat} is the saturation voltage and depends on source/drain resistance model which is used in the model. When *rdsMod = 0*:

$$V_{dsat} = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \tag{3.47}$$

where

$$a = A_{bulk}^2 W_{eff} VSAT C_{oxe} R_{DS} + A_{bulk}\left(\frac{1}{\lambda} - 1\right) \tag{3.48}$$

$$\begin{aligned} b = \left(V_{gsteff} + 2V_t\right)\left(1 - \frac{2}{\lambda}\right) - \frac{2A_{bulk} L_{eff} VSAT}{\mu_{eff}} \\ - 3A_{bulk} W_{eff} C_{oxe} R_{DS} VSAT \left(V_{gsteff} + 2V_t\right) \end{aligned} \tag{3.49}$$

$$c = 2\left(V_{gsteff} + 2V_t\right)\left(\frac{2L_{eff} VSAT}{\mu_{eff}} + W_{eff}\ C_{oxe} R_{DS} VSAT\right) \tag{3.50}$$

λ is introduced to model the non-saturation effects which are found for p-MOSFETs [2] and is described as:

$$\lambda = A1 \times V_{gsteff} + A2 \tag{3.51}$$

where $A1$ and $A2$ are two model parameters to describe first non-saturation effect parameter and second non-saturation factor, respectively. When *rdsMod = 1* equation (3.47) can be rewritten as follows:

$$V_{dsat} = \frac{E_{sat} L_{eff} \left(V_{gsteff} + 2V_t\right)}{A_{bulk} E_{sat} L_{eff} + V_{gsteff} + 2V_t} \tag{3.52}$$

where

$$E_{sat} = \frac{2VSAT}{\mu_{eff}} \tag{3.53}$$

where *VSAT* is a model parameter to account for the saturation velocity at nominal temperature. The complete single equation channel current model is given by:

$$I_{DS} = \frac{I_{ds0}}{1+\frac{R_{DS} I_{ds0}}{V_{dseff}}}\left[1+\frac{\ln\left(\frac{V_A}{V_{ASAT}}\right)}{C_{clm}}\right]\left(1+\frac{V_{DS}-V_{dseff}}{V_{ADIBL}}\right)\left(1+\frac{V_{DS}-V_{dseff}}{V_{ADITS}}\right)\left(1+\frac{V_{DS}-V_{dseff}}{V_{ASCBE}}\right) \tag{3.54}$$

where V_A is called the Early voltage and has two components: the Early voltage at the saturation voltage point and the Early voltage from the channel length modulation (CLM) effect:

$$V_A = V_{ASAT} + V_{ACLM} \tag{3.55}$$

VASAT and *VACLM* are given by:

$$V_{ASAT} = \frac{\left(E_{sat} L_{eff}\right)+V_{dsat}+2R_{DS} C_{oxe} W_{eff} V_{gsteff} VSAT\left(1-\frac{A_{bulk} V_{dsat}}{2\left(V_{gsteff}+2V_t\right)}\right)}{\left(R_{DS} C_{oxe} W_{eff} A_{bulk} VSAT\right)-1+\frac{2}{\lambda}} \tag{3.56}$$

$$V_{ACLM} = C_{clm}\left(V_{DS}-V_{dsat}\right) \tag{3.57}$$

$$C_{clm} = \frac{F_{pocket}}{PCLM \times litl}\left(L_{eff}+\frac{V_{dsat}}{E_{sat}}\right)\left(1+PVAG\frac{V_{gsteff}}{E_{sat} L_{eff}}\right)\left(1+\frac{R_{DS} I_{ds0}}{V_{dseff}}\right) \tag{3.58}$$

$$litl = \sqrt{\frac{\varepsilon_{si}\, TOXE \cdot XJ}{EPSROX}} \tag{3.59}$$

$$F_{pocket} = \frac{1}{1 + FPROUT \dfrac{\sqrt{L_{eff}}}{V_{gsteff} + 2V_t}} \tag{3.60}$$

where *PCLM*, *PVAG* and *FPROUT* are BSIM4 parameters for channel length modulation parameter, gate bias dependence of Early voltage and effect of pocket implant on output resistance degradation, respectively. The Early voltage contributed by DIBL is described as:

$$V_{ADIBL} = \left(\frac{V_{gsteff} + 2V_t}{\theta_{rout}\left(1 + V_{bseff}\, PDIBLCB\right)}\right)\left(1 - \frac{A_{bulk}\, V_{dsat}}{\left(A_{bulk}\, V_{dsat}\right) + V_{gsteff} + 2V}\right) \left(1 + PVAG \frac{V_{gsteff}}{E_{sat}\, L_{eff}}\right) \tag{3.61}$$

$$\theta_{rout} = PDIBLC2 + \frac{PDIBLC1}{2\cosh\left(\dfrac{L_{eff}\, DROUT}{l_{t0}}\right) - 2} \tag{3.62}$$

where *PDIBLC*1, *PDIBLC*2 and *DROUT* are model parameters that have been introduced to correct the DIBL effect in the strong inversion region. Other contributions of the Early voltage due to substrate current (V_{ASCBE}) effect and drain-induced threshold shift (V_{ADITS}) which exist in equation (3.54), are given in [3].

Finally, it should be noted that the different BSIM4 parameters have different impact on the electrical characteristics of the transistors. Some of them have strong impact on sub-threshold region (i.e., *VOFF*), some others can change the drain current characteristics in above-threshold region (i.e., *RDSW*) and some critical parameters like *VTH0* and *DSUB* can change I_d-V_g characteristics in both regions. The impact of important parameters on the MOSFET characteristics is visualized in Appendix 2 where a 20% increase or decrease is applied on the

nominal value of each parameter and the resultant curve is compared with the nominal transistor characteristics.

3.2 Template MOSFETs

The template devices which will be used as examples for the development our BSIM4 parameter extraction and optimization methodology are 35nm physical gate length poly-gate n- and p-MOSFETs. These are based on TCAD simulations and have been carefully designed to match the performance of recently published state-of-the art high performance 45nm technology generation MOSFETs [113,114]. Their design was initially based on the 35nm gate length transistor published by Toshiba in 2002 [115] but its structure was updated to incorporate the latest technology features embedded in 45nm CMOS technologies which were reported by Intel in 2007 and published in [113]. This use of TCAD has two major advantages. Firstly, it allows investigation of MOSFETs which include the effects of technology enhancements such as strain engineering, where all details, including extensive experimental measurements, are not yet in the public domain. Secondly TCAD allows aspects of these technologies to be examined in a way not always easy or practical with physical devices, in order to better understand their effects. Of course, the quality of the TCAD based models is critically dependent on the quality of TCAD tool and model calibration to existing experimental data.

Creation of a TCAD device model consists of two major steps: process simulation and device simulation [6,59]. The MOSFET structure and doping profiles are generated through process simulation and calibrated against measured or calculated structural information. Device simulation predicts device electrical performance based on this MOSFET structure and models of charge transport in the device. Calibration of device electrical characteristics adjusts these transport models, and can require additional adjustment of device structure or doping to obtain accurate modelling over the whole range of device operation.

In the 35nm poly gate n- and p-MOSFETs, the device structure features silicon oxynitride (SiON) as a gate dielectric, with 'equivalent oxide thickness (EOT)' of 1nm, and ultra-shallow S/D junction extensions with 20nm depth for n-MOSFET and 28nm for p-MOSFET. A retrograde channel doping profile is achieved by three channel implementations, and a 30^{o} tilt halo implementation is employed to further reduce short channel effects. Strain engineering is introduced in the last steps of the process simulation by using a tensile 'contact etch stop layer (CESL)' to increase the electron mobility in the channel for n-channel MOSFET, and with the formation of a SiGe epitaxial layer in the S/D regions of p-MOSFET and the deposition of compressive CESL layer over the MOSFET (introducing compressive strain) to increase the hole mobility. Figure 3.6 illustrates the doping profiles of the n- and p-MOSFETs after process simulation. The details about the accuracy of process calibration can be found in [59].

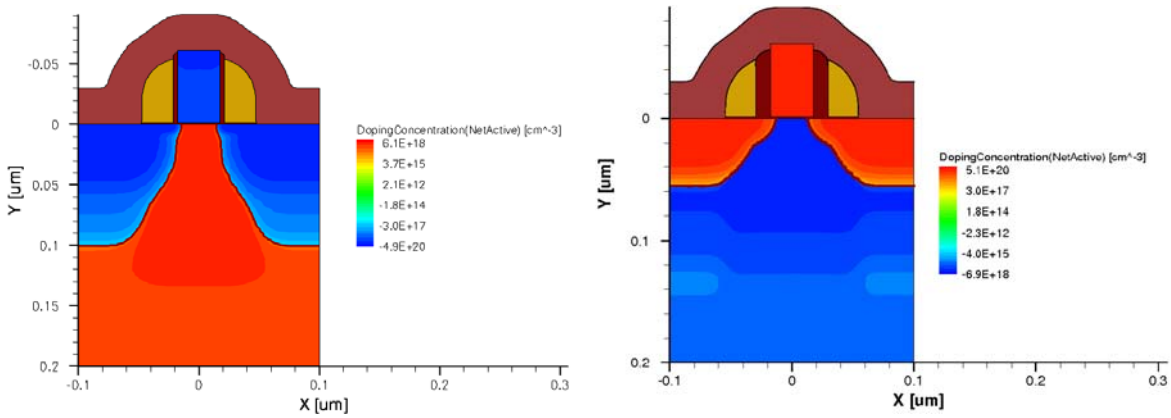


Figure 3.6: p-MOSFET (left) 35nm; n-MOSFET (right) 35nm device doping profile simulated using Sentaurus based on the standard modern process flow. These devices are enhanced with strain engineering to match the performance of 45nm technology generation counterparts. The areas adjacent to the gate are called spacers.

In device simulation, as noted above, calibration focuses on adjusting the models of charge transport, and specifically on parameters of the charge mobility models. Low field mobility is tuned to match the low drain voltage I_D-V_G characteristics from the measurement. By adjusting material saturation velocities and the critical field dependent

mobility model, high field mobility is tuned to produce calibrated I_D-V_G characteristics under the high drain bias conditions.

Table 3.1 presents key drain current figures of merit of the uniform 35nm n-MOSFET under study and compares the drive and leakage current of our TCAD device with real devices in the same technology node fabricated by Intel [113], TSMC [114] and IBM [116]. While the leakage current is exactly identical between real fabricated devices and the device under study, the drive current of the our device is 6.6% less than the Intel device, 5.8% more than TSMC device and 10.4% more than IBM device. Full sets of I_D-V_G and I_D-V_D characteristics for the TCAD device are shown in Figures 3.7 to 3.9 where they are compared with our BSIM4 model results.

Table 3.1: Drain current characteristic of n-MOSFET device illustrated in Figure 3.6 and comparison with real devices

Current / Device	TCAD	Intel	TSMC	IBM
I_d (V_d=V_g=1V)	1.27 mA	1.36 mA	1.20 mA	1.15 mA
I_d (V_d=1V&V_g=0V)	0.10 µA	0.10 µA	0.10 µA	0.10 µA

3.3 Parameter Extraction and Optimization Methodology

Having obtained an accurate electrical model of 35nm gate length devices from TCAD, BSIM4 parameters can be extracted to create an accurate SPICE model of the MOSFETs. The compact model parameter extraction methodology is the key to the results of this thesis, and so the uniform device extraction strategy is presented in detail below for 35nm devices. The same TCAD modelling and parameter extraction strategy was repeated for devices with 18nm channel length and the results are also recorded at the end of this section.

To obtain compact model parameters, two overarching extraction approaches can be used: single device extraction and group extraction. In single device extraction, measured data from a single device is used to extract a compact set of model parameters [3]. There are two problems in this approach: fitting to the single device will be accurate, but fitting to devices with differing geometries may not be. In addition there is no guarantee that those extracted parameters which are introduced to be related to the device dimensions will have physical meaning, as the geometry dependencies of these parameters cannot be determined.

Group extraction is used in this study, with both long channel and short channel devices used in order to extract compact model parameters. The long channel devices are used to extract parameters which are independent of short channel effects. Specifically, these are: mobility, the large-sized device threshold voltage *VTH*0, and the body effect coefficients *K*1 and *K*2 which depend on the vertical doping concentration distribution. A set of devices with shorter channel lengths are used to extract parameters which are related to the short channel effects, and the parameters which are geometry dependent. 30nm, 35nm, 40nm and 200nm devices are used to extract parameters of devices based on a technology generation with nominal channel lengths of 35nm. All devices are simulated under the same bias conditions. The resulting fit might not be absolutely perfect for any single device but will be better for the group of devices under consideration, and more useful in practice.

The basic transistor current-voltage characteristics needed in compact model extraction are:

1. I_D-V_G at low V_D bias and different V_B

2. I_D-V_G at high V_D bias and different V_B

3. I_D-V_D at different V_G bias and $V_B=0$

There are two main compact model parameter optimization strategies: global optimization and local optimization [3]. Global optimization relies on global fitting to find

one set of model parameters which will best fit the available experimental data. This methodology may give the minimum average error between measured and simulated data points, but it also treats each parameter as a fitting parameter. Physical parameters extracted in such a manner might yield values that are not consistent with their physical intention.

In local optimization, sequential groups of parameters are extracted independently of one another. Parameters are extracted from device bias conditions which correspond to dominant physical mechanisms. Parameters which are extracted in this manner might not fit experimental data in all bias conditions. Nonetheless, these extraction methodologies are developed specifically with respect to a given parameter's physical meaning. If properly executed, it should, predict device performance well, and values extracted in this manner will have physical relevance. In this work, local optimization is used.

Commercial optimisation software called 'Aurora' has been used for parameter extraction purpose [117]. Aurora extracts model parameters that produce a least-squares fit to the data in reference to physical device characteristics. Least-square optimisation algorithms are the standard method of optimization in many parameter extraction softwares [101].

Before extraction begins, compact model parameters which need no fitting, because they are directly obtained from experimentally measured or process simulated structural values are provided [117]. These are typically parameters related to gate oxide thickness and dielectric constant (*TOXE*, *TOXP*, *DTOX* or *EPSROX*), doping concentration in the channel (*NDEP*), temperature at which the measurements are performed (*TNOM*), mask level channel length (L_{drawn}) and mask level channel width (W_{drawn}) and junction depth (*XJ*).

Since the core of the BSIM4 model is based on long channel device physics (1D description of gate control), basic device parameters that describe long channel device operation such as threshold voltage parameters *VTH*0, *K*1, *K*2, sub-threshold region parameters *NFACTOR*, *VOFF*, source/drain resistance parameter *RDSW*, mobility

parameters *U*0, *Ua*, *Ub*, *Uc* and middle inversion parameter *MINV* are extracted first. The required device data-sets at this stage are the long channel device I_D-V_G characteristics at low drain bias condition.

At the second stage of extraction, short channel effect parameters such as *DVT*0, *DVT*1, *DVT*2, *LPE*0, *LPEB* and *CDSC*; bias dependent source/drain resistance parameters *PRWG* and *PRWB* and Channel length dependent mobility parameters *UP*, *LP* will be extracted. Some parameters such as *K*1, *K*2, *VOFF*, *Ua*, *Ub*, *Uc*, *RDSW* and *MINV* which were extracted at long channel length step will be re-extracted here as well to obtain better fitting for short channel devices. The required device data-sets are the long and short channel devices I_D-V_G characteristics at low drain bias condition.

At the third stage of extraction, high drain bias short channel effect parameters such as DIBL related parameters *DSUB*, *ETA*0, *ETAB*; drain induced threshold shift parameters *DVTP*0, *DVTP*1; bulk charge effect parameters *A*0, *AGS*, *KETA*; velocity saturation parameter *VSAT*; saturation output conductance parameters such as *PDIBLC*1, *PDIBLC*2, *PCLM*, *PVAG* and non-saturation effect parameters *A*1, *A*2 will be extracted. Some parameters at previous steps such as *VTH*0, *RDSWMIN*, *VOFF*, *NFACTOR*, *RDSW* and *MINV* will be re-extracted here. The required device data-sets are the long and short channel devices I_D-V_G and I_D-V_D characteristics.

The results of this parameter extraction strategy for 30nm, 35nm and 40nm devices are illustrated in Figures 3.7, 3.8 and 3.9, respectively. The root mean square (RMS) errors of the final BSIM4 compact model with respect to device TCAD simulations are presented in Table 3.2 where RMS error (in percent) is defined by:

$$E_{RMS} = 100 \times \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{I_{TCAD,i} - I_{BSIM,i}}{I_{TCAD,i}}\right)^2} \tag{3.63}$$

$I_{TCAD,i}$ is the data of drain current obtained from TCAD, $I_{BSIM,i}$ is the corresponding point simulated using the extracted compact model and N is total number of data points.

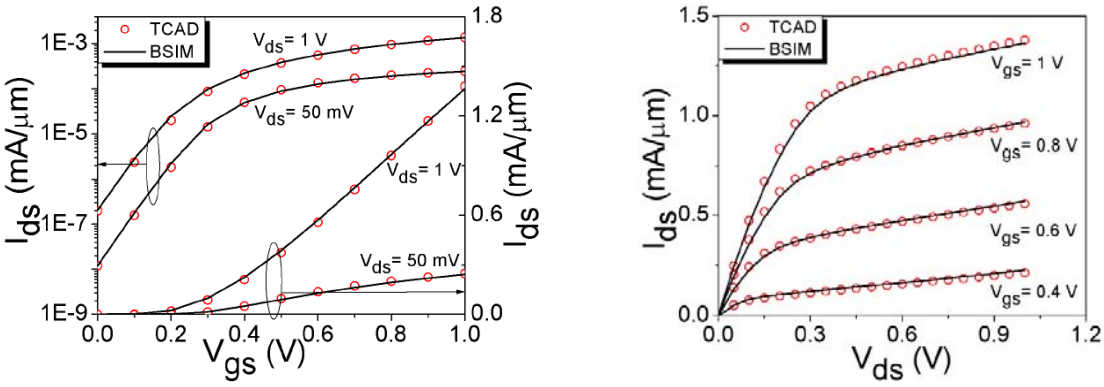


Figure 3.7: Device I_D-V_G and I_D-V_D characteristic comparison between TCAD simulation and BSIM4 compact model extraction results for 30nm n-MOSFET.

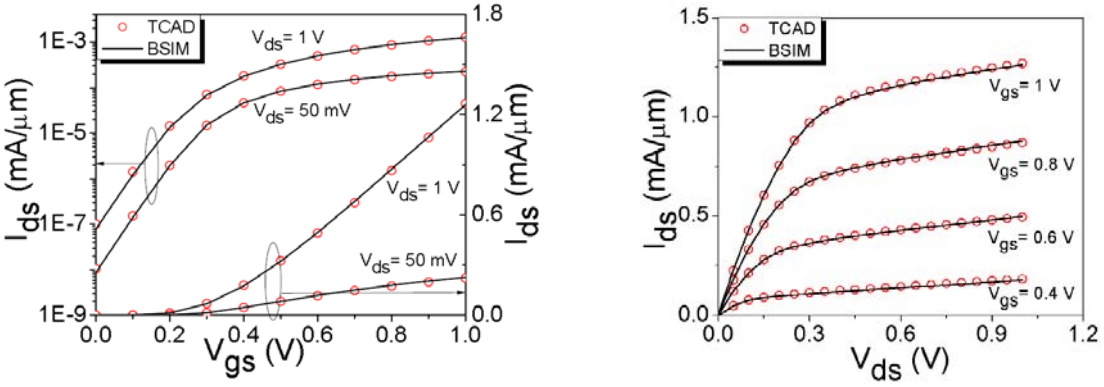


Figure 3.8: Device I_D-V_G and I_D-V_D characteristic comparison between TCAD simulation and BSIM4 compact model extraction results for 35nm n-MOSFET which was shown in Figure 3.6.

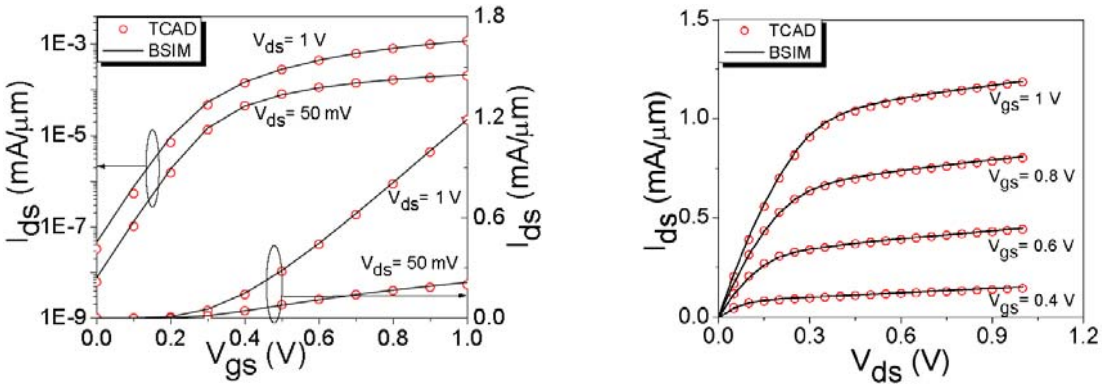


Figure 3.9: Device I_D-V_G and I_D-V_D characteristic comparison between TCAD simulation and BSIM4 compact model extraction results for 40nm n-MOSFET.

Table 3.2: RMS errors of the final BSIM4 compact model in respect to device TCAD simulations

Gate Length / RMS error (%)	I_D-V_G @ V_D=0.05V	I_D-V_G @ V_D=1V	I_D-V_D
30nm	0.77	0.90	3.52
35nm	0.41	0.66	2.18
40nm	0.77	1.27	2.26

The identical parameter extraction strategy is employed for 18nm gate length transistor BSIM4 parameter extraction as well. MOSFETs with 14nm, 22nm and 90nm gate length are used to extract parameters of 18nm device. Figures 3.10 to 3.12 illustrate the results of parameter extraction for 14nm, 18nm and 22nm gate length n-channel MOSFET. The accuracies of BSIM4 parameter extraction in reference with TCAD are 1.66% for I_D-V_G and 1.16% for I_D-V_D characteristics for 18nm device. The RMS errors of the final BSIM4 compact model in respect to device TCAD simulations are shown in Table 3.3.

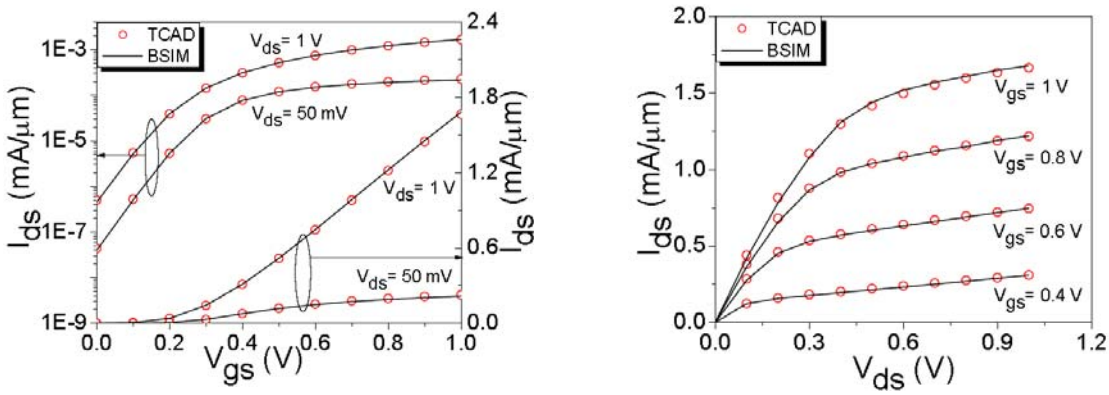


Figure 3.10: Device I_D-V_G and I_D-V_D characteristic comparison between TCAD simulation and BSIM4 compact model extraction results for 14nm n-MOSFET.

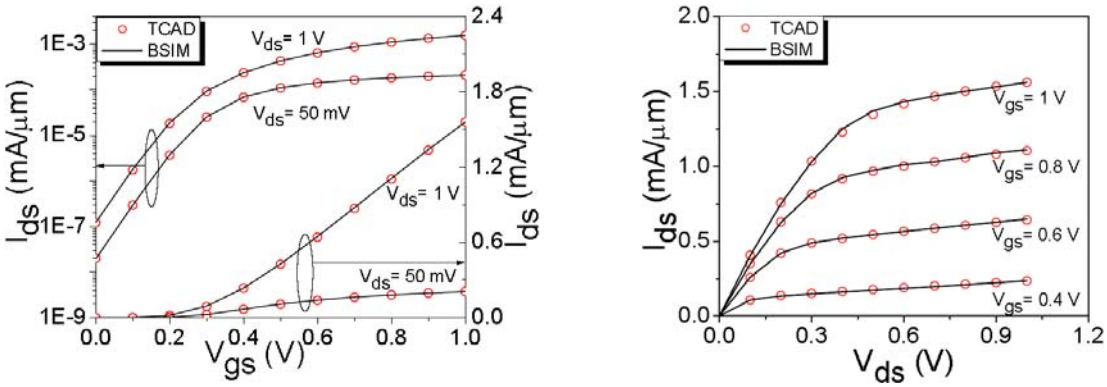


Figure 3.11: Device I_D-V_G and I_D-V_D characteristic comparison between TCAD simulation and BSIM4 compact model extraction results for 18nm n-MOSFET.

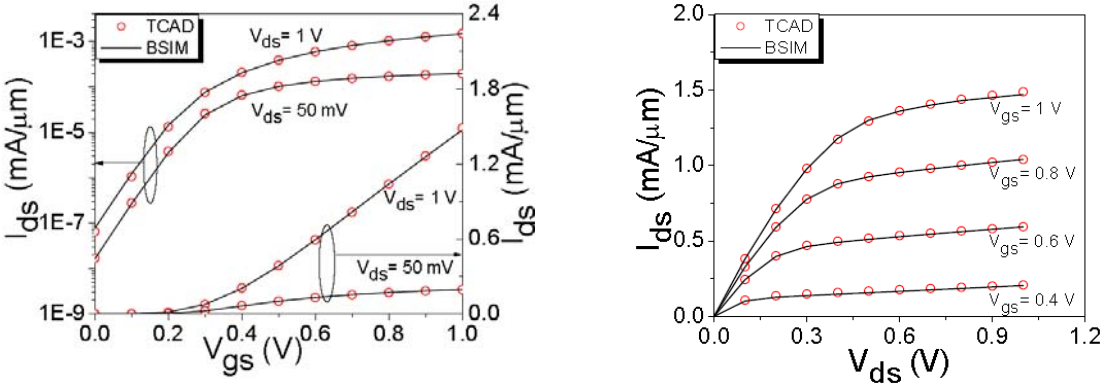


Figure 3.12: Device I_D-V_G and I_D-V_D characteristic comparison between TCAD simulation and BSIM4 compact model extraction results for 22nm n-MOSFET.

Table 3.3: RMS errors of the final BSIM4 compact model in respect to device TCAD simulations

Gate Length / RMS error (%)	I_D-V_G @ V_D=0.05V	I_D-V_G @ V_D=1V	I_D-V_D
14nm	4.29	3.86	2.16
18nm	1.34	1.93	1.16
22nm	3.62	7.45	2.99

3.4 Summary

In this chapter, an overview of the structure and equations of BSIM4 compact model was provided, this includes gate dielectric model, effective channel length and width model, threshold voltage model, channel charge model, mobility, source/drain resistance and drain current model. Then a 35nm poly-gate MOSFET was introduced and used as a test-bed device for the statistical compact model extraction and generation strategies represented in this PhD project. A group device extraction strategy was described based on I_D-V_G and I_D-V_D TCAD simulation results. This group includes data set of 30nm, 35nm, 40nm and 200nm MOSFETs. Similarly nominal compact model was extracted for an 18nm channel length template transistor. The benefit of this extraction approach is that it gives more flexibility to circuit designers to choose devices with different gate length. The worst RMS error of proposed parameter extraction strategy on the group of test-bed devices is less than 3.60% for the 35nm MOSFET and 7.45% for the 18nm MOSFET. The extracted nominal compact models are the basis for the statistical compact model extraction and generation techniques in this work.

Chapter 4

Statistical Compact Model Extraction

Since compact models act as a bridge between IC designer and foundry, statistical compact modelling is the only way to effectively communicate device statistical variability information to designers [118]. The investigation and development of flexible and accurate, yet economical strategies for capturing statistical variability in industrial standard compact models is of great importance for variability aware design. Integrating the effect of these fluctuations into compact models can result in the prediction of the statistical circuit in DC, AC and transient regimes. The overall accuracy of this circuit/system simulation is determined by the accuracy of the statistical compact models need in the simulations.

In order to obtain compact models which can accurately predict the statistics of real circuit operation, without any presumption of parameter distribution, correlation and sensitivity, a two-stage direct statistical compact model extraction procedure [119] is employed. In the first stage, as described in Chapter 3, a combination of group extraction and local optimization strategy has been applied to obtain the complete set of BSIM parameters for a *uniform* device. The resulting compact model card serves as the base model card for the second stage: the *statistical* extraction. In this second step a small number of key BSIM parameters are chosen, and this small number of parameters re-

extracted for each member of a large ensemble of device *IV* curves – obtained either from measurement or physics based 3D TCAD device simulation.

As a result, within the accuracy of the compact model fitting, this approach will be the most accurate representation of the current voltage characteristics of each device, and the statistics of the key parameters will capture the statistics of device operation. It has been found [119] that just the gate characteristics at low and high drain bias are required to perform these statistical extractions, since most of variability information can be captured by gate characteristics.

This chapter starts with a description of the statistical variability simulation of 35nm gate length n-MOSFETs and p-MOSFETs which are the test bed devices of this thesis. The next section discusses how the choice of key parameters for BSIM compact model extraction is made including the use of parameter sensitivity analyses. Then the behaviour of BSIM compact models with these sets of parameters will be investigated, including comparison for devices *in situ* in circuits. In the last section we will focus on BSIM compact model fitting and the accuracy of statistical parameter extraction. This chapter is an introduction for chapter 5, which investigates correlation between parameters and different statistical compact model generation approaches.

4.1 Simulation of Statistical Variability

Intrinsic parameter fluctuations associated with the discreteness of charge and granularity of matter are now one of the major factors limiting scaling, integration and the reduction of supply voltages and power consumption in ULSI applications [7]. The accurate modelling and simulation of such effects is very important for the development of present and future generations of semiconductor devices and the integration of giga-transistor count chips. In order to achieve reasonable performance and yield in contemporary CMOS design, statistical variability has to be accurately represented by industry standard compact models [77]. As explained in chapter 2, the major sources of statistical variability are: Random Discrete Dopants (RDD); Line Edge Roughness (LER);

Poly Gate Granularity (PGG); High-k Granularity and Interface Roughness and Oxide Thickness Variations (OTV). These fluctuations can be simulated by using Drift Diffusion (DD), Monte Carlo (MC) and Quantum Transport (QT) techniques [43]. All sources of variability are three dimensional in nature and therefore in order to correctly capture their effects 3D simulation should be carried out [46].

Drift Diffusion (DD) represents the simplest model used in multi-dimensional numerical simulations and captures the lowest order moments of the BTE [73]. In this model the electron and hole current densities are approximated using two components [120]: a drift component driven by the electric field and a diffusion component driven by the carrier density gradient. The current density is sum of these components. A major benefit of the drift diffusion approach is low computational cost. It is therefore suitable for carrying out the large sets of statistical simulations needed to characterize the impact of various sources of variability.

Although full quantum mechanical simulations are prohibitive in terms of computational time it is possible to include some quantum mechanical effects into otherwise classical simulations using the so called quantum corrections [46]. Quantum corrections allow the quantum confinement effects to be approximated in the simulation. These quantum effects play an increasingly important role as devices are aggressively scaled into the nanometer regime. Using quantum corrections significantly improves the accuracy of drift diffusion simulation for nano-scale MOSFETs where quantum confinement effects influence device threshold voltage. Their use also becomes essential in resolving the influence of individual discrete dopants [121].

In this work, instead of experiment data, ensembles of 1000 statistical device simulation results, obtained using the predictive Glasgow ‘atomistic’ 3D drift-diffusion simulator, are used [42]. The simulator simultaneously employs density gradient quantum corrections for both electrons and holes to resolve the impact of individual impurities [122]. The test bed devices are 35nm gate length n-MOSFETs and p-MOSFETs designed to match the performance of state-of-the art 45nm technology devices [113,114]. The combined impacts

of RDD, LER and PGG in n-MOSFETs and RDD and LER in p-MOSFETs have been simulated, and devices with $W/L=1$ are used. RDD are generated based on a normal continuous doping profile by placing dopant atoms on silicon lattice sites with probability determined by the local ratio between dopant and silicon atom concentration [123]. LER is introduced through 1D Fourier synthesis with a power spectrum corresponding to a Gaussian auto correction function with correlation length of 30nm and RMS amplitude of 1.3nm [56]. PGG is introduced by importing a random section of a large template polycrystalline silicon grain image for the whole gate region with the average grain size of 65nm obtained through X-ray-diffraction measurements. Due to the presence of acceptor type interface states along the grain boundaries, the Fermi level remains pinned at the silicon band-gap at these boundaries [124], in this case 200mV above the middle of the band gap. Since Fermi level pinning only occurs in NMOS devices, PGG in PMOS devices is not simulated [124]. An example of the impact of variability sources on the potential distribution in a 35nm gate length bulk n-MOSFET is illustrated in Figure 4.1.

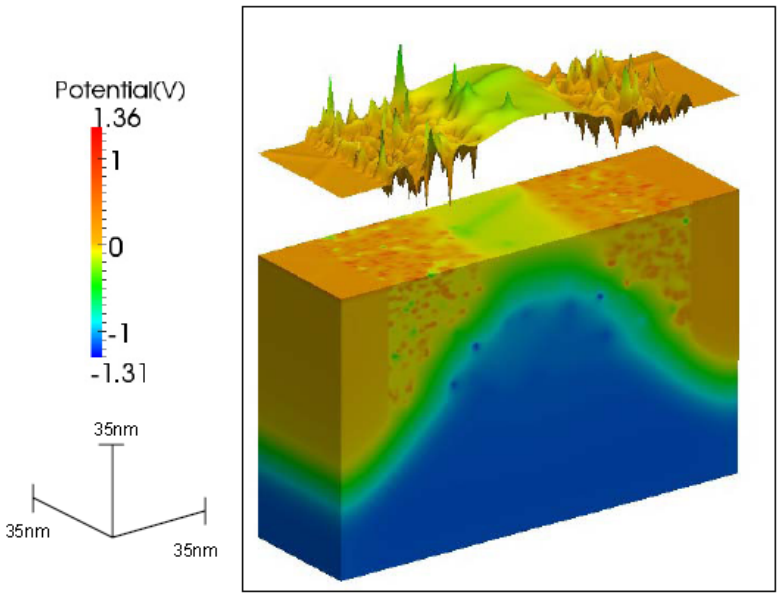


Figure 4.1: Potential distribution in a 35nm n-MOSFET subject to sources of variability. The drain bias is 1V.

Figures 4.2 and 4.3 illustrate the spread in I_D-V_G characteristics obtained from

'atomistic' simulation due to the combined effects of variability for a 35nm gate length n- and p-MOSFET device, respectively.

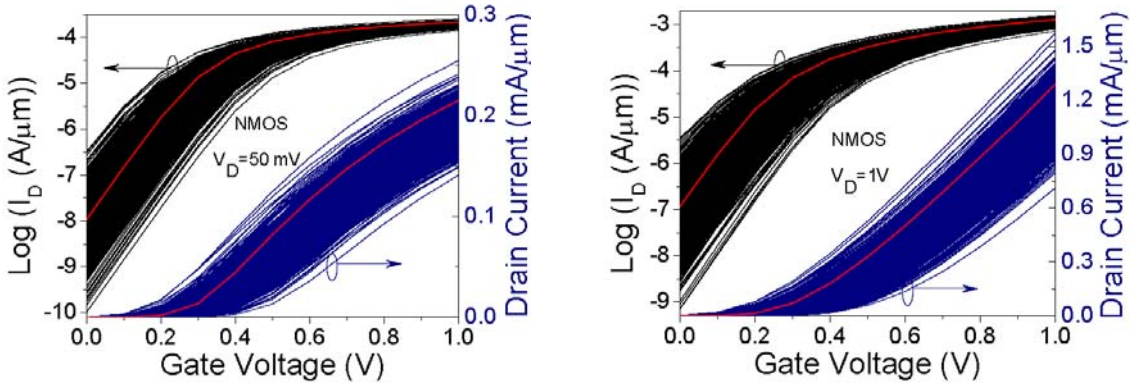


Figure 4.2: Variability in the current voltage characteristics of a statistical sample of 1000 microscopically different 35nm gate length square ($W = L$) n-NMOSFETs at $V_D = 50mV$ and $V_D = 1V$. Red line shows characteristic of uniform device.

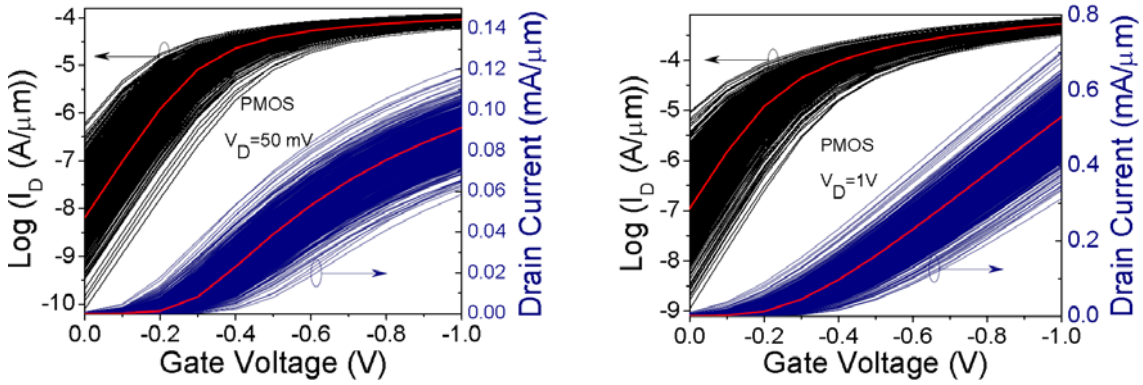


Figure 4.3: Variability in the current voltage characteristics of a statistical sample of 1000 microscopically different 35nm gate length square ($W = L$) p-MOSFETs at $V_D = 50mV$ and $V_D = 1V$. Red line shows characteristic of uniform device.

It is clear from Figures 4.2 and 4.3 that the magnitude of variation introduced by statistical variability sources for devices in the deca-nanometer gate length regime is considerable. The off-current variation spans approximately 3 orders of magnitude, indicating that statistical variability has a strong impact on the electrostatically dominated sub-threshold behaviour of the devices. It is well known that drift-diffusion simulations

can underestimate drive current variability [125]. However the observed deviation between maximum and minimum drive current, even in these drift-diffusion simulations, can still be 45% of its mean.

4.2 Statistical Set of BSIM Parameters

Compact models are key components of the interface between technology and design. Although the initial drive for compact model development was to obtain accuracy in modelling circuit components in analogue IC design, compact models are also extensively used in transistor-level digital circuit design and verification, especially in the methodologies for standard cell characterization [15].

As the BSIM compact model has approximately 400 independent parameters, a reduction in the number of parameters involved in the statistical phase of the compact model extraction is necessary for the development of any scalable compact modelling strategy. A compact model parameter sensitivity analysis was carried out to provide a solid ground for the selection of a subset of the BSIM parameters to serve as the statistical compact model parameter set. Although the actual statistical compact parameter extraction strategy will need extraction steps that involve multiple-variables, nevertheless, an individual parameter sensitivity analysis can provide vital first-order information regarding each individual parameter's capability to capture the impact of statistical variability on device characteristics.

Ten parameters were identified as possible candidates for the statistical parameters set, based on our understanding of underlying device physics. These are: *VTH*0, *VOFF*, *NFACTOR*, *U*0, *RDSW*, *DSUB*, *MINV*, *VSAT*, *PVAG* and *LINT* all identified in chapter 3. *VTH*0 is selected to account for traditional threshold variation introduced by statistical variability; *NFACTOR* and *VOFF* are selected to account for sub-threshold slope and leakage current variation; *U*0 is selected to account for current factor variation caused by statistical variability; RDSW is selected to account for dopant variation in the source/drain regions; *DSUB* is selected to account for DIBL variation introduced by statistical

variability; *MINV* is selected to account for variation in the moderate inversion regime; *VSAT* is selected to account for the velocity saturation phenomena which occur in short channel transistors; *PVAG* is selected to investigate channel length modulation effects on the drain current and *LINT* is selected to model effective channel length variation caused by LER. In order to evaluate the impact of each parameter on the drain current, a first order sensitivity analysis is carried out. Parameter sensitivity strength is defined in a normalized fashion as:

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

where I is the drain current, P refers to a typical BSIM4 parameter and ΔP is the increment of a parameter from its nominal value in uniform model. A relative increment of $\Delta P/P$=0.01 is chosen for each parameter and the corresponding increment in the drain current, ΔI, is measured using HSPICE simulation.

In digital applications, transfer current-voltage characteristics from both high and low drain bias conditions provide enough device characteristic variation information for accurate simulation of device switch behaviour under the influence of statistical variability. As a result, only device bias conditions associated with the transfer characteristics were considered in the parameter sensitivity strength analysis. Figure 4.4 presents the parameter sensitivity strength as a function of gate bias under low drain voltage conditions ($V_D = 50mV$), with the corresponding I_D-V_G curve also presented as a reference.

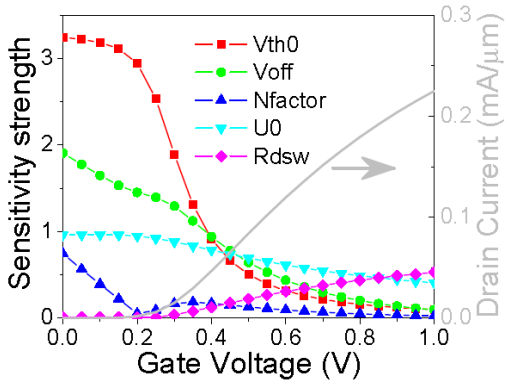


Figure 4.4: The sensitivity strength of BSIM parameters as a function of gate voltage, low drain bias ($V_D = 50mV$).

Only parameters that exhibit meaningful sensitivity (i.e, with sensitivities more than 0.1) are presented in Figure 4.4, which clearly demonstrates that in the sub-threshold regime, *VTH*0, *VOFF*, *U*0 and *NFACTOR* are the most sensitive parameters. However, there is subtle difference between the impact of *VTH*0, *U*0 and that of *VOFF*, *NFACTOR* in the sub-threshold regime. For *VTH*0 and *U*0, the values of sensitivity strength are almost constant in sub-threshold, which means that sub-threshold slope variation cannot be effectively captured by *VTH*0 and *U*0. While for *VOFF* and *NFACTOR*, the change of sensitivity strength against gate voltage in sub-threshold provides a means (in combination with *VTH*0) for BSIM models to capture both first and second order effects of variation in the sub-threshold regime. The drain current in sub-threshold follows an exponential relationship against gate bias, whilst the drain current in inversion follows a sub linear relationship. However, the sensitivity strength is calculated on a linear scale. As expected, *RDSW* starts to play an important role with increasing gate voltage.

Figure 4.5 presents the parameter sensitivity strength as a function of gate bias at high drain voltage ($V_D = 1V$), with the corresponding I_D-V_G curve presented as a reference.

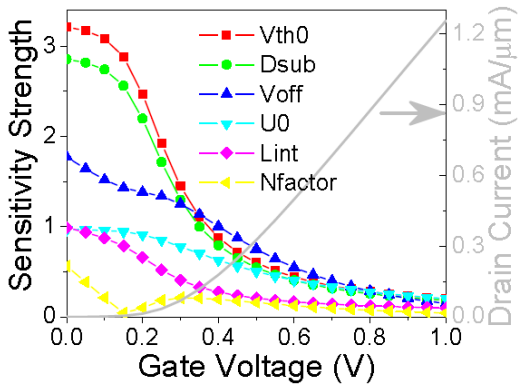


Figure 4.5: The sensitivity strength of BSIM parameters as a function of gate voltage under high drain bias ($V_D = 1V$).

Apart from *VTH*0, *VOFF*, *U*0 and *NFACTOR* noted above, *DSUB* and *LINT* also come to play a role in sub-threshold regime at high drain bias. *DSUB* especially becomes one of the dominant parameters.

Based on this sensitivity strength analysis, the parameters *VTH*0, *VOFF*, *NFACTOR*, *U*0, *RDSW* and *DSUB* have been identified as the most important six parameters to be used in BSIM to capture the device variability. Starting from *VTH*0 as the most sensitive parameter in both low drain and high drain bias conditions, we will add other parameters incrementally to the statistical parameter set based on their significance in RMS error reduction. This is due to the fact that first, the order of sensitivity of parameters is different at low and high drain bias conditions as considered in Figures 4.4 and 4.5 and second, the sensitivity of parameters has been measured individually. In other words, the correlations between parameters are not considered in the sensitivity analysis. Measuring the RMS error incrementally after fitting each parameter will help to decide which parameter should be chosen as second, third, fourth, fifth or sixth important parameter in the final statistical set.

4.2.1 Impact of parameter Set on the Accuracy of Individual Device

Having chosen the key subset of BSIM parameters, the typical approach is to perform parameter extraction on each member of a large ensemble of devices, keeping all the BSIM parameters constant aside from our key chosen subset, to obtain a large ensemble of key BSIM parameter subset which directly captures the variability of the device ensemble. We refer to this as the *exhaustive* or *direct* approach. The advantage of direct approach is twofold: Firstly, it does not require that the variation of device electrical performance parameter follows any particular distribution. Secondly, it does not presume any statistical compact model parameter distribution or correlation which naturally arises in the process of performing multiple parameter extractions. As a result, within the accuracy of the compact model fitting, this approach will be the most accurate representation of the current voltage characteristics from the physical 3D simulations or from measurement. The accuracy in representing each one of statistical I_D-V_G characteristics from the device ensemble depends on the choice of key BSIM parameters and on the number of parameters used. Figure 4.6 shows the reduction of mean and standard deviation of errors with the increase in the number of key statistical parameters for 1000 n-MOSFET and p-MOSFET

devices. These statistical parameter sets are selected according to parameters significance on the error reduction. The accuracy criterion is RMS error (in percent) which is defined by:

$$E_{RMS} = 100 \times \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{I_i - I_{i,SIM}}{I_i}\right)^2} \tag{4.2}$$

where N is the number of data points from both high drain bias ($V_D = 1V$) and low drain bias ($V_D = 50mV$) I_D-V_G characteristics, I_i is the data point from physical simulation and $I_{i,SIM}$ is the corresponding data point which is simulated using extracted compact model.

Figure 4.6 shows the trend in RMS error of the extracted models using different numbers of key parameters. As expected, the mean and standard deviation of the RMS error decreases as the number of key parameters is increased. For n-MOSFETs the mean is reduced from 24.2% with a 1-parameter set, to 2.0% for a 6-parameter set and the standard deviation is reduced from 6.8% with a 1-parameter set, to 0.88% for a 6-parameter set. For p-MOSFETs the mean is reduced from 15.8% with a 1-parameter set, to 2.8% for a 6-parameter set and the standard deviation is reduced from 5.3% with a 1-parameter set, to 0.77% for a 6-parameter set.

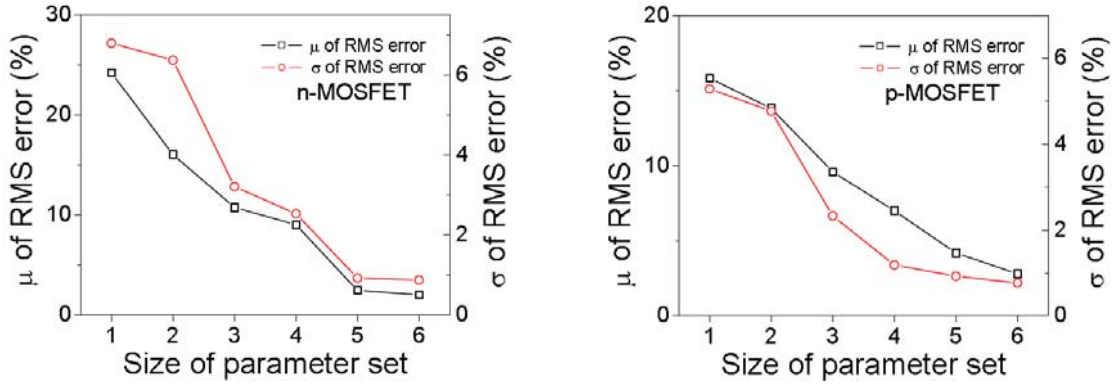


Figure 4.6: The impact of statistical parameter set size on RMS error of a statistical compact model. (Left) n-MOSFET and (Right) p-MOSFET. The parameters are selected based on the order presented in Table 4.1.

As an attempt to increase compact model accuracy further, the 6-parameter set was augmented by including the parameter *MINV*. However, from sensitivity strength analysis, the physical impact of this parameter on device characteristics is weak and improvements in fitting accuracy actually come from using artificially large values of *MINV*, an undesirable technique in analytic compact modelling. By using such large values of *MINV* the mean of error of extraction can be reduced to 1.5% for n-MOSFETs and 2.36% for p-MOSFETs. In Figure 4.7 characteristic comparison between TCAD simulation and a typical set of BSIM4 compact model extraction results for 35nm n-MOSFETs are illustrated. When using the 7-parameter set over the 6-parameter set, the RMS extraction error decreases from 3.0% to 1.9% for this particular n-MOSFET device. By using the 7-parameter set for 1000 devices, the average RMS error will be decreased from 2.0% to 1.5% for n-MOSFETs and from 2.8% to 2.4% for p-MOSFETs.

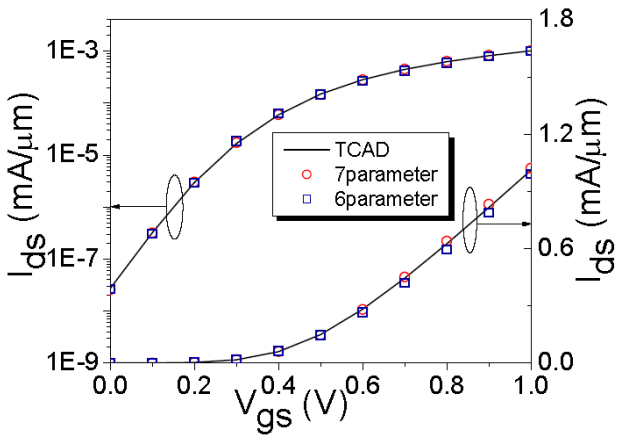


Figure 4.7: Device I_D-V_G characteristic (high drain condition) comparison between TCAD simulation and BSIM4 compact model extraction results for 35nm n-MOSFET.

The impacts of the different choices of the statistical parameter set on the statistical compact models accuracy and histogram of RMS error are shown in Figure 4.8. Table 4.1 represent the mean and standard deviation of RMS error for different number of parameters in each statistical set.

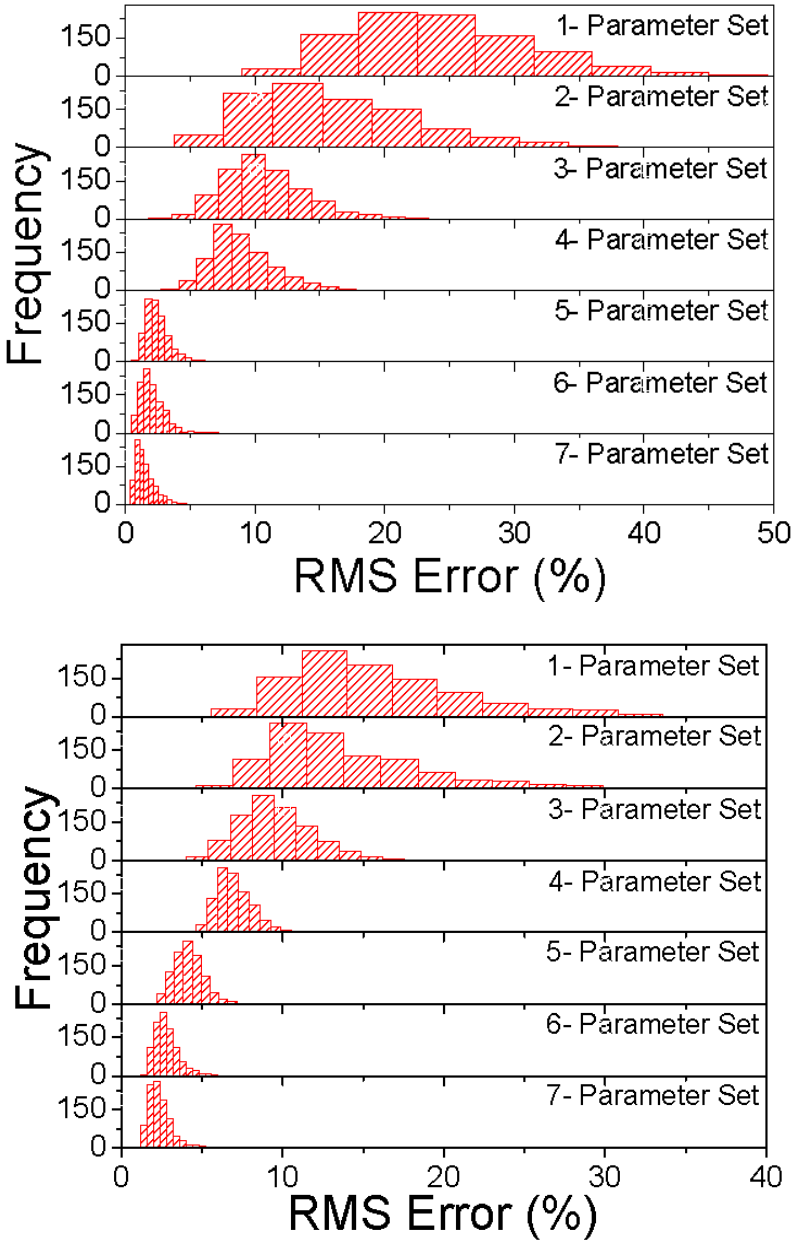


Figure 4.8: The impact of statistical parameter set size on relative RMS error of BSIM statistical compact model. (Up) n-MOSFET and (Down) p-MOSFET

Table 4.1: Statistical parameter extraction RMS errors for 1000 BSIM compact models

Parameters Error	n-MOSFET		p-MOSFET	
	μ	σ	μ	σ
*VTH*0	24.21	6.79	15.84	5.29
*VTH*0, *U*0	16.05	6.37	13.81	4.77
*VTH*0, *U*0, *DSUB*	10.75	3.21	9.60	2.33
*VTH*0, *U*0, *DSUB*, *VOFF*	9.04	2.53	7.00	1.19
*VTH*0, *U*0, *DSUB*, *VOFF*, *NFACTOR*	2.47	0.93	4.20	0.93
*VTH*0, *U*0, *DSUB*, *VOFF*, *NFACTOR*, *RDSW*	2.03	0.88	2.81	0.77
*VTH*0, *U*0, *DSUB*, *VOFF*, *NFACTOR*, *RDSW*, *MINV*	1.50	0.77	2.36	0.68

Figure 4.9 shows the correlation between a typical parameter *VTH*0 at different collection of statistical parameters. It indicates that the physical meaning of *VTH*0, as the most important parameter in the BSIM compact model, is preserved during each stage of parameter extraction. The increased variance of the parameter between the 1- and 2-parameter sets is due to the fact that in the 1-parameter set *VTH*0 alone must account for all points of the I_D-V_G spread.

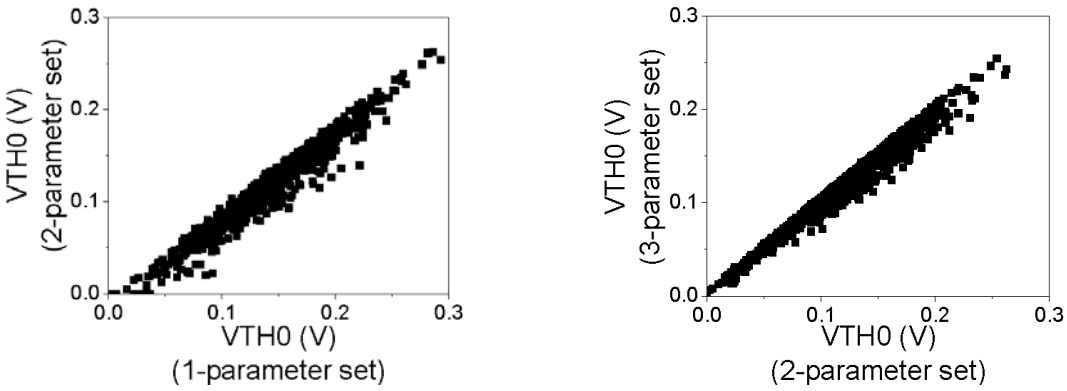


Figure 4.9: The correlation between a typical statistical parameter at different parameter sets.

4.2.2 Impact of Parameter Set on the Statistical Property of Device Figures of Merit

So far we have carried out a quantitative analysis of the impact of various numbers of key parameters on the accuracy of BSIM statistical set in respect to physical simulation data. However, the real devices are often judged by figures of merit which are often found at the external point of the devices operation, and it is important to examine the accuracy of our extraction technique in respect to match those figures of merit between BSIM statistical compact models and physical devices.

The key device figures of merit (FOM) for MOSFETs are: threshold voltage (V_{th}); drive current (I_{on}); source/drain leakage current (I_{off}); Drain Induced Barrier Lowering (DIBL) and sub-threshold slope (SS). These parameters are defined as follows:

1. The threshold voltage is the gate voltage when the device starts to turn on [101]. The accurate modelling of threshold voltage is important for accurate circuit simulation. Since V_{th} has profound effect on circuit operation, it is often used to monitor process variations. The threshold voltage is extracted for each of I_D-V_G atomistic simulations with constant threshold current criteria of $I_D = 100nA \times W/L$ where W and L are the width and length of devices, respectively, and the threshold voltage is extracted under conditions of both high drain ($V_D = 1V$) and low drain ($V_D = 50mV$) bias.

2. The drive current (I_{on}) is the drain current of a transistor when the device is turned fully on. For n-MOSFETs it is the current when the gate and drain are connected to the supply voltage and the source and bulk terminals are grounded. For p-MOSFETs the gate and drain are connected to negative supply and the source is grounded.

3. The channel leakage current (I_{off}) is the drain current when the drain terminal is connected to the supply voltage and the device is fully turned off. For n-MOSFETs

this requires all other terminals to be grounded. For p-MOSFETs gate and source are grounded and the drain is connected to negative supply voltage.

The definition of I_{on} and I_{off} for n-MOSFETs is shown in Figure 4.10.

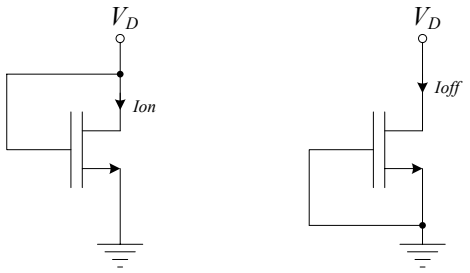


Figure 4.10: A simple circuit for description of the I_{on} and I_{off}.

4. Another important device parameter which is extracted from atomistic simulations is the DIBL parameter. This short-channel effect has been attributed to the penetration of the drain junction electric fields into the channel region, causing barrier lowering, which in turn leads to V_{th} reduction [101,126]. The DIBL is measured as the threshold voltage reduction due to drain bias increase divided by the corresponding increment in drain voltage.

5. The reciprocal of the slope of the *Log* (I_D) versus V_G is the sub-threshold slope (*SS*). It is an important device parameter which determines how well the MOSFET functions as a switch [127,112]. It is the change in the gate voltage required to change the drain current by one order of magnitude in the sub-threshold region.

The correlation between key device figures of merit under high drain bias ($V_D = 1V$) is shown in Figure 4.11. As expected, there is a strong correlation between threshold voltage and leakage current of devices. However, the correlation between threshold voltage and drive current is not perfect, for a fixed threshold voltage value, the drive current can have more than a 10% spread.

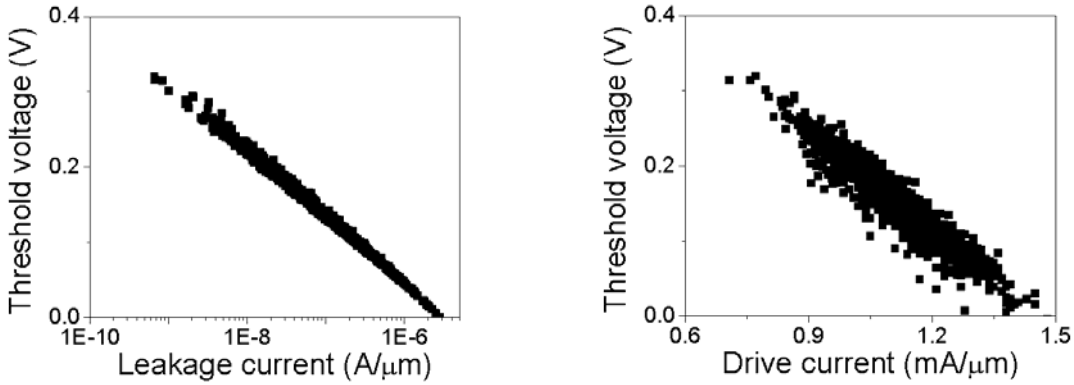


Figure 4.11: the correlation between typical electrical figures of merit for a statistical sample of 1000 microscopically different, 35nm gate length, and square ($W = L$) n-MOSFETs at $V_D = 1V$.

The impact of variability on the range of the figures of merit discussed above, under different drain voltage conditions, is summarized in Table 4.2. The devices under investigation are those of Figure 4.2.

Table 4.2: Summary of simulation results for 1000 microscopically different, 35nm gate length, square ($W = L$) n-MOSFETs including all statistical variability sources.

FOM V_D	High Drain Voltage (V_D = *1*V)			Low Drain Voltage (V_D = *50* mV)		
	Standard Deviation (σ)	Mean (μ)	$\frac{\sigma}{\mu}$%	Standard Deviation (σ)	Mean (μ)	$\frac{\sigma}{\mu}$%
V_{th} (V)	0.06	0.15	40	0.05	0.25	20
I_{on} (mA/μm)	0.12	1.13	10.62	0.02	0.19	10.53
$LogI_{off}$(A/μm)	0.62	-7.17	8.65	0.56	-8.17	6.85
DIBL (V/V)	0.02	0.10	30	0.02	0.10	20
Slope (mV/dec)	5.92	91.97	6.44	2.48	89.43	2.77

Clearly V_{th} and DIBL have largest spread when considering normalized standard deviations (σ/μ). This indicates that statistical variability has a strong impact on the

device electrostatic-dominated sub-threshold behaviour, introducing noticeable modulation of the short channel effects. Since the sub-threshold region in I_D-V_G curves follows log normal distribution, we used the logarithm of I_{off} instead of I_{off}.

In order to evaluate the impact of key parameter selection on device figures of merit, different size of key parameter set are used to obtain BSIM compact models, from which the mean and standard deviation of I_{on}, I_{off} and V_{th}, are calculated. These are then compared with the same results obtained directly from the original physical TCAD simulations.

Figure 4.12 and 4.13 show the mean and standard deviation of the leakage current, drive current and threshold voltage for different parameter sets at high drain bias ($V_D = 1V$) for n-MOSFETs and p-MOSFETs, respectively. The mean and standard deviation of the most accurate (5-, 6- and 7-parameter) sets have reasonable errors compared to results taken directly from physical simulation (shown as a reference in these figures by the dashed horizontal lines). Figure 4.12 shows that selection of a 5-parameter set for the n-MOSFETs will be enough to settle the error trends of I_{on} and using more parameters will not further reduce the mean and standard deviation of error. In addition, a 5-parameter set will be enough to settle the error trends of I_{off} and gives around negligible (0.1%) error in its mean and standard deviation. Selection of this 5-parameter set result in an 0.6% error for the mean and 0.2% error in the standard deviation of V_{th}. As for I_{on} and I_{off}, increasing further the number of parameters does not affect on the accuracy of the compact model.

Figure 4.13 shows that selection of a 6-parameter set for the p-MOSFETs will be enough to settle the error trends of I_{on} and using more parameters will not further reduce the mean and standard deviation of error. The 5-parameter set will be enough to settle the error trends of I_{off} and gives an 0.1% error in the mean and negligible (0.05%) error in the standard deviation of I_{off}. Selection of this 5-parameter set results in a 1.6% error for the mean and 1.4% error for standard deviation of V_{th}. As for I_{off}, increasing further the number of parameters does not affect on the accuracy of compact model.

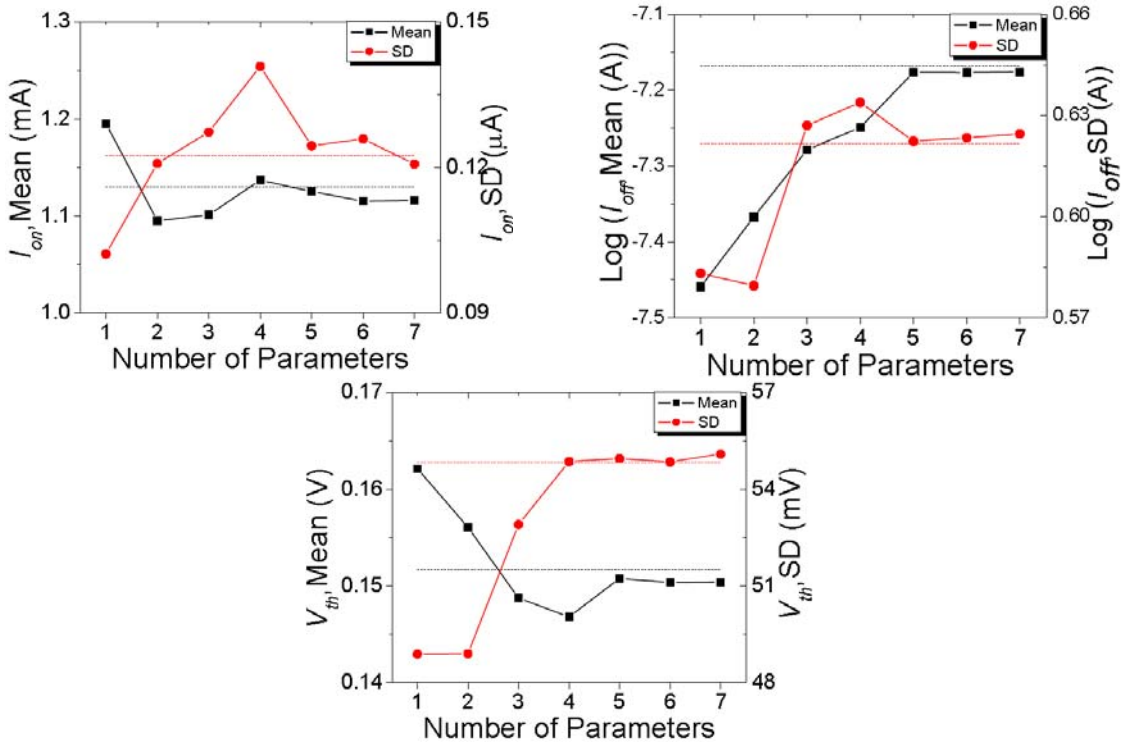


Figure 4.12: Impact of parameter set selection in n-MOSFET device; dashed line shows the results of the original physical atomistic simulations.

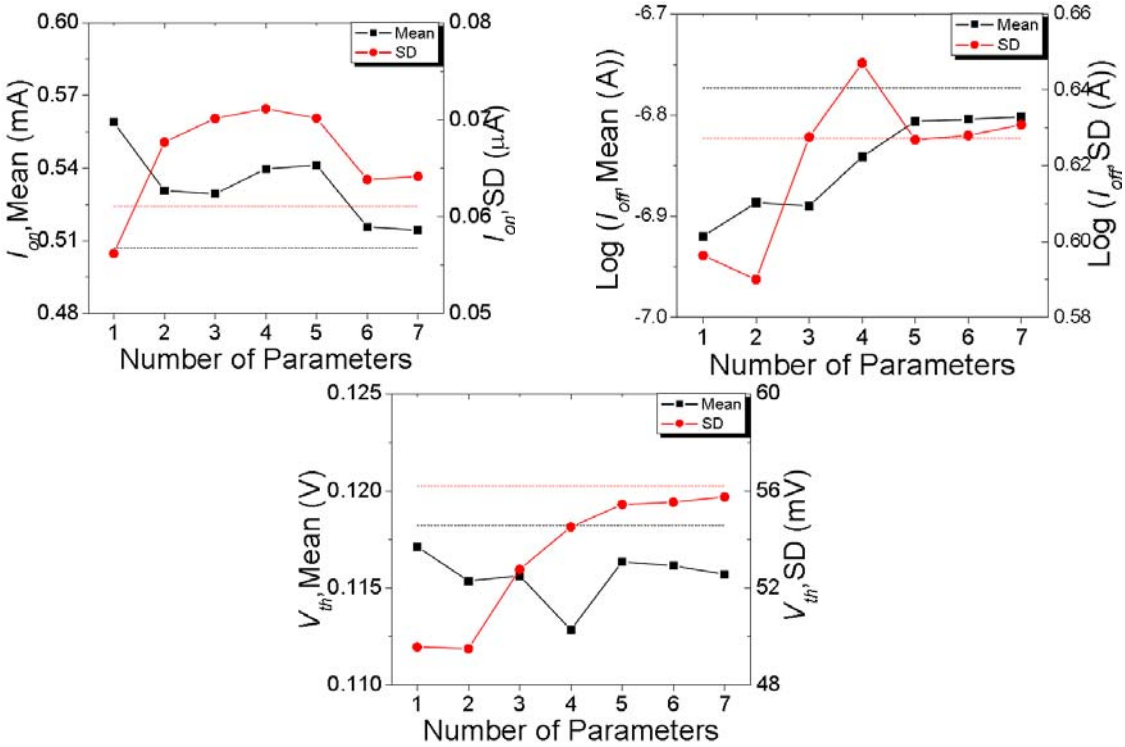


Figure 4.13: Impact of parameter set selection in p-MOSFET device; dashed line shows the results of the original physical simulations.

Since the direct extraction targets are the full set of gate characteristics, the overall monotonic reduction of the total RMS error with increasing parameter-set size, demonstrated in Figures 4.12 and 4.13, does not guarantee a monotonic error reduction for a particular device figure of merit, although the overall tendency is one of reducing errors.

Figure 4.14 shows a comparison of drive current distribution using 7-, 6- and 5-parameter statistical sets. As illustrated in Figure 4.14, using 5 or 6 parameters in statistical compact model results in little deviation in the tail of normal distribution compared with the distribution obtained by use of 7-parameter set.

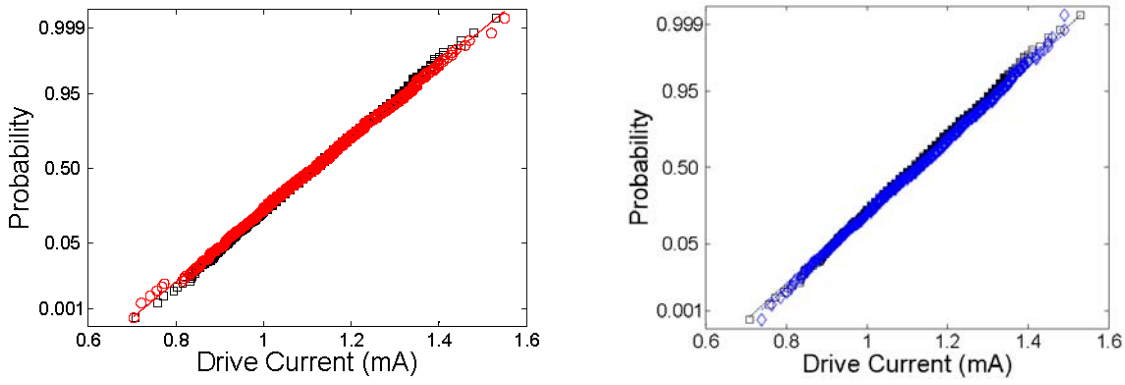


Figure 4.14: Probability plot of drive current in high drain bias. Black square: 7-parameter set; Red circle: 6-parameter set and Blue lozenge: 5-parameter set.

4.2.3 Impact of Parameter Set on Statistical Circuit Simulation

Having considered the ability of various numbers of parameters to capture device *I-V* characteristics and figures of merit, we have developed an appropriate statistical compact model which can be used to predict the statistical behaviour of circuits in presence of variability. Therefore, in order to investigate the accuracy of different parameter set selections in real circuit simulations, a CMOS inverter has been considered, constructed from both n-MOSFET and p-MOSFET devices, using 35nm gate length devices. The n-MOSFET width is eight times of length, while the p-MOSFET has a width of 2.3 times of n-MOSFET to properly balance the drive currents. A 2GHz input signal with 50ps rise and fall time is considered as V_{in}. With this selection of rise/fall time and under heavy

capacitive loads (i.e., fan out of 10), the output of the inverter can still be settled within half period. Simulations are carried out for two cases: first, CMOS inverter without a load which corresponds to highest variability in the output and second, CMOS inverter in a chain with fan out of 4. The above test conditions and the waveform of their input voltage are shown in Figure 4.15. Monte Carlo SPICE simulations are carried out for 1000 inverter samples while p-MOSFET devices in circuit are fixed and n-MOSFETs are selected from directly extracted statistical compact model library.

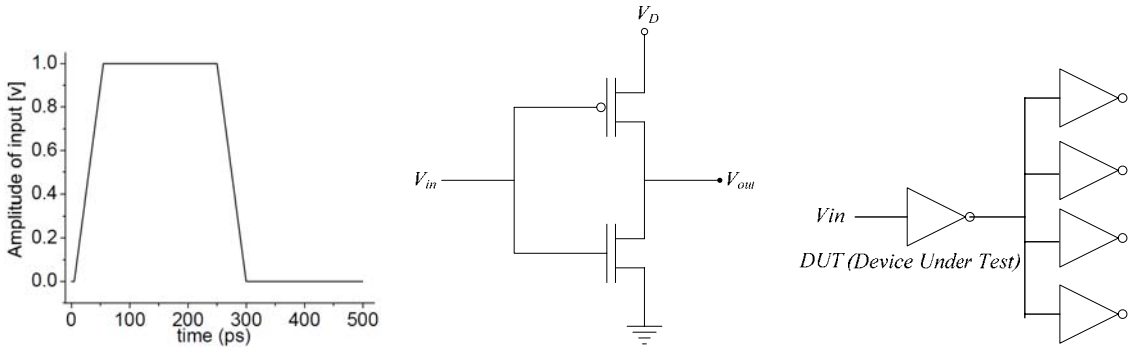


Figure 4.15: The input signal used for the simulation of inverter test beds (Left), the schematic of the CMOS inverter used for the simulation (middle) and the chain of inverters in fan out of 4 (FO4) used as a realistic test bed (Right).

By defining the rise time delay as the time difference between 50% points of transition between initial and final values of input and output voltages and by assuming the effect of statistical variability for the n-MOSFET device, the rise time delay will be a statistical variable. When the delay is measured in a no load condition, the intrinsic capacitances of devices making up the inverter will strongly affect on the delay values. This condition will give a worst case scenario with the inverter operation most sensitive to modelling inaccuracies, and hence will most clearly show the effects of differences in BSIM key parameter set size. Figure 4.16 shows the mean and the standard deviations of rise time, calculated over the ensemble of 1000 devices, for each of the different sizes of BSIM parameter sets and in both no load and FO4 conditions. It is assumed that the 7-parameter set will produce the most accurate results, and as the number of parameters is increased, the results do seem to approach a settled value. Using a 5-parameter set gives 0.2% error in the mean value of the delay compared to 7-parameter. Moreover, using 5 parameter settles

the delay normalized standard deviation (σ/μ) to 27% for no load conditions while and to about 21% for FO4 test bed in similar conditions. The dissipated energy in a CMOS inverter is divided into two parts: static and dynamic energy [112]. The static part is a result of MOSFET leakage current when the input is constant. The dynamic part occurs in the input transitions. Since the trend of static part is exactly similar to I_{off} for one device as we discussed in Figure 4.12, the trend for dynamic energy is investigated here. The comparison of dynamic energy distribution in the test bed inverters using statistical compact models with 5- and 7-parameter sets is illustrated in Figure 4.17. Figures 4.16 and 4.17 shows that 5-parameter for digital circuit applications like an inverter are sufficient for 35nm gate length MOSFETs.

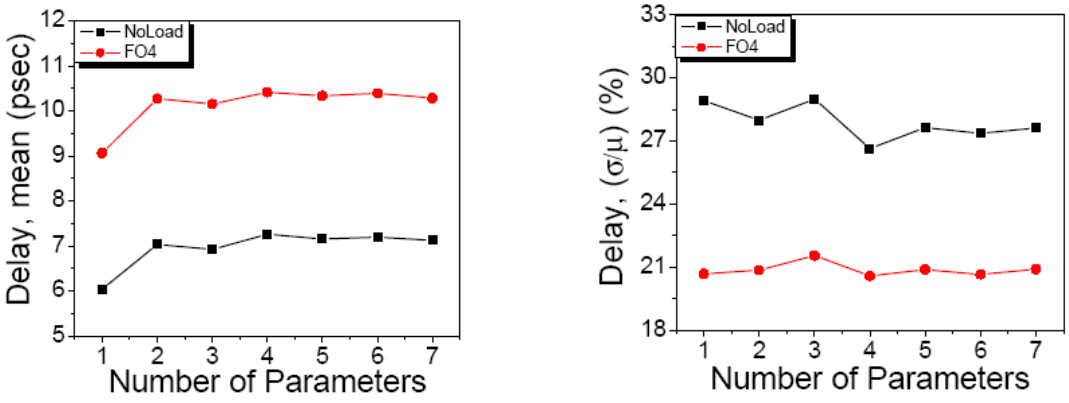


Figure 4.16: Impact of statistical parameter set selection on the mean and standard deviation of rise time delay of CMOS inverter in no load and FO4 conditions.

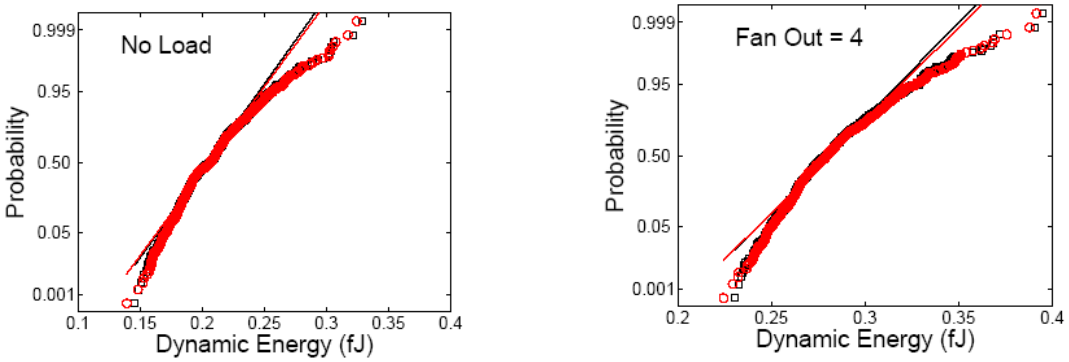


Figure 4.17: Probability plot of dynamic energy of inverter. Black square: 7-parameter set; Red circle: 5-parameter set statistical compact models. (Left) no load inverter, (Right) FO4.

4.3 Accuracy of Full Parameter Set Extraction

This section will focus on statistical compact models with a set of key BSIM parameters. The correlations between key electrical and statistical compact model parameters are shown in Figure 4.18. A very high correlation indicates that the physical meaning of the compact model parameters is maintained during statistical extraction. More importantly, such correlations may provide guidelines for developing techniques to generate statistical compact model sets based on the distributions of the figures of merit of device characteristics. There are a few stray points in (V_{th}-VTH0) plot. For example, the points with VTH0 more than 0.3V are results of numerical errors caused by optimization in the parameter extraction procedure. This owes to the contribution of many other parameters to model V_{th} of short channel devices, as discussed in chapter 3.

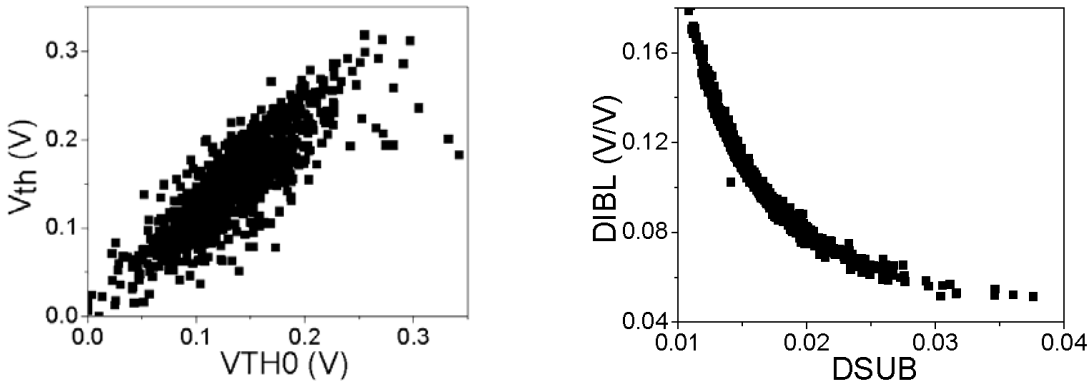


Figure 4.18: The correlation between typical electrical parameters and BSIM statistical parameters.

For the purpose of investigating accuracy of statistical parameter extraction strategy on device electrical characteristics, the distribution of figures of merit will be compared with the results of atomistic simulations. These figures of merit are simulated with HSPICE using statistical model cards which are selected from different compact models. Since in real digital circuit operation the device characteristics at high drain bias condition are more important than at low drain bias condition, here, only the result of high drain are shown.

Figure 4.19 shows a comparison between figures of merit from atomistic simulations and directly extracted statistical compact model simulations and Table 4.3 shows the statistical parameter of figures of merit and the error between atomistic simulations and direct compact models. Relative error is defined as:

$$E = 100 \times \left| \frac{F_{CompactModel} - F_{Atomistic}}{F_{Atomistic}} \right| \tag{4.3}$$

where F is a typical electrical MOSFET figure of merit. As expected, the results of physical simulations and statistical compact models match well.

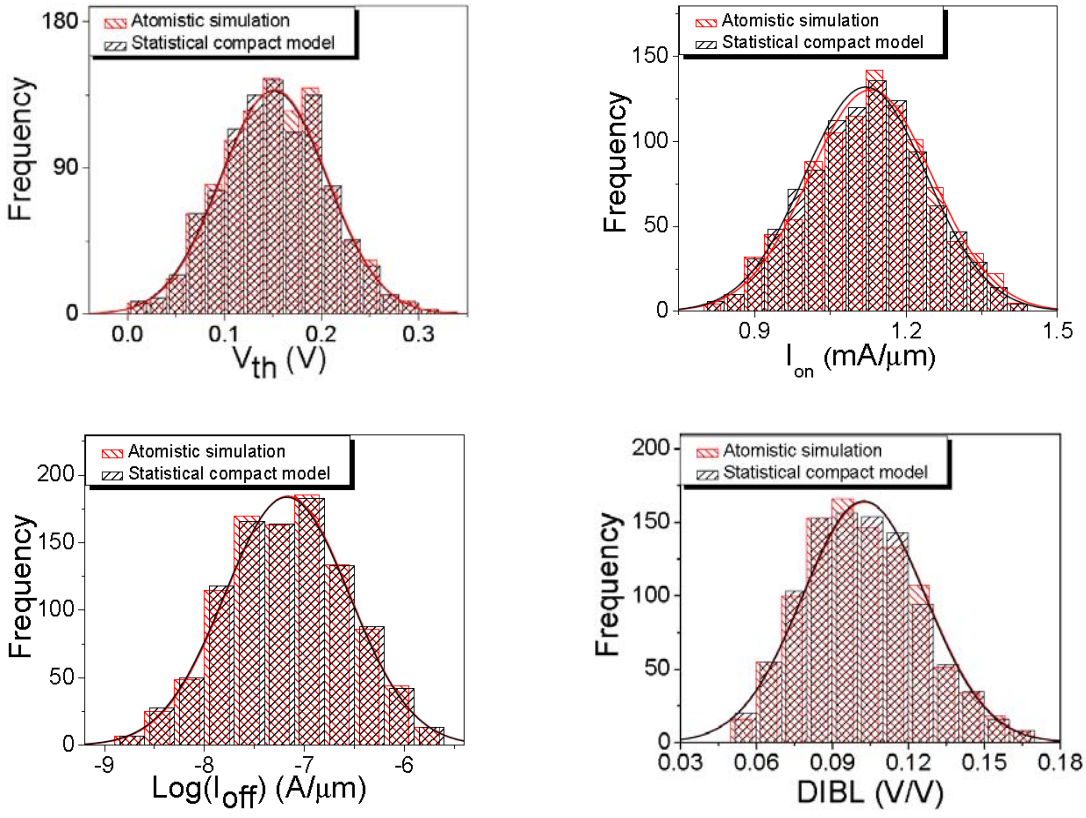


Figure 4.19: Comparison between figure of merits from 1000 atomistic simulations and directly extracted statistical compact model simulations.

Table 4.3: Mean and standard deviation of figures of merit from atomistic simulations and direct extracted statistical compact model simulations and the corresponding errors

FOM σ , μ	Atomistic simulations		Compact models		Relative error (%)	
	σ	μ	σ	μ	σ	μ
V_{th} (mV)	54.85	151.62	55.09	150.42	0.44	0.79
I_{on} (mA/μm)	0.12	1.13	0.12	1.12	0	0.88
$Log\ I_{off}$ (A/μm)	0.62	-7.17	0.62	-7.18	0	0.14
DIBL (mV/V)	24.37	102.73	24.26	102.43	0.45	0.29
SS (mV/dec)	5.92	91.97	5.74	91.69	3.04	0.30

Using drive current as an example, the error in standard deviation is zero, and the error in mean is only 0.88%.

Figure 4.20 illustrates a comparison between scatter plots of figures of merit from physical simulations and directly extracted statistical compact model simulations. It demonstrates that the direct parameter extraction approach can closely reproduce statistical 'atomistic' simulation results. Moreover, it demonstrates that although the threshold voltage variation is a good indicator for sub-threshold leakage current variation due to the very strong correlation between them, for a given threshold voltage value, I_{on} value can scatter more than ±10% around its mean. This indicates that just considering threshold voltage variation in statistical variability study cannot provide a full statistical variability picture particularly when timing variability is of major concern.

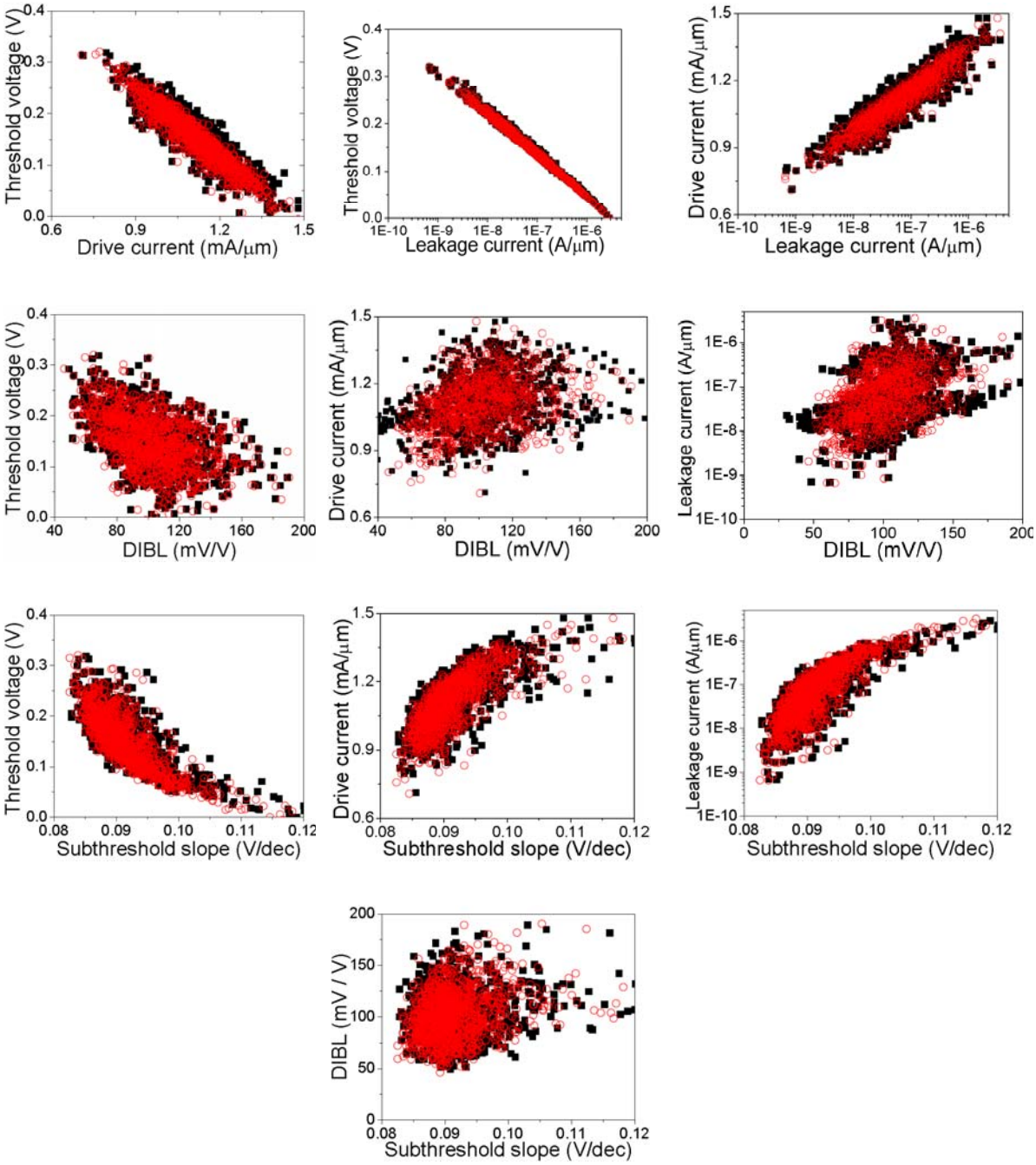


Figure 4.20: Comparison between scatter plots of physical simulations and direct extracted statistical compact models. (Black square: Physical simulations, Red circle: Direct extracted statistical compact models)

4.4 Summary

Statistical compact modelling strategies focusing on the selection of the number and type of key BSIM parameters and the corresponding statistical compact model accuracy, were studied in this chapter. We have shown that it is possible to select a limited number of BSIM parameters to describe with sufficient accuracy the effects of all main sources of variability in 35nm gate length n-MOSFETs and p-MOSFETs. Individual parameter sensitivity analysis can provide vital first-order information regarding the capability of an individual parameter to capture the impact of statistical variability on device characteristics. By selecting 7 parameters we have achieved average RMS error of 2% across I_D-V_G curves for n-MOSFETs and 2.8% for p-MOSFETs which is an acceptable level of error. Statistical compact modelling of device figures of merit and circuit operation, with focus on the impact of different number of parameters selection on the accuracy, was also performed. It was shown that for most applications 5-parameter sets can suffice.

Chapter 5

Statistical Compact Model Parameter Generation Techniques

The extracted ensemble of statistical compact model parameter sets is fundamental to the statistical simulation of circuits or systems. The direct statistical parameter extraction approach, where a set of compact model parameters is obtained for each TCAD simulated or measured device, gives the most accurate results for a given number of physically simulated ensemble size. However, this approach has two major disadvantages. Firstly, the accuracy of any Monte Carlo circuit or system simulation has a pre-determined limit, determined by the size of the underpinning compact model ensemble/library (and the circuit size). There are situations where the distribution of a parameter of interest obtained by Monte Carlo simulations to an accuracy of 2.5σ needs to be extended to 4σ or 5σ, due to additional knowledge of device physics. However such extrapolations are difficult to make in practice, using Monte Carlo simulations based on a fixed compact model ensemble size. Secondly, the direct statistical parameter extraction approach usually requires the extraction, storage and manipulation of databases containing extremely large ensembles of statistical compact models.

Common practice in Monte Carlo circuit simulation is to generate statistical compact model parameter values on the fly, based on the statistics and the correlations of a limited set of directly extracted parameters and various degrees of simplifying assumptions. In this chapter the focus will be on the statistical compact model parameter generation strategies that, can still accurately represent the distribution of, and correlations between, important device electrical parameters in nano-scaled transistors. As a result of the work described in chapter 4 in relation to the statistical parameter extraction, 6-statistical-parameter-sets are found to be of sufficient accuracy, and will be used in this chapter.

Two typical statistical parameter generation approaches are investigated in comparison with the direct statistical extraction results reported in chapter 4. The first approach is to generate statistical compact model parameters on the fly, assuming independent normal distribution for each extracted parameter [128]. This will be called “Naïve Approach”. The second approach is based on Principal Component Analysis (PCA), which preserves the correlation between extracted parameters [87]. In this “PCA Approach”, a covariance matrix, S, is generated based on the normalized direct parameter extraction results. Investigations will be carried out to compare the accuracy of digital circuit simulations carried out by Monte Carlo analysis using each of these two statistical parameter generation strategies.

In the last section of this chapter, the Nonlinear Power Method will be introduced for statistical compact model generation. The accuracy of this novel approach will be compared with the Naïve and PCA approaches by investigating figures of merit for different compact model sets.

5.1 Correlation of BSIM Parameters in Direct Statistical Parameter Extraction Strategy

Compact model parameters extracted from '*ab initio*' TCAD simulation or experiment do not always follow a normal distribution. The distribution of a number of BSIM4 parameters extracted from 35nm gate length NMOS devices are shown in Figure 5.1. It clearly demonstrates that the tails of the extracted statistical parameters deviate from a normal distribution. In addition, the extracted parameters are rarely statistically independent due to the complex physical mechanisms involved in device operation in the deca-nanometer regime, and some unavoidable aspects of the empirical nature of compact models.

Figure 5.1 illustrates normality plots of seven directly extracted BSIM parameters capturing the statistics of the underlying device ensemble. Employing the 7^{th} parameter, *MINV*, reduces the RMS error of the compact models (compare with the *IV* curves they capture) from 2.0% to 1.5% for n-MOSFETs, and from 2.8% to 2.4% for p-MOSFETs. Since the accuracy improvement of direct extraction including this 7^{th} parameter, *MINV*, is limited, only a 6-parameter set is actually employed on the parameter generation process described below. However as demonstrated in Figure 5.2, *MINV* does give a perfect example of a BSIM parameter which diverges significantly from normal distribution. The physical reason behind this distribution is that *MINV* is an empirical parameter added to BSIM compact model to provide fitting in moderate inversion condition [3]. It does not replicate a physical parameter of the device and the distribution of *MINV* cannot be reproduced by current parameter generation approaches. Using *MINV* in direct statistical compact model strategy leads to very negative or very positive value for some samples and results in a discontinuous like behaviour in the normal probability plot. However, having more data samples in expense of more time and computational resources will help to reduce this type of discontinuity.

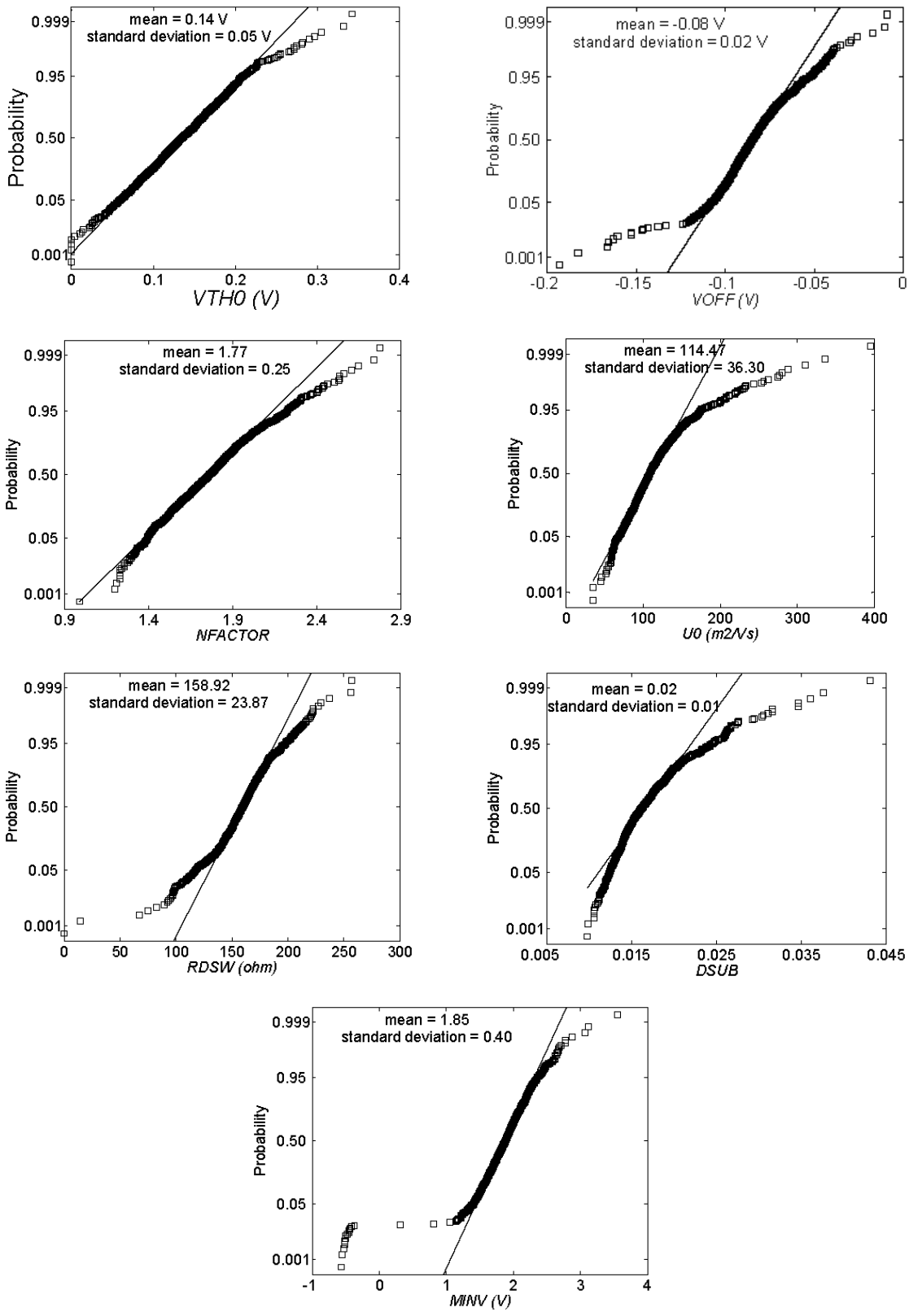


Figure 5.1: Normality plots of BSIM statistical parameters.

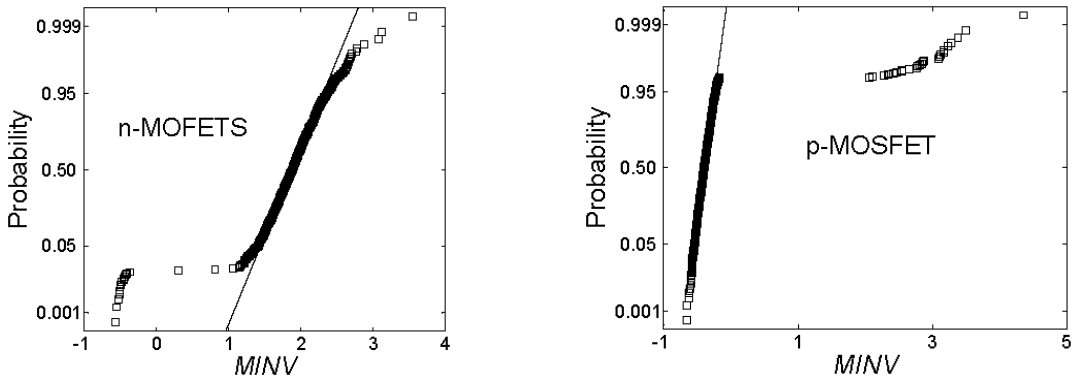


Figure 5.2: Probability plot of *MINV* for n-MOSFET and p-MOSFET.

Figures 5.3 and 5.4 show scatter plots between each pairing of the directly extracted statistical compact models parameters, for 35nm physical gate length MOSFETs. As well as this graphical indication of correlation, the calculated correlation coefficients, ρ, are also shown on the up right side of the plot. The correlation coefficient of two given statistical sets X and Y is given by:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} \tag{5.1}$$

where $Cov(X,Y)$ is covariance between X and Y and $Var(X)$ is variance of X. It should be pointed out here that there are a few additional clusters or sub-clusters in the pair plots of Figures 5.3 and 5.4. The main reason for those points is the locality of the extraction strategy and the impact of initial conditions. Using local optimization algorithms embedded in the parameter extraction softwares does not necessarily lead to the best answer. On the other hand, using the global optimization strategy will not help due to non-physical results obtained by global algorithms. The only way to improve the local strategy is through the use of appropriate initial conditions. We used the uniform or template device to set the initial conditions prior to statistical extraction because it has I_d-V_g characteristics in the middle of spread. One might use another set of appropriate initial conditions which will have a direct impact on the results and pair plots.

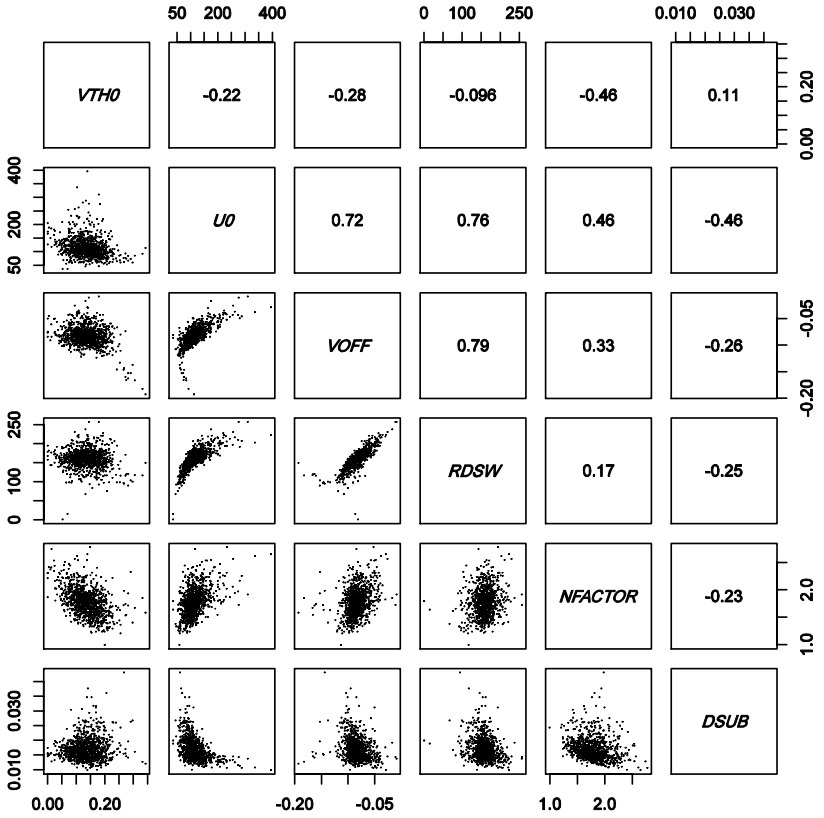


Figure 5.3: Scatter plots of selected BSIM parameters from statistical extraction of 35nm channel length of 1000 n-MOSFET devices.

These Figures clearly demonstrate that the extracted parameters are not statistically independent. The effect of correlations between parameters should be maintained in statistical circuit simulation in order to guarantee that devices used in circuit simulation are 'real'.

The importance of taking correlations into account should be obvious from an understanding of the complex nature of compact models, where each parameter makes a contribution to drain current variations. Even by assuming two statistical variables X and Y and defining a new variable $Z = X + Y$, linearly dependent on them and representing a device figure of merit, it can be shown that:

$$Var(Z) = Var(X) + Var(Y) + 2Cov(X,Y) \tag{5.2}$$

by substituting equation 5.1 into equation 5.2:

$$Var(Z) = Var(X) + Var(Y) + 2\rho\sqrt{Var(X)Var(Y)} \tag{5.3}$$

It is clear that the correlation coefficient of two variables (X and Y) plays an important role in the variance of new parameter (Z), even in the simplest case of linear dependence.

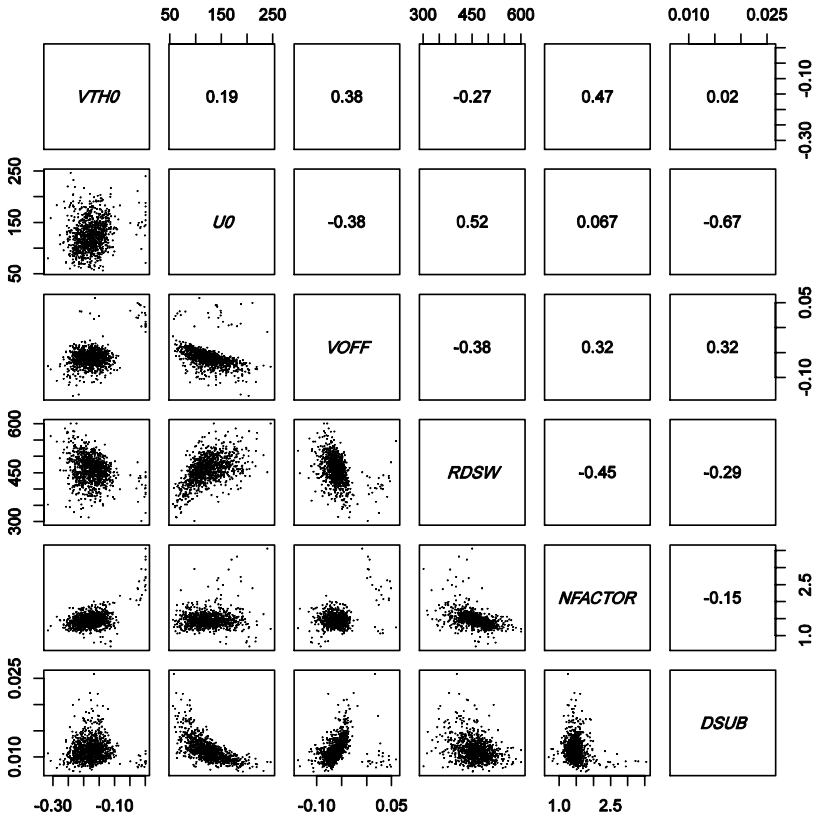


Figure 5.4: Scatter plots of selected BSIM parameters from statistical extraction of 35nm channel length of 1000 p-MOSFET devices.

5.2 Statistical Parameter Generation Based on the Naïve Approach

There are several standard statistical parameter generation strategies which can be used to transfer statistical variability information into compact models, and their accuracy is essential for achieving reliable variability aware design. One of these approaches is the Naïve approach, which is the standard approach used in most SPICE simulators [1,129]. In this approach, parameters are treated as statistically independent, and the inter-parameter correlations are ignored. The values of statistical parameters are generated by Gaussian random number generator with the mean and standard deviation coming from direct extraction results.

The probability plots of typical BSIM4 parameters are shown in Figure 5.5, which clearly demonstrate the deviation in the tails of the distribution compared to the directly extracted values due to the Naïve assumption of normal parameter distribution. Although the Naïve approach preserves the mean and standard deviation of the distribution with high accuracy, considerable errors are generated in these tails.

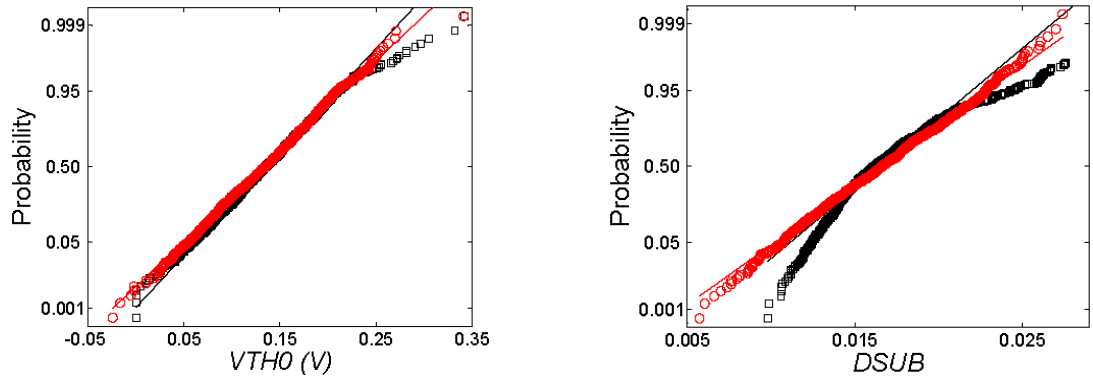


Figure 5.5: Probability plot of typical BSIM4 statistical parameters generated by Naïve approach. Black square: Direct extraction; Red circle: Naïve approach.

The scatter plots between parameters generated by Naïve approach in comparison with the direct extracted parameters are shown in Figure 5.6. The correlations between the statistical parameters are lost by using the Naïve approach, as is clearly demonstrated in the scatter plots. For instance, the correlation between *U0* and *VOFF* in the direct approach is 0.72 (according to Figure 5.3) and this correlation reduces to 0.008, i.e. approximately zero as expected, in the Naïve approach.

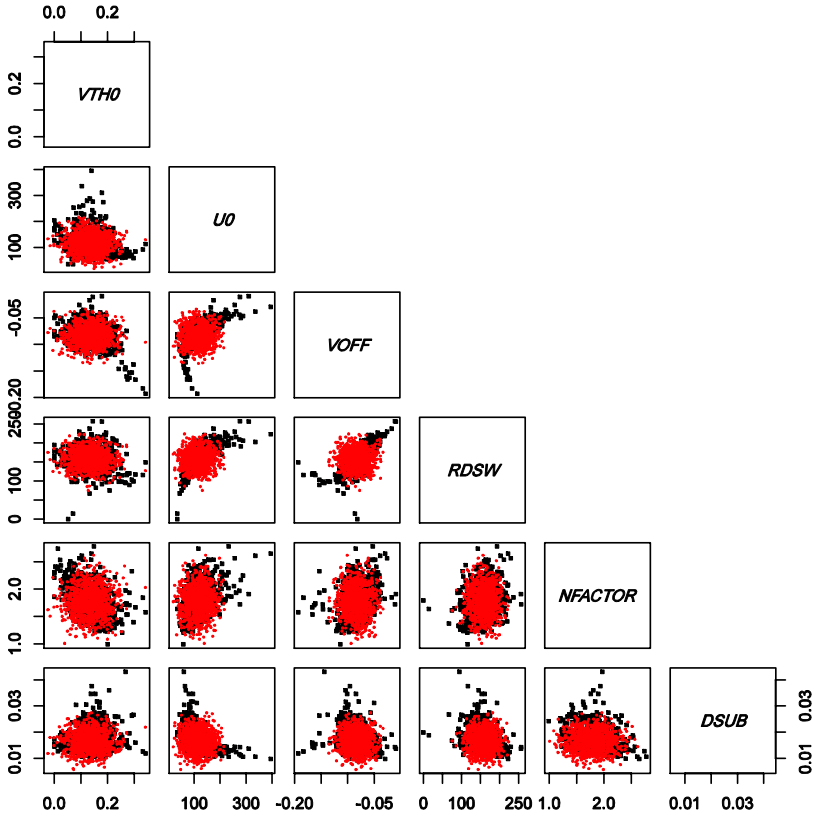


Figure 5.6: Scatter plots between BSIM4 statistical parameters. Black square: Direct extraction parameters; Red circle: parameters generated by Naïve approach.

5.3 Statistical Parameter Generation Based on the PCA Approach

PCA is a useful statistical technique that has found application in fields such as face recognition and image compression, and is a common technique for finding patterns in data of high dimension. Generally PCA involves the mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible [130,131,132]. The main purpose of employing PCA approach in statistical compact modelling is to decouple the correlations between parameters. The covariance matrix for PCA is generated from direct extraction results. In the PCA approach, the parameter values are generated by statistically independent principal components (PC) using a Gaussian random number generator. Before applying PCA on a parameter set, all parameter distributions have been normalized to a mean of 0 and standard deviation of 1 and the covariance matrix S is generated based on the normalized parameter set distributions [133]. The key step of PCA is to find the eigenvalues and eigenvectors of S, which follows:

$$U'SU = L \tag{5.4}$$

where U lists the eigenvectors, and L orders the eigenvalues. The transformed variables

$$Z = U'x \tag{5.5}$$

are the principal components, where x are the original variables. PCA itself does not require that the original multi-dimension data follow a particular distribution. However, in order to reconstruct the original data from statistically independent principal components,

it is desirable that the original data closely approximate Gaussian distributions. They can be recovered by following operation:

$$x = Uz \tag{5.6}$$

where the corresponding principal components follow Gaussian distributions with mean of 0 and variances equal to the eigenvalues L. Therefore in the PCA approach, we assume that parameters follow normal distributions but because the distributions of the BSIM4 parameters are not always normal (as shown in Figure 5.1) this inevitably introduces errors in the values of the generated statistical compact model parameters after the PCA process is completed.

A comparison of statistical correlation of parameters generated using the PCA technique, and the directly extracted parameters is given in Figure 5.7 and the correlation coefficients between parameters generated by PCA approach and the correlation coefficients between the directly extracted parameters are given in Table 5.1 for n- and p-MOSFET, respectively. These results clearly demonstrate that the correlation between parameters is well preserved by PCA approach. Since the statistical parameter extraction is carried out based on setting the initial conditions of all parameters in the model to an appropriate value from the template 35nm n- and p- transistors and the template transistors are different in terms of electrostatic (i.e., doping profile, gate overlap) as discussed in Chapter 3, the results of the statistical extraction will be different for n- and p-devices. This justifies the different correlation coefficient of identical parameter pais for n- and p-devices in Table 5.1. However, in the tails of these parameter's distributions, similar errors to those observed in the results for the Naïve approach are seen, since both approaches are based on the normal distribution assumption that parameters. A more intuitive example is shown in Figure 5.8. It illustrates the distribution of parameter *VTH0* generated from PCA process and compares it with its original distribution obtained from direct approach. Depending on the particular application of the PCA generated statistical compact models, this kind of error may give pessimistic or optimistic results in circuit simulation – and crucially loss of predictive power of the circuit simulations in the tails of the distribution.

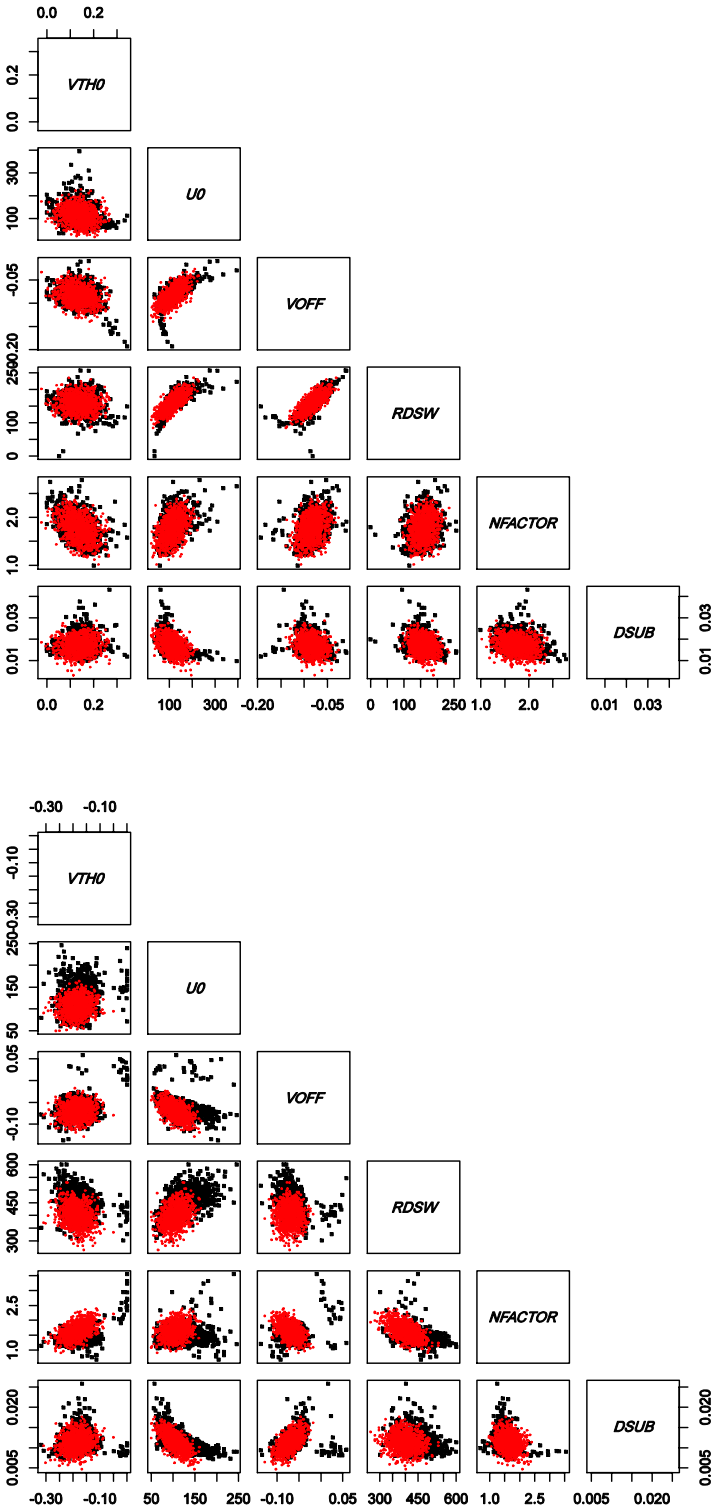


Figure 5.7: Scatter plots between BSIM4 statistical parameters for n- and p-MOSFET, respectively. Black square: Direct extraction parameters; Red circle: parameters generated by PCA technique.

Table 5.1: The correlation coefficient between parameters of n- and p-MOSFET; Down-left: Direct parameter; Up-right: PCA parameters.

n-MOSFETs					
VTH0	-0.18	-0.27	-0.077	-0.43	0.073
-0.22	***U0***	0.69	0.77	0.45	-0.44
-0.28	0.72	***VOFF***	0.78	0.33	-0.21
-0.096	0.76	0.79	***RDSW***	0.19	-0.24
-0.46	0.46	0.33	0.17	***NFACTOR***	-0.20
0.11	-0.46	-0.26	-0.25	-0.23	***DSUB***
p-MOSFETs					
VTH0	0.13	0.051	-0.23	0.5	0.17
0.19	***U0***	-0.59	0.36	0.28	-0.68
0.38	-0.38	***VOFF***	-0.026	-0.49	0.6
-0.27	0.52	-0.36	***RDSW***	-0.57	-0.17
0.47	0.067	0.32	-0.45	***NFACTOR***	-0.25
0.02	-0.67	0.32	-0.29	-0.15	***DSUB***

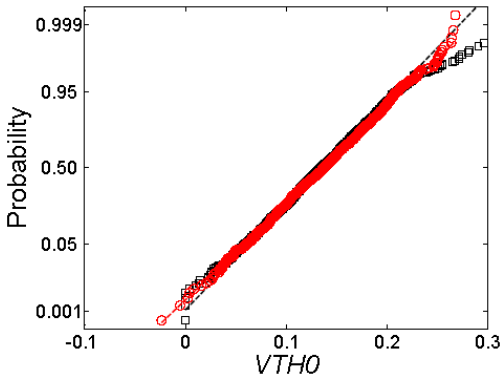


Figure 5.8: Probability plot of typical BSIM4 statistical parameters generated by PCA approach. Black square: Direct extraction; Red circle: PCA approach.

5.4 MOSFET Figures of Merit Obtained from Statistical Parameter Generation

The ability of the Naïve and the PCA approaches to reproduce the distributions and the correlations between key device figures of merit is indicative of their usefulness in statistical circuit simulation. The ability of these approaches to reproduce figures of merit is shown in Figures 5.9 and 5.10 for n- and p- MOSFETs, respectively. The results clearly indicate deviations in the tails of the distributions of each figure of merit due to the shared assumption of the Naïve and PCA approaches that the extracted parameters are normally distributed. Both approaches preserve most of the leakage current and threshold voltage distribution, aside from the tail regions. However, when considering drive current variation, both approaches produce considerable errors across the entire region. For n-MOSFET, the PCA and naïve are accurate for leakage current and threshold voltage away from the tails, and equally less accurate for drive current. From a circuit simulation point of view this implies that for highly driven components which spend a lot of their critical operation in the saturation regime, both PCA and naïve approaches are not accurate, whereas for circuits where leakage current is important like low power circuits in mobile phones, both approaches seem to be useful. Moreover, the drive current distribution in

Figure 5.9 it is concluded that PCA does provide some improvement in the accuracy of the distribution, although it is still not as accurate as directly extracted results.

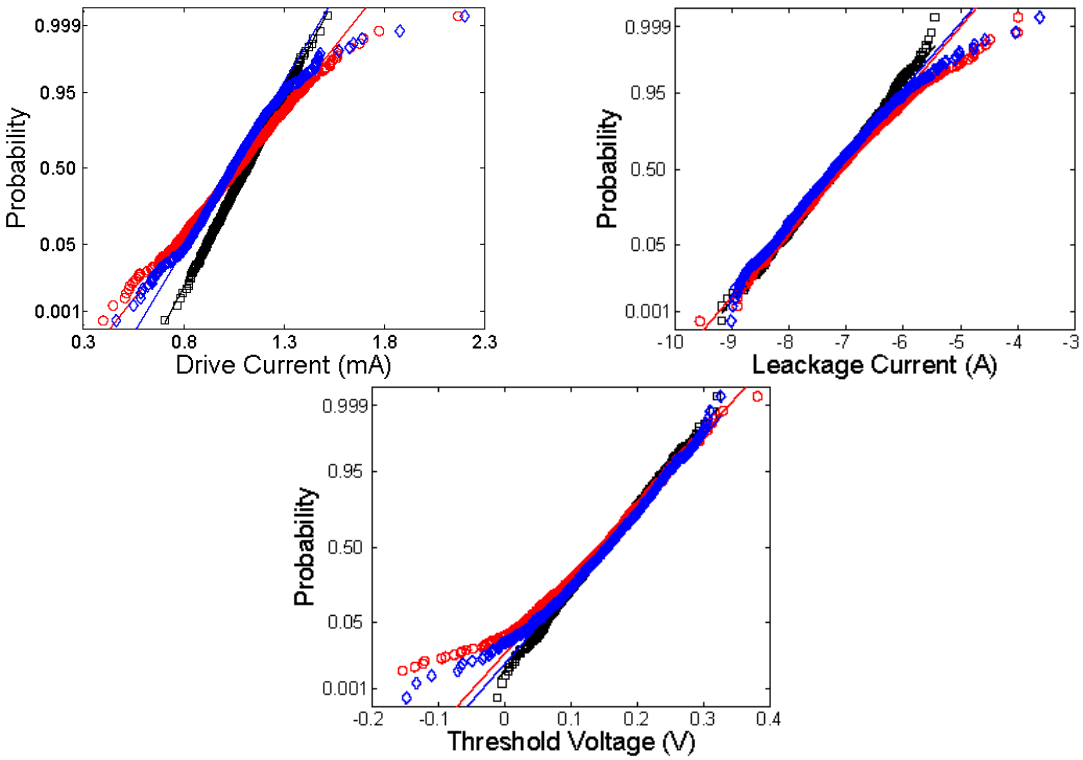


Figure 5.9: Probability plot of figures of merit of n-MOSFETs generated by different statistical approaches in high drain conditions $V_D = 1V$. Black square: direct extraction; red circle: Naïve approach and blue lozenge: PCA approach.

For p-MOSFETs if the small deviation in the upper tail of drive current in the PCA and naïve are ignored they can be used as accurate generation techniques for the devices operating in saturation regime. However, for the p-MOSFET leakage current, PCA does better prediction compared with naïve approach while for the threshold voltage the lower tails of both approaches are equally away from directly extracted results while for upper tail, PCA shows more close prediction compared with naïve results. The mean and standard deviation of figures of merit are presented in Table 5.2.

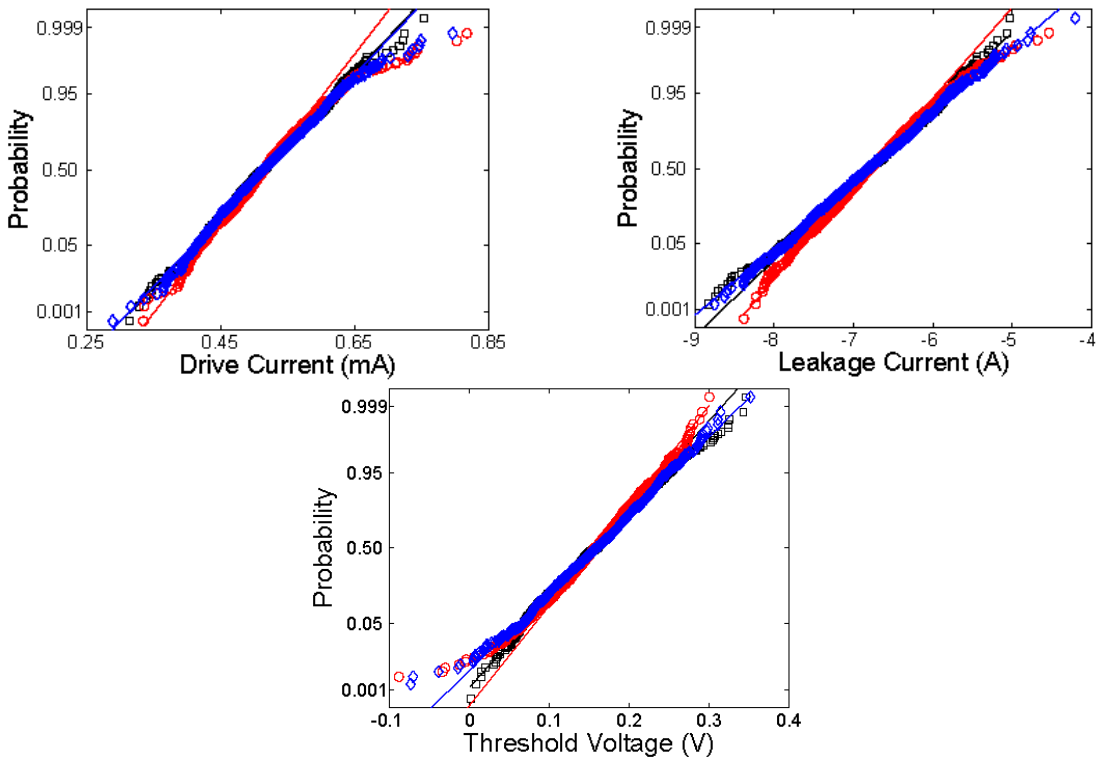


Figure 5.10: Probability plot of figures of merit of p-MOSFETs generated by different statistical approaches in high drain conditions $V_D = 1V$. Black square: direct extraction; red circle: Naïve approach and blue lozenge: PCA approach.

Table 5.2: The statistical results of figures of merits for n- and p-MOSFET

Device Type	Mean/Standard Deviation	Approaches	Figures of Merit		
			I_{on} (mA)	Log(I_{off} (A))	V_{th} (V)
n-MOSFET	Mean	Direct	1.11	-7.18	0.15
		Naïve	1.07	-7.10	0.14
		PCA	1.05	-7.20	0.15
	Standard Deviation	Direct	0.12	0.62	0.06
		Naïve	0.20	0.74	0.07
		PCA	0.16	0.72	0.06

Device Type	Mean/Standard Deviation	Approaches	Figures of Merit		
			I_{on} (mA)	Log(I_{off} (A))	V_{th} (V)
p-MOSFET	Mean	Direct	0.516	-6.80	0.157
		Naïve	0.518	-6.77	0.154
		PCA	0.517	-6.81	0.157
	Standard Deviation	Direct	0.064	0.63	0.055
		Naïve	0.065	0.56	0.060
		PCA	0.067	0.65	0.060

Another interesting question is to what extent the Naïve and PCA approaches preserve the correlation between the transistors, important figures of merit. Here, we focus on the sub-threshold regime, which is particularly important for low-power operation, and we compare the correlation between V_{th}, I_{off} and the sub-threshold slope *SS*. Figure 5.11shows the results of extracting these figures of merit from BSIM4 simulations using the Naïve and the PCA approaches and comparing them with the directly extracted data. Clearly, the PCA approach does a somewhat better job in preserving the physical correlation between the three figures of merit, whereas the Naïve approach results capture less of this physical correlation. Note that a comparison between the scatter plots of physical simulations and directly extracted statistical compact models was presented in Figure 4.20.

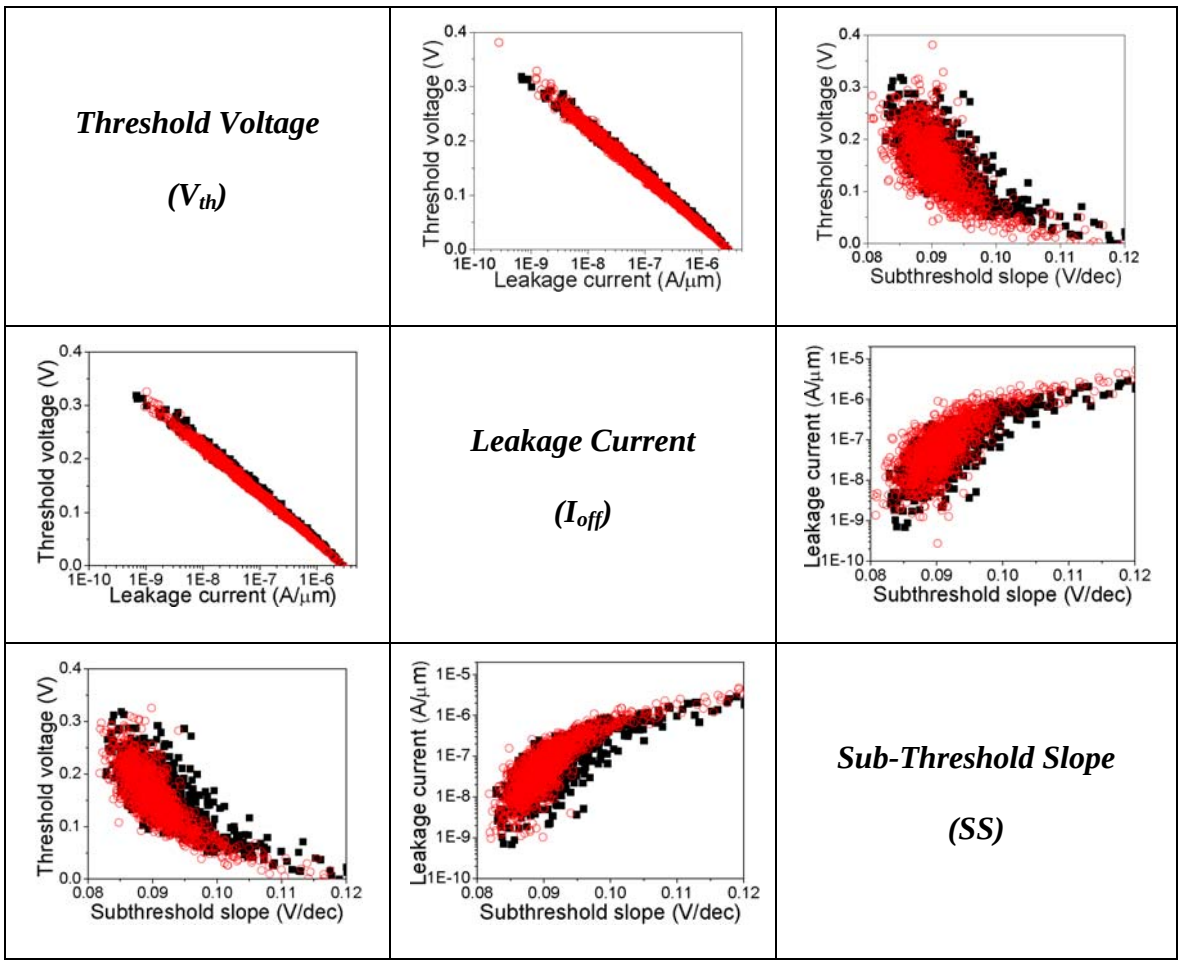


Figure 5.11: Scatter plots between figures of merit for 1000 statistical BSIM compact models. Black square: physical simulation; Red circle: PCA approach (down-left) and the Naïve approach (up-rigth)

5.5 Impact of Statistical Compact Modelling Generation Approaches on the Accuracy of Circuit Simulation

In order to investigate the impact of statistical variability on circuits, a statistical compact model card library can be built based on the direct statistical parameter extraction results described in chapter 4, and devices in each circuit can be randomly selected from

the library during statistical circuit simulation. Although this is the most accurate method to do statistical circuit simulation, the statistical sample size is pre-determined by size of compact model library. As noted in section 5.4, common practice in Monte Carlo circuit simulation is to generate statistical parameter values on the fly to improve computational efficiently. Since the accuracy of circuit simulation results are determined by the accuracy of compact models of the device, it is critically important for the designer to understand the limitations of the statistical compact modelling techniques that they employ, when using simulations to make design decisions. The Naïve and PCA statistical parameter generation approaches are now investigated with respect to circuit simulation, in comparison with directly extracted statistical results.

Circuit fundamental to analogue and digital systems are often referred to as standard cells, and the creation and the accurate characterization of libraries of these standard cells is central to digital design. A reliable strategy to capture intrinsic statistical variability in the SPICE based tools used to characterize these standard cell libraries is essential for the practical transfer of variability information from transistor-level to circuit and system simulation [119]. Characterization of standard cell libraries requires SPICE style simulation because full transistor characteristics play an important role during circuit switching. In addition, the trajectory of the current during switching is affected by the input signal slew rate [134], with a range of differing input slews to any cell requiring analysis before the cell can be accurately characterised. Since the impact of statistical variability on device characteristics is strongly bias dependant, the shape of the input waveform can also modify the statistical behaviour of a cell's properties. The simplest standard cell, and therefore the key circuit on which our statistical parameter generation approaches can be tested, is the basic inverter. In this section, based on an accurate direct statistical compact modelling approach, the effect of input waveform on the variation of delay and power dissipation of an inverter at 45nm technology generation are therefore investigated in detail under various load conditions. Results from this analysis can provide guidelines for reliable statistical standard cell characterization – currently an area of great research interest.

A CMOS inverter using minimum device size of W/L equal to 35nm/35nm for n-MOSFET and W/L of 70nm/35nm for p-MOSFET is employed to highlight the variability trend, since minimum width devices will show the largest statistical variability. The larger width p-MOSFET is necessary to correctly balance the inverter. The supply voltage is 1V, the unit load is 0.105fF (equivalent to fan-out of 1 under minimum size configuration) and various load conditions (0, 1, 2, 4 and 6 units) are considered. In order to explicitly demonstrate the effect of input slew on circuit performance variations, instead of generating the inverter input signal through an inverter chain, a 2GHz ideal symmetrical clock pulse with various rise/fall times (10, 20 and 50ps) is considerd as shown in Figure 5.12.

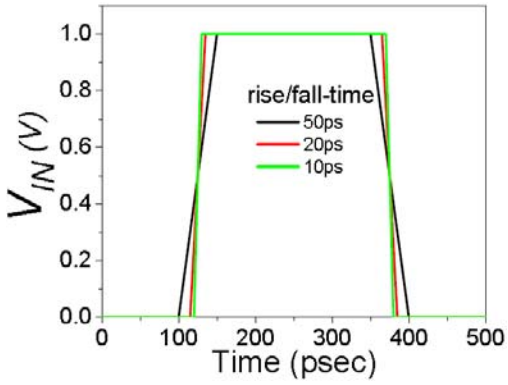


Figure 5.12: Inverter input signal with various rise/fall time.

Both the n-MOSFET and p-MOSFET devices were used in statistical simulations using: direct, naïve and PCA model cards. 1000 inverter samples are simulated and the corresponding transistors are randomly selected from statistical compact model libraries. Figure 5.13 illustrates definition of rise/fall time delays denoted by t_{dLH} and t_{dHL}, respectively.

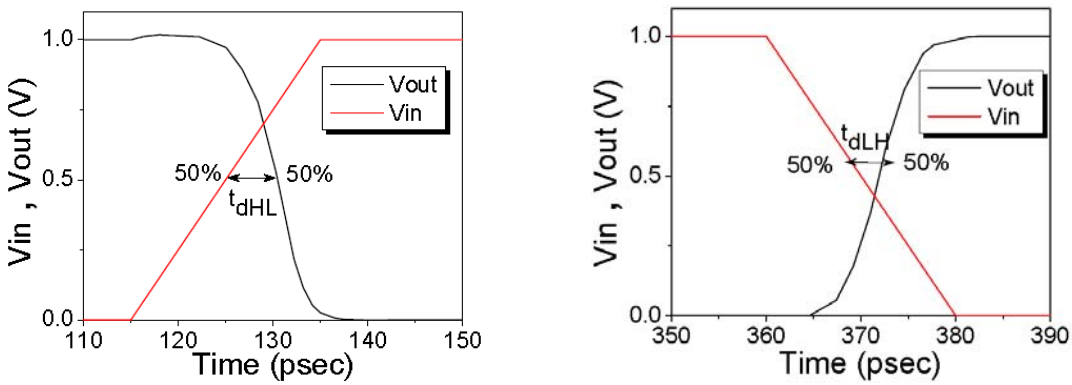


Figure 5.13: Definitions of t_{dHL} and t_{dLH}.

Figure 5.14 shows the mean and standard deviation of the fall time delay (t_{dHL}) characteristics of the inverter versus different input slew rates for a fixed load capacitance of twice the unit load of the inverter. The increased amount of the variability in higher input rise times is a result of the short circuit current which flows from the supply to ground through n- and p-MOSFET devices operating in saturation. For higher input rise times both devices spend more time in the saturation regime and the variability of the delay will be increased.

Figure 5.15 shows the same delay versus different load capacitances for a fixed input fall time of 20ps. As expected, the mean values increase almost linearly with the increase of the load capacitance and the error on the mean values of different approaches is negligible. Increasing load capacitance will improve the variability behavior because the impact of device intrinsic capacitances will be reduced compared with the external load capacitance. For the external load capacitances greater than two times of the unit load capacitance, the normalized standard deviations of t_{dHL} levels off. Considering standard deviations in Figures 5.14 and 5.15 implies the fact that naïve approach produces more error compared with PCA approach in respect to direct approach. This is justified by the fact that the naïve approach does not consider the correlation of the parameters, and as a result, the drive current of naïve approach in the n-MOSFET is far away from the direct results as shown in Figure 5.9. It is well known that the fall time delay is mostly dependent

on the n-MOSFET drive current and using the naïve results adds in considerable amount of error in the fall time delay.

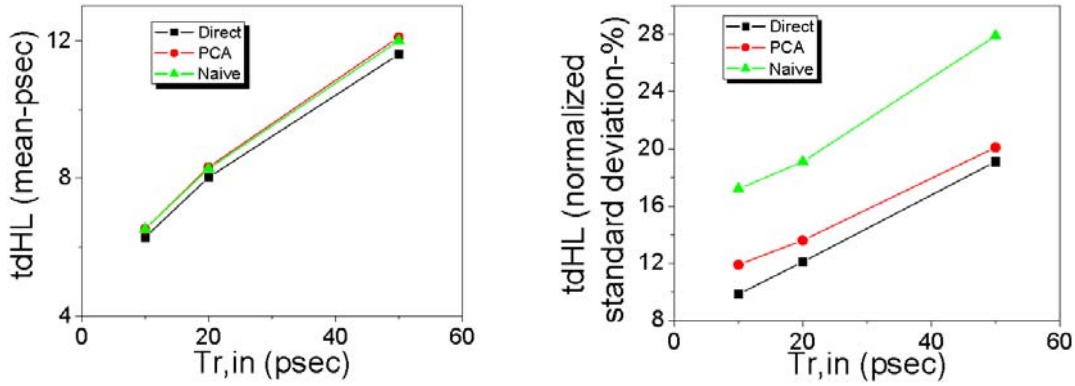


Figure 5.14: Statistical trend of inverter fall time delay versus input rise time, mean values (left), coefficient of variation values (right)

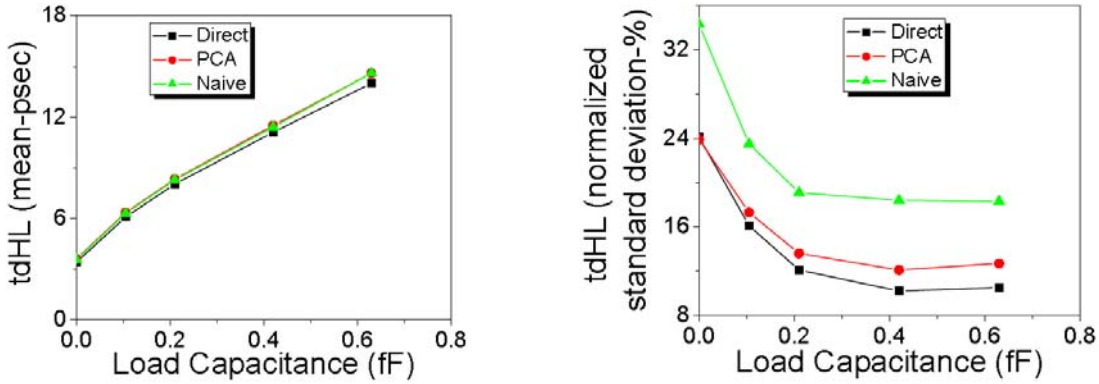


Figure 5.15: Statistical trend of inverter fall time delay versus load capacitance, mean values (left), coefficient of variation values (right)

The results for rise time delay (t_{dLH}) are presented in Figures 5.16 and 5.17. As expected, their behaviour is similar to those of fall time delay but with reduced variability due to the larger size of the p-MOSFET. Since the dominant device in determining the rise time delay is the p-MOSFET and in this transistor the naïve and PCA techniques have produced closer match in the drive current distribution compared to the n-MOSFET (as discussed in Figure 5.10) the rise time delay behaviour versus different input rise times or

different load capacitances for the naïve and the PCA approaches are better matched compared to the fall time delay where the n-MOSFET was dominant.

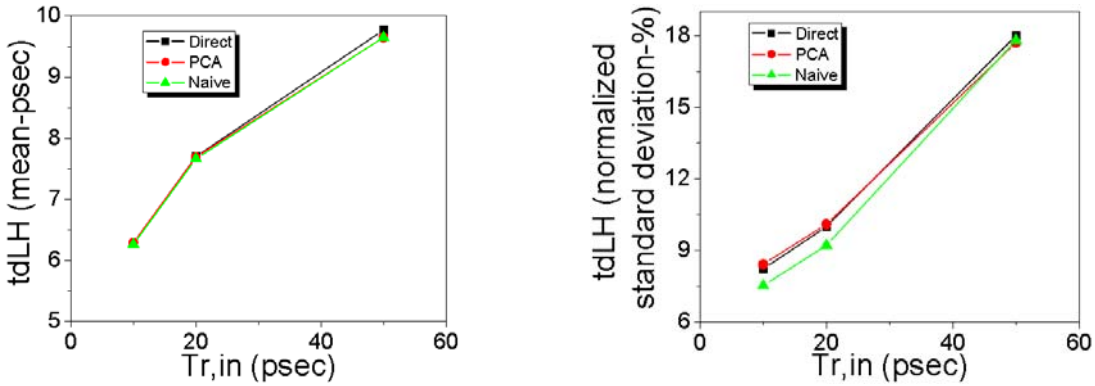


Figure 5.16: Statistical trend of inverter rise time delay versus input rise time, mean values (left), coefficient of variation values (right)

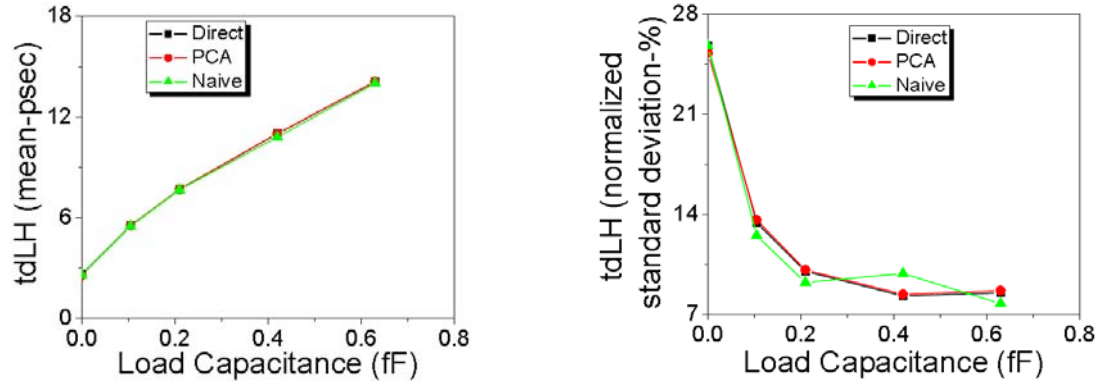


Figure 5.17: Statistical trend of inverter rise time delay versus load capacitance, mean values (left), coefficient of variation values (right)

Power dissipation in digital circuits can be separated in static and dynamic components. Static dissipation is due to sub-threshold and gate leakage current flowing from the supply trough the transistors that are nominally off. Dynamic dissipation depends on the size of the capacitive loads. There is a second part in the dynamic power dissipation determined by the short circuit current through the two transistors during the switching period [135]. This is due to the existence of a DC path for the current flowing from supply to ground during the switching. Since input signal rise/fall time will determine the length of time that

inverter can stay in short circuit condition, it will have a dramatic impact on power dissipation variation. Figures 5.18 and Figure 5.19 show the total energy dissipation during a full input cycle. Variation in the energy dissipation at this particular case is introduced by the short-circuit component since the leakage power dissipation is negligible. The conclusion is that the mean values of the energy in different approaches are very close together. Clearly as the input signal rise time increases, the total switching energy increases due to short circuit dissipation, and as the load capacitance increases, the first component of dissipation increases, as expected. Considering the trend of energy standard deviation versus load capacitance in Figure 5.18, it is concluded that the error between the PCA or naïve approache (which both give vary similar results) with respect to the directly extracted approach is almost constant. Considering energy versus load capacitance in Figure 5.19, all approaches converge together for large load capacitance. This is due to the fact that for larger load capacitances the contribution of dynamic power will be increased compared to short circuit power. This results in less impact of the device parameters on the dissipated energy.

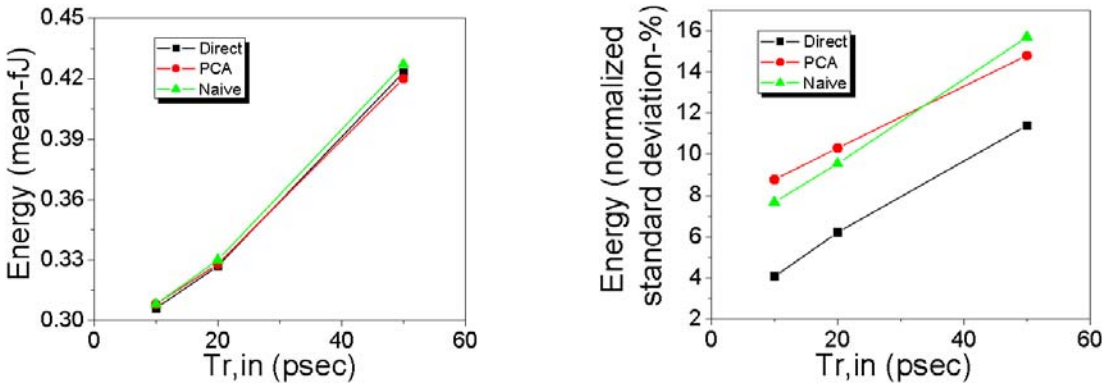


Figure 5.18: Statistical trend of inverter consumed energy versus input rise time, mean values (left), coefficient of variation values (right)

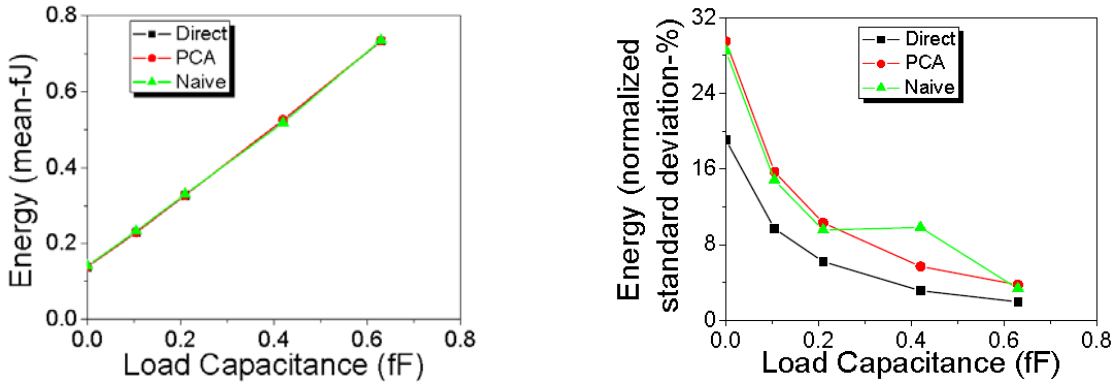


Figure 5.19: Statistical trend of inverter consumed energy versus load capacitance, mean values (left), coefficient of variation values (right)

Compared with the accurate direct statistical compact modelling approach, we have investigated the effect of simplified approaches to compact model parameter generation, combined with the effect of input signal slew rate on the statistical properties of an inverter at the 45nm technology generation. Our results demonstrate that, depending on the size of load capacitance, the input slew rate can have a dramatic impact on the variation of circuit performance.

5.6 Statistical Parameter Generation Based on the Nonlinear Power Method

There is a challenge in the statistical compact models related to statistical variability which is associated with the statistical generation of compact model parameters preserving the shapes and the correlations of the originally extracted statistical compact model parameters. Previously it has been demonstrated that the accuracy of the statistical circuit simulation can be somewhat improved in a number of cases by using PCA in generating statistical compact model parameters. However, PCA assumes normal distributions of the extracted compact model parameters, which adversely affects the tails of the distributions of the statistical circuit simulation results. It was shown that in a number of cases it is as accurate as the naïve approach.

In this section, we introduce a Nonlinear Power Method (NPM) approach for statistical compact model parameter generation. It can accurately reproduce the shapes and tails of non-normally distributed directly extracted statistical compact model parameters preserving also the correlations. The accuracy of this approach is compared with the previously used PCA method.

The key advantage of the NPM method stems from the capability to generate univariate or multivariate non-normal distributions with an arbitrary covariance matrix [133] from a set of analytical equations. The NPM can be described as follows. Let Y_i denote the standard non-normal random variable with zero mean and unit variance representing the chosen normalized i*th* directed extracted statistical compact model parameter that needs to be reconstructed with the non-normal distribution property preserved. NPM generates the non-normal random variable Y_i using the s*th* order polynomial transformation of the standard normal random variable $Z_i \sim N(0,1)$ as $Y_i = c_i^T Z_i$ where $c_i^T = (c_{0i}, c_{1i}, ..., c_{si})$ are unknown constants and $Z_i^T = \left(1, Z_i, Z_i^2, ..., Z_i^S\right)$. Setting $s = 2$ allows controlling the degree of skew and setting $s = 3$ controls degree of both skew and kurtosis. Therefore, we derive the expressions for the first four moments of Y_i in order to determine the constants C_{ki}. This requires knowing the even central moments of Z_i up to the 12*th* order. The odd central moments of Z_i are equal to zero. Substituting the values of the central moments of Z_i into the moment formulas of Y_i leads to an algebraic system of nonlinear equations setting $s = 3$ as follows [136]:

$$E[Y_i] = c_i^T E[Z_i] \tag{5.7}$$

$$VAR[Y_i] = E\left[\left(c_i^T Z_i\right)^2\right] - \left(E\left[c_i^T Z_i\right]\right)^2 \tag{5.8}$$

$$y_{1i} = \frac{E\left[\left(c_i^T Z_i\right)^3\right] - 3E\left[\left(c_i^T Z_i\right)^2\right]\left(E\left[c_i^T Z_i\right]\right) + 2\left(E\left[c_i^T Z_i\right]\right)^3}{\left(VAR\left[Y_i\right]\right)^{3/2}} \tag{5.9}$$

$$\begin{aligned} y_{2i} = & \frac{E\left[\left(c_i^T Z_i\right)^4\right] - 4E\left[\left(c_i^T Z_i\right)^3\right]\left(E\left[c_i^T Z_i\right]\right) - 3\left(E\left[\left(c_i^T Z_i\right)^2\right]\right)^2}{\left(VAR\left[Y_i\right]\right)^{3/2}} \\ & + \frac{12E\left[\left(c_i^T Z_i\right)^2\right]\left(E\left[c_i^T Z_i\right]\right)^2 + 6\left(E\left[c_i^T Z_i\right]\right)^4}{\left(VAR\left[Y_i\right]\right)^{3/2}} \end{aligned} \tag{5.10}$$

where $E[Y_i]$ is a mean value, $VAR[Y_i]$ is a variance, y_{1i} is the sample skew and y_{2i} is the sample kurtosis which are given in [137]. In the case of the first three moments, setting $s = 2$ reduces the algebraic system of nonlinear equations to the system of equations (5.7), (5.8) and (5.9) where the constant c_{3i} is set to zero. This system of equations (5.7), (5.8), (5.9) and (5.10) is simultaneously solved to provide the constants c_{ki}. In order to maintain the correlations between directly extracted statistical compact model parameters, it is necessary to calculate the intermediate correlation matrix between non-normal random variables Y following the procedure described in [133] and using Isserlis's theorem [138,139]. The elements of the intermediate correlation matrix for the setting $s = 3$ can be computed using the following expression

$$\rho_{Y_i Y_j} = E\left(c_i^T Z_i c_j^T Z_j\right) \tag{5.11}$$

where $\rho_{Y_i Y_j}$ is the desired correlation between two chosen statistical compact model parameters Y_i and Y_j and $\rho_{Y_i Y_j} = E\left(Z_i Z_j\right)$ is called the intermediate level correlation coefficient between two standard normal random variables Z_i and Z_j. *S*etting $s = 2$ replaces coefficients c_{3i} and c_{3j} with zero in equation (5.10). A total of $(N-1)\times N/2$ polynomial equations now need to be solved in order to obtain a complete intermediate

correlation matrix. In this case, setting $s = 3$ and $N = 6$ indicates that the roots of 15 cubic polynomials need to be obtained. Finally, the multivariate non-normal distribution of the random variable Y_i will be generated using a combination of Singular Value Decomposition of the intermediate correlation matrix and the NPM approach. As a result, either the first three or the first four moments of the extracted statistical compact model parameter distributions will be preserved in the generated distribution, dependent on including the first three or four moments of the initial distribution respectively, and the correlations between these statistical compact model parameters will be preserved.

Applying the NPM approach, statistical BSIM parameters have been generated based on directly extracted results. For two of the statistical compact model parameters of the n-MOSFET (*VOFF* and *RDSW*), Figure 5.20, compares randomly generated parameter distributions using PCA and NPM with the direct parameter extractions. It is clear that the NPM approach shows significant improvement in reproducing the shape and tail of the direct parameter distributions compared with the PCA approach.

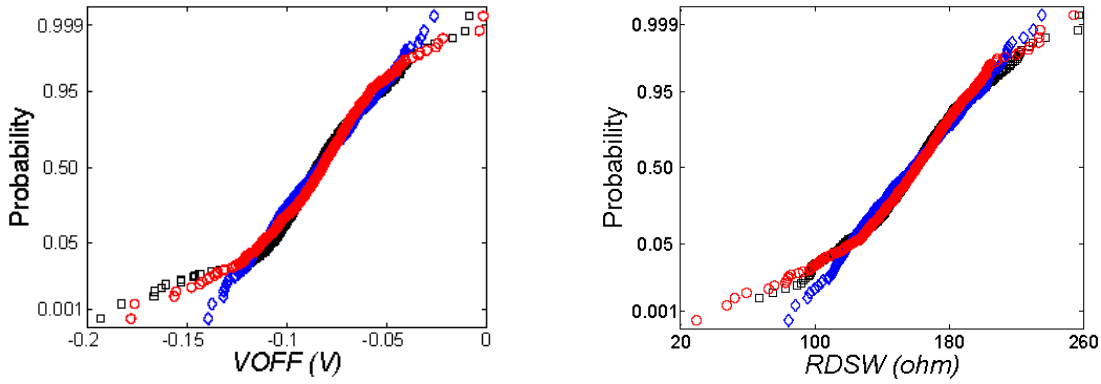


Figure 5.20: Probability plot of BSIM parameters from direct statistical parameter extraction compared with PCA and NPM generated parameter approach. Black square: direct extraction; red circle: NPM approach and blue lozenge: PCA approach.

The correlations between extracted and generated statistical compact model parameters using NPM approach for 35nm n-MOSFETs are presented in Table 5.3. It can be clearly seen from these results that the correlations between all generated statistical parameters have been well preserved.

Table 5.3: The correlation coefficient between parameters; Down-left: Direct parameter; Up-right: NPM parameters.

VTH0	-0.14	-0.29	-0.074	-0.42	0.059
-0.22	***U0***	0.57	0.68	0.44	-0.33
-0.28	0.72	***VOFF***	0.77	0.26	-0.27
-0.096	0.76	0.79	***RDSW***	0.13	-0.23
-0.46	0.46	0.33	0.17	***NFACTOR***	-0.17
0.11	-0.46	-0.26	-0.25	-0.23	***DSUB***

Next, we illustrate the advantages of the NPM in relation to the accuracy of statistically generated current-voltage characteristics. The comparison with the PCA approach is based on distributions of three key figures of merit of the statistically generated current-voltage characteristics - V_{th}, I_{on} and I_{off} distributions and their correlations. Figure 5.21 compares for the n-MOSFET the correlation between the selected figures of merit obtained from the original target current-voltage characteristics used in the statistical compact model parameter extraction and from current voltage characteristics based on statistical compact model parameter sets generated by using the PCA and the NPM approaches. It is clear that the shape of the correlation clouds is much better preserved by using NPM approach. In order to illustrate the impact of the NPM statistical compact model parameter generation approach on statistical circuit simulation, the propagation of the delay and energy in a simple CMOS inverter is again studied. We investigate the statistical accuracy of the PCA and NPM approaches in reproducing the distributions of key figures of merit of the inverter in comparison to simulations using the directly extracted statistical parameter sets as a benchmark. Probability plots of the fall time propagation delay (t_{dLH}), rise time propagation delay (t_{dHL}) and dissipated energy are compared together in Figure 5.22. In all cases the best agreement with the direct simulation results are obtained using the NPM approach.

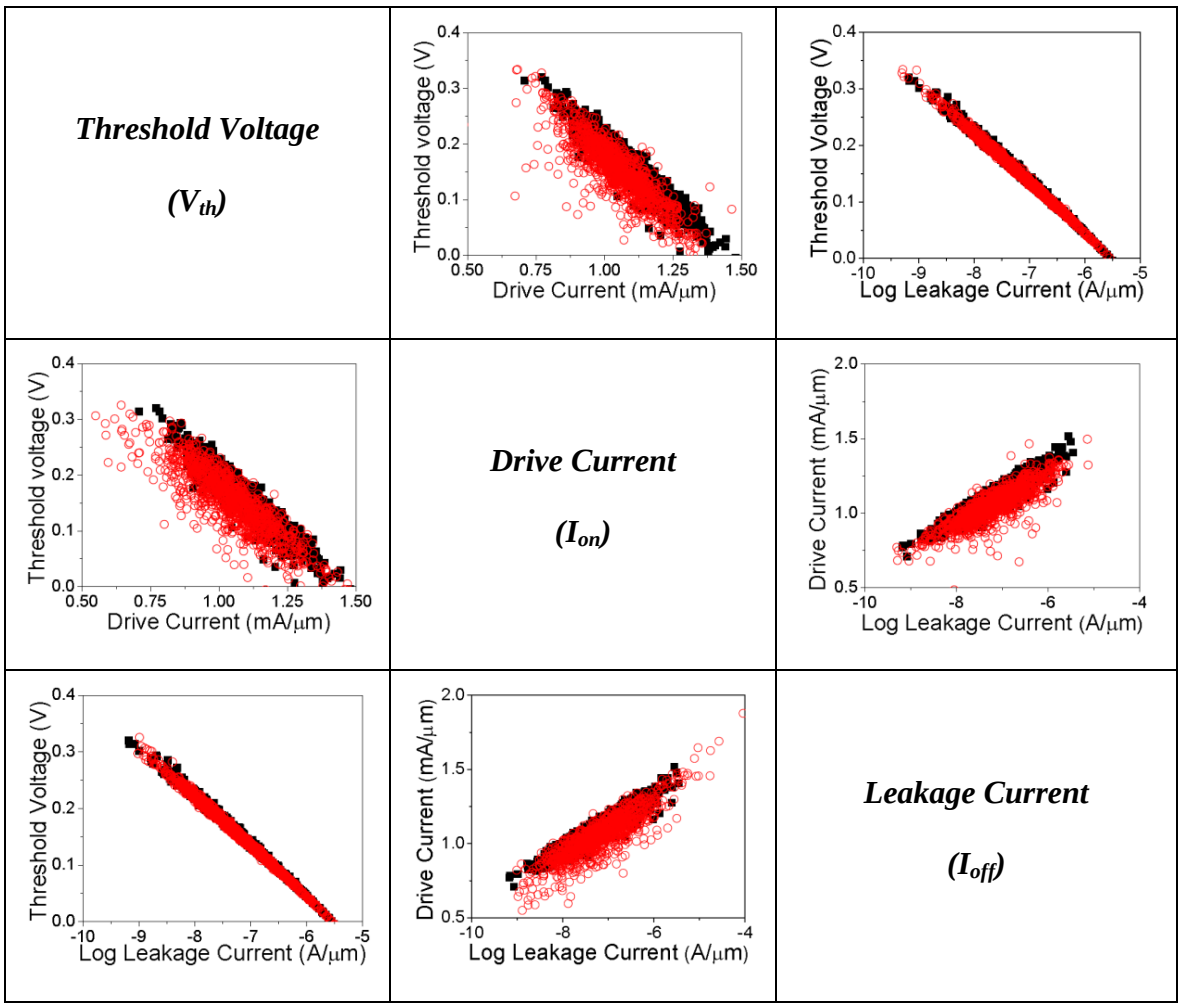


Figure 5.21: Correlation between electrical parameters. Bottom-left: Comparison between results from direct statistical compact model and PCA; Top-right: Comparison between results from direct statistical compact model and NPM approach.

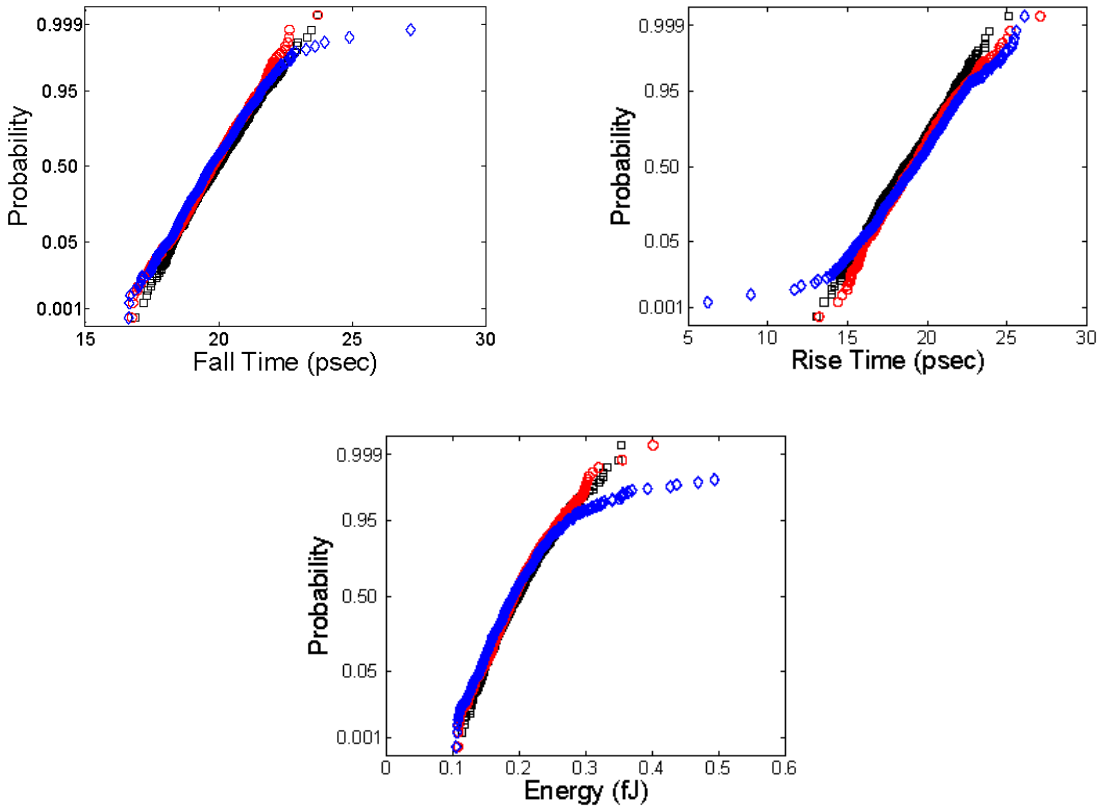


Figure 5.22: Probability plot of rise, fall time delays and dissipated energy of inverter simulations using statistical directly extracted, PCA and NPM generated statistical compact model parameters approach. Black square: direct extraction; red circle: NPM approach and blue lozenge: PCA approach.

5.7 Summary

Based on simulated statistical variability in a state-of-the-art 35nm gate length device, the benchmarking of statistical compact model strategies is carried out. The results indicate that the naïve approach, which generates statistical compact model parameters assuming independent normal distribution for each extracted parameter, produces considerable errors in circuit simulation compared to directly extracted parameters.

The PCA approach for the generation of statistical compact model parameters can produce better results than that of the naïve approach, but is limited by its assumption of

normally distributed parameters. An accurate treatment of the statistical compact model distribution, the NPM approach, not only maintains the correlations between generated statistical compact model parameters, but also accurately captures the tails and the nonlinear shape of their distributions.

The NPM approach can provide the accurate and reliable statistical compact model generation that is required by the design community. Although the direct parameter extraction approach gives the best accuracy, NPM is a step toward the development of a computationally efficient general statistical compact modelling approach in the presence of purely statistical variability or in combination with process induced variability.

Chapter 6

Effect of Device Geometry on Statistical Device Characteristics

Device variability is a function of device geometry, and this chapter considers the effect of the transistor gate width and length, on the statistics of MOSFET electrical characteristics. Although these are the simplest possible geometry effects to study, they are of considerable practical importance. Typical gate library in modern ICs has around 30 combination and layouts of *W/L* for each of n- and p-MOSFETs. The minimum width, square device was investigated in previous chapters and in this chapter, we extend the study to wider devices, and devices with tuned gate lengths.

6.1 Effect of Width on Variability

Two approaches will be used to study the impact of channel width on the device parameters: 1) a parallel component approach where a circuit consisting of a parallel combination of square devices is used to represent a wider device, and 2) full statistical atomistic simulation. The first method has the advantage of computational efficiency as '*ab initio*' device simulation (or device measurement) is the most costly part of the

compact model characterisation process, and obtaining a statistical ensemble of only a single transistor width results is a significant save in computational effort. Disadvantages of this method are a lack of discrimination between transport in the middle of the transistor compared with transport at the edge of its width (unless specific 'edge' building block transistors are included, or it is already known that such edge effects are negligible) and a limitation to circuit transistor widths that are multiples of the base transistor width (unless interpolation is employed). These problems will not be considered here. More fundamentally, the method fails to capture correlations between various points across the width of wider transistors, or where transport charge percolates across the width of devices to a significant extent in travelling from the source to the drain. A detailed consideration of the sources of such errors is given in [140]. The second method, full statistical simulation of each required transistor width, whilst a significant computational (or experimental) burden, does not suffer from these problems. A goal of this chapter is to evaluate quantitatively the errors in a range of device figures of merit resulting from the use of the first, significantly faster approach, and assess whether it is accurate enough to be usefully employed.

6.1.1 Parallel Component Approach

In the parallel component approach, I_d-V_g characteristics of 1000 devices with nominal gate width and length of 35nm are simulated using the Glasgow University atomistic simulator. The simulations include combined sources of variability: RDD, LER and PGG. I_d-V_g characteristic sets for different width (70, 140, 210, 280nm) transistors are constructed from random combinations of this base set of 1000 devices using a Python script to collate the results and calculate resultant drain currents in mA/μm. Theory suggests that if the base transistors are statistically independent in their device variability the resultant average drain currents (per μm) should be identical to those of the base device within statistical errors of $\sqrt{1000}/1000 \approx 3\%$, and the statistical variability of the drain current should drop by a factor of $\sqrt{N}$, where N is the multiple of the base device gate width.

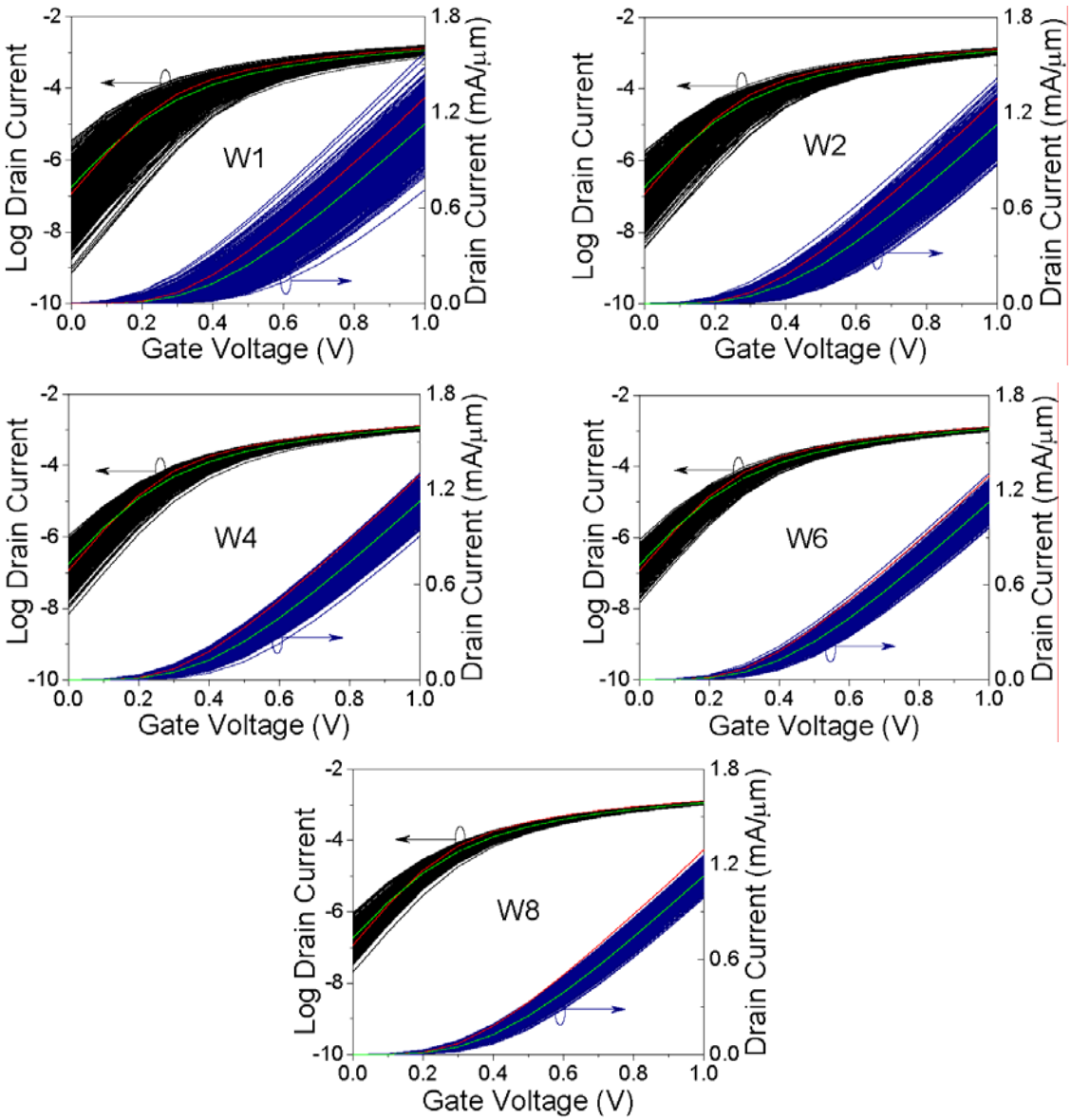


Figure 6.1: Variability in the I_d-$V_g$characteristics of a statistical sample of 1000 microscopically different 35nm n-MOSFETs whose gate widths are multiples of 35nm (V_D = *1V*). Red curves show idealised devices. Green curves are the average currents for each statistical set device.

The resultant I_d-V_g curves for different channel width of high drain bias (V_d =*1V*) are shown in Figure 6.1. The gate width of each device (*W1*, *W2*, etc.) is noted in multiples of 35nm. Red curves show the equivalent curve for a uniform/idealised transistor, while the green curve is the statistical average of the currents in the statistical set.

Table 6.1: Statistical average of the drain current for different gate width n-MOSFETs

	Drain Current (A/μm)				
Width (nm) / Gate Voltage (V)	35	70	140	210	280
0.0	1. 77E-4	1. 78E-4	1. 78E-4	1. 69E-4	1. 80E-4
0.1	1.80E-3	1.81E-3	1.81E-3	1.74E-3	1.82E-3
0.2	1.21E-2	1.21E-2	1.22E-2	1.18E-2	1.22E-2
0.3	4.83E-2	4.83E-2	4.84E-2	4.75E-2	4.85E-2
0.4	0.13	0.13	0.13	0.12	0.13
0.5	0.24	0.24	0.24	0.24	0.24
0.6	0.39	0.39	0.39	0.39	0.39
0.7	0.56	0.56	0.56	0.56	0.56
0.8	0.74	0.74	0.74	0.74	0.74
0.9	0.93	1.13	1.13	1.13	1.13
1.0	1.13	1.13	1.13	1.12	1.13

The average drain currents of different width transistors, in mA/μm, at each gate bias point are collected in Table 6.1 and plotted in Figure 6.2. As expected, increasing the width of the device under consideration has little effect on the average drain current per μm gate width. Above threshold, any discrepancy is less than 3% as expected from the theory. However, below $V_g = 0.2V$ there is some effect, with the largest discrepancy being up to 6% of the drain current of wide devices at $V_g = 0V$. This shows that 35nm gate length devices do show correlated effects across their widths for distances greater than 35nm. We would expect that such differences would be most apparent where charge percolation from source to drain is more effected due to lack of screening in the channel, below threshold, and this is indeed the case. In addition, the effects of electrostatics on drain current in subthreshold are exponential in nature, amplifying their effect.

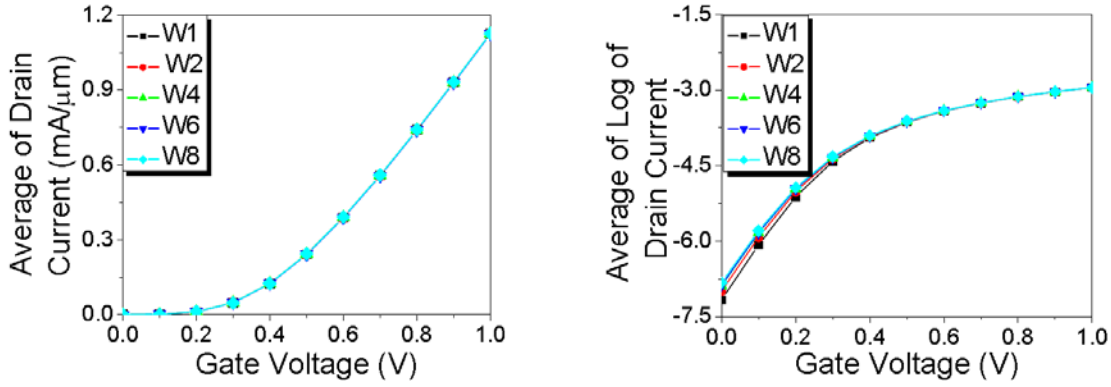


Figure 6.2: Average drain currents obtained from Figure 6.1.

Table 6.2: Standard deviation of the drain current for different gate width n-MOSFETs

	Standard deviation of drain current (A/μm)				
Width (nm) / Gate Voltage (V)	35 ($N=1$)	70 ($N=2$)	140 ($N=4$)	210 ($N=6$)	280 ($N=8$)
0.0	3.29E-7	2.32E-7	1.63E-7	1.26E-7	1.24E-7
0.1	2.55E-6	1.79E-6	1.27E-6	1.01E-6	9.32E-7
0.2	1.17E-5	8.16E-6	5.80E-6	4.70E-6	4.14E-6
0.3	3.11E-5	2.17E-5	1.54E-5	1.26E-5	1.09E-5
0.4	5.50E-5	3.85E-5	2.71E-5	2.24E-5	1.92E-5
0.5	7.59E-5	5.32E-5	3.74E-5	3.10E-5	2.66E-5
0.6	9.18E-5	6.44E-5	4.53E-5	3.76E-5	3.23E-5
0.7	1.03E-4	7.24E-5	5.09E-5	4.24E-5	3.64E-5
0.8	1.125E-4	7.82E-5	5.50E-5	4.59E-5	3.95E-5
0.9	1.18E-4	8.23E-5	5.81E-5	4.84E-5	4.18E-5
1.0	1.22E-4	8.55E-5	6.04E-5	5.04E-5	4.35E-5

Table 6.2 presents the standard deviations of the drain current at different values of the gate voltage for different width devices. The results clearly indicate a decrease at the rate of $\sqrt{N}$ for wider devices, where N is the multiple of the base device gate width. Figure 6.3

illustrates the trend of standard deviations of drain current as a function of $1/\sqrt{W}$ for different gate bias points. As evident from this figure, the trend is almost linear for the standard deviation of drain current versus $1/\sqrt{W}$ for each gate bias point. Since the variation of drain current in subthreshold region (V_g = *0, 0.1* and *0.2V*) is not obvious from this figure, the standard deviations on a logarithmic scale has been used to prepare a more clear figure showing the trend for subthreshold region.

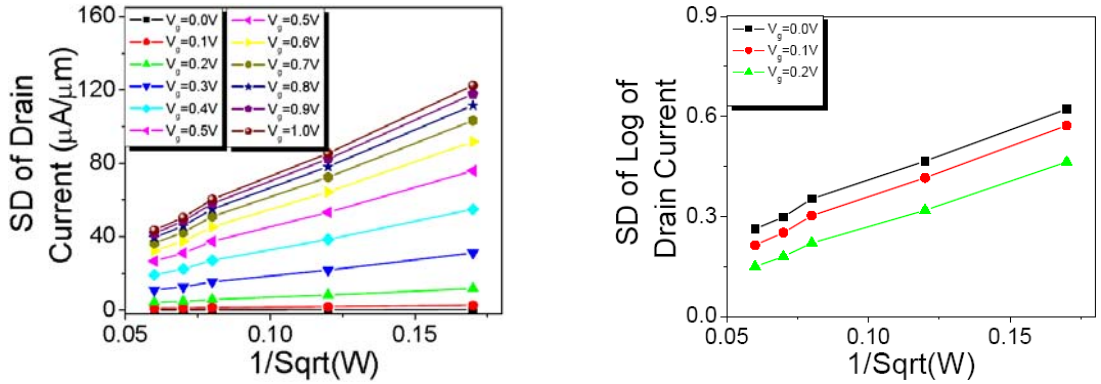


Figure 6.3: Standard deviation of the drain current for different values of the gate voltages in different width set devices (The x-axis dimension is $nm^{-0.5}$).

6.1.2 Full Statistical Atomistic Simulation Approach

Simulations, accounting for combined sources of variability: RDD, LER and PGG, are now carried out for devices of 70, 140, 210 and 280nm, and compared with the results for a 35nm square device. 1000 devices are simulated by the Glasgow University atomistic simulator. Again, simple statistical theory predicts that if the correlation lengths of any of the atomistic variability effects are significantly smaller than 35nm, then the average drain currents (per μm) should be identical to those of the base device within statistical errors of $\sqrt{1000}/1000 \approx 3\%$, and the statistical variability of the drain current should drop by a factor of $\sqrt{N}$, where N is the multiple of the base device gate width.

The resultant I_d-V_g curves for high drain bias (V_d=*1V*) are shown in Figure 6.4. The gate width of each device (*W1*, *W2*, etc.) is noted in multiples of 35nm. Red curves show the

equivalent curve for a uniform/idealized device, while the green curve is the statistical average of the currents in the statistical set.

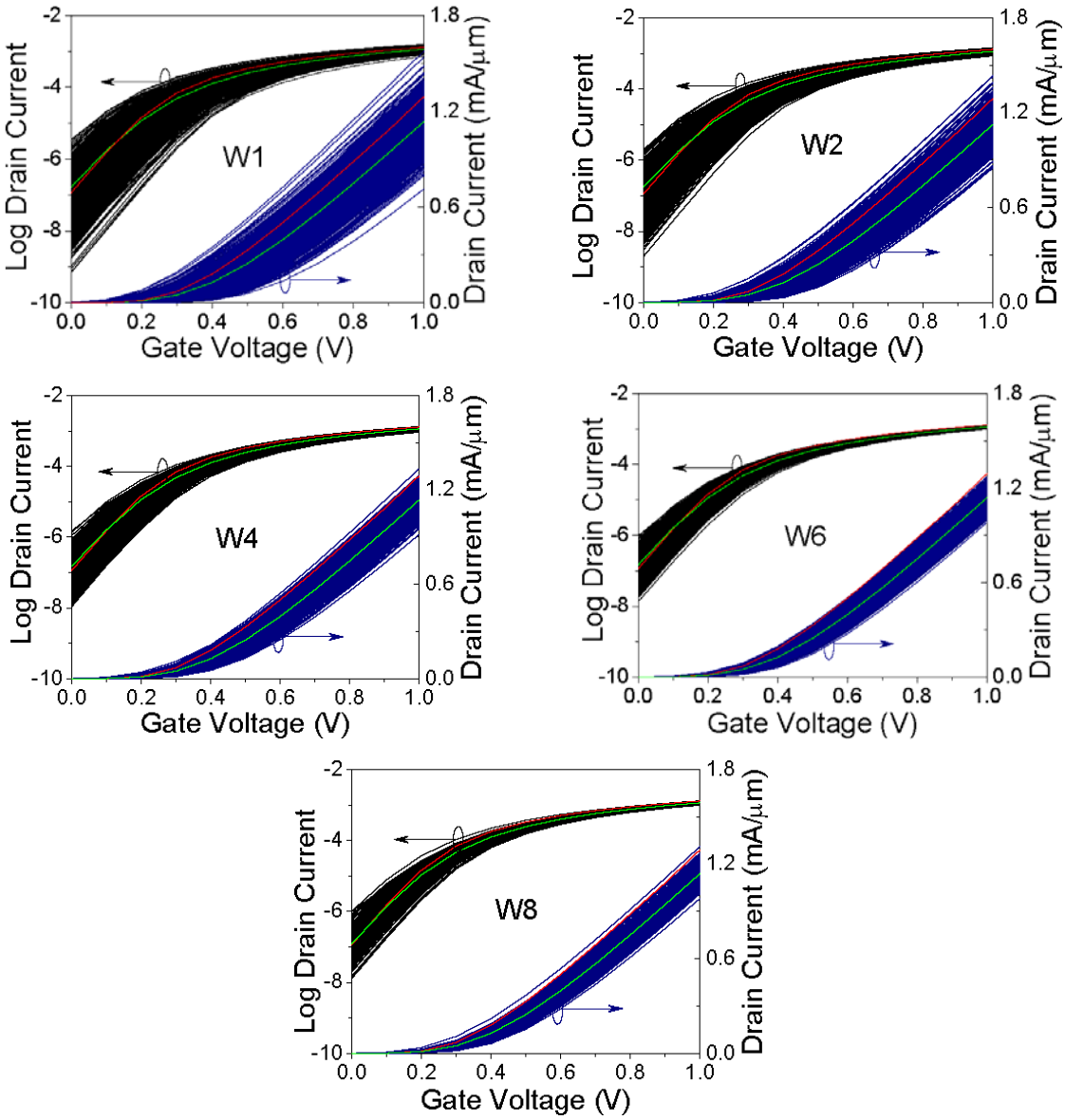


Figure 6.4: Variability in the I_d-$V_g$characteristics of a statistical sample of 1000 microscopically different 35nm n-MOSFETs whose gate widths are multiples of 35nm ($V_D = 1V$). Red curves described idealized devices. Green curves are the average currents for each statistical set.

The average of drain currents per μm gate width, for different device widths, are collected in Table 6.3 and plotted in Figure 6.5. While the averages are close together with less than 3% fluctuations in above threshold, there is 9% fluctuation in their values for

subthreshold region. The higher fluctuation in subthreshold region can be explained based on the correlation effects of drain current in wider devices as discussed in section 6.1.1.

Table 6.3: Statistical average of the drain current for different gate width n-MOSFETs

	Average of drain current (mA/μm)				
Width (nm) / Gate Voltage (V)	35	70	140	210	280
0.0	1.63E-4	1.78E-4	1.51E-4	1.47E-4	1.45E-4
0.1	1.71E-3	1.81E-3	1.61E-3	1.60E-3	1.57E-3
0.2	1.19E-2	1.21E-2	1.15E-2	1.15E-2	1.14E-2
0.3	4.85E-2	4.84E-2	4.79E-2	4.81E-2	4.80E-2
0.4	0.13	0.13	0.13	0.13	0.13
0.5	0.25	0.24	0.25	0.25	0.25
0.6	0.40	0.40	0.40	0.40	0.40
0.7	0.56	0.56	0.57	0.57	0.57
0.8	0.75	0.74	0.75	0.75	0.75
0.9	0.94	0.93	0.94	0.94	0.94
1.0	1.13	1.13	1.14	1.14	1.14

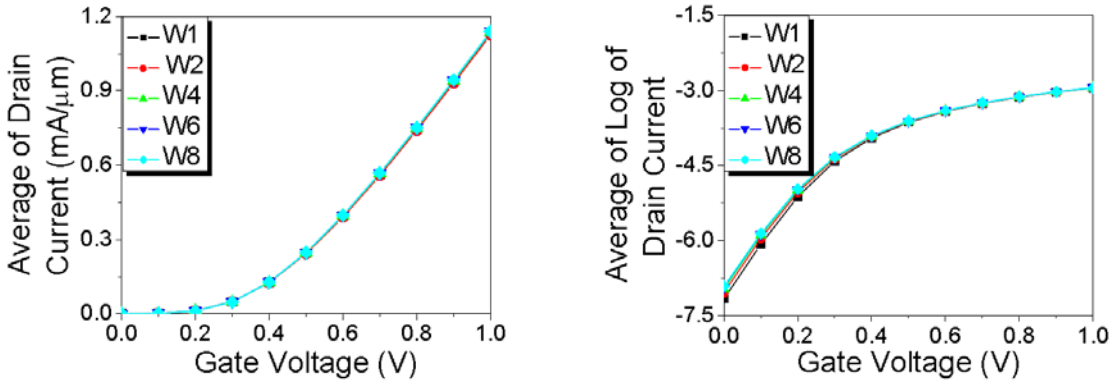


Figure 6.5: Average drain currents obtained from Figure 6.4.

Table 6.4: Standard deviation of the drain current for different gate width n-MOSFETs

	Standard deviation of drain current (A/μm)				
Width (nm) / Gate Voltage (V)	35 ($N=1$)	70 ($N=2$)	140 ($N=4$)	210 ($N=6$)	280 ($N=8$)
0.0	3.29E-7	2.18E-7	1.50E-7	1.14E-7	1.03E-7
0.1	2.55E-6	1.76E-6	1.22E-6	9.57E-7	8.37E-7
0.2	1.17E-5	8.35E-6	5.86E-6	4.70E-6	4.10E-6
0.3	3.11E-5	2.24E-5	1.58E-5	1.28E-5	1.12E-5
0.4	5.50E-5	3.97E-5	2.78E-5	2.26E-5	1.7E-5
0.5	7.59E-5	5.46E-5	3.81E-5	3.10E-5	2.68E-5
0.6	9.18E-5	6.58E-5	4.58E-5	3.72E-5	3.21E-5
0.7	1.03E-4	7.39E-5	5.13E-5	4.16E-5	3.58E-5
0.8	1.12E-4	7.98E-5	5.54E-5	4.49E-5	3.85E-5
0.9	1.18E-4	8.42E-5	5.84E-5	4.72E-5	4.04E-5
1.0	1.22E-4	8.76E-5	6.07E-5	4.91E-5	4.19E-5

Table 6.4 presents the standard deviations of the drain currents at different values of the gate voltage for different width devices. The results are in agreement with $\sqrt{N}$ falloff decay rule as discussed in section 6.1.1, particularly above threshold. Figure 6.6 illustrates the trend of standard deviations of drain currents as a function of $1/\sqrt{W}$ for different gate bias points. As evident from this figure, the trend is almost linear for the standard deviation of drain current versus $1/\sqrt{W}$ for each gate bias point. Since the variation of drain current in the subthreshold region (V_g = *0, 0.1* and *0.2V*) is not obvious from this figure, the standard deviations of logarithmic scale has been used to construct a more clear figure showing the trend for subthreshold region.

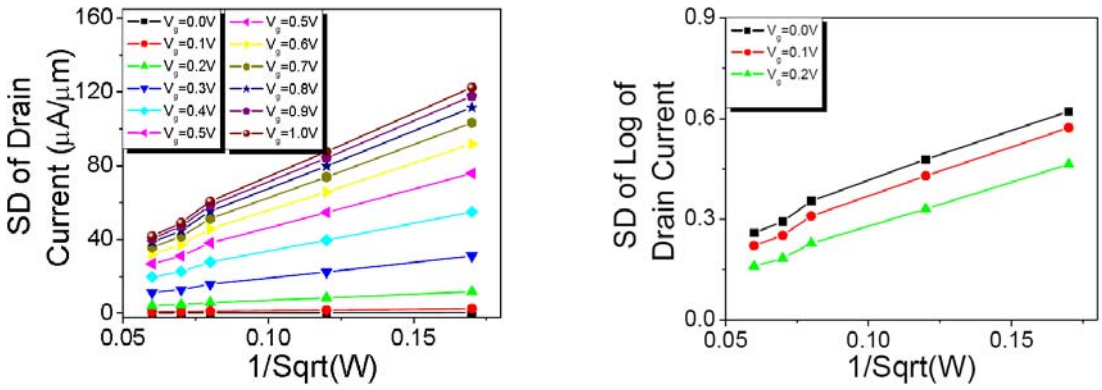


Figure 6.6: Standard deviation of the drain current for different values of the gate voltages in different width set devices (The x-axis dimension is $nm^{-0.5}$).

6.1.3 MOSFET Figures of Merit: Comparison of Parallel Component Approach and Full Atomistic Simulation

Having considered the fundamental I_d-V_g characteristics of devices with varying width, the device figures of merit of industrial importance which are typically extracted from these characteristics are now considered. The extraction of on-current I_{on}, off-current I_{off}, threshold voltage V_{th} and DIBL, averages and standard deviations are carried out for I_d-V_g characteristics associated with different width devices, using both the parallel component and full statistical simulation approaches. These results are laid out in Figure 6.7. As expected, the average values of most of these figures of merit are almost constant with device width, with the standard deviations decreasing at approximately $\sqrt{N}$ and both of the approaches are very close together. For I_{off}, the trend of the mean values versus width is almost constant with 5% increase of the average for the widest devices compared with basic width devices using parallel component approach. This increase will be reduced to 3% using atomistic simulation results and they are consistent with what the discussion related to Figures 6.2 and 6.5. The errors remain the same for the low drain bias results. There is a monotonic decreasing trend in the SD of I_{off} as a function of channel width. The results of the parallel component approach are within 1% tolerance of atomistic simulation results in high drain bias while there is more error about 4% at low drain bias conditions.

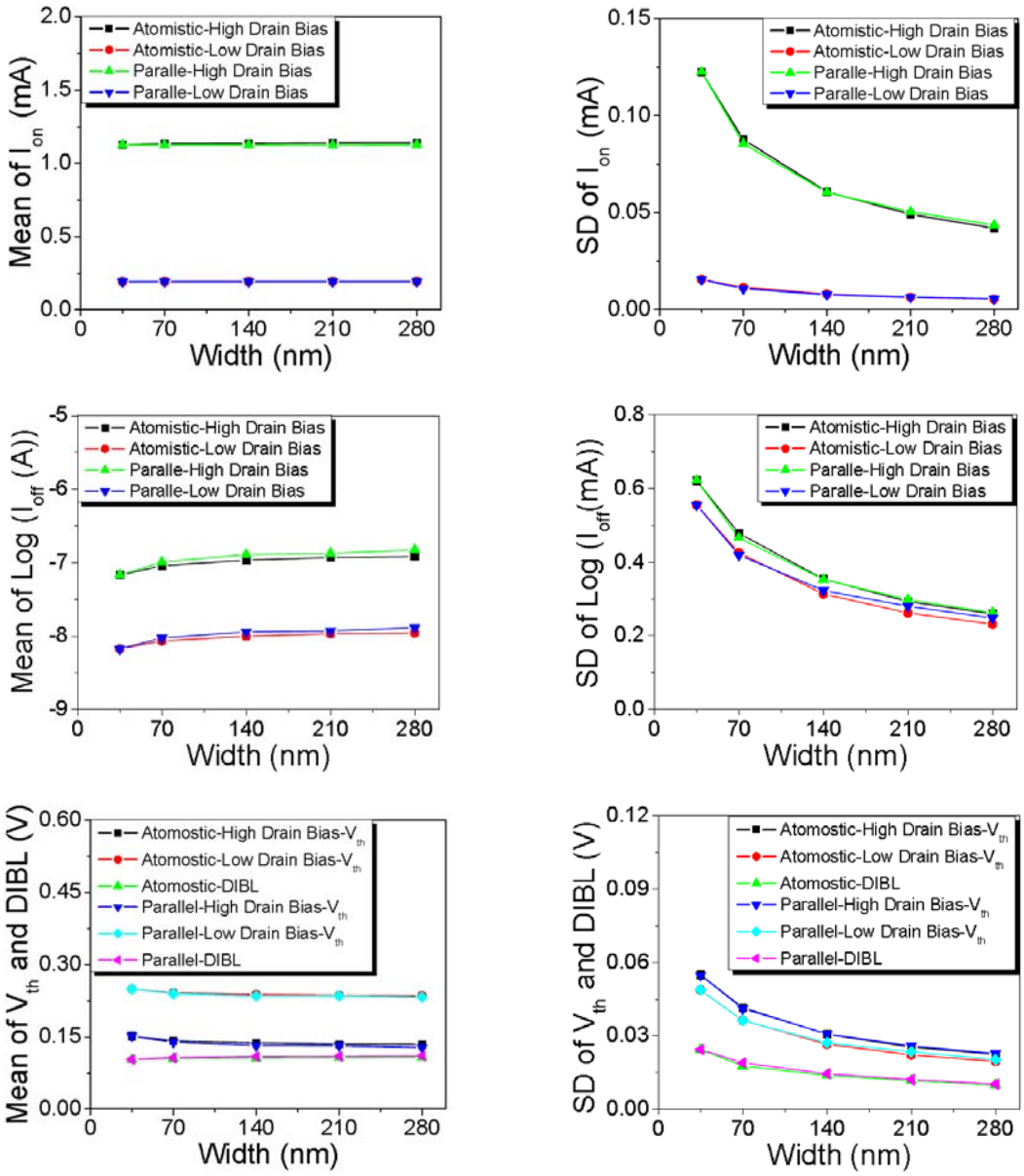


Figure 6.7: Trend of mean and standard deviation of figures of merit versus width.

For V_{th} and *DIBL*, the constant trend of the average versus width and the decreasing trend of the SD versus width is evident from Figure 6.7 and the error introduced by using parallel component approach remains less than 1% in respect to atomistic simulation results. This is due to the fact that LER pattern is discontinuous in parallel component approach but, full atomistic simulations provides continuous LER pattern in simulations.

Figure 6.8 shows the standard deviation of threshold voltage as a function of $1/\sqrt{W}$. This is called a Pelgrom plot [141]. The extrapolation of the standard deviation for very wide devices results in zero variability as expected.

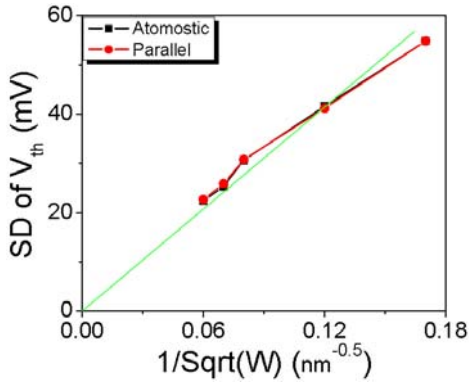


Figure 6.8: Trend of standard deviation of V_{th} versus $1/\sqrt{W}$

6.1.4Compact Model Parameter Extraction as a Function of Device Width

Finally, we consider the extraction of variability aware compact model parameters (as discussed in chapter 4) and how those compact model parameters are dependent on device width when adopting both the parallel component and full statistical extraction approaches. From the results above, we would expect that the parameters associated with the above threshold behaviour will show near constant average values and $\sqrt{N}$ standard deviation reductions as a function of device width, whilst parameters associated with the subthreshold performance (or at worst parameters dependent on second order subthreshold effects) to vary from the predictions of the simple statistical theory.

Statistical extraction of parameters for BSIM compact models are carried out using the Aurora script. The supplied data to the script are the I_d-V_g sets for 35, 70, 140, 210 and 280nm width devices described above. Figure 6.9 shows that the average RMS error in performing these statistical extractions remains less than 3%. Therefore the extraction error is less than, or of the same order as, the statistical error associated with the size of our

statistical ensemble. However, there is a systematic trend of slight increase in the error as width increases. This increase can be understood based on the impact of initial conditions in the statistical parameter extraction. The uniform device has been used to set the initial conditions for all transistors with different width prior to the statistical parameter extraction. By looking into Figures 6.1 and 6.4, the uniform 35nm transistor shown by red colour is located close to the middle and inside of I_d-V_g spread in linear scale of W1 devices. As the width increases, the spread moves downward or equivalently the uniform device moves to the upper edge of the spread, as can be seen for the W8 devices in the same figure. As a result, the uniform transistor will better represent the most of devices in W1 compared with W8. Hence, a gradual increase of the error while width increases will be expected.

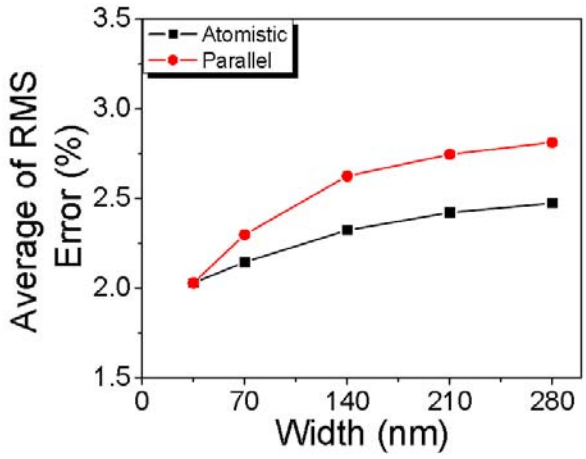


Figure 6.9: The average value of the RMS error between statistically extracted BSIM compact models.

The behaviour of the mean and standard deviation of six critical atomistic variability aware compact model parameters are plotted in Figures 6.10 and 6.11 for parallel component and full statistical simulation approaches, respectively. It can be seen that the mean values of almost all the parameters are almost independent of the device width, as expected, and almost all the standard deviations have a monotonically decreasing behaviour. Calculation shows that for almost all of the parameters, the difference between the parallel component and full atomistic simulation approaches is less than 10%.

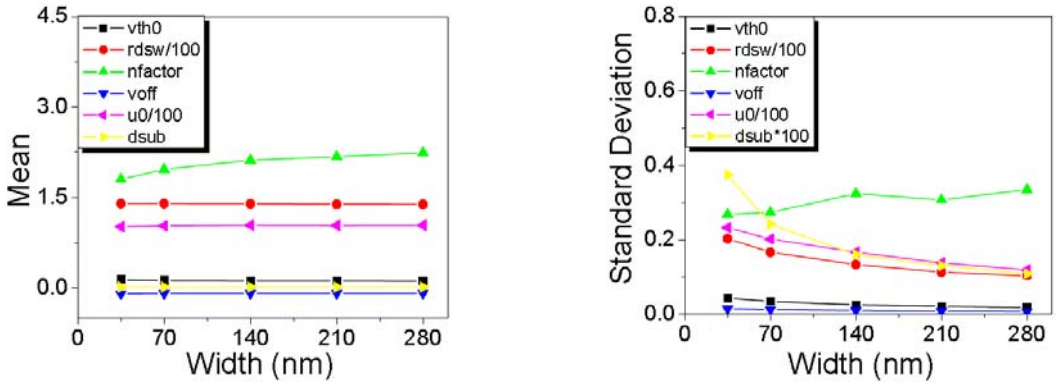


Figure 6.10: Value of different parameters of statistical parallel BSIM compact models versus width, mean values (left), standard deviations (right).

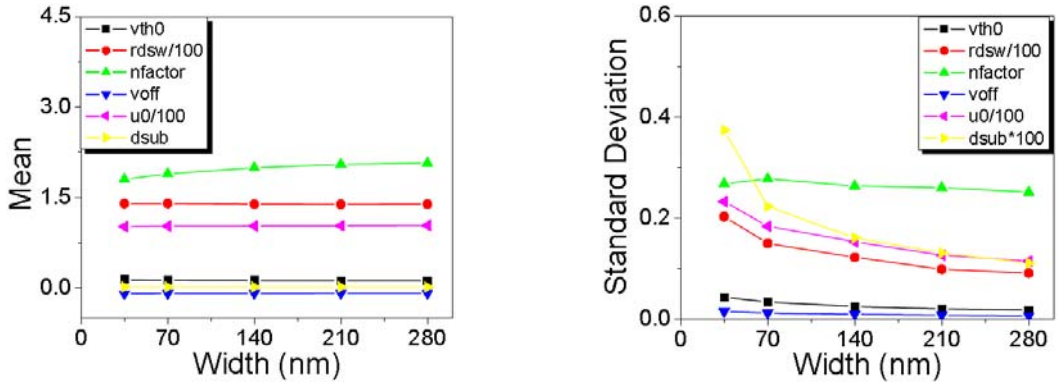


Figure 6.11: Value of different parameters of statistical atomistic BSIM compact models versus width, mean values (left), standard deviations (right).

However, the parameter *NFACTOR* shows a varying mean value as a function of device width. Therefore it is the only parameter which shows a non-decreasing standard deviation in both approaches. It also shows a maximum difference of 32% in its standard deviation when comparing the parallel component approach with full atomistic results. The problem can be investigated by plotting the distribution of this parameter versus width as shown in Figure 6.12. It is clear that this parameter has an increasing skew versus width thus the mean and standard deviations are not sufficient moments to completely characterise the statistical distribution of *NFACTOR*.

Table 6.5 presents the mean and standard deviation of statistical parameters obtained from atomistic and parallel component approaches in different width devices.

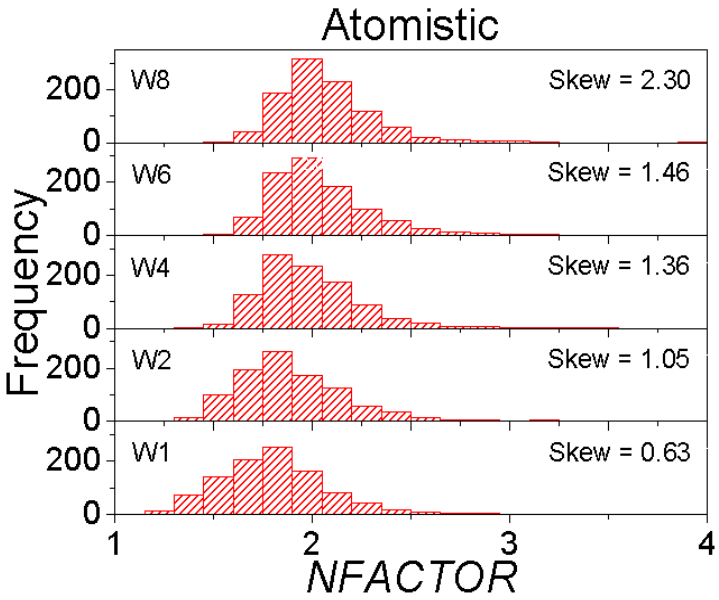


Figure 6.12: The impact of width of device on *NFACTOR* parameter of statistical compact models.

Table 6.5: Mean and standard deviation of statistical parameters obtained from atomistic and parallel component approaches in different width devices.

Parameters/Width	Mean/SD	Atomistic		Parallel	
		Mean	Standard Deviation	Mean	Standard Deviation
VTH0 (mV)	35nm	138.19	43.79	138.19	43.79
	70nm	128.12	33.91	124.33	34.52
	140nm	123.67	25.17	118.7	25.56
	210nm	121.92	20.34	118.9	21.7
	280nm	121.25	17.86	116.16	18.4
U0 (m^2/Vs)	35nm	101.81	23.29	101.81	23.29
	70nm	102.70	18.395	103.24	20.20
	140nm	102.98	15.332	104.25	16.66
	210nm	103.23	12.672	103.85	13.82
	280nm	103.77	11.48	104.08	11.92

VOFF (mV)	35nm	-93.64	15.17	-93.64	15.17
	70nm	-89.23	12.17	-87.07	13.07
	140nm	-87.63	9.84	-84.71	10.06
	210nm	-87.28	8.04	-84.75	8.32
	280nm	-86.94	6.92	-84.62	7.46
RDSW (ohm)	35nm	139.83	20.32	139.83	20.32
	70nm	139.97	15.00	139.88	16.65
	140nm	138.88	12.24	139.33	13.36
	210nm	138.67	9.90	138.86	11.34
	280nm	138.89	9.13	138.68	10.36
NFACTOR	35nm	1.80	0.27	1.80	0.27
	70nm	1.89	0.28	1.97	0.27
	140nm	1.99	0.26	2.12	0.32
	210nm	2.05	0.26	2.17	0.31
	280nm	2.07	0.25	2.24	0.33
DSUB	35nm	0.0169	0.0037	0.0169	0.0037
	70nm	0.0162	0.0022	0.0161	0.0024
	140nm	0.0159	0.0016	0.0156	0.0016
	210nm	0.0158	0.0013	0.0155	0.0013
	280nm	0.0157	0.0011	0.0154	0.0011

Tables 6.6 to 6.9 provide the correlation coefficients between the compact model parameters extracted for both the fully atomistic and the parallel component approaches. The correlation coefficient of base width devices was illustrated in Figure 5.3. The correlation coefficients (ρ) of statistically extracted parameters are calculated based on Equation (5.1). The main message of these tables is that the significant correlations between parameters that are present in the base 35nm results are retained in longer width devices, with only small fluctuations (within a few percent) compared to the correlation

between 35nm devices. As expected, parallel component approach results in correlation coefficients which are very close to those of the basic width transistors.

Table 6.6: The correlation coefficient between parameters of devices with 70nm gate width; Down-left: Atomistic compact models; Up-right: Parallel compact models.

VTH0	-0.14	-0.14	0.01	-0.56	0.11
-0.22	***U0***	0.68	0.75	0.25	-0.69
-0.14	0.66	***VOFF***	0.65	0.20	-0.56
0.03	0.68	0.63	***RDSW***	-0.17	-0.55
-0.60	0.32	0.22	-0.23	***NFACTOR***	-0.22
0.16	-0.66	-0.47	-0.45	-0.30	***DSUB***

Table 6.7: The correlation coefficient between parameters of devices with 140nm gate width; Down-left: Atomistic compact models; Up-right: Parallel compact models.

VTH0	-0.02	-0.04	0.19	-0.58	0.08
-0.13	***U0***	0.70	0.73	0.10	-0.74
-0.15	0.69	***VOFF***	0.58	0.06	-0.59
0.14	0.73	0.58	***RDSW***	-0.40	-0.49
-0.61	0.19	0.20	-0.31	***NFACTOR***	-0.26
0.15	-0.71	-0.59	-0.48	-0.31	***DSUB***

Table 6.8: The correlation coefficient between parameters of devices with 210nm gate width; Down-left: Atomistic compact models; Up-right: Parallel compact models.

VTH0	0.06	0.06	0.30	-0.58	0.00
-0.14	***U0***	0.71	0.73	0.02	-0.74
-0.10	0.69	***VOFF***	0.61	-0.07	-0.56
0.10	0.75	0.62	***RDSW***	-0.48	-0.47
-0.57	0.20	0.13	-0.30	***NFACTOR***	-0.21
0.11	-0.74	-0.60	-0.52	-0.29	***DSUB***

Table 6.9: The correlation coefficient between parameters of devices with 280nm gate width; Down-left: Atomistic compact models; Up-right: Parallel compact models.

VTH0	0.07	0.10	0.32	-0.54	0.00
-0.12	***U0***	0.68	0.73	0.00	-0.76
-0.10	0.71	***VOFF***	0.64	-0.20	-0.50
0.10	0.76	0.62	***RDSW***	-0.55	-0.43
-0.55	0.17	0.08	-0.34	***NFACTOR***	-0.25
0.12	-0.75	-0.61	-0.53	-0.28	***DSUB***

6.2 Impact of Gate Length on Variability

Two forms of MOSFET gate length alteration are typical. Firstly relatively small variations about the design length (in this case of 35nm) due to process fluctuations between wafers, or between dies within a wafer. Such variations are likely to degrade the

operation of the transistor, either by making it leakier and less controlled (so called short channel effects) at smaller gate lengths, or by lowering the drive current away from the designed drive current at longer gate lengths. The second gate length alteration is often made by circuit designers on purpose. Significantly lengthening the gate length of the transistor will lower its performance, but also significantly reduce leakage in circuits which are not time critical, thus reducing overall circuit power draw. We investigate the effects of both types of length alterations on atomistic variability.

6.2.1. Statistical Atomistic Simulation

Atomistic simulation of different gate length devices, with gate lengths of 30, 35 and 40nm has been carried out using Glasgow University atomistic simulator for 1000 samples at each length. The simulations include combined sources of variability: RDD, LER and PGG. In order to highlight the impact of length on the statistical variability, a width of 35nm is considered for all of these devices.

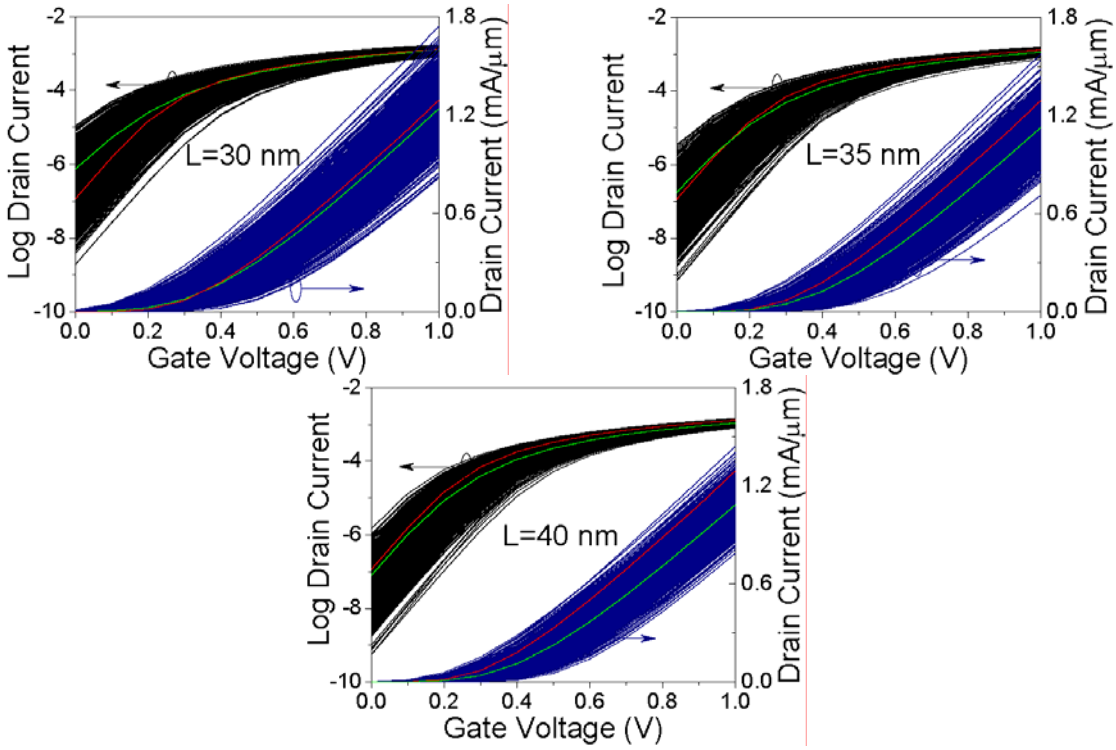


Figure 6.13: Variability in the I_d-$V_g$characteristics of a statistical sample of 1000 microscopically different length n-MOSFETs. $V_D = 1V$. Red curves described idealised devices. Green curves are the average currents for each statistical set.

The I_d-V_g characteristics obtained from atomistic simulation at high drain bias (V_d =1V) are shown in Figure 6.13. Red curves show the I_d-V_g response of an idealised/uniform device whilst the green curves are the statistical averages of the currents for each statistical set. It can be seen that the average on- and off-currents for 35 and 40nm devices are similar, and whilst variations in on-current seem broadly similar, there is slightly less leakage variation in longer devices. 30nm devices show substantially greater off-current and substantially more variation in on-current, disadvantages which outweight slightly improved drive current for these devices. The average drain currents for different values of the gate voltage in different width devices are illustrated in Figure 6.14. These results confirm the introductory comments to section 6.2, that the leakage / off-current in 35nm devices significantly degrades as we move to shorter gate lengths than the transistor was designed for.

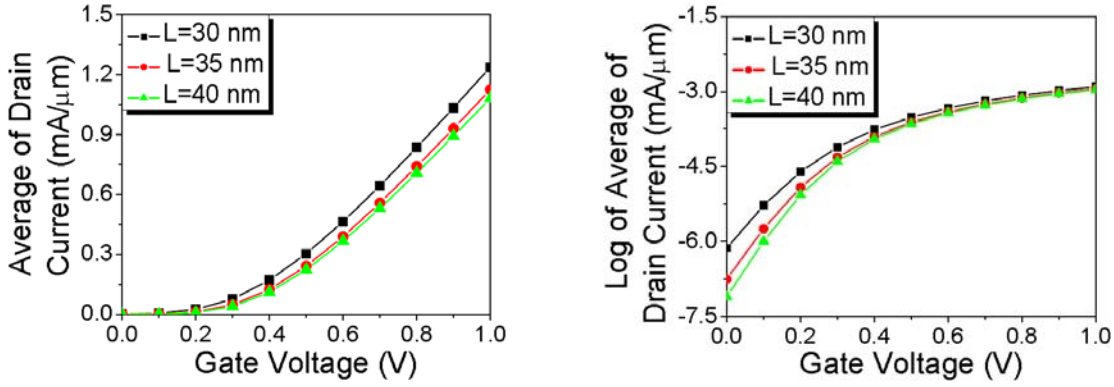


Figure 6.14: Collected average drain currents from Figure 6.13.

6.2.2. Trends in Figures of Merit versus Gate Length

Atomistic simulation of different gate length devices was extended to L_g = 50, 70 and 100nm gate length devices to study the impact of gate length on MOSFET figures of merit. The uniform device electrical figures of merit were extracted and are shown in Figure 6.15. Drive current monotonically increases as gate length shortens, while the leakage current drastically increases at around 40nm. This is the point at which the threshold voltage peaks due to the impact of halo doping and associated reverse short channel effects [142,143,144,145,146,147,148] before rolling off sharply below 35nm. In other words, the

threshold enhancement which appears around 40nm in Figure 6.15 is the result of non-uniform channel dopant distribution along the channel region [101].

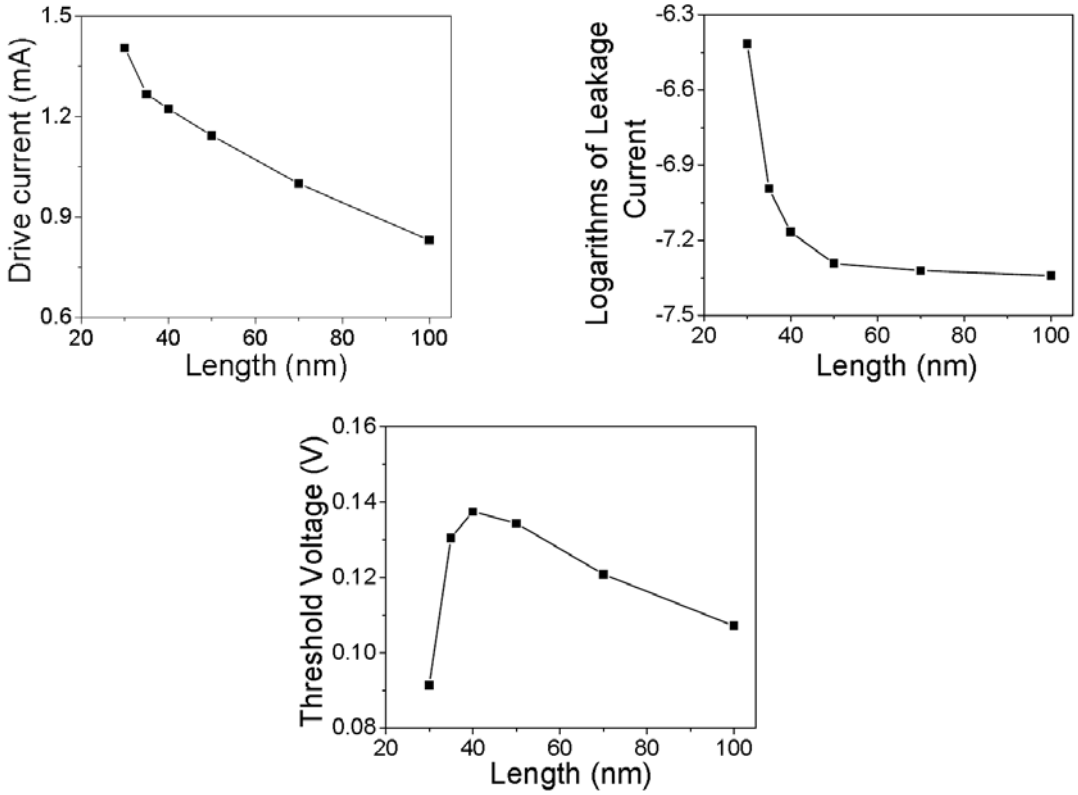


Figure 6.15: Figures of merit of uniform devices with different length.

The electrical figures of merit for 1000 statistical devices have been extracted and the statistical trends of the mean and standard deviation values are shown in the Figure 6.16 and Table 6.10. All the values are extracted at high drain bias conditions (V_d=1V). Theoretically, assuming that the most significant source of fluctuations in bulk MOSFETs is RDD, the total channel charge is proportional to L_g and the variation in channel charge is proportional to $\sqrt{L_g}$. Thus, a first order dependence of I_d on $\sqrt{L_g}$ is expected. In Figure 6.16, plotting the values of parameter variation (red traces) against $1/\sqrt{L_g}$ clearly shows deviation from a $\sqrt{L_g}$ dependence as deviation from a straight line on the graph. To first order, the variations of I_{on}, I_{off} and V_{th} indeed follow this trend, with I_{on} and V_{th} deviating from it somewhat at the longest and shortest channel lengths.

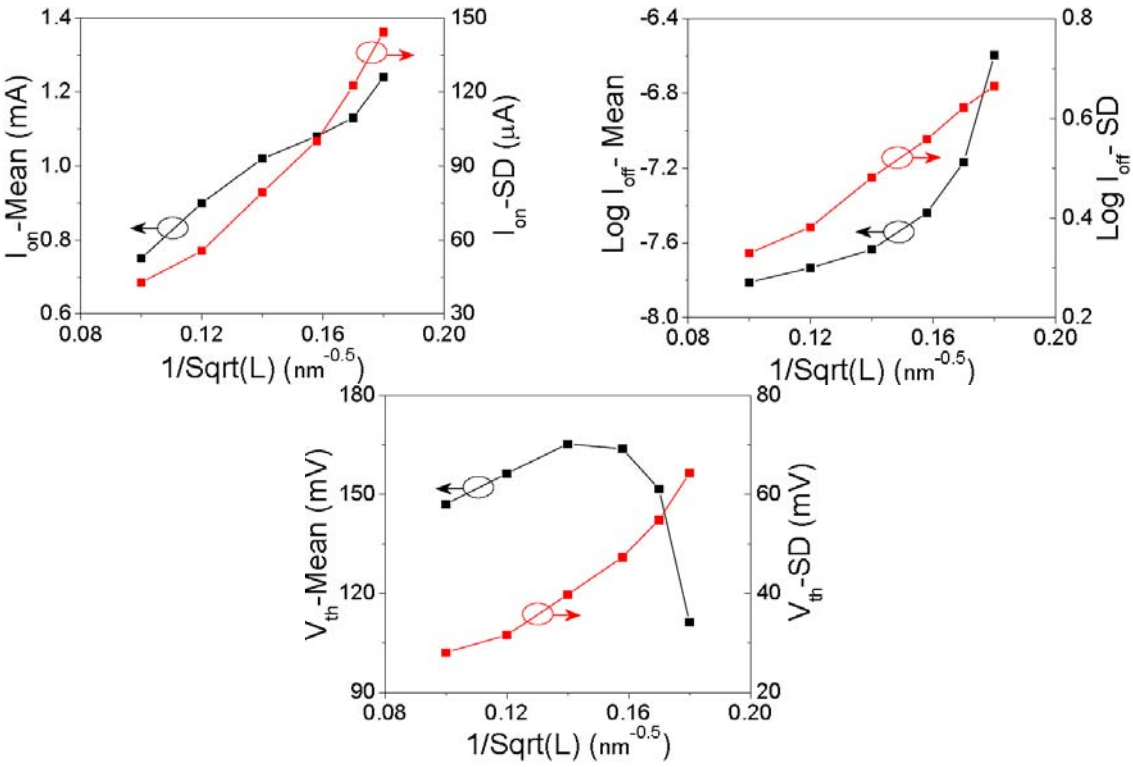


Figure 6.16: Trend of mean and standard deviation of figures of merit versus width.

Table 6.10 shows the statistical information related to the figures of merit of 1000 devices with different length.

Table 6.10: Means and standard deviations of MOSFET figures of merit

Length of(nm) Figures Merit	I_{on} (mA)		Log (I_{off}(A))		V_{th} (mV)	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
30	1.24	0.14	-6.59	0.66	111.26	64.28
35	1.13	0.12	-7.17	0.62	151.62	54.85
40	1.08	0.10	-7.44	0.56	163.74	47.34
50	1.02	0.08	-7.64	0.48	165.26	39.75
70	0.90	0.06	-7.73	0.38	156.27	31.58
100	0.75	0.04	-7.81	0.33	147.05	28.12

6.2.3. Parameter Extraction Trend versus Length

As described in chapter 3, a group extraction methodology is used in this study. Therefore, to investigate the effect of gate length on the extraction of compact model parameters, uniform devices of gate length L_g = 30, 35 and 40nm are used. Figure 6.17 shows the statistical trend of three typical parameters in close range of different length devices.

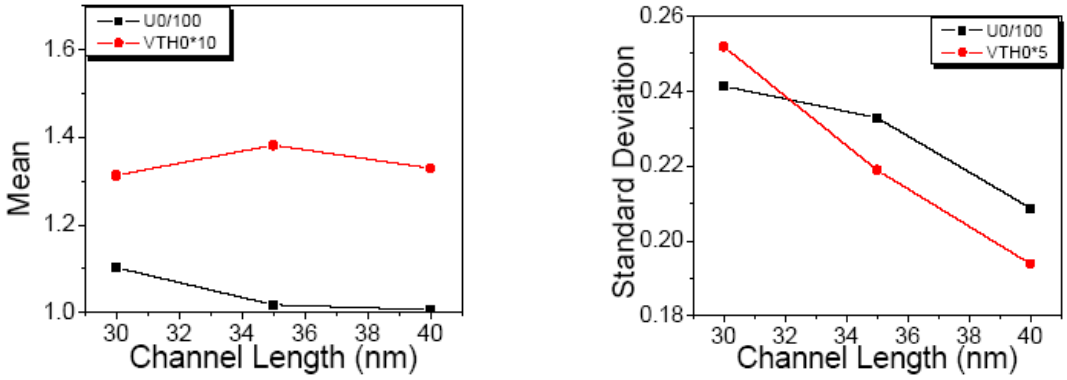


Figure 6.17: Value of different parameters of statistical parallel in different length BSIM compact models versus width, mean values (left), standard deviations (right).

These parameters have a monotonically decreasing behaviour in the standard deviation and a flat behaviour in the mean as expected.

6.3 Summary

By increasing the width of MOSFETs, the statistical variability in drain current, transistor figures of merit and compact model parameters decrease, typically as the square root of the device width. Figures of merit which are obtained from the fully atomistic and parallel component approaches are typically within 3% of each other with exceptions being found in those parameters representing subthreshold device behaviour, or second order effects in subthreshold. Therefore by using the parallel component approach (including interpolation between devices of fractional width), I_d-V_g characteristics of single width

devices *can* be used to generate statistical compact model parameters, and to build appropriate libraries of SPICE model cards, that cover a wide range of device widths. Variability of I_d-V_g characteristics as a function of device gate length, both due to process variation and circuit design choice was also investigated and increasing gate length again found to decrease variability, typically as the square root of the device length.

Chapter 7

Conclusions

Statistical variability introduced by the discreteness of charge and granularity of matter has become a major obstacle in scaling of bulk MOSFETs. Accurate statistical compact models are needed to achieve reliable variability aware design. In this thesis, a statistical BSIM compact model strategy was developed to capture the impact of variability at the early design stage of integrated circuits. The BSIM statistical compact model strategy was based on statistical extraction of a limited subset of parameters from the full compact model to reproduce the shape of I_d-V_g characteristics for microscopically different MOSFETs in presence of various sources of statistical variability including RDD, LER and PGG. Predictive 'atomistic' simulations were used to provide benchmark data for developing the statistical compact model framework.

7.1 Summary of Results

In Chapter 2, the statistical variability was introduced as a major challenge in the scaling of deca-nano meter bulk MOSFETs. The major sources of statistical variability were then introduced including RDD, LER, PGG, OTF and high-k granularity. Three different methods for physical simulation of statistical variability were introduced including DD, MC and QM techniques. DD is the most efficient method from

computational point of view and therefore, is a good choice for predictive atomistic simulations of statistical variability, particularly if DG quantum corrections are included. However, DD underestimates the drive current variations although it provides very high accuracy in subthreshold region of transistor operation. The results of physical simulation of statistical variability for a 35nm gate length MOSFET were presented including a set of statistical I_d-V_g characteristics and it was observed that the leakage current variation had a spread of almost 3 orders of magnitude while the drive current variation showed a spread of almost 50% of its mean. This level of variations cannot be ignored in circuit design and to verify this fact, simulation of a CMOS inverter were performed under influence of statistical variability sources. Finally, the existing statistical MOSFET models from the literature were reviewed and their main limitations and deficiencies were outlined.

In chapter 3, an overview of the BSIM4 compact model equations was presented. The operation and the design of 35nm gate length template n- and p-channel MOSFETs were reviewed. The development of parameter extraction and optimization strategy of a template MOSFET was among the main contributions of this chapter. The input data includes I_d-V_g characteristics at high and low drain bias points and I_d-V_d characteristics. The BSIM parameters which need to be extracted at each stage of the extraction strategy were explored and the final RMS error of the fitting in both I_d-V_g and I_d-V_d characteristics were presented in detail for both n- and p-channel MOSFETs. The RMS errors of parameter extraction remain less than 1.3% for I_d-V_g and less than 3.6% for I_d-V_d data fitting of 30, 35 and 40nm gate length MOSFETs utilizing a group extraction strategy.

In chapter 4, predictive ‘atomistic’ simulations were performed to obtain two sets of statistical I_d-V_g characteristics at low and high drain bias conditions for macroscopically identical but microscopically different p- and n-channel 35nm MOSFETs under the combined influence of the relevant sources of statistical variability. Important contribution of this chapter is the optimum set of statistical BSIM parameters based on the first order sensitivity analysis of drain current to capture the impact of statistical variability on transistor characteristics. The impact of different parameter set size on the accuracy of statistical compact model was investigated for the first time and the trend of the RMS error

mean and standard deviation as function of number of parameters in each statistical set was studied. The MOSFET electrical figures of merit were simulated using statistical compact model library and the mean and standard deviation of distributions showed less than 0.9% absolute error in respect to corresponding distributions obtained from 'atomistic' simulations. Scatter plots were used to illustrate the correlations between MOSFET figures of merit. The highest correlation coefficient links V_{th} and I_{off} and the lowest links DIBL and SS. It was clearly demonstrated that the correlation of figures of merits are well maintained in the directly extracted compact models. Moreover, the extracted statistical compact models were used to simulate a CMOS inverter in 35nm technology node. The delay and dissipated energy of the inverter were simulated as function of number of parameters in each statistical set. It was observed that the trend of reducing the error in the mean and standard deviation of the delay settles for more than 5-parameters set. Using 5 or 6-parameter gives less than 1% error in the mean and standard deviation of delay and energy dissipation of the inverter in respect to the most accurate results obtained using 7-parameter set. This means that with respect to circuit design, the use of more than 5-parameter provides diminishing return. A detailed study on the impact of the number of statistical parameters on the mean and standard deviation of the n- and p-MOSFET electrical figures of merit was also reported in this chapter.

Chapter 5 introduces statistical parameter generation strategies. In the case of the direct parameter extraction approach, larger statistical parameter set generally produce better distributions of the statistical compact model errors. In respect to the different statistical parameter generation approaches, a relatively small parameter set size is desirable because it reduces the complexities associated with preserving the parameter correlations during the statistical parameter generation process. The statistical properties, including the distributions and correlations of the directly extracted parameter set play an important role ensuring the statistical accuracy of the generated compact model. The results indicate that the naïve approach, which generates statistical compact model parameters assuming independent normal distribution for each extracted parameter, will produce considerable error in circuit simulation. The PCA approach to the generation of statistical compact model parameters is limited by its assumption of normally distributed parameters. The

accuracy of the naïve and the PCA approaches was assessed in circuit simulation using a CMOS inverter. The impact of different input rise/fall times and different output load capacitances on the distribution of inverter delay and energy was simulated using both extracted and generated statistical compact models. The simulation results indicate that the statistical variability will have the largest impact on the spread of delay and energy of the inverter, for the longest input rise/fall time and when there is no load in the output. It was found that the statistical compact models generated with the naïve and the PCA approaches result in low error (less than 4%) in the mean delay and energy distributions in respect of the corresponding values obtained using directly extracted statistical compact models. The standard deviations of the delay and energy for the naïve and the PCA have larger error in respect to the direct approach. The accurate treatment of higher order moments of the statistical compact model distribution in the proposed NPM approach not only maintains the correlations between the generated statistical compact model parameters, but also accurately captures the tails and the nonlinear shape of their distributions. The simulations have shown that using NPM generated compact models produces less than 3% error in the mean and the standard deviation of the inverter delay and energy in respect to the corresponding values obtained from the direct approach.

In chapter 6, statistical compact modeling strategies for transistors with different geometry were investigated. Due to strong geometry dependence of the statistical variability the focus was on the impact of transistor channel length and width on the electrical characteristics variability. Two methods were used to study the impact of the channel width on the variability in the transistor parameters. In the first approach atomistic simulations were performed for a set of microscopically different transistors with different width while they are subject to combined sources of statistical variability. In the second approach which is computationally more efficient, wider devices were generated from the atomistic simulation results for a square transistor by randomly connecting them in parallel in SPICE simulations. Both approaches were compared based on BSIM statistical parameter trends and MOSFET electrical figures of merit versus device geometry. As expected by increasing width and length, a decrease in the variability was observed. The average values of MOSFET figures of merit are almost constant (within 5% fluctuations)

with device width, but the standard deviations decrease at approximately $\sqrt{N}$ rate, where N is the multiple of the square base device gate width. The parallel component approach is within 1% of the benchmark for I_{on} and within 4% error for I_{off}. Since I_{on} is the most important figure of merit when calculating circuit speed, this approach will reproduce accurately the circuit speeds in simulations. The full atomistic simulation of 1000 samples of different width transistors showed that the average drain current remains unchanged above threshold, within 3% fluctuations due to finite statistical sample errors. However, the discrepancies are larger in the subthreshold region, up to 9%. This is due to the fact that non-square 35nm gate length transistors exhibit correlated effects across their widths because charge percolation from source to drain is least affected by screening in the channel at subthreshold region and the effects of electrostatics on drain current in subthreshold are exponential in nature, amplifying this effect.

7.2 Future Work

There are several areas where the work presented in this thesis can be extended. First, the BSIM statistical compact model strategy developed in this thesis can be applied to the next generation of scaled planar bulk CMOS devices. Second, the sensitivity analysis of parameters performed in chapter 4 to identify responsible parameters to capture the impact of statistical variability can be extended and applied to the next generation CMOS technology and larger statistical parameter sets may become necessary to reproduce the impacts of statistical variability in truly nano-CMOS regime. Third, application of a reduced set of statistical parameters can be investigated for particular applications. For example, in circuit designs where the transistors are used in sub-threshold or above-threshold regions, using a sub-set of the statistical parameters will be sufficient.

Another area in which the work in this thesis can be extended is the development of BSIM statistical compact models for new device architectures such as SOI and multi-gate devices. This will help the designers to take into account the corresponding effects of statistical variability at the design stage. Proposing computationally efficient and accurate

parameter generation techniques is another interesting but challenging area in the context of statistical compact models.

In a longer term plan, since we have shown that it is possible to extract statistical information from a set of device parameters, and collect that statistical information through accurate parameter generation strategies, a collaboration with the industry can lead to development of higher level variability aware CAD tools, place and route tools for IC layout or synthesis tools to turn VHDL into transistors, which can make use of this statistical data for transistors directly.

Appendix I

Compact Model Parameters

A0	Coefficient of channel length dependence of bulk charge effect
A1	First non-saturation effect parameter
A2	Second non-saturation factor
AGS	Coefficient of V_{gs} dependence of bulk charge effect
B0	Bulk charge effect coefficient for channel width
B1	Bulk charge effect width offset
CDSC	Coupling capacitance between source/drain and channel,
CDSCB	Body-bias sensitivity of *CDSC*
CDSCD	Drain bias sensitivity of *CDSC*
CIT	interface trap capacitance
DELTA	Smoothing parameter
DLC	Channel length offset parameter for *CV* model
DSUB	Length dependent DIBL behavior
DTOX	Difference between *TOXE* and *TOXP*

DVT0	First coefficients of SCE on V_{th}
DVT0W	First coefficient of NWE on V_{th} for small channel length
DVT1	Second coefficients of SCE on V_{th}
DVT1W	Second coefficient of NWE on V_{th} for small channel length
DVT2	Body-bias coefficients of SCE on V_{th}
DVTP0	First coefficient of drain-induced V_{th} shift due to for long-channel pocket devices
DVTP1	Second coefficient of drain-induced V_{th} shift due to for long-channel pocket devices
DWB	substrate bias effects
DWC	Channel width offset parameter for *CV* model
DWG	gate bias effects
DWJ	Offset of the source/drain junction width
EPSROX	Gate Dielectric Constant Relative to Vacuum
ETA0	DIBL coefficient in sub-threshold region
ETAB	Body bias coefficient for the sub-threshold DIBL effect
EU	Exponent for mobility degradation
FPROUT	Effect of pocket implant on output resistance degradation
K1	First body bias coefficient
K2	Second body bias coefficient
K3	Narrow width coefficient
K3B	Body effect coefficient of *K3*
KETA	Body bias coefficient of bulk charge effect
LINT	Channel-length offset parameter

LL	Coefficient of length dependence for length offset
LLC	Coefficient of length dependence for *CV* channel length offset
LLN	Power of length dependence for length offset
LP	Mobility channel length exponential coefficient
LPE0	Lateral non-uniform doping parameter at $V_{bs} = 0$
LPEB	Lateral non-uniform doping effect on *K1*
LW	Coefficient of width dependence for length offset
LWC	Coefficient of width dependence for *CV* channel length offset
LWLC	Coefficient of length and width cross-term dependence for *CV* channel length offset
LWN	Power of width dependence for length offset
LWL	Coefficient of length and width cross-term dependence for *CV* channel length offset
mobMod	Flag of the effective mobility model
MINV	V_{gsteff} fitting parameter for moderate inversion condition
NFACTOR	Sub-threshold swing factor
NDEP	Doping concentration at the edge of the channel depletion layer at $V_{BS} = 0$
NGATE	*D*oping concentration in the gate
NSD	Doping concentration in the source/drain regions
PCLM	Channel length modulation parameter
PHIN	Non-uniform vertical doping effect on surface potential
PRT	*T*emperature coefficient for *RDSW*
PRWB	Body-bias dependence of the low-doped drain resistance

PRWG	Gate-bias dependence of the low-doped drain (LDD) resistance
PVAG	Gate bias dependence of Early voltage and e
rdsMod	Flag to select different source/drain resistance models
RDSW	Zero bias LDD resistance per unit width for *rdsmod=0*
RDSWMIN	LDD resistance per unit width at high V_{gs} and zero V_{bs} for *rdsmod=0*
TNOM	Nominal temperature
TOXE	Gate Oxide Thickness
TOXP	Physical Gate Oxide Thickness
U0	Low-field mobility
UA	Coefficient of first-order mobility degradation due to vertical field
UB	Coefficient of second-order mobility degradation due to vertical field
UC	Coefficient of mobility degradation due to body-bias effect
UP	Mobility channel length coefficient
VACLM	Early voltage
VASAT	Early voltage at $V_{ds}=V_{dsat}$
VFB	Flat-band voltage
VOFF	Offset voltage in subthreshold region for large *W* and *L*
VOFFL	Channel length dependence of *VOFF*
VTH0	Long-channel threshold voltage at $V_{BS} = 0$
VSAT	*S*aturation velocity at nominal temperature
W0	Narrow width parameter
WINT	Channel-width offset parameter
WL	Coefficient of length dependence for width offset
WLN	Power of length dependence for width offset

WR	Channel width dependence parameter of the LDD resistance
WW	Coefficient of width dependence for width offset
WWC	Coefficient of width dependence for *CV* channel width offset
WWL	Coefficient of length and width cross-term dependence for width offset
WWLC	Coefficient of length and width cross-term dependence for *CV* channel width offset
WWN	Power of width dependence of width offset
XJ	Source/drain junction depth
XL	Channel length offset
XW	Channel width offset

Appendix II

Values of the Extracted Parameters

This appendix aims to tabulate the final value of the extracted parameters for the uniform 35nm n-MOSFET as discussed in chapter 3 and then visualize the impact of some of compact model parameters on device characteristics.

Parameter Name	Final Quantity (35nm)	Dimension
A0	1.81	-
A1	0	V^{-1}
A2	0.6412	-
AGS	0	V^{-1}
CDSC	0.003399	F/m^2
DSUB	0.01745	DROUT
DVT0	0.3241	-
DVT1	0.264	-
DVT2	-0.01176	V^{-1}
ETA0	0.0002269	-
ETAB	-8.398e-5	V^{-1}
K1	0.3662	$V^{1/2}$

K2	-0.02669	-
LP	1e-9	M
LPE0	3.396e-8	M
LPEB	0	M
MINV	2.286	-
NFACTOR	1.039	-
PCLM	0.3395	-
PDIBLC1	0	-
PDIBLC2	0.016	-
PRWB	-0.09093	$V^{-1/2}$
PRWG	0.2168	V^{-1}
PVAG	0.1144	-
RDSW	131.3	$Ohm(\mu m)^{WR}$
RDSWMIN	0	$Ohm(\mu m)^{WR}$
U0	112.7	$M^2/(Vs)$
UA	-4.734e-10	m/V
UB	1.166e-19	m^2/V^2
UC	-1.323e-10	m/V^2
UP	0.01482	-
VOFF	-0.09517	V
VOFFL	4.149e-10	mV
VTH0	0.1168	V
VSAT	1.561e5	m/s

Four typical parameters are selected to visualize their impact on the I_d-V_g characteristics: *DSUB*, *RDSW*, *VOFF* and *VTH0*. Every parameter is deviated by 20% from its nominal value in the uniform model card and the resultant I_d-V_g characteristics are then plotted and compared with the uniform device characteristics. For more discussion, please refer to Chapter 3, page 45.

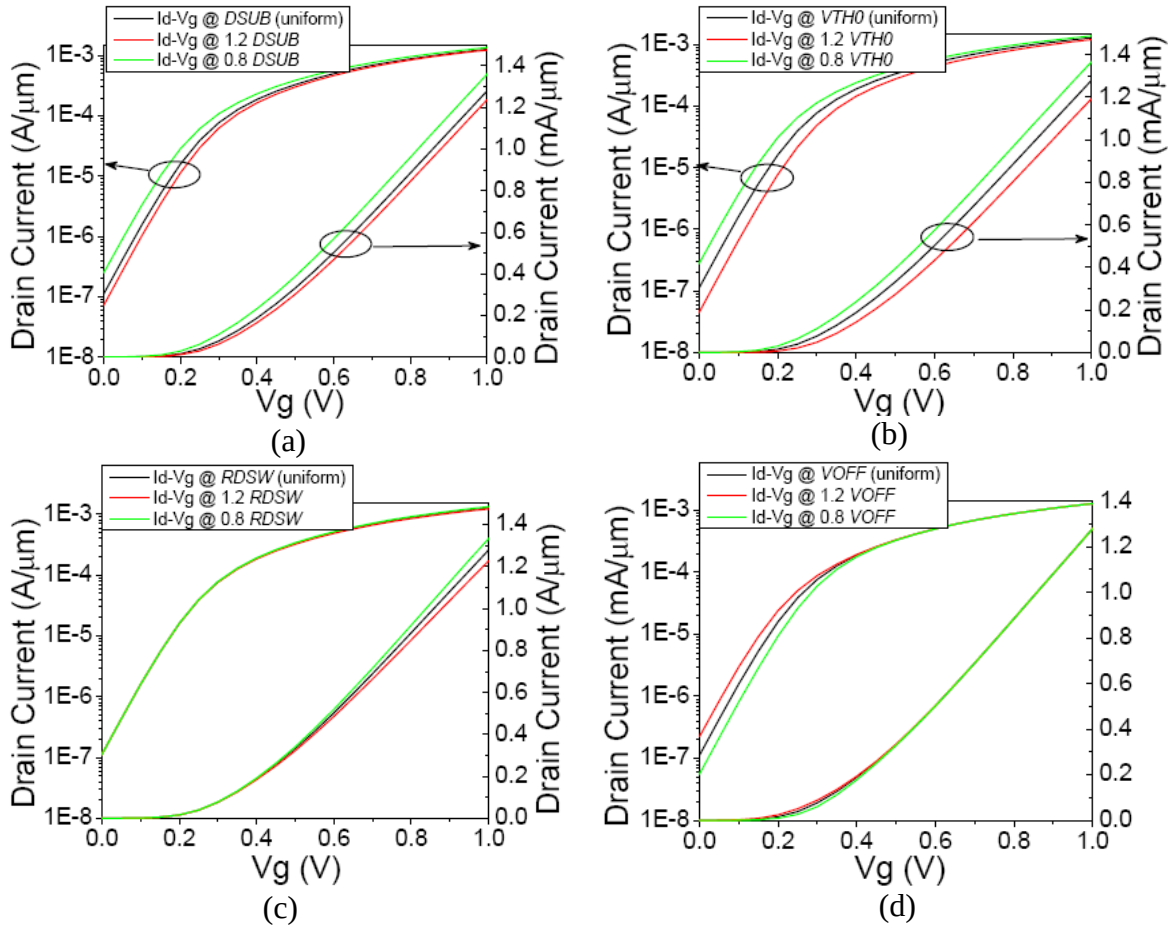


Figure A2: Impact of four typical parameters on I_d-V_g characteristics: (a)-*DSUB*, (b)-*VTH0*, (c)-*RDSW*, (d)-*VOFF* .

Appendix III

Confidence Intervals

Estimation of statistical parameters like mean, standard deviation or a proportion of a distribution has some amount of uncertainty. A confidence interval specifies a confidence level or a measure of reliability in statistical calculations. Moreover, it provides additional information about the distribution function based on frequency theory of probability [149]. In theory, a confidence interval for parameter θ is an interval which is calculated from sample values by a procedure such that if a large number of independent samples is taken, $(1-\alpha)\%$ of the intervals obtained will contain θ [150]. It can be written as:

$$P(L \le \theta \le U) = 1-\alpha \tag{A3.1}$$

where α is called the significance level and the bounds L and U are called lower and upper confidence limits, respectively. Equation (A3.1) states that there is a probability of $1-\alpha$ of selecting a sample for which the confidence interval will contain parameter θ. Assuming a normal distribution with a known variance σ^2, it can be proved that $100(1-\alpha)\%$ confidence interval for the mean μ can be found from [151]:

$$\bar{x} - (z_{\alpha/2})\sigma/\sqrt{n} \le \mu \le \bar{x} + (z_{\alpha/2})\sigma/\sqrt{n} \tag{A3.2}$$

where $\bar{x}$ is the sample mean, *n* is the number of samples and $z_{\alpha/2}$ is the upper $(\alpha/2)$ percentage point of the standard normal distribution. In this thesis with 1000 samples for simulations, the 95% confidence interval for the mean will be $\bar{x} \pm 0.062\sigma$. Even in case when the distribution is not known to be normal and the variance is also unknown, the central limit theorem may be used to give an approximate confidence interval following Equation (A3.2), if *n* is reasonably large ($n \geq 30$) [150,151]. For example, 95% confidence intervals for the mean of 1000 samples will be $\bar{x} \pm 0.062s$ where *s* is the sample standard deviation. Formulas for the confidence intervals of the variance or standard deviation are also given in [150,151].

The confidence interval for a population proportion can be approximated by [151]:

$$\hat{p} - (z_{\alpha/2})\sqrt{\hat{p}(1-\hat{p})/n} \leq p \leq \hat{p} + (z_{\alpha/2})\sqrt{\hat{p}(1-\hat{p})/n} \qquad \text{(A3.3)}$$

where $\hat{p}$ is a point estimator of the proportion of the population and other terms were already defined. For example, 95% confidence interval for half proportion of the distribution with 1000 samples will be 0.5 ± 0.03. This indicates that estimation of half proportion may have 3% error around its point estimation. Finally, the confidence region for high quantiles of a heavy tailed distribution is discussed in details in [152] and the interested reader can refer to it.

Bibliography

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

[2] B. J. Sheu, D. L. Scharfetter, P-K Ko, and M-C Jeng, "BSIM: Berkley Short-Channel IGFET Model for MOS Transistors", *IEEE Journals of Solid-State circuits*, Vol. Sc-22, No. 4, pp. 558-566, 1987.

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

[4] M. Chan, P. Su, H. Wan, C.-H. Lin, S. K.-H. Fung, A. M. Niknejad, C. Hu, P.K. Ko, "Modelling the Floating-Body Effects of Fully Depleted, Partially Depleted, and Body-Grounded SOI MOSFETs", *Solid-State Electronics,* Vol. 48, No. 6, pp. 969-978, 2004.

[5] M. V. Dunga, C.-H. Lin, D. D. LU, W. Xiong, C. R. Cleavelin, P. Patruno, J.-R. Hwang, F.-L. Yang, A. M. Niknejad, C. Hu, "BSIM-MG: A Versatile Multi-Gate FET Model for Mixed-Signal Design", *Digest of Technical Papers, symposium on VLSI Technology*, pp.60-61, 2007.

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

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

[8] S. K. Saha, "Modelling Process Variability in Scaled CMOS Technology", *IEEE J. Design and Test of Computers*, Vol.27, Iss. 2, pp. 8-16, 2010.

[9] S. Nassif, K. Bernstein, D. J. Frank, A. Gattiker, W. Haensch, "High Performance CMOS Variability in 65nm Regime and Beyond", *IEEE International Conference on Electron Devices Meeting*, pp. 569-571, 2007.

[10] 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 Modelling and Characterization in Standard Cell in 45nm CMOS with Stress Enhancement Technique", S*ymposium on VLSI technology*, pp. 90-91, 2008.

[11] J-B Shyu, G. C. Temes, K. Yao, "Random Errors in MOS Capacitors", *IEEE Journals of Solid-State Circuits*, Vol. SC-17, No. 6, pp. 1070–1076, 1982.

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

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

[14] C. C. McAndrew, "Statistical Modelling for Circuit Simulation", *Proceedings of the fourth ISQED*, pp. 357-362, 2003.

[15] K. R. Lakshikumar, R. A. Hadaway, and M. A. Copeland, "Characterization and Modelling of Mismatch in MOS Transistors for Precision Analogue Design", *IEEE Journal of Solid-State Circuits*, Vol. 21, No. 6, pp. 1057-1066, 1986.

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

[17] C.-H. Lin, M. V. Dunga, D. Lu, A. M. Niknejad and C. Hu, "Statistical Compact Modelling of Variations in Nano MOSFETs", *VLSI-TSA*, pp. 165-166, 2008.

[18] X. Li, C. C. McAndrew, W. Wu, S. Chaudhry, J. Victory and G. Gildenblat, "Statistical Modelling with the PSP MOSFET Model", *IEEE Trans. Computer-*

Aided Design of Integrated Circuits and System, pp. 599-606, 2010.

[19] N. Telang and J. M. Higman, "Statistical Modelling Techniques: FPV vs. BPV", *Proc. IEEE Int. Conference on Microelectronic Test Structures*, Vol. 14, pp. 71-75, 2001.

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

[21] http://en.wikipedia.org/wiki/Moore's_law

[22] K. Kuhn, M. D. Giles, D. Becher, P. Kolar, A. Kornfeld, R. Kotlyar, S. T. Ma, A. Maheshwari, and S. Mudanai, "Process Technology Variation", *IEEE Transactions on Electron Devices*, Vol. 58, No. 8, pp. 2197-2208, 2011.

[23] S. M. Sze, K. K. NG, "Physics of Semiconductor Devices", 3rd ed., John Wiley & Sons, 2007.

[24] C. L. Alexander, G. Roy, and A. Asenov, "Random Impurity Scattering Induced Variability in Conventional Nano-Scaled MOSFETs: Ab initio Impurity Scattering Monte Carlo Simulation Study", *International Electron Devices Meeting*, pp. 1-4, 2006.

[25] O. Semenov, A. Pradzynski, and M. Sachdev, "Impact of Gate Induced Drain Leakage on Overall Leakage of Submicrometer CMOS VLSI Circuits", *IEEE Transactions on Semiconductor Manufacturing*, Vol. 15, No. 1, pp. 9-18, 2002.

[26] T. Y. Chan, J. Chen, P. K. Ko, and C. Hu, "The Impact of Gate-Induced Drain Leakage Current on MOSFET Scaling", *International Electron Devices Meeting*, pp. 718-721, 1987.

[27] Y. Taur, C. H. Wann, and D. J. Frank, "25nm CMOS Design Considerations", *International Electron Devices Meeting*, pp. 789-792, 1998.

[28] J. Welser, J.L. Hoyt, S. Takagi, and J.F. Gibbons, "Strain Dependence of the Performance Enhancement in Strained-Si n-MOSFETs", *International Electron Devices Meeting*, pp. 373-376, 1994.

[29] K. Rim, J. Welser, J.L. Hoyt, and J.F. Gibbons, "Enhancement Hole Mobilities in Surface-Channel Strained-Si p-MOSFETs", *International Electron Devices Meeting*, pp. 517-520, 1995.

[30] N. A. Kasmani, "Statistical Circuit Simulations-From 'Atomistic' compact Models to Statistical Standard Cell Characterisation", *Ph.D. dissertation*, University of Glasgow, 2011.

[31] J. Frank, R. H. Dennard, E. Nowak, P. M. Solomon, Y. Taur, and H.-S. P. Wong, "Device Scaling Limits of Si MOSFETs and Their Application Dependencies", *Proceedings of the IEEE*, Vol. 89, No. 3, pp. 259-288, 2001.

[32] M. A. Quevedo-Lopez, S. A. Krishnan, D. Kirsch, C. H. J. Li, J. H. Sim, C. Huffman, J. J. Peterson, B. H. Lee, G. Pant, B. E. Gnade, M. J. Kim, R. M. Wallace, D. Guo, H. Bu, and T. P. Ma, "High Performance Gate First HfSiON Dielectric Satisfying 45nm Node Requirements", *IEEE International Electron Device Meeting*, pp. 425–428, 2005.

[33] T. Mizuno, M. Iwase, H. Niiyama, T. Shibata, K. Fujisaki, T. Nakasugi, A. Toriumi, and Y. Ushiku, "Performance Fluctuations of 0.10 μm MOSFETs-Limitation of 0.1 μm ULSIs", *Symposium on VLSI Technology*, pp. 13-14, 1994.

[34] T. Mizuno, J. Okamura, and A. Toriumi, "Experimental Study of Threshold Voltage Fluctuation Due to Statistical Variation of Channel Dopant Number in MOSFET's", *IEEE Transaction on Electron Devices*, Vol. 41, No. 11, pp. 2216-2221, 1994.

[35] T. Yamaguchi, H. Namatsu, M. Nagase, K. Yamazaki, and K. Kurihara, "Nanometer-Scale Linewidth Fluctuations Caused by Polymer Aggregates in Resist Films", *Applied Physics Letter*, Vol. 71, No. 16, pp. 2388-2390, 1997.

[36] M. Pelgrom, H. Tuinbout, M. Vertregt, "Modelling of CMOS Matching", In: G. Gildenblat (ed.) "Compact Modelling: Principles, Tecgniques and Applications", Spring, pp.453-490, 2010.

[37] K. Bernstein, D. J. Frank, A. E. Gattiker, W. Haensch, B. L. Ji, S. R. Nassif, E. J.

Nowak, D. J. Pearson, and N. J. Rohrer, “High-performance CMOS variability in the 65-nm regime and beyond”, *IBM Journal of Research and Development*, Vol. 50, No. 4.5, pp. 433-449, 2006.

[38] J. N. Randall and A. Tritchkov, “Optically induced mask critical dimension error magnification in 248nm lithography”, *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, Vol. 16, No. 6, pp. 3606-3611, 1998.

[39] E. Morifuji, H. Aikawa, H. Yoshimura, A. Sakata, M. Ohta, M. Iwai, and F. Matsuoka, “Layout Dependence Modelling for 45-nm CMOS With Stress Enhanced Technique”, *IEEE Transactions on Electron Devices*, Vol. 56, No. 9, pp. 1991-1998, 2009.

[40] T. Kanamoto, Y. Ogasahara, K. Natsume, K. Yamaguchi, H. Amishiro, T. Watanabe, and M. Hashimoto, “Impact of Well Edge Proximity Effect on Timing”, *Solid State Circuit Conference*, pp. 115–118, 2007.

[41] U. Kovac, “3D Drift Diffusion and 3D Monte Carlo Simulation of On-Current Variability Due to Random Dopant”, *Ph.D. dissertation*, University of Glasgow, 2010.

[42] 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”, *Electron Device Letters*, Vol. 29, No. 6, pp. 609-611, 2008.

[43] A. Asenov, S. Roy, R. A. Brown, G. Roy, C. Alexander, C. Riddet, C. Millar, B. Cheng, A. Martinez, N. Seoane, D. Reid, M. F. Bukhori, X. Wang, U. Kovac, “Advanced Simulation of Statistical Variability and Reliability in Nano CMOS Transistors”, *IEEE International Electron Devices Meeting,* pp.1-1, 2008.

[44] D. J. Frank, Y. Taur,M. Ieong, and H.S. P.Wong, “Monte-Carlo Modelling of Threshold Variation Due to Dopant Fluctuations”, *Symposium on VLSI Technol. Digest of Technical Papers*, pp. 169-170, 1999.

[45] P. A. Stolk, and D. B. M. Klaassen, "The Effect of Statistical Dopant Fluctuations on MOS Device Performance", *International Electron Devices Meeting*, pp. 627-630, 1996.

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

[47] http://userweb.eng.gla.ac.uk/scott.roy/atomistic.html

[48] K. Takeuchi, T. Tatsumi, and A. Furukawa, "Channel Engineering for The Reduction of Random-Dopant-Placement-Induced threshold voltage Fluctuations", *International Electron Device Meeting*, pp. 841-844, 1997.

[49] R. Difrenza, P. Llinares, and G. Gibaundo. "The Impact of Short Channel and Quantum Effects on MOS Transistor Mismatch", *Solid-State Electronics*, Vol. 47, No. 7, pp. 1161-1165, 2002.

[50] Y. Taur, D. Buchanan, W. Chen, D. Frank, K. Ismail, S. H. Lo, G. Sai- Halasz, R. Viswanathan, H. J. C. Wann, S. Wind, and H. S. Wong. "CMOS Scaling into the Nanometer Regime", *Proceeding of the IEEE*, Vol. 85, No. 4, pp. 486-504, 1997.

[51] T. Hagivaga, K. Yamaguchi, and S. Asai, "Threshold Voltage Variation in Very Small MOS Transistors Due to Local Dopant Fluctuations", *Symposium on VLSI Technology*, pp. 46-47, 1982.

[52] K. Nishinohara, N. Shigyo, and T. Wada, "Effects of Microscopic Fluctuations in Dopant Distributions on MOSFETs Threshold Voltage", *IEEE Trans*actions *on Electron Device*, Vol. 39, No. 3, pp. 634-639, 1992.

[53] A. Asenov, "Random Dopant Induced Threshold Voltage Lowering and Fluctuations in Sub 50nm MOSFET's: A 3-D 'Aatomistic' Simulation Study", *Nanotechnology*, Vol. 10, No. 2, pp. 153-158, 1999.

[54] A. Asenov, "Random Dopant Induced Threshold Voltage Lowering and Fluctuations in Sub-0.1µm MOSFET's: A 3-D 'Aatomistic' Simulation Study", *IEEE Transactions on Electron Devices*, Vol. 45, No. 12, pp. 2505-2513, 1998.

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

[56] A. Asenov, S. Kaya, and 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, 2003.

[57] C. H. Diaz, H.-J. Tao, Y.-C. Ku, A. Yen, and K. Young, "An Experimentally Validated Analytical Model for Gate Line-Edge Roughness (LER) Effects on Technology Scaling", *IEEE on Electron Device Letters*, Vol. 22, No. 6, pp. 287–289, 2001.

[58] http://web.eng.gla.ac.uk/groups/dev_mod/research/atomistic.php

[59] X. Wang, "Simulation Study of Scaling Design, Performance Characterization, Statistical Variability and reliability of decananometer MOSFETs", *Ph.D. dissertation*, University of Glasgow, 2010.

[60] G. Roy, A. R. Brown, F. Adamu-Lema, S. Roy, and A. Asenov, "Simulation Study of Individual and Combined Sources of Intrinsic Parameter Fluctuations in Conventional NanoMOSFETs", *IEEE Transactions on Electron Devices*, Vol. 53, No. 12, pp. 3063-3070, 2006.

[61] T. D. Drysdale, A. R. Brown, G. Roy,S. Roy, and A. Asenov. "Capacitance Variability of Short Range Interconnects", *Journal of Computational Electronics*, Vol. 7, No. 3, pp. 124-127, 2008.

[62] 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.

[63] A. Asenov, and S. Kaya, "Effect of Oxide Interface Roughness on the Threshold Voltage Fluctuations in Decanano MOSFETs with Ultrathin Gate Oxides", *International Conference on Simulation of Semiconductor Processors and*

devices, pp. 135-138, 2000.

[64] A. Asenov S. Kaya, J. H. Davies, and S. Saini. "Oxide Thickness Variation Induced Threshold Voltage Fluctuations in Decanano MOSFET's: A 3D Density Gradient Simulation Study", *Superlattices and Microstructures*, Vol. 28, No. 5-6, pp. 507-515, 2000.

[65] A. R. Brown, J. R. Watling, A. Asenov, G. Bersuker, and P. Zeitzoff, "Intrinsic Parameter Fluctuations in MOSFETs Due to Structure Non-Uniformity of High-κ Gate Stacks Materials", *International Conference on Simulation of Semiconductor Processors and Devices*, pp. 27-30, 2005.

[66] A. R. Brown, J. R. Watling, and A. Asenov, "Intrinsic parameter fluctuations due to random grain orientations in high-κ gate stacks", *Journal of Computer Electronics*, Vol. 5, No. 4, pp. 333–336, 2006.

[67] S. Saito, D. Hisamoto, S. Kimura, and M. Hiratani, "Unified Mobility Model for High-κ Gate Stacks", *International Electron Device Meeting,* pp. 33.3.1-33.3.4, 2003.

[68] A. Asenov and S. Saini, "Polysilicon Gate Enhancement of the Random Dopant Induced Threshold Voltage Fluctuations in Sub-100nm MOSFETs With Ultrathin Gate Oxide", *IEEE Transactions on Electron Devices*, Vol. 47, No. 4, pp. 805–812, 2000.

[69] 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, 2007.

[70] 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.

[71] U. Ravaioli, "Hierarchy of Simulation Approaches for Hot Carrier Transport in Deep Sub-Micron Devices", *International Workshop on Computational Electronics*, pp. 100-103, 1998.

[72] D. K. Ferry, R. Akis, and D. Vasileska. "Quantum effects in MOSFETs: Use of an Effective Potential in 3D Monte Carlo Simulation of Ultra Short Channel Devices", *International Electron Device Meeting,* pp. 287-290, 2000.

[73] S. Selberherr, "Analysis and Simulation of Semiconductor Devices", Springer-Verlag, 1984.

[74] M. G. Ancona and G. J. Iafrate, "Quantum Correction to the Equation of State of an Electron Gas in a Semiconductor", *Physical Review B*, Vol. 39, No. 13, pp. 9536-9540, 1989.

[75] 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.

[76] J. Taiault, J. Foucher, J. H. Tortai, O. Jubert, S. Landis and S. Pauliac, "Line Edge Roughness Characterization With Three-Dimensional Atomic Force Microscope: Transfer During Gate Patterning Process", *Journal of Vacuum Science & Technology B*, Vol. 23, No. 6, pp. 3075-3079, 2005.

[77] 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.

[78] B. Cheng, S. Roy, G. Roy, F. Adamu-Lema, and A. Asenov, "Impact of Intrinsic Parameter Fluctuations in Decanano MOSFETs on Yield and Functionality of SRAM Cells", *Solid-State Electronics*, Vol. 49, pp. 740–746, 2005.

[79] R. Spence and R. S. Soin, "Tolerance Design of Electronic Circuits", Addison-Wesley Publishing Co., Reading, MA, 1988.

[80] P. Yang and P. Chatterjee, "Statistical Modelling of Small Geometry MOSFET", *International Electron Device Meeting,* pp. 286-289, 1982.

[81] P. Cox, P. Yang, S. S. Mahant-Shetti, and P. Chatterjee, "Statistical Modelling for Efficient Parametric Yield Estimation of MOS VLSI Circuits", *IEEE Journal of Solid-State Circuits*, Vol. 20, No. 1, pp. 391-398, 1985.

[82] N. Herr and J. J. Barnes, "Statistical Circuit Simulation Modelling of CMOS VLSI", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 5, No. 1. pp. 15-22, 1986.

[83] J. P. Spoto, W. T. Coston, and C. P. Hernandez, "Statistical Integrated Circuit Design and Characterization", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 5, No. 1. pp. 90-103, 1986.

[84] T. K. Yu, S. M. Kang, I. N. Hajj, and T. N. Trick, "Statistical Performance Modelling and Parametric Yield Estimation of MOS VLSI", *IEEE Transactions on Computer-Aided*, Vol. 6, No. 6. pp. 1013-1022, 1987.

[85] P. Tuohy, A. Gribben, A. J. Walton, and J. M. Robertson, "Realistic Worst-Case Parameters for Circuit Simulation", *IEEE Journal of Solid-State Circuits*, Vol. 134, No. 5, pp. 137-140, 1987.

[86] G. Rappitsch, E. Seebacher, M. Kocher and E. Stadlober, "SPICE Modelling of Process Variation Using Location Depth Corner Models", *IEEE Transactions on Semiconductor Manufacturing*, Vol. 17, No.2, pp. 201-213, 2004.

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

[88] C. Michael, M. Ismail, "Statistical Modelling for Computer-Aided Design of MOS VLSI Circuits", Kluwer Academic Publishers, 1993.

[89] M. Balsi, F. Centurelli, A. Forte, G. Scotti, P. Tommasino, and A. Trifiletti, "Non-Linear Statistical Modelling of GaAs FET Integrated Circuits Using Principal Component Analysis", *GaAs Applications Symposium*, 2003.

[90] 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.

[91] 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. 30, No. 8, pp. 873-875, 2009.

[92] C.C. McAndrew, I. Stevanovic, X. Li, G. Gildenblat, "Extensions to Backward Propagation of Variance for Statistical Modelling", *IEEE Design and Test of Computers*, Vol. 27, No. 2, pp. 36-43, 2010.

[93] B. Bindu, B. Cheng , G. Roy, X. Wang, S. Roy, A. Asenov, "Parameter Set and Data Sampling Strategy for Accurate yet Efficient Statistical MOSFET Compact Model Extraction", *Solid-State Electronics*, Vol. 54, No. 3, pp. 307-315, 2010.

[94] S. V. Kumar, C. H. Kim and S. S. Sapatnekar, "An Analytical Model for Negative Bias Temperature Instability", *IEEE/ACM International Conference on Computer-Aided Design*, pp. 493-496, 2006.

[95] W. Wang, V. Reddy, A. T. Krishnan, R. Vattikonda, S. Krishnan and Y. Cao, "Compact modelling and simulation of circuit reliability for 65nm CMOS technology", *IEEE Transaction on Device and Materials Reliability*, Vol. 7, No. 4, pp. 509-517, 2007.

[96] W. Wang, V. Balakrishnan, B. Yang, Y. Cao, "Statistical Prediction of NBTI-induced Circuit Aging", *International Conference on Solid-State and Integrated-Circuit Technology*, pp. 416-419, 2008

[97] J. P. Campbell, P. M. Lenahan, C. J. Cochrane, A. T. Krishnan, S. Krishnan, "Atomic-Scale Defects Involved in the Negative-Bias Temperature Instability", *IEEE Transactions on Device and Materials Reliability*, Vol. 7, No. 4, pp. 540-557, 2007.

[98] V. Huard, C. Parthasarathy, C. Guerin, T. Valentin, E. Pion, M. Mammasse, N. Planes and L. Camus, "NBTI Degradation: From Transistor to SRAM Arrays", *IEEE International Symposium on Reliability Physics*, pp. 289-300, 2008.

[99] P. M. Lee, M. M. Kuo, P. K. Ko, and C. Hu, "Circuit Aging Simulator (CAS)", *IEEE International Conference on Electron Devices Meeting*, pp. 134-137, 1988.

[100] P. M. Lee, T. Garfinkel, P. K. Ko, and C. Hu, "Simulating the Competing Effects of P- and N-MOSFET Hot-Carrier Aging in CMOS Circuits", *IEEE Transaction on Electron Devices*, Vol. 41, No.5, pp. 852-853, 1994.

[101] N. Arora, "MOSFET Modelling for VLSI Simulation Theory and Practice", World Scientific, 2007.

[102] D. K. Schroder, "Semiconductor Material and Device Characterization", John Wiley & Sons Inc., New York, 1990.

[103] B. J. Sheu, D. L. Scharfetter, P.-K. Ko, M.-C. Jeng, "BSIM: Berkeley Short-Channel IGFT Model for MOS Transistors", *IEEE Journal of Solid-State*, Vol. 22, No. 4, pp. 558-566, 1987.

[104] Y. Cheng, M-C. Jeng, Z. Liu, J. Huang, M. Chan, K. Chen, P. K. ko and C. Hu, "A Physical and Scalable I-V Model in BSIM43v3 for Analog/Digital Circuit Simulation", *IEEE Transactions on Electron Devices*, vol. 44, No. 2, pp. 277-287, 1997.

[105] T. Ytterdal, Y. Cheng and T. Fjeldly, "Device Modelling for Analog and RF CMOS Circuit Design", John Wiley & Sons Inc., New York, 2003.

[106] http://www.freeda.org/doc/elements/mosnBSIM4.pdf

[107] L. D. Yau, "A Simple Theory to Predict the Threshold Voltage of Short-Channel IGFET's", *Solid-State Electronics*, Vol. 17, No. 10, pp. 1059-1063, 1974.

[108] X. Wang, S. Roy, A. Asenov, "Impact of Strain on the Performance of High-K/Metal Replacement Gate MOSFETs", *International Conference on Ultimate Integration of Silicon*, pp.289-292, 2009.

[109] Z.-H. Liu, C. Hu, J.-H. Huang, T.-Y. Chan, M.-C. Jeng, P. K. Ko, and Y. C. Cheng, "Threshold Voltage Model for Deep-Submicron MOSFETs", *IEEE Transactions on Electron Devices*, Vol. 40, No. 1, pp. 86-95, 1993.

[110] Y. Cheng, K. Chen, K. Imai, and C. Hu, "A Unified MOSFET Channel Charge Model for Device Modelling in Circuit Simulation", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 17, No. 8, pp. 641-644, 1998.

[111] A. G. Sabnis and J. T. Clemens, "Characterization of the Electron Mobility in the Inverted <100> Si Surface", *International Electron Device Meeting*, pp.18–21, 1979.

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

[113] K. Mistry, C. Allen, C. Auth, B. Beattie, D. Bergstrom, M. Bost, M. Brazier, M. Buehler, A. Cappellani, R. Chau, C.-H. Choi, G. Ding, K. Fischer, T. Ghani, R. Grover, W. Han, D. Hanken, M. Hattendorf, J. He, J. Hicks, R. Huessner, D. Ingerly, P. Jain, R. James, L. Jong, S. Joshi, C. Kenyon, K. Kuhn, K. Lee, H. Liu, J. Maiz, B. McIntyre, P. Moon, J. Neirynck, S. Pae, C. Parker, D. Parsons, C. Prasad, L. Pipes, M. Prince, P. Ranade, T. Reynolds, J. Sandford, L. Shifren, J. Sebastian, J. Seiple, D. Simon, S. Sivakumar, P. Smith, C. Thoms, T. Troeger, P. Vandervoorn, S. Williams, K. Zawadzki, "A 45nm Logic Technology With High-k+Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-Free Packaging", *IEEE International Electron Device Meeting*, pp. 247-250, 2007.

[114] K. L. Cheng, C.C. Wu, Y.P. Wang, D.W. Lin, C.M. Chu, Y.Y. Tarng, S.Y. Lu, S.J. Yang, M.H. Hsieh, C.M. Liu, S.P. Fu, J.H. Chen, C.T. Lin, W.Y. Lien, H.Y. Huang, P.W. Wang, H.H. Lin, D.Y. Lee, M.J. Huang, C.F. Nieh, L.T. Lin, C.C. Chen, W. Chang, Y.H. Chiu, M.Y. Wang, C.H. Yeh, F.C. Chen, C.M. Wu, Y.H. Chang, S.C. Wang, H.C. Hsieh, M.D. Lei, K. Goto, H.J. Tao, M. Cao, H.C. Tuan, C.H. Diaz, and Y.J. Mii, "A Highly Scaled, High Performance 45nm Bulk Logic CMOS Technology With 0.242μm^2 SRAM Cell", *IEEE International Electron Device Meeting*, pp. 2007.

[115] S. Inaba, K. Okana, S. Matsuda, M. Fujiwara, A. Hokazono, K. Adachi, K.

Ohuchi, H. Suto, H. Fukui, T. Shimizu, S. Mori, H. Oguma, A. Murakoshi, T. Itani, T. Iinuma, T. Kudo, H. Shibata, S. Taniguchi, M. Takayanagi, A. Azuma, H. Oyamatsu, K. Suguro, Y. Katsumata, Y. Toyoshima, and H. Ishiuchi, "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, 2002.

[116] Z. Luo, N. Rovedo, S. Ong, B. Phoong, M. Eller, H. Utomo, C. Ryou, H. Wang, R**.** Stierstorfer, L. Clevenger, S. Kim, J. Toomey, D. Sciacca, J. Li, W. Wille, L. Zhao, L. Teo, T. Dyer, S. Fang, J. Yan, O. Kwon, O. Kwon, D. Park, J. Holt, J. Han, V. Chan, J. Yuan, T. Kebede, H. Lee, S. Kim, S. Lee, A. Vayshenker, Z. Yang, C. Tian, H. Ng, H. Shang, M. Hierlemann, J. Ku, J. Sudijono, M. Ieong, "High Performance Transistors Featured in an Aggressively Scaled 45nm Bulk CMOS Technology", *IEEE Symposium on VLSI Technology*, pp. 16-17, 2007.

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

[118] Y. Cao, C. McAndrew, "MOSFET Modelling for 45nm and Beyond", *International Conference on Computer-Aided Design*, pp. 638-643, 2007.

[119] B. Cheng, S. Roy, G. Roy, F. Adamu-Lema and A. ASenov, "Impact of Intrinsic Parameter Fluctuations in Decanano MOSFETs on Yield and Functionality of SRAM Cells", *Solid-State Electronics*, Vol. 49, No. 5, pp. 740-746, 2005.

[120] T. Grasser, T.-W. Tang, H. Kosina, and S. Selberherr, "A Review of Hydrodynamic Energy-Transport Models for Semiconductor Device Simulation", *Proceeding of the IEEE*, Vol. 91, No. 2, pp. 251-274, 2003.

[121] D. T. Reid, "Large-Scale Simulations of Intrinsic Parameter Fluctuations in Nano-Scale MOSFETs", PhD Thesis, University of Glasgow, 2010.

[122] A. Asenov, "Simulation of Statistical Variability in Nano MOSFETs", *IEEE Symposium on VLSI Technology*, pp. 86-87, 2007.

[123] A. Asenov, G. Slavcheva, A. R. Brown, J. H. Davies and S. Saini, "Increase in

The Random Dopant Induced Threshold Fluctuations and Lowering in Sub-100nm MOSFETs Due to Quantum Effects: A 3-D Density-Gradient Simulation Study", *IEEE Transactions on Electron Devices*, Vol. 48, No.4, pp. 722-729, 2001.

[124] A. Asenov, A. Cathignol, B. Cheng, K. P. McKenna, A. R. Brown, A. L. Shluger, D. Chanemougame, K. ochereau, 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.

[125] C. Alexander, G. Roy and A. Asenov, "Random-Dopant-Induced Drain Current Variation in Nano-MOSFETs: A Three-Dimensional Self-Consistent Monte Carlo Simulation Study Using 'Ab Initio' Ionized Impurity Scattering," *IEEE Transactions on Electron Devices*, Vol. 55, No. 11, pp. 3251-3258, 2008.

[126] A. Ruangphanit, S. Harnsoongnoen, R. Muanghlua, N. Phongphanchantra, and A. Polyai, "Drain Induced Barrier Lowering (DIBL) Effect in Submicrometer NMOS Devices", *IEEE Transactions on Electron Devices*, Vol. 39, No. 4, pp. 908-915, 1999.

[127] S. Aghtar, J. W. Haslett, and F. N. Trofimenkoff, "Subthreshold Analysis of an MOS Analog Switch", *IEEE Transactions on Electron Devices*, Vol. 44, No. 1, pp. 89-96, 1997.

[128] G. E. P. Box, M. E. Muller, "A Note on the Generation of Random Normal Deviation", *Annals of Mathematical Statistics*, Vol. 29, No. 2, pp. 610-611, 1958.

[129] Q. Y. Tang, "Uncertainty Propagation in Transistor-Level Statistical Circuit Analysis", *Ph.D. dissertation*, University of California at Berkeley, May 2011.

[130] L. I. Smith, "A Tutorial on Principal Components Analysis", 2002.

[131] K. Takeuchi, and M. Hane, "A Highly Efficient Statistical Compact Model Parameter Extraction Scheme", *International Conference on Simulation of Semiconductor Processes and Devices*, pp. 135-138, 2005.

[132] J. Shlens, "A Tutorial on Principal Components Analysis", from: www.snl.salk.edu/~shlens/pub/notes/pca.pdf

[133] C. D. Vale, V. A. Maurelli, "Simulating Multivariate Non-Normal Distributions", *Psychometrika*, Vol. 48, No. 3, pp. 465-471, 1983.

[134] M. H. Na, E. J. Nowak, W. Haensch and J. Cai, "The Effective Drive Current in CMOS Inverter", *International Electron Devices Meeting,* pp. 121-124, 2002.

[135] V. Adler, E. G. Friedman, "Delay and Power Expressions for a CMOS Inverter Driving a Resistive-Capacitive Load", *Analog Integrated Circuit and Signal Processing*, Vol. 14, No 1-2, pp. 29-39, 1997.

[136] A. I. Fleishman, "A Method for Simulating Non-normal Distributions", *Sychometrika*, Vol. 43, No. 4, pp. 521-532, 1978.

[137] A. Stuart and K. Ord, "The Advanced Theory of Statistics", Volume 1: Distribution Theory, fourth Edition, Macmillan, 1977

[138] G. C. Wick, "The Evaluation of the Collision Matrix", *Physics Review Online Archive*, Vol. 80, pp. 268 -272, 1950.

[139] L. Isserlis and D. Sc., "On a Formula for the Product-Moment Coefficient of Any Order of a Normal Frequency Distribution in Any Number of Variables", *Biometrika*, Vol. XI, Part III, pp.134-139, 1918.

[140] Y. Ye, F. Liu, M. Chen, S. Nassif, and Y. Cao, "Statistical Modelling and Simulation of Threshold Variation under Random Dopant Fluctuations and Line-Edge Roughness", *IEEE Transactions on VLSI systems*, Vol. 19, No. 6, pp. 987-996, 2011.

[141] M. J. M. Pelgrom, H. P. Tuinhout, M. Vertregt "Transistor Matching in Analog CMOS Applications", *International Electron Device Meeting,* pp. 915-918, 1998.

[142] C. R. Viswanathan, B. C. Burkey, G. Lubberts, and T. J. Tredwell, "Threshold Voltage in Short Channel MOS Devices", *IEEE Transactions on Electron Devices*, Vol. 32, No. 5, pp. 932-940, 1985.

[143] T. Toyabe and S. Asai, "Analytical Models of Threshold Voltage and Breakdown Voltage of Short-Channel MOSFET's Derived from Two-Dimensional Analysis", *IEEE Journal of Solid-State*, Vol. 14, No. 2, pp. 375-383, 1979.

[144] C. Y. Wu, S. Y. Yang, H. H. Chen, F. C. Tseng, and C. T. Shih, "An Analytic and Accurate Model for the Threshold Voltage of Short Channel MOSFETs in VLSI", *Solid-State Electronics*, Vol. 27, No. 7, pp. 651-658, 1984.

[145] K. N. Rantnakumar, and J. D. Meindel, "Short-Channel MOST Threshold Voltage Model", IEEE Journal of Solid-State Circuits, Vol. 17, No. 5, pp. 937-948, 1982.

[146] T. N. Nguyen, and J. D. Plummer, "Physical Mechanisms Responsible for Short Channel Effects in MOS Devices", *International Electron Devices Meeting*, pp. 596-599, 1981.

[147] D. R. Poole, and D. L. Kwong, "Two-Dimensional Analytical Modeling of Threshold Voltage of Short-Channel MOSFET's", *IEEE Electron Device Letters*, Vol. 5, No. 11, pp. 443-446, 1984.

[148] S. Chakraborty, A. Mallik, C. K. Sakar, and V. R. Rao "Impact of Halo Doping on the Subthreshold Performance of Deep-Submicrometer CMOS Devices and Circuits for Ultralow Power Analog/Mixed-Signal Applications", *IEEE Transactions on Electron Devices*, Vol. 54, No. 2, pp. 241-248, 2007.

[149] Kendall and Stuart, "The Advanced Theory of Statistics", Vol. 2, 4th ed., 1977.

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

[151] Montgomery, Runger, Hubele, "Engineering Statistics", 5th ed., John Wiley & Sons, 2011.

[152] L. Peng, Y. Qi, "Confidence Regions for High Quantiles of a Heavy Tailed Distribution", *The Annals of Statistics*, Vol. 34, No. 4, pp.1964-1986, 2006.