# Robustness Analysis of Controllable-Polarity Silicon Nanowire Devices and Circuits

THÈSE N$^{O}$ 6924 (2016)

PRÉSENTÉE LE 8 AVRIL 2016

À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS
LABORATOIRE DES SYSTÈMES INTÉGRÉS (IC/STI)
PROGRAMME DOCTORAL EN INFORMATIQUE ET COMMUNICATIONS

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Hassan GHASEMZADEH MOHAMMADI

acceptée sur proposition du jury:

Prof. B. Falsafi, président du jury
Prof. G. De Micheli, directeur de thèse
Prof. M. Sonza Reorda, rapporteur
Prof. M. Nicolaidis, rapporteur
Prof. D. Atienza Alonso, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2016

To Maryam for her endless love,
and to the memory of my father, Ahmad

# Acknowledgements

This thesis was a unique adventure during my study at EPFL. This work could not be done without the help and support of many people. Collaborating, discussing, and working with such great colleagues were invaluable experience during the last 4 years.

I would particularly like to thank my supervisor Prof. Giovanni De Micheli for all of his supports, guides and time for constructive discussions. He was a constant source of support and encouragement during the challenges of my research.

Working with a knowledgeable person like Prof. Pierre-Emmanuel Gaillardon is always a remarkable experience. I would like to thank him for his tremendous guides and supports.

I would also appreciate my examiners, Prof. Babak Falsafi, Prof. David Atienza Alonso, Prof. Matteo Sonza Reorda, and Prof. Michael Nicolaidis, who provided encouraging and beneficial feedbacks.

My colleagues at EPFL have a significant role in my project. I wish to thank Dr. Majid Yazdani and Dr. Sara S. Ghoreishizadeh. It has been a great pleasure working with them.

My greatest thanks go to all the members of *Integrated System Laboratory* (LSI) for the wonderful moments I spent with them. Moreover, I specially thank Christina Govoni for all tremendous help and administrative support during my PhD and also my dear friends in LSI.

My stay at Lausanne would not have been pleasing without the great time spent with Somayyeh, Mahdi, Amin, Nima, Sahar, Ardalan, Samira, Afsaneh, Mohammad-Javad, Glareh, Peyman and all of my lovely friends. I wish the best for you in your life.

I would like to thank my family, my mother, brother and sister, for their unconditional support, and encouragement, without which I would not have come this far. Maryam, with her endless love and support, has always been the greatest inspiration in my life. Without her companionship and love, I would not have been able to reach this end.

*Hassan Ghasemzadeh Mohammadi*
*Lausanne, January 2016*

# Abstract

Substantial downscaling of the feature size in current CMOS technology has confronted digital designers with serious challenges including short channel effect and high amount of leakage power. To address these problems, emerging nano-devices, e.g., *Silicon NanoWire FET* (SiNWFET), is being introduced by the research community. These devices keep on pursuing Moore's Law by improving channel electrostatic controllability, thereby reducing the *Off*-state leakage current. In addition to the improvements in conventional device performances, recent developments introduced devices with enhanced capabilities such as *Controllable-Polarity* (CP) SiNWFETs. In particular, these transistors support in-field reconfiguration which makes them very interesting for compact logic cell and arithmetic circuits.

At advanced technology nodes, fabrication-induced variations are expected to significantly affect the yield of complex circuits. Indeed, the amount of physical controls, during the fabrication process of nanometer devices, cannot be precisely determined because of technology fluctuations. Consequently, the physical and structural parameters of fabricated circuits can be significantly different from their nominal values. Moreover, giving an *a-priori* conclusion on the variability of advanced technologies for emerging nanoscale devices, with novel geometries and different fabrication process, is a difficult task and novel estimation methodologies are required. This is a necessity to guarantee the performance and the reliability of future integrated circuits.

Statistical analysis of process variation requires a great amount of numerical data for nanoscale devices. This introduces a serious challenge for variability analysis of emerging technologies due to the lack of fast simulation models. One the one hand, the development of accurate compact models entails numerous tests and costly measurements on fabricated devices. On the other hand, *Technology Computer Aided Design* (TCAD) simulations, that can provide precise information about devices behavior, are too slow to timely generate large enough data set. In this research, a fast methodology for generating data set for variability analysis is introduced. This methodology combines the TCAD simulations with a learning algorithm to alleviate the time complexity of data set generation for emerging nano-devices.

Another formidable challenge for variability analysis of the large circuits is growing number of process variation sources in deeply nanoscale technologies. Utilizing parameterized models is becoming a necessity for chip design and verification. However, the high dimensionality of parameter space imposes a serious problem. Unfortunately, the available dimensionality reduction techniques cannot be employed for three main reasons of lack of accuracy, distribution dependency of the data points, and finally incompatibility with device and circuit

simulators. We propose a novel technique of parameter selection for modeling process and performance variation. The proposed technique efficiently takes into account the nonlinearity among process and performance parameters which is necessary for high precision variation analysis.

Appropriate testing, to capture manufacturing defects, plays an important role on the quality of integrated circuits. Compared to conventional CMOS, emerging nano-devices such as CP-SiNWFETs have different fabrication process steps. Therefore, the type of defects for these technologies is different from CMOS devices. In this case, current fault models must be extended for defect detection of emerging technologies. In this research, we investigated the inefficiency of the current CMOS fault models for detecting the fabrication defects of CP-SiNWFET technology. Considering the fabrication steps, we extracted the possible fabrication defects, and then proposed a fault model for this technology. We also provided a couple of test methods for detecting the manufacturing defects in various types of CP-SiNWFET logic gates. Finally, we used the obtained fault model to build fault tolerant arithmetic circuits with a bunch of superior properties compared to their competitors.

Key words: Controllable-Polarity Silicon Nanowire, 3-D TCAD Simulation, Process Variation, Statistical Analysis, Circuit Modeling, Fault Models, Arithmetic Circuits.

# Résumé

L’importante réduction de la taille aus la technologie CMOS actuelle a confronté les concepteurs numériques à de sérieux défis tels que l’effet de canal court, ainsi qu’une des transistors à importante puissance de fuite. Pour résoudre ces problèmes, de nouveaux nanodispositifs, p.ex., Nanofils de Silicium (SiNWFET), ont été présentés par la communauté de recherche. Ces dispositifs continuent de suivre la loi de Moore tout en améliorant la contrôlabilité électrostatique du canal, réduisant ainsi le courant de fuite à l’état ouvert. En plus des améliorations de performances des dispositifs classiques, une évolution récente a introduit des dispositifs qui sont dotés de fonctionnalités améliorées comme les SiNWFETs polarité contrôlable (CP). En particulier, ces transistors supportent une reconfiguration dynamique ce qui les rendent très intéressants pour la réalisation de cellules logique compactes, ainsi que pour les circuits arithmétiques.

A des nœuds technologiques avancés, les variations induites par la fabrication devraient affecter de manière significative le rendement des circuits complexes. En effet, la quantité des contrôles physiques, durant le processus de fabrication des dispositifs nanométriques ne peut être déterminée avec précision en raison des fluctuations de la technologie. Par conséquent, les paramètres physiques et structurelles des circuits fabriqués peuvent être sensiblement différents de leurs valeurs nominales. En outre, donner une conclusion a-priori sur la variabilité des technologies de pointes utilisant de nouvelles géométries et des processus de fabrication différents, est une tâche difficile et de nouvelles méthodes d’estimation sont nécessaires. C’est une nécessité pour garantir les performances et la fiabilité des futurs circuits intégrés.

L’analyse statistique des variations de fabrication pour les dispositifs nanométriques nécessite une grande quantité de données à traiter. Ceci introduit un sérieux défi pour l’analyse de la variabilité des technologies émergentes en raison de l’absence de modèles de simulation rapide. D’une part, le développement de modèles compacts et précis implique de nombreux tests et des characterisation coûteuses sur les dispositifs fabriqués. D’autre part, les simulations de type TCAD, qui pourraient fournir des informations précises sur le comportement des dispositifs sont trop lentes pour produire un ensemble de données suffisamment large. Dans ces travaux de recherche, une méthode rapide de création de l’ensemble des données pour l’analyse de la variabilité est introduite. Cette méthode combine les simulations TCAD avec un algorithme d’apprentissage pour atténuer la complexité associée a la création de l’ensemble des données adaptées aux nano-dispositifs émergents.

Un autre défi de taille pour l’analyse de la variabilité des grands circuits est le nombre croissant

des sources de variation de fabrication dans les technologies nanométriques. Utiliser des modèles paramétrés devient une nécessité pour la conception de puces et de vérification. Toutefois, la grande dimensionnalité de l'espace des paramètres pose un sérieux problème. Malheureusement, les techniques de réduction de l'espace de paramèters disponibles ne peuvent être employées pour trois principales raisons : le manque d'exactitude, la dépendance de distribution des points de données, et enfin l'incompatibilité avec les simulateurs de dispositifs et de circuits. Nous proposons une nouvelle technique de sélection de paramètres pour la modélisation du la variabilité. La technique proposée prend efficacement en compte la nonlinéarité entre et les paramètres de performances qui est nécessaire pour l'analyse de variation de haute précision.
Des tests appropriés pour capturer les défauts de fabrication jouent un rôle important sur la qualité de circuits intégrés. Par rapport aux systèmes CMOS classiques, les nano-dispositifs émergents tels que CP-SiNWFETs ont différentes étapes de fabrication. Par conséquent, le type de défaut pour ces technologies est différent des dispositifs CMOS. Dans ce cas, des modèles de faute adaptés doivent être employés pour la détection des défauts dans les technologies émergentes. Dans cette recherche, nous avons identifié l'inefficacité des modèles de faute CMOS actuels pour la détection des défauts de fabrication de la technologie CP-SiNWFET. En considérant les étapes de fabrication, nous avons extrait les défauts de fabrication possibles, et nous avons ensuite proposé un modèle de faute pour cette technologie. Nous avons également fourni deux méthodes d'essai pour détecter les défauts de fabrication dans divers types de portes logiques à base de CP-SiNWFET. Enfin, nous avons utilisé le modèle de faute obtenu pour construire des circuits arithmétique insensibles aux défaillances avec beaucoup de propriétés supérieures par rapport à leurs concurrents.

Mots cléfs : Transistors à Polarité Contrôlable à base de Nanofils de Silicium , Simulation 3-D TCAD, Variabilité, Analyse Statistique, Modélisation du Circuit, Modèles de Faute, Circuits arithmétiques.

# Contents

# List of Figures

# List of Tables

# 1 Introduction

*Integrated Circuit* (IC) industry has experienced a tremendous growth during the last few decades. The demand for consumer electronic products has increased the market size of semiconductor industry. Companies require keeping the cost of products low while trying to add more functionality on a single chip. Over the past decades, this aim has been successful. Indeed, the number of transistors per chip has increased exponentially, as well captured by Moore's law. The downscaling trend has enabled semiconductor industries to both put more transistors on a constant chip area and boost the chip performance [1]. The capability of employing more transistors allows us to perform designs that satisfy specific goals like low power, high performance, and high memory capacity. As a result, the development of portable multi-function products such as tablets and smart phones has become feasible.

As the shrinkage passed below 90nm over a decade ago, the scaling trend faced several challenges such as *Short Channel Effect* (SCE) [2], lithography wavelength limitation [3], random dopant fluctuations [4]. SCE caused higher Off-state leakage currents of the devices, and therefore prevented the higher integration of transistors on a single die. Various techniques such as thinning gate oxide, using shallow source/drain junctions, and employing thin-film SOI MOSFET proposed to alleviate this difficulty [2]. The light wavelength used for the projection of device dimensions on a die was also a serious challenge since it is much longer than the actual dimensions of the device. Therefore, the projected device image was distorted, which complicated the control of variability on the critical dimensions. Researchers proposed various techniques to improve the lithography resolution, e.g., *Optical Proximity Correction* (OPC) [5]. OPC enhances the impact of light distortion by shifting edges or adding further polygons to the mask pattern. The challenge was successfully addressed through *Double Patterning Lithography* (DPL) [6]. In DPL the dense lithography patterns are separated into two sets, for which the exposure and the depth of focus is remarkably improved for each set.

As the technology downscaling pushed towards the 22nm node and beyond, the industry moved away from the planar structures which had successfully used very well for several decades [7]. The 14nm lithography is expected to heavily depend on DPL, *Triple Patterning Lithography* (TPL) [8], and even *Quadruple Patterning Lithography* (QPL) [9], as shown

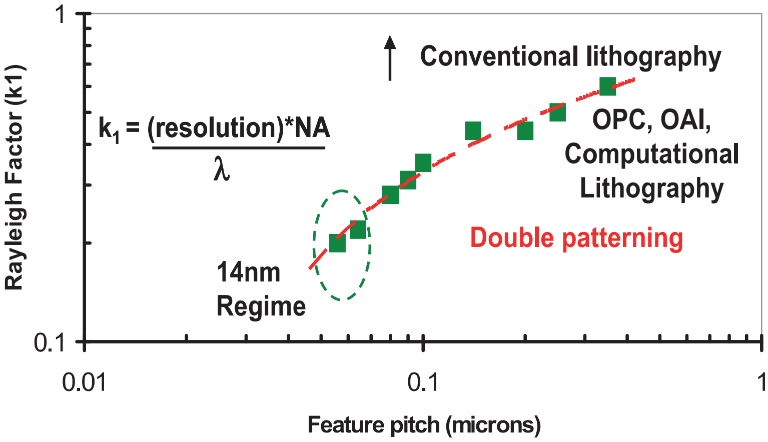


Figure 1.1: The complexity of lithography for 14nm scaling and beyond [7].

in Figure 1.1, due to the fact that the feature size is scaled faster than the wavelength of the lithography light. The cumulative impacts of emerging device structures, performance requirements, manufacturability, and DPL, will impose significant challenges for the implementation of efficient design solutions. Although some variability sources such as *Random Dopant Fluctuation* (RDF) is expected to be improved by exploiting novel devices [10], new sources of variability such as height variation in FinFET arises due to the complex geometry and process of future devices [7]. Additionally, variability that originates from *Line Edge Roughness* (LER) will become more challenging when the dimensions are decreased further. Moreover, heavily exploiting of DPL leads to color-dependent systematic shifts capacitance because of correlated/anti-correlated effects of miss-alignment [7].

As both device technology and lithography compete against scaling restrictions, the future technology generation will continue massive changes in both transistor structure and lithography procedures. In order to sustain the improving performance and efficiency trend according to Moore's law, some new alternative processing device and elements are being extensively researched. These investigations are in progress for advanced and evolutionary devices, which allows further scaling. The ever increasing complexity of structure and process for such devices, along with performance, reliability, and yield requirements will result in many new challenges for future circuit generations.

## 1.1 Evolution of Transistor Technologies

In this section, we briefly review a number of advanced and emerging devices, proposed by VLSI community, which try to address the currents issues to continue scaling of transistors.

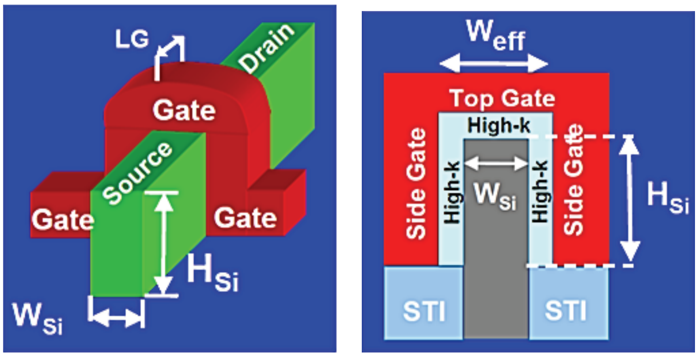
### 1.1.1 CMOS Advanced Devices

#### FD-SOI Technology

*Fully Depleted Silicon On Insulator* (FD-SOI) is a promising technology that addresses the SCE problem beyond 65 nm. FD-SOI transistors utilize an ultra thin channel deposited on a thin layer of insulator so-called *Buried Oxide* (BOX). This technology provides a better electrostatic control over the device channel and considerably reduces the leakage current when compared to the bulk technology (Figure 1.3). Accordingly, it presents a more efficient body-biasing control owing to no floating body. Thanks to the tiny height of the channel, the channel doping is not required. This can reduce the variation of $V_{Th}$ as well. After all, the SOI technology traditionally has been faced with the heat conduction from device channel to substrate due to the very low thermal conduction of the buried oxide. The generated heat degrades the carrier mobility leading to $V_{Th}$ shift.

#### FinFET Technology

As depicted in Figure 1.2, FinFETs employ vertical channels by which the leakage surface area between Source and Drain is greatly reduced. Using spacer lithography process enables Self-aligned gate deposition by which Thin-body (Fin width) devices is successfully fabricated [11]. The obtained device therefore has lower Source/Drain parasitic capacitance and better immunity to SCE. FinFET maintaining a steep subthreshold slope and represents good matching due to low doping concentration in the channel [12].

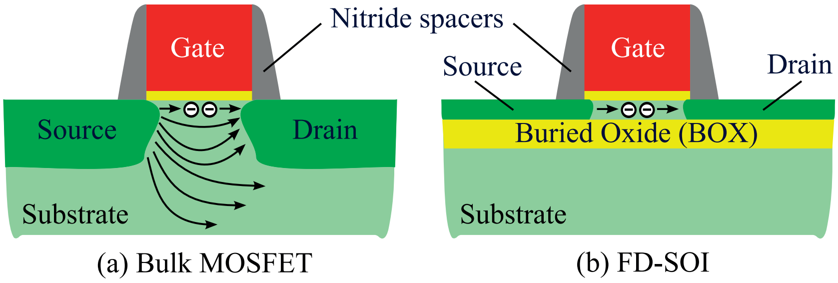


(a) Top view of a FinFET transistor [13].

(b) Cross view of a FinFET transistor [13]

Figure 1.2: 3-D structure of a FinFET Transistor.

The quantization of transistor width in FinFET circuit design is completely different than the continuous values of the planar technology. This property can significantly increase the complexity of the design. Therefore, new CAD tools are required to consider device width for custom circuit optimization. Indeed, this is a critical requirement for memory arrays and analog circuits.



(a) Bulk MOSFET

(b) FD-SOI

Figure 1.3: Comparison of the of the carrier transportation in the channel: (a) bulk MOSFET, and (b) FD-SOI [14].

### Multi-Gate Technology

Improving carrier mobility can be accomplished either by exploiting materials with higher mobility, i.e., *GaAs* or by modifying the device geometry to improve the electrostatic control over the device channel. Utilizing more than one gate, i.e., *Double-Gate* (DG) FETs, provides better channel modulation and reduces the SCE effects. Extra gate(s) can be used for $V_{Th}$ adjustment and leakage current control; however, these significantly increase fabrication complexity and requires introspection in circuit design and verification (Figure 1.4). However, the complexity can be mitigated by going to FinFET structures.

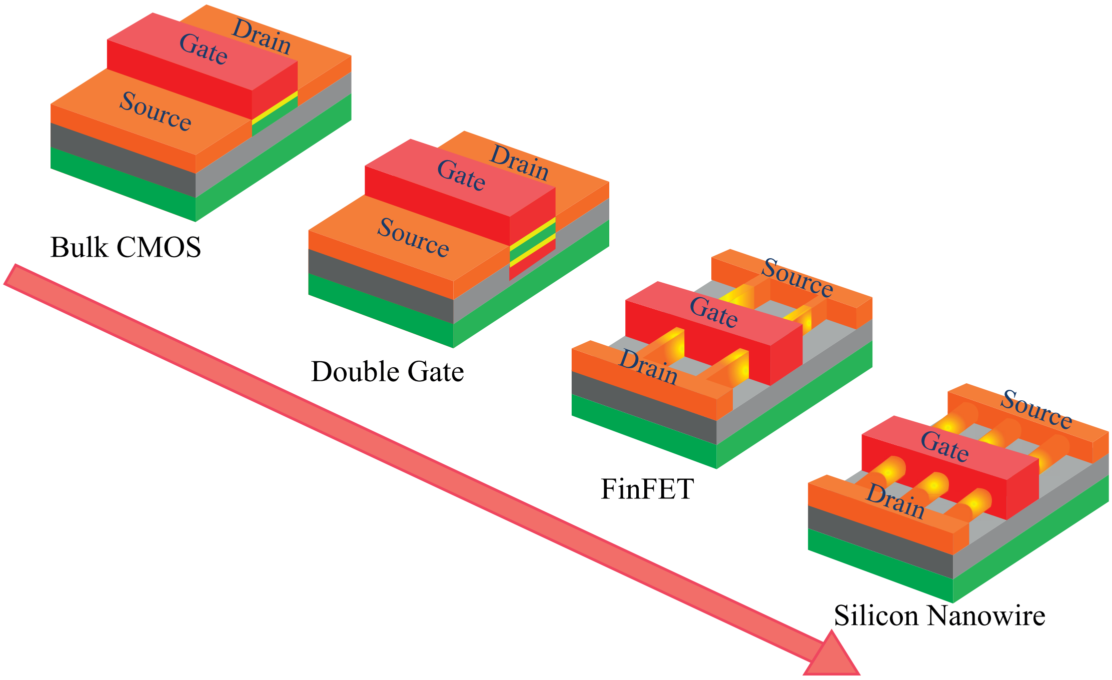


Figure 1.4: The increase of geometrical complexities for future nano-devices [11].

### 1.1.2 Emerging Devices

#### Silicon Nanowire FETs

As a promising candidate of replacing current CMOS technology, *Silicon Nanowire FET* (SiNWFET) has received considerable attention. SiNWFET employs a very thin nanowire as the device channel which fully covered by the gate. This *Gate-All-Around* (GAA) structure provides great electrostatic control and high device $On/Off$ current ratio, using the best feasible control of the gate over the channel (Figure 1.4). Accordingly, it is cost effective and CMOS compatible manufacturing process, promote it as high interest to continue scaling below 10 nm [15]. Because of the tiny dimensions of the channel, variability plays an important role on the performance and the functionality of SiNWFETs. Various sources of variation, i.e, LER, can cause carrier scattering, and therefore change the device characteristics such as *On*-current ($I_{On}$) and $V_{Th}$. Therefore, controlling the variability is an important feature to make this technology achievable.

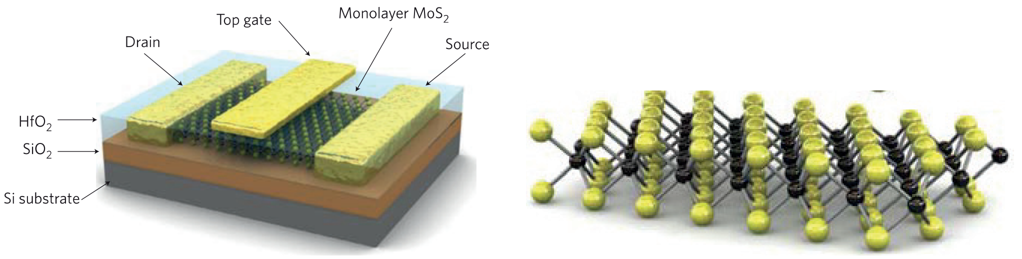
#### Carbon Nanotube FETs

*Carbon Nanotube FET* (CNTFET) possesses superior characteristics for future integrated circuits, in terms of carrier mobility and heat conductivity. Carbon Nanotubes are rolled up cylinders made of graphene. Depending on their physical configurations, CNTs can show semiconductor properties. Thanks to the cylinder structure and strong covalent bound among carbon atoms, CNTs have quasi-ballistic carrier transport, even for large distances [16]. They have high thermal conductivity, along with no boundary scattering [17]. These properties make them interesting for fabricating very fast switches. However, they suffer from several challenges including lifetime degradation, complication of mass production, and reliability issues [18].

#### Two-Dimensional Transistors

Two-dimensional materials are attractive to be exploited for future nano-devices because fabrication of complex structures is simple [19]. Recently, graphene has been intensively researched owing to its superior mobility. However, graphene does not have a bandgap which is necessary for building semiconductor devices. Engineering a bandgap for graphene considerably reduces its mobility and imposes fabrication complexity. It has been shown that for other single-layer materials like $MoS_2$, the fabricated device has either an acceptable bandgap while it has superior mobility [20]. Similar to previous mentioned devices, graphene based FETs suffer from the same fundamental problems, i.e., gate and source-to-drain tunneling, variability, parasitic resistances/capacitances [19]. Figure 1.5 depicts the structure of a typical $MoS_2$ transistor and its channel structure.

Along with the several promising opportunities of emerging nanoscaled technologies for future ICs, a number of problems such as process variation, manufacturing defects, and reliability



(a) A typical $MoS_2$ transistor [20].

(b) The channel structure of a typical $MoS_2$ [20].

Figure 1.5: Three-dimensional structure of a typical $MoS_2$ transistor and the channel structure.

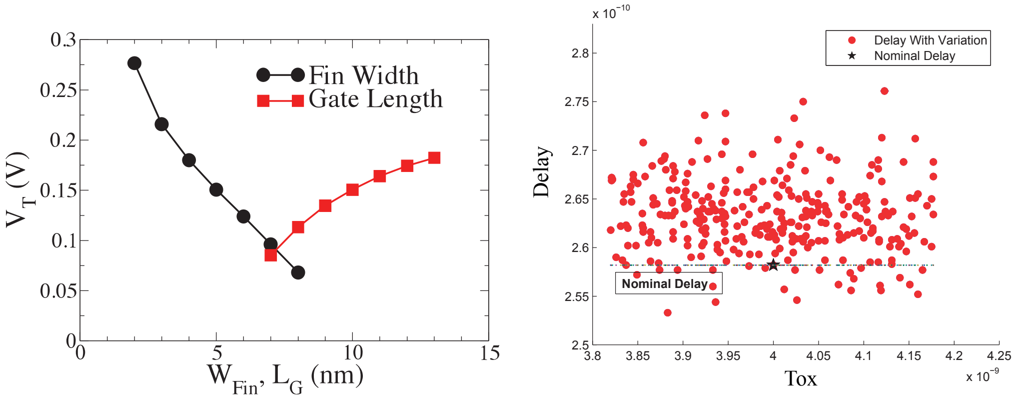
remains as a serious challenge. In the following section, we briefly review the major technical challenges of the emerging technologies from device fabrication level to system level.

## 1.2 Challenges of Emerging Technologies

Moving towards emerging nonascaled devices for mass production of ICs, manufacturers and designers are confronted with many challenges for design, verification, and test. Manufacturing uncertainties and process variation along with fabrication irregularities and defects adversely affect the circuit's yield and dependability for deeply scaled technologies. To reach high yield products, the issues of manufacturing process have to be considered from the early steps of the design. These issues originate either form physical flaws, i.e., *Gate Oxide Short* (GOS) or variability of the critical parameters, i.e., gate length. With the increasing fabrication uncertainty for future ICs, the use of efficient statistical variability analysis for precise yield estimation is inevitable. Moreover, developing precise defect models through the analysis of manufacturing process is necessary for enhancing product test. In the following, we briefly review the various sources of manufacturing imperfections and variations, which are relevant to our variation and defect analysis.

### 1.2.1 Process Variation

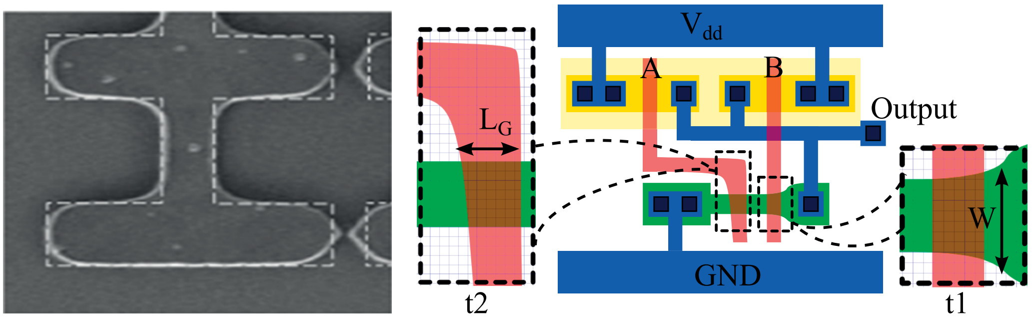
*Process Variation* (PV) is referred to unwanted deviations from the nominal or expected values of physical and structural parameters for fabricated transistors, i.e., gate length, $V_{Th}$, gate oxide thickness, and number of dopant atoms. PV therefore changes the performance or even the functionality of the fabricated devices. Figure 1.6a depicts the impact of gate length/width variations on the $V_{Th}$ of FinFET in 10 nm technology. The $V_{Th}$ sensitivity on length and width shows a diverse trend. Such device variations then impact circuits' performance or functionality and lead to yield loss, as shown in Figure 1.6a.



(a) The impact of gate length and width variations on the $V_{Th}$ of 10 nm FinFET [21].

(b) The impact of oxide thickness ($T_{ox}$) variation on the timing of the ISCAS89 C71 benchmark circuit in [22].

Figure 1.6: The impact of parameter variation on the performance of a device and circuit.



(a) Corner rounding as a result of mask restriction [23].

(b) The impact of corner rounding on the gate length and transistor width variations.

Figure 1.7: Mask patterning restriction as a source of variation.

### Variation Sources

Various sources of variation contribute to change the characteristics of the fabricated devices. These sources either significantly impact the critical dimension of a transistor including Gate Length ($L_G$), and Gate Oxide Thickness ($T_{Ox}$) or affect the device physical properties like the number of dopants in Source/Drain.

The main source of critical dimension variation consists of mask positioning and alinement, photoresist effect, and chemical/mechanical polishing. Mask pattern restriction is a result of low pass filter behavior of lithography lens, and can make a major difference between the real layout and the printed image on die. This anomaly can be seen as corner rounding and line end shortening (Figure 1.7). Both of these effects, can seriously impact the behavior of the transistor. As illustrated by Figure 1.7b, the corner rounding can change the effective width of the transistor (t1) leading to remarkable increase of transistor transistor leakage

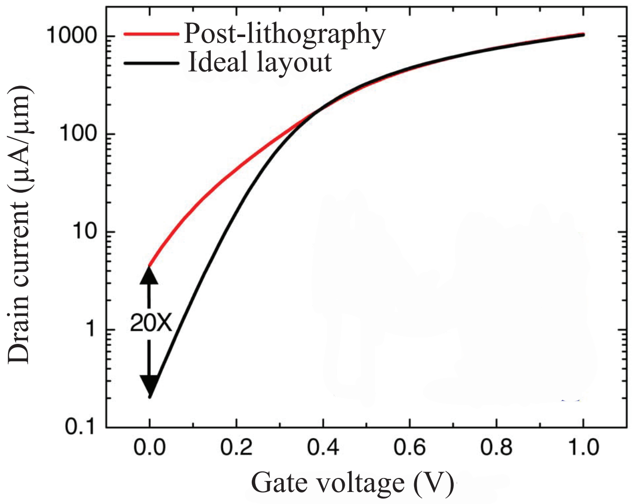


Figure 1.8: Mask patterning restriction causes rounding effect on the transistor gate and vastly increases the leakage current in 65 nm technology [23].

(Figure 1.8) [23]. Beside, the effective $L_G$ is increased by the corner rounding for transistor (t2). As another source of variation, chemical resist effect can adversely affect edge roughness of the patterns. Chemical photoresist processing is a series of chemical reactions on the wafer surface and it causes random variation on the edge of patterns so-called *Line Edge Roughness* (LER). LER does not scale with the technology and becomes a serious barrier for further scaling beyond 45 nm. Along with the mentioned variability sources, chemical and mechanical polishing, which is used to remove excess parts of the device such as extra parts of polysilicon and oxide, can change the effective width and length of the device.

On the other hand, there exist other sources of variation, such as *Random Dopant Fluctuations* (RDF), which impact the physical properties of the transistors. RDF is generated by the variability on the number and location of dopant atoms in the device channel. This phenomenon becomes crucial for deeply nanoscaled devices since the charge discreteness prevent reaching a uniform doping concentration.

### Types of Variations

Fabrication of the ICs involves usually many steps, in which several sources of variation affect the final product. The variation may be caused by photolithography equipment fluctuation, environmental variations, and material properties. Discovering all the actual variation sources of each IC is almost impossible. In order to be able to analyze the performance and functionality of circuits, designers classify the various types of process variations to make the analysis tractable.

Variations can be generally divided into two categories of *inter-die* and *intra-die* variations.

*Inter-die* variations refer to common variations within the die and include different levels, i.e., die to die, wafer to wafer, and lot to lot [24]. These variations can be taken into account by assigning a random variable to each die, by which a deviation from nominal value for a device parameter is represented. Fluctuations such as variation of critical dimensions, owing to mask misalignment or unequal exposure time, are considered in this category. Here, the number of necessary variables for *inter-die* variation modeling of a die is limited by the number of variation parameters of a device. *Intra-die* variations, on the other way, refer to random and spatial correlated variations within a die. These variations are becoming very challenging at advanced technology node since the overall impact of fluctuations can result in substantial yield loss. In contrast to *inter-die* variations, the number of required variables for *intra-die* variation modeling can be equal to the total number of variation parameters in a die. Consequently, high dimensional parameter space of large circuit makes the *intra-die* variation modeling very challenging.

### Variation Modeling

Measuring performance uncertainties, especially for *inter-die* variations, has been traditionally performed by corner-based analysis. Corner models explore additional extra device models such as *fast* and *slow* and estimate the boundary of performance deviation. The utilization of this model for variability analysis of nanoscale technologies has become very limited. This model simply assumes that the all components of a circuit are correlated and they reach to worst-case at the same time. This can lead to pessimistic performance estimation. Moreover, this model does not take to account the *intra-die* variations, for which it may result in erroneous performance estimation. Although a number of modifications have been proposed to incorporate *intra-die* variations into corner model, a huge number of corners are necessary for precise performance analysis that makes it impractical for large circuits.

In contrast to the corner model, statistical methods are able to deal with *intra-die* variations, on-chip spatial correlations and *inter-die* variations, and therefore they are widely used for circuit performance analysis. *Monte Carlo* (MC) simulation is among the most important methods for statistical circuit analysis. MC numerically computes the distribution of the desired performance parameter through repeated sampling. In each iteration, numerical samples of the process parameters are randomly selected and injected to a device-level or circuit-level simulator. The obtained values are then aggregated to form the probability distribution of the desired performance parameter. One of the good points about MC simulation is that the accuracy of the obtained result is independent of the problem dimension. As a result, we do not need to increase the number of the input samples as the problem dimension increases. The major drawback of this method is that reaching an acceptable precision requires a huge number of sampling point. As the simulation time is a serious bottleneck in performance analysis, the practical application of MC is limited to devices and small-size circuits.

To overcome such computational costs, two other classes of statistical performance analysis are commonly utilized in VLSI community. First, *response surface* performance modeling which tries to reduce the problem complexity by estimating the desired performance using a kernel (i.e., polynomial) regressor. Second, multivariate dimension reduction methods that reduce the problem dimensions either by removing correlation among process parameters or finding the most dominant process parameters. Although these methods can obtain the statistical dependencies between device parameter variations and circuit performance, they require an extensive revision to be applicable for future emerging nanotechnologies.

### Variation Modeling Requirements for Emerging Technologies

To perform an efficient PV analysis for emerging nanotechnologies, the following requirements for statistical methods are necessary to be considered.

- Simulation complexity: Precise circuit simulation in current CMOS technology is performed through compact models. Due to the tiny dimension and complex geometry of the future nano-devices, statistical compact models are inevitable for PV analysis. The development and verification of these models requires both experimental data of the fabricated devices and enormous amount of TCAD simulations. TCAD uses physical equations of the target device, and find its characteristics through numerical solution of the equations. This substantially increases the complexity of model computations so that the TCAD simulation may take a couple of days or even months for a single device. Therefore, methodologies that reduce the computational complexity of variation analysis for emerging future devices are becoming very important.

- Nonlinearity between process and performance parameters: The process and performance parameters have strong nonlinear dependencies. These dependencies and correlations among these parameters are key factors to simplify variability models. The correlations are generally divided into two types: *intra-set* correlations which refer to the correlations among process parameters, and *inter-set* correlations which refer to the correlations between process and performance parameters. Reducing complexity is obtained by removing these correlations. Consequently, the nonlinear dependencies of parameters should be taken to account for optimally reducing the complexities of variability models.

- Statistical distributions: Most statistical variation analyses have strong assumptions on the probability distribution of the process parameters. They assume that the distributions of the all parameters are Gaussian or Non-Gaussian to avoid more computational complexities. Although these assumptions were acceptable for bulk CMOS, they are not valid for emerging nano-devices. This can negatively impact the accuracy of such analyses. Thus, statistical methods that can address a combination of different probability distributions are highly required.

### 1.2.2 Manufacturing Defects

Faulty fabrication can cause manufacturing defects, and finally lead to circuit failure. These defects result in faulty circuits, where the behavior may differ from the correct one. As a result, circuit testing is necessary to ensure that the functionalities of fabricated circuits are correct. In order to develop high-coverage test, faults should be modeled based on their behavior. The importance of fault modeling is highlighted by the fact that a defect may lead to a different behavior in various technologies. For instance, an open defect in a bulk CMOS transistor results in switching failure, while the same defect on the back gate of a Double-Gate transistor only leads to delay and leakage variation. As emerging nano-devices have different geometrical structures and dissimilar physics of operation rather than bulk CMOS technologies, they need a precise defect and fault modeling for test.

The faults can be classified in two groups. The first group is parametric faults that originate from process parameter variations which reviewed in the previous section. The second group is defective faults that arise from unwanted connections or isolations in different parts of a transistor due to the lack of control during manufacturing. The type of defective faults, in a specific technology, depends on the different steps of fabrication along with the materials used in each step. Both groups are becoming very challenging in deeply-nanoscaled technologies, since the control on the physical and structural parameters of devices is reduced by dimensional downscaling. Using cross section analysis to extract the characteristics of the defects for fabricated circuits is very expensive, and sometimes is almost impossible. Therefore, inductive fault analysis should be used for defect and fault modeling of the nano-devices.

#### The Need for Defect and Fault Modeling in Nano Era

Defect tolerant design for future technologies is faced by several challenges. Changing the process flow and the materials can change the population and distribution of the defects. Moreover, the different design of logic cells and functional components may change the sensitivity of the design for a specific group of faults. Consequently, making progress in reliable circuit design and test necessitates both to develop efficient fault models for emerging technologies and to revise the current test method according to the related fault model models and circuits sensitivities.

- Process and layout dependent defects: The impact of a manufacturing defect can be influenced by the materials, structure of logic gates or circuits layout. The structure of logic gates in various technologies such as FinFET, and *Controllable-Polarity Silicon Nanowire FETs* (CP-SiNWFETs) are significantly different. This information is essential for extracting the effects of physical failures on circuit functionality. Moreover, the specific faults for each technology can be ranked according to its importance or probability of occurrence. We require this information for both increase the yield and keep the overheads of fault-tolerant design as low as possible.

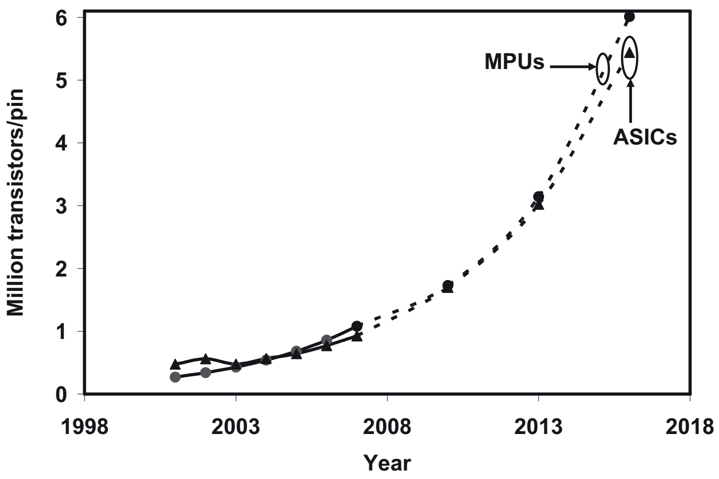


Figure 1.9: The exponential increase of the number of transistor per IO pin for microprocessors and ASICs as a challenge of test for future technologies [25].

- Test complexity: For many years, the aim of IC testing was to verify its functionality by the minimum number of input test vectors to guarantee full fault coverage. Moving towards nano era, test development needs to address several serious challenges. Accessing to transistors on a chip becomes more complicated owing to the exponential growth of the devices rather than the IO pins (Figure 1.9). Second, circuit sensitivity variation in each technology can make a defect into a very aggressive one while it was negligible in the past technologies. Moreover, the new manufacturing PVs that influence the functionalities in each technology should be covered by test algorithms. Consequently, all aspects of the test methodologies, including defect and fault modeling, input test generation, and coverage evaluation should handle these problems.

## 1.3 Dealing with Process Variation and Fabrication Defects

The inadequacies in the traditional methodologies to deal with the manufacturing threats in nanometer technologies, have led to advanced DFM techniques. As stated in [26], design flows for nano-devices require tighter integration between design and manufacturing. Unlike earlier technologies, design and manufacturing cannot remain independent of each other.

In order to address the aforementioned issues, statistical design approach has been investigated as an effective method to ensure certain yield criteria [27]. In this approach, the design space is explored to optimize performance parameters (such as power, reliability) to meet timing and power requirements. However, most of these investigations cannot handle both objectives simultaneously. Unfortunately, most of these method cannot be utilized for emerging technologies owing to the lack of fast simulation models like compact models for emerging nano devices.

To increase the yield and fault tolerance, some methods also point towards adding extra components on the chip, irrespective of fault detection, as a design based tolerance. This is a pessimistic technique adding unnecessary area on chip. Despite the high redundancy overhead, the nanocomputing community has proposed the use of $T$riple- and $N-$modular redundancy techniques and shown that they provide high reliability even in the presence of high device failure [28]. Such techniques increase the reliability of designs, but the main assumption made in these redundancy techniques is that the voter block is free from the same failures that the rest of the design is facing. While such assumption seems true for designs based on traditional silicon technologies, this assumption will not hold for designs based on nanotechnology hindering the applicability of $T$riple- and $N-$modular redundancy techniques for nano era designs.

Figure 1.10 summarizes the main difficulties and hindrances have been highlighted in the reports of *International Technology Roadmap for Semiconductors* (ITRS). Addressing these challenges plays a key role on the performance and the yield of future circuits. A number of these challenges are addressed in this thesis.

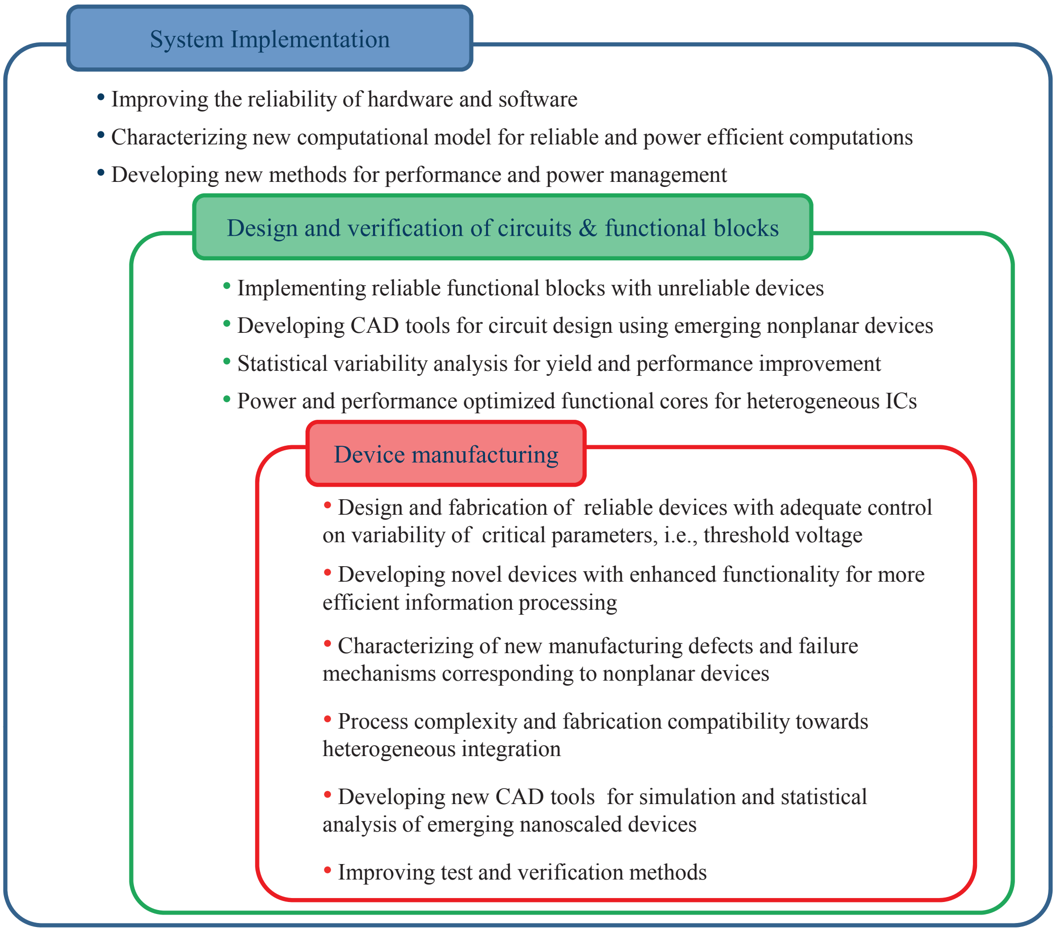


Figure 1.10: Challenges of emerging technologies for future integrated circuits.

## 1.4 Thesis Contribution

The contribution of this thesis consists of five major parts:

- **Robustness analysis for controllable-polarity silicon nanowires:** this part of research includes the detection of important source of variations, and impact of process parameters variation on the performance of the device and logic gates. Here, we investigate the variation of critical parameters of the nanowires, such as gate length, on the characteristics of the device such as *On*-current, channel leakage current, threshold voltage and subthreshold slope.

- **Learning-based methodology to reduce the computational complexity for device level PV analysis:** Generating data set for variability analysis of novel devices is a difficult task due to the lack of fast and precise compact model. To address this problem, we combine TCAD simulations with a learning algorithm and propose a methodology to speed up the data set generation for variability analysis.

- **A statistical method for fast PV analysis of emerging technologies:** The increasing number of parameters for variability modeling of integrated circuits imposes a remarkable computational complexity. We propose a parameter selection technique than can effectively select the most important parameters of the target circuit for variability analysis. Therefore, a considerable speed up is obtained.

- **Inductive fault analysis:** We perform inductive fault analysis on controllable-polarity silicon nanowires to find an appropriate defect model that helps to test development for this technology. We extract the possible defects from device manufacturing process steps and analysis the impact of all extracted defects for the performance and functionality of the device. Based on the results, we propose a guide line for test of the logic circuits in this technology.

- **Fault tolerant arithmetic circuit design:** we exploit the obtained defective fault model for designing fault tolerant arithmetic circuits. We investigate the behavior of adder components in presence of fault and then we propose a fault tolerant novel structure for carry-ripple adder.

## 1.5 Thesis Organization

The reminder of this thesis is organized as follows. In Chapter 2, an overview of the CP-SiNWFETs as an emerging technology is presented. The promising characteristics of this technology for device and logic circuit development are then outlined. The fabrication process, simulation methodology, and the structure of logic gates are also reviewed. Finally the major sources of variation in this technology are investigated and the results of variation on device and logic cells are then represented.

In Chapter 3, a learning based methodology that can substantially reduce the time complexity and cost of variation analysis for emerging technologies is introduced. The components of the proposed methodology such as a prediction module are then represented. Finally, experimental results are presented that show the proposed methodology effectively accelerates TCAD-based PV simulations close to compact-model-based simulations.

Chapter 4 introduces a feature selection method to reduce the circuit modeling complexity. Next, the mathematical description of this method is explained in detail. The characteristics of this feature selection for VLSI modeling and simulation is then discussed. Finally, the application of this method is demonstrated in digital circuit timing analysis in both FinFET and CP-SiNWFET technologies.

In Chapter 5, the possible manufacturing defects of CP-SiNWFET technology is explored through analyzing the fabrication steps and the layout structure of logic gates. Afterwards, the impacts of obtained defects are investigated on the performance and the functionality of CP-SiNWFET logic gates. Out of the results, the current fault model is extended to a new a hybrid model can be efficiently used to test the logic circuits in this technology. The inefficiency of current CMOS test methods for covering all faults in CP-SiNWFET technology is then shown. Finally, an appropriate test method to capture such faults is proposed as well.

Chapter 6 is dedicated to the intrinsic capability of fault tolerate in CP-SiNWFET technology. Based on this result, a fault tolerant and scalable adder is introduced. Next, the robustness of the proposed circuit is investigated, and finally, the efficiency of proposed circuit is verified with experimental results.

Chapter 7 summarizes the contribution of this thesis and concludes it. In addition, a number of ideas for future work are provided.

Overall, variability and fabrication defects are the two major challenges of the deeply scaled integrated circuits. In this thesis, we investigate these two topics for CP-SiNWFETs as an emerging technology for future logic circuits. The robustness of this technology is carefully analyzed in presence of these difficulties. Then, several techniques to handle these problems are proposed. The outcome of this thesis will help to achieve more dependable circuits in nano era.

# 2 Controllable-Polarity Silicon Nanowire FET: Background and Robustness Analysis

## 2.1 Introduction

*Silicon NanoWires* (SiNWs) have attracted growing interest as a promising candidate for future ICs owing to their unique physical and electrical properties. SiNWs with *Gate-All-Around* (GAA) structures [29] provide an even better electrostatic control over the channel (and reduce leakage current) as compared to current technologies such as FinFET [30] and FDSOI [31]. Since short-channel effect and leakage power are the most significant challenges of further feature-size scaling, SiNWs are very a promising technology to pursue the scaling trend towards higher performance and functionality. In this chapter, we introduce the SiNWs with *Controllable-Polarity* (CP) characteristics. We review the device characteristics, and then we introduce our framework for device and circuit simulation in this technology. Next, we look at fabrication process of these devices and then we investigate the structure of logic gates in this technology. This is required for variability analysis, defect modeling and fault-tolerant circuit design. Finally, we analysis the robustness of this technology in presence of manufacturing variability.

Beyond 45 nm, many devices exhibit ambipolar characteristic at source and drain contacts such as SiNWs [32], carbon nanotubes [33], and graphene [34]. These devices have ambipolar behavior, i.e., they support the flow of both $n$-type and $p$-type carriers. While ambipolarity is usually suppressed by fabrication process to provide unipolar devices [35], it can be used to enhance the logic functionality, i.e., the capability of implementing more complex functions using smaller number of transistors [36].

Ambipolar conduction of the nanoscaled devices can be controlled by adjusting the device polarity online. Such transistors with Controllable-Polarity have been successfully implemented in *Silicon NanoWires* (SiNWs) [37, 38], Carbon Nanotube [39], graphene [40], and FinFET [41]. These devices have been successfully used for the fabrication of controllable-polarity logic gates that provide compact hardware realization with remarkable circuit design flexibility. In such transistors, one electrode gate, the *Control Gate* (CG), works like conventional MOSFETs, and provides the conduction by controlling potential barriers. At least one another electrode gate, the *Polarity Gate* ($PG$), is needed to control the $n$-type or $p$-type characteristics of the device. Indeed, the type of carriers that flow in the device channel is adjustable through the applied voltage on $PG$. Among various materials used for CP transistors, SiNWs have a CMOS-compatible fabrication process. Different architectures have been proposed for their implementations [37, 38, 42, 43]. As an example, Figure 2.1 represents a *Three-Independent-Gate* (TIG) SiNWFET with a $CG$ and two $PG$s ($PG_S$ and $PG_D$). Here, the side regions ($PG_S$ and $PG_D$) determine the majority carriers through adjusting the Schottky barrier height at the source/drain junctions. Thus, the device can exhibit controllable $n$-type and $p$-type characteristics. DG-SiNWFETs [37] exploit a similar structure than TIG-SiNWFETs where the two polarity gates are connected together. Reconfigurable SiNW [42] is another example in which only one polarity gate is utilized for configuration of transistor to desired polarity.

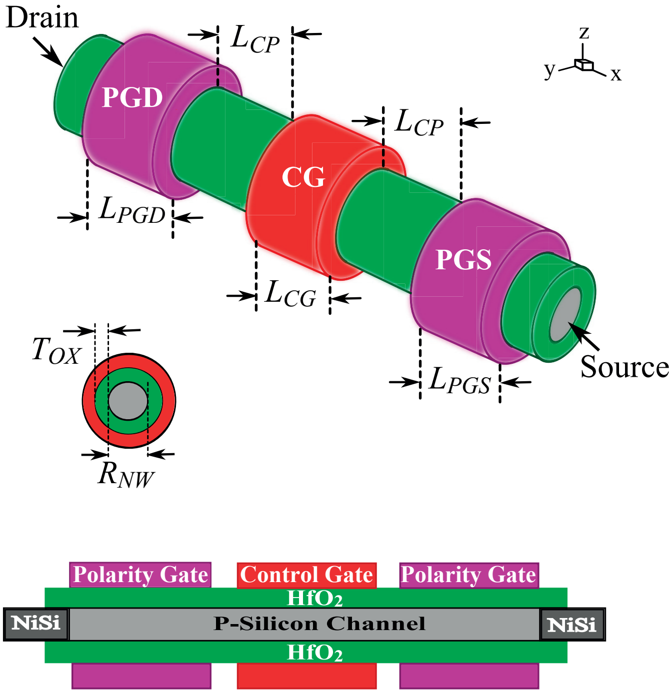
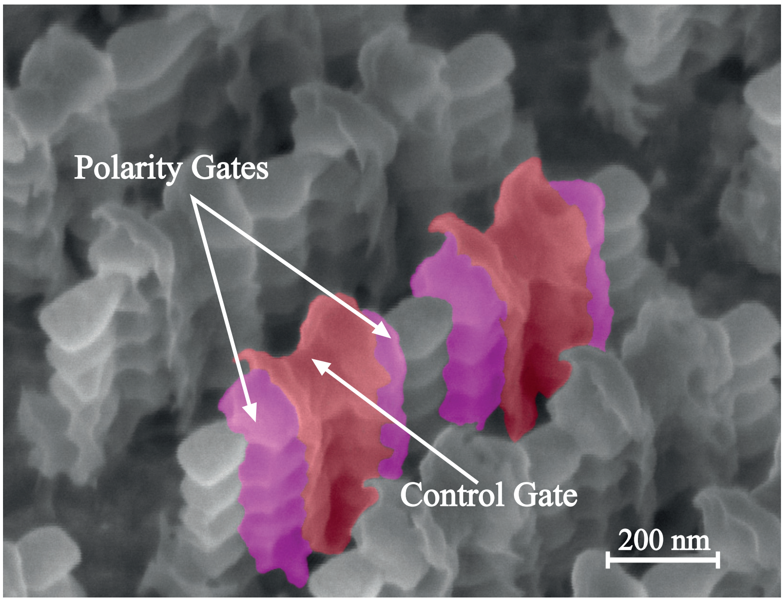
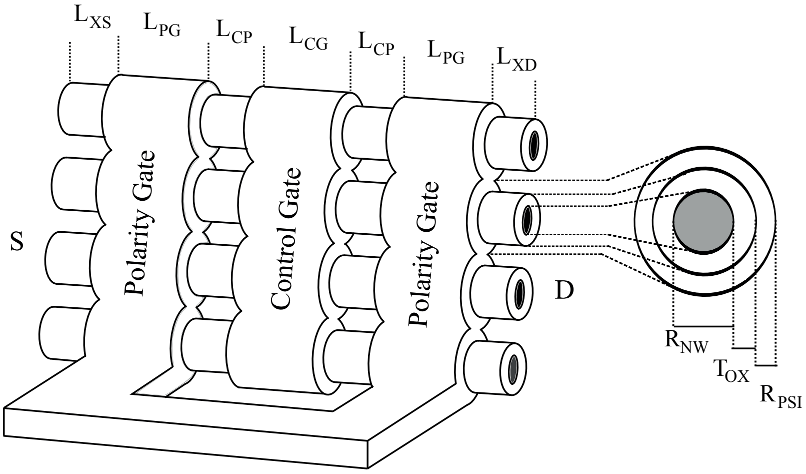


Figure 2.1: The 3D structure of TIG-SiNWFET along with the corresponding geometrical parameters.



(a) The SEM figure of the fabricated DG-SiNWFET [37]. The red and purple colors highlight the control and polarity gates respectively.



(b) The 3D model of a DG-SiNWFET along with the geometrical parameters. Here, the both polarity gate are connected to each other.

Figure 2.2: Vertically stacked DG-SiNWFET as a CP-SiNW.

## 2.2 Fabrication Process

The TIG-SiNWFET devices are fabricated in a top-down approach. Table 2.1 summarizes the fabrications process of the device along with the outcome of each step. The Bosch etching process [44] is utilized to form the nanowire stack. An high-$\kappa$ gate dielectric is then deposited over each patterned nanowire and provides a thin oxide layer ($\leq 5nm$) around the channel.

Oxidation process is followed by a conformal metal deposition to shape the polarity gates around the nanowires. Finally, the control gate structure is self-aligned to polarity gates. Thus, a three-gate device is obtained in which polarity gates are electrically isolated by the controlled deposition of the gate oxide.

## 2.3 Electrical Simulation

In order to first model the possible defects of this technology, and then to investigate their impact on the performance and the functionality of logic gates, we need to setup a simulation framework that brings together device and circuit simulators. Consequently, a two-step simulation environment, that integrates the Sentaurus TCAD and HSPICE simulators into a single framework, is used to facilitate high-level simulations. First, we build a TCAD model of the TIG-SiNWFET, for which the I-V curves are calibrated with those of our fabricated devices. The typical parameters of the TIG-SiNWFET, shown in the Figure 2.1, are listed in Table 2.1.

Table 2.1: TIG-SiNWFET structural and physical parameters.

| Device Parameter | Value |
|---|---|
| $Length of Control Gate (L_{CG})$ | $22nm$ |
| $Length of Polarity Gates (L_{PGS}, L_{PGD})$ | $22nm$ |
| $Length of Spacer (L_{CP})$ | $18nm$ |
| $Channel Dopping Concentaration$ | $10^{15}cm^{-3}$ |
| $Schottky Barrier Height$ | $0.41eV$ |
| $Oxide Thickness (T_{Ox})$ | $5.1nm$ |
| $Radius of NanoWire (R_{NW})$ | $7.5nm$ |

Next, circuit level simulations are realized by a simple compact model in Verilog-A (Figure 2.3). The result of the TCAD simulations from the previous step, makes a look-up table that characterizes the channel conductivity as a function of the $V_{CG}$, $V_{PGS}$, and $V_{PGD}$. A sample of such table model is depicted by Figure 2.3. Moreover, it provides the value of the parasitic capacitance among various terminals and the access resistance corresponding to the source and drain. This model is used in our simulations to efficiently implement the functional behavior of the TIG-SiNWFET. The general flow of the circuit simulation is illustrated in Figure 2.4.

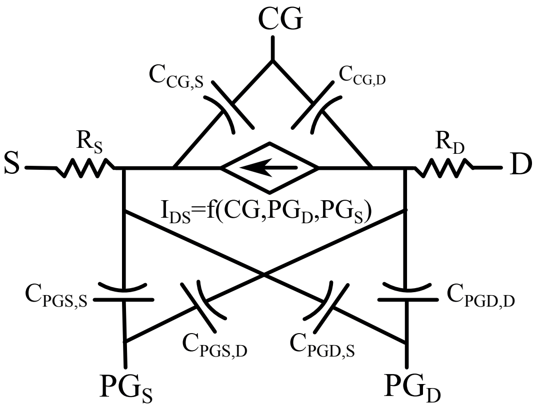


Figure 2.3: The lumped model of the device and a sample of table model.

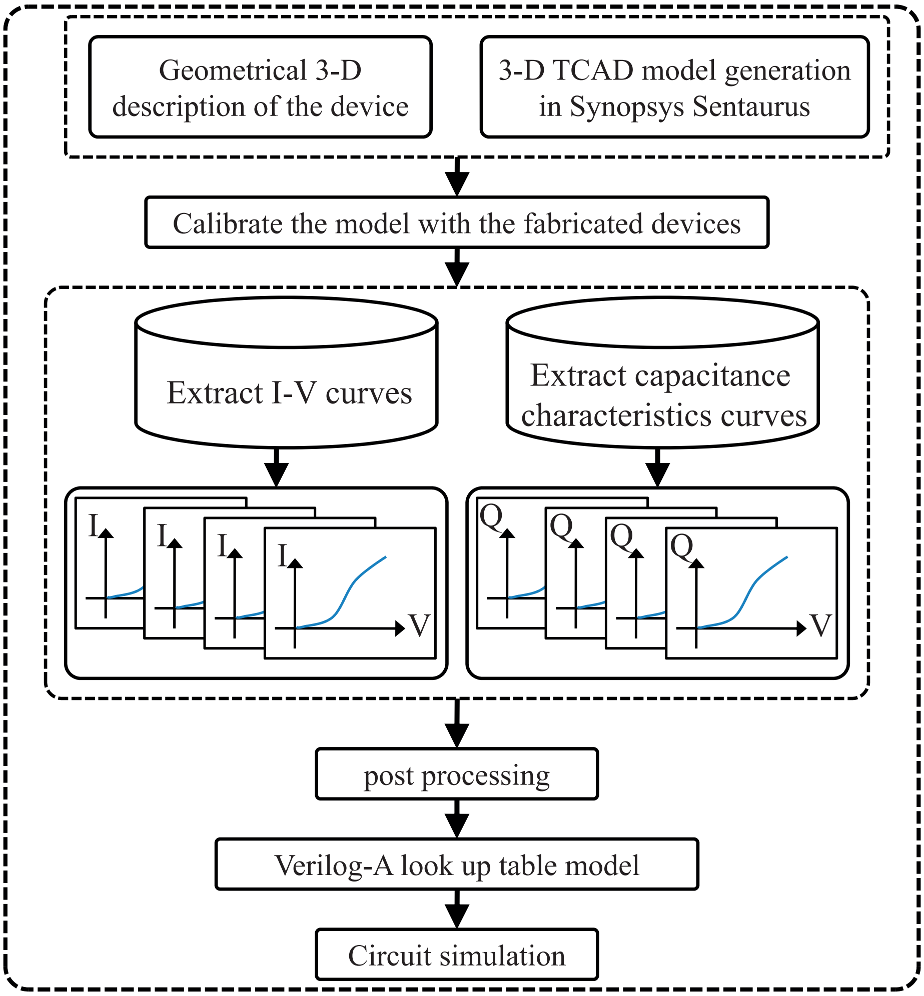
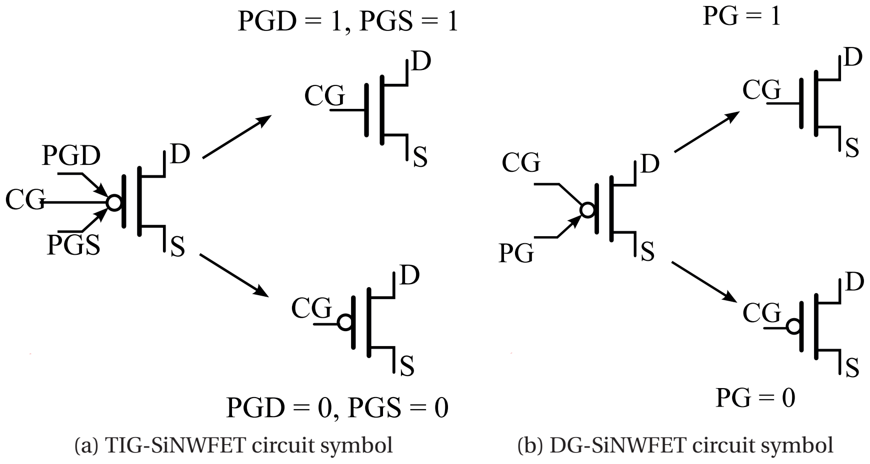


Figure 2.4: Circuit simulation methodology for controllable-polarity silicon nanowire.

Figure 2.5 illustrates the circuit symbol of TIG-SiNWFET and DG-SiNWFET. When the polarity gate(s) is polarized to $V_{dd}$, the transistor can be replaced by a $n$-type device. Accordingly, when the polarity gate(s) is polarized to $GND$, the device can be substituted by a $p$-type one.



(a) TIG-SiNWFET circuit symbol (b) DG-SiNWFET circuit symbol

Figure 2.5: The circuit symbols for various controllable-polarity transistors.

Figure 2.6 also represents how polarity gate bias can be utilized for realization of either $n$- or $p$-type devices. In comparison with other unipolar technologies, this characteristic enriches this technology to reach a very compact implementation of binate logic cells, as detailed in the following.

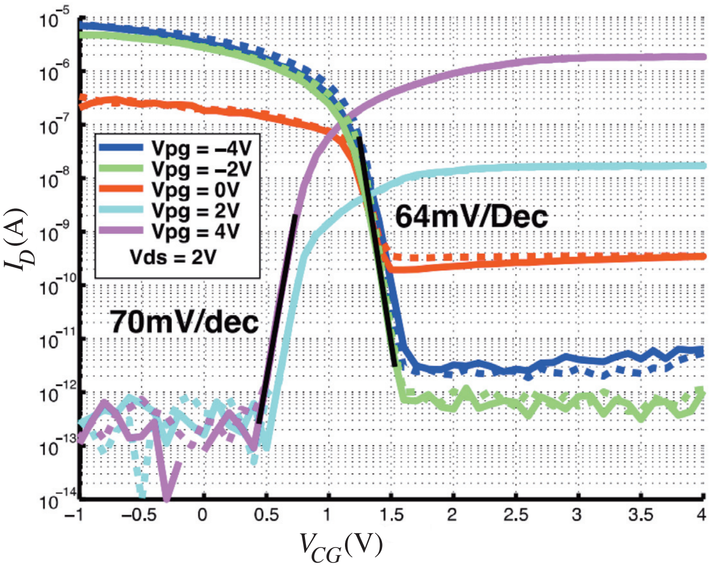


Figure 2.6: $I-V$ of a fabricated controllable-polarity silicon nanowire [37].

## 2.4 Logic Gate Realization in CP-SiNWFET Technology

The CP transistors as configurable structures can be efficiently utilized to implement logic gates. The polarity terminals are used to select the device polarity. According to the different configuration of polarity gates, the CP logic gates are divided into two categories.

The first group, called *Static Polarity* (SP), is characterized by the $PG$ gates directly connected to either power supply ($V_{dd}$) or ground ($GND$) rails to provide the desired polarity. In SP logic gates, the polarity of all devices remains the same during the whole device life-time. Figure 2.7 illustrates the three examples of SP logic gates (INV, NAND, and NOR gates) realized in TIG-SiNWFET technology.

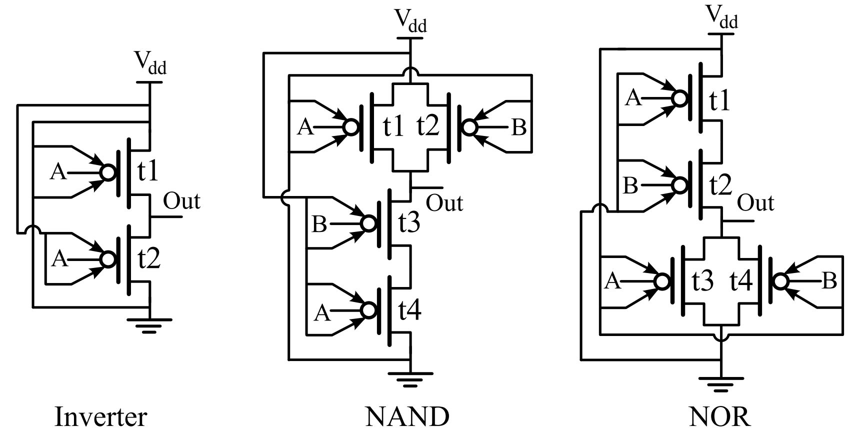


Figure 2.7: Logic gate realization in TIG-SiNWFET. Inverter, NAND, and NOR gates are the examples of the SP logic gates.

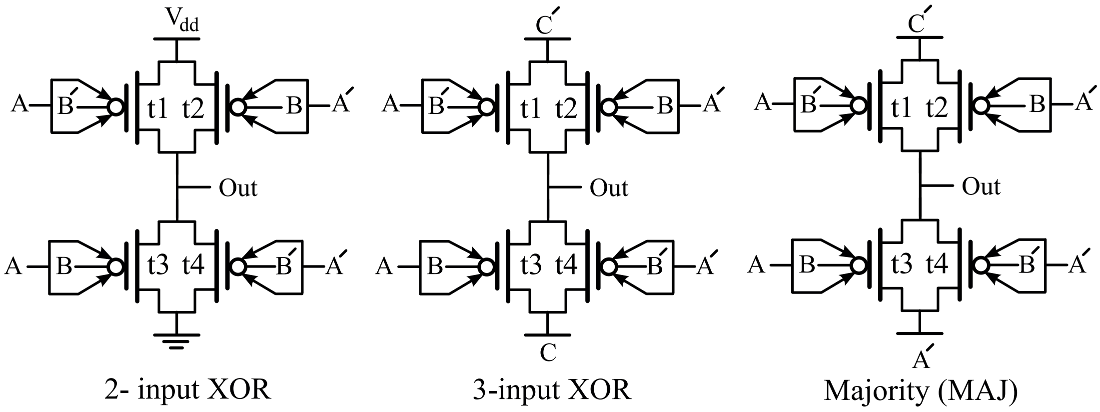


Figure 2.8: Logic gate realization in TIG-SiNWFET. 2-input XOR, 3-input XOR and Majority (MAJ) logic gates are those of DP logic structures.

The second group, called *Dynamic Polarity* (DP), consists of logic gates in which the polarity gates are treated as an extra logic variable. Indeed, the polarity of transistors is dynamically configured by a logic signal during the logic gate operation. Since the conductivity of CP transistors can be controlled by control and polarity gates, this property provides more flexibility for compact realization of binate logic gates. The conductivity of a CP transistor is possible when $CG$, $PG_s$, and $PG_d$ have the same values ($'1'$ for the $n$-type and $'0'$ for the $p$-type. Similarly the transistor has no conductivity when $CG \oplus (PG_S \cdot PG_D) = 1$. This property expresses the intrinsic XOR characteristics of the CP transistors, which has been used for the compact realization of the binate functions such as XOR gate [36]. Figure 2.8 represents the implementation of three DP logic gates (XOR, XOR-3, and MAJ) in TIG-SiNWFET technology. The complementary pull-up and pull-down with parallel transistors leads to static full-swing logic gates and prevents threshold drops in the output.

## 2.5 Variation Sources

In this technology, variations can affect the device during five major fabrication steps, namely:

- *Nanowire patterning through e-beam lithography:* Nanowire patterning is obtained using *Hydrogen SilsesQuioxane* (HSQ) which is very sensitive to variation in electron dose and temperature fluctuation [45]. Exposure to higher temperatures will cause the HSQ to self-expose, thus reducing the necessary exposure electron dose and reducing pattern sharpness. This directly translates into variations of the total length of the nanowires ($L_{NW}$) and *Line Edge Roughness* (LER).

- *Nanowire formation by Bosch process etching* [46]: Variability, in the case of dry etching with Bosch process, originates from various sources. One example is pattern sharpness. Low pattern sharpness, in the case of HSQ, will lead to a tapered nanowire stack, in which bottom nanowires are thicker than top nanowires. This influences the radius of the nanowires ($R_{NW}$). As depicted in Figure 2.9, the iterative etching decrease the width on the top of the stack more than bottom. Therefore, after oxidation step, the upper nanowires will have radii lower than the others.

- *Gate oxide formation and gate deposition:* Variability in this case mainly refers to the thickness of the grown oxide ($T_{OX}$) and the length of the deposited gates ($L_{CG}$). Gate oxide thickness will be influenced by the presence of contaminants on the silicon surface and its shape. This includes the surface roughness and nanowire radius. Moreover, variation of the thickness can influence the time of the gate depositions, thus modifying the gate lengths.

- *Metal-gate deposition:* The alignment of the polarity gates are the main target of variability in this process step. This phenomenon can alter the distances between polarity gates and S/D regions, and finally impact the $On$-current of the SiNWs.

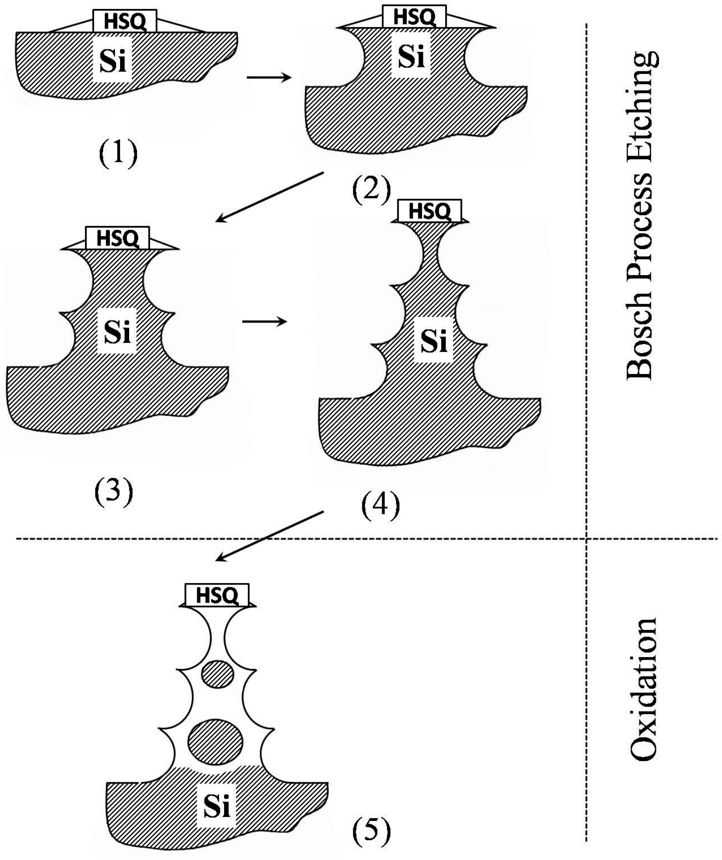


Figure 2.9: Variation on $R_{NW}$ due to the dry etching with Bosch process. In top of the stack, nanowires have the smaller radius.

- *Nickel silicide deposition:* As the crystal structure of the Nickel is not similar to the Silicon, the final arrangement of the obtained silicide after silicidation process determines the characteristics of the Schottky contact like the barrier height. In this case, the properties of the fabricated junctions may differ from device to device, finally lead to SiNWs with dissimilar conductivity.

Table 2.2 summarizes the possible fabrication-induced variations for controllable-polarity SiNWFETs. Figure 2.10 also represents an example of variation in nanowire radius and length of polarity gates for a fabricated DG-SiNWFET. Extracting the possible variations of the target device provides the opportunity of both (i) tuning the process technology early in the development (and thus at low cost) and (ii) studying the device advantages compared to other competitors.

Table 2.2: Fabrication induced variations in controllable-polarity SiNWFET.

| Controllable-polarity SiNWFET | | | |
|---|---|---|---|
| | **Fabrication process** | **Outcome** | **Possible variations** |
| 1 | HSQ-based nanowire patterning | Initial pattern of nanowires | • Variation in length of nanowires<br>• Line edge roughness |
| 2 | Bosch process | Nanowire formation | • Variation in radius of nanowires |
| 3 | Oxide deposition | Dielectric formation | • Variation in oxide thickness |
| 4 | Metal-gate stack deposition | Polarity and control gates | • Variation in polarity gates alignment<br>• Variation in length of polarity and control gates |
| 5 | Nickel silicide deposition | Drain/Source formation | • Variation in barrier height |

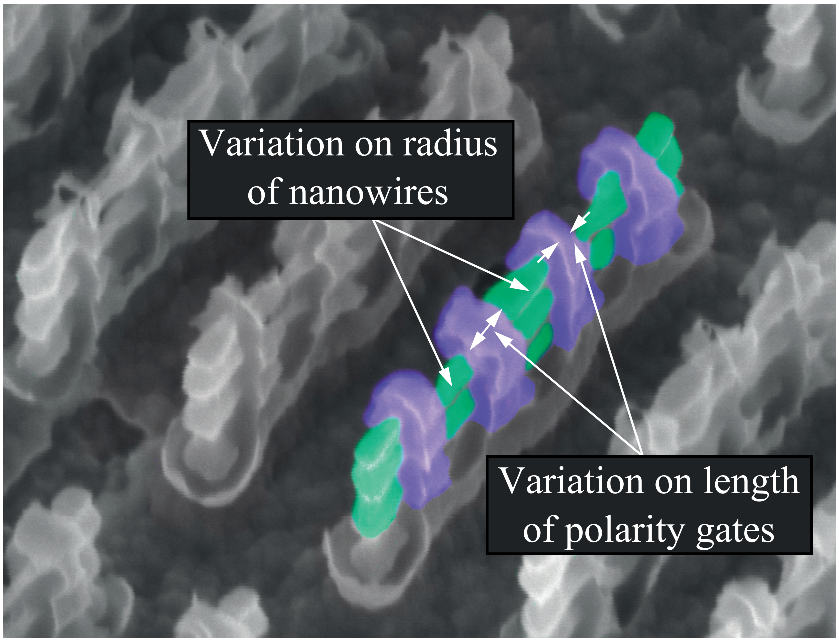


Figure 2.10: An SEM figure of the fabricated DG-SiNWFET which clearly shows the variations on radius of nanowire and length of polarity gates.

## 2.6 Devices Robustness

In this section, we apply the proposed methodology to SiNWs and we investigate the role of process variation for both Schottky-type and doped-type of SiNWFETs.

### 2.6.1 Experimental Setup and Simulation Results

The device robustness for DG-SiNWFET is evaluated and compared to an equivalent single-gate 22nm CMOS technology. The selected CMOS device is SiNWFET with doped Source/Drain contacts (doped SiNWFET). This technology has been chosen as ultimate extension of FinFETs. Here, we only study the behavior of $n$-type devices to have a first insight about the variability of geometrical parameters on the performance of the devices. The study for $p$-type devices can be done easily by following the applied methodology.

The nominal voltage is 0.95V. Transistors are both $n$-type (doped S/D SiNWFETs are obtained using $Sb$ dopants, adapted from [47], while DG-SiNWFETs are polarized with $V_{PG} = V_{DD}$). Variations, with a standard deviation of 30% from the nominal value, are applied to the structural parameters of the devices such as *Nanowire Length* ($L_{NW}$ - 107nm nominal for the DG-SiNWFET and 58nm nominal for a doped SiNWFETs), the *Nanowire Radius* ($R_{NW}$ - 7.5nm nominal), the *Oxide Thickness* ($T_{OX}$ - 12nm nominal) and the *Control Gate Length* ($L_{CG}$ - 22nm nominal). These parameters have been selected in order to represent the impact of process steps on the characteristics of devices. The characteristics, that are investigated here, are *On-current* ($I_{On}$), *Off-current* ($I_{Off}$), *Threshold Voltage* ($V_{Th}$) and *Sub-threshold Slope* ($SS$). For each case study, 100 simulations runs were performed, and non-convergent runs were discarded. The following section provides the numerical results of our study for both DG-SiNWFET and doped SiNWFET.

Table 2.3: The characteristics of DG-SiNWFET and doped-SiNWFET devices in presence of process variation on the structural parameters.

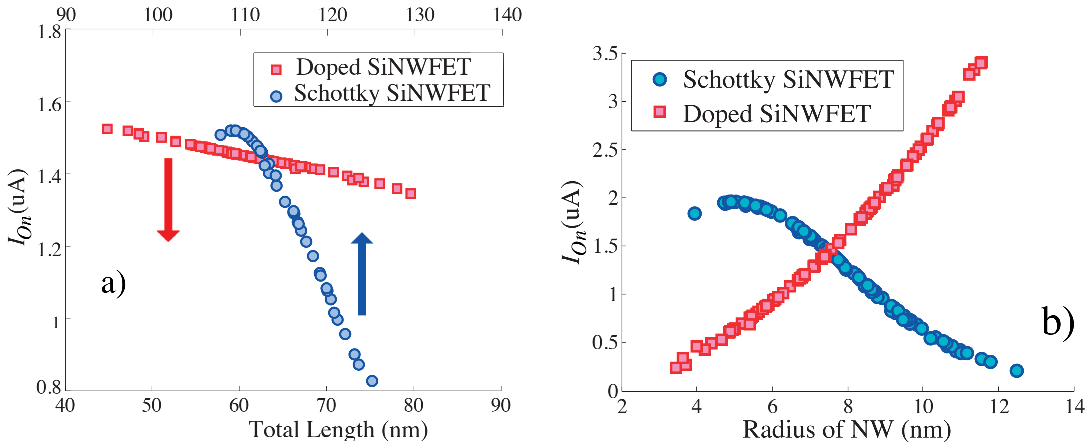
| | | | DG-SiNWFET | | | | Doped SiNWFET | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | $L_{NW}$ | $R_{NW}$ | $T_{OX}$ | $L_{CG}$ | $L_{NW}$ | $R_{NW}$ | $T_{OX}$ | $L_{CG}$ |
| $I_{On}$ | Mean | $(\mu A)$ | 1.34 | 1.28 | 1.45 | 1.44 | 1.44 | 1.64 | 1.44 | 1.44 |
| | Std | $(nA)$ | 193.0 | 528.7 | 48.7 | 0.5 | 21.1 | 789.4 | 0.2 | 15.7 |
| $I_{Off}$ | Mean | $(fA)$ | 0.70 | 0.90 | 0.730 | 0.72 | 6.61 | 9.59 | 6.07 | 7.08 |
| | Std | $(fA)$ | 0.06 | 0.53 | 0.02 | 0.006 | 2.13 | 8.06 | 0.43 | 2.58 |
| $V_{Th}$ | Mean | $(mV)$ | 338 | 330 | 331 | 333 | 387 | 387 | 388 | 386 |
| | Std | $(mV)$ | 17 | 55 | 21 | 9 | 5 | 9 | 3 | 8 |
| $SS$ | Mean | $(\frac{mV}{dec})$ | 79 | 80 | 80 | 80 | 64 | 64 | 64 | 64 |
| | Std | $(\frac{mV}{dec})$ | 2.4 | 2.4 | 4.00 | 0.9 | 1.0 | 1.7 | 0.3 | 0.6 |

The simulation results of the variation analysis for DG-SiNWFETs and doped S/D SiNWFETs are summarized in Table 2.3. Looking at the characteristics of the devices, the DG-SiNWFET demonstrates lower $I_{off}$ value than the doped S/D SiNWFET for the same $I_{On}$ currents. This means that the DG-SiNWFET exhibits superior $\frac{I_{On}}{I_{Off}}$ ratios. Therefore, this makes DG-SiNWFETs extremely well suited for low power applications.

### 2.6.2 $I_{ON}$ of DG-SiNWFET and Doped SiNWFET in Presence of Variation on Length of Nanowires

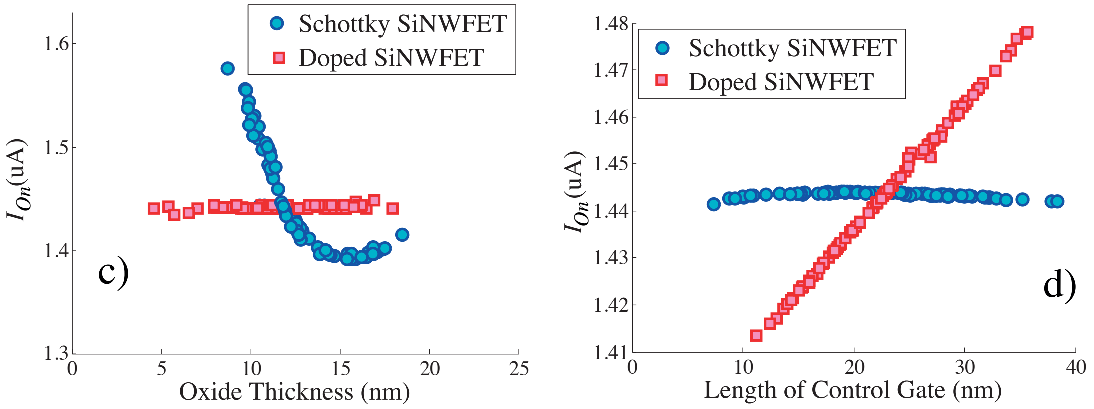
Figure 2.11 depicts the $I_{ON}$ alteration in presence of $L_{NW}$, $R_{NW}$, $L_{CG}$ and $T_{OX}$ variations. In Figure 2.11a, we observe that $L_{NW}$ can affect $I_{ON}$ of the DG-SiNWFET more than that of the doped device (9.2x more). This can be the direct result of characteristic distinction between the two switching mechanisms. The DG-SiNWFET is controlled by two Schottky barrier contacts at the source/drain regions. An increase in $L_{NW}$ does not influence the amount of carriers injected in the channel, however it creates undoped intrinsic regions between the gates. This phenomenon starts to be critical for the DG-SiNWFET even after a variation of 8nm. This confirms the importance for self-aligned structures that relax the constraints on mask alignments for control gate deposition step. Mask alignments for polarity gate deposition through lithography have a considerable impact on the conduction of the DG-SiNWFET. In doped devices, an increase in the total length moves the carrier reservoirs away from the control of gate, leading to the slight decrease in $I_{ON}$. A decrease in total length causes small increase in $I_{ON}$ of doped device thanks to the better electrostatic integrity, while there is no remarkable change on that of the Schottky device.

Figure 2.11b represents the impact of $R_{NW}$ variation on the $I_{On}$ of both SiNWFET types. Variation of $I_{On}$ for the doped SiNWFET is almost linear with nanowire radius variations. In this case, the growth of the nanowire radius corresponds to an increase in transistor width. However, the growth of $R_{NW}$ leads to an slight decrease in the $I_{On}$ of the DG-SiNWFET. Here, the carrier transport is more complex than for a doped SiNWFET owing to the added complication of Schottky barriers at source and drain. Here, the control of injected carriers into and out of the channel is critically depended upon the bias conditions of the source and drain. Indeed, the transportation is obtained through a balance of thermionic emission and tunneling from source to channel and drift and diffusion in the channel. Increasing $R_{NW}$ makes the control of carrier less precise and fewer carriers are injected inside the channel. This highlights the importance of process variation on channel sizing that can greatly impact the conduction of Schottky device.



(a) Impact of $L_{NW}$ variation on $I_{On}$

(b) Impact of $R_{NW}$ variation on $I_{On}$



(c) Impact of $T_{OX}$ variation on $I_{On}$

(d) Impact of $L_{CG}$ variation on $I_{On}$

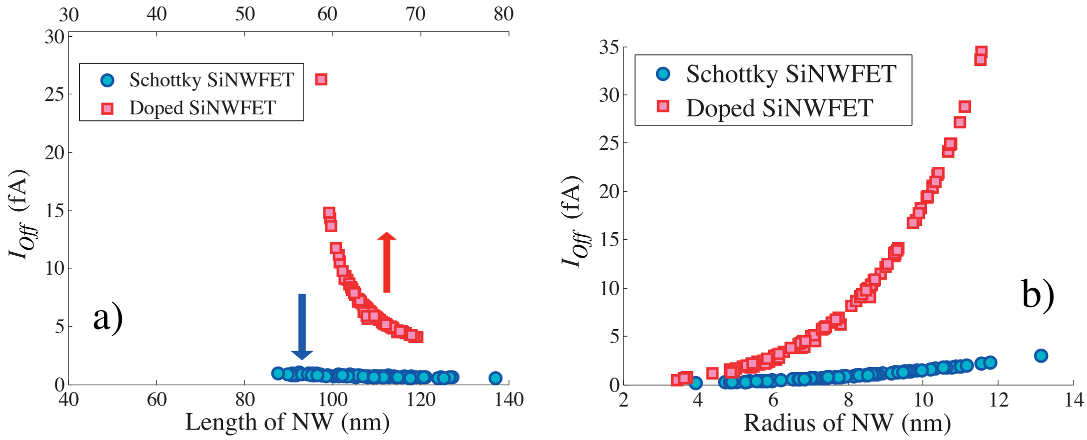
Figure 2.11: Impact of parameter variations on $I_{On}$ for both DG-SiNWFET and Doped SiNWFET.

Figure 2.11c illustrates the outcome of the $T_{OX}$ fluctuation on the $I_{On}$ of DG-SiNWFET and doped SiNWFET. Indeed, when the $T_{OX}$ increases, the $I_{On}$ of doped devices is not affected significantly, while it is heavily influenced for the DG-SiNWFETs. The $T_{OX}$ thickness changes the active region volume of the device. Thus, it becomes more important in DG-SiNWFETs where the number of carriers in the channel is much lower than that of its doped counterpart.

The effect of $L_{CG}$ variation on $SS$ is depicted in Figure 2.11d. The variation on length of control gate is not a dominant factor on the *On*-current of the DG-SiNWFETs. However, it causes a linear increase in $I_{On}$ for the doped SiNWFET. For the DG-SiNWFET, the gates simply control the Schottky barriers, while in the doped SiNWFET, the gate creates the conductive channel. Schottky barrier control is extremely localized; therefore small variations on the gate length do not considerably influence their control (3× less).

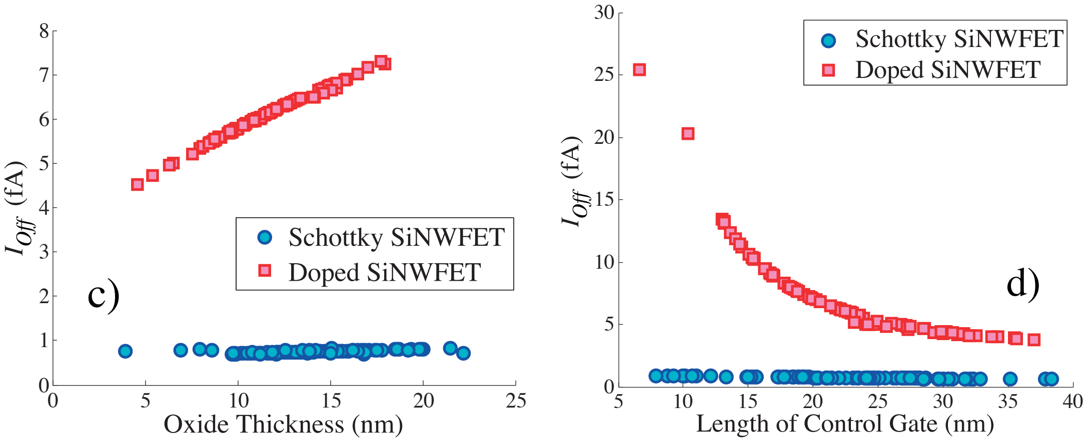
### 2.6.3 $I_{Off}$ of DG-SiNWFET and Doped SiNWFET in Presence of Variation on Length of Nanowires

Figure 2.12 provides several interesting facts about the $I_{Off}$ properties of the nanowires. Figure 2.12a represent the effect of $L_{NW}$ variation on the *Off*-current for both DG- and doped devices. Based on the Schottky switching, a few number of carriers are able to leak into the channel during the *Off* condition. Moreover, utilizing two extra gates over the channel decreases the conductivity when the device is *Off*. However, the doped SiNWFET includes two high-concentration doped regions. Therefore more carriers are prone to enter and pass the channel, leading to higher value and variations on $I_{Off}$.



(a) Impact of $L_{NW}$ variation on $I_{Off}$

(b) Impact of $R_{NW}$ variation on $I_{Off}$



(c) Impact of $T_{OX}$ variation on $I_{Off}$

(d) Impact of $L_{CG}$ variation on $I_{Off}$

Figure 2.12: Impact of parameter variations on $I_{Off}$ for both DG-SiNWFET and Doped SiNWFET.

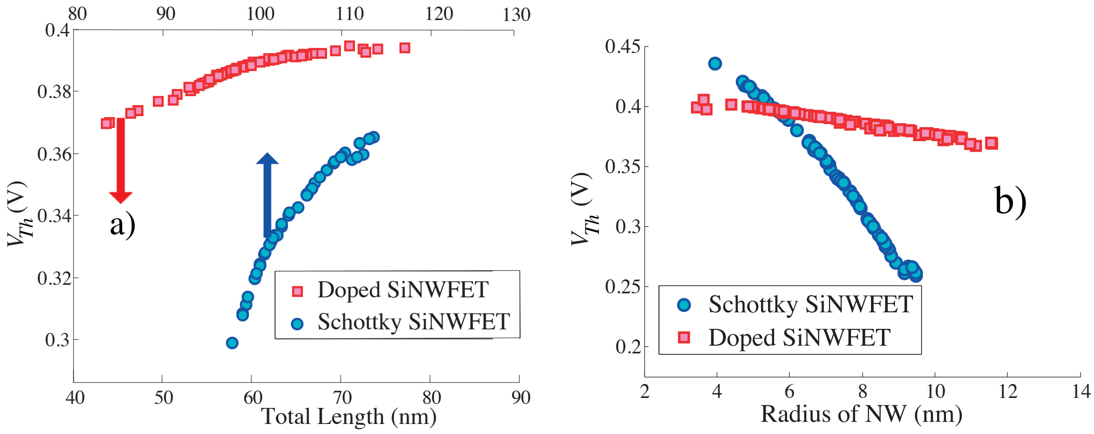
Figure 2.12b demonstrates the effect of $R_{NW}$ variation on the *Off*-current for both DG- and doped SiNWFETs. Fluctuations in nanowire radius slightly change the $I_{Off}$ of the DG-SiNWFET, while it can seriously affect the doped transistor (16× more). Increase in $R_{NW}$ leads to an exponential increase in $I_{Off}$ for the doped device. This is the direct result of a weaker electrostatic control on the channel. On the other hand, $I_{Off}$ of DG-SiNWFETs is less affected by $R_{NW}$ as the carrier injection is blocked through barriers.

Figure 2.12c and Figure 2.12d also depict the influence of $T_{OX}$ and $L_{CG}$ variations on the *Off*-current for both DG- and doped SiNWFETs respectively. The increase on $T_{OX}$ has a linear impact on the *Off*-current of doped devices, while there is no considerable change on that of DG-SiNWFETs. In case of $L_{CG}$ variation, the *Off*-current fluctuates considerably for the doped device (more than 5×), while the variation has negligible impact on the *Off*-current of the DG-SiNWFET. The conductivity of the channel in doped device mainly relies on the electrostatic integrity obtained by the CG. Therefore, fluctuation on $L_{CG}$ can remarkably change the amount of $I_{Off}$. On the contrary, the DG-SiNWFET exploits barrier height control which make them less vulnerable to channel conductivity fluctuations when $L_{CG}$ varies.

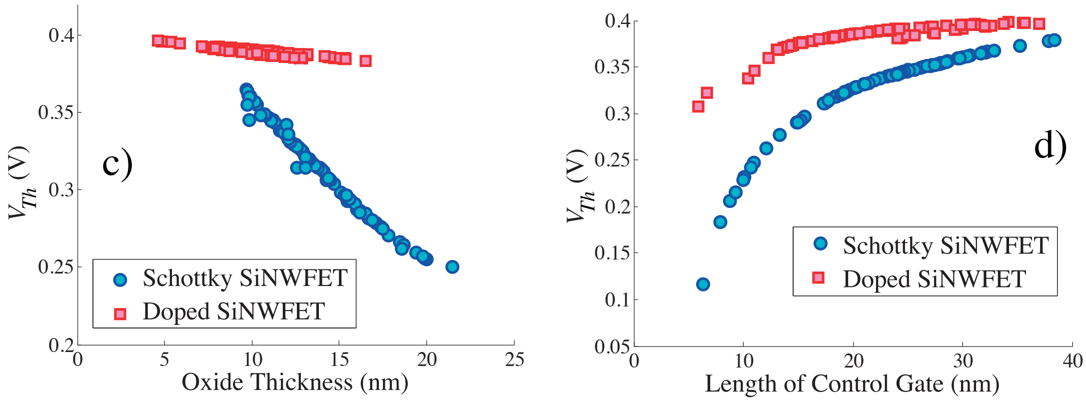
### 2.6.4 $V_{Th}$ of DG-SiNWFET and Doped SiNWFET in Presence of Variation on Length of Nanowires

The $V_{Th}$ is an important parameter that is necessary for some further analysis of a device. $V_{Th}$ is defined as the gate voltage at which the inversion layer charge density is equal but opposite the bulk charge density. However, Schottky devices introduce more complexity on $V_{Th}$ definition due to the complex carrier injection from the S/D regions through the Schottky barrier. It is possible to alter the $V_{Th}$ to a new value by fabricating through fabricating device with smaller or larger barrier heights. For a large barrier height, the current flow from drain to source is limited by the Schottky barrier and, therefore, the $V_{Th}$ is increased. Figure 2.14 provides a couple of interesting information about the $V_{Th}$ properties of nanowires. Basically, the $V_{Th}$ fluctuations is larger in the DG-SiNWFET.

It is observed that $V_{Th}$ degrades under larger $R_{NW}$ and $T_{OX}$ due to the reduced field controllability (Fig. 2.13b and Fig. 2.13d). This results in reduced field emission and finally leads to reduction of $V_{Th}$. Moreover, the large increase of $V_{Th}$ under slight growth of $L_{NW}$ and $L_{CG}$ also is the result of changes in DG-SiNWFETs tunneling mass. These facts are depicted in Fig. 2.13a and Fig. 2.13c.



(a) Impact of $L_{NW}$ variation on $V_{Th}$ (b) Impact of $R_{NW}$ variation on $V_{Th}$



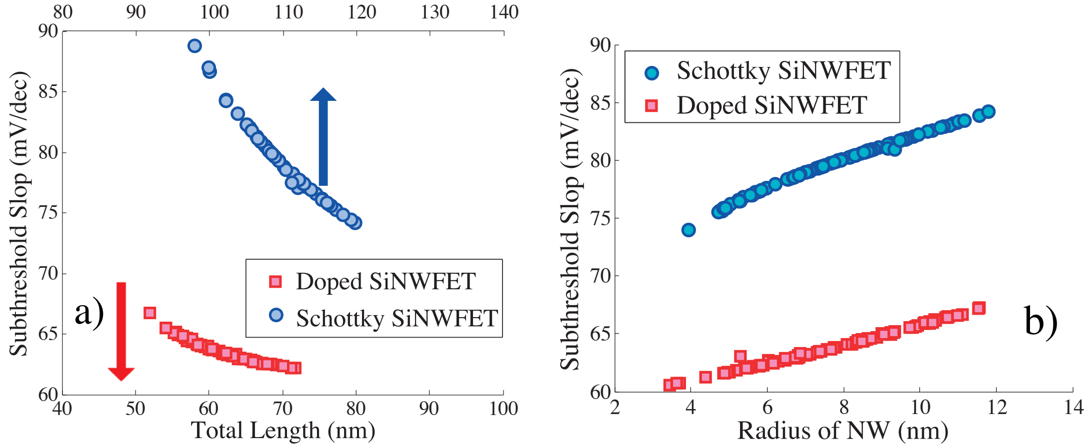
(c) Impact of $T_{OX}$ variation on $V_{Th}$ (d) Impact of $L_{CG}$ variation on $V_{Th}$

Figure 2.13: Impact of parameter variations on $V_{Th}$ for both DG-SiNWFET and Doped SiNWFET.

### 2.6.5 Subthreshold Slope of DG-SiNWFET and Doped SiNWFET in Presence of Variation on Length of Nanowires

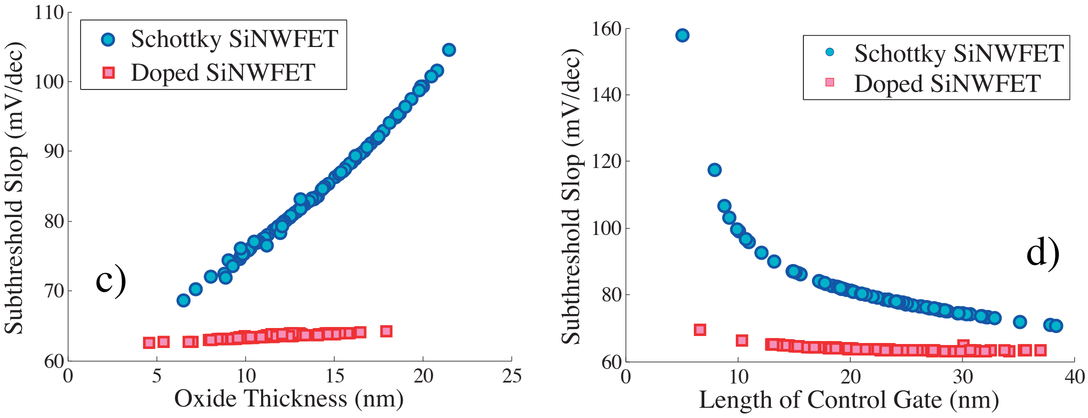
The *Subthreshold Slope* (SS), as an important characteristic of a device, is defined as the slope of drain current in $log$ scale versus $V_{GS}$. The subthreshold slope is an important characteristic for two main reasons. First, it determines the maximum switching speed of the device. Second, a great subthreshold slope allows a low threshold voltage device realization, which is highly important for deeply scaled nano-devices.

Figure 2.14 represents the fluctuations of $SS$ when the geometrical parameters of the both DG- and doped SiNWFETs change. Principally, DG-SiNWFETs show worse Subthreshold-Slope (avg = 80 $mV/dec$) while the doped devices have the smaller value (mean = 64 $mV/dec$). This can be attributed to the barrier height of Schottky contacts. The variations on $L_{NW}$ and $L_{CG}$ cause larger fluctuations on the $SS$ value of the DG-SiNWFETs, which is shown by Figure 2.14a and Figure 2.14c. The impact of $R_{NW}$ for both devices is almost the same, however increase in the oxide thickness linearly degrades the $S$ of DG-SiNWFETs (Figure 2.14b and Figure 2.14d).



(a) Impact of $L_{NW}$ variation on $SS$

(b) Impact of $R_{NW}$ variation on $SS$



(c) Impact of $T_{OX}$ variation on $SS$

(d) Impact of $L_{CG}$ variation on $SS$

Figure 2.14: Impact of parameter variations on $SS$ for both DG-SiNWFET and Doped SiNWFET.

In general, both polarity-controlled and doped devices show variations on the performance parameters such as $I_{On}$, $I_{Off}$, $V_{Th}$ and $SS$ when the value of each geometrical parameter of the devices deviates form its nominal value. The simulation results clearly depicts that the polarity-controlled SiNWFETs are very robust against the variation of $I_{Off}$. This vivdly shows that the polarity-controlled devices are very robust against short channel effects. This would be very important factor for scaling these devices into very tiny dimensions.

## 2.7 Summary

In this chapter, the controllable-polarity silicon nanowire technology is briefly introduced as a promising candidate for future integrated circuits. The device and its characteristics obtained from simulations are then reviewed. In order to perform the simulations, an efficient methodology for electrical simulation, using TCAD and Spice, is introduced. This methodology utilizes TCAD for precise device level modeling. Utilizing the results of the device-level simulations, we established an electrical simulation framework for circuit simulations. The model of the device includes a table mode in Verilog-A. This model is then embedded into Spice simulators for doing the circuit-level simulations in controllable-polarity SiNWFET technology. We also briefly reviewed the structure of the logic gates for this technology. Two two different category of the SP and DP logic gates are introduced. We also looked at the high expressive power of compact realization of logic cells, i.e., 2-input XOR gate, in this technology.

We studied the robustness of the device in presence of variation on the structural parameters of the device. Moreover, the process variation impacts for both polarity-controlled and doped nanowires are comprehensively investigated. The device-level simulation results show that the polarity-controlled SiNWFETs are more robust as compared to doped devices against the variation of a number of structural parameters, such as 16 × smaller standard deviation of *Off*-current on the variation of nanowire length. The simulation results revealed that these devices are very robust against short channel effects and can be utilized for low-power applications. Thus, this technology can be an suitable alternative one for future semiconductor technology.

# 3 A Fast Methodology for Variation Analysis of Emerging Nano-Devices

## 3.1 Introduction

As we have shown previously, *Process Variations* (PVs) and fabrication defects are largely dominating the performance and reliability in advanced technology nodes. The use of compact-models no longer help, to get fast prediction of the device electrical behavior. Indeed, developing and verifying a dependable compact-model for emerging nano-devices is a costly procedure, requiring mature device fabrication, test, and measurement. Therefore, the generation of statistical information for variability analysis is becoming a tough challenge for novel semiconductor technologies.

With the lack of mature devices and compact-models availability, TCAD appears as the most appropriate candidate for PV analysis of emerging technologies. However, the simulation time requirements of potential TCAD-based PV analysis of nano-devices critically depends on the complexity of the physical models, the size of parameter space, and the underlying optimization algorithms [48, 49]. In addition, convergence of computations for small feature size also represents a key limitation for commercial TCAD software. Current TCAD simulators benefit from different types of optimization techniques at the various level of simulation [50]. Nevertheless, long TCAD simulation time for nano-devices remains an important barrier, preventing their use for fast PV analysis.

In this chapter, we propose a novel methodology intended to effectively speed up the variation analysis for nano emerging devices and circuits. A prediction technique is proposed to estimate the output of TCAD simulation. The technique is based on a “*Feed Forward Neural Network*" (FFNN) and provides a multivariate non-linear regression method for faster transistor I-V prediction as compared to TCAD simulations. This technique learns the fundamental relations between device parameters and its functionality. In order to optimize the I-V curves fitting, we modify the FFNN structure through adding a number of extra nodes called *Control Nodes*. These nodes modify the weights of the network in a way that the predicted I-Vs have minimum error value corresponding to *On Current* ($I_{On}$), *Off Current* ($I_{Off}$), *Threshold Voltage* ($V_{Th}$), and *Sub-threshold Slope* ($S$). This methodology enables us to produce a large enough data set of a target device in presence of variations and handles the convergence issues related to complex numerical models. To evaluate the capability of the developed methodology, *Double-Gate Silicon Nanowire* FETs (DG-SiNWFETs) were used as a target device. Simulation results proved that the proposed methodology makes the time complexity of device simulation comparable to compact-model simulation. After training, the average runtime for I-V estimation of a DG-SiNWFET is 0.12 $ms$ which is considerably lower than TCAD simulation execution time.

## 3.2 Background and Motivation

### 3.2.1 Problem: Difficulty of Generating Statistical Information

Statistical information for transistors can be obtained by the three following ways: device fabrication, TCAD simulation, and compact-model simulation. One important issue of nano-devices fabrication, in addition to the lack of maturity, resides in the increase of processing complexity. Moreover, the contribution of each device parameter to the functionality variation cannot be determined through cross-section analysis. On the other side, reliance on device compact-models is a hardly feasible solution since the development of accurate models mainly depends on tests and extensive measurements on mature devices.

Thanks to their precise physical models, TCAD simulation provides comprehensive information about devices behavior and alleviates expensive fabrication and inspection process. Hence, TCAD simulations are an invaluable solution for the evaluation and minimization of PV impacts, thus enabling technologists to early identify the main sources of parametric yield loss in manufacturing. However, it suffers from several significant challenges such as computational complexity of the DC/AC device simulations and convergence problem for complex models. Thus, device-level *Monte Carlo* analysis of nano-structures using TCAD also will result in prohibitive time complexity. Therefore, a fast methodology for PV analysis is required, especially when new processes or devices are introduced. In the following, a novel methodology which can be integrated with TCAD to enhance simulation performance of PV analysis is introduced.

### 3.2.2 Previous Work

TCAD simulation are extensively exploited for evaluating the fabrication process and the electrical behavior of semiconductors [51]. During the recent years, TCAD has been used for the simulation of emerging nano-devices such as *Multigate FETs* [52] and CMOS/Flash devices [53]. Moving towards deep submicron technologies, the process simulation and physical model of devices becomes more complex. Hence, optimization technique have been utilized for improving TCAD performance.

Process simulation requires geometrical information and mesh generation for the model under test. Moreover, it accurately estimates structural layers and active dopant distributions at the end of a procedure run. Techniques to enhance process simulation precision and complexity have been investigated in [54, 55].

Physical models are based on complex equations, describing the semiconductor conduction mechanisms. In order to reach desirable accuracy at advanced technology nodes, the complexity of model keeps increasing. Due to computational complexity, several optimization techniques such as cluster computation and gradient-based optimizations have been proposed [56]. However, the downscaling of semiconductors are correlated to increasingly complex models and therefore the efficiency of the mentioned techniques are counteracted.

Most recently, [57] proposed a methodology in which redundant physics computations are removed for common parts of the model and accordingly repetitive 3-D simulations are discarded for various parts of a circuits/layouts. This methodology accelerates TCAD simulations and makes it reachable for analysis of small circuits and logic blocks. However, it is still very limited for large layouts or for the repetitive simulations in the context of PV analysis. For example, one TCAD simulation run of a 6T SRAM cell in FinFET technology takes more than 17 hours in this methodology.

## 3.3 Learning-based Methodology for Fast Process Variation Analysis of Emerging Devices

In this section, we introduce the proposed methodology and its key component: a prediction module. The prediction module is used, after training, to predict the device variations instead of using TCAD simulator, thereby improving significantly the simulation process speed.

### 3.3.1 PV Analysis Using Neural Networks

To model the underlying relations of multivariate parameters in TCAD simulations, a precise definition of the functional relations among the variables is required. These functional relations are based on solid-state physics and quantum-mechanics equations. Thus, the TCAD simulation for only a single device is very slow ($O(hours)$) [57], and also the simulation does not converge for a non-negligible number of data points.

To overcome this issue, our method uses a predictor which learns how to approximate the simulator results. The predictor is used for data generation and deal with the missing data resulting from failed simulation runs. Without loss of generality, we will focus on single device simulation in this work. Circuit-level modeling using this method is out of this paper's scope.

Figure 3.1 represents the general flow of the mentioned methodology. First, our methodology considers the TCAD model for the target device, using the nominal values of the device parameters. This set of parameters is determined to precisely reproduce the normal I-V characteristics of the target device. In the next step, the profile of model parameter variations such as gate length, channel length, and oxide thickness, is applied to the TCAD model generator, and then a large data set of various device models is created. Only a small subset of these TCAD models is randomly selected and simulated. The cardinality of this subset represents the minimum necessary TCAD simulations for training the predictor. Predictor uses the obtained I-V curves as a ground truth and learns the underlying relations between the device parameters and its functionality. In order to speed up the estimation of I-V curves, the prediction module uses a regression technique to prevent repetitive time-consuming TCAD simulations. After training, the TCAD simulator can be replaced by the predictor, in order to speed up the PV simulations. In the following subsections, we explain the predictor structure and the training algorithm.

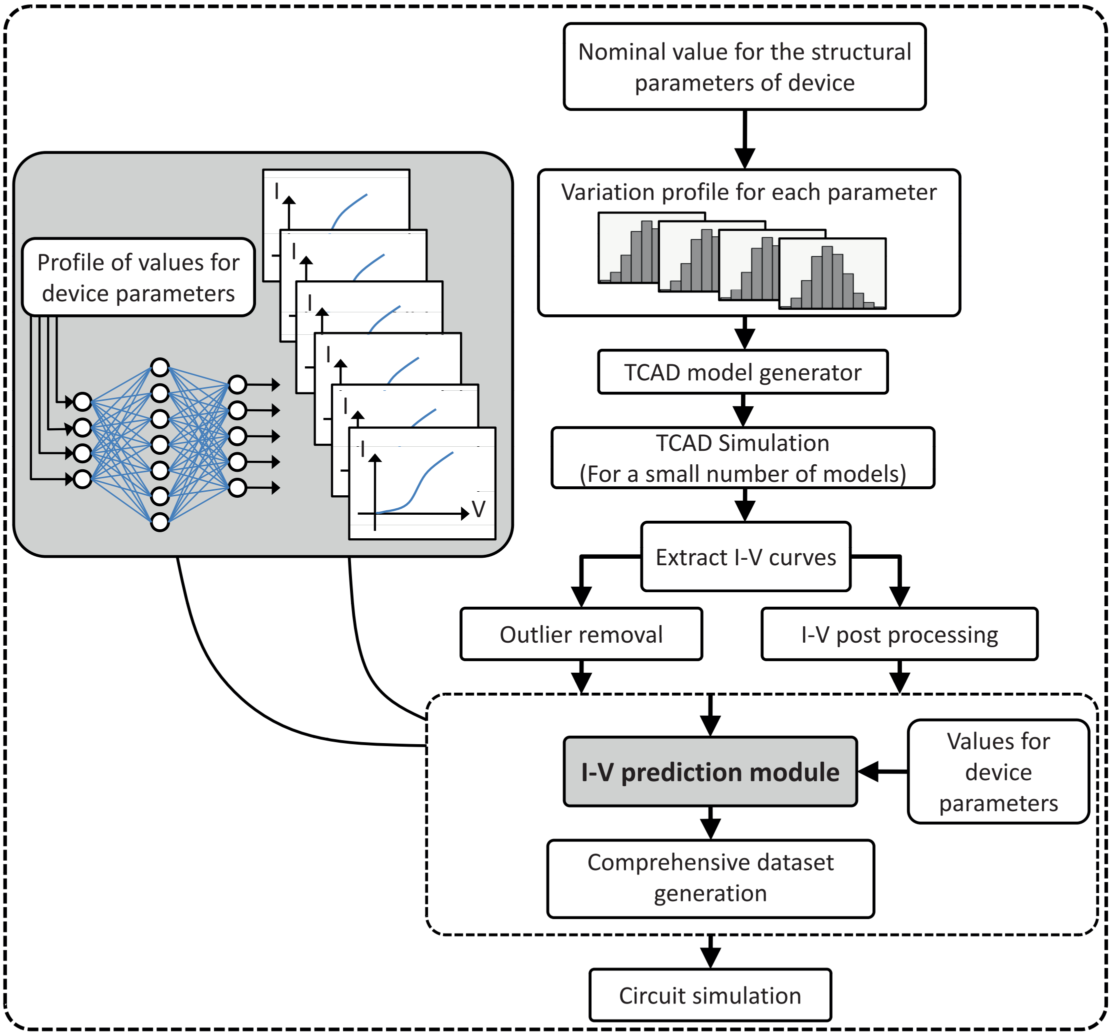


Figure 3.1: Flowchart of the proposed methodology.

### 3.3.2 Neural Network Structure

Among nonlinear regression techniques, *Feed-Forward Neural Network* (FFNN) [58] is mostly used in practice. Compared to its competitors, such as *Super Vector Regression* (SVR) [59], FFNN has smaller number of hyper parameters and smaller model-size. Moreover, FFNN is fast while SVR becomes very slow when many support vectors are created. Although we utilize FFNN for prediction purpose here, it should be noted that other non-linear regression methods can also be applied to estimate the $I-V$ values. The performance comparison for these methods is out of this paper's scope.

Figure 3.2 provides a graphical representation for the proposed three layer FFNN. The set of device parameters such as *Channel Length*, *Oxide Thickness*, and *Gate Length* makes the FFNN input vector ($\mathbf{x} = [x_1, x_2, \cdots, x_j]$). The set of $\mathbf{h} = [1, 2, \cdots, h]$ also represents the hidden layer nodes, where $h$ is the size of hidden layer. The FFNN output vector ($\hat{\mathbf{i}}_d = [\hat{i}_{d1}, \hat{i}_{d2}, \cdots, \hat{i}_{dv}]$) approximates the observed values of the I-V curves (sampled *Drain Current* $\mathbf{i}_d$ set) for the given input. An estimated $i_{dv}$ value ($\hat{i}_{dv}$) for the given input is calculated by:

$$\hat{i}_{dv}(\mathbf{x}) = b'_v + \sum_{m=1}^{h} w'_{mv} \cdot \phi(b_m + \sum_{k=1}^{j} x_k \cdot w_{km})$$

where $\{b', b\}$ and $\{\mathbf{W}', \mathbf{W}\}$ are the network parameters and are called bias set and weight set respectively. These parameters are learned during the training step. $\mathbf{W} \in \mathbb{R}^{j \times h}$ is a matrix which transforms the input data into the hidden space, and $\mathbf{W}' \in \mathbb{R}^{h \times v}$ is a matrix which transforms data from the hidden space to the output space. The activation function $\phi$ is conventionally chosen as *Sigmoid logistic function* ($\frac{1}{1+e^{-x}}$) or *hyperbolic tangent* ($tanh(x)$) in non-linear regression problems [60]. It is known that FFNN is a universal approximator, which is expressed in the following theorem reported from [61].

**Theorem 1:** Let $\phi(.)$ be a bounded, and monotonically-increasing continuous function and $I_m$ denote the $m$-dimensional unit hypercube $[0,1]^m$. The space of continuous functions on $I_m$ is denoted by $C(I_m)$. Then, given any function $f \in C(I_m)$ and $\varepsilon > 0$, there exist an integer $N$ and real constants $\alpha_i, b_i \in \mathbb{R}$ and $w_i \in \mathbb{R}^m$, where $1 \leq i \leq N$ such that: $F(x) = \sum_{i=1}^{N} \alpha_i \phi(w_i^T x + b_i)$ and we have $|F(x) - f(x)| < \varepsilon$.

The hidden space dimension, $h$, determines the learning capability of the network. An FFNN with a sufficiently large single hidden layer can approximate the function showing the relations among the parameters and their impact on device functionality. It should be noted that the excessive number of hidden nodes makes networks more prone to learn noise and memorize the training data. Therefore the prediction ability of the network is then reduced. In the next section, we discuss how to determine the number of hidden units practically.

**Control Nodes:** Information obtained from I-V curves are not equally valuable and a number of key statistical information such as $I_{On}$, $I_{Off}$, $V_{Th}$, and $S$ are more important. Therefore, we introduce a set of control nodes within the output layer in order to approximate specifically

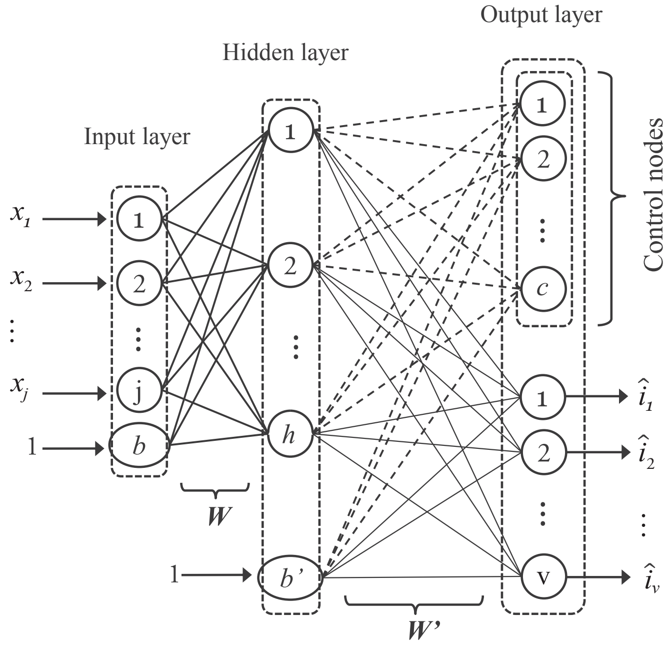


Figure 3.2: Graphical representation of the FFNN for I-V curve regression

these valuable information. Note that this information is already presented in the output vector, but we distinctively add control nodes to insist on the accuracy of I-V estimation and modify the degree of fitting. By utilizing control nodes, the weights of the nodes which contribute to estimate this information are modified in a way that they have more impact on the regression. Control nodes prevent model from overfitting and reduce the impact of flat data on the I-V prediction. Our experimental results verify that control nodes enhance the performance of regression.

### 3.3.3 Training of the Neural Network

In order to make our ground truth for training purpose, the TCAD generated I-V curves are discretized. The *Gate Voltage* ($V_G$) value is limited to $[0, V_{dd}]$ along a TCAD simulation. This interval is divided to a number of equidistant subintervals, and then current values are sampled at the end point of these subintervals. Since the $V_G$ values are alike for the whole I-V curves, the correspondent $\mathbf{i}_d$ values have to be estimated through the prediction unit. Let the set $Tr = ((\mathbf{x}_1, \mathbf{i}_{d1}), (\mathbf{x}_2, \mathbf{i}_{d2}) \cdots (\mathbf{x}_n, \mathbf{i}_{dn}))$ represents the training set which $n$ is the number of simulations used in the training set. The set of the device input parameters is $\mathbf{x}_k$ and $\mathbf{i}_{dk}$ is the observed values of the sampled $\mathbf{i}_d$ set for the given input. We minimize the loss function which is defined as the least squares error between the approximation values and the observed values, over the training set:

$$min\ l = \sum_k (\sum_j (\mathbf{i}_{dj}(\mathbf{x}_k) - \hat{\mathbf{i}}_{dj}(\mathbf{x}_k))^2 + (\mathbf{w}^T(\mathbf{c}(\mathbf{x}_k) - \hat{\mathbf{c}}(\mathbf{x}_k)))^2)$$

where $\mathbf{c}(\mathbf{x}_k)$ represents the vector of control nodes in output layer and $\mathbf{w}$ is a vector of size $\|\mathbf{c}\|_1$ which shows the importance of each control nodes. The backpropagation training algorithm plus *Levenberg-Marquardt* algorithm [62] is used to find the weights of the network.

Cross validation is a practical method to find the number of hidden units and prevent overfitting. To perform cross validation, a part of data set is randomly selected to test the generalization of the network during learning process. The error value of cross validation can be used a criterion for the number of training cycles and the network size. As the network size increases, the error value on the training set becomes smaller. However, the error value on the test set reach to a minimum value for the optimized network size and training iterations. In the following section, we study the experimental results of the cross validation.

## 3.4 Result and Discussion

In this section, the experimental results are presented. First, the setup of the experiments is explained and, then, the simulation results are discussed.

### 3.4.1 Setup of Experiments

As a case study, *Double-Gate Silicon Nanowire* FETs (DG-SiNWFETs) was used as a target model in our experiments. Table 3.1 presents the main geometrical parameters of device and their nominal values.

Table 3.1: DG-SiNWFET geometrical parameters.

| Parameters | Acronyms | Nominal values (nm) |
|---|---|---|
| Length of Control Gate | $L_{CG}$ | 22 |
| Length of Spacer | $L_{CP}$ | 18 |
| Length of Polarity Gate | $L_{PG}$ | 22 |
| Length of Drain/Pillar Extensions | $L_{XD}$ | 2.5 |
| Length of Source/Pillar Extensions | $L_{XS}$ | 2.5 |
| Radius of NW | $R_{NW}$ | 7.5 |
| Polysilicon Thickness | $R_{PSI}$ | 2 |
| Oxide Thickness | $T_{OX}$ | 12 |

For the sake of experiments, a 30% variation along normal distribution is used for each geometrical parameters ($\sigma = 30\%$), in order to show the impact of PVs involved in the device fabrication. In our experiments, *Sentaurus* was used as the TCAD simulator. As a case study, 2300 TCAD simulations were performed to provide enough statistical information, and then analytical metrics such as the On-current ($I_{On}$), the Off-current ($I_{Off}$), the threshold voltage ($V_{Th}$) and the sub-threshold slope ($S$) were extracted through I-V curves post processing. In order to find the minimum number of data points for training, we also trained the predictor with the various sizes of a training set. Figure 3.3 depicts the training performance when the size of training set is increases. Our study revealed that almost 600 input data is enough for training the predictor.

The proposed FFNN is implemented in *MATLAB*. In our experiments, all the input vectors are applied to the FFNN in each iteration which is called epoch and then the estimation error is computed. The new values of the weights are computed by applying gradient descent on the error function. The learning method utilizes *Mean Square Error* (MSE) error function. In order to validate the proposed prediction I-V module, we divided the data set, i.e., the outputs of TCAD simulations into three parts: 75% for FFNN training, 15% for validation, and 15% for test.

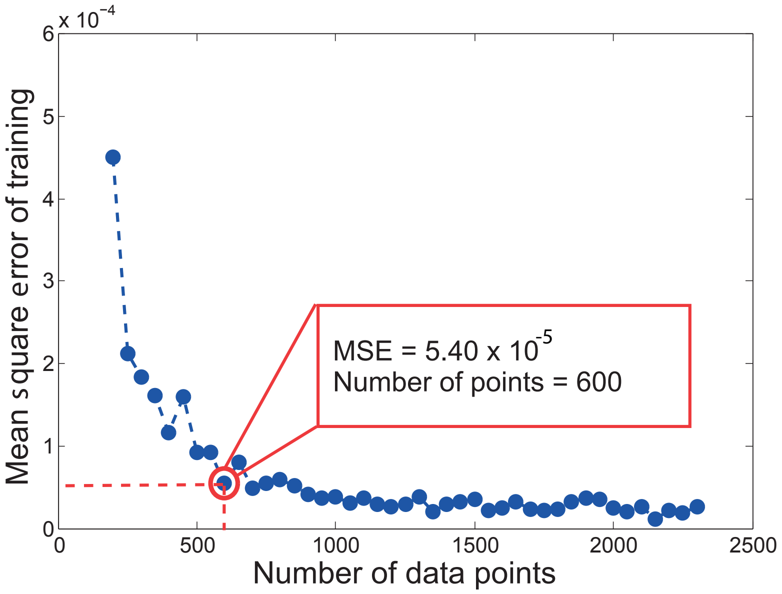


Figure 3.3: Necessary number of data points for training the predictor.

Before the learning process, we need to determine the hyper parameters of the model such as error function, activation function and the number of neurons in hidden layer. The selection of the first two hyper parameters was discussed before. In order to find the optimum number of nodes in the hidden layer, we performed the following procedure. The number of hidden nodes is increased until the performance of the FFNN on the test set starts to decrease. Figure 3.4 depicts the error value of the network on the test and training sets. As shown in the figure, the minimum error on the test set is achieved when hidden layer has 50 nodes. The figure clearly depicts that how the performance of the larger networks are degraded on the validation set.

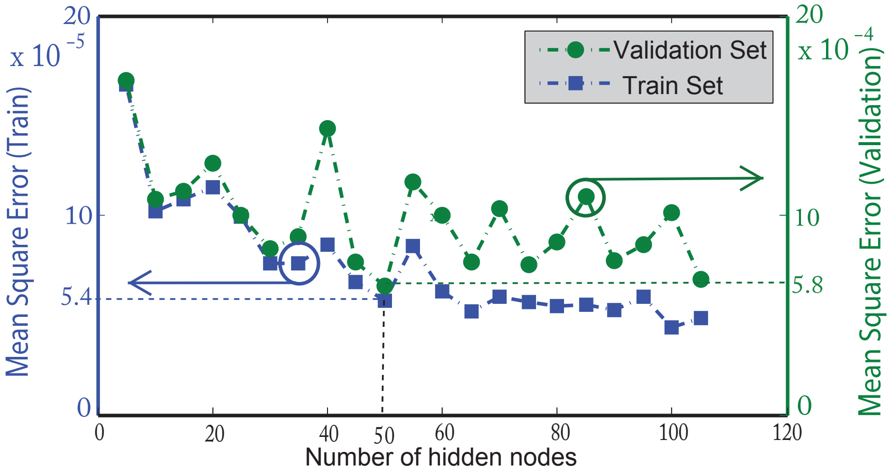
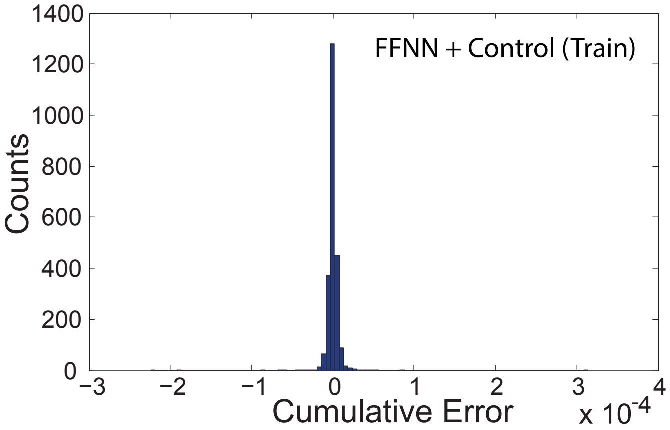
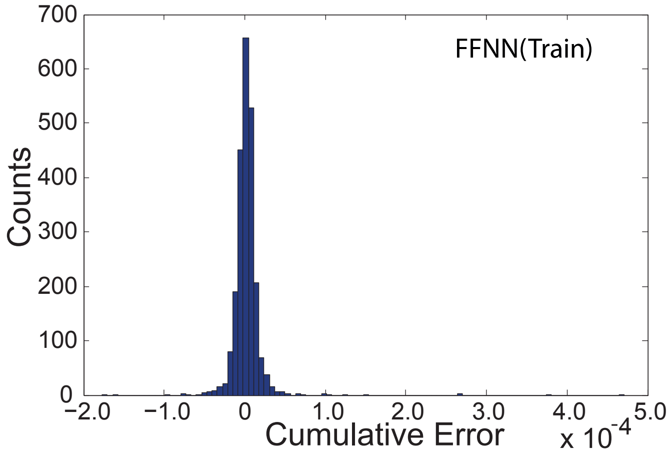


Figure 3.4: Mean square error of training and test sets for networks with various number of hidden nodes.



(a) Training error with control nodes.



(b) Training error without control nodes.

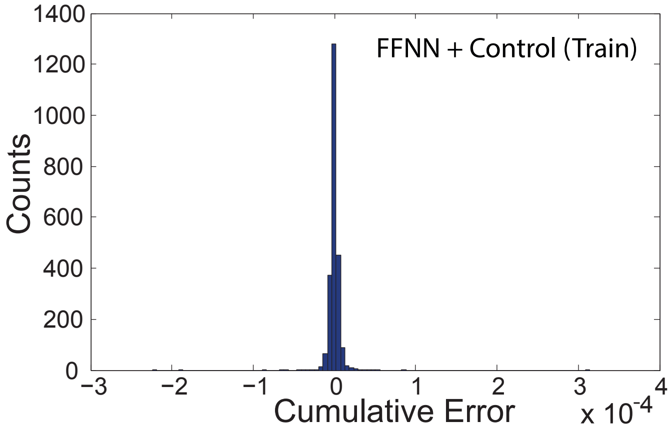
Figure 3.5: Error distribution for the training of networks with and without control nodes.

### 3.4.2 Experimental Results

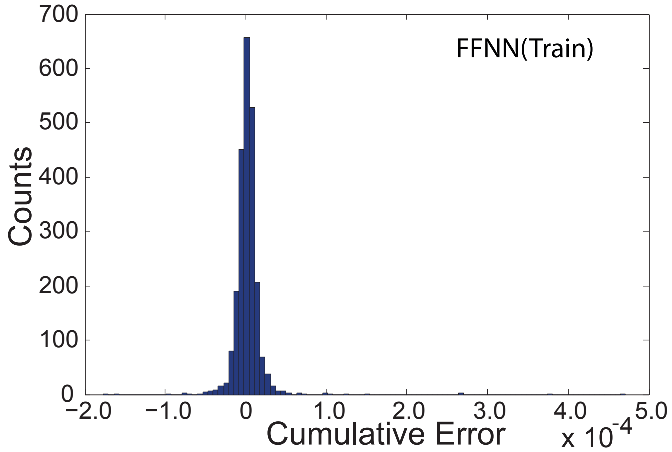
In the following, the simulation results of the proposed methodology are represented. Precision and time complexity are the main aspects of our study.

#### Precision of Prediction

To show that the control nodes can enhance the results of regression and prevent overfitting over the training set, we built two separate FFNNs: one with control nodes and the other one without control nodes. Figure 3.5 illustrates the error distribution of the training for the two FFNNs. Figure 3.6 also depicts the error distribution of the test for the two FFNNs. The best performance of the FFNN with control nodes is bounded to the MSE of $4.3 \times 10^{-4}$, while the ordinary FFNN reaches the error bound of $6.0 \times 10^{-4}$. Thus, the addition of control nodes, controls the error bound of estimator.



(a) Test error with control nodes.



(b) Test error without control nodes.

Figure 3.6: Error distribution for the test of networks with and without control nodes.

Figure 3.7 depicts a sample of I-V regression for the ordinary and proposed FFNNs. The error value of $V_{Th}$ and $I_{On}$ are much smaller when control nodes are added to FFNN. Therefore, control nodes enable the FFN to learn the shape of the output while, simultaneously, the error of the important outputs is kept minimum. This approach of estimation can be useful for the statistical analysis that use the information of post proceed I-Vs such as $I_{On}$, $I_{Off}$, $V_{Th}$, and $S$ for the purpose of PV analysis.

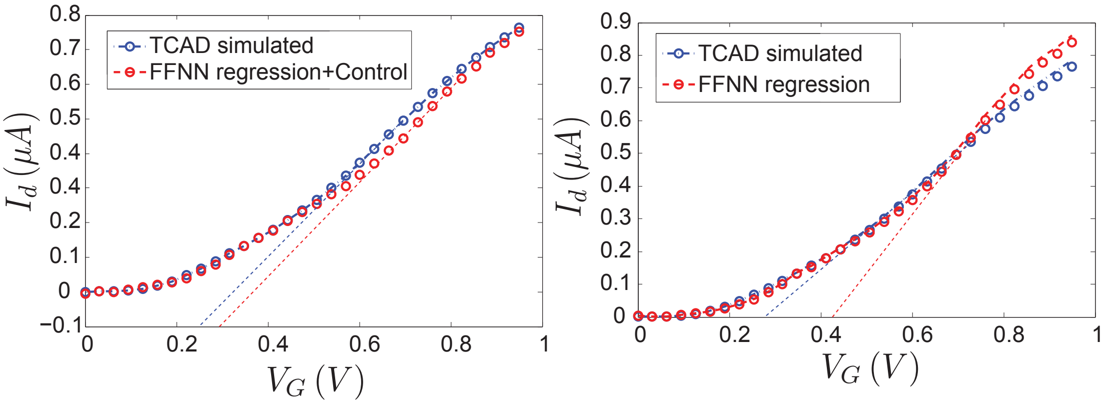


Figure 3.7: Two sample I-V regression through different FFNNs. Adding control nodes to FFNNs results in more precise prediction.

In order to verify our methodology, we used it for PV study of DG-SiNWFET by applying a 30% Gaussian random variations for each studied parameters. The same experiments were performed with TCAD simulations for comparison purpose. Table 3.2 represents the effect of $R_{NW}$ ,$T_{OX}$, and $L_{CG}$ variations on $I_{On}$, $I_{Off}$ and $V_{Th}$. Looking at the obtained values, our methodology demonstrates a good correlation (average error of less than 2% in mean values of $I_{On}$, $I_{Off}$ and $V_{Th}$) with the results of TCAD simulation.

Table 3.2: DG-SiNWFET PV analysis using TCAD and proposed methodology.

| | | TCAD | | | Proposed methodology | | |
|---|---|---|---|---|---|---|---|
| | | $R_{NW}$ | $T_{ox}$ | $L_{CG}$ | $R_{NW}$ | $T_{ox}$ | $L_{CG}$ |
| $I_{On}$ | $mean\,(\mu A)$ | 1.28 | 1.45 | 1.44 | 1.29 | 1.43 | 1.43 |
| | $std\,(nA)$ | 528.7 | 48.7 | 0.5 | 536.2 | 57.7 | 0.7 |
| $I_{Off}$ | $mean\,(fA)$ | 0.90 | 0.73 | 0.72 | 0.88 | 0.76 | 0.72 |
| | $std\,(fA)$ | 0.53 | 0.02 | 0.006 | 0.49 | 0.03 | 0.01 |
| $V_{Th}$ | $mean\,(mV)$ | 330 | 331 | 333 | 321 | 346 | 327 |
| | $std\,(mV)$ | 55 | 21 | 9 | 59 | 18 | 12 |

### Time Complexity

The runtime of the PV analysis for DG-SiNWFET through the proposed methodology is shown in Table 3.3. The prediction methodology is clearly faster than TCAD simulation. The time complexity of the FFNN predictor is related to the learning step which is done once. This methodology is useful when a large data set is needed. This is a common case for statistical PV analysis. For example, consider a usual Monte Carlo method for variation analysis of DG-SiNWFET in which 5000 data points is required. Only a small part of the data points (nearly 12%) is necessary to reveal the variation distribution on the device output. In this case, only 600 TCAD simulations can be performed to set up the predictor instead of 5000 overkilling simulations. In the next step, TCAD simulator is replaced by predictor which can quickly produces the necessary data set in $O(sec)$ for 4500 remaining simulation (with average execution time of 0.12 $ms$ per sample). Figure 3.8 compares the execution time of TCAD and proposed methodology for PV analysis. After a number of TCAD simulation which is necessary for the training, FFNN considerably accelerates the simulations runtime.

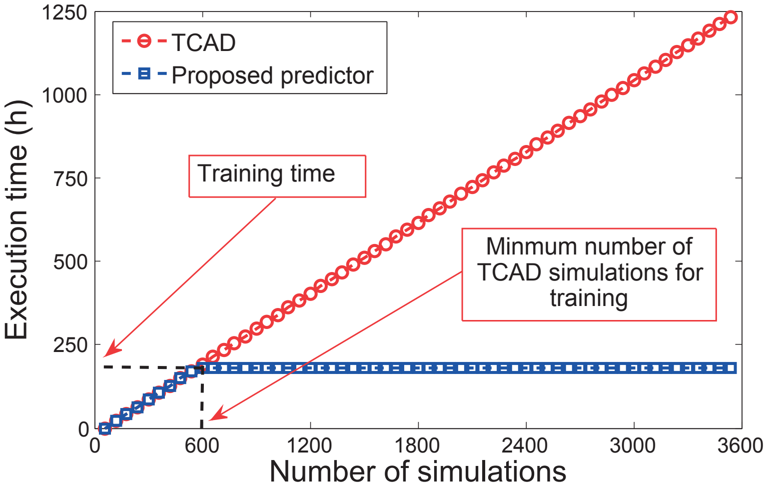


Figure 3.8: Runtime comparison of the TCAD simulation with proposed methodology.

Table 3.3: Execution time comparison between TCAD-based and proposed learning-based I-V curve estimators.

| Method | Average execution time for one simulation (CPU: Dual-Xeon X5650, Memory: 24 GB ) |
|---|---|
| TCAD | 19 min |
| Proposed methodology | $1.23 \times 10^{-4}$sec (Learning time: 18 min) |

## 3.5 Summary

In this chapter, we introduced an efficient methodology for variation analysis of emerging nano-devices. This methodology consists of a prediction module which can be added to TCAD simulator to improve the time complexity of the PV analysis. The prediction module is based on feed-forward neural network, which is capable of estimating the underlying relation of device parameters and device functionality. Learning this relation enables us to replace repetitive and time consuming TCAD simulation by a fast estimation methodology. Our simulation results revealed that the proposed methodology can effectively be used for PV analysis, while reducing the time complexity of TCAD simulations of device to $O(ms)$ with the maximum error bound of $4.3 \times 10^{-4}$.

# 4 Parameter Selection for Nonlinear Modeling of Process Variation

*Process Variation* (PV) analysis through simulation is the most realistic approach for the comprehensive study of variation impact for both circuit static timing and leakage power. Parametric variation analysis is performed by means of *Monte Carlo* (MC) simulation and is widely used in microelectronics industries, even if it is extremely time-consuming for large circuits. Considering the variety of local and global variations in device and circuit simulations would need up to some thousands or millions of variation variables to represent the distributions of the geometrical and physical parameter quantities [63]. Moreover, for practical reasons circuits are usually characterized with relatively small number of parameters through compact models. Scaling beyond 22nm forces a transition from bulk CMOS technologies to others like FD-SOI, FinFET, or SiNW, for which statistical compact models are inevitable for variability aware design. Statistical compact models exploit *Technology Computer Aided Design* (TCAD) to predict the impact of fluctuations on device performance [64, 65]. The results of TCAD simulation can be fed to SPICE-like simulator for MC simulations of circuits. Nevertheless, the high dimensionality of the parameters space and the computational complexity of TCAD simulation make the PV analysis very costly and even sometimes infeasible. Therefore, new tools which speed up the variation analysis for deeply nano-scaled circuits are required.

The efficiency of current methods for performance analysis, e.g., statistical timing verification techniques, critically relies on the dimension of the parameter space [48, 66, 67]. Most of the existing techniques such as *Principal Component Analysis* (PCA) and *Independent Component Analysis* (ICA) use a linear transformation to reduce the number of input parameters by decorrelating the input space [68], and [69]. In spite of their popularity, they are inherently limited because they only consider the relations among the input parameters and ignore the impact of each input on the circuit outputs. This limitation becomes important when either some critical parameters, which significantly affect the output, are ignored or a large set of transformed parameters may still be produced after redundancy removal. Moreover, although statistical methods, such as *Reduced Ranked Regression* (RRR) and *Canonical Correlation Analysis* (CCA), consider the correlation between the input parameters and the circuit outputs, they ignore the correlation among the input parameters [70]. Therefore, they may lead to a large set of correlated parameters while the input space can be compressed by considering inter-parameter correlation. Last but not least, the mentioned methods put strict assumptions on the distribution of the model parameters such as Gaussian distribution which limits their applicability to recently proposed nano-devices where parameters have mixed Gaussian and non-Gaussian distributions.

In this chapter, we introduce a novel multi-objective parameter selection method capable of addressing the aforementioned limitations. This method takes into account the inter-set (among inputs) and intra-set (between input and output sets) correlations. The objective function is modified to be distribution free and minimize the error of output estimation. The major contributions of the method can be summarized as follows:

- High precision by considering nonlinear dependencies between inter-set and intra-set parameters.
- Distribution-free feature selection which can be used for any model or parameter set with unknown statistical distributions.
- Feature selection in the input parameter space which preserves the meaning of the parameters and highlights the major contributors on device or circuit variability.

We show that such parameter selection approach leads to more feasible PV analysis of complex design. Therefore, parameterized models are built with a smaller set of statistically significant parameters.

To validate the technique, we performed two sets of experiments on two different target technologies. First, we use FinFET 20nm technology as a contemporary integrated circuit technology. Based on that, we analyzed the delay variation of the longest path for a couple of ITC'99 and ISCAS benchmark circuits. Here, 5× speed up in *Monte Carlo* (MC) is obtained for timing variation analysis with the average variance error of 4.1%.

We also applied the proposed technique on a ISCAS benchmark circuit in DG-SiNWFET technology. The simulation results for timing analysis of the combinational logic ISCAS89 benchmark circuit s27 proved the performance of this technique for selecting relevant parameters. Indeed, up to 2.5× speed up in *Monte Carlo* (MC) is obtained for timing variation analysis with the variance error of 11.7%.

## 4.1 Background and Motivation

In the nanoscale era, modeling and simulation of VLSI circuits have been facing a significant challenge called "curse of dimensionality". Due to the extra process complexity required to build deeply scaled devices, the number of device parameters affected by inter-die and intra-die variations dramatically grows [71]. The variation modeling requires distinct variables for each physical and structural parameter in order to represent the effect of PV. Exploiting modeling techniques such as *Response Surface Model* (RSM) technique is not applicable anymore because the complexity of the model is exponential with respect to the number of parameters [72]. Fortunately, all of these parameters are not independent and therefore they can be partitioned into several sets of correlated parameters. By considering the correlation among parameters of each set and understanding the contribution of each set on the output, it is possible to substantially reduce the model complexity by selecting the most statistically significant parameters. Thus, new methodologies are required to reduce the number of variables while keeping the estimation error fairly small.

### 4.1.1 Principal Component Analysis (PCA)

PCA has been widely used in the field of device compact modeling [68] and statistical static timing analysis [73]. The PCA performs a linear transformation through the conversion of correlated parameters into a smaller set of new uncorrelated parameters, called principal components. Indeed, the parameter space is transformed to new coordinates in which the largest variance of the data is projected to the first few principal components. Then, the principal components, which have the maximum variations in the parameter space, are selected as it follows.

Let us consider an $n$-dimensional input vector $\mathbf{x} = [x_1, x_2, \cdots, x_n]^\top$, which has zero mean and multivariate Gaussian distributions. Assume that the correlation of components in $\mathbf{x}$ is represented by the covariance matrix $\Sigma$. Using eigen-decomposition procedure, PCA computes $\Sigma$ as:

$$\Sigma = \mathbf{E}\,\mathbf{A}\,\mathbf{E}^\top \tag{4.1}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$ is a diagonal matrix of $\Sigma$ eigenvalues, and $\mathbf{E} = [\mathbf{e}_1, \mathbf{e}_1, \cdots, \mathbf{e}_n]$ contains the corresponding orthogonal eigenvectors. Fig. 4.1 illustrates the projection of a multivariate Gaussian distribution for which the vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are corresponding orthogonal eigenvectors computed by PCA. By including few eigenvectors of $\mathbf{E}$ that have the largest eigenvalues into the projection matrix $\mathbf{E}_{red} \in \mathbb{R}^{n\times d}$ $(d \ll n)$, the new parameter set that has a smaller dimension than that of the original set can be obtained by:

$$\mathbf{x}_{red} = \mathbf{E}_{red}^\top\,\mathbf{x} \tag{4.2}$$

where $\mathbf{x}_{red}$ is a $d$ dimensional encoding of the $\mathbf{x}$.

As a main limitation of the PCA, it only focuses on the correlation among the input parameters and discards the dependency between the input parameters and the corresponding outputs. Therefore, a set of parameters may be selected that have no considerable impact on the output of the model. Moreover, when the underlying statistical information about the distribution of the input parameters is unknown, PCA fails to select the relevant parameters that contribute to the model output. Last but not least, the maximum performance can be obtained when the distribution of input parameters is Gaussian [74].

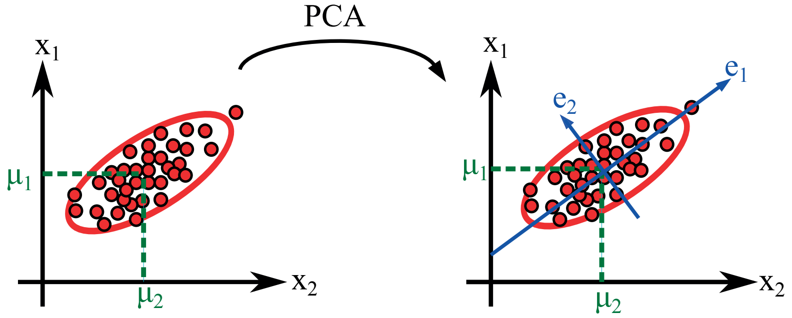


Figure 4.1: The visual representation of the PCA. PCA finds the eigenvectors of the covariance matrix Σ. The obtained orthogonal vectors represent the directions in which the variations of input parameters are maximized.

### 4.1.2 Independent Component Analysis (ICA)

For a Gaussian distribution, uncorrelatedness implies statistical independence which means that the principal components are also statistically independent. However, such a property does not hold for general non-Gaussian distributions. In Equation (2), the random vector **x** consists of correlated non-Gaussian random variables, and a PCA transformation would not guarantee statistical independence for the components of the transformed input parameters. Since the PCA technique focuses only on second order statistics, it can only ensure uncorrelatedness, and not the much stronger requirement of statistical independence.

ICA is a statistical technique that precisely transforms a set of non-Gaussian correlated parameters to a set of parameters that are statistically as independent as possible, through a linear transformation. Given a linear mixture of $n$ independent components such as $\mathbf{x} = [x_1, x_2, \cdots, x_n]^\top$, that are the correlated non-Gaussian parameters, the $n$ statistically independent components like $\mathbf{s} = [s_1, s_2, \cdots, s_n]^\top$ can be obtained as follows:

$$\mathbf{x} = \mathbf{A}\,\mathbf{s} \tag{4.3}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$ is a transformation matrix. Similar to PCA, the independent components of vector **s** are mathematical abstractions that cannot be directly observed. The ICA technique requires centering and whitening the vector **x**, and leads to variables with zero mean and unit variance. The goal of ICA is to estimate the elements of unknown transformation matrix **A**, and the samples of statistically independent components of vector **s** given only the samples of the observed vector **x**. Equation (4.3) can also be written as:

$$\mathbf{s} = \mathbf{W}\,\mathbf{x} : s_i = \mathbf{w}_i\,\mathbf{x} = \sum_{j=1}^{n} w_{ij} x_i \; for \; i = 1, \cdots, n \tag{4.4}$$

Here, $\mathbf{W} \in \mathbb{R}^{n\times n}$ is the inverse of the unknown mixing matrix $\mathbf{A}$. Figure 4.2 represents a hypothetical multivariate distribution along with the corresponding independent components ($s_1$ and $s_2$). It is obvious that ICA has found the original components by relaxing the constraint that all the identified directions have to be orthogonal. Algorithms for computing ICA estimate the vectors $\mathbf{w}_i$ that maximize the non-Gaussianity of $\mathbf{w}_i\, \mathbf{x}$ by solving a nonlinear optimization problem. This can be performed by using *kurtosis*, *neg-entropy*, and *mutual information* as typical methods measuring non-Gaussianity [75].

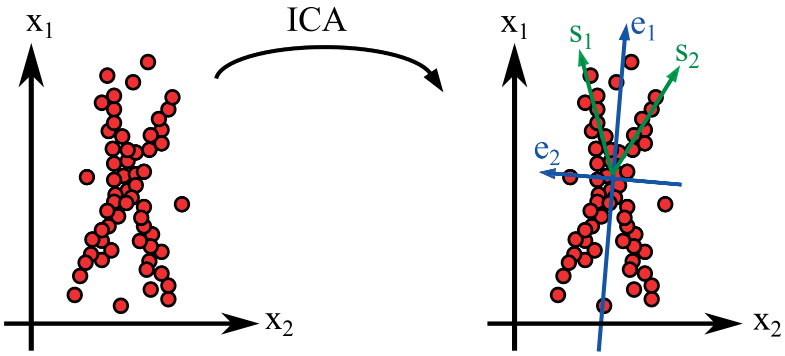


Figure 4.2: The visual representation of the ICA. ICA computes the independent components like $s_1$ and $s_2$ which can be used for parameter selection of non-Gaussian distributions. Here, the obtained components are not necessarily orthogonal.

In contrary to PCA, ICA is used for feature reduction of non-Gaussian parameters. When more than two parameters follow the Gaussian distribution, ICA fails to find the constructive components [76]. ICA like PCA is *output ignorant* which means that the parameters with minor impacts on the outputs may be selected, and important information may be lost during the dimensionality reduction.

### 4.1.3 Canonical Correlation Analysis (CCA)

As an output sensitive statistical method, CCA is capable of reducing the parameters which have major impact on the output. Suppose that the relationship between model parameters $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n]$, and model outputs $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_m]$, can be estimated by the following regression:

$$\mathbf{Y}_k = \mathbf{X}_k\, \mathbf{A}_k + \varepsilon_k \tag{4.5}$$

where $\mathbf{X}_k \in \mathbb{R}^{k\times n}$ and $\mathbf{Y}_k \in \mathbb{R}^{k\times m}$ are matrices containing the samples of the $\mathbf{x}$ and the corresponding $\mathbf{y}$, $\mathbf{A}_k \in \mathbb{R}^{n\times m}$ is a matrix to project the $n$-dimensional parameter space onto an $m$-dimensional output space, and $\varepsilon$ is a zero-mean random error of the regression. As we assume that the components are correlated in the input space, CCA modifies the Equation (5) to:

$$\mathbf{Y}_k = \mathbf{X}_k\, \mathbf{B}_k\, \mathbf{C}_k + \varepsilon_k \tag{4.6}$$

where $\mathbf{B}_k \in \mathbb{R}^{n\times r}$ projects the input parameters to an $r$-dimensional space so that $r \le n$ and $\mathbf{C}_k \in \mathbb{R}^{r\times m}$ is the transforming matrix of regression for the projected input parameters. Therefore, the dimensionality of the input parameter space is reduced from $n$ to $r$ and then a smaller linear model relates the projected space to the output space. The problem is reduced to the estimation projection matrix ($\mathbf{B}_k$), via the following objective function:

$$\max_{\mathbf{w}_i} \mathbf{b}_i^\top \mathbf{X}_k^\top \mathbf{Y}_k \mathbf{Y}_k^\top \mathbf{X}_i \mathbf{b}_i \tag{4.7}$$

where $\mathbf{X}_i \mathbf{b}_i$ are the components of projected input parameters. Fig. 4.3 shows how CCA selects a direction in the input space which has the maximum correlation with the projected output space.

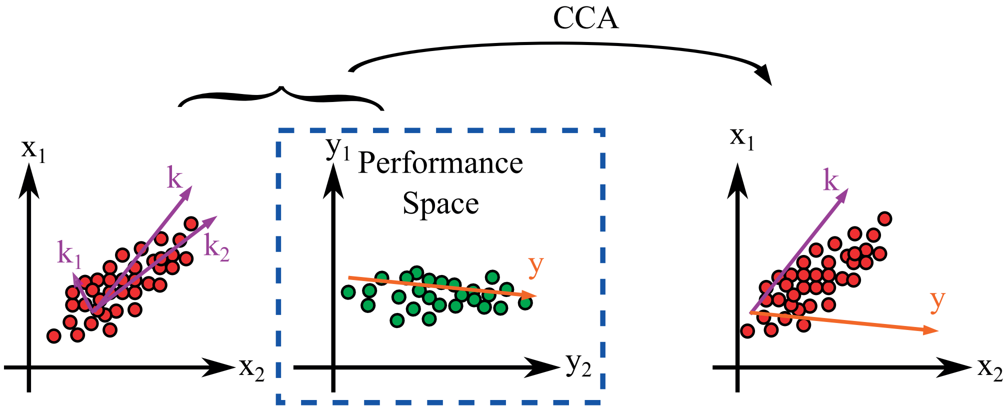


Figure 4.3: The visual representation of the CCA. CCA considers the variation of output parameters in the performance space and tries to find the components that have larger impacts on the output variation. Indeed, CCA performs an output aware parameter reduction using linear correlations.

Similar to previous methods, CCA strictly requires a Gaussian distribution of input variables to significantly enhance the performance of feature reduction. However, variation analysis of deeply nanometer scaled technologies has revealed that the distribution of several parameters, such as $V_{th}$, does not follow a Gaussian distribution [77]. Thus, the performance of feature selection may be considerably affected by the distribution of input parameters. Furthermore, in CCA like other linear models input parameters are considered independent, while several geometrical parameters of the transistor, e.g., gate length and $V_{th}$ are correlated to one another [76].

### 4.1.4 Sparse Linear Regression via $\ell_1$-Norm Regularization

Sparsity via $\ell_1$-norm regularization is a learning-based feature selection method [78]. This method focuses on the cases where the number of samples is less than the number of coefficients. In this case, the solution (i.e., the model coefficients) is not unique, unless exploiting several additional constraints. As a result, sparsity can be used to uniquely determine the values. For a vector of input parameters such as $\mathbf{w}$, $\ell_1$-norm regularization technique is used to find the most important parameters subject to the following objective function:

$$min\ l = \|\mathbf{wx} - \mathbf{y}\|_2^2 + \lambda\|\mathbf{w}\|_1 \tag{4.8}$$

where $\|\cdot\|_2$ and $\|\cdot\|_1$ represent the $\ell_2$-norm and $\ell_1$-norm of a vector, respectively. The $\ell_1$-norm ($\|\mathbf{w}\|_1$) gives us the sum of the absolute non-zero elements of the $\mathbf{w}$. Indeed, it measures the sparsity of $\mathbf{w}$ in the regression model. Therefore, $\ell_1$-norm regularization attempts to find a sparse solution that minimizes the least-square error. $\lambda$ is a hyper parameter in Equation (4.8) that controls the tradeoff between the sparsity of the input parameters and the minimal value of the loss function $\|\mathbf{wx} - \mathbf{y}\|_2^2$. For example, a large $\lambda$ value will result in a small error function, but it will increase the number of non-zero elements in $\mathbf{w}$. It is important to note that a small error function does not necessarily mean a small modeling error. Although this method can find linear dependencies between input and output parameters, it suffers from lack of modeling nonlinear relations among parameters.

### 4.1.5 Parameter Reduction for PV Analysis

In order to handle the high dimensionality of the circuits' models in presence of process variation, parameter reduction is necessary to find the intrinsic dimensionality of the models. The intrinsic dimensionality of the models is the minimum number of PV parameters needed to account for variation analysis. In the framework of PV analysis, the applicable parameter reduction method needs to capture the nonlinearity among process parameters and performance parameters. The simple construction of process parameters from the reduced space is necessary for experimental simulations. Last but not least, the reduction method should be able to handle PV variables with different statistical distributions.

The methods mentioned above are linear. Considering nonlinear dependencies can remarkably increase the precision of parameter reduction. Many modifications [76] have been proposed to alleviate this problem, e.g., *Function Driven Component Analysis* (FCA), quadratic RRR, Kernel PCA, and Kernel ICA. Kernel-based methods try to address this issue by using fixed nonlinear kernels, e.g., quadratic, polynomial, and exponential functions. They map the input space to a higher dimensional space, and then linearly relate the model to the output space. This has several limitations: it increases the dimensionality of problem before reducing it, and, more importantly, it assumes a known nonlinear relationship between the input and the output spaces.

Moreover, these methods perform dimensionality reduction, meaning that the problem is transformed from an input parameter space to a reduced parameter space. Since these transformations change the meaning of physical parameters, either we need to reconstruct the original parameters from the reduced parameters, or modify the PV simulator to work with the new set of parameters. Modifying device and process simulators like TCAD simulators is very challenging. Moreover, due to the nonlinearity of these transformations, it is extremely costly to reconstruct the original parameters from the lower dimension space. While the above nonlinear methods increase the precision, but they can not be used efficiently in our applications. Therefore, a parameter selection method in the input space, that considers the nonlinear relation between the input and the output spaces, is then proposed in the following section. The method accelerates statistical PV analysis and addresses the major drawbacks of the previous work.

## 4.2 Learning-based Parameter Reduction for Fast Variation Analysis of Emerging Devices

In this section, we present a learning-based feature selection method adapted to VLSI modeling and simulation. We overview the framework of parameter selection, and then discuss the method in detail.

### 4.2.1 Parameter Reduction towards Low Dimensional Device and Circuit Models

In order to achieve fast PV analysis for digital ICs, large designs have to be partitioned into a set of logic cells. The size of each logic cell should be small enough such that the parameter selection can be efficiently performed. After extracting the variation parameters, logic cells are hierarchically clustered to form the initial large circuit. Then, the parameter selection can be performed again on each cluster with the new reduced parameter set to completely cover the targeted large circuit. In most cases, the circuits that we want to model are known to be structured in the sense that their physical parameters are highly correlated and therefore the associated models are compressible. Considering the correlation among parameters provides an opportunity by which the circuit functionality can be estimated with smaller number of parameters which leads to a lower computational complexity.

Fig. 4.4 illustrates the general flow of the proposed parameter reduction method for circuit PV analysis. First, input and output parameter sets are selected according to the hierarchy level at which the parameter reduction is performed. The input parameter set can be obtained from three different sources: compact model parameters of the device, parameters of the TCAD model, or measured characteristics of the fabricated devices such as *Threshold Voltage* ($V_{Th}$), $I_{On}$, $I_{Off}$, and *Subthreshold Slope* ($SS$). The output parameter set can also be selected among delay, power consumption, or any other functionality criteria of the logic cells and circuit blocks. In the next step, a learning-based statistical multivariate regression is used to predict the relations among the input and output parameter sets. The objective function of the regression is modified to minimize the error of the output prediction while discarding the unnecessary parameters. Here, training the regressor under the constraint of a limited error bound is the major step toward parameter reduction. Finally, the most significant parameters are only considered for the PV analysis of the target circuit, whereby increasing the evaluation speed.

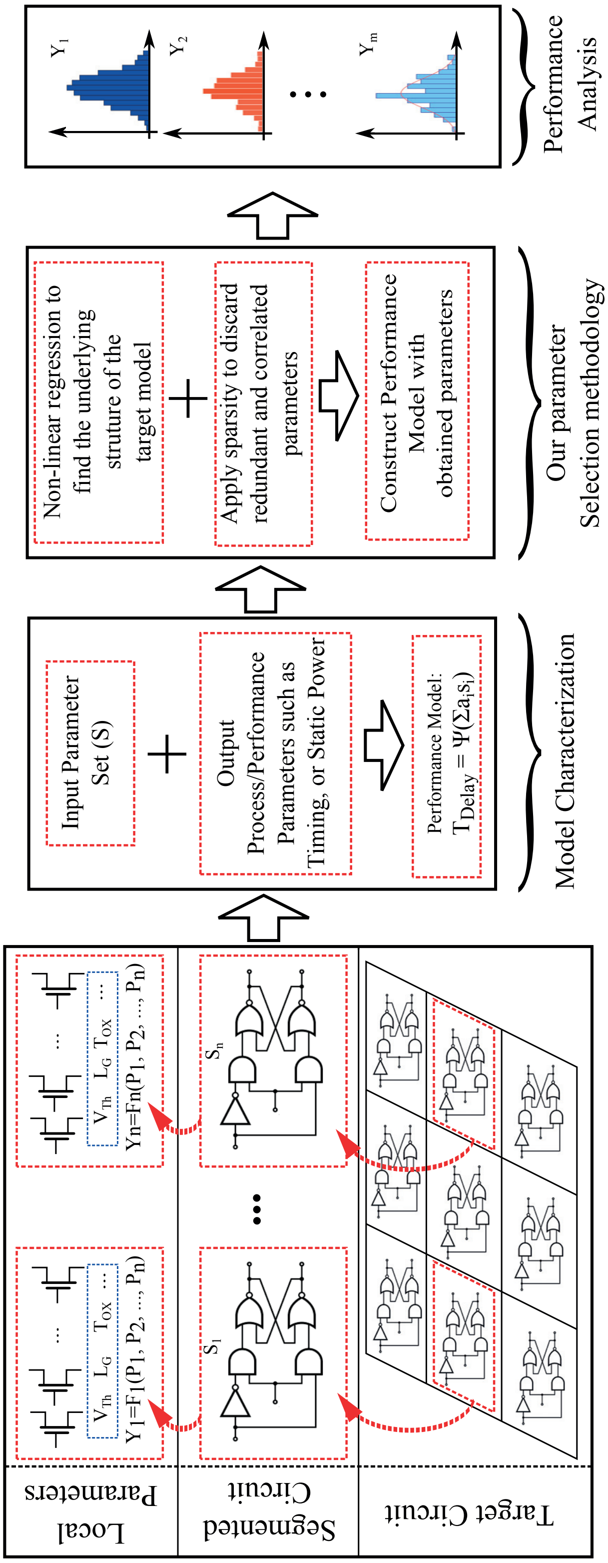


Figure 4.4: General flow of the parameter reduction towards a fast and efficient process variation analysis.

### 4.2.2 Nonlinear Regression via Feed Forward Neural Network

*Feed Forward Neural Network* (FFNN) is a powerful nonlinear regressor known to be a universal approximator by increasing the size of hidden layer [61]. We adopt FFNN here as our regressor to consider the nonlinear relations among the parameters. The regression model is formulated as:

$$\mathbf{y} = \mathbf{W}' \, tanh(\mathbf{W}\mathbf{x}^T) + \epsilon \tag{4.9}$$

where $\mathbf{x} \in \mathbb{R}^{1\times m}$ is a vector that represents a sample value of the input set. The vector $\mathbf{y} \in \mathbb{R}^{1\times n}$, of length $n$, represents the corresponding output. $\mathbf{W} \in \mathbb{R}^{k\times m}$ is a transformation matrix in which $k$ is the size of the hidden layer. It transforms each input feature to a space formed by hidden units. $\mathbf{W}' \in \mathbb{R}^{n\times k}$ is a matrix that forms the output from the hidden layer. Vector $\epsilon$ represents the error of estimation in comparison with target objectives and $tanh$ is a nonlinear activation function that is chosen conventionally. In the case of multiple outputs, $(\mathbf{Y} \in \mathbb{R}^{r\times n},\ r > 1)$, we handle each output independently.

To find the best fitting model we perform the following optimization over the objective function:

$$\underset{\mathbf{W},\mathbf{w}'}{\operatorname{argmin}} L(\mathbf{W}, \mathbf{W}') = \frac{1}{2} \left\| \mathbf{y} - \mathbf{W}' \, tanh(\mathbf{W}\mathbf{x}^T) \right\|_2^2 \tag{4.10}$$

The above optimization minimizes the prediction error of the model using all $m$ parameters.

### 4.2.3 Learning Algorithm for Nonlinear Optimization

The *Levenberg-Marquardt* (LM) algorithm is used for learning the parameters of the FFNN [79]. LM benefits the *Steepest Descent* (SD) and *Gauss-Newton* (GN) algorithms to avoid finding local optimums. The LM algorithm combines the SD and GN in the following manner:

$$\mathbf{x}_k = \mathbf{x}_{k-1} - (\mathbf{J}_{k-1}^{\top} \, \mathbf{J}_{k-1} + \mu \mathbf{I})^{-1} \, \mathbf{J}_{k-1}^{\top} \, \mathbf{e} \tag{4.11}$$

where $\mathbf{J}$ is the Jacobian matrix which contains first derivatives of the FFNN errors, $\mathbf{e}$ is a vector of FFNN errors in the last step, $\mathbf{x}$ is a vector of unknown parameters $w_{i,j}, k_i, b_i$ that are obtained after training, and $\mu$ is a hyper parameter that offers a balance between SD and GN during the learning iterations. The Jacobian matrix can be computed via back-propagation technique that is much less complex than computing the Hessian matrix.

In Section 4.2.6, we design a function to reward sparsity over input parameters and add that function to the above optimization. Thus, we can find the set of significant parameters that can predict the output precisely.

### 4.2.4 Validation via 10-fold Cross-Validation

Cross-Validation is a frequently used method to avoid overfitting on training set and improve the quality of trained model [80]. In order to perform 10-fold cross validation, the training set is randomly divided into 10 separate sub-sets of equal size. Then, training procedure is performed 10 times, each time discarding one set as a test set, and the average error over all the runs is computed. Finally, the trained model with the lowest error is selected. This has the additional benefit of avoiding local optimums for the trained model.

### 4.2.5 Column-Wise Sparse Parameter Selection

Our proposed parameter reduction technique inspired by $\ell_1$-norm regularization method. If $x$ represents one input sample, we can reformulate the $\ell_1$-norm regularization loss function as the following:

$$\mathbf{y} = \mathbf{W}' tanh(\sum_{i=1}^{m} \mathbf{W}_i \mathbf{x}^\top) + \epsilon \tag{4.12}$$

in which the vector $\mathbf{W}_i$ represents the column $i$ of the matrix $\mathbf{W}$. The contribution of each input parameter corresponds to a column of the matrix $\mathbf{W}$. To select few number of parameters, we need to learn $\mathbf{W}$ as a column-wise sparse matrix. If the matrix is column-wise sparse, it means that there are several columns of all zeros and the corresponding parameters do not have any contribution in the model. Consequently, the significant parameters are the ones with the corresponding non-zero columns.

To achieve the column-wise sparsity, we measure the sparsity on the vector consisting of the maximum of the columns: $||max(\mathbf{W}_1)\cdots max(\mathbf{W}_m)||_0$. If the entry with maximum value in a column is pushed towards zero, we expect all the other values in a column become zeros. It is a common practice to approximate norm-zero with norm-one to achieve a sparse answer while making the optimization easier. But, still the optimization is almost impossible because of the discrete $max$ function applied on the columns. Large norm function provides a continuous approximation of the maximum function (infinity norm is equal to $max$). Therefore, we approximate the $max$ function with the $p$-norm function ($p \geq 2$):

$$\|\mathbf{v}\|_p = (\sum_{i=1}^{n} |v_i|^p)^{\frac{1}{p}} \tag{4.13}$$

We choose $p$ large enough that achieves a column-wise sparse answer on a held-out data set. Similarly in group lasso [81] combination of norm 1 and 2 is used to achieve a linear group-wise sparse model.

Fig. 4.5 schematically represents the concept of column-wise sparsity. The norm-$p$ ($p$ is selected reasonably big) is applied to $\mathbf{W}$ in order to compute the maximum element of each column. Then, norm-one is applied to the vector of obtained values to impose the sparsity.

Thus, the column-wise sparsity is measured by $||||\mathbf{W}_1||_p \cdots ||\mathbf{W}_m||_p||_1$. In the following, we present how the column-wise sparsity is applied on an FFNN regressor to form a feature selector.

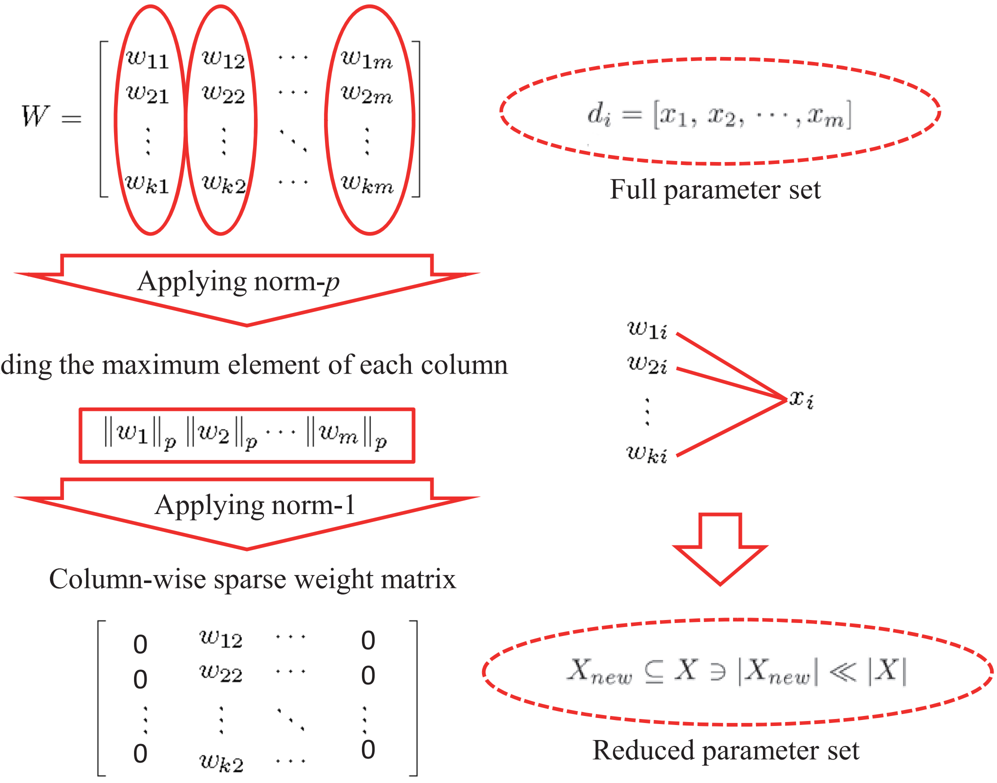


Figure 4.5: The role of norm-p regularization in weight matrix for feature selection.

### 4.2.6 Nonlinear Column-Wise Sparse Parameter Selection

In order to find the reduced input set, the sparsity objective function is added to the regressor. Putting the FFNN regressor and column-wise sparsity together, the objective function becomes:

$$\underset{\mathbf{W},\mathbf{W}'}{\operatorname{argmin}} L(\mathbf{W},\mathbf{W}') = \frac{1}{2}\left\|\mathbf{y} - \mathbf{W}'\,tanh(\mathbf{W}\mathbf{x}^T)\right\|_2^2 + \lambda||||\mathbf{W}_1||_p \cdots ||\mathbf{W}_m||_p||_1$$

The first term of the objective function is called *loss function* and tries to minimize the error of regression. The second term is called *regularization term* which controls the number of parameters in regression.

Feature selection can be used whenever the values of the $\mathbf{W}$ and $\mathbf{W}'$ are obtained. Algorithm 1 represents the steps of learning for the column-wise sparse feature selection method. In each iteration, the gradient of objective function is computed to update **W** and $\mathbf{W}'$ (Algorithm 1 - l. 6). The algorithm continues either to reach the defined bound on the error or to end at the maximum learning iterations (Algorithm 1 - l. 3). Thus **W** and $\mathbf{W}'$ are learned during the training process. The $\lambda$ and $p$ are model hyper parameters. The $\lambda$ value controls the number of parameters in the regression model. As the $\lambda$ value increases, the objective function shrinks the weights in **W** in a column-wise manner towards zero. Thus, the bigger $\lambda$ value forces more parameters toward zero and reduces the parameter space.

```
Algorithm 1: Nonlinear Multi-Objective Parameter Selection
input  : $\mathbf{x}_i$ = {Input vector}, $\mathbf{y}_i$ = {Output vector}, $\epsilon$ = Error bound, $\lambda$ = Regularization
         parameter, $\mathbf{M}$ = Maximum number of iterations
output : $\mathbf{W},\mathbf{W}'$ = Matrix of transition weights
1:  Initialize $\mathbf{W},\mathbf{W}'$;
2:  $Iter \leftarrow \emptyset$, $E \leftarrow \emptyset$;
3:  while $|E| \leq \epsilon$ or $Iter \leq M$ do
4:      $E \leftarrow \emptyset$;
5:      for $i$=1:$n$ do
6:          Set the objective function $L \leftarrow \frac{1}{2}\|\mathbf{y}-\mathbf{W}' tanh(\mathbf{W}\mathbf{x}^T)\|_2^2 + \lambda\| \|\mathbf{W}_i\|_P \|_1$;
7:          Compute the error ($E \leftarrow \frac{1}{2}\|\mathbf{y}-\mathbf{W}' tanh(\mathbf{W}\mathbf{x}^T)\|_2^2$);
8:          Calculate the gradient of objective function in order to update weights;
9:          $\mathbf{W},\mathbf{W}'$ ← Gradient-based optimization ($\mathbf{W}, \mathbf{W}', \frac{\partial L}{\partial W}, \frac{\partial L}{\partial W'}$);
10:     $Iter \leftarrow Iter + 1$;
11:     $E \leftarrow \frac{E}{n}$;
12: return $\mathbf{W}$;
```

## 4.3 Experimental Results

This section evaluates the proposed method by applying the column-wise feature selection to a set of combinational and sequential logic benchmark circuits in the context of regular and emerging technologies. The focus of our study is on the timing variation analysis. The use of this method is motivated by the lack of intuition that a skilled designer may have to identify the critical parameters of novel devices. We first look at FinFET as a cutting edge technology. Design verification in FinFET technology is complex because of the novel three-dimensional structure of the devices. We then look at a further interesting technology, silicon nanowires, that are also three-dimensional structures with specific features.

### 4.3.1 PV analysis for FinFET Technology

In this section we present the application of described method for timing variation of FinFET-based circuits.

### Setup of Experiments

To evaluate the proposed parameter selection technique, we exploit a number of combinational and sequential logic circuits from ITC'99 and ISCAS benchmarks. We study the timing variation of the longest path for each benchmark circuits. The longest path of each benchmark circuit is extracted using Synopsys PrimeTime [82] as exemplified in Fig. 4.6. In our analysis, the input parameter set includes the $V_{Th}$ of each transistor within the circuit. Here, $V_{Th}$s are selected as they can significantly reflect the variation of physical parameters of each transistor on its performance. A transistor pool, which contains 5000 $n$-type and $p$-type transistors in 20nm FinFET technology, is generated by applying a 5% Gaussian variation on the $V_{Th}$ of each transistor. To build each circuit instance, the transistors are randomly selected from the pool and are added to the SPICE model of the target circuit. Using the obtained SPICE model, we can assess the timing variation through MC simulations but require tremendous amount of time. By applying the proposed parameter selection method, we show how this sampling space can be limited to the most important input parameters that mainly impact the timing of the circuits.

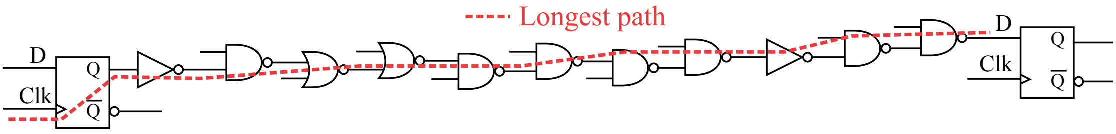


Figure 4.6: Longest path in ITC'99 b03 benchmark.

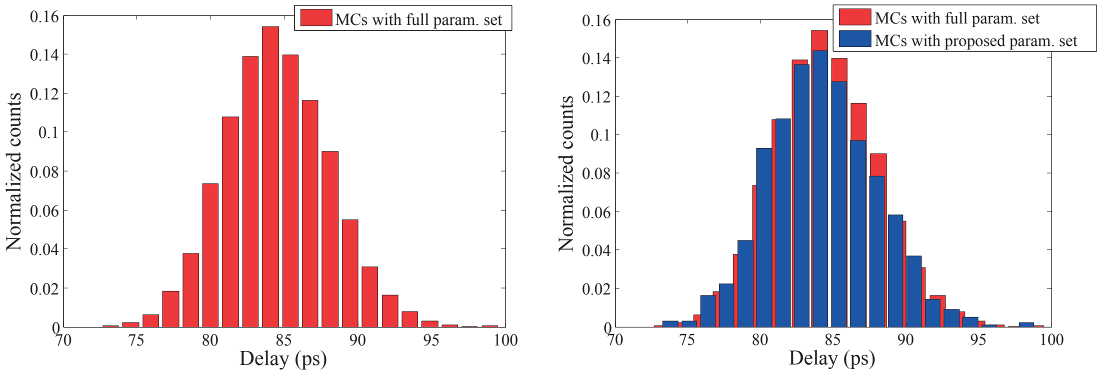
### Parameter Reduction and Simulation Speed-up

We performed 10,000 MC simulations to extract the distribution of the delay for each benchmark by applying variation on all the parameters (The total number of parameters are listed in Table 4.1). We then trained our parameter selection method over 1,000 MC simulations to pick out the 20% most important parameters from each benchmark. In our method, the reduced set of input parameters is achievable through increasing the value of $\lambda$ till reaching a desirable balance between the performance estimation accuracy and the number of selected input parameters. After extracting the important parameters, we perform 1,000 SPICE simulations for each target benchmark by applying variations on the selected parameters. Table 4.1 demonstrates the mean and standard deviation of the longest path delay for each benchmark before and after parameter reduction. The results reveal average errors of 1.2% and 3.2% on the mean and the standard variation values respectively.

Table 4.1: Comparison of mean and variance of various ITC'99 and ISCAS benchmarks on the delay variation of longest path before and after applying proposed parameter reduction.

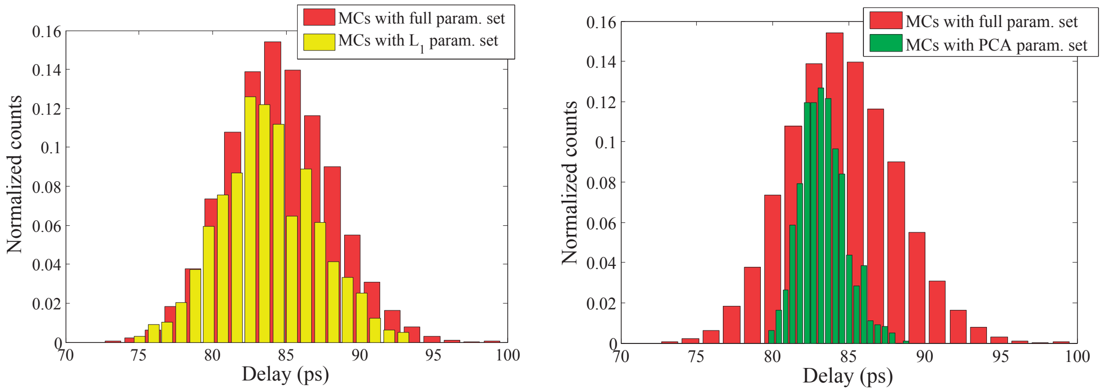
| | | | Full-parameter set | | Reduced-parameter set | | | |
|---|---|---|---|---|---|---|---|---|
| | Benchmark | # of param. | Mean | Std | Mean | Std | Mean-error (%) | Std-error (%) |
| ITC'99 | b01 | 74 | 5.372e-11 | 3.160e-12 | 5.335e-11 | 3.208e-12 | 0.69 | 1.52 |
| | b02 | 80 | 5.506e-11 | 2.980e-12 | 5.457e-11 | 2.955e-12 | 0.88 | 0.86 |
| | b03 | 88 | 8.451e-11 | 3.547e-12 | 8.429e-11 | 3.686e-12 | 0.26 | 3.91 |
| | b04 | 116 | 1.237e-10 | 4.156e-12 | 1.226e-10 | 3.847e-12 | 0.84 | 7.44 |
| | b05 | 96 | 9.031e-11 | 3.707e-12 | 8.970e-11 | 3.615e-12 | 0.68 | 2.49 |
| | b06 | 76 | 6.611e-11 | 3.113e-12 | 6.528e-11 | 2.894e-12 | 1.25 | 7.05 |
| ISCAS | C432 | 238 | 2.687e-10 | 1.529e-11 | 2.660e-10 | 1.433e-11 | 0.99 | 6.28 |
| | C880 | 242 | 2.424e-10 | 7.569e-12 | 2.398e-10 | 7.262e-12 | 1.08 | 4.06 |
| | S386 | 82 | 7.544e-11 | 7.733e-12 | 7.480e-11 | 3.722e-12 | 0.86 | 0.30 |
| | S641 | 172 | 1.485e-10 | 1.369e-11 | 1.480e-10 | 1.272e-11 | 0.32 | 7.09 |

The distributions of the longest path delay for ITC b03 benchmark are depicted in Fig. 4.7 before and after parameter reduction. The distribution in Fig. 4.7a is obtained through 10,000 MC simulations over 88 variation parameters. We reduced the size of the input parameter space using proposed parameter selection method, and two baselines $\ell_1$-Norm regularization and PCA. We did not perform our experiments using ICA and CCA baseline methods since ICA failed to find the reduced input parameters having Gaussian distributions, and $\ell_1$-norm regularization surpasses the CCA method [83]. The distributions in Figs. 4.7b, 4.7c, and 4.7d are attained through 1,000 MC simulations but using only 17 parameters selected by three mentioned parameter selection methods. The variation sampling using the most relevant parameters, obtained by our method, is capable of estimating the timing variation distribution with a smaller error ($\sigma = 3.68ps$ as compared to $\sigma = 3.54ps$) than $\ell_1$-Norm regularization (($\sigma = 3.30ps$ as compared to $\sigma = 3.54ps$) and PCA (($\sigma = 1.48ps$ as compared to $\sigma = 3.54ps$). Therefore, the proposed parameter selection can efficiently reproduce the timing variation with a small subset of input parameters. The number of required MC simulations is reduced by 5× (2,000 simulations for training and reproducing the distribution of timing variation vs. 10,000 MC simulations without parameter selection).



(a) Distribution of longest path delay for benchmark b03 using 10,000 MC simulations over 88 parameters.

(b) Distributions of the longest path delay for benchmark b03 using 1,000 MC simulations using 17 parameters obtained by the proposed method (blue bars) vs. 10,000 MC simulations using full parameter set (red bars).



(c) Distributions of the longest path delay for benchmark b03 using 1,000 MC simulations using 17 parameters obtained by the $\ell_1$-norm regularization method (yellow bars) vs. 10,000 MC simulations using full parameter set (red bars).

(d) Distributions of the longest path delay for benchmark b03 using 1,000 MC simulations using 17 parameters obtained by the PCA method (green bars) vs. 10,000 MC simulations using full parameter set (red bars).

Figure 4.7: Distribution of longest path delay for benchmark b03 before and after parameter selection: (a) full parameter set, (b) proposed method parameter set, (c) $\ell_1$-norm regularization parameter set, and (d) PCA parameter set.

### 4.3.2 PV analysis for Double-Gate Silicon Nanowire Technology

Finally, we demonstrate the result of parameter selection for the variation analysis of a benchmark circuit in double-gate silicon nanowire technology.

#### Double-Gate Silicon Nanowire Technology

To perform variation analysis, we first characterize a population of devices by TCAD simulation using a 30% Gaussian variation on each geometrical parameter ($\sigma = 30\%$). In our case study, 2,500 3-D TCAD simulations were performed to provide statistical information of the DG-SiNWFET device. Fig. 4.8 depicts the distinctive analytical metrics of the device such as $I_{on}$, $I_{off}$, $V_{Th}$, and $SS$. Only the distribution of $V_{Th}$ can be approximated by a Gaussian distribution contrary to the remaining metrics. This result highlights that the distributions of all parameters are not necessarily Gaussian. Thus, the distribution free parameter reduction techniques are required for future technologies.

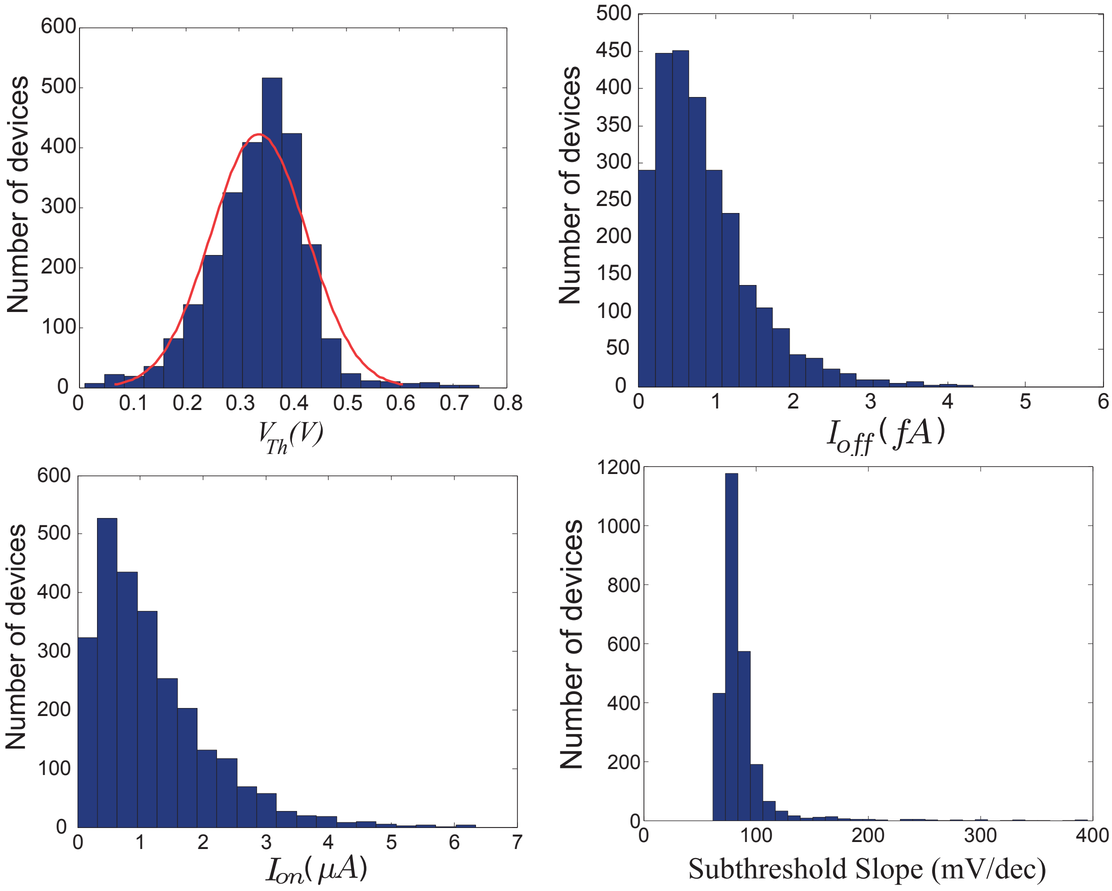


Figure 4.8: Distribution of $V_{Th}$, $I_{off}$, $I_{on}$, and $SS$ for DG-SiNWFET ($\sigma$=30% for structural parameters). Only the variation of $V_{Th}$ follows a Gaussian distribution.

### Setup of Experiments

For evaluating the proposed parameter selection method, the small size benchmark circuit ISCAS89-s27 is selected as a case study. Without loss of generality, the method can be used for any other circuits. The main reason to select such a small size circuit is the long computation time of the TCAD simulations to produce the DG-SiNWFET device data set due to the lack of a mature compact model. In other technologies, compact models can be used to accelerate the data set generation. The schematic of the circuit is shown in Fig. 4.9. All the gates use DG-SiNWFET transistors. The PG of each transistor is appropriately configured to provide the correct functionality in the pull-up and pull-down of the gates. The considered circuit is comprised of 30 transistors leading to 300 geometrical parameters. Normal MC simulation to evaluate the performance variation requires tremendous amount of time, considering that no intuitions on the fundamental parameters can be done in the context of unconventional device mechanisms. By applying the proposed method, we show how this sampling space can be restricted to the main parameters that considerably affect the performance of the circuit.

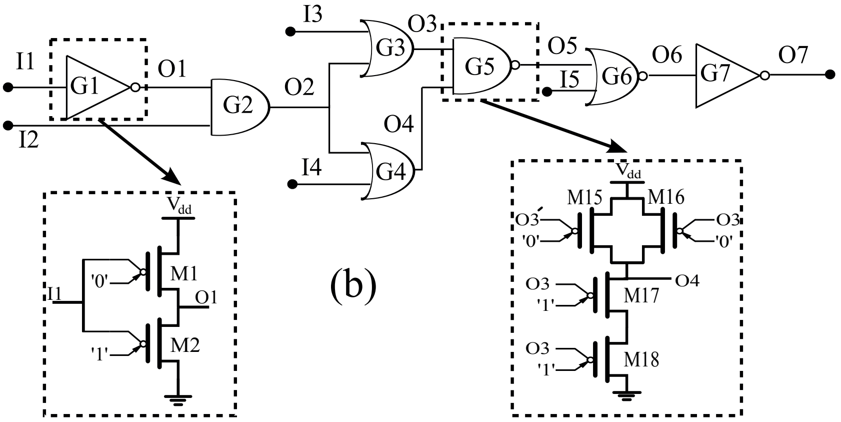


Figure 4.9: ISCAS89 benchmark circuit s27 using DG-SiNWFET technology.

Among various performance metrics, we select the delay of circuit to form the output set. For the sake of keeping a reasonable complexity for the experiments, a reduced subset of geometrical parameters of the transistors (50 parameters) is randomly considered as the input set. Here, the goal is to determine how much the parameter reduction can improve the circuit performance evaluation, while the estimation error is bounded by a certain threshold.

To simulate the characteristics of the target circuit, the obtained $I-V$ curve of the transistors, are injected in a Verilog-A table model. This model is run with HSPICE to perform the MC simulations for the timing analysis purpose.

### Parameter Reduction and Simulation Speed-up

After applying column-wise sparse parameter selection, we can reduce the number of parameters to improve the computational complexity of the simulations. Decreasing the number of parameters can be obtained by increasing the $\lambda$ value which results in larger delay estimation error. In this case, the performance of the circuit can be evaluated with a smaller number of parameters which really contribute to the MC simulations, but results in a higher performance estimation error. The capability of bounding the error by changing the numbers of parameters enables the designers to trade-off evaluation precision with computation complexity. In our case study, reducing the number of parameters to 10 (from 50) is obtained with the variance of delay estimation error of 11.7%.

We compared the proposed technique with PCA as a well-known parameter reduction methods for estimating the delay of ISCAS89-s27. For PCA, 20% of the new features were selected according to their highest eigenvalues. To be able to perform the MC simulations without any change in the underlying model or simulator, the reverse of these transformations are applied to produce the exact values of the input space parameters. In our method, $\lambda$ value was tuned to select the same number of parameters in input space.Using reduced input parameter sets obtained by PCA and the proposed method, we performed 1,000 MC simulations for each set to estimate the delay distribution of ISCAS89-s27. The proposed method shows a better performance as compared to its competitor with lower variance of delay estimation error (11.7% vs. 13.5%).

To verify the accuracy and the performance improvement of doing such reduction, we evaluate the delay of the target circuit in the presence of variations. We perform the MC simulations in both cases of reduced and non-reduced input parameter set with 10 and 50 parameters respectively. Fig. 4.10 represents the *Probability Density Function* (PDF) of the ISCAS89-s27 delay in both cases. The figure depicts a high correlation between two sets. We observe that the proposed column-wise sparsity is able to estimate the major parameters for delay variation analysis with tiny amount of error on each test samples ($\sigma = 8.91ps$ as compared to $\sigma = 10.10ps$ leading to a variance error of 11.7%). Thus, the method is able to efficiently evaluate the delay variation of the circuit, while reducing the number of parameters. A reduced input set results in less MC simulations which is very critical in the case of execution time. As we used 100 random samples for each parameter, the parameter reduction reduces the number of required MC runs by 2.5× (5,000 simulations without feature selection vs. 2,000 simulations for training and feature selection).

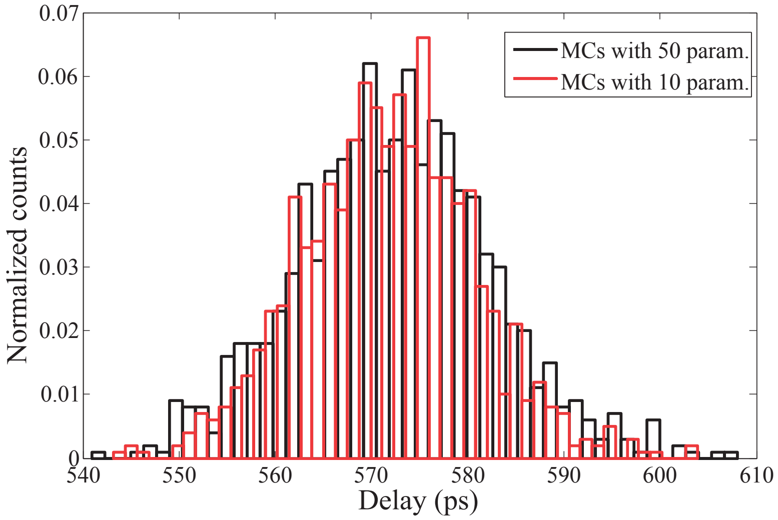


Figure 4.10: Delay distribution comparison of the full and the reduced parameter models.

## 4.4 Summary

We introduced an efficient parameter selection method which can be used for performance evaluation of the emerging technologies like *FinFET* and *Silicon Nanowires*. Using this method, we are able to accurately evaluate the process variations while reducing the computation complexity by utilizing the obtained reduced parameter set. This method is based on Feed Forward Neural Network regression, and employs column-wise sparsity to reduce the size of parameters space. Unlike the widely used feature reduction methods, this method is able to take to account the mixed Gaussian and non-Gaussian parameters. Moreover, it considers the nonlinear dependencies between input parameters and outputs which lead to effective parameter reduction. We applied this method to a couple of FinFET-based combinational and sequential benchmarks from ITC'99 and ISCAS to study the variation of delay of the longest path for each circuit. In this case, experimental results show 5× speed up and estimate the delay distribution with the average variance error of 4.1% in presence of 5% variation on each parameter. Applied to ISCAS89-s27 benchmark exploiting DG-SiNWFET technology as well, experimental results show 2.5× speed up in timing analysis and estimation of the delay distribution with the variance error of 11.7% in presence of 30% variation on each parameter.

# 5 Defect Analysis and Fault Modeling for Controllable-Polarity Silicon Nanowires

## 5.1 Introduction

Testing of integrated circuits has always been an important research area for researchers in *Very Large Scale Integration* (VLSI) community. The test engineering methods to verify the proper operation of manufactured circuits has been changed over the past decades. The migration from one manufacturing process to another shrinking process induces new types of failure mechanisms and fabrication defects which were unknown or even less important ones in previous technologies. Consequently, improved or completely new defect models and testing methods are inevitable for each technology.

To reveal fabrication defects and circuits malfunctioning, a number of structural fault models for planar single-gate CMOS and FinFET technologies have been proposed and proved to be efficient. For instance, stuck-at [84], delay [85], stuck-open [86], and bridging fault [87] are among the most commonly-used models for CMOS technology. For FinFETs, a few number of studies have been conducted in modeling defects such as floating gates and shorts [88, 89], stuck-open/stuck-on [90, 91], and *Gate Oxide Short* (GOS) [92]. These studies revealed the deficiency of current CMOS fault models for detecting all the defects in FinFET circuits, and required the introduction of new fault models for test generation purpose.

In this chapter, we perform an inductive fault analysis to investigate the specific malfunctions of CP-SiNWFETs. We used *Three-Independent-Gate Silicon NanoWire FETs* (TIG-SiNWFETs), as were previously introduced in Chapter 2. For a new technology such as CP-SiNWFET, which has different geometrical structure and physics of operation rather than CMOS technology, it is not known a priori how the manufacturing defects will impact the device and logic circuits. Considering the technology process, the possible defects that can change the functionality of the CP-SiNWFETs and occur during fabrication process are modeled. Using the obtained defect model, we investigate the functionality and the performance of various logic gates in the presence of defects. Out of the obtained results for CP-SiNWFET technology, we extend the current CMOS fault models to consider also stuck-at $p$-type and $n$-type. The results also confirm the inefficiency of traditional test methods for covering the defects, such as open defects on polarity terminals of the device, in CP-SiNWFET technology.

## 5.2 Background and Motivation

In nano-scaled technologies, process variation negatively affects the driving current of transistors and consequently results in delay faults. Moreover, undetected design rules violations increase the chance of bridging faults as unintended resistive connections between two or more conductive parts. Bridging faults could be efficiently diagnosed by supply current monitoring through IDDQ test [93] for bulk planar CMOS, but the test is becoming less effective for the deeply nanoscaled technologies [94]. *Line Edge Roughness* (LER) is an inevitable limitation of etching process and leads to non-homogeneous deposition of dielectrics when the dielectric thickness goes beyond 5nm. This may result into the *Gate Oxide Short* (GOS) [95]. Last but not least, twin boundaries during forming the nanowires may strongly influence the *On* current of the device that finally causes to channel break [96, 97].

Proper fault modeling for testing manufacturing defects plays a significant role for quality of circuits and their correct functionality. Currently, few researches have been carried out on fault modeling of the FinFET. The authors in [88, 89] investigated open and short faults on FinFETs, and they showed that *Stuck-Open Faults* (SOFs) on the back gate of FinFET have a unique effect on the leakage and delay. In [92], the GOS defect on the FinFET dielectric has been studied. The amount of *Saturation Drain Current* ($I_{D(SAT)}$) vastly increases with GOS at the front gate dielectric. However, GOS occurrence in the back-gate dielectric causes much lower carrier density in the device channel. The SOF, Stuck-on and GOS on different number of Fins in a FinFET have been examined by [90, 91]. The results manifested that when the number of faults is large enough, the defect can be captured by SOF or delay fault tests. In [86], authors presented the problem of SOF detection for small nanometer technologies. They proposed a new multiple test vector mechanism to enhance the probability of SOF detection.

While stuck-at, SOF, bridging, and delay faults efficiently model the defects in CMOS, FinFET, and CNT devices, the particular structure of devices with controllable-polarity necessitates further study to see whether these fault models can properly capture the manufacturing defects. There is no available comprehensive fault model for circuits designed with these devices. In the following, we analyze the possible manufacturing defects of CP transistors, and then investigate their impact of the functionality of the various types of CP logic gates.

## 5.3 Manufacturing Defects of TIG-SiNWFETs: from Device to Logic Cells

In this section, the possible defects of TIG-SiNWFETs are analyzed by considering major manufacturing steps. The defect model then is used for inductive fault analysis. We follow the device fabrication process and consider the layout structure of logic cells to provide the opportunity of finding the most probable possible defects. This defect model helps us to find a realistic fault model for non-classical CMOS devices.

### 5.3.1 Device Manufacturing Defects

During the nanowire patterning and etching, variations along with LER contribute to lowering the pattern sharpness that may lead to nanowire break. The defect can drastically limit the driving current of the device or lead to a stuck-open fault. When the dielectric thickness is scaled beyond the 5nm, the control for conformal oxide formation is reduced, and this eventually leads to poor insulator coverage. Consequently, *Gate Oxide Short* (GOS) defects may happen. GOS may degrade the performance of the device or even malfunctioning according to the defect size. Finally, the similar mechanism may result in bridging defect between the control gate and each adjacent polarity gate. These are the most possible defects during TIG-SiNWFETs fabrication, which are also summarized in Table 5.1.

Table 5.1: TIG-SiNWFET fabrication process steps and related defect model.

| **Controllable-polarity SiNWFET** | | | |
|---|---|---|---|
| | **Fabrication process** | **Outcome** | **Possible defects** |
| 1 | HSQ-based nanowire patterning | Initial pattern of nanowires | • Nanowire break |
| 2 | Bosch process | Nanowire formation | • Nanowire break |
| 3 | Oxide deposition | Dielectric formation | • Gate oxide short |
| 4 | Metal-gate stack deposition | Polarity and control gates | • Bridging among two or more terminals |
| 5 | Metal layer(s) deposition | Interconnections | • Bridging among interconnects<br>• Floating gates |

### 5.3.2 Logic Cell Defects

Different sources of defects such as under/over polishing, poorly planarized surfaces, and scratches contribute to a combination of open and bridging defects in logic cells during polysilicon and metal deposition [98]. Open defects form floating regions that are very challenging for test in emerging technologies since they may affect either the performance or the functionality of logic cells. Bridging defects cause unwanted cell-internal connections between a logic cell input and the power rails ($V_{dd}/Gnd$) or any other adjacent inputs. Bridging defects similarly may have local impact (logic cell performance degradation) or global impact (deteriorate the functionality of several logic cell) and become very challenging for emerging technologies. The bridging defects are extracted by considering the proximity of interconnects in the layout of the logic cells. The logic cell defects then are added to the device manufacturing defects to form our defect model for TIG-SiNWFET technology. Figure 5.1 represents the layouts of two TIG-based logic cells ((a) NOR gate and (b) 2-input XOR) with some highlighted possible defects. Note that the presented layouts rely on the *Sea-of-Tile* (SoT) physical design methodology presented in [99, 100].

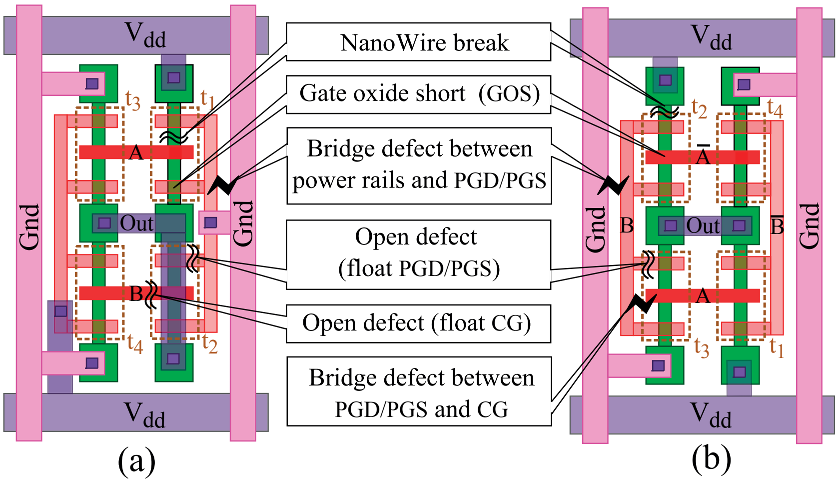


Figure 5.1: Layouts of the NOR (a) and 2-input XOR (b), based on the SoT method [99, 100], with several possible defects.

## 5.4 Fault-Modeling in Controllable polarity Silicon Nanowire Circuits

In this section, we study the behavior of SP and DP TIG-SiNWFET logic gates in the presence of fabrication defects, discussed in the previous section.

### 5.4.1 Gate Oxide Short in CP-SiNWFETs

We report the effect of *Gate Oxide Short* (GOS) occurrence in TIG-SiNWFETs. Then, we propose a new model for GOS which enables us to inject circuit level faults, and finally we study the effect of GOS on SP and DP dynamic logic circuits.

#### GOS in TIG-SiNWFETs

A GOS is a manufacturing defect happening in the oxide around the nanowire. The impact of a GOS mainly depends on its size. The defect may considerably decrease the impedance between the gate and the channel, and entirely change the electrical behavior of the transistor for a big enough GOS. For the TIG-SiNWFET, three gates ($PG_S$, $CG$, and $PG_D$) contribute to the functionality of the device. Therefore, three locations are possible for the GOS defect. The defect injection on TCAD model of the device is accomplished through replacing a tiny cuboid (10nm × 10nm) of the dielectric layer with the channel material. Thus, a conductive path is created between the defective gate and the channel. The effects of GOS on the performance of the TIG-SiNWFETs [101], and can be summarized as: 1) the parasitic current from the defective gate to the drain/source, which is proportional to the gate-drain/gate-source voltage; 2) the degradation of transistor driving current ($I_{D(SAT)}$) and the increase of *Threshold Voltage* ($V_{Th}$); 3) a sharp rise in the leakage current of the defective gate, which is proportional to the gate voltage. Figures 5.2a–5.2c depict the behavior of *n*-type TIG-SiNWFETs in the presence of GOS under $PG_S$, $CG$, and $PG_D$ respectively. Here, GOS occurrence in $PG_S$ and $CG$ results in a significant reduction of $I_{D(SAT)}$ similar to bulk CMOS. Moreover, the defect causes a weak inversion and tightens the channel for carriers, that leads to a slight increase of $V_{Th}$ ($\Delta V_{Th} = 170mV$ and $\Delta V_{Th} = 140mV$ for defective $PG_S$ and CG respectively). However, GOS effect in $PG_D$ is negligible (Figure 5.2c). In the following, we propose an equivalent lumped model for GOS defect.

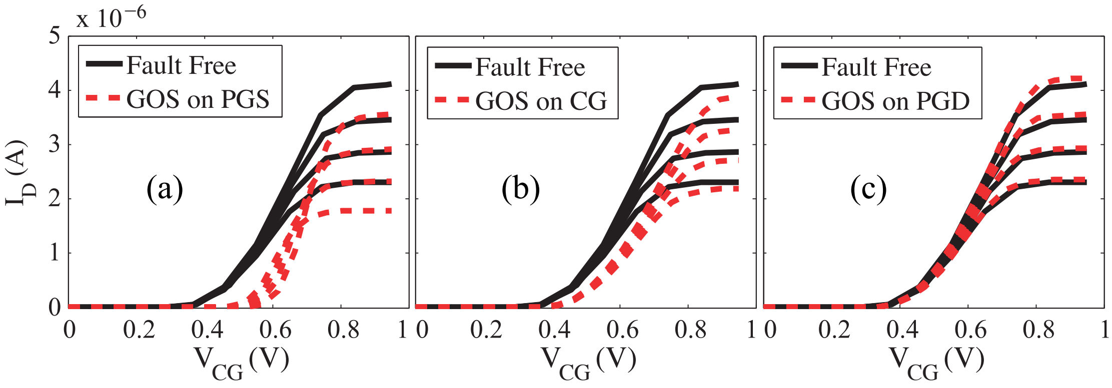


Figure 5.2: Behavior of defective *n*-type TIG-SiNWFETs in the presence of GOS: (a) GOS under $PG_S$, (b) GOS under $CG$, and (c) GOS under $PG_D$.

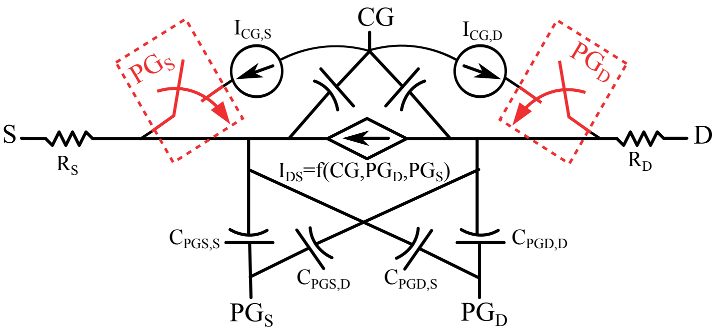
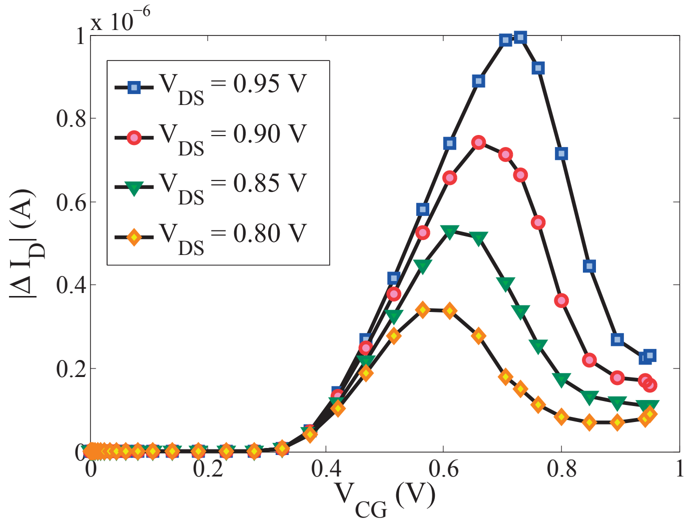


Figure 5.3: Nonlinear piecewise model of GOS under $CG$ for n-type TIG-SiNWFETs. The switches are necessary to realize the impact of the polarity gates on the leakage of defective device in DP logics.

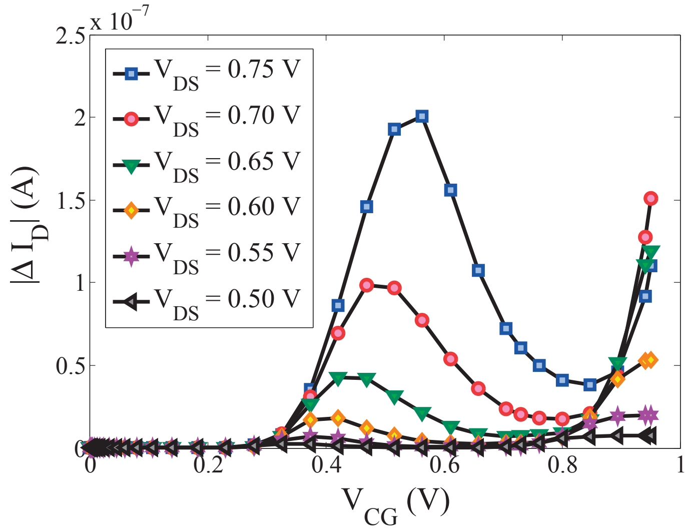
### GOS Modeling in TIG-SiNWFET

The proposed GOS model for TIG-SiNWFETs consists of two nonlinear piecewise *Voltage Controlled Current Sources* VCCSs ($I_{GD}$, and $I_{GS}$) that connect the faulty gate to the drain and source respectively (Figure 5.3). These VCCSs might be followed by extra switches that show the effect of other gates on the current flow. For example, the GOS in $CG$ is modeled by two extra switches that are controlled by the $PG_S$ and $PG_D$ (red blocks of Figure 5.3). These switches highlight that the leakage current between $CG$ and source/drain depends on the polarity gate voltages. Here, we explain how the model is extracted for a GOS in $CG$. The same procedure is then used for the modeling of the GOS in $PG_S$ and $PG_D$ respectively.

In the proposed model, the variation of $I_{DS}$ for a TIG-SiNWFET with and without GOS under CG is used to model the current loss (Figure 5.4). Here, $|\Delta(I_D)|$ represents the difference of $I_{DS}$ for a faulty (GOS under CG) and non-faulty device, in which $I_{DS}$ depends on both $V_{CG}$ and $V_{DS}$ values ($|\Delta(I_D)| = |I_{D(faulty)} - I_{D(non-faulty)}|$). In the proposed model, $|\Delta(I_D)|$ determines the amount of $I_{CG,D}$. For $V_{DS}$ >= 0.80 (Figure 5.4 (a)), $|\Delta(I_D)|$ is $V_{CG}$ controlled and we use a second-order polynomial of $V_{CG}$ to represent $I_{CG,D}$ as shown in Eq. (5.1). For $V_{DS} < 0.80$ (Figure 5.4 (b)), $|\Delta(I_D)|$ shows a rapid increase when $V_{CG}$ is approaching $V_{dd}$. This behavior is related to the negative $I_{D(sat)}$ when $V_{DS}$ decreases. Thus, $|\Delta(I_D)|$ is modeled by a third-order polynomial of $V_{CG}$ to represent $I_{CG,D}$ as shown in Eq. (5.1). In both cases, the impact of GOS on the current loss is above $V_{Th}$. For $V_{CG} < V_{Th}$, $I_{CG,D}$ current is limited to the device leakage as well as the drain-to-gate leakage when $V_{dd}$ is high. Therefore, the $I_{CG,D}$ is obtained by the following equation:



(a) Samples of $|\Delta I_D|$-$V_{CG}$ for $V_{DS} \geq 0.8$. The curves can be modeled by a degree-2 polynomial above the threshold.

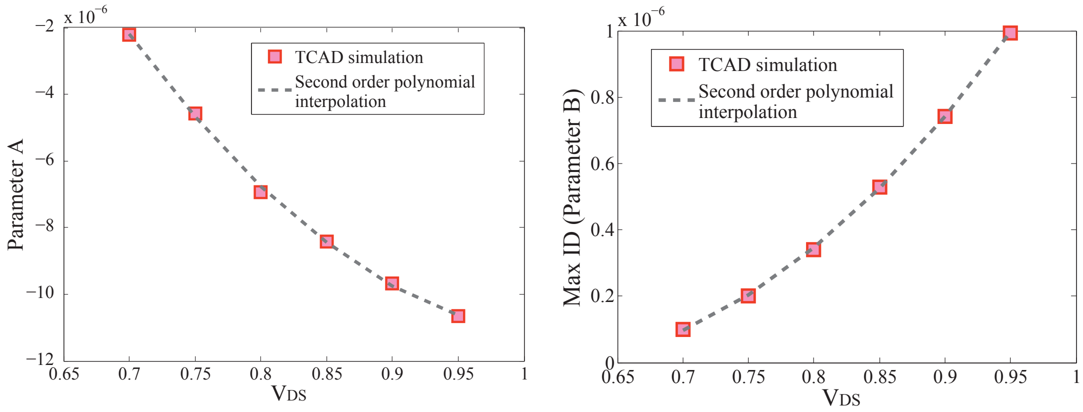


(b) Samples of $|\Delta I_D|$-$V_{CG}$ for $V_{DS} \leq 0.8$. The curves can be modeled by a degree-3 polynomial above the threshold.

Figure 5.4: Difference of $I_D$ with and without GOS under CG versus $V_{CG}$.

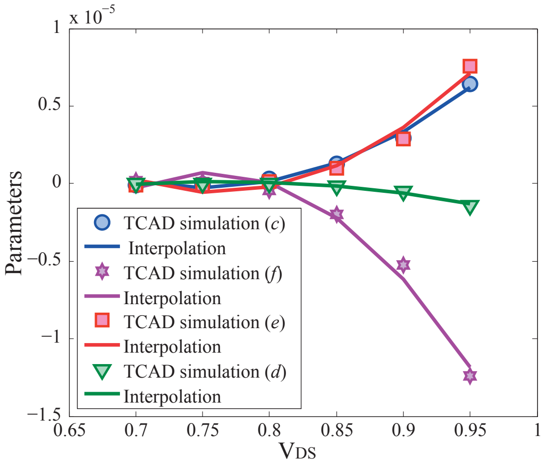
$$I_{CG,D} = \begin{cases} a(V_{CG} - b)^2 & : V_{DS} \geq 0.8 \\ \\ c(V_{CG})^3 + d(V_{CG})^2 + e(V_{CG}) + f & : V_{DS} < 0.8 \end{cases} \tag{5.1}$$

where $a$, $b$, $c$, $d$, $e$, and $f$ are the fitting parameters, which are all shown in Figure 5.5. All the parameters in both equations are a function of $V_{DS}$. Figures 5.5a–5.5c represent the relation of these parameters with $V_{DS}$. All the parameters can be estimated by a second order polynomial of $V_{DS}$ (The curves show the interpolation of the parameters with corresponding second-order polynomial). The same procedure is utilized to find the $I_{CG,S}$ model. We built this model in Verilog-A and we used it for circuit simulations.



(a) Interpolation of parameter $a$ as a function of $V_{DS}$.

(b) Interpolation of parameter $b$ as a function of $V_{DS}$.



(c) Second order interpolation for parameters of the 3rd order polynomial as a function of $V_{DS}$.

Figure 5.5: Parameters of the GOS model estimator as a function of $V_{DS}$. All the parameters can be interpolated with second order polynomials.

### GOS in SP and DP Logic Circuits

Here, we present a case study for circuit level GOS injection. We performed the circuit level GOS simulations for SP (INV) and DP (2-input XOR) logic gates. Here, the defect is injected under the CG gate. Figure 5.6 represents the transient simulation of an inverter with a defect on the pull-down transistor. Note that we only depict the inverter response when its input changes from 0 to 1, since the faulty $n$-type transistor affects the inverter response more significantly during this transition. The result demonstrates a significant reduction of noise margin ($NM_L$). A similar experiment is done on the pull-down transistor (t4) of a 2-input XOR. The result is shown in Figure 5.7. Here, the GOS-impacted defect result in degraded logic levels that can cause logic gate functionality failure during output low-to-high transition ($AB = (00 \rightarrow 01)$) and noise margin reduction during output high-to-low transition ($AB = (10 \rightarrow 11)$).

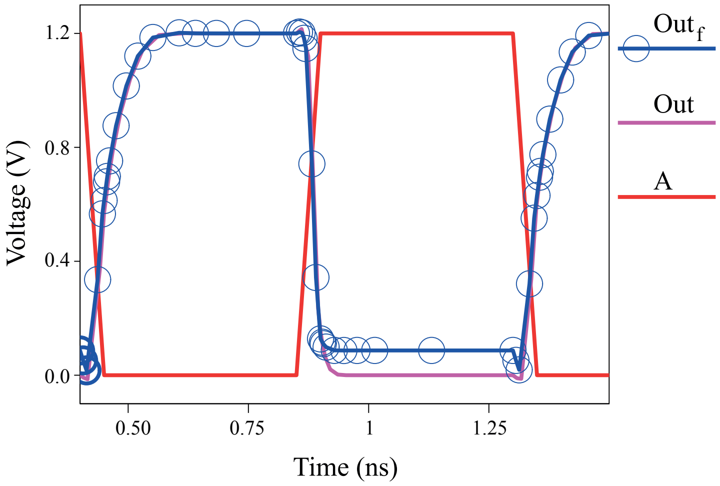


Figure 5.6: GOS occurrence in CG of the pull-down transistor of an Inverter.

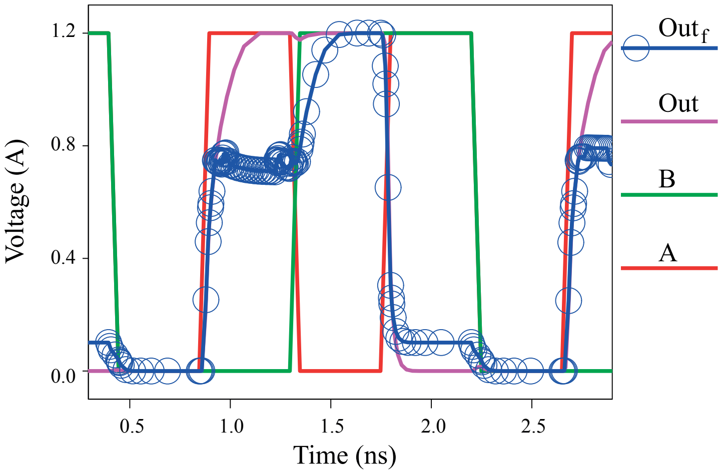
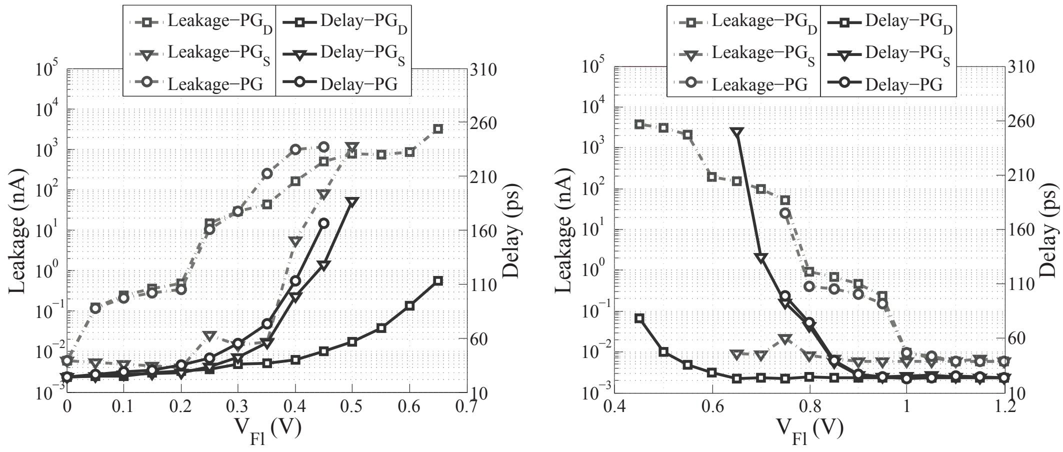


Figure 5.7: GOS occurrence in CG of a pull-down transistor (t3) of an XOR.

### 5.4.2 Open CGs and PGs on the SP and DP logic gates

In this section, we investigate the impact of open defects on the functionality and the performance of CP logic gates. When an open defect happens on a node, that should be treated as a floating node. According to the capacitances that couple to the floating node, and transitions that occur across the coupling capacitances, the floating node may acquire the intended original value, or vary dynamically. Due to the coupling effects, it is necessary to analyze the logic gates for a range of possible voltages which may be exhibited on the floating node. The voltage value for a floating node, $V_{Fl}$, is varied from $V_{Lo}$ to $V_{Hi}$. ($V_{Lo}$, $V_{Hi}$) is a subset of ($GND$, $V_{dd}$) in which the functionality of the logic gate under test is correct. We simulated SP (INV and NAND) and DP (2-input XOR, 3-input XOR, and MAJ) logic gates with open faults on the polarity gates of the pull-up and pull-down transistors. The defect-free PG biases are set to their nominal values (for SP logic gates $PG_S = PG_D = '0'$ and $PG_S = PG_D = '1'$ for pull-up and pull-down devices respectively, and for the DP gates $PG_S = PG_D =$ *appropriate input signals*).

Figures 5.8a and 5.8b illustrate the leakage-delay variation with respect to $V_{Fl}$ for the polarity gates of the pull-up and pull-down transistors in a inverter (INV) gate. Here, floating on $PG_D$, floating on $PG_S$, and floating on both polarity gates are denoted by $PG_D$, $PG_D$, and $PG_D$ respectively. In Figure 5.8a, the delays of $PG_D$ and $PG_S$ stay relatively constant up to $V_{Fl} = 0.3V$. When $V_{Fl}$ further increases to $0.50V$, the delay of output low-to-high transition for floating $PG_S$ rises exponentially (7×). However, the delay of $PG_D$ increases slightly, since $PG_D$ plays a less important role in control of carrier concentrations. The leakage shows a drastical increase for both cases (5×). This is due to the fact that, here, the leakage is dominated by the *p*-type transistor. Finally, beyond $V_{Fl} = 0.50$, the *p*-type is always *Off*. Figure 5.8b also shows a similar trend to pull-up transistor as discussed. Consequently, open defects on the TIG-SiNWFETs in INV logic gates get along with several fault models, corresponding to the voltage of $V_{Fl}$. For $V_{Fl}$ below $0.50V$, the pull-down transistor (t3) is polarized to *n*-type. Here, the floating defect impacts the timing of the INV gate but does not change the functionality. Beyond this threshold, the INV gate exhibits an incorrect functionality that can be captured by the common *Stuck-Open Fault* (SOF) model.

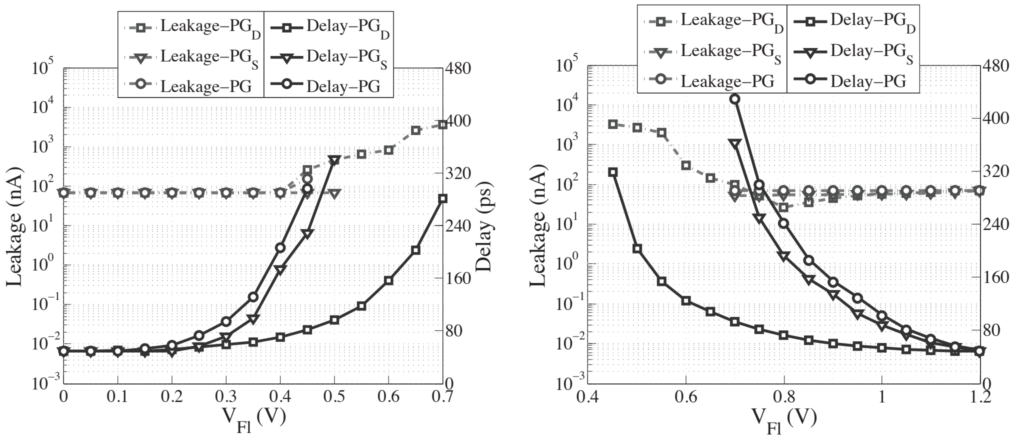


(a) INV (transistor t1)

(b) INV (transistor t2)

Figure 5.8: Leakage-delay variation with different biases on $PG_S$ and $PG_D$ of an Inverter gate in TIG-SiNWFET technology.

In Figures 5.9a, and 5.9b, the variation in leakage and delay of the pull-up (t1) and pull-down (t3) transistors of the TIG-SiNWFET NAND are shown. A drastic increase in delay occurs as $V_{Fl}$ changes from its intended bias similar to what observed for the INV gate. For t3, the leakage represents a relatively small variation . This is due to the fact that leakage of $n$-type device (t3) is dominated by the other transistor (t4) of the pull-down in NAND gate. Therefore, the open defect in TIG-SiNWFET NAND can be detected using the combination of delay fault and SOF.

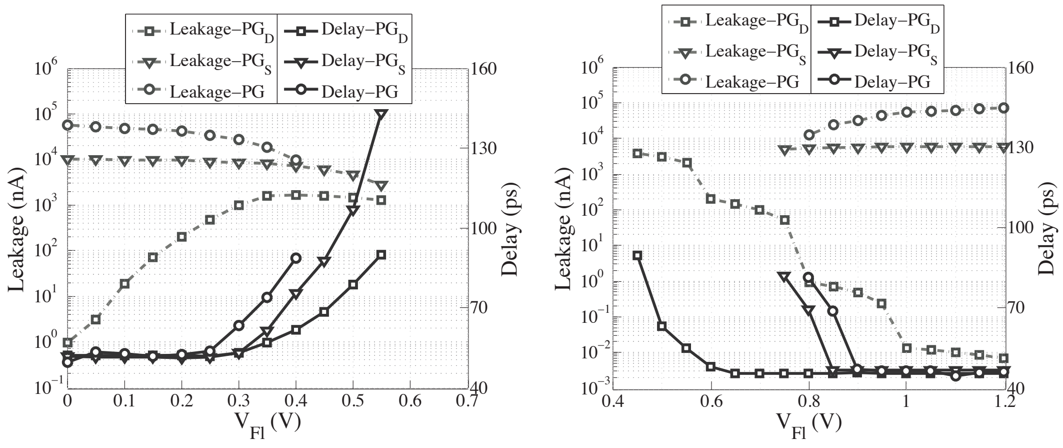


(a) Nand (transistor t1)

(b) Nand (transistor t3)

Figure 5.9: Leakage-delay variation with different biases on $PG_S$ and $PG_D$ of a Nand gate in TIG-SiNWFET technology.

The leakage-delay characteristics of the 2-input XOR are illustrated in Figures 5.10a, and 5.10b. Against the SP gates, here only the leakage represents a considerable variation (5 orders of magnitude) for the various $V_{Fl}$ (Figure 5.10a). Thus, the defect can be tested only with stuck-on fault model. Figure 5.10b also represents the behavior of XOR when the open fault happens in the pull-down transistor (t3). Here, the leakage variation (6 orders of magnitude) contributes to detect the faulty device, while the delay represents a slight variation. Therefore, the test of open defect for TIG-SiNWFET XOR requires a combination of SOF, and stuck-on fault models. For the 3-input XOR and MAJ logic gates, there exists at least an input pattern that can reveal the defective device. Thus, the SOF is enough for these logic gates.



(a) XOR (transistor t1)

(b) XOR (transistor t3)

Figure 5.10: Leakage-delay variation with different biases on $PG_S$ and $PG_D$ of a XOR gate in TIG-SiNWFET technology.

### 5.4.3 Bridging Defects Characterization in the SP and DP Logic Gates

Among various types of bridging defects, short between polarity terminals and supply voltage is exclusive to CP logic gates. In SP logic gates, the bridging connection between polarity controls and $V_{DD}$ in pull-up network changes the desired polarity of the device from $p$ to $n$. Similarly, bridging defect between polarity controls and $GND$ in pull-down network leads to the polarity change from $n$-type to $p$-type. This defect in SP logic gates represents similar behavior to channel break which can be easily covered by SOF. For the DP logic circuit, this defect can be masked depending on the location of the faulty transistor in the circuit. As polarity gates come from input signals, and polarity terminals are accessible from circuit inputs, it is possible to define a logic level fault model for this defect to facilitate the test process. We define the stuck-at $n$ fault to represent the bridging defect in the pull-up network. The stuck-at $n$ defect can be applied on the circuits using $V_{stuck-at\ n} = [PG_D : '1' \;\; PG_S : '1']$. Similarly, the stuck-at $p$-type defect is defined by $V_{stuck-at\ p} = [PG_D : '0' \;\; PG_S : '0']$.

In order to evaluate the performance of this model, we analyzed the TIG-SiNWFETs XOR by exhaustive fault injection. If the faulty device is located in pull-down network, the wrong output of the logic gate reveals the fault. For the pull-up network, the fault detection is only possible by leakage observation. Here, the leakage variation is more than $\times 10^6$. This variation is high enough to be sensible by the $I_{DDQ}$ test.

### 5.4.4 Channel Break in the SP and DP Logic Gates

Channel break demonstrates a similar behavior like stuck-open faults in SP logic gates. The detection of this defect requires to employ a two-pattern test. The first vector initializes the gate output and the second one evaluates the wrong output value in the presence of a fault. For example, a NAND gate as a SP logic contains three vectors of two-pattern tests ($v1 = (11 \rightarrow 01)$, $v2 = (11 \rightarrow 10)$, $v3 = (00 \rightarrow 11)$), by which all the channel break defects for the TIG-SiNWFET NAND can be detected. Although it is possible to detect all faults related to the SP logic gates such as FinFET NAND gates, detection of open faults in DP logic gates is non-trivial. When a channel break happens on a transistor of a DP TIG-SiNWFET gate, the redundant structure of the transistors masks the impact of faulty transistor. Here, the fault masking depends on the capacitances that couple to the output node and the polarity of the fault free transistor. In order to simulate this situation, we performed fault injection using vectors $V_{P_{Off}}$ = [CG: $'1'$ PG: $'0'$] ($p$-type off transistor) and $V_{N_{Off}}$ = [CG: $'0'$ PG: $'1'$] ($n$-type off transistor). When the negated value of the CG signal is applied to PGs, the transistor goes to the turn off mode. The vectors have been applied on the 2-input TIG-SiNWFET XOR (FO4) to evaluate the channel break on the DP logic gates. Here, all the injected faults are masked by the transistors in the pull-up and pull-down networks. Indeed, the channel break defect does not change the logic gate functionality. The defect only affects the performance parameters of the gate such as delay and leakage. Our simulation results on 2-input XOR revealed that the variation of performance parameters are too low for the purpose of fault detection ($\Delta leakage \leq 100\%$, and $\Delta delay \leq 58\%$). The challenging part to reveal this defect is limited access to the polarity terminals in DP logic gates, since they are utilized as logic gate inputs. In the following, we propose a procedure which can be efficiently used for channel break detection of DP logic gates.

Unlike the SP logic gates, the detection of channel break defects in DP logic gates requires a pair of two-pattern test vectors. Suppose that we have a channel break defect in a pull-up transistor (t1) of an 2-input XOR gate. In this case, the transmission gate structure that connects the output to $V_{dd}$ is degraded to a pass transistor (t2). The first vector, ($v1 = (00 \rightarrow 01)$), initializes the gate output to $'0'$ and then connects the gate output to $V_{dd}$ by polarizing the fault-free device (t3) to a *ON* $n$-type transistor (Figure 5.11). The second vector, ($v2 = (00 \rightarrow 10)$), initializes the gate output to $'0'$ again and then connect the gate output to $V_{dd}$ by oppositely polarizing the fault-free device (t3) to a *ON* $p$-type transistor (Figure 5.11). The $p$-type transistor, in pull-up, passes a weak $'1'$ that leads to a large delay penalty in the XOR output to switch from $'0'$ to $V_{IH}$ (The $V_{IH}$ value for this technology is $0.60V$). Here, we observe a delay up to 10× for the XOR gate. In case of no channel break, an *ON* transmission gate in pull-up network passes a strong $'1'$ to the gate output for both test vectors without introducing any extra delay. The same scenario can be utilized for channel break detection in pull-down network as well. Therefore, the delay can be used a criteria for the detection of channel break.

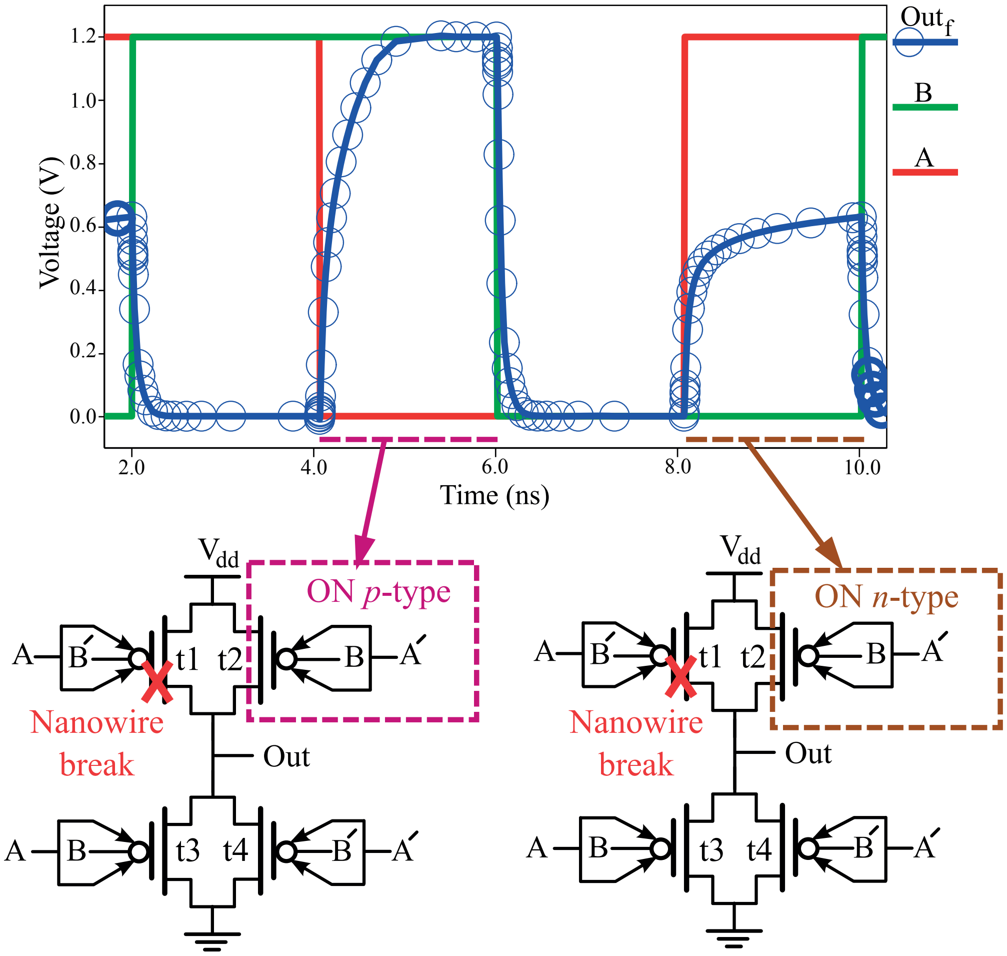


Figure 5.11: Nanowire break defect detection using a pair of two-pattern test.

## 5.5 Summary

Further scaling of the planar CMOS technology has been confronted with serious challenges such as increased leakage and process variation. Among the alternative technologies, controllable-polarity silicon nanowires such as TIG-SiNWFETs are promising owing to their lower leakage and great electrostatic control. Moreover, their reconfigurable structures provide the opportunity of implementing logic gates with enhanced functionality, i.e., implementing logic gates with fewer number of devices than that of the current technology. As one of the necessary design steps, fault modeling is needed for this new technology. In this chapter, we performed an inductive fault analysis on the TIG-SiNWFETs. According to the fabrication process steps, a defect model was extracted. This model contains channel break, gate oxide short, bridging and floating defects. We simulated the effect of these defects on various categories of CP circuits. Then, we extended the current CMOS fault model to a new hybrid model, including stuck-at *n*-type and stuck-at *p*-type, which can be efficiently used for the detection of defects in CP logic gates. The experimental results revealed that the gate oxide short and floatings on the polarity gates are detectable by analyzing the performance parameters like delay and leakage. We also illustrated that the current CMOS test methods are not able to capture all faults in CP logic gates, i.e., stuck- Open. Finally, we proposed an appropriate test method to cover such faults.

# 6 Design of a Fault Tolerant Ripple-Carry Adder with Controllable-Polarity Transistors

## 6.1 Introduction

All digital integrated circuits are susceptible to functionality failures. These failures are not tolerable since they adversely affect the fabrication yield. Moreover, safety-critical applications require a certain level of reliability. Therefore methods for increasing reliability and fault tolerant ability of such systems are inevitable. Fault tolerant can be achieved by using redundancy. This can be in the form of hardware, time, information, software, and even system redundancy, or any of their combinations.

In this chapter, we first perform an analysis on the effects of possible permanent faults affecting a generic controllable-polarity device based on the inductive fault analysis that was presented in Chapter 5. In order to take into account the specific characteristics of the controllable-polarity devices, this analysis has to be performed at the transistor level. We propose a new fault model that takes into account the specific characteristics of these devices, extending the popular stuck-open/stuck-short fault model traditionally used at that level. Then, we analyze the behavior of circuits based on controllable-polarity devices when permanent device faults are present, and identify the conditions for their detection/masking. Results show that a high number of faults are masked, thus making this new technology particularly interesting from a reliability point of view. Performing this analysis at the transistor level allowed us to express the behavior of each gate when any of the possible faults affecting each of its transistors arise. We use this information to forecast the behavior of more complex circuits consisting of the above gates, thereby achieving the same precision than a transistor level analysis but with much lower computational complexity.

Based on the results of the previous analysis, we also propose a fault-tolerant ripple-carry adder architecture exploiting XOR/MAJ logic gates built entirely with controllable-polarity transistors. In order to guarantee a high degree of resiliency with respect to single and double permanent faults in every single stage, we combine the intrinsic resiliency of the controllable-polarity-based circuits with the usage of the *Triple Modular Redundancy* (TMR) architecture. Although faster solutions are often adopted to implement adders, e.g., based on the Kogge-

Stone architecture [102] and its fault tolerant version [103], the TMR version of the ripple-carry adder still represents the reference to compare with, especially when the parallelism is limited and power is not a major issue (in the latter case, solutions based on reversible logic are often adopted [104]).

Experimental validation shows that the full-adder architecture we propose is able to tolerate all possible single faults and a very high percentage of the double ones. In addition, it proves that the proposed solution provides a 15%, 18% and 12% gain in area, performance and leakage power with respect to similar architectures implemented in FinFET technology at 22-nm technology node. Finally, the proposed architecture is significantly cheaper with respect to solutions based on hardening the circuit at the transistor level, such as those proposed in [105], whose area overhead 4× the unhardened circuit.

## 6.2 Background and Motivation

Adders are the necessary part of data processing systems in safety-critical applications [106]. They are commonly found in the critical path of arithmetic and logical units (ALUs) and address generation units. Therefore, the design of adder structures with error detection and correction capabilities is an important research topic.

Existing methods usually use the parity prediction technique in combination with a two-rail code [107, 108]. In [107], the two-rail code is used for carries and the parity prediction is used for the outputs. Both the parity prediction and checkers of two-rail codes decrease the performance drastically, because both of them need XOR-tree structures. In [108], the parity code is exploited to detect input operands error and the two-rail code is used for the output error detection. In addition to these techniques, duplication with comparison [108] can also be used. This technique has about 100% area and power overheads. All of the mentioned techniques only detect errors and simply use re-execution, the most popular and simplest correction technique. The re-execution cannot correct permanent faults, and it increases the delay about two times. Furthermore, many single-bit errors cannot be detected in CLA adders, which are checked by arithmetic codes [109]. An alternative to re-execution is TMR technique which is an almost traditional technique to cope with error occurrence in all circuits as well as the adders. This technique has high power consumption and area overheads; thus, a fault tolerant adder is achieved at the cost of highly increasing the power consumption and the area. Moreover, TMR is suffers from the problem of single point of failure for voting operation. So, single-bit faults can also escape to other part of the system.

Using controllable-polarity devices, it is possible to build very compact arithmetic logic gates, such as *eXclusive OR* (XOR) and *MAJority* (MAJ). This compactness can be leveraged in adder implementations, as reported in Figure 6.1, where we show a full adder composed of only 8 controllable polarity transistors. This circuit exploits 3-input XOR and MAJ gates to implement the sum and the carry, respectively. Note that the proposed cells exploit a transmission-gate design. We will see, in the following, that this introduces a degree of redundancy at the gate level, that is beneficial from a robustness perspective. A self-checking ripple-carry adder architecture, exploiting this adder structure, is proposed in [110]. This architecture is far less expensive in terms of area than comparable CMOS architectures. Here, we make one step forward with respect to [110]. In addition to exploit the reduced area cost offered by controllable-polarity devices, we take into account their intrinsic capabilities in masking, i.e., tolerating, faults and use them to build a fault-tolerant adder.

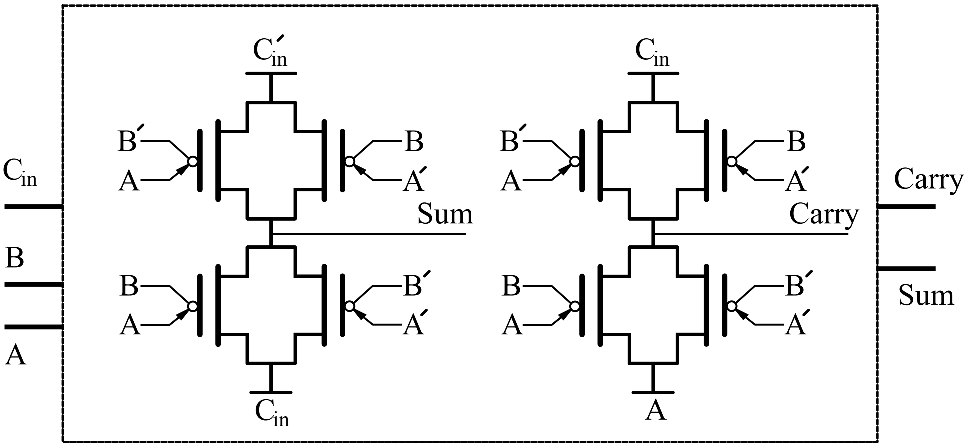


Figure 6.1: Realization of 1-bit Full adder in CP-SiNWFET technology.

## 6.3 Evaluating Controllable-Polarity Circuit Robustness

In this section, we introduce a new fault-model suited to controllable-polarity transistors, study the robustness of operations for XOR and MAJ logic gates exploiting controllable-polarity transistors, and extend the results of this analysis to circuits composed of different gates.

### 6.3.1 Fault Model for Functionality Evaluation

The robustness evaluation of circuits based on controllable-polarity devices cannot be performed by relying on usual fault models and tools, e.g., working at the gate level [111]. Indeed, when new technologies are introduced, it is common to envisage a lower-level approach, e.g., resorting to transistor-level fault models [112]. In such a case, the most common solution lies in inductive fault analysis of the device as well as layout-based defect map extraction for feasible fabrication shortcomings.

In this chapter, we only consider the defects that completely change the functionality, e.g., change the polarity of a transistor from p-type to n-type. Defects affecting the performance but keeping the functionality untouched are not considered. These defects can be modeled by generalizing bridge defects to the two gates composing our transistors. Therefore, we introduce a new fault model that generalizes the stuck-at model for the mentioned bridge defects:

- stuck-at-0 on CG (CG/0), stuck-at-1 on CG (CG/1): This defect is similar to what happens in the current technology. Depending on the polarity of the transistor, such defect will lead to a Stuck-Open or Stuck-Short behavior of the device.
- stuck-at-0 on PG (PG/0), stuck-at-1 on PG (PG/1): This defect affects the polarity of the device. The device will be either stuck-at-n or stuck-at-p, affecting the logic operation.

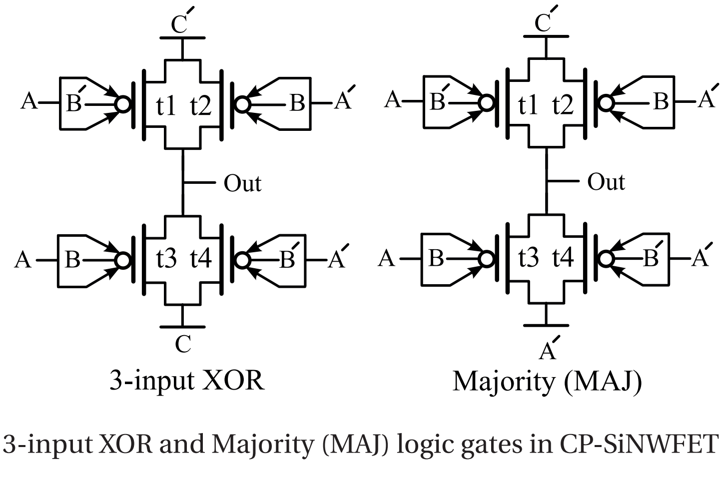
The new fault model straightforwardly extends the traditional transistor-level fault model, where the gate can be either stuck-at-0 or stuck-at-1, and takes into account the specific characteristics of controllable polarity transistors. Each of these faults corresponds to forcing to 0 or 1 the value of the corresponding controllable-polarity transistor input signal. We denote this fault model as CG/PG fault model.

### 6.3.2 XOR/MAJ Gates Robustness

In order to evaluate the robustness of a circuit implemented with controllable-polarity devices, we need to evaluate the behavior of the basic logic primitives, when a CG/PG fault occurs in any of the transistors of the gate.

#### Simulation of Logic Gates

The 3-input XOR and MAJ logic gates, shown in Figure 6.2 (repeated for reader convenience), respectively, have been characterized using electrical simulations.



3-input XOR

Majority (MAJ)

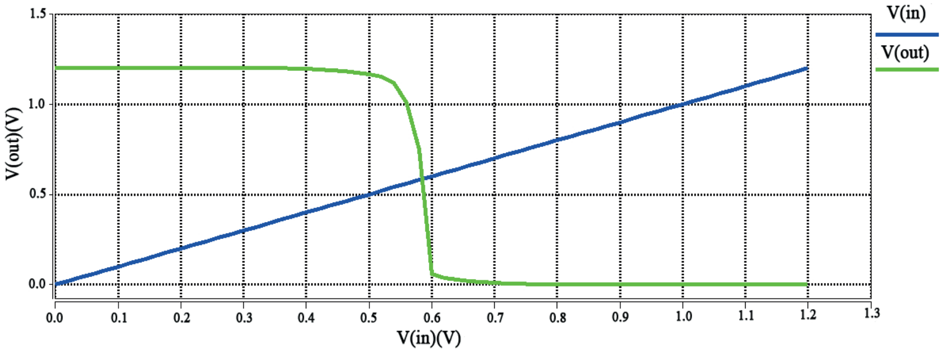
Figure 6.2: 3-input XOR and Majority (MAJ) logic gates in CP-SiNWFET technology.

First, we identify the input voltage ranges associated to Boolean input 0 ($V_{IL}$) and 1 ($V_{IH}$). This is done by performing Spice simulation on the logic cells of our library, and then selecting the worst case. Figure 6.3 depicts two example for a 2-input XOR gate. Defining the input voltage boundaries will help us to identify a faulty gate behavior in presence of a transistor-level fault. We report the obtained points for the logic gates:

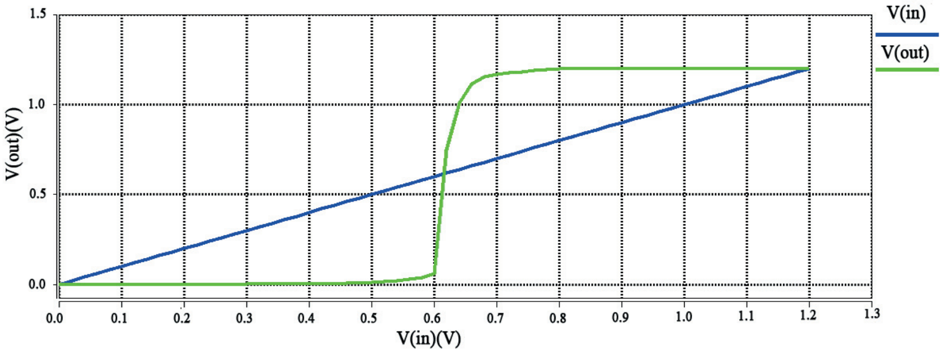
- $V_{IH} = 0.600V$
- $V_{IL} = 0.540V$

Therefore, the two logic gates will correctly behave when the output voltages for Boolean output $'0'$ ($V_{OL}$) and $'1'$ ($V_{OH}$) are in the following ranges:

- $0.600V < V_{OH} < 1.2V$
- $0V < V_{OL} < 0.540V$



(a) An example of transfer characteristic of a 2-input XOR in CP-SiNWFET technology (input A=sweep from 0 to 1 and input B=1)



(b) An example of transfer characteristic of a 2-input XOR in CP-SiNWFET technology ((input A=sweep from 0 to 1 and input B=0)

Figure 6.3: Transfer characteristic of a CP-SiNWFET 2-input XOR gate.

The identified ranges are used to classify the output values of the different gates. Then, the behavior of the logic gates under all possible CG/PG faults is computed by using DC operating points analyses for all possible input conditions.

### XOR/MAJ Gates Behavior under CG/PG Faults

Tables 6.1 and 6.2 report the simulated DC operating points of the 3-input XOR and MAJ gates, respectively, when the gates are fault-free and when each of the CG/PG faults are injected in the different transistors (t1 to t4).The CG/PG fault injection induces different behaviors classified under three categories:

- Correct behavior (highlighted in green) when a CG/PG fault is not excited by the applied input vector;

- Masked-fault behavior (highlighted in blue) when a CG/PG fault is excited and induces a reduction of the noise margin at the output of the gate, but does not induce a faulty gate behavior as the output voltage is still in the correct VOH and VOL range;

- Faulty behavior (highlighted in red) when a CG/PG fault induces an incorrect value at the output of the gate.

Considering the 3-input XOR (Table 6.1), the results indicate that 8 CG/PG faults out of 16 lead to a faulty gate behavior that is observable at the gate output for at least one input combination. The remaining 8 CG/PG faults are always masked. Moreover, the 8 detectable faults produce a faulty output when 4 out of 8 possible input values are applied ("001", "010", "100", and "111"). With the other 4 input combinations ("000", "011", "101", and "110"), the circuit always produces the correct output no matter the presence of a fault. Controllable-polarity transistors have 4 different modes of operations: *On n*-type, *Off n*-type, *On p*-type and *Off p*-type. A CG/PG fault restricts the number of operations of the device but does not fully lock it in a unique mode. This property is unique to the class of controllable-polarity transistors and unachievable with standard transistors. This has a positive impact on the fault tolerance of the overall gate circuit. As an example, we can consider the PG/0 fault on t4 in the 3-input XOR under input values "000". Under fault-free conditions, the bottom transmission-gate is on, with t3 configured as p-type and t4 as n-type. In this case, t4 propagates properly the logic 0. However, when a PG/0 fault affects t4, t4 polarity switches to *p*-type. In this condition, the logic 0 cannot be fully propagated, but is still transmitted with limited voltage degradation. Such degradation reduces the noise margin of the gate, but does not induce a faulty behavior.

Table 6.1: Output voltage values of the 3-input XOR gate in DG-SiNWFET technology.

| | | t1 | | | | t2 | | | | t3 | | | | t4 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Input | Fault-Free | CG/0 | CG/1 | PG/0 | PG/1 | CG/0 | CG/1 | PG/0 | PG/1 | CG/0 | CG/1 | PG/0 | PG/1 | CG/0 | CG/1 | PG/0 | PG/1 |
| 000 | 0.0 | 0.32 | 0.0 | 0.0 | 0.27 | 0.0 | 0.27 | 0.32 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.14 | 0.0 | 0.13 | 0.0 |
| 001 | 1.2 | 0.83 | 1.2 | 1.2 | 0.36 | 1.2 | 0.36 | 0.83 | 1.2 | 1.2 | 1.1 | 1.2 | 1.1 | 1.2 | 1.2 | 1.2 | 1.2 |
| 010 | 1.2 | 1.2 | 1.1 | 1.2 | 1.1 | 1.2 | 1.2 | 1.2 | 1.2 | 0.83 | 1.2 | 1.2 | 0.36 | 1.2 | 0.36 | 0.83 | 1.2 |
| 011 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.14 | 0.0 | 0.13 | 0.0 | 0.32 | 0.0 | 0.0 | 0.27 | 0.0 | 0.27 | 0.32 | 0.0 |
| 100 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.1 | 1.2 | 1.1 | 1.2 | 0.36 | 0.83 | 1.2 | 0.83 | 1.2 | 1.2 | 0.36 |
| 101 | 0.0 | 0.14 | 0.0 | 0.13 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.27 | 0.32 | 0.0 | 0.32 | 0.0 | 0.0 | 0.27 |
| 110 | 0.0 | 0.0 | 0.27 | 0.32 | 0.0 | 0.32 | 0.0 | 0.0 | 0.27 | 0.14 | 0.0 | 0.13 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 111 | 1.2 | 1.2 | 0.36 | 0.83 | 1.2 | 0.83 | 1.2 | 1.2 | 0.36 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.1 | 1.2 | 1.1 |

Table 6.2: Output voltage values of the 3-input MAJ gate in DG-SiNWFET technology.

| | | t1 | | | | t2 | | | | t3 | | | | t4 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Input | Fault-Free | CG/0 | CG/1 | PG/0 | PG/1 | CG/0 | CG/1 | PG/0 | PG/1 | CG/0 | CG/1 | PG/0 | PG/1 | CG/0 | CG/1 | PG/0 | PG/1 |
| 000 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 001 | 0.0 | 0.32 | 0.0 | 0.0 | 0.27 | 0.0 | 0.27 | 0.32 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.14 | 0.0 | 0.13 | 0.0 |
| 010 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 011 | 1.2 | 1.2 | 1.1 | 1.2 | 1.1 | 1.2 | 1.2 | 1.2 | 1.2 | 0.83 | 1.2 | 1.2 | 0.36 | 1.2 | 0.36 | 0.83 | 1.2 |
| 100 | 0.0 | 0.14 | 0.0 | 0.13 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.27 | 0.32 | 0.0 | 0.32 | 0.0 | 0.0 | 0.27 |
| 101 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 |
| 110 | 1.2 | 1.2 | 0.36 | 0.83 | 1.2 | 0.83 | 1.2 | 1.2 | 0.36 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.1 | 1.2 | 1.1 |
| 111 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 | 1.2 |

Similarly, for the 3-input MAJ (Table 6.2), the results indicate that the number of faulty behaviors is very small:12 faults out of 16 are always masked. Moreover, the 4 faulty conditions do produce a difference in the output voltage only when 2 of the 8 possible input combinations are applied ("011" and "110"). For the 6 remaining input combinations, the CG/PG faults never produce any output misbehavior.

#### Circuit-level Analysis

Based on the results of the detailed transistor-level analysis presented so far, we can describe the behavior of each possible logic gate for each possible input combination and for each possible fault affecting each of the internal transistors. Therefore, we build a detailed model of the fault-free and faulty behavior of each gate. Using these models, we then determine the fault-free and faulty behavior of a larger circuit composed of different gates working at the logic level. This allows us to ignore the details of the underlying transistor-level structure, without loosing accuracy. In the sequel, we develop VHDL models for each gate (with suitable control signals to inject each possible fault), and combine them to extensively analyze the behavior of larger circuits.

## 6.4 Fault-tolerant Ripple-carry Adder Architecture

Knowing the behavior of the 3-input XOR and MAJ gates exploiting controllable-polarity transistors, and using the approach we just described, we now investigate the possibility to implement a fault-tolerant ripple-carry adder architecture based on these primitives.

We consider the 1-bit adder circuit represented in Fig 6.4 and consisting of two 3-input XOR gates for the sum generation and two 3-input MAJ gates for the carry generation. As compared to the simpler adder with one 3-input XOR and a Majority, this circuit generates both the sum and carry signals and their inverted versions in a unique logic level. This allows us to create ripple-carry adder structures without adding any inverters to drive the next stage. Note that, due to the transmission gates, buffers will be required every 4 stages.

Using the approach described in the previous section, we first create a logic model of the proposed adder in VHDL combining the models of each composing gate, and use it to gather simulation results for every one of the proposed faults during an exhaustive simulation. Results show that most of the possible CG/PG faults (40 out of 64) are masked in this structure. In order to make the adder fault-tolerant with respect to all the possible faults and under all the input conditions, we propose a TMR-based 1-bit adder architecture, shown in Figure 6.5. Note that, in this figure, the inverted input and internal signals are not represented.

The key characteristics of the proposed fault tolerant adder are:

- Each 1-bit adder of Figure 6.4 is triplicated and voted. In this way, any single fault

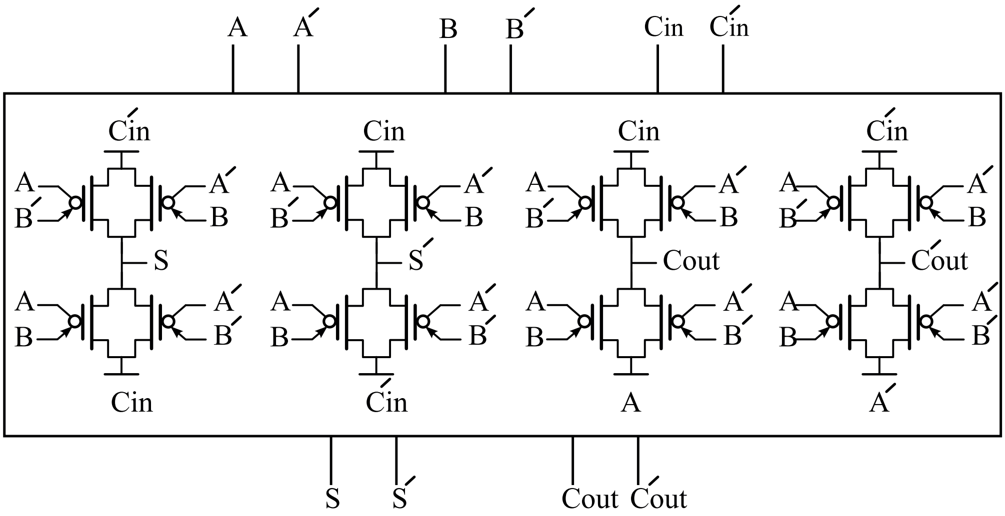


Figure 6.4: 1-bit adder with generation of inverted sum and carry.

affecting a single adder can be tolerated and does not propagate to the following stages of the adder.

- The inputs to each replica, labeled from 1 to 3, are permuted. In this way, even if the same fault affects more than a single replica, this does not evolve into a common mode fault, and the circuit behaves correctly.

- Each of the 3 output signals, i.e., the sum and the carry signals (regular and complemented to cascade further stages) is voted. Since the majority voter has been shown in the previous section to never fail with the "000" and "111" input combinations, the voter never fails when the three 1-bit adders are fault-free. This guarantees that the 1-bit adder never produces a faulty output in the presence of a single fault affecting an adder or a voter.

- Given the above properties, one can easily build a ripple-carry adder out of the proposed 1-bit adder, knowing that fault effects cannot propagate from one stage to another.

The proposed architecture provides significant benefits in terms of fault tolerance. Thanks to triplication, it can mask any single fault in the three 1-bit adders and in the voter, which is never stimulated with an input value able to excite a fault inside it, unless another fault exists in an adder. Clearly, faults on the inputs A and B of the single cell of the full adder cannot be tolerated (unless they are in turn triplicated). Thanks to the input permutation and to the robustness of the 1-bit adder, less than 1% of the possible 4,032 (64 × 63) double faults affecting a couple of replicas produce a faulty behavior.

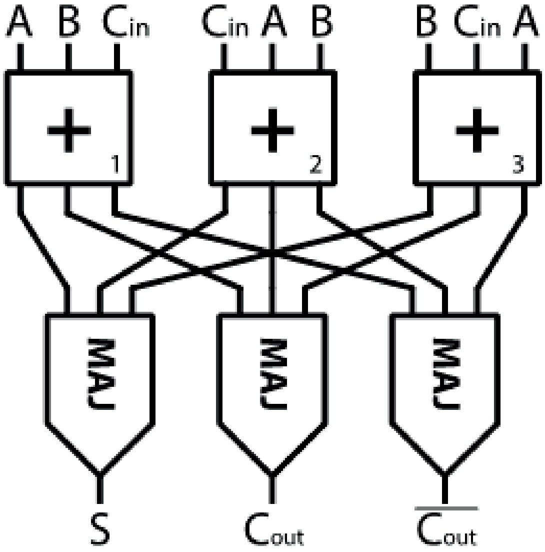


Figure 6.5: (TMR)-based 1-bit adder architecture

## 6.5 Quantitative Analysis

In order to provide the reader with some more details about the performance and characteristics of the proposed architecture, we first performed some experimental analysis, aimed at checking its behavior in the presence of single and double faults.

Results of this analysis (performed by combining at the gate level the results reported in the previous Sections) confirmed that all single faults are masked, either by the characteristics of the controllable-polarity gate implementation, or by the TMR architecture.

From the circuit-level performance perspective, we compared the proposed circuit implemented using SiNWFETs with its equivalent CMOS FinFET 20-nm LSTP counterpart using electrical simulations. The load capacitance for the two circuits is set to 1fF. We consider the area, the worst-case delay and the leakage power. The circuit-level results are summarized in Table 6.3.

Table 6.3: Fault-tolerant 1-bit adder characteristics

| 20 nm node | # of transistors | Area ($\mu m^2$) | Delay ($ps$) | Leakage power ($nW$) |
|---|---|---|---|---|
| FinFET LSTP | 108 | 5.98 | 371 | 23.84 |
| DG-SiNWFET | 60 | 4.98 | 304 | 21.06 |
| Gain | 44% | 15.5% | 18.1% | 11.6% |

The proposed implementation requires 16 controllable-polarity transistors for each 1-bit adder, plus 12 transistors for the 3 majority voters. Hence, 60 transistors are required for the proposed fault-tolerant 1-bit adder. By applying the same design principles with transmission-gate CMOS, we obtained 24 transistors for a 1-bit full adder. Note that the reference structure also generates all the inverted signals required to cascade the different adder stages. Smaller implementations can be identified for both CMOS and controllable-polarity transistors if dedicated inverters are used to generate the inverted signals. Then, a TMR-based implementation in CMOS technology would require 3 × 24 transistors, plus the cost for the majority voting on the data output, accounting for 3 × 12 transistors. In total, 108 transistors would thus be required. Hence, the proposed solution requires 44% fewer transistors. When considering the area of the two adders, the proposed solution requires $4.98 \mu m^2$ as compared to $5.89 \mu m^2$ for its equivalent FinFET implementation. This leads to a gain of 15% in area. The gain is reduced compared to the simple transistor count, as controllable-polarity transistors are bigger than FinFETs, due to the additional polarity terminals. The proposed solution is also significantly less expensive than the one proposed in [113], which proposes a fault-tolerant architecture for the voter consisting of an XOR and a multiplexer.

Finally, the proposed solution can be easily used to build up an adder with any data parallelism $n$, whose total cost scales linearly with $n$. Since we demonstrated that single faults affecting one stage do not propagate to the following ones, the level of fault tolerance of the final adder is not affected by its parallelism.

From a performance perspective, the proposed implementation is shown to be faster by a 18% reduction of the worst case delay. This is accounted to the reduced number of stacked transistors coming from the use of controllable polarity transistors. Finally, the leakage power is reduced by 12%, thanks to the good electrostatic control offered by the NWFETs.

## 6.6 Summary

Controllable-polarity transistors offer many advantages to implement arithmetic logic gates at a reduced implementation cost. Besides the implementation compactness, an important parameter to consider is the robustness with respect to possible faults. In this chapter, we performed such an analysis and showed that circuits based on controllable-polarity transistors can tolerate a large number of faults. Thanks to this property, they can be used to build effective structures demonstrating great fault tolerance, in addition to area, power and speed improvements. In particular, we showed that the SiNWFET implementation of a fault tolerant 1-bit adder (that can be easily used to build an adder of any size) is 15% smaller, 18% faster, and 12% less power consuming than the corresponding CMOS solution. This module can be used to build a ripple-carry adder of any length able to tolerate any single permanent fault and most of the possible double faults in any of its stages.

# 7 Conclusion

## 7.1 Thesis Conclusion

The aggressive shrinkage trend of the feature size in current CMOS technology has confronted digital designers with serious challenges including short channel effect and high amount of leakage power. To address these problems, emerging nano-devices, e.g., Silicon NanoWire FETs (SiNWFETs) have been introduced by the research community. These devices keep on pursuing Moore's Law by improving channel electrostatic controllability, thereby reducing the *Off*-state leakage current. In addition to the improvements in conventional device performances, recent developments introduced devices with enhanced capabilities such as *Controllable-Polarity* (CP) SiNWFETs. In particular, these transistors represent the ability of in-field reconfiguration which makes them very interesting for compact logic cell and arithmetic circuits.

For deeply-scaled nano-devices, variability and fabrication defects are expected to significantly affect the reliability of complex systems. Indeed, the amount of physical controls, during the fabrication process of nanometer transistors, cannot be precisely determined because of technology fluctuations. Furthermore, due to the novel device geometries, a number of new variation sources may arise during the fabrication. Therefore, fabrication parameters can be very different from their nominal values. On the other hand, novel devices rely on the use of unconventional switching mechanisms with regard to conventional doped source/drain transistors, such as Schottky barrier injection at channel interfaces. These new mechanisms allow technologists to build devices that are less relying on statistical process steps and are therefore prone to exhibit the desired robustness. Hence, giving an a-priori conclusion on the variability of advanced technologies is a difficult task and novel estimation methodologies are required.

Statistical analysis of process variation requires a great amount of numerical data for nanoscale devices. Large data sets can capture the characteristics of the population of fabricated devices. Generating large data sets, introduces a serious challenge for variability analysis of emerging technologies due to the lack of fast simulation models. On the one hand, the development of accurate compact models entails numerous tests and costly measurements on fabricated devices. Compact models are widely used for fast and precise simulations of circuits in current CMOS technology. However, this possibility is not available for novel devices. On the other hand, *Technology Computer Aided Design* (TCAD) simulations, that can provide precise information about devices behavior, are too slow to timely generate large enough data set. TCAD simulations are used to predict the electrical behavior of a semi-conductor device given its physical description. The physical description indicates the geometry and materials involved. TCAD simulations provide the first order opportunity to design devices without developing or purchasing expensive new fabrication equipment, at a cost of heavy simulation runtime and tough convergence difficulties for complex models. In this thesis, we proposed a fast methodology for generating data set for variability analysis is introduced. This methodology combines the TCAD simulations with a learning algorithm to alleviate the time complexity of data set generation for emerging nano-devices. Here, a *Feed Forward Neural Network* is exploited to predict the I-V curve of a CP-SiNWFET for a given set of parameters. The simulation results revealed that this method can considerably alleviate the time complexity of generation large data set for variability analysis.

Another formidable challenge for variability analysis of the large circuits is the growing number of process variation sources in deep nanoscale technologies. Utilizing parameterized models is becoming a necessity for chip design and verification. However, the high dimensionality of parameter space imposes a serious problem, i.e. computational complexity for variability analysis. Unfortunately, the available dimensionality reduction techniques cannot be employed for three main three reasons: lack of accuracy, distribution dependency of the data points, and finally incompatibility with device and circuit simulators. Most of the current methods are linear. As a result they cannot capture the strong non-linearity among process and performance parameters. Furthermore, these methods represent their maximum efficiency when the distributions of process parameters are Gaussian, which is not correct for devices such as FinFET or SiNWFETs. We proposed a novel technique of parameter selection for modeling process and performance variation. This method utilizes a universal non-linear regression model which can efficiently model all the linear and non-linear dependencies among process and performance parameters. This is necessary for high precision variation analysis. Our simulation results on various benchmarks in FinFET and DG-SiNWFET proved the efficiency of this method.

Appropriate testing, to capture manufacturing defects, plays an important role on the quality of integrated circuits. Compared with conventional CMOS, emerging nano-devices such as CP-SiNWFETs have different fabrication process steps. Therefore, the type of defects for these technologies can be different from the CMOS devices. In this case, it is not clear that the current fault models can be efficiently used for defect detection of emerging technologies. In this research, we investigated the inefficiency of the current CMOS fault models for detecting the fabrication defects of CP-SiNWFET technology. Considering the fabrication steps, we extracted the possible fabrication defects. We studied the impacts of extracted defects on the performance and functionality of devices, and then proposed a hybrid fault model for this technology. We also provided a couple of test methods for detecting the manufacturing defects in various types of CP-SiNWFET logic gates.

Designing fault-tolerant circuits is very important for safety-critical applications. We studied the behavior of various arithmetic logic cells in presence of faults. Based on the obtained result we proposed a fault-tolerant adder that can mask all the single faults. This design is based on using triple modular redundancy while efficiently address the single point of failure problem of the voters. The proposed architecture has superior characteristics in terms of power consumption, speed, and area rather than the competitive architectures in FinFET technology.

## 7.2 Further Directions

With continuing downscaling trend for integrated circuits, the impact of variations on circuit performances becomes increasingly important. In order to satisfy the required yield and performance, statistical analysis and design will be necessary for digital circuits. Statistical compact modeling can be a feasible solution for deeply scaled circuit design. This can be achieved by the extension of the proposed methodology in Chapter 3. In a statistical compact model, the variability parameters of the target device are not constant anymore. Indeed, the statistical behavior of these variables is considered during the circuit simulation. Consequently, the variability on the circuit performance parameters can be specified because the statistical behavior of each device parameter is accurately mapped on the functionality of the circuit. This procedure can be performed through a learning-based $I-V$ curve predictor. This approach requires a powerful outlier removal to guarantee the accuracy of each $I-V$ curve as well as a mechanism for regenerating missing curves.

Fast and accurate performance analysis of the large designs is needed for precise yield estimation. Adding topology information of the circuit under test to the proposed technique in Chapter 4, can considerably increase its performance and accuracy. This can be performed by extracting the circuit topology graph which includes the map of connections between the logic gates. As the proposed technique keep the meaning of the parameters, branches and reconvergent fanouts do not increase the complexity of variability analysis. Accordingly, the most important parameters can be selected simply, and then circuits' performance can be efficiently computed from primary inputs to the primary outputs.

Reliability evaluation of the CP-SiNWFETs is also an important problem. Modeling the aging phenomena in this technology is a requirement for such an evaluation. TCAD simulation, similar to what we have done in Chapter 5, can be done to extract the deviation of performance and the functionality of a CP device. This model can be used for reliable circuit design in CP-SiNWFET technology.

Soft errors play an important role on the reliability of the safety-critical applications. Dimension shrinkage increases the susceptibility of the circuits to soft errors since the tinier devices use a smaller amount of charge for storing the logic states. Therefore, the hit of a high energy particle can generate spike currents that may propagate in the circuit and even change the logic values of the memory elements. Modeling soft errors in CP-SiNWFETs is a two-fold problem. First, the model of the soft error for a single device has to be extracted through TCAD simulations. Second, the fault injection and soft error rate estimation have to be done for various circuits to reveal the susceptibility of such technology. The outcome of this step is then used for dependable circuits.

# Bibliography

[1] M. Bohr, "The Evolution of Scaling from the Homogeneous Era to the Heterogeneous Era," in *Proceeding of IEEE International Electron Devices Meeting (IEDM)*, 2011, pp. 1–1.

[2] A. Chaudhry and M. J. Kumar, "Controlling Short-Channel Effects in Deep-Submicron SOI MOSFETs for Improved Reliability: A Review," *IEEE Transactions on Device and Materials Reliability*, vol. 4, no. 1, pp. 99–109, 2004.

[3] R. F. Pease and S. Y. Chou, "Lithography and Other Patterning Techniques for Future Electronics," *Proceedings of the IEEE*, vol. 96, no. 2, pp. 248–270, 2008.

[4] A. Asenov, "Random Dopant Induced Threshold Voltage Lowering and Fluctuations in sub-0.1 $\mu$m MOSFET's: A 3-D "atomistic" Simulation Study," *IEEE Transactions on Electron Devices*, vol. 45, no. 12, pp. 2505–2513, 1998.

[5] M. Cho, K. Yuan, Y. Ban, and D. Z. Pan, "ELIAD: Efficient Lithography Aware Detailed Router with Compact Post-Opc Printability Prediction," in *Proceedings of the Design Automation Conference (DAC)*, 2008, pp. 504–509.

[6] A. B. Kahng, C.-H. Park, X. Xu, and H. Yao, "Layout decomposition for double patterning lithography," in *Proceedings of IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2008, pp. 465–472.

[7] J. Warnock, "Circuit and PD challenges at the 14nm technology node," in *Proceedings of ACM international symposium on International symposium on physical design*, 2013, pp. 66–67.

[8] B. Yu, K. Yuan, D. Ding, and D. Z. Pan, "Layout Decomposition for Triple Patterning Lithography," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, vol. 34, no. 3, pp. 433–446, 2015.

[9] B. Yu and D. Z. Pan, "Layout Decomposition for Quadruple Patterning Lithography and Beyond," in *Proceedings of the Design Automation Conference (DAC)*, 2014, pp. 1–6.

[10] X. Wang, B. Cheng, A. R. Brown, C. Millar, and A. Asenov, "Statistical Variability in 14-nm Node SOI FinFETs and Its Impact on Corresponding 6T-SRAM Cell Design," in *Proceedings of Solid-State Device Research Conference (ESSDERC)*, 2012, pp. 113–116.

[11] Y.-K. Choi, T.-J. King, and C. Hu, “Nanoscale CMOS Spacer FinFET for the Terabit Era,” *IEEE Transaction on Electron Device Letters*, vol. 23, no. 1, pp. 25–27, 2002.

[12] M. Jurczak, N. Collaert, A. Veloso, T. Hoffmann, and S. Biesemans, “Review of FINFET technology,” in *Proceedings of IEEE International SOI Conference*, 2009, pp. 1–4.

[13] D. Hisamoto, W.-C. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T.-J. King, J. Bokor, and C. Hu, “FinFET-a Self-Aligned Double-Gate MOSFET Scalable to 20 nm,” *IEEE Transactions on Electron Devices*, vol. 47, no. 12, pp. 2320–2325, 2000.

[14] N. Planes, O. Weber, V. Barral, S. Haendler, D. Noblet, D. Croain, M. Bocat, P.-O. Sassoulas, X. Federspiel, A. Cros *et al.*, “28nm FDSOI Technology Platform for High-Speed Low-Voltage Digital Applications,” in *Proceedings of Symposium on VLSI Technology (VLSIT)*, 2012, pp. 133–134.

[15] Y. K. Ryu, P. A. Postigo, F. Garcia, and R. Garcia, “Fabrication of sub-12 nm thick silicon nanowires by processing scanning probe lithography masks,” *Applied Physics Letters*, vol. 104, no. 22, p. 223112, 2014.

[16] R. Martel, V. Derycke, J. Appenzeller, S. Wind, and P. Avouris, “Carbon Nanotube Field-Effect Transistors and Logic Circuits,” in *Proceedings of Design Automation Conference (DAC)*, 2002, pp. 94–98.

[17] P. Avouris, “Molecular electronics with carbon nanotubes,” *Accounts of Chemical Research*, vol. 35, no. 12, pp. 1026–1034, 2002.

[18] J. Zhang, A. Lin, N. Patil, H. Wei, L. Wei, H.-S. P. Wong, and S. Mitra, “Carbon Nanotube Robust Digital VLSI,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, vol. 31, no. 4, pp. 453–471, 2012.

[19] F. Schwierz, “Graphene Transistors: Status, Prospects, and Problems,” *Proceedings of the IEEE*, vol. 101, no. 7, pp. 1567–1584, 2013.

[20] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, “Single-layer MoS2 transistors,” *Nature nanotechnology*, vol. 6, no. 3, pp. 147–150, 2011.

[21] X. Wang, A. Brown, B. Cheng, and A. Asenov, “Statistical Variability and Reliability in Nanoscale FinFETs,” in *Proceeding of IEEE International Electron Devices Meeting (IEDM)*, 2011, pp. 5–4.

[22] P. Mishra, A. Bhoj, and N. Jha, “Die-level Leakage Power Analysis of FinFET Circuits Considering Process Variations,” in *International Symposium on Quality Electronic Design (ISQED)*, 2010, pp. 347–355.

[23] B. Calhoun, Y. Cao, X. Li, K. Mai, L. Pileggi, R. Rutenbar, and K. L. Shepard, “Digital Circuit Design Challenges and Opportunities in the Era of Nanoscale CMOS,” *Proceedings of the IEEE*, vol. 96, no. 2, pp. 343–365, 2008.

[24] S. R. Nassif, “Design for variability in dsm technologies [deep submicron technologies],” in *Proceedings of International Symposium on Quality Electronic Design (ISQED)*, 2000, pp. 451–454.

[25] “International Technology Roadmap for Semiconductors [Online],” http://www.itrs.net/reports.html.

[26] M. H. Tehranipoor, *Emerging nanotechnologies: test, defect tolerance, and reliability*. Springer Science & Business Media, 2007, vol. 37.

[27] A. Zjajo, *Stochastic Process Variation in Deep-Submicron CMOS*. Springer, 2014.

[28] C. Huang, *Robust Computing with Nano-scale Devices: Progresses and Challenges*. Springer Science & Business Media, 2010, vol. 58.

[29] S. Bangsaruntip, G. Cohen, A. Majumdar, Y. Zhang, S. Engelmann *et al.*, “High Performance and Highly Uniform Gate-All-Around Silicon Nanowire MOSFETs with Wire Size Dependent Scaling,” in *Proceeding of IEEE International Electron Devices Meeting (IEDM)*, 2009, pp. 1–4.

[30] B. Yu, L. Chang, S. Ahmed, H. Wang, S. Bell *et al.*, “FinFET Scaling to 10 nm Gate Length,” in *Proceeding of IEEE International Electron Devices Meeting (IEDM)*, 2002, pp. 251–254.

[31] N. Planes, O. Weber, V. Barral, S. Haendler, D. Noblet *et al.*, “28nm FDSOI Technology Platform for High-Speed Low-Voltage Digital Applications,” in *Proceeding of Symposium on VLSI Technology (VLSIT)*, 2012, pp. 133–134.

[32] A. Colli, A. Fasoli, S. Pisana, Y. Fu, P. Beecher *et al.*, “Nanowire Lithography on Silicon,” *Nano letters*, vol. 8, no. 5, pp. 1358–1362, 2008.

[33] H.-S. Wong, S. Mitra, D. Akinwande, C. Beasley, Y. Chai, H.-Y. Chen, X. Chen, G. Close, J. Deng, A. Hazeghi *et al.*, “Carbon Nanotube Electronics-Materials, Devices, Circuits, Design, Modeling, and Performance Projection,” in *Proceeding of IEEE International Electron Devices Meeting (IEDM)*, 2011, pp. 23–1.

[34] A. K. Geim and K. S. Novoselov, “The Rise of Graphene,” *Nature materials*, vol. 6, no. 3, pp. 183–191, 2007.

[35] J. Wan, C. Le Royer, A. Zaslavsky, and S. Cristoloveanu, “Tunneling FETs on SOI: Suppression of Ambipolar Leakage, Low-Frequency Noise Behavior, and Modeling,” *Elsevier Solid-State Electronics*, vol. 65, pp. 226–233, 2011.

[36] M. Ben-Jamaa, K. Mohanram, and G. De Micheli, “An Efficient Gate Library For Ambipolar CNTFET Logic,” *IEEE Transaction on Computer-Aided Design of Integrated Circuits and Systems*, vol. 30, no. 2, pp. 242–255, 2011.

[37] M. De Marchi, D. Sacchetto, S. Frache, J. Zhang, P.-E. Gaillardon *et al.*, “Polarity Control in Double-gate, Gate-All-Around Vertically Stacked Silicon Nanowire FETs,” in *Proceeding of IEEE International Electron Devices Meeting (IEDM)*, 2012, pp. 8–4.

[38] J. Zhang, P.-E. Gaillardon, and G. De Micheli, “Dual-threshold-voltage configurable circuits with three-independent-gate silicon nanowire fets,” in *Proceeding of IEEE International Symposium on Circuits and Systems (ISCAS)*, 2013, pp. 2111–2114.

[39] Y.-M. Lin, J. Appenzeller, J. Knoch, and P. Avouris, “High-Performance Carbon Nanotube Field-Effect Transistor with Tunable Polarities,” *IEEE Transaction on Nanotechnology*, vol. 4, no. 5, pp. 481–489, 2005.

[40] S. Nakaharai, T. Iijima, S. Ogawa, S.-L. Li, K. Tsukagoshi, S. Sato, and N. Yokoyama, “Electrostatically Reversible Polarity of Dual-Gated Graphene Transistors,” *IEEE Transaction on Nanotechnology*, vol. 13, no. 6, pp. 1039–1043, 2014.

[41] J. Zhang, M. De Marchi, P.-E. Gaillardon, and G. De Micheli, “A Schottky-Barrier Silicon FinFET with 6.0 mV/dec SubthresholdSlope over 5 Decades of Current ,” in *Proceeding of IEEE International Electron Devices Meeting (IEDM)*, 2014, pp. 339–342.

[42] A. Heinzig, S. Slesazeck, F. Kreupl, T. Mikolajick, and W. M. Weber, “Reconfigurable silicon nanowire transistors,” *Nano letters*, vol. 12, no. 1, pp. 119–124, 2011.

[43] M. Mongillo, P. Spathis, G. Katsaros, P. Gentile, and S. De Franceschi, “Multifunctional devices and logic gates with undoped silicon nanowires,” *Nano letters*, vol. 12, no. 6, pp. 3074–3079, 2012.

[44] M. Zervas, D. Sacchetto, G. De Micheli, and Y. Leblebici, “Top-down fabrication of very-high density vertically stacked silicon nanowire arrays with low temperature budget,” *Elsevier Microelectronic Engineering*, vol. 88, no. 10, pp. 3127–3132, 2011.

[45] M. Borg, H. Schmid, K. E. Moselund, D. Cutaia, and H. Riel, “Mechanisms of template-assisted selective epitaxy of InAs nanowires on Si,” *Journal of Applied Physics*, vol. 117, no. 14, p. 144303, 2015.

[46] E. Buitrago, M. Fernández-Bolaños, S. Rigante, C. F. Zilch, N. S. Schröter, A. M. Nightingale, and A. M. Ionescu, “The top-down fabrication of a 3D-integrated, fully CMOS-compatible FET biosensor based on vertically stacked SiNWs and FinFETs,” *Sensors and Actuators B: Chemical*, vol. 193, pp. 400–412, 2014.

[47] S. Maheshwaram, S. Manhas, G. Kaushal, B. Anand, and N. Singh, “Vertical Silicon Nanowire Gate-All-Around Field Effect Transistor Based Nanoscale CMOS,” *IEEE Electron Device Letters*, vol. 32, no. 8, pp. 1011–1013, 2011.

[48] W. Hong, W.-K. Sun, Z.-H. Zhu, H. Ji, B. Song, and W. W.-M. Dai, “A Novel Dimension-Reduction Technique for the Capacitance Extraction of 3-D VLSI Interconnects,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 46, no. 8, pp. 1037–1044, 1998.

[49] C. Heitzinger and S. Selberherr, “An Extensible TCAD Optimization Framework Combining Gradient based and Genetic optimizers,” *Elsevier Microelectronics Journal*, vol. 33, no. 1, pp. 61–68, 2002.

[50] J. K. Lorenz, E. Bär, T. Clees, R. Jancke, C. P. Salzig, and S. Selberherr, “Hierarchical Simulation of Process Variations and Their Impact on Circuits and Systems: Methodology,” *IEEE Transactions on Electron Devices*, vol. 58, no. 8, pp. 2218–2226, 2011.

[51] G. Pei, J. Kedzierski, P. Oldiges, M. Ieong, and E. C.-C. Kan, “FinFET Design Considerations Based on 3-D Simulation and Analytical Modeling,” *IEEE Transactions on Electron Devices*, vol. 49, no. 8, pp. 1411–1419, 2002.

[52] M. Nawaz, W. Molzer, S. Decker, L.-F. Giles, and T. Schulz, “On the Device Design Assessment of Multigate FETs (MuGFETs) Using Full Process and Device Simulation with 3D TCAD,” *Elsevier Microelectronics Journal*, vol. 38, no. 12, pp. 1238–1251, 2007.

[53] A. Zaka, J. Singer, E. Dornel, D. Garetto, D. Rideau, Q. Rafhay, R. Clerc, J.-P. Manceau, N. Degors, C. Boccaccio *et al.*, “Characterization and 3D TCAD Simulation of NOR-type Flash Non-Volatile Memories with Emphasis on Corner Effects,” *Elsevier Solid-State Electronics*, vol. 63, no. 1, pp. 158–162, 2011.

[54] P. Fleischmann, R. Sabelka, A. Stach, R. Strasser, and S. Selberherr, “Grid Generation for Three-Dimensional Process and Device Simulation,” in *Proceedings of IEEE International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, 1996, pp. 161–166.

[55] W. Wessner, J. Cervenka, C. Heitzinger, A. Hössinger, and S. Selberherr, “Anisotropic Mesh Refinement for the Simulation of Three-Dimensional Semiconductor Manufacturing Processes,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, vol. 25, no. 10, pp. 2129–2139, 2006.

[56] R. Strasser and S. Selberherr, *Parallel and Distributed TCAD Simulations Using Dynamic Load Balancing*. Springer, 1998.

[57] A. N. Bhoj, R. V. Joshi, and N. K. Jha, “Efficient Methodologies for 3-D TCAD Modeling of Emerging Devices and Circuits,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, vol. 32, no. 1, pp. 47–58, 2013.

[58] D. F. Specht, “A General Regression Neural Network,” *IEEE Transactions on Neural Networks*, vol. 2, no. 6, pp. 568–576, 1991.

[59] D. Basak, S. Pal, and D. C. Patranabis, “Support Vector Regression,” *Neural Information Processing-Letters and Reviews*, vol. 11, no. 10, pp. 203–224, 2007.

[60] E. M. Biganzoli, F. Ambrogi, and P. Boracchi, “Partial Logistic Artificial Neural Networks (PLANN) for Flexible Modeling of Censored Survival Data,” in *International Joint Conference on Neural Networks (IJCNN)*, 2009, pp. 340–346.

[61] K. Hornik, “Approximation capabilities of multilayer feedforward networks,” *Neural networks*, vol. 4, no. 2, pp. 251–257, 1991.

[62] L. S. Ngia and J. Sjoberg, “Efficient Training of Neural Nets for Nonlinear Adaptive Filtering Using a Recursive Levenberg-Marquardt Algorithm,” *IEEE Transactions on Signal Processing*, vol. 48, no. 7, pp. 1915–1927, 2000.

[63] Z. Feng, P. Li, and Y. Zhan, “An on-the-Fly Parameter Dimension Reduction Approach to Fast Second-Order Statistical Static Timing Analysis,” *IEEE Transaction on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, vol. 28, no. 1, pp. 141–153, 2009.

[64] R. Wang, J. Zhuge, R. Huang, T. Yu, J. Zou *et al.*, “Investigation on Variability in Metal-Gate Si Nanowire MOSFETs: Analysis of Variation Sources and Experimental Characterization,” *IEEE Transaction on Electron Devices*, vol. 58, no. 8, pp. 2317–2325, 2011.

[65] 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,” *Microelectronics Journal*, vol. 44, no. 1, pp. 7–14, 2013.

[66] Y. Zhan, A. J. Strojwas, X. Li, L. T. Pileggi, D. Newmark, and M. Sharma, “Correlation-Aware Statistical Timing Analysis with Non-Gaussian Delay Distributions,” in *Proceedings of Design Automation Conference (DAC)*, 2005, pp. 77–82.

[67] C. Visweswariah, K. Ravindran, K. Kalafala, S. G. Walker, S. Narayan, D. K. Beece, J. Piaget, N. Venkateswaran, and J. G. Hemmett, “First-Order Incremental Block-Based Statistical Timing Analysis,” *IEEE Transaction on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, vol. 25, no. 10, pp. 2170–2180, 2006.

[68] B. Cheng, D. Dideban, N. Moezi, C. Millar, G. Roy, X. Wang, S. Roy, and A. Asenov, “Statistical-Variability Compact-Modeling Strategies for BSIM4 and PSP,” *IEEE Design & Test of Computers*, vol. 27, no. 2, pp. 26–35, 2010.

[69] A. Agarwal, D. Blaauw, and V. Zolotov, “Statistical Timing Analysis for Intra-Die Process Variations with Spatial Correlations,” in *Proceedings of IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2003, p. 900.

[70] Z. Feng and P. Li, “Performance-Oriented Parameter Dimension Reduction of VLSI Circuits,” *IEEE Transaction on Very Large Scale Integration Systems (VLSI)*, vol. 17, no. 1, pp. 137–150, 2009.

[71] K. Chopra, C. Zhuo, D. Blaauw, and D. Sylvester, “A Statistical Approach for Full-Chip Gate-Oxide Reliability Analysis,” in *Proceedings of IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2008, pp. 698–705.

[72] D. S. Boning and P. K. Mozumder, “DOE/Opt: A System for Design of Experiments, Response Surface Modeling, and Optimization Using Process and Device Simulation,” *IEEE Transaction on Semiconductor Manufacturing*, vol. 7, no. 2, pp. 233–244, 1994.

[73] D. Blaauw, K. Chopra, A. Srivastava, and L. Scheffer, “Statistical Timing Analysis: From Basic Principles to State of the Art,” *IEEE Transaction on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, vol. 27, no. 4, pp. 589–607, 2008.

[74] C. M. Bishop, *Pattern Recognition and Machine Learning.* springer New York, 2006.

[75] A. Hyvärinen and E. Oja, “Independent Component Analysis: Algorithms and Applications,” *Neural networks*, vol. 13, no. 4, pp. 411–430, 2000.

[76] L. Cheng, P. Gupta, and L. He, “Accounting for Non-Linear Dependence Using Function Driven Component Analysis,” in *Proceedings of Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2009, pp. 474–479.

[77] R. Huang, R. Wang, J. Zhuge, T. Yu, Y. Ai *et al.*, “Variability Investigation of Gate-All-Around Silicon Nanowire Transistors from Top-Down Approach,” in *Proceedings of IEEE International Conference on Electron Devices and Solid-State Circuits (EDSSC)*, 2010, pp. 1–4.

[78] R. Tibshirani, “Regression Shrinkage and Selection via the Lasso,” *Journal of the Royal Statistical Society*, pp. 267–288, 1996.

[79] J. J. Moré, “The Levenberg-Marquardt Algorithm: Implementation and Theory,” in *Numerical analysis*, 1978, pp. 105–116.

[80] M. Stone, “Cross-Validation: a Review 2,” *Statistics: A Journal of Theoretical and Applied Statistics*, vol. 9, no. 1, pp. 127–139, 1978.

[81] N. Simon, J. Friedman, T. Hastie, and R. Tibshirani, “A Sparse-Group Lasso,” *Journal of Computational and Graphical Statistics*, vol. 22, no. 2, pp. 231–245, 2013.

[82] *Synopsys PrimeTime [Online]*, http://www.synopsys.com.

[83] Y. Zhang, S. Sankaranarayanan, and F. Somenzi, “Sparse Statistical Model Inference for Analog Circuits under Process Variations,” in *Proceedings of Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2014, pp. 449–454.

[84] E. J. McCluskey and C.-W. Tseng, “Stuck-fault tests vs. actual defects,” in *Proceedings of IEEE International Test Conference (ITC)*, 2000, pp. 336–342.

[85] P. Franco and E. J. McCluskey, “Three-Pattern Tests for Delay Faults,” in *Proceedings of IEEE VLSI Test Symposium (VTS)*, 1994, pp. 452–456.

[86] V. Champac, J. V. Hernandez, S. Barcelo, R. Gomez, C. Hawkins *et al.*, “Testing of Stuck-Open Faults in Nanometer Technologies,” *IEEE Design & Test of Computers*, vol. 29, no. 4, pp. 80–91, 2012.

[87] T. M. Storey and W. Maly, “CMOS Bridging Fault Detection,” in *Proceedings of IEEE International Test Conference (ITC)*, 1990, pp. 842–851.

[88] M. O. Simsir, A. Bhoj, and N. K. Jha, “Fault Modeling for FinFET Circuits,” in *Proceedings of IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH)*, 2010, pp. 41–46.

[89] A. N. Bhoj, M. O. Simsir, and N. K. Jha, “Fault Models for Logic Circuits in the Multigate Era,” *IEEE Transaction on Nanotechnology*, vol. 11, no. 1, pp. 182–193, 2012.

[90] Y. Liu and Q. Xu, “On Modeling Faults in FinFET Logic Circuits,” in *Proceedings of IEEE International Test Conference (ITC)*, 2012, pp. 1–9.

[91] G. Harutyunyan, G. Tshagharyan, V. Vardanian, and Y. Zorian, “Fault Modeling and Test Algorithm Creation Strategy for FinFET-Based Memories,” in *Proceedings of IEEE VLSI Test Symposium (VTS)*, 2014, pp. 1–6.

[92] C.-W. Lin, M.-T. Chao, and C.-C. Hsu, “Investigation of Gate Oxide Short in FinFETs and the Test Methods for FinFET SRAMs,” in *Proceedings of IEEE VLSI Test Symposium (VTS)*, 2013, pp. 1–6.

[93] R. Rajsuman, “Iddq Testing for CMOS VLSI,” *Proceedings of the IEEE*, vol. 88, no. 4, pp. 544–568, 2000.

[94] Yu WeiP’ng and Moo Kit Lee and Peng Weng Ng and Chin Hu Ong, “IDDQ Test Challenges in Nanotechnologies: A Manufacturing Test Strategy,” in *Proceedings of IEEE Asian Test Symposium (ATS)*, 2007, pp. 211–211.

[95] T. G. Dziura, B. Bunday, C. Smith, M. M. Hussain, R. Harris *et al.*, “Measurement of High-k and Metal Film Thickness on FinFET Sidewalls using Scatterometry,” in *Proceedings of SPIE Advanced Lithography*, 2008, pp. 69 220V–69 220V.

[96] N. Shin, M. Chi, J. Y. Howe, and M. A. Filler, “Rational defect introduction in silicon nanowires,” *Nano letters*, vol. 13, no. 5, pp. 1928–1933, 2013.

[97] N. Shin, M. Chi, and M. A. Filler, “Interplay Between Defect Propagation and Surface Hydrogen in Silicon Nanowire Kinking Superstructures,” *ACS nano*, vol. 8, no. 4, pp. 3829–3835, 2014.

[98] C.-C. Yang, C. Witt, P.-C. Wang, D. Edelstein, and R. Rosenberg, “Stress Control During Thermal Annealing of Copper Interconnects,” *Applied Physics Letters*, vol. 98, no. 5, pp. 051 911–051 911, 2011.

[99] S. Bobba, P.-E. Gaillardon, J. Zhang, M. De Marchi, D. Sacchetto, Y. Leblebici, and G. De Micheli, “Process/Design Co-Optimization of Regular Logic Tiles for Double-Gate Silicon Nanowire Transistors,” in *Proceedings of IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH)*, 2012, pp. 55–60.

[100] O. Zografos, P.-E. Gaillardon, and G. De Micheli, “Novel Grid-Based Power Routing Scheme for Regular Controllable-Polarity FET Arrangements,” in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS)*, 2014, pp. 1416–1419.

[101] H. G. Mohammadi, P.-E. Gaillardon, and G. De Micheli, "Fault modeling in controllable polarity silicon nanowire circuits," in *Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2015, pp. 453–458.

[102] P. M. Kogge and H. S. Stone, "A parallel algorithm for the efficient solution of a general class of recurrence equations," *IEEE Transactions on Computers*, vol. 100, no. 8, pp. 786–793, 1973.

[103] S. Ghosh, P. Ndai, and K. Roy, "A novel low overhead fault tolerant Kogge-Stone adder using adaptive clocking," in *Proceedings of the Conference on Design, Automation and Test in Europe (DATE)*, 2008, pp. 366–371.

[104] S. K. Mitra and A. R. Chowdhury, "Minimum cost fault tolerant adder circuits in reversible logic synthesis," in *Proceedings of International Conference on VLSI Design (VLSID)*, 2012, pp. 334–339.

[105] L. Anghel and M. Nicolaidis, "Defects tolerant logic gates for unreliable future nanotechnologies," in *Computational and Ambient Intelligence*, 2007, pp. 422–429.

[106] C.-I. H. Chen and A. Kumar, "Area-time optimal digital BiCMOS carry look-ahead adder," in *Proceedings of Asia-Pacific Conference on Circuits and Systems (APCCAS)*. IEEE, 1994, pp. 115–120.

[107] M. Nicolaidis, "Carry checking/parity prediction adders and ALUs," *IEEE Transactions on Very Large Scale Integration Systems (VLSI)*, vol. 11, no. 1, pp. 121–128, 2003.

[108] E. Sogomonyan, V. Ocheretnij, and M. Gössel, "A new code-disjoint sum-bit duplicated carry look-ahead adder for parity codes," in *Proceedings of Asian Test Symposium (ATS)*, 2001, pp. 365–370.

[109] G. Langdon and C. Tang, "Concurrent error detection for group look-ahead binary adders," *IBM Journal of Research and Development*, vol. 14, no. 5, pp. 563–573, 1970.

[110] O. Turkyilmaz, F. Clermidy, L. G. Amarù, P.-E. Gaillardon, and G. De Micheli, "Self-checking ripple-carry adder with ambipolar silicon nanowire FET," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS)*, 2013, pp. 2127–2130.

[111] M. Bushnell and V. D. Agrawal, *Essentials of electronic testing for digital, memory and mixed-signal VLSI circuits*. Springer Science & Business Media, 2000, vol. 17.

[112] F. Hapke, W. Redemund, A. Glowatz, J. Rajski, M. Reese, M. Hustava, M. Keim, J. Schloeffel, and A. Fast, "Cell-Aware Test," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, vol. 33, no. 9, pp. 1396–1409, 2014.

[113] T. Ban and L. de Barros Naviner, "A Simple Fault-tolerant Digital Voter Circuit in TMR Nanoarchitectures," in *Proceedings of IEEE International NEW Circuits And Systems (NEWCAS)*, 2010, pp. 269–272.

# Curriculum Vitae

## HASSAN GHASEMZADEH MOHAMMADI

Rue de la Délèze, 18
1920 Martigny, Switzerland
Mobile : +41786028254
Email : ghasemzadeh@gmail.com

## EDUCATION

| | |
|---|---|
| **Ph.D.** | **Computer Science** - Sept. 2011 - Dec. 2015<br>École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland<br>**Thesis:** Robustness Analysis of Controllable-Polarity Silicon Nanowire Devices and Circuits |
| **M.Sc.** | **Computer Engineering** - Sept. 2006 - Sept. 2008<br>Sharif University of Technology (SUT), Tehran, Iran<br>**Thesis:** Design and Evaluation of a Fault-Tolerant Control Unit for an Embedded Microprocessor |
| **B.Sc.** | **Computer Engineering** - Sept. 2000 - Sept. 2005<br>Iran University of Science & Technology (IUST), Tehran, Iran<br>**Thesis:** FPGA-Based Classification of Fingerprint Images Using Neural Networks |

## TECHNICAL EXPERTISE

| | |
|---|---|
| **Digital Circuit Design** | VHDL, Verilog, SystemVerilog, familiar with UVM |
| **Simulation Environments** | Synopsys Design Compiler & PrimeTime, Mentor Modelsim & Questa, Hspice, Xilinx ISE, Altera Quartus, MATLAB, Simulink |
| **Programming Languages** | C, C++, Java and C#, Linux shell scripting |

## TEACHING ASSISTANT

| | |
|---|---|
| **2012 – 2014** | **Design Technologies for Integrated Circuits (DTIS),** Instructor: Prof. Giovanni De Micheli<br>➢ The focus of this course was on the synthesis algorithms and optimization of digital circuits<br>School of Computer and Communication Sciences (IC), EPFL |
| **Fall 2007** | **Digital Logic Circuit Lab.,** Instructor: Prof. Afshin Hemmatyar<br>School of Computer Engineering, SUT |
| **Fall 2005** | **VLSI,** Instructor: Prof. Farshad Safaei<br>School of Computer Engineering, IUST |

## RESEARCH & JOB EXPERIENCE

| | |
|---|---|
| **Sept. 2011 – present** | **Research Assistant**, Integrated System Laboratory (LSI)<br>Supervisor: Prof. G. De Micheli<br>➢ Study the effect of process variation on the performance parameters (e.g., timing) of logic circuits for novel technologies such as FinFET, and silicon nanowires<br>➢ Two novel techniques to speed up process variation and timing analysis of large designs<br>➢ Developing gate-level fault models, and test techniques for logic circuits in controllable-polarity silicon nanowire technology<br>➢ Implementation of a low overhead and fault tolerant adder in controllable-polarity silicon nanowire technology |
| **Sept. 2007 – Sept. 2009** | **Research Assistant**, Dependable Systems Laboratory (DSL)<br>Supervisors: Prof. S.G. Miremadi, and Prof. A. Ejlali<br>➢ Implementation of an embedded microprocessor based on SPARC-V8 ISA for automotive applications: (As a member of a team of 8persons, I was working on the pipeline core of the microprocessor)<br>➢ Front-end implementation using VHDL<br>➢ Synthesis, verification, and post-synthesis simulation for ASIC (using TSMC 130 nm technology)<br>➢ Implementation of a novel fault detection technique for random logic (control logic) part of the microprocessor<br>➢ Research on the low power design technologies for developing secure logics against side channel attack |
| **Fall 2009** | **Lecturer**, Faculty of Electrical and Computer Engineering, QIAU, Qazvin, Iran<br>**Course title**: Electronic Circuits |
| **Fall 2008** | **Lecturer**, Faculty of Computer Engineering, Payame Noor University, Qazvin, Iran<br>**Course title**: Introduction to Microprocessor Systems |
| **Oct. 2005 – July. 2006** | **C++ Programmer**, Dadeh Pardazan Aftabe Shargh Co., Tehran, Iran<br>➢ Developing a Content Manager System (CMS) as a "distributed educational environment" using C++ Builder |

## RELEVANT COURSES

| | |
|---|---|
| **EPFL** | Synthesis and optimization of digital circuits, Microelectronic for System-on-Chips, Co-design of Systems-on-Chip on Reconfigurable Hardware, Advanced algorithms, Machine Learning |

**SUT** Advanced computer architecture, Low power digital circuit design, Fault-tolerant system design, Embedded system design, Advanced VLSI, Digital signal processing

## LANGUAGE SKILLS

**English(**full-working proficiency**)** **French(**A2, in progress**)** **German(**A1, in progress**)** **Persian(**mother tongue**)**

## EXTRA

**Status** 33 years old, Married, Owner of Swiss permit-B

**Other activities** **Sports:** Squash, Swimming**, Hobbies:** Playing "Setar" (an Iranian instrument)

## PUBLICATIONS

**Journals**

1. **H. Ghasemzadeh Mohammadi**, P-E. Gaillardon, and G. De Micheli, "From Defect Analysis to Gate-Level Fault Modeling of Controllable-Polarity Silicon Nanowires," IEEE Transactions on Nanotechnology (**TNANO**), Vol. 14, pp. 1-10, 2015.
2. **H. Ghasemzadeh Mohammadi,** P-E. Gaillardon, and G. De Micheli, "Efficient Statistical Parameter Selection for Nonlinear Modeling of Process/Performance Variation," submitted to IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (**TCAD**).
3. **H. Ghasemzadeh Mohammadi**, P-E. Gaillardon, J. Zhang, G. De Micheli, E. Sanchez, M. Sonza Reorda, "A Fault Tolerant Adder with Controllable-Polarity Transistors," submitted to ACM Journal on Emerging Technologies in Computing (**JETC**).

**Conferences**

1. **H. Ghasemzadeh Mohammadi**, P-E. Gaillardon, J. Zhang, G. De Micheli, E. Sanchez, M. Sonza Reorda, "On the Design of a Fault Tolerant Ripple-Carry Adder with Controllable-Polarity Transistors," IEEE Computer Society Annual Symposium on VLSI (**ISVLSI**), 2015.
2. **H. Ghasemzadeh Mohammadi**, P-E. Gaillardon, and G. De Micheli, "Fast Parametric Fault Modeling of Nanoscale Integrated Circuits," DATE'15 Workshop on Designing with Uncertainty - Opportunities & Challenges, 2015.
3. **H. Ghasemzadeh Mohammadi**, P-E. Gaillardon, and G. De Micheli, "Fault Modeling in Controllable-Polarity Silicon Nanowire Circuits," in Proc. the Design, Automation & Test in Europe (**DATE**), 2015.
4. **H. Ghasemzadeh Mohammadi**, P-E. Gaillardon, M. Yazdani, and G. De Micheli, "Fast process variation analysis in nano-scaled technologies using column-wise sparse parameter selection," IEEE/ACM International Symposium on Nanoscale Architectures (**NANOARCH**), 2014.
5. **H. Ghasemzadeh Mohammadi**, P-E. Gaillardon, and G. De Micheli, "Fast process variation analysis in emerging nano-scaled technologies," Designing with Uncertainty - Opportunities & Challenges workshop, 2014.
6. **H. Ghasemzadeh Mohammadi**, P-E. Gaillardon, M. Yazdani, and G. De Micheli, "A fast TCAD-based methodology for Variation analysis of emerging nano-devices," IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (**DFT**), 2013.
7. P-E. Gaillardon, **H. Ghasemzadeh Mohammadi**, and G. De Micheli, "Vertically-stacked silicon nanowire transistors with controllable polarity: A robustness study," IEEE 14th Latin American Workshop (**LATW**), 2013.

8. M. Khatir, **H. Ghasemzadeh Mohammadi**, and A. Ejlali, “Sub-threshold charge recovery circuits,” in Proc. IEEE International Conference on Computer Design (**ICCD**), 2010.
9. **H. Ghasemzadeh Mohammadi**, S.G. Miremadi, and A. Ejlali, “Signature Self Checking (SSC): A Low-Cost Reliable Control Logic for Pipelined Microprocessors,” in Proc. IEEE Pacific Rim International Symposium on Dependable Computing (**PRDC**), 2009.
10. **H. Ghasemzadeh Mohammadi**, H. Tabkhhi, S.G. Miremadi, and A. Ejlali, “A Cost-Effective Error Detection and Roll-back Recovery Technique for Embedded Microprocessor Control Logic,” in Proc. IEEE International Conference on Microelectronics (**ICM**) 2008.