1-1-2011

Robust Circuit & Architecture Design in the Nanoscale Regime

Rehman Ashraf
Portland State University

Robust Circuit & Architecture Design in the

Nanoscale Regime

by

Rehman Ashraf

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Electrical and Computer Engineering

Dissertation Committee:
Malgorzata Chrzanowska-Jeske, Chair
Siva Narendra
W. Robert Daasch
Douglas V. Hall
Tugrul U. Daim

Portland State University
©2011

ABSTRACT

Silicon based integrated circuit (IC) technology is approaching its physical limits. For sub 10nm technology nodes, the carbon nanotube (CNT) based field effect transistor has emerged as a promising device because of its excellent electronic properties. One of the major challenges faced by the CNT technology is the unwanted growth of metallic tubes. At present, there is no known CNT fabrication technology which allows the fabrication of 100% semiconducting CNTs. The presence of metallic tubes creates a short between the drain and source terminals of the transistor and has a detrimental impact on the delay, static power and yield of CNT based gates.

This thesis will address the challenge of designing robust carbon nanotube based circuits in the presence of metallic tubes. For a small percentage of metallic tubes, circuit level solutions are proposed to increase the functional yield of CNT based gates in the presence of metallic tubes. Accurate analytical models with less than a 3% inaccuracy rate are developed to estimate the yield of CNT based circuit for a different percentage of metallic tubes and different drive strengths of logic gates. Moreover, a design methodology is developed for yield-aware carbon nanotube based circuits in the presence of metallic tubes using different CNFET transistor configurations. Architecture based on regular logic bricks with underlying hybrid CNFET configurations are developed which gives better trade-offs in terms of performance, power, and functional yield.

In the case when the percentage of metallic tubes is large, the proposed circuit level techniques are not sufficient. Extra processing techniques must be applied to remove the metallic tubes. The tube removal techniques have trade-offs, as the removal process is not perfect and removes semiconducting tubes in addition to removing unwanted metallic tubes. As a result, stochastic removal of tubes from the drive and fanout gate(s) results in large variation in the performance of CNFET based gates and in the worst case open circuit gates. A Monte Carlo simulation engine is developed to estimate the impact of the removal of tubes on the performance and power of CNFET based logic gates. For a quick estimation of functional yield of logic gates, accurate analytical models are developed to estimate the functional yield of logic gates when a fraction of the tubes are removed.

An efficient tube level redundancy (TLR) is proposed, resulting in a high functional yield of carbon nanotube based circuits with minimal overheads in terms of area and power when large fraction of tubes are removed. Furthermore, for applications where parallelism can be utilized we propose to increase the functional yield of the CNFET based circuits by increasing the logic depth of gates.

DEDICATION

To my parents (Muhammad Ashraf Butt and Sakina Begum) who taught me the value of education

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

AFM:	Atomic Force Mircoscopy
ALD:	Atomic Layer Deposition
ASIC:	Application Specific Integrated Circuit
BSIM4:	Berkeley Short-channel IGFET Model
BTBT	Band to Band Tunneling
CNFET:	Carbon Nanotube Field Effect Transistor
CNT:	Carbon Nanotube
CPU:	Central Processing Unit
CVD:	Chemical Vapor Deposition
DIBL:	Drain Induced Barrier Lowering
EDP:	Energy Delay Product
GIDL:	Gate Induced Drain Lowering
IGFET:	Insulated Gate Field Effect Transistor
ITRS:	International Technology Roadmap for Semiconductors
LNA:	Low Noise Amplifier
MC:	Monte Carlo
MIM:	Metal Insulator Metal
MUX:	Multiplexer
N-CNFET:	N-type CNFET
NWFET:	Nanowire Field Effect Transistor
OPC:	Optical Proximity Correction
P-CNFET:	P-type CNFET
PECVD:	Plasma Enhanced Chemical Vapor Deposition
SB:	Schottky Barrier
SCE:	Selective Chemical Etching
SP:	Static Power

SWCNT: Single-Walled Carbon Nanotube

TLR: Tube Level Redundancy

TRCS: Tube Removed and Charge Screening Considered

VLSI: Very Large Scale Integration

VMR: VLSI Compatible Metallic Carbon Nanotube Removal

1   Introduction

1.1   Need for CMOS Alternatives

The driving force for semiconductor industry growth has been the elegant scaling nature of CMOS technology. In nanoscale CMOS technology nodes, supply voltage (V_{DD}) and threshold voltage (V_{th}) must continually scale in order to sustain a performance increase, limit energy consumption, control power dissipation, and maintain reliability. The scaling of CMOS technology has sustained over the last four decades, but is now approaching atomistic and quantum-mechanical physics limits [1]. Some of the main challenges faced by the Si CMOS technology are large short channel effects resulting in an exponential increase in leakage power, process variations resulting in large deviations in the performance of the circuits and technological limitations.

In nano-scale CMOS devices, leakage power is the major contributor to total power consumption. Figure 1-1 shows the six mechanisms which contribute to total leakage power in the short channel devices. In Figure 1-1, I_1 is the leakage current due to the reverse-bias pn junction, I_2 is the leakage current due to the subthreshold leakage, I_3 is the current due to the tunneling of carriers through the thin gate oxide, I_4 is the current flowing in the gate because of an injection of hot carriers, I_5 is the current because of Gate Induced Drain Lowering (GIDL) and finally I_6 is the current because of a channel punch through.
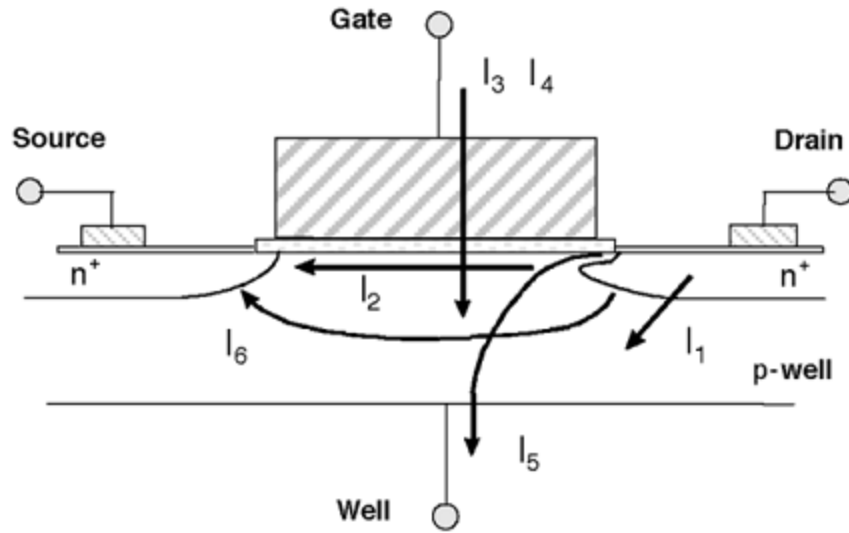
Figure 1-1: Different sources of leakage currents in nano-scale transistors.

The process variations result in the increased parametric variation of the CMOS devices [1]. The major impacts of process variations are on the variation in the channel length L, width W, and threshold voltage V_{th}. The variations in the L, and W are mainly caused due to a limited resolution of photolithography effects. Similarly, variations in threshold voltage are caused by the variation in both the doping concentration in the channel, and in the oxide thickness. The process variations result in a significant deviation in the performance and the power of digital circuits from their nominal values.

The fabrication of CMOS transistors is obtained by patterning, which is achieved by a combination of photolithography and masks. Therefore, the size of the smallest feature size that can be patterned is dependent on the wavelength of light. The patterning of feature sizes that are smaller than the wavelength of light, although possible, result in an increase of complexity and costs of the masks. Figure 1-2 shows the evolution of optical masks starting

2

from 180nm technology nodes to less than 45nm technology nodes [2]. From the figure it can be observed that migration from 180nm to 130nm technology nodes required the Rule/Model based optical proximity correction (OPC) techniques. Similarly, for technology nodes 65nm or below, the complexity of the masks required more advanced techniques. It is reported in [2] that for technology nodes with a feature size of <22nm the complexity of the processes involved in the masks may result in these approaches to be economically unviable.
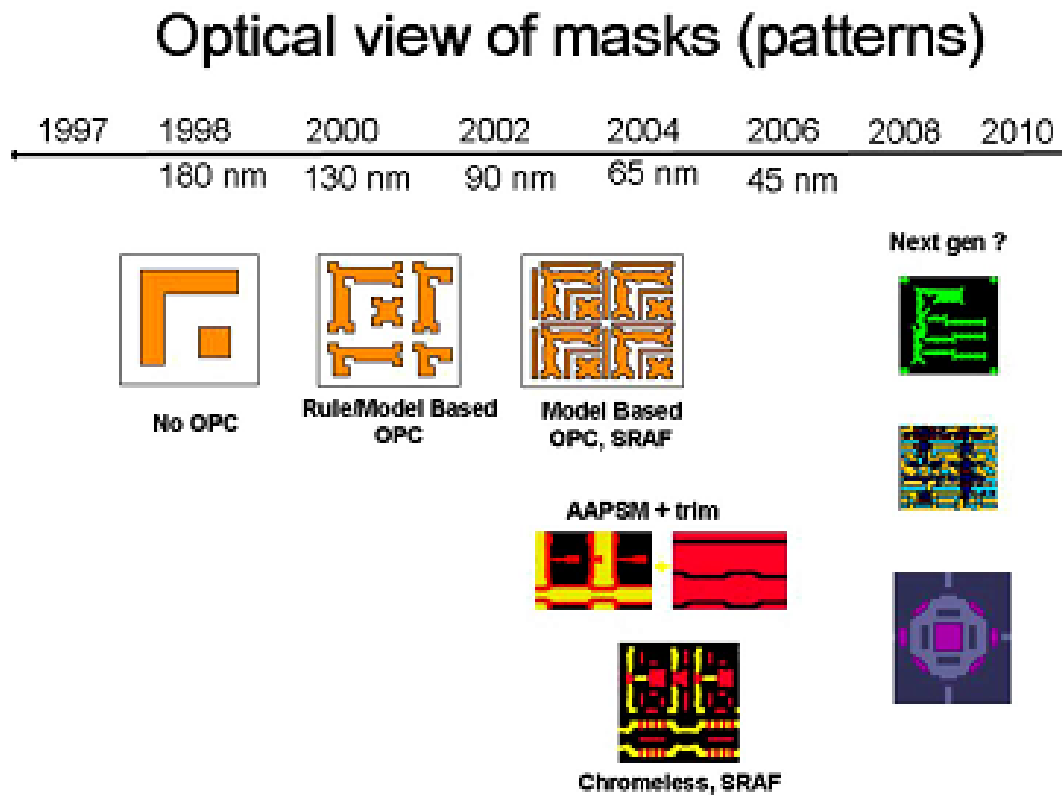
Figure 1-2: Evolution of optical masks for patterning of different technology nodes.

Scaling of Si CMOS is continued by innovations like the use of strained-Si channels, high K-dielectrics and metal gate electrodes. The application of strain on silicon increases the

mobility of carriers, which in turn improves the drive current, and performance of the transistor. Different strain mechanisms are required for NMOS and PMOS transistors. For an NMOS transistor, an insulating film of silicon nitride (SiN) is applied on the gate of the NMOS transistor which creates a tensile stress on the channel. In the PMOS transistor, a compressive stress is applied by putting the epitaxial layer of silicon germanium (SiGe) in the source and drain slots of transistors [3]. The application of strain increases the mobility of NMOS transistors by 40%, and PMOS transistors by 100% as compared to transistors without the application of strain [4]. The application of high-K gate dielectrics and metal gate electrodes help to significantly reduce the gate leakage. Figure 1-3 shows the reduction in gate leakage by incorporating the high-k gate dielectric and metal gate [5]. From the figure it can be observed that the application of high-k gate dielectric and metal gate reduced the gate leakage by 25X, while migrating the technology from 65nm to 45nm. The leakage reduction is due to the use of high-k gate dielectric.
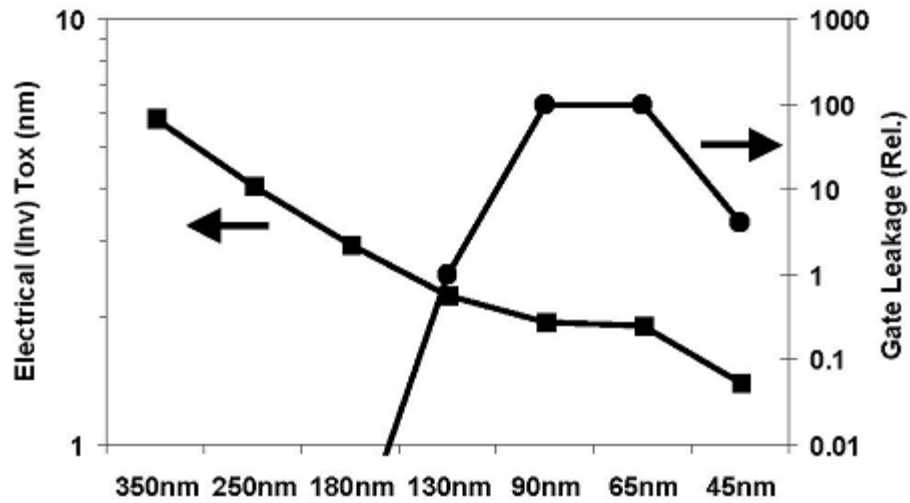
Figure 1-3: The scaling trend of Intel's inversion electrical T$_{OX}$ and gate leakage for different technology nodes [5].

To reduce the short channel effects, researchers have proposed double-gate MOSFETs and finfets/tri-gate devices [6], [7]. In tri-gate devices the gate is placed on the three sides of the channel as shown in Figure 1-4 . This results in a better control on the channel and significant reduction in the drain to source subthreshold leakage current.
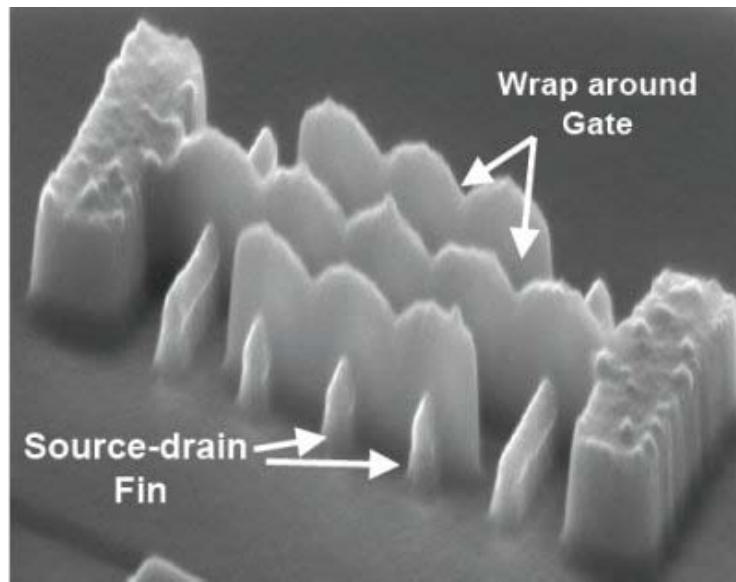
Figure 1-4: Micrograph of a tri-gate transistor developed by Intel.

1.2 Emerging Logic Devices

Researchers have also started exploring new devices and channel materials in the sub-10nm technology nodes that have the potential to become the successor of Si-CMOS. According to ITRS [8] some of the emerging logic devices which have the potential to replace Si in the post Si era are:

a) Nanowire Field-Effect Transistors(NWFETs) [9]

b) III-V compound semiconductor Field-Effect Transistors [10-13]

c) Graphene Nanoribbon Field-Effect Transistors

d) Carbon nanotube Field-Effect Transistors (CNFETs) [14]

Nanowire Transistors: In Nanowire transistors a semiconducting nanowire of diameter around 0.5nm is used as a channel material. The nanowire material can be of silicon (Si),

6

germanium (Ge), III-V, In_2O_3, ZnO or SiC semiconductors. The remaining structure of the silicon nanowire transistor is similar to conventional CMOS. Nanowire transistors have been reported by several groups [15][16].Figure 1-5 shows the schematic of a fabricated NWFET where silicon nanowire is used as a channel material [17]. The main advantage of using the small diameter nanowire is to obtain 1-D conduction, minimizing the short channel effects. The fundamental challenge faced by nanowire based transistors is the fabrication of conventional diffused P-N junctions in nanowire devices. Current technologies use metal source drain junctions, resulting in ambipolar conduction [18]. This produces a large OFF state current in the nanowire devices.
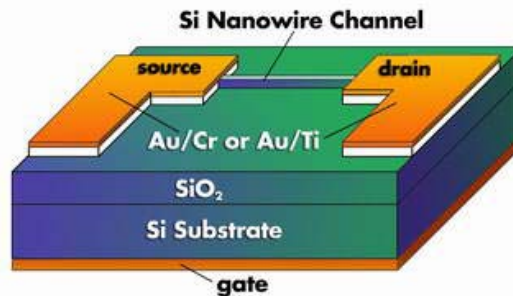
Figure 1-5: Schematic diagram of nanowire FET. Silicon is used as the channel material for nanowire.

III-V compound semiconductor FET: The III-V compound semiconductor FET uses III-V compound semiconductor such as InSb, InAs, InGaAs as a channel material. Higher performance can be obtained from these devices because of the high mobility of carriers in these materials compared to CMOS devices. These III-V compound semiconductor FETs have the potential to deliver 3X higher performance than silicon at iso-power consumption, or can deliver the same performance as obtained by silicon transistors at one-tenth the power consumption of silicon [19]. Figure 1-6 shows the schematic of an n-type MOSFET

[20]. Here, ZrO2 is used as the gate dielectric and InGaAs is used as the channel material. The mobility of the device is reported to be 3000 cm^2/Vs[20].There are two main challenges faced by the III-V semiconductor devices, *1)* the III-V materials have lower bandgaps, resulting in excessive leakage and large static power consumption in III-V semiconductor devices, *2)* the problem of forming a compatible high-k dielectric interface [21] which is essential in the electrostatic control of the device.
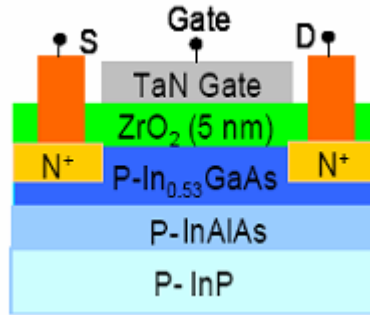
Figure 1-6: Schematic of a n-type MOSFET with InGaAs used as the channel material and ZrO₂ as the gate dielectric [20].

Graphene Nanoribbon Transistor: In the graphene nanoribbon transistor, a monolayer of carbon atoms, packed into a two-dimensional honeycomb lattice is used as the channel material. Figure 1-7(a) shows the schematic diagram of nanoribbons FET transistor fabricated with nanoribbons with a width of ~2nm [22]. Figure 1-7(b) shows the AFM image of the graphene nanoribbon FET, where 10nm thick SiO_2 is used as dielectric and Pd is used as source and drain contacts and P^{++} is used as the backgate. An advantage of using graphene as channel material is a very high mobility (15,000 cm^2/Vs) [23], high carrier velocity which will result in fast switching, monolayer thin body for optimum electrostatic scaling, and excellent thermal conductivity. It is expected that the integrated circuits

8

fabricated with graphene based transistors can be 100X to 1000X faster than silicon devices [24]. Graphene based transistors are reported by different research groups [22], [25], [26]. In 2010, a transistor operating at 100 GHz has been reported by IBM [27]. Graphene transistors are expected to beat the performance of the fastest transistors fabricated with other materials, if researchers can overcome the challenges faced by the graphene technology. The major challenge faced by graphene based transistors is the comparatively low I_{ON}/I_{OFF} ratio of ~7[28], a measure of how much power is consumed by the circuit when it is in the standby state. In case of low I_{ON}/I_{OFF} ratio, the integrated circuit made of billions of graphene transistors will consume an enormous amount of energy [29].

(a)   **(b)**
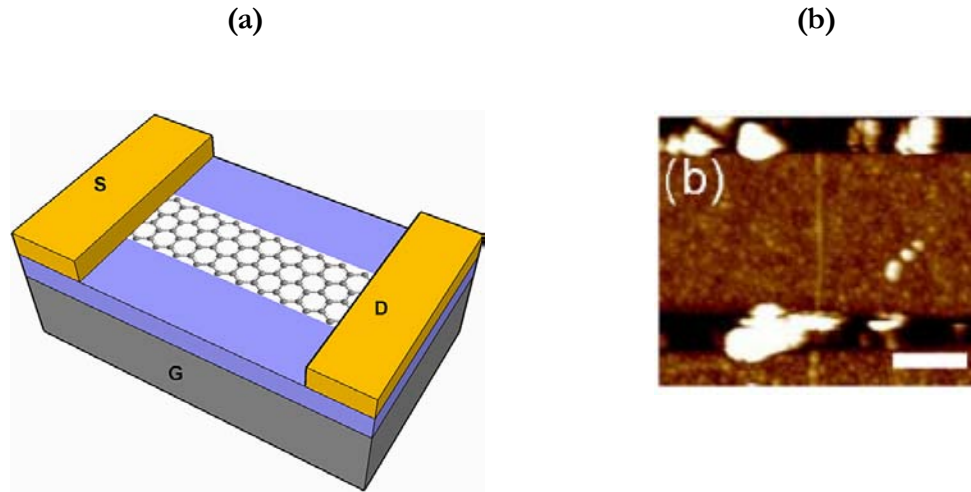
Figure 1-7: (a) Schematic of a graphene nanoribbons FET (b) AFM image of graphene nanoribbon FET with w~2±0.5nm.

Carbon Nanotube based FET (CNFET): The CNFET has the potential to become the channel material of future nanoscale transistors because of the excellent electronic properties of carbon nanotubes, such as near ballistic transport [30], high carrier mobility

$(10^3 \sim 10^4 \text{cm}^2/\text{Vs})$, in semiconducting CNTs [31], and easy integration of high-k dielectric material [32] resulting in better gate electrostatics.

CNFET uses a single-walled carbon nanotube (SWCNT) as channel material. The control electrode (gate) is placed above the conduction channel and separated from it by a thin layer of dielectric (gate oxide). Figure 1-8 shows the side and top view of a CNFET where an array of four single-walled CNTs is used as a channel. The first carbon nanotube based transistor was demonstrated by Dekker et al. [33] and by IBM in 1998 [34]. After that demonstration, significant progress was made in the fabrication of carbon nanotube based devices and circuits. Physical implementations of inverters[35],5 stage ring oscillator[36], NAND, NOR gates and SRAM cells [37] built with CNFETs have been demonstrated by various research groups. In 2006, IBM announced that they built the first integrated circuit using a single-walled carbon nanotube [38]. Rogers et al. demonstrated medium scale integrated circuits built with CNFET based transistors on a thin plastic substrate [39].At present, the fundamental challenges faced by carbon nanotube based technology are the unwanted growth of metallic tubes, and the placement and alignment of an array of aligned carbon nanotubes [40].
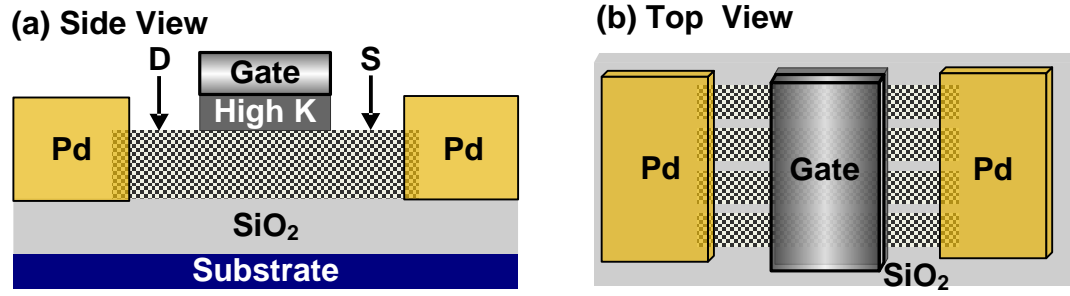
Figure 1-8: (a) Cross section view of CNFET (b) Top view of CNFET layout with an array of four parallel CNTs.

The above mentioned logic devices have the potential to replace silicon in the post silicon era. NWFET and CNFET are 1-D devices, graphene FET is a 2-D device, and the III-V compound semiconductor transistor is a 3-D device. Out of these, 1-D devices (NWFET and CNFET) allow the ballistic transport of carriers in the channel without any scattering. As a result, performance of these devices is superior to 2-D and 3-D devices. The absence of dangling bonds at the CNT surface allows an easy integration of high-K dielectric resulting in better gate electrostatics, which in turn results in lower sub-threshold slopes and lower *OFF* current in CNT based devices. As previously discussed, the mobility of carriers in III-V semiconductor FET, graphene FET, CNFET and in NWFET (depending on the channel material used in NWFET) is higher than silicon resulting in higher carrier velocities. The mobilities of CNFET and graphene are in the same order of magnitude (10,000 cm^2/Vs to 15,000 cm^2/Vs) making them a strong candidate for future devices. Moreover, according to ITRS 2009[8], carbon nanotube and graphene based transistors show the highest potential of being part of future giga-scale integrated circuits. When this research started, R&D in CNFETs was leading, whereas graphene was recently introduced. Therefore, the focus of this work is on carbon nanotube based devices and circuits.

11

1.3　Contributions of This Work

In the previous section the challenges faced by CNT based technology were briefly mentioned. This dissertation focuses on the two aspects of such challenges: *a)* the impact of spacing among CNTs and the variability in diameter and spacing of CNTs on the performance of CNFETs, and *b)* the impact of the presence of metallic tubes on the performance, power, and yield of CNFET based circuits. Two different approaches are proposed depending upon the percentage of metallic tubes in a given CNT fabrication technology.

1.3.1　Spacing among CNTs and variability in diameter and spacing of CNTs

The spacing between adjacent CNTs impacts the performance of CNFETs. In chapter 4 the impact of spacing among adjacent CNTs on the performance of CNTs is analyzed. Moreover, the fabrication of CNT results in variation in their diameter, as well as spacing among them. The analysis of variation in the diameter and spacing is done in order to examine their impact on the performance of CNFETs. A tool is developed to stochastically estimate the spacing impact on the performance of drive strength of CNFETs. Finally a methodology is developed for variation-tolerant CNFET based circuit design.

1.3.2　When the Percentage of Metallic Tubes is Smaller

As discussed in Section 1.2, one of the major challenges faced by the CNT technology is the unwanted growth of metallic tubes, which severely impacts the yield of CNFET based circuits. Initially, Monte Carlo simulations are used to evaluate the impact of metallic tubes on the performance, power, and yield of CNFET based circuits. A set of novel CNFET configurations are proposed in order to increase the yield of logic gates in the presence of

metallic tubes. Analytical models are developed to accurately estimate the functional yield of logic gates instead of going through the computationally intensive Monte-Carlo simulations.

A yield aware methodology is developed, offering better trade-off among performance, power and yield of CNFET based circuits by using our proposed architecture level solutions. Similarly, for ASIC design style, an implementation of circuits with regular logic bricks composed of hybrid configurations of transistors are proposed. The proposed configurations allow the designers to obtain optimal trade-off between performance, power and yield.

1.3.3   When the Percentage of Metallic Tubes is Higher

The circuit and architecture level solutions proposed in the previous sub-section to mitigate the impact of the presence of unwanted metallic tubes is not sufficient when the percentage of metallic tubes produced by the CNT growth process is higher(>7%). In this case, researchers have   proposed to remove the metallic tubes by extra processing techniques such as Selective Chemical Etching (SCE) or VLSI-compatible metallic carbon nanotube removal (VMR). The trade-off of these extra processing techniques is that in addition of removing metallic tubes, they also remove a finite portion of required semiconducting tubes. Monte-Carlo simulations are used to obtain the impact of tube removal processing techniques on the performance, power and yield of logic gates. We have derived analytical expressions for the quick estimation of the impact of tubes removed on the yield of logic gates. We propose an efficient Tube Level Redundancy (TLR) technique which reduces the impact of tube removal, and helps in achieving high yield.

1.4   Organization of the Dissertation

The rest of the dissertation is organized as follows:

Chapter 2 provides a formal discussion of CNFETs, different CNTs growth methods, and the fabrication flow of CNFETs. The chapter also highlights the different challenges faced by CNFETs such as spacing, variation in diameter and spacing, misalignment of CNTs, Schottky Barrier contacts between nanotube and metal junctions, the unwanted growth of metallic tubes, and their impact on the CNFET based circuits.

Chapter 3 introduces the CNFET device modeling, estimation of performance, power and area and of CNFET based circuits and their functional yield.

Chapter 4 focuses on the impact of diameter, and spacing among adjacent CNTs, and variations in the diameter and spacing on the performance of parallel tube CNFETs. Furthermore, a novel methodology is presented to stochastically estimate the impact of spacing among adjacent tubes because of the removal of tubes on CNFETs with different drive strengths. In this chapter we propose a set of strategies for variation-tolerant CNFET based circuit design.

Chapter 5 provides the circuit level solutions to address the challenges due to the presence of unwanted growth of metallic tubes. Different transistor configurations are proposed for CNFETs, and the yield results of different configurations of CNFET based logic gates are presented. In addition, analytical models are developed to quickly estimate the yield of logic gates. Finally, architecture level solutions are presented using our proposed set of transistor

level configurations in order to obtain a better trade-off among delay, power, and yield in the presence of metallic tubes.

Chapter 6 contains the analysis for the yield of CNFET based circuits when the metallic tubes are removed by extra processing techniques such as SCE, and VMR. Yield results of logic gates are obtained from Monte Carlo simulations by considering the impact of tubes removed from the drive and fanout gates. Analytical models are developed to estimate the functional yield of gates when a large number of tubes are removed. In this chapter we also report an efficient Tube Level Redundancy technique to increase the functional yield of CNFET based circuits when a large number of tubes are removed. The analysis also shows the impact of path depth on the yield.

Chapter 7 summarizes this work while also suggesting future work.

2   CNFETs, Advantages and Fabrication Challenges

The excellent electronic properties of CNFETs make them a potential candidate of future integrated circuits. The major difference between a CNFET and Si CMOS is the channel material, which in the case of a CNFET, is a single-walled carbon nanotube (SWCNT).

2.1   Carbon Nanotube

SWCNTs are hollow cylinders in which carbon atoms are arranged in the honeycomb lattice [41] as shown in Figure 2-1, and were first demonstrated by Bethune [42] and Iijima [43] in 1993. For the purpose of visualization, SWCNTs are obtained by rolling a sheet of graphene. The band structure of SWCNT can be defined by the chiral vector as given by:

$$C = na_1 + ma_2$$
<div align="right">(2.1)</div>

Here n and m are integers that specify the chirality of the tube, and a_1 and a_2 are the unit vectors of the graphene lattice. Figure 2-1 shows the pictorial representation of the chiral vectors of a SWCNT. The values of n and m determine the characteristic of the carbon nanotube i.e., metallic or semiconducting. It is observed that *a)* when *n=m* the carbon nanotube is metallic, and *b)* when *n-m=3i*, where *i* is an integer, the carbon nanotube is semiconducting with a *small bandgap* [44], and *c)* when n-m≠3i then CNTs are semiconducting with a *large bandgap* [45].
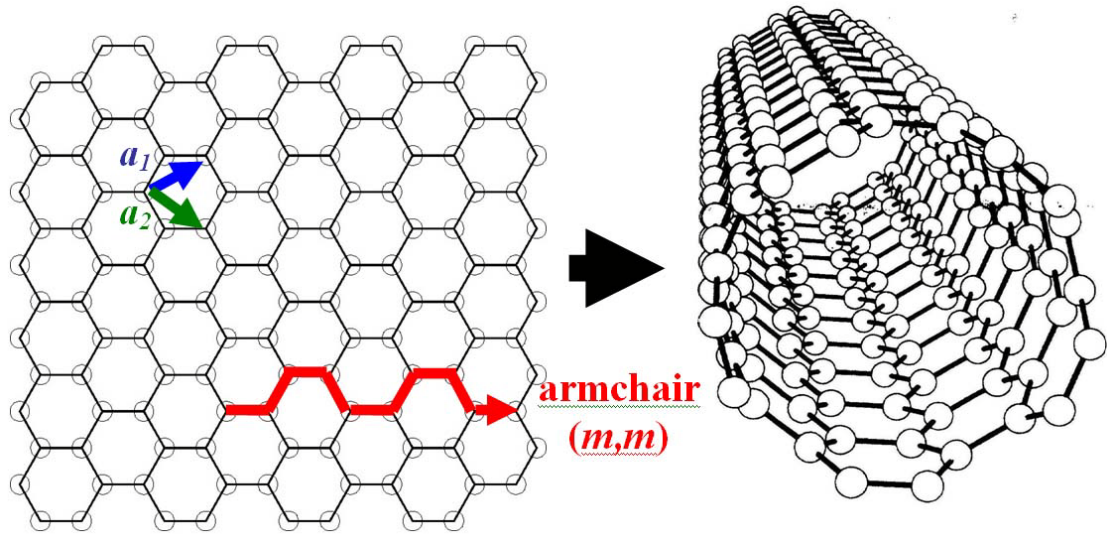
Figure 2-1: Pictorial representation of a single-walled carbon nanotube (SWCNT) with chiral vector.

Figure 2-2 shows the schematic of a metallic carbon nanotube (armchair) and semiconducting (zigzag) SWCNTs. Both metallic and semiconducting CNTs have found many applications in Nanoelectronics. Ballistic transport of carriers can be achieved in single-walled carbon nanotubes because of their quasi 1-D structure which restricts the movement of carriers only along the axis of the tube. This eliminates the wide angle scatterings of carriers and results in a ballistic transport of carriers. The 1-D structure also restricts the wave vector K_C to certain values by fulfilling the condition $K_C.C=2j$. Where j is a constant and can take only integer values. Therefore, each band of graphene can split into a number of 1-D sub-bands.

Armchair: metallic

Zigzag: semi-conducting
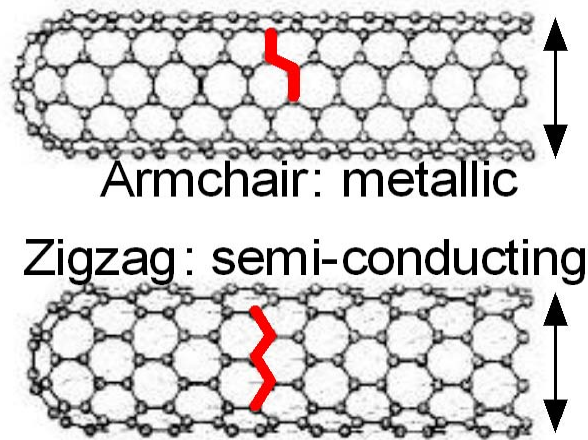
Figure 2-2: A single-walled carbon nanotube (SWCNT) obtained by rolling a sheet of graphite. Depending upon the angle with which the sheet of graphite is rolled metallic (upper) or a semiconducting (lower) CNT is obtained.

2.2 Applications of Carbon Nanotubes in Nanoelectronics

As described in the previous section, the 1-D structure of metallic carbon nanotubes allows the electrons to travel without scattering for longer distances. The mean free path of metallic CNTs is estimated to be 1000nm [41], much longer than 40nm obtained for copper interconnects (which is 25X larger than copper interconnects) at room temperature. Moreover, the metallic carbon nanotubes current carrying capacity are almost 10^{10} (A/cm^2) [46] which is several orders of magnitude larger than the current carrying capacity of copper interconnects. These potential advantages of metallic carbon nanotubes make them a suitable candidate for future interconnects as well as vertical vias. In 2008, Wong et al. [47] demonstrated an integrated circuit in which SWCNT was used as an interconnect. The circuit operated at a frequency of greater than 1GHz.

On-chip capacitors are required by certain analog circuits, and for decoupling purposes in digital circuits. Current integrated circuit technology uses metal-insulator-metal (MIM) and MOS capacitors as decoupling capacitors in the integrated circuits. However, the major

18

problem with these capacitors is the small capacitance per unit area. Carbon nanotubes, because of their low resistivity at nano-scale dimensions, make them a potential candidate to be used as integrated capacitors in future integrated circuits [48]. Researchers have shown that the use of CNT based integrated capacitors results in a significant increase in capacitance per unit area, and larger quality factors than capacitors fabricated with MIM and MOS capacitors [49], [50].

Carbon nanotubes can also be used as on-chip inductors [51], [52] because of their smaller footprint, higher drive current and smaller curvatures. Recent research works have shown promising results for the use of CNTs as passive inductors in Low Noise Amplifiers (LNA) [53].

Carbon nanotubes have excellent mechanical properties, in addition to the excellent electronic properties. Their strong mechanical strength makes them potential candidates for being used in the fabrication of flexible electronics. Various groups have reported that the fabrication of CNFETs [54] and CNFET based circuits [55] on flexible substrates with performance ranging from 40MHz-6GHz.

2.3   Carbon Nanotube Growth Methods

Different methods have been used by researchers for the growth of SWCNTs such as arc discharge, laser ablation and chemical vapor deposition (CVD). Out of the different CNT fabrication methods, CVD produces the most cleanly and untangled tubes, in addition the process of CVD is compatible with the present IC fabrication process. Therefore, in this work we are assuming that CNTs are produced by CVD.

CNTs can be fabricated first, and then deposited on the substrate, on which CNFETs are later fabricated [56]. Kocabas et al. [57] demonstrated the growth of carbon nanotubes onto single-crystal substrates of sapphire or quartz which are later transferred on a plastic substrate for the fabrication of CNFETs.

CNTs can also be fabricated at desired locations on the substrate on which CNFETs are later fabricated. Kong et al. demonstrated the growth of SWCNTs on SiO_2/Si wafers [58].

Figure 2-3 shows the setup used to fabricate CNTs using CVD. In this process, a catalyst material (Fe, Co, Pt) is heated in the furnace in the presence of hydrocarbon gas. The reaction of the hydrocarbon gas with catalyst material results in the growth of CNTs which are only grown on the places where the catalyst particles are deposited, therefore no further cleaning or detangling action is required.
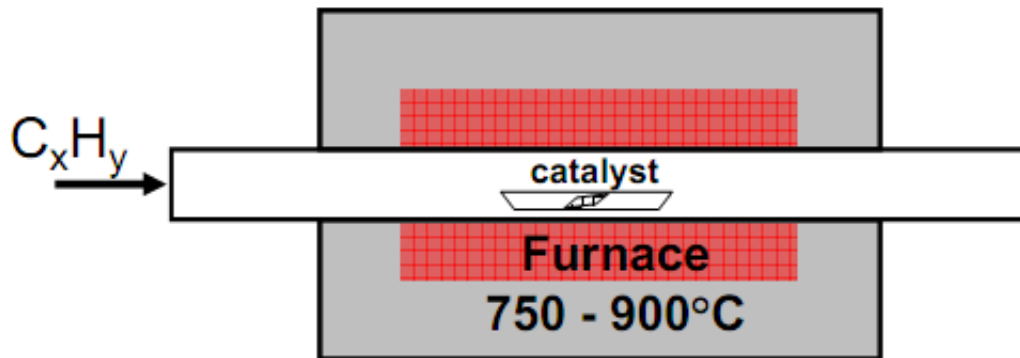
Figure 2-3: Setup used to fabricate CNTs using chemical vapor deposition (CVD) [59].

2.4   Carbon Nanotube Field Effect Transistors

In the previous section, CNT synthesis techniques were described. In this section, we focus on the different types of CNFETs using carbon nanotubes. Two main types of CNFETs

are being explored by researchers are, the Schottky Barrier (SB) CNFETs [60], and the MOSFET-type CNFETs [61].Figure 2-4(a) shows the device structure and Figure 2-5(a) shows the conduction band profile of a SB transistor. Similarly, Figure 2-4(b) shows the device structure and Figure 2-5(b) shows the conduction band profile of MOSFET-type of CNFETs. In a SB CNFET, the gate voltage controls the width of the Schottky Barrier at the source end of the channel as shown in Figure 2-5(a); therefore, the presence of the tunneling barrier at the source side of the channel controls the *ON* current of the SB CNFET. SB devices exhibit ambipolar conduction. Therefore both the n-type and p-type CNFETs can be obtained by a proper selection of the work function in metals for source/drain contacts.
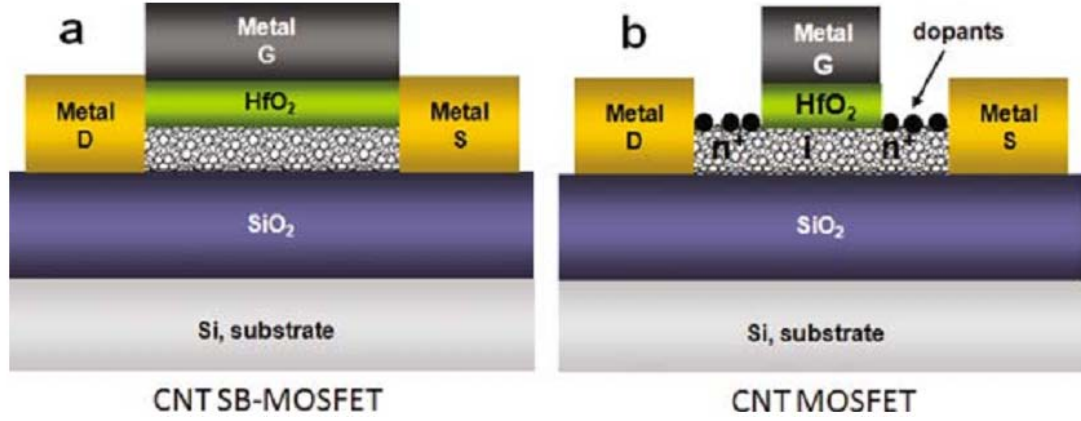
Figure 2-4: Device structure of a (a) Schottky Barrier (SB) CNFET, (b) MOSFET-type of CNFET [59].
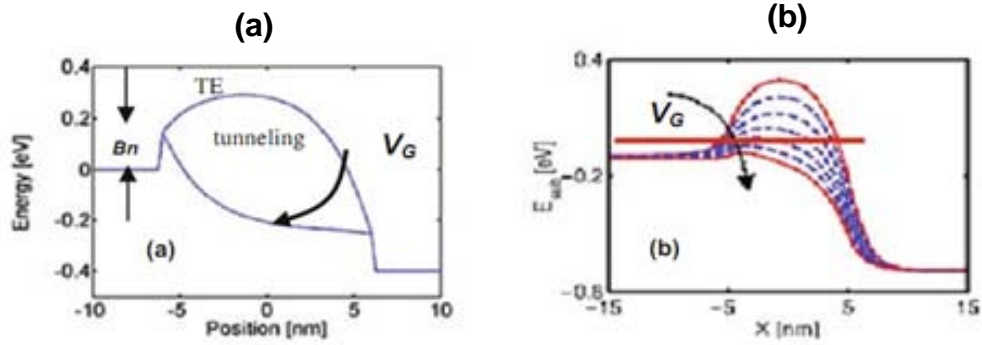
Figure 2-5: Conduction band profile of (a) Schottky Barrier (SB) CNFET (b) MOSFET-type of CNFET [59].

Figure 2-4(b) shows the schematic of a MOSFET-type N-CNFET in which source and drain regions are chemically doped with potassium (K). Similarly, the MOSFET-type of P-CNFETs are demonstrated by Chen et al. [32], [62], [63] in which the source and drain regions of the transistor are doped with tri-ethyloxonium hexachloroantimonate $(C_2H_5)3O+SbCl_6$ (OA). In the MOSFET-type of CNFETs, the conductance of the channel is controlled by the gate voltage as shown in Figure 2-5(b). The doping of source and drain

regions in MOSFET-type CNFETs suppresses the transport of either the electrons or the holes, and hence results in unipolar conduction characteristics.

The main problem with SB transistors is that the formation of SB between the CNT and source/drain contact results in a large subthreshold slope and ambipolar conduction in nanoscale devices. The large subthreshold slope and ambipolar conduction severely limits the *ON* current, and exponentially increase the *OFF* current, both of which are unacceptable in high performance and low power digital applications. Therefore, doping methods and using different doping materials (as described in the previous paragraph) are used to fabricate MOSFET-type of CNFETs with small subthreshold slopes and uni-polar conduction characteristics. Because of the high performance and the low *OFF* current of MOSFET-type of CNFETs, the focus of this work is on the MOSFET-type of CNFETs. For the sake of simplicity we will refer the MOSFET-type of CNFETs in the rest of the thesis as CNFETs.

2.5   Fabrication Flow of CNFETs

Figure 2-6 shows the sample fabrication process flow of MOSFET-type of CNFET. Here, first SiO_2 is thermally grown on the Si wafers as shown in Figure 2-6(a). Then lithographically defined alignment markers are patterned on regions as shown in Figure 2-6(b), where CNTs are later grown. Afterward, windows are opened in the photo resist to deposit the catalyst (Fe, Co, Pt) on the substrate, as shown in Figure 2-6(c). The catalyst is deposited either in the form of liquid drops or thin layers of metal catalyst films at specific locations on the substrate as shown in Figure 2-6(d), and then the photo resist is etched

away as shown in Figure 2-6(e).Then SWCNTs are synthesized on catalytically patterned areas of the Si/SiO$_2$ substrate by CVD as shown in Figure 2-6(f).

After the growth of CNTs, source/drain contacts are patterned by either using photolithography or e-beam lithography. In case of e-beam lithography, the metal films with a thickness of ~7-30nm are deposited, as shown in Figure 2-6(g). Palladium (Pd) is used as the source and drain contacts for both n-type [61] and p-type CNFETs [62]. The top gate stack consisting of high-k dielectric (HfO$_2$, ZrO$_2$), and the metal gate is fabricated by atomic layer deposition (ALD) and the lift-off technique without overlapping the metal source and drain contacts [64]. Figure 2-6(h) shows the patterned high-k dielectric and metal gate using ALD and lift-off. During this process, the segments of nanotubes where doping is performed are remained fully exposed as it can be observed in Figure 2-6(h). The source and drain regions between the gate stack and metal source and drain contacts are doped as depicted in Figure 2-6(i). For n-type CNFETs the source and drain regions are exposed to Potassium (K) vapor in vacuum [61] and for p-type CNFETs the source and drain regions are exposed to tri-ethyloxonium hexachloroantimonate (C2H5)3O+SbCl6 (OA) [65] .

(a)

SiO₂
Si

(f)

CNT
SiO₂
Si

(b)

Alignment
marker
SiO₂
Si

(g)

m   m
SiO₂
Si

(c)

Photo
resist
SiO₂
Si

(h)

Gate
High K
Pd   Pd
SiO₂
Si

(d)

Catalyst
(Fe,Co,Pt)   Photo
resist
SiO₂
Si

(i)

D   Gate   S
High K
Pd   Pd
SiO₂
Si

(e)
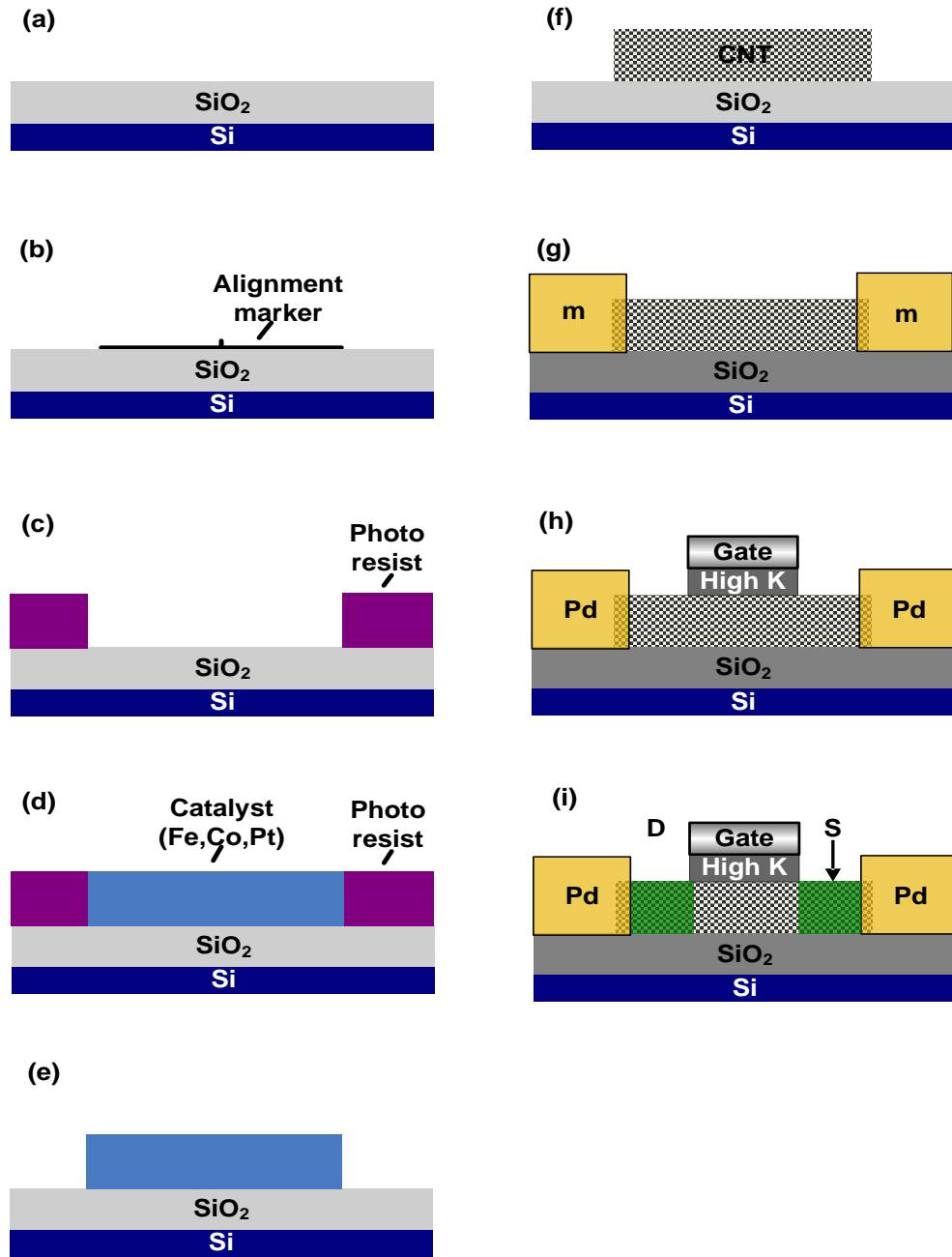
SiO₂
Si

Figure 2-6: Sample CNFET fabrication flow (a) Thermal growth of SiO₂ on Si wafer (b) Patterning of alignment markers (c) Opening of windows in the photo resist (d) Deposition of catalyst resist (e) Etching of photo resist (f) CNT grown by chemical vapor deposition (g) Fabrication of metallic electrodes (h) Formation of top gate stack consisting of high-k gate dielectric and metal gate (i) Doping of CNTs.

2.6 Advantages of CNFETs

Single-walled CNFETs are promising candidates for future integrated circuits [66], [41] because of their excellent properties, like long scattering mean free path(MFP) $>1\mu m$[65], resulting in near ballistic transport [30], high carrier mobilities ($10^3\sim10^4 cm^2/Vs$) in semiconducting CNTs [31], and the easy integration of high-k dielectric material such as HfO_2[32], or ZrO_2[67] resulting in better gate electrostatics. Because of the aforementioned properties, CNFETs have a potential to deliver higher performance and lower power as compared to FETs built in silicon technology [68], [69]. The theoretical analysis results show that CNFETs is thirteen times faster than a PMOS transistor and six times faster than an NMOS transistor [70] using 32nm technology node.

2.7 Challenges Faced by CNFETs

Since the first demonstration of carbon nanotube field effect transistors by researchers at Delft University [56], [71] in 1998, tremendous progress has been made in CNT based technology. However, fabrication of CNFET-based circuits still faces major challenges which are needed to be solved for making the CNFET technology commercially viable. These challenges are as follows:

1. Variation in the diameter of CNTs [72]

2. Packing Density of CNTs [73]

3. Spacing and variation in spacing among adjacent CNTs

4. Misalignment of CNTs [74], [57]

5. Schottky Barrier contact between source and drain and CNT

6. Unwanted growth of metallic CNTs

2.7.1 Variation in the Diameter of CNTs

The diameter of the SWCNT can be approximated as

$$D_{CNT} = \frac{\sqrt{3}a_0}{\pi}\sqrt{n^2 + m^2 + nm} \tag{2.2}$$

Where a_0=0.142nm is the carbon to carbon atom distance. In case of semiconducting nanotubes, the bandgap of the CNT is inversely proportional to the diameter of the carbon nanotube. The relationship between the bandgap and diameter of a CNT is given as

$$E_G \propto \frac{0.8ev}{d\ (nm)} \tag{2.3}$$

Where E_G is the bandgap of the CNT and d is the diameter of the carbon nanotube in nm. This means that tubes with smaller diameters have a larger bandgap, and tubes with larger diameters have a smaller bandgap. The fabrication of CNTs results in the variation in the diameter of the tubes where normally fabricated CNTs have diameters within 1nm to 2nm [75]. Furthermore, experimental results show that the diameter of CNTs shows a Gaussian distribution [76]. Figure 2-7 shows the diameter distribution of CNTs with μ and 3σ diameter of 1.5nm and 0.5nm respectively.
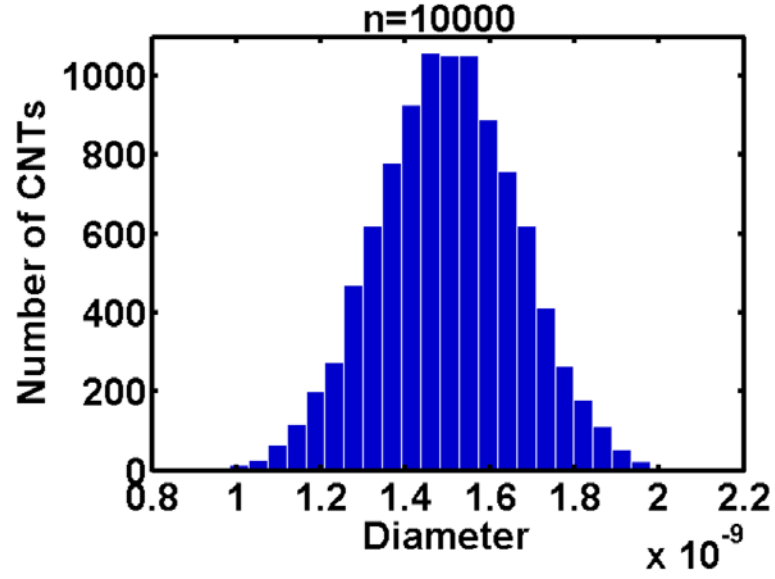
Figure 2-7: Diameter distribution of CNTs with respect to μ and 3σ CNT diameter of 1.5nm and 0.5nm.

In [77] the authors developed the first order model to approximate the threshold voltage of CNFET as a function of the diameter of the CNT.

$$V_{th} = \frac{\sqrt{3}a_0 V_\pi}{3ed} \tag{2.4}$$

Here V_π is the carbon bonding energy, e is the charge on the electron, and d is the diameter of the carbon nanotube. Please note the inverse dependence of the threshold voltage on tube diameter, as a result Therefore CNFETs with tubes of smaller diameter have both *ON* and *OFF* currents smaller. The *ON* current of a semiconducting tube in a CNFET can be expressed by the equation (2.5) taken from [78], in which g_{CNT} represents the transconductance of a CNFET. V_{DD} is the supply voltage and according to ITRS guidelines, it is projected at 900mV [79] for 32 nm technology node. L_S is the length of a

doped CNT which acts as a source region, and ρ_s is the resistance per unit length of the source region.

$$I_{ons} = \frac{g_{CNT}(V_{DD} - V_{th})}{1 + g_{CNT} L_S \rho_S} \qquad (2.5)$$

From equations (2.4) and (2.5) , it can be observed that there is a linear dependence of the ON current on the diameter of CNTs. The OFF current (I_{offs}) of a semiconducting tube has the exponential dependence on the threshold voltage, and on the subthreshold slope of the device. Equation (2.6) is the approximation used to obtain the OFF current of a semiconducting nanotube, where $I_{on\mu}$ is the mean value of the ON current of a semiconducting tube. A fitting parameter "r" is used to obtain the desired ratio of I_{on}/I_{off} ratio. S is the subthreshold of the device and its value varies between 63mV/decade [80] to 100mV/decade [61]. From equations (2.4) and (2.6) it can be observed that there is an exponential dependence of CNFETs diameter on the OFF current of the transistor.

$$I_{offs} = \frac{I_{ON\mu}}{r}(10^{-Vth/S}) \qquad (2.6)$$

2.7.2 Packing Density of CNTs

Single-tube CNFET's are not very feasible for circuit applications because of their low drive currents and small active areas. To produce scalable devices, an array of densely packed CNTs is considered as a possible solution, resulting in multiple parallel transport paths that can deliver large drive currents.

A CNFET using an array of tubes has been demonstrated in [65], [81]. Present CNT synthesis technologies allow us to pack almost 10-50 CNTs/µm [81], [82]. However, Patil et al. [78] analyzed that in order to obtain delay and energy gains over Si-CMOS with future technology nodes, almost 250 CNTs/µm are required corresponding to spacing (S) between adjacent tubes of 2.5nm for a tube diameter of 1.5nm. The spacing between adjacent tubes in parallel tube CNFETs impacts the channel capacitance due to *charge screening effects* from adjacent tubes, thus impacting the current delivered by individual CNTs. The packing of almost 250 CNTs/µm also gives us the optimal number of CNTs as further increase in the density of tubes results in a reduction in the drive current from the parallel tube CNFETs.

Deng et al.[83] showed a reduction of almost 2X in the *ON* current of a parallel tube CNFET when the spacing between adjacent tubes is reduced to 1nm. On the one hand, the increase in the number of parallel tubes in the channel improves the drive current of the transistor because of the increase in the number of conducting channels. On the other hand, there will be a reduction in the drive current of parallel CNTs due to the increase in charge screening because of the reduction in the spacing between adjacent tubes.

2.7.3   Spacing and Spacing Variation

The drive current of a parallel tube CNFET depends upon the gate to channel capacitance. The parallel tubes in the CNFET have screening effects on the potential profile in the gate region and therefore effects the overall gate to channel capacitance of the parallel tube CNFET[84]. The amount of screening from adjacent tubes in parallel tube CNFETs is a function of the spacing between adjacent CNTs. The spacing between adjacent tubes is

inversely proportional to the gate to channel capacitance. Therefore, less spacing between adjacent CNTs decreases the channel capacitance which implies a reduction in the drive strength of parallel tube CNFETs. Moreover, for fixed width CNFETs, the variation in the spacing between adjacent tubes can also result in variation in the density of CNTs. The variation in the charge screening because of variation in the spacing and in the density of CNTs, results in a large variation in the drive current of CNFETs, which will be presented in detail in Chapter 4.

2.7.4   Misalignment of CNTs

The lack of precise control on the positioning of CNTs during the fabrication of CNFETs can result in a misalignment of the tubes [74], [57]. Significant progress has been made in the fabrication of aligned CNTs, and less than 0.5% of CNTs fabricated on the single-crystal quartz substrate are misaligned [81]. The misaligned tubes can cause either a short between the output and the supply rail, or an incorrect logic function. Figure 2-8(a) shows a NAND cell in which the misaligned tube causes a short between the V_{DD} and output because the entire CNT is a doped p-type. Similarly, Figure 2-8(b) shows the layout of the gate in which the misalignment of the tube results in the incorrect logic functionality of the gate. Therefore, even less than 0.5% of misaligned tubes can have a significant impact on the yield of CNFET based circuits.
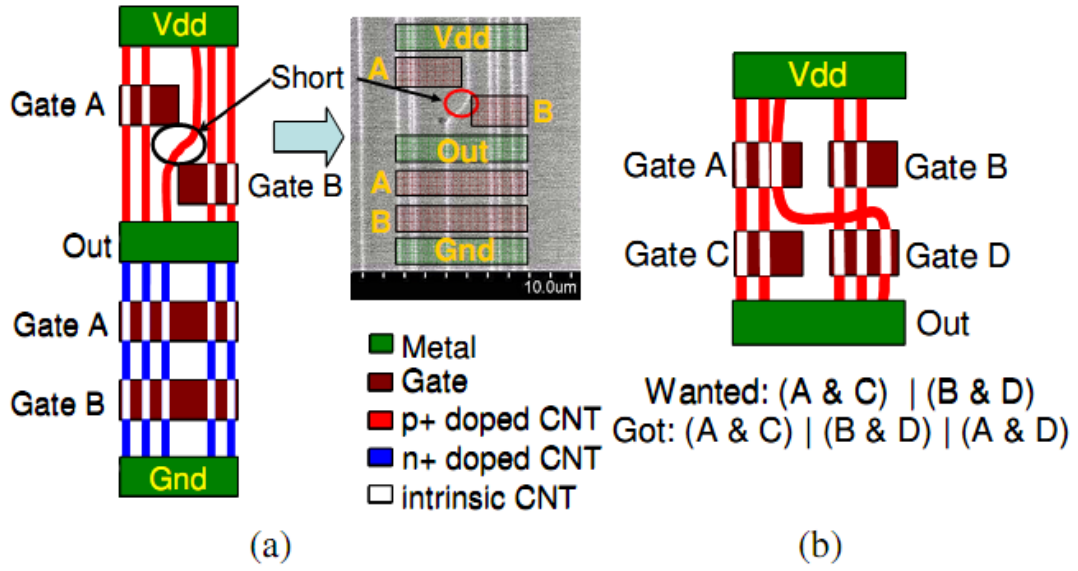
Figure 2-8: (a) Short inside NAND gate caused by misaligned CNT (b) Incorrect logic function due to misaligned CNT [85].

2.7.5   Schottky Barrier Contact

The interface between the carbon nanotubes and metals that are used as source/drain of a

CNFET forms a Schottky Barrier (SB). The formation of these energy barriers for injection

of electrons and holes due to Schottky contacts are reported by [60], [86-88]. The height of

the SB strongly depends upon the work function of the metal and the annealing conditions

used during the fabrication of CNFETs [14], [89]. The SBs at the source and drain side of

transistors results in a significant reduction in the drain current in the transistors.

Therefore, for a high performance operation of the CNFET devices, suitable metals are

required, which can be used as source and drain contacts and also provide ohmic source

and drain contacts.

2.7.6   Unwanted Growth of Metallic Tubes

To use CNTs as the channel material, semiconducting CNTs are required. Depending on

the chirality, a SWCNT can be either metallic or semiconducting. At present, there is no

CNT synthesis technique that can produce 100% semiconducting tubes. The percentage of semiconducting and metallic tubes obtained by different CNT fabrications techniques is shown in Table 2-1.In HiPco method, the SWCNTs are grown by thermal decomposition of catalyst $Fe(CO)_5$ in the heated flow of CO at temperatures of 800^0C to 1000^0C. The fabricated tubes result in almost 68% semiconducting tubes. In a Plasma-Enhanced CVD, catalyst Fe is heated in the presence of CH_4 gas at 600^0C resulting in almost 90% semiconducting tubes. In fast heating using the PECVD method, the SWCNTs are grown by heating the catalyst Fe to 750^0C in the presence of C_2H_2 gas. In all of the above three mentioned methods, the CNTs are grown on SiO_2/Si wafers.

Table 2-1: Percentage of semiconducting tubes produced by different CNT synthesis processes.

CNT synthesis process	Semiconducting CNTs (%)
HiPco with CO gas[90]	61±7.6
Plasma-Enhanced CVD with CH4[91]	89.3±2.3
Fast heating with PECVD[92]	96%

In the case of metallic tubes, the gate terminal has no control over the channel due to an ohmic short between the source and drain. Therefore, the presence of metallic tubes in complementary CNFET circuits has a dramatic impact on static current, static noise margin delay, and yield of CNT based circuits.

2.8   Extra Processing Steps to Remove Unwanted Metallic CNTs

In case of a large percentage of metallic tubes, other processing techniques are required to remove the metallic tubes. The main techniques proposed by researchers to remove the metallic tubes are:

- Current-Induced Electrical Burning [93].

- Selective Chemical Etching(SCE) [94]

- VLSI-Compatible Metallic Carbon Nanotube Removal (VMR)[95]

2.8.1   Current-Induced Electrical Burning

The conductance of a metallic CNT is *independent* of the gate voltage, whereas the conductance of a semiconducting tube *depends* upon the gate voltage. Therefore, semiconducting tubes charge carriers can be depleted by applying the appropriate gate voltage. The independence of conductance of metallic tubes from the gate voltage is used by [93] to eliminate the metallic tubes in an ensemble of metallic and semiconducting tubes. In the electrical burning technique, a high voltage is applied at the gate and across the source and drain side of the CNFET consisting of multiple parallel CNTs. The voltage applied at the gate is such that it reverse biases the transistor. For example, a positive voltage is applied at the gate of PMOS device, depleting the semiconducting CNTs of carriers, and no current will flow through the semiconducting CNTs in the presence of the voltage across the source and drain terminals. On the other hand as the conductance of metallic tubes is independent of gate voltage, a high bias across the source and drain terminals results in a large current to flow through the metallic tubes. A sufficient large current breaks down the metallic tubes electrically. The electrical burning technique can remove almost all of the metallic tubes, but faces some major limitations. First, it requires a high gate voltage (~10V), and because of reliability concerns, a thick gate oxide will be required. The thick gate oxide will reduce the performance of CNT based circuits. Second, the electrical burning technique requires a contact with each individual transistor which is not scalable and therefore not suitable for ultra large scale VLSI systems. And third, in

complex logic gates, internal contacts are not accessible and as a result some metallic tubes will not be removed, causing detrimental power and performance impacts.

2.8.2   Selective Chemical Etching

The Selective Chemical Etching (SCE) technique, proposed by Zhang et al. [94], selectively etches and gasifies the metallic nanotubes. The main advantage of SCE is that it is scalable, and can be applied to future ultra large scale integrated circuits. In this technique, an ensemble of metallic and semiconducting tubes on the substrate are subjected to methane plasma, followed by an annealing process, hydrocarbonating the tubes depending on their cutoff diameters. The cutoff diameters are different for metallic (D_{CM}) and semiconducting tubes (D_{CS}). Therefore, depending on the diameter range of CNTs, it may not completely remove all metallic tubes, while it can remove some of the needed semiconducting tubes.

Table 2-2 shows the percentages of tubes removed for different percentages of metallic tubes present, ranging from 31% to almost 50%, when the percentage of metallic tubes varies from 5% to 30%. CNTs have a Gaussian diameter distribution with μ of 1.5nm and 3σ of 0.5nm. The tubes which are removed by selective etching process are represented by making their *ON* and *OFF* current equal to zero.

Table 2-2: Percentage of CNTs (metallic and semiconducting) removed by the Selective Chemical Etching (SCE) process.

P_m	% of metallic CNTs Removed		% of semiconducting CNTs removed	% of total CNTs removed	% of metallic CNTs remaining
	d <1.4nm	1.4≤d ≤2nm	d <1.4nm		
5%	1.38	3.63	25.92	30.93	0.06
10%	2.76	7.28	24.54	34.58	0.02
15%	4.19	10.79	23.11	38.09	0.02
20%	5.46	14.56	21.84	41.86	0.03
25%	6.88	18.05	20.42	45.35	0.02
30%	8.13	21.87	19.18	49.17	0.04

2.8.3   VLSI-Compatible Metallic Carbon Nanotube Removal (VMR)

Recently, Patil et al. [95] presented a VLSI-compatible metallic CNT removal technique, which is an extension of the current-induced electrical burning technique. The main advantages of the VMR technique are that it is scalable, and is compatible with Ultra Large Scale Integrated Circuit (ULSI) processing. In VMR, first a special inter-digitated electrode structure is applied with minimal metal pitch, and then a high voltage at the back-gate is applied to turn off the semiconducting CNTs all at once. After that, a high voltage is applied on the supply lines which results in high current to flow through the metallic tubes and electrically breaks down the unwanted metallic tubes. Based on the final design, unwanted areas of CNTs, and unwanted sections of electrodes are etched away. This technique eliminates almost all the metallic tubes. The trade-off of using VMR is that, it also removes some of the required semiconducting tubes, similar to the SCE technique.

2.9   Other Work on CNFETs

To make the CNT based technology commercially viable, we need to overcome the challenges faced by the CNT technology. The details of CNT manufacturing challenges were discussed in the previous section. Significant work has been done by researchers in terms of analyzing the impact of fabrication imperfections of the performance and energy of CNFET based circuits, and various solutions have been proposed to overcome the challenges faced by CNT technology.

Javey et al. demonstrated CNFETs with doped source and drain regions and high-K gate dielectrics [32], [61]. CNFETs with a perfect array(s) of aligned CNTs have been demonstrated by different research groups [65], [81], [82], [96]. CNFET based integrated circuits are reported by [55], [36], [97]. Researchers have demonstrated CNFETs with ohmic contact between CNTs and source and drain contacts [14], [61]. [32] reported that the use of Palladium(Pd) results in ohmic contact between the Pd electrode and valence band of CNT in a p-type CNFET. Similarly, [98] reported that the use of Scandium (Sc) results in ohmic contact between the Sc electrode and conduction band of CNT of an n-type CNFET.

The impact of diameter and density variation of CNTs on the performance of CNFETs is analyzed by [99], [70], [73]. Processing techniques, such as Current-Induced Electrical Burning [93], Selective Chemical Etching(SCE) [94], and VLSI-Compatible Metallic Carbon Nanotube Removal(VMR) [95] have been proposed to eliminate the unwanted metallic tubes. Patil et al. proposed a design technique to design circuits that function correctly even

in the presence of misaligned tubes [85], [100] with minimal overhead in terms of area and performance.

2.10 Conclusions

The high mobility and easy integration of high-K dielectrics with CNT based technology makes them a potential candidate device in the post silicon era. However, the challenges faced by CNT technology which are discussed in this chapter make it difficult to fabricate large scale CNFET based circuits. Solutions to some of these challenges are addressed by researchers, while one of the major challenges of handling the unwanted growth of metallic tubes is presented in this work.

3   Delay, Power, Area, and Yield Modeling

CNFET based circuits with parallel tubes have the potential to give 2X to 10X higher performance, 7X to 2X lower energy consumption per cycle, and 15X to 20X lower energy delay product as compared to silicon CMOS circuits[101]. However, fabrication challenges associated with CNTs adversely impact the performance and power of CNFET based circuits. To evaluate the power and performance of a CNFET based circuits, models are required, which allows to analyze the impact of the fabrication imperfections on the design parameters. Since these fabrication imperfections are unique to CNFET based devices, innovative models are required to be developed to estimate these fabrication imperfections. In this chapter, we model the delay/performance and power consumption of CNFET based logic gates, and estimate the functional yield of gates in the presence of fabrication imperfections.

3.1   CNFET Device Modeling

To find the *ON* and *OFF* currents of CNFETs, we used a circuit compatible model developed by [83], [101] . The model considers the practical device non idealities, such as quantum confinement effects, acoustic/optical phonon scattering, elastic scattering, resistance of the source and drain, the resistance of Schottky Barrier, and parasitic gate capacitance for the computation of current vs. voltages of CNFET based circuits. The circuit compatible model allows simulating CNFET-based circuits with multiple parallel tubes as transistor channels and with a large range of tube diameters. The current vs. voltage results obtained from the circuit compatible model [102] are in close agreement with the experimental CNFET data [103].

The equivalent circuit model of the n-type CNFET is shown in the Figure 3-1. Because of the symmetric band structure of the CNT, a p-type CNFET model is similar to n-type CNFET, only when the polarity of the voltages are required to be changed. It consists of two current sources, one resistance and four capacitances between the different terminals of the transistor. In Figure 3-1, I_{semi} is the current flowing in the semiconducting CNT because of the existence of semiconducting energy sub-bands and is given by:

$$I_{semi}\left(V_{ch,DS}, V_{ch,GS}\right) \approx \frac{4e^2}{h}\sum_{\substack{k_m \\ m=1}}^{M}T_m\left[V_{ch,DS} + \frac{kT}{e}\ln\left(\frac{1+e^{\left(E_{m,0}-\Delta\Phi_B\right)/kT}}{1+e^{\left(E_{m,0}-\Delta\Phi_B+eV_{ch,DS}\right)/kT}}\right)\right] \quad (3.1)$$

Here $V_{ch,DS}$ and $V_{ch,GS}$ are the Fermi potential difference in the channel near the source side, e is the charge on the electron, h is the Plank's constant, k is the Boltzmann constant, T is the temperature in Kelvin degrees, M is the number of sub-bands, T_m is the transmission probabilities of the carriers, $\Delta\Phi_B$ is the change in the channel surface potential with respect to voltage at the gate and the drain terminals of CNFET. $\Delta\Phi_B$ is obtained from Spice, and its value depends on the diameter of the carbon nanotube and the gate to channel capacitance. $E_{m,0}$ is the carrier energy at the mth subband and 0th sub-state respectively and is obtained from equation (3.2).

$$Em,0 \approx \frac{3a_0V_\pi\lambda}{d} \quad (3.2)$$

λ is a constant, and its value depends on the values of integers n and m as given by

$$\lambda = \begin{cases} \dfrac{6C-3-(-1)^m}{12} & C=1,2,... \quad ,\text{mod(n-m,3)} \neq 0 \\ C & C=0,1... \quad ,\text{mod(n-m,3)=0} \end{cases} \quad (3.3)$$

I_{semi} is the major source of current flowing in the semiconducting tubes.

Similarly the current in metallic nanotubes can be obtained from equation(3.4)

$$I_{metal} \approx (1 - m0)\frac{4e^2}{h}T_{metal}V_{ch,DS}$$

(3.4)

Here m0 is the metallic sub-band. Similarly the band to band tunneling (BTBT) current in the CNT is given as:
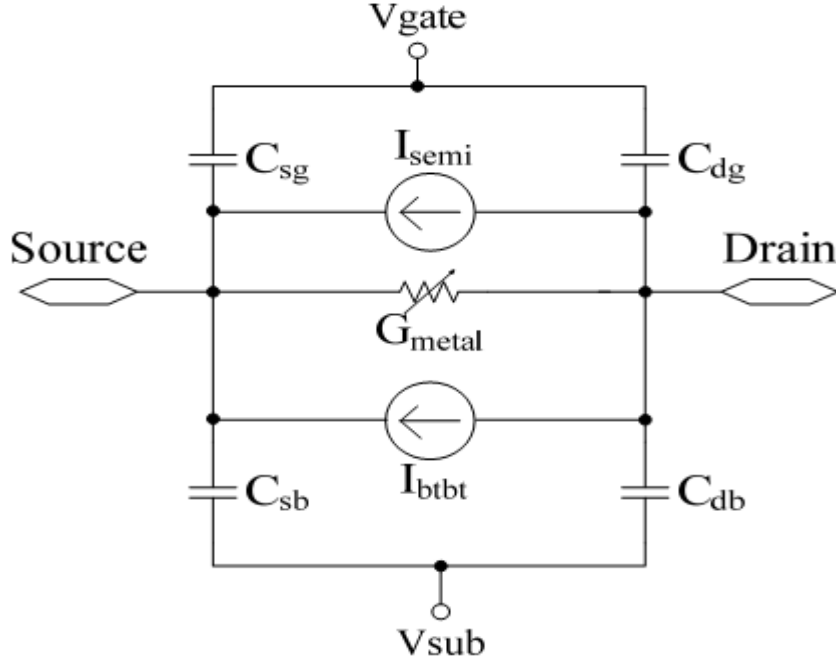
Figure 3-1: Equivalent circuit model of the CNFET [83].

$$I_{btbt} \approx \frac{4e^2}{h}kT\sum_{\substack{m=1 \\ k_m}}^{M}\left[\frac{T_{btbt}\ \ln\left(\frac{1 + e^{\left(eV_{ch,DS} - E_{m,0} - E_f\right)/kT}}{1 + e^{\left(E_{m,0} - E_f\right)/kT}}\right) \times}{\frac{\max\left(eV_{ch,DS} - 2E_{m,0}, 0\right)}{eV_{ch,DS} - 2E_{m,0}}}\right]$$

(3.5)

Here T_{btbt} is the band-to-band tunneling (BTBT) probability, while E_f is the Fermi level of the doped source and drain nanotube in electron volt. In Figure 3-1, C_{Sg} is the capacitance between gate to source terminal of the transistor, C_{dg} is the capacitance between drain and gate terminal, C_{sb} is the capacitance between source and bulk, and C_{db} is the capacitance between the drain and bulk terminal of the transistor.

3.2   Delay

In a complementary CMOS-based circuit that has pull-up and pull-down networks, the delay of logic gate D_g is given by

$$D_g \propto \frac{C_L V_{DD}}{\left(I_{ON} - I_{OFF}\right)} \tag{3.6}$$

The same equation can be used for a complementary circuit build with n-type and p-type CNFETs [104]. Here V_{DD} is the supply voltage, and . I_{ON} and I_{OFF} are the currents flowing in the *ON* and *OFF* network of the gate. In logic gates implemented with transistors that have channels built with parallel tubes, the *ON* current is a function of the number of tubes in the *ON* network, and the *OFF* current, I_{OFF} depends upon the number of tubes in the *OFF* network. In cases where metallic tubes are present, there will be a large current flowing in the *OFF* network of the gate. The procedure to obtain the *ON* and *OFF* currents of multi-channel CNFETs is presented in the next section. The delay of the gate is also a strong function of the correlation among tubes used in the pull-up and pull-down networks. The

impact of correlation between tubes used in the pull-up and pull-down network is presented in Chapters 5 and 6 of this dissertation.
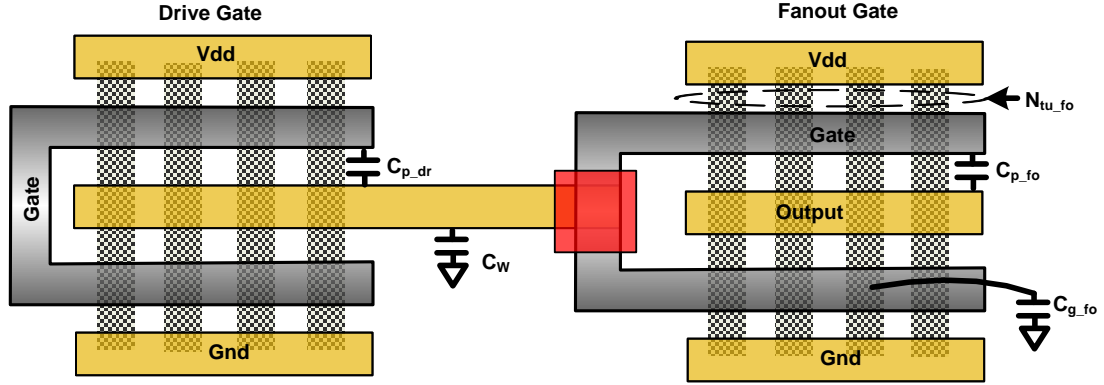
Figure 3-2: Layout of a CNFET based inverter driving another inverter.

In equation(3.6), C_L is the load capacitance of the gate, and is composed of different components as given by:

$$C_L = C_{p\_dr} + C_w + C_{p\_fo} + N_{tu\_fo}C_{g\_fo}$$

(3.7)

Figure 3-2 shows the layout of an inverter driving another inverter as a load. The figure shows the different components of the load capacitance. In equation(3.7), C_{p_dr} is the parasitic capacitance of the driving gate and is a function of the W/L of the driving gate, and C_W is the capacitance of the interconnects. To keep the analysis simple only local interconnects are considered. Therefore, the impact of this capacitance is negligible. C_{p_fo} is the parasitic capacitance of the fanout gate(s) and depends on the size of the fanout gate(s). C_{g_fo} is the gate capacitance of the fanout gate and is a function of the number of tubes in the fanout gate. In cases where extra processing steps are used to remove the metallic tubes,

the removal of tubes results in a large variation in the load capacitance of the gate, resulting in large variation in the performance of the gates.

3.3 Power

CNFET based logic gates are very power efficient compared to logic gates implemented with CMOS technology because of less switching capacitance. However, because of fabrication imperfections, if metallic tubes are present in the parallel tube CNFETs, then a large short circuit current will flow in the *OFF* networks of the logic gates, resulting in large static power consumption in the CNFET based gates. Figure 3-3(a) shows the distribution of *OFF* current of CNFETs when all the tubes in the CNFET are semiconducting and Figure 3-3(b) shows the distribution in *OFF* current when 90% of the tubes are semiconducting and 10% of the tubes are metallic. In Figure 3-3(b) we observe two separate distributions of *OFF* current, one having small *OFF* current that is due to semiconducting tubes, and the other with large *OFF* current due to the presence of metallic tubes. The detailed impact of the presence of unwanted metallic tubes on the static power is presented in Chapter 5 of the dissertation. Similarly, when metallic tubes are removed it will result in a variation in the dynamic power dissipation. This will be addressed in the Chapter 6 of the dissertation.
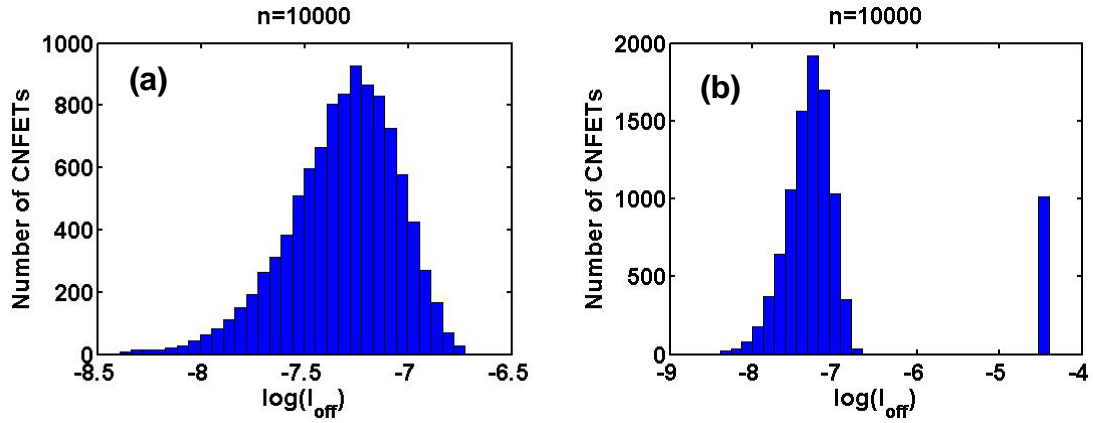
Figure 3-3: I_{OFF} **distribution of CNFETs with respect to** μ **and** 3σ **diameter of 1.5nm and 0.5nm (a) when all the tubes are semiconducting and (b)when 90% of the tubes are semiconducting and 10% tubes are metallic a**

3.4   Area

It is anticipated that circuits implemented with CNTs will be more area efficient as compared with circuits implemented with silicon CMOS. The area advantage for CNT based circuits is due to two reasons. First, in case of CNFETs, the band structure of a CNT is symmetrical for the conduction and valence band, and the same size N-CNFET and P-CNFET results in symmetric performance. Because in Si-CMOS the mobility of holes is lower than the mobility of electrons, for symmetrical performance PMOS devices have to be sized almost 3X larger than NMOS. Second, PMOS devices are implemented in an n-well, and design rules require at least 12λ separation between n-well and NMOS device. For a 32-nm technology node with tube lengths used is 32-nm and with tube density of 250CNTs/μm, it is observed that multi-channel CNFET based gates results in significant area advantages over gates implemented with Si-CMOS. Figure 3-4(a) shows the layout of an inverter implemented with n-type and p-type CNFETs, and Figure 3-4(b) shows the Si-CMOS inverter implemented with NMOS and PMOS transistors. Please note that Figure

45

3-4(a) and Figure 3-4(b) are not drawn to scale. Almost 6X and 3X improvement in area is observed for inverter and 2-input NAND gate implemented using parallel tube CNFETs.
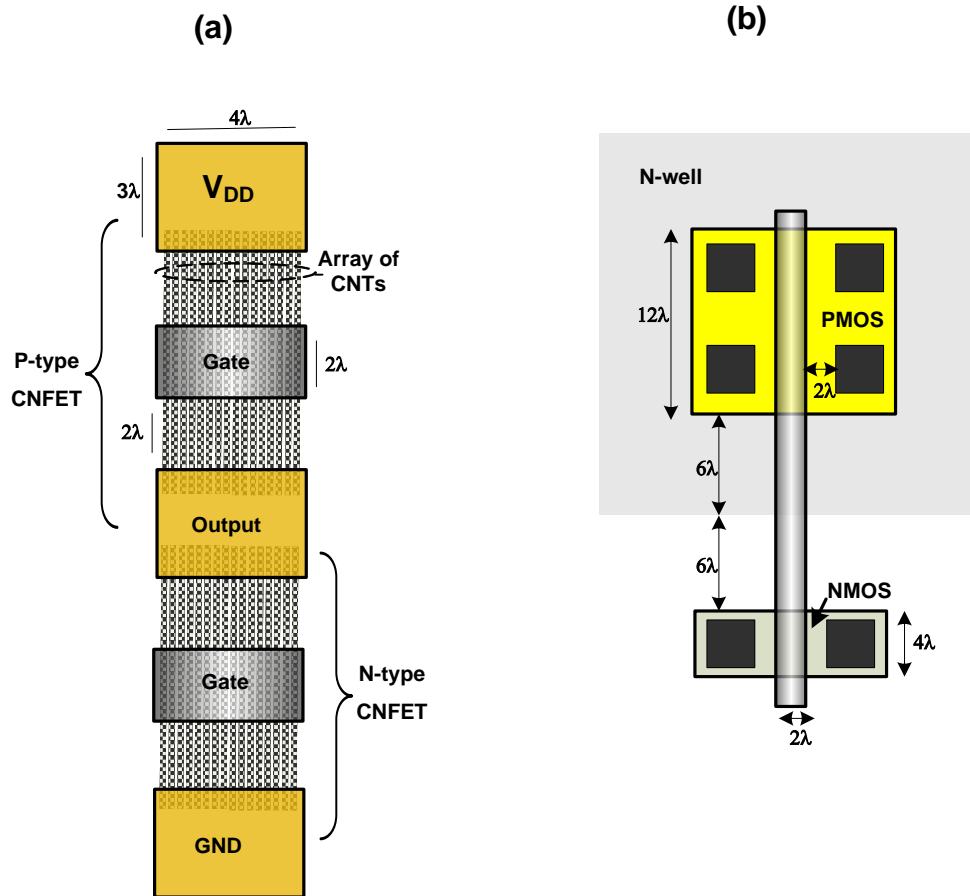
Figure 3-4 :(a) Layout of a CNT based inverter implemented with parallel tubes (b) Layout of a Si CMOS inverter. Note: Figures are not drawn to scale

In Chapter 5 and 6 we will be analyzing the area trade-offs in terms of circuit level techniques and tube level redundancy techniques.

3.5   Functional Yield

The fabrication imperfections associated with the synthesis of CNTs mainly impacts the performance and power consumption of CNT based circuits. The presence of metallic

tubes creates a short between the source and drain of a transistor and result in an increased delay, and static power of CNFET based gates. Similarly, if the metallic tubes are removed by extra processing techniques as discussed in Chapter 2, then removal of tubes results in a large variation in the delay of gates. Also if some of the metallic tubes remained that results in large static power.

Figure 3-5 shows the Monte Carlo simulation for parallel tube inverters with number of tubes in the gate (N_{tug}) is equal to16 and when 10% of the tubes is metallic. The presence of metallic tubes impacts both the delay and static power of logic gates. We define the maximum allowable delay and static power constraints for the gates. In the figure we define a window of acceptable delay of 1.3X and static power constraint of 200X in the presence of metallic tubes. These variations in delay and static power are common in nanoscale CMOS technologies [105]. If the gates have delay and static power within the defined constraints the gates are considered functional, and if the delay and static power of the gates due to imperfections is greater than the defined limits, they are considered as non-functional. The functional yield (Y_f) of logic gates is obtained as a function of the drive strength of the gates, percentage of metallic tubes, and percentage of tubes removed if tube removal process is applied. For a gate to be functional, its delay and static power after the removal of tubes must be less than 1.3X delay as that of the fastest gate with no tubes removed, and within 200X static power as that of the lowest static power gate under the absence of any metallic tubes in the gate. The functional yield is then defined as the ratio of a number of functional gates to the total number of gates.
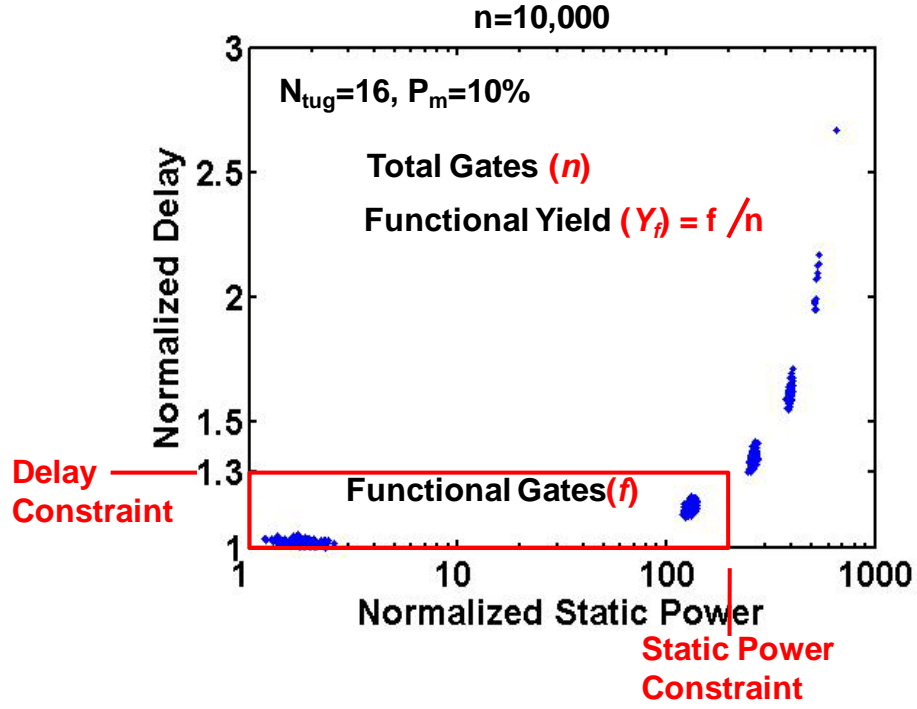
Figure 3-5: Monte Carlo simulation for parallel tubes inverters with N$_{tug}$=16, showing normalized delay vs. static power for 10% metallic tubes

$$Y_{f-gate} = \frac{\text{No. of func gates}}{\text{total No. of gates}} \qquad (3.8)$$

3.6   Conclusions

The performance and power advantages of CNFET based circuits are hampered by the CNT fabrication imperfections. In this chapter, we discussed the parameters i.e. power and performance (delay) which will be used in the subsequent chapters to analyze the impact of fabrication imperfections such as variation in the diameter and spacing among tubes, unwanted growth of metallic tubes and variation resulting because of the removal of tubes on CNFET based circuits. We also propose solutions in the subsequent chapters which

help to mitigate the impact of fabrication imperfections and result in CNT based circuits with acceptable levels of functional yield.

4 Diameter and Spacing Variation

Part of this chapter will be submitted to Rehman Ashraf, Malgorzata Chrzanowska, Siva G. Narendra," Performance Analysis of CNFET based Circuits in the Presence of Fabrication Imperfections", IEEE NANO,2011

The synthesis of SWCNTs results in a variation in the diameters of the tubes. The diameter of the tubes impacts the bandgap of CNFETs which in turn impacts the drive current of CNFETs. Therefore, variation in the diameter of tubes results in the variation in the drive current of CNFET. These variations in the diameter impacts the performance of CNFET based circuits. Similarly, for scalable devices there is a need to fabricate CNFETs with a dense array of parallel tubes as a channel. The spacing between adjacent tubes in a parallel tube CNFET also impacts its drive strength because of variable charge screening from the neighboring tubes. This chapter analyzes the impact of variation in the diameter and spacing on the performance of parallel tube CNFETs. The flowchart of Monte Carlo simulation setup used in this chapter is provided in Section 10.3.1 of Appendix B.

4.1 Variation in the Diameter of CNTs

Significant progress has been made in the synthesis of CNTs with controlled diameter. Typically CNTs have a diameter range of 1nm to 2nm, and the variation in the diameter follows a Gaussian distribution. Therefore, we assume that CNTs have μ diameter of 1.5nm and 3σ diameter variation is 0.5nm. This results in a mean I_{ON} and I_{OFF} of 38μA and 0.84nA respectively for a single-tube CNFET. Here we expressed the variations in the *ON* and *OFF* currents in terms of sigma-to-mean (σ/μ) ratio which is called the coefficient of variation. The advantage of using coefficient of variation is that the variation is expressed

relative to the mean. Figure 4-1 shows the variations in sigma-to-mean I_{ON} and I_{OFF} currents as a function of number of parallel CNTs (N_{tur}) in a transistor with the assumption that all the tubes are present and semiconducting. The spacing (S) between adjacent CNTs is assumed to be large enough that the tubes have no impact from the adjacent tubes. Charge screening impact from adjacent tubes is negligible at this spacing value. Sample size (n) of 1000 transistors is used for Monte Carlo simulations. From Figure 4-1 we observe that the σ/μ variation in the ON and OFF currents is decreasing with the increasing number of tubes in the transistor. The maximum variation in the ON and OFF currents is almost 10%, and 3.5X and is for $N_{tur}=1$. Similarly, the minimum variation in the ON and OFF currents is almost 2% and 0.5X, and is for $N_{tur}=32$. This decrease in variations while increasing the number of parallel CNTs is due to statistical averaging of currents among the multiple tubes of the transistor. As the maximum variation in the OFF current is 3.5X, which is more than three orders of magnitude less than that observed in nanoscale devices. Therefore in the rest of the chapter we will be focusing on the variation in the ON current. We present the impact of OFF current in Chapter 5 where a large current in the OFF network of CNFETs is observed because of the presence of metallic tubes.
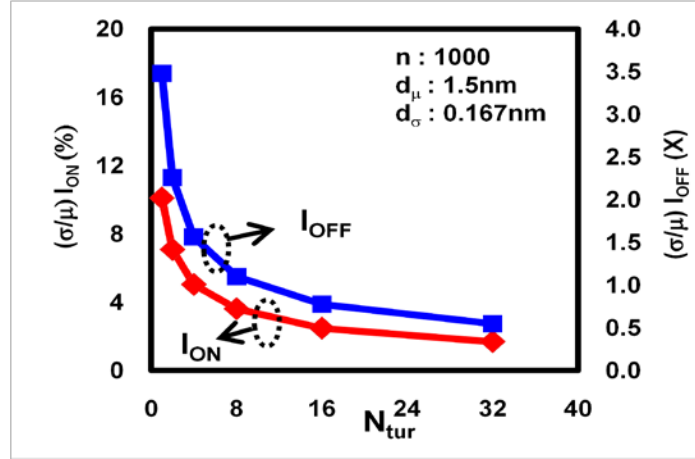
Figure 4-1: Impact of CNT diameter variation on the (σ/μ) *ON* and *OFF* currents for different drive strength as measured by number of tubes (N_{tur}) in a transistor. The mean (d_μ) and sigma (d_σ) of diameter distribution is 1.5nm and 0.167nm respectively.

4.2 CNT Spacing

The spacing between the adjacent parallel tubes in a CNFET impacts the drive strength of parallel tube transistors due to screening of charge from the adjacent tubes. In [84], the authors calculated the gate capacitance of multichannel CNFETs by considering the coupling capacitance between the gate and one isolated CNT (C_{gc_inj}) and the equivalent capacitance (C_{gc_sn}) due to charge screening from the adjacent tubes as given in(4.1) . In equation (4.1), ε_0 is the permittivity of free space, r is the radius of a CNT, h_{ox} is the gate dielectric thickness between the gate and the center of CNT, and k_1 and k_2 are the dielectric constants of gate and bulk oxide. *P* is the pitch between the centers of two adjacent CNTs.

$$C_{gc\_inf} = \frac{2\pi k_1 \varepsilon_0}{\cosh^{-1}\left(\dfrac{2h_{ox}}{d}\right) + \lambda_1 . \ln\left(\dfrac{2h_{ox} + 2d}{3d}\right)} , \qquad\qquad \lambda_1 = \frac{k_1 - k_2}{k_1 + k_2}$$

$$C_{gc\_sr} = \frac{4\pi k_1 \varepsilon_0}{\ln\left(\dfrac{P^2 + 2(h_{OX} - r).\left[h_{OX} + \sqrt{h_{OX}^2 - r^2}\right]}{P^2 + 2(h_{OX} - r).\left[h_{OX} - \sqrt{h_{OX}^2 - r^2}\right]}\right) + \lambda_1 \ln\left(\dfrac{(h_{OX} + 2r)^2 + P^2}{9r^2 + P^2}\right) \tanh\left(\dfrac{h_{OX} + r}{P - 2r}\right)}$$

(4.1)

Figure 4-2 shows a CNFET channel composed of three CNTs arranged in parallel. Two CNTs at the both edges of the channel are getting charge screening from only one adjacent tube each, and the middle tube is experiencing the charge screening from adjacent tubes on both sides. In [84] the authors assumed that all spacing's between adjacent tubes are the same, as for example the distances S_1 and S_2 shown in Figure 4-2.This assumption results in an accurate estimation of charge screening from the neighboring tubes when all the tubes are present because the small variation in the spacing's from adjacent sides. However, it will be shown in the Section 4.4 of this chapter that when extra processing techniques are used to remove unwanted metallic tubes, the assumption of considering the same charge screening from adjacent tubes on both sides will result in an overestimation of charge screening from neighboring tubes. In this work we calculate the equivalent capacitance of the edge tubes in the same way as described in [84], with the modification of equation for the equivalent capacitance of the tubes which are not at the edges, with adjacent tubes on both sides (C_{gc_m}), due to the influence of the spacing variations between tubes on the charge screening effect. Spacing variations will play a major role after metallic tubes are removed, as they can change significantly. This modified equation for equivalent capacitance of the middle tubes is given in equation (4.2)

$$C_{gc\_m} = \frac{C_{gc\_inf}C_{gc\_sr}(P_1)C_{gc\_sr}(P_2)}{C_{gc\_sr}(P_1).C_{gc\_sr}(P_2) + C_{gc\_sr}(P_2)C_{gc\_inf} + C_{gc\_sr}(P_1)C_{gc\_inf}} \tag{4.2}$$
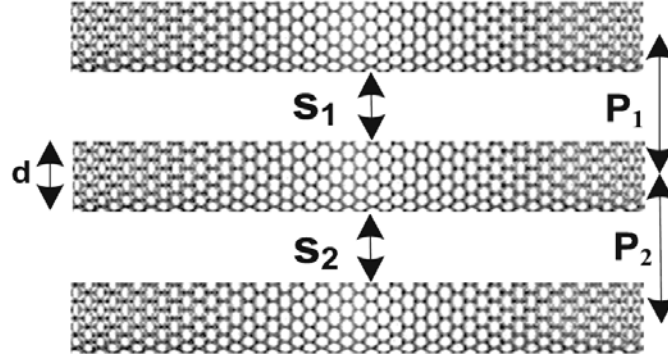
Figure 4-2: An array of three parallel CNTs where *d* is the diameter of tubes, *S₁ and S₂ is* the spacing between adjacent tubes, and *P₁ and P₂* is the pitch between the center of two adjacent CNTs.

Current CNT synthesis technology allows packing of almost 10-50 CNTs/μm, which for 1.5nm diameter tubes corresponds to spacing of almost 100nm – 20nm between adjacent tubes [81], [82]. Deng et al. [83] analyzed that when spacing between adjacent parallel CNTs are greater or equal to 20nm, the charge screening impact from adjacent tubes is negligible and there will be negligible impact on the drive strength of CNTs due to charge screening from adjacent tubes. However, it was evaluated that almost 250 CNTs/μm are required to obtain performance and energy gains over silicon CMOS [78]. This corresponds to spacing (*S*) between adjacent tubes of 2.5nm for a tube diameter (*d*) of 1.5nm. This spacing of 2.5nm between adjacent tubes will result in significant charge screening from adjacent tubes.

Figure 4-3 shows the impact of charge screening from adjacent tubes on the mean drive current (μ)I_{ON}. As the amount of charge screening is a function of spacing (*S*) between the

tubes, Case S=20nm is considered as the reference case when screening from adjacent tubes is negligible. It can be observed that reducing the spacing between adjacent tubes, from 20nm to 2.5nm reduces the drive current of parallel tube CNFET by almost 30% irrespective of the number of tubes in the channel.
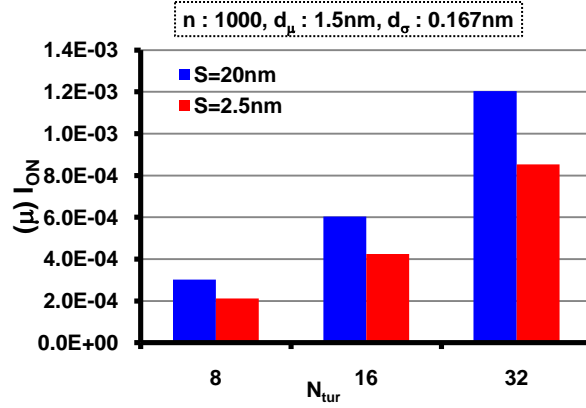
Figure 4-3: Mean drive (*ON*) current in the CNFET for two values of spacing (*S*) between adjacent tubes and three values of N$_{tur}$. *S*=20nm is considered as reference case when screening from adjacent tubes is negligible.

4.3   Variations in the Pitch of CNT

The CNTs fabrication process results in certain variability in terms of spacing between adjacent parallel tubes. The combined variations in spacing and diameter of tubes result in the pitch variation between adjacent parallel tubes. For fixed channel width CNFET devices, the pitch variations will result in the variation in charge screening and in tube density variations among different CNFETs. Researchers in [106] analyzed the impact of the density variations of tubes on the yield of CNFET devices. They considered the devices to be functional if there were at least a single CNT present in the parallel tube CNFET. Our analysis in this work shows that for the required density of 250 CNTs/μm, the maximum possible spacing variation before the tubes crossover in the CNTs will result in

density variations, but the probability of having a transistor with no tube present in its channel is negligible. However, the density variation results in variation in the total drive current of the transistors because of variation in the number of conducting channels available.

Figure 4-4 shows the impact of the pitch variation on the drive current of CNFETs for various transistor drive strengths represented by the number of tubes in the transistor. The combined impact of diameter and spacing variation results in less than 8% variation in the mean drive current of parallel tube CNFETs. Moreover, the variation decreases by increasing the drive strength of CNFETs. From the analysis it is observed that both the diameter and spacing variations can be tolerated in parallel tube CNFETs. This is mainly due to statistical averaging among multiple parallel channels.
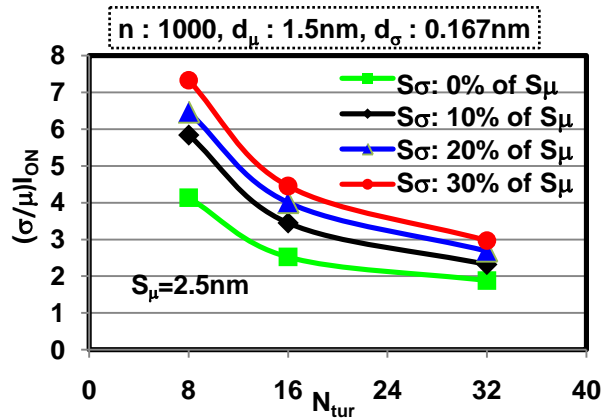
Figure 4-4: Impact of pitch (diameter and spacing) variation between adjacent CNTs on the (σ/μ) *ON* **current as a function of number of parallel tubes (N_{tur}) in a CNFET.**

4.4 Removal of Metallic CNTs

It is mentioned in Chapter 2 that unwanted growth of metallic tubes is one of the biggest challenges faced by the CNT technology. It will be shown in Chapter 5 that if the

percentage of metallic tubes is larger than 5%, then post processing techniques are required to remove the metallic tubes to build robust CNT based circuits. Based on the discussion in Chapter 2, two post processing techniques for tube removal are of main interest, SCE and VMR. Both of these techniques remove almost all of the metallic tubes, but as a side effect, they also remove some of the needed semiconducting tubes.

In this work we used data based on the SCE technique and our evaluation methodology can also be applied to the VMR technique. The removal of tubes results in large delay variations, and in the worst case, open-circuit gates can be created due to all the tubes being removed. Since open-circuit devices significantly reduce yield, our primary objective is to find the minimum number of tubes (N_{turmin}) needed in a CNFET prior to Selective Chemical Etching, that produce less than 0.001% probability of open circuit CNFETs. Now if P_r is the probability of tube being removed by SCE, and N_{tur} is the number of tubes in the CNFET, then the probability of all the tubes removed from the transistor is equal to $P_r^{N_{tur}}$. Based on this N_{turmin} can be obtained as

$$N_{turmin} = \log\left(\frac{10^{-5}}{P_r}\right)$$ (4.3)

P_r can be obtained from equation (4.4)

$$P_r = P_s P_{sr} + P_m P_{mr}$$ (4.4)

Here P_{sr} is the conditional probability that the tube is semiconducting and it is removed. The probability of P_{sr} is obtained by

$$P_{sr} = \int\limits_{-\infty}^{D_{CS}} P_s(x)\,dx \qquad\qquad (4.5)$$

Similarly, P_{mr} is the conditional probability that the tube is metallic and is removed. P_{mr} can be obtained by

$$P_{mr} = \int\limits_{-\infty}^{2nm} P_m(x)\,dx \qquad\qquad (4.6)$$

P_m is the percentage of metallic tubes and P_s is the percentage of semiconducting tubes before the application of SCE. After the application of SCE process, the number of remaining tubes left in the channel will be N_{tur}, resulting in a reduction of the drive current and significant increase in the variation in drive current. On the other hand, because of removal of tubes, there will be, on average, an increase in spacing between adjacent tubes resulting in a reduction in the charge screening effect that will cause an increase in the drive current of CNFETs.

Figure 4-5(a) shows a CNFET composed of N_{tur}=8 tubes prior to the application of SCE. Tubes T_1-T_8 are arranged in parallel in the channel of CNFET and S_1-S_7 is the spacing between adjacent tubes, and P_1-P_7 is the pitch of tubes T_1-T_8. Figure 4-5(b) shows one snapshot of Monte Carlo simulations of 1000 transistors when P_m=5% and with 31% of tubes removed through the SCE process. Here, tubes T_1, T_3 and T_6 are removed after the application of SCE. The distance between tubes T_2 and T_3 before SCE was S_2 and after SCE, because of the removal of tube T_3 the distance between T_2 and T_4 is $S_2+S_3+d(T_3)$. On the one hand, the removal of tubes T_1, T_3 and T_6 will results in a decrease in the overall current of CNFET because of reduction in the number of conducting channels, while on

the other hand, the current from the remaining CNTs will increase due to reduced charge screening from adjacent tubes. Overall, the removal of tubes by SCE and the resulting spacing variations between remaining tubes results in large variation in the *ON* current of transistors.
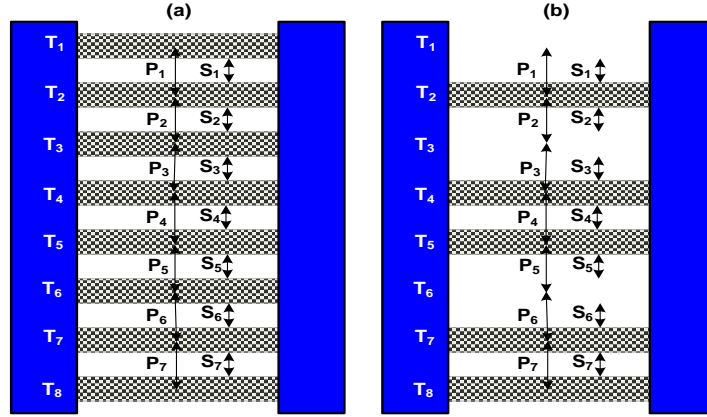
Figure 4-5: CNFET consisting of parallel CNTs (a) Reference case with N_{tur}=8 when all tubes are semiconducting (b) Random sample taken from Monte Carlo simulations after applying SCE. 5 tubes are remained with large variation in spacing between adjacent tubes.

Figure 4-6 shows the normalized mean drive(*ON*) current of parallel tube CNFETs for three values of tubes (N_{tur}) in a channel, P_m=0% ,no SCE and no charge screening from neighboring tubes (green), P_m=5% and SCE is applied and impact of reduction in charge screening from adjacent tubes due to removal of tubes is considered(blue). We refer the approach of tubes removed and considering the resulting variation in charge screening from adjacent tubes as TRCS (Tube Removal and Charge Screening) considered, P_m=5% and SCE is applied but no TRCS is considered (red).
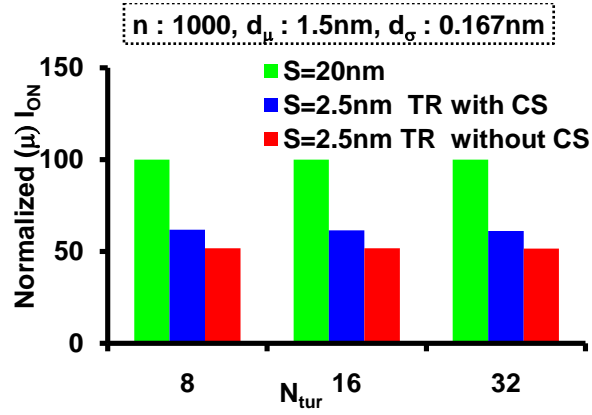
Figure 4-6: Normalized mean drive current when P_m=0% and no SCE is applied, when P_m=5% and SCE is applied and reduction in charge screening from adjacent tubes is considered, and when P_m=5% and SCE is applied but no reduction in charge screening from adjacent tubes is considered because of the removal of tubes.

Without TRCS, an almost 50% reduction in the μ drive current is observed, and when TRCS is included (realistic case) the reduction in the μ drive current is 40% ,which is 10% lower than without considering TRCS. Figure 4-7shows the combined impact of diameter variation, spacing variation and removal of tubes on the (σ/μ) drive current as a function of the number of parallel tubes in the CNFET. It can be observed that with a realistic example of 5% of tubes being metallic, a less than 15% (σ/μ) variation in performance can be obtained when $N_{tur} \geq 32$.
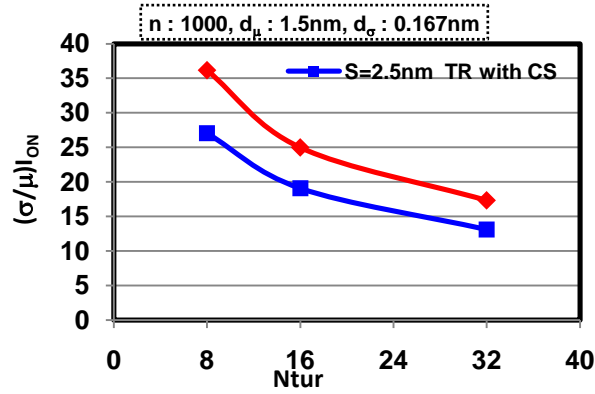
Figure 4-7: Impact of diameter (*d*) , spacing (*S*) variations between adjacent CNTs and tube removal on the (σ/μ) *ON* current as a function of the number of parallel tubes (*N_{tur}*) in a CNFET.

Assuming that technology allows us to fabricate CNTs with a density of 250 CNTs/μm, from the three discussed challenges; diameter variation, spacing variation and metallic tube removal, the first two will have negligible impacts on the performance of CNFET based circuits. The removal of metallic tubes, however, will be a significant source of performance degradation and variation.

4.5 Summary

In this chapter, we have analyzed the impact of different fabrication imperfections, such as variation in the diameter, impact of spacing and variation in spacing among adjacent tubes on the performance of parallel tube CNFETs. Our analysis shows that both the diameter and the spacing variations make a negligible impact on the performance of CNFET based devices due to statistical averaging among adjacent tubes. However, the existence and removal of metallic tubes is shown to have a significant effect resulting not only in a large performance reduction, but also in a large increase in performance variability. The charge screening effect between adjacent tubes in a CNFET channel has the opposite, to tube

61

removal effect on the performance of CNFETs. Therefore, considering the charge screening effect make the evaluation of device performance and its variation more accurate.

5   Circuit Level Solutions for CNFETs with Metallic Tubes Present

Part of this chapter has been published in:

- Rehman Ashraf, Malgorzata Chrzanowska, Siva G. Narendra," Functional Yield Estimation of Carbon Nanotube based Logic Gates in the Presence of Defects", *IEEE Trans. Nanotechnology, 2010*
- Rehman Ashraf, Malgorzata Chrzanowska, Siva G. Narendra," Design Methodology for Carbon Nanotube based Circuits in the Presence of Metallic Tubes", NANOARCH,2010
- Rehman Ashraf, Malgorzata Chrzanowska, Siva G. Narendra, "Carbon Nanotube Circuit Design Choices in the Presence of Metallic Tubes", ISCAS, 2008

Semiconducting tubes are required for the fabrication of CNFET based circuits. However, there is no known CNT fabrication method which can produce 100% semiconducting tubes. Current CNT synthesis techniques yield between 4% to 40% [91], [92] metallic tubes as discussed in Chapter 2. In the case of metallic tubes, the gate terminal has no control over the channel due to an ohmic short between the drain and the source of a transistor. Therefore, complementary CNFET based circuits with metallic tubes have a detrimental impact on static power, delay, noise margin, and yield of CNFET based circuits because of the contention current from the metallic tubes present in the *OFF* network of a gate. For small percentage of metallic tubes i.e. less than 5%, circuit level techniques can be used to handle the detrimental impact of metallic tubes. In this chapter, two CNFET configurations are proposed [99], which reduces the statistical probability of a short between the source and the drain terminals of a transistor in the presence of metallic tubes. The circuit level techniques help to increase the functional yield of gates in the presence of metallic tubes but the trade-off is in terms of reduction in the performance of the gates.

Also in this chapter, we present a methodology for yield-aware circuit design in the presence of metallic tubes using different CNFET transistor configurations. Similarly for ASIC design styles, we propose to implement CNT based circuits using regular logic blocks (bricks) proposed by [107] to reduce the systematic lithographic related variations associated with the nanoscale fabrication technologies.

5.1 CNFET Configurations Proposed in the Literature

Two CNFETs configurations have been proposed in the literature. Shared Tube (*ST*) configuration was demonstrated experimentally by [65] and is shown in Figure 5-1(a). In this configuration one long tube with alternating source and drain contacts is used to create four parallel channels. The Parallel Tube (*PT*) configuration, shown in Figure 5-1(b), was theoretically evaluated by [77], and practically demonstrated by [81]. In this configuration four separate parallel tubes (channels) are arranged in parallel and all tubes have shared source and drain terminals.
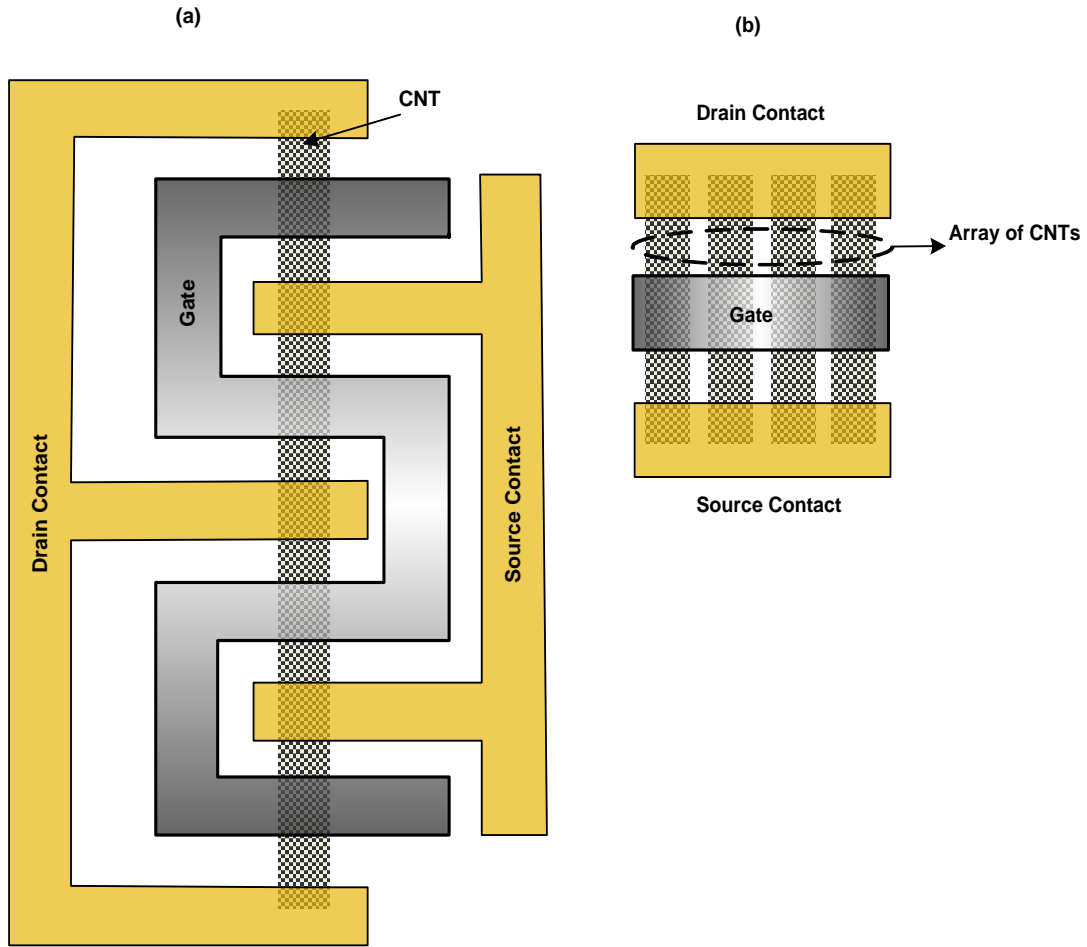
Figure 5-1: Tube configurations for (a) Shared Tube (*ST*) and (b) Parallel tube (*PT*) CNFET. Both the configurations have the same number of channels to present iso-input capacitance.

Parallel tube CNFETs can be fabricated so that multiple transistors share the same tubes (correlated tubes), and where performance of the transistors is highly correlated with respect to tube variations. The transistors or tubes can also be arranged such that each transistor has a separate set of tubes (un-correlated tubes) in which case the performance of these CNFETs are un-correlated with respect to tube variations[99], [108], [109].

In the case of *ST* configuration, each transistor has only one tube, the same tube, therefore all the channels will be highly correlated, and if that tube is metallic, an ohmic short

65

between the source and the drain terminals is created. Functional yield calculation of gates implemented in *ST* is deterministic once the percentage of metallic tubes is given. If a tube is metallic it will result in a non functional gate.

5.2   Proposed Tube Configurations in CNFETs

In *PT* and *ST* configurations, the presence of metallic tubes will result in an ohmic short between the source and the drain of a transistor. To reduce the statistical probability of a short between the source and the drain we proposed two new tube configurations.

Figure 5-2 (a) shows Transistor Stacking (*TrS*) configuration where two transistors with un-correlated (different) parallel tubes are stacked through a common intermediate node between the power and output. In Figure 5-2(b) Tube Stacking (*TuS*) configuration is shown in which each stacked parallel path from the output to power is isolated from each other by not having a shared intermediate node. These stacking configurations help to reduce the probability of an ohmic short between the power and the output. Clearly, in the stacked configurations, more than one tube has to be metallic to create an ohmic short between the power and the output. To maintain iso-input capacitance, the total number of tubes is kept the same in both stacking configurations as it is in *PT* configuration. This will result in same load on the driving gate and gates with either parallel tube or stacking configurations can be used interchangeably. While the stacked configurations could possibly reduce the number of ohmic shorts, it comes with a performance penalty. This performance penalty is because of two reasons, (a) the number of parallel tubes in the stacking configuration is reduced by half, measure of the drive strength, therefore the drive strength is reduced by 2X, (b) and two transistors or tubes are stacked which increases the

overall resistance of the channel by almost 2X or decreases the drive strength by 2X. Therefore, overall stacking configurations can result in up to 4X performance penalty.
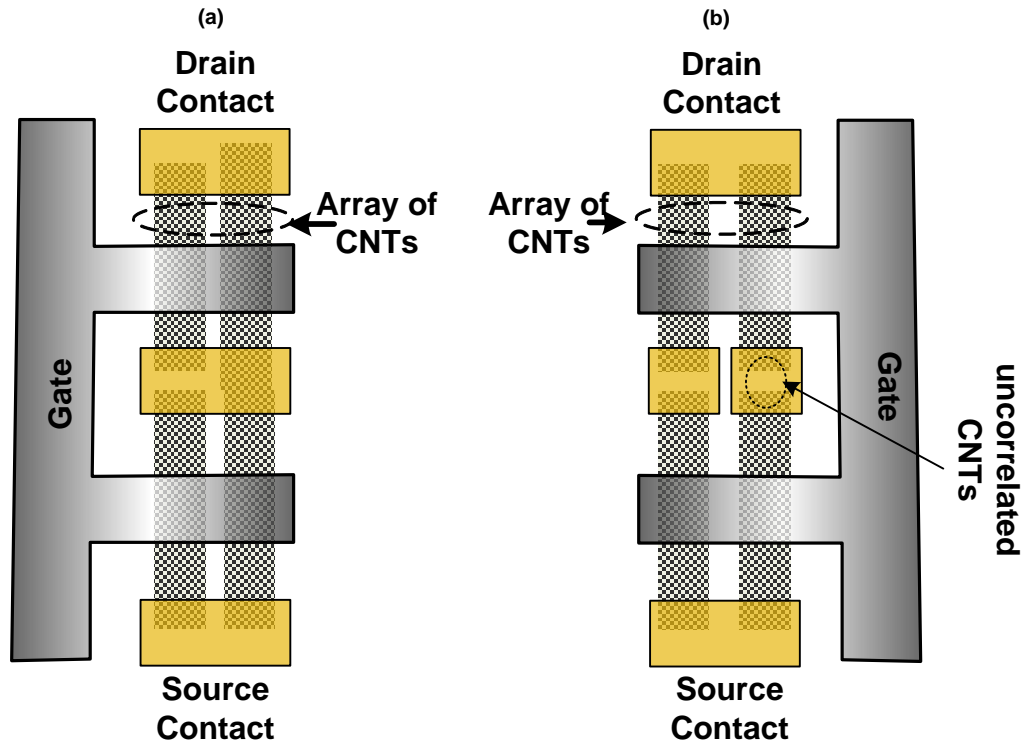
Figure 5-2: Tube configurations for (a) Transistor Stacking (*TrS*) and (b) Tube Stacking (*TuS*) CNFET. Both the configurations have the same number of channels to present iso-input capacitance.

A probability of an ohmic short between the drain and the source of a transistor in the *TuS* configuration, as compared to *TrS* configuration, is lower. Fabrication of *TuS* configuration, however, requires more precise control in terms of tube alignment and positioning of contacts. Figure 5-3 shows examples of a transistor implemented in stacked configurations in the presence of metallic tubes. Figure 5-3(a) shows the *TrS* case in which tubes T_3 and T_5 are metallic, resulting in a direct short between the drain and source of the CNT transistor. Figure 5-3(b) and Figure 5-3(c) show a transistor implemented in *TuS* configuration with

two contacts, C_2 and C_3, shorted due to lack of precision. In Figure 5-3(b), tubes T_3 and T_5 are metallic, as in *TrS* configuration of Figure 5-3(a), but in this case there is no ohmic short between the drain and source, and *TuS* configuration is maintained with slightly changed performance. In Figure 5-3(c), contacts C_2 and C_3 are shorted as in Figure 5-3(b) but a different pair of tubes, T_3 and T_6, is metallic such that there is an ohmic short between the drain and source. The worst-case, very unlikely, contact-positioning situation in *TuS* would be when all contacts are shorted. Such a case is not shown in Figure 5-3 but it would be equivalent to *TrS* case. In many cases, however, contact overlaps are not critical as even with the presence of metallic tubes in a transistor, they will not necessarily create a direct short between the drain and source terminals of a CNT transistor in the presence of metallic tubes.
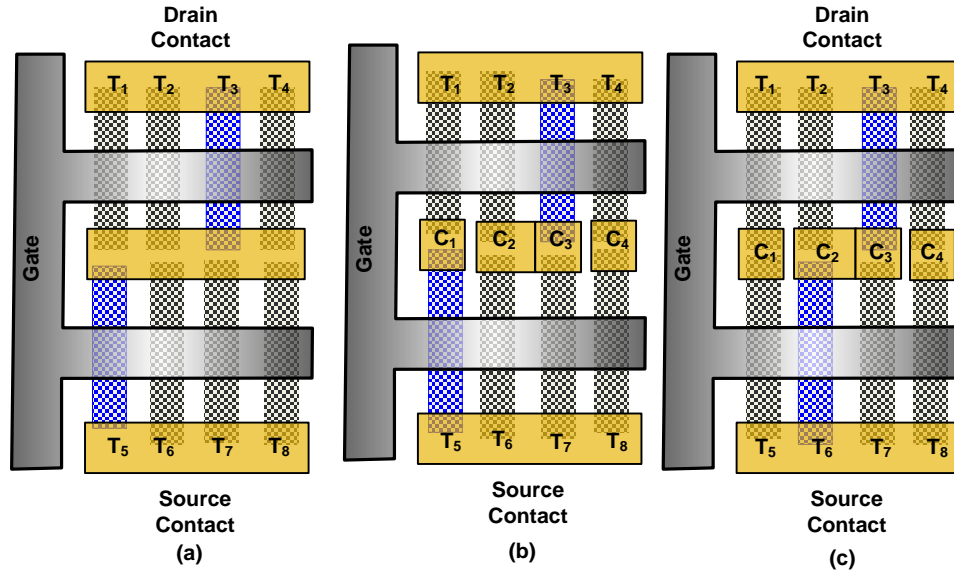
Figure 5-3: (a) Transistor Stacking (*TrS*) configuration with tubes T_3 and T_5 being metallic (b) Tube Stacking (*TrS*) configuration with tubes T_3 and T_5 being metallic and shortened contacts C_2 and C_3 (c) Tube Stacking (*TuS*) configuration with tubes T_3 and T_6 being metallic and shortened contacts C_2 and C_3.

From the above analysis and examples in Figure 5-3, we can conclude that when considering possible contact overlaps in *TuS* in the presence of metallic tubes, only in some percentage of cases, the yield and performance of *TuS* transistor will be reduced to that of *TrS* case.

The other important manufacturing challenge specific to CNFET technology is the alignment of carbon nanotubes in arrays of parallel tubes. The misaligned tubes in stacked parallel paths in *TuS* configuration may result in opens in the stacked channels and hence negatively impact the yield of the gates. This impact of un-contacted tubes on yield is analyzed in the next section of this chapter.

It is worth noting that if there are two semiconducting tubes stacked in series from output to power, then there will be leakage reduction due to stack effect [110-113]. This assumes that the CNTs are used to realize traditional FETs with unipolar conduction characteristics. The probability of stack effect based leakage reduction is higher in *TuS* because there is no node sharing in *TuS* configuration as compared to *TrS* configuration of a CNFET.

5.3   Monte Carlo Simulation for Functional Yield of Logic Gates

Monte Carlo simulations are used to generate functional yield for an inverter and NAND gate built of CNFET transistors with different configurations of tubes. The flowchart of Monte Carlo simulation setup used is provided in Section 10.1.1 of Appendix B. The yield results are used to validate those obtained from analytical models developed in Section 5.4 to 5.6 of this chapter.

The functional yield is calculated as a function of drive strength of the gate as required by the circuit design and as a function of the percentage of metallic tubes as defined by the synthesis process. As demonstrated by available technologies the percentage of metallic tubes is between 4% [92] and 40% [91]. The impact of the presence of metallic tubes for different configurations of an inverter is analyzed in [99], and it is observed that by increasing the drive strength of the gate, the functional yield of *PT* inverter asymptotically approaches to 0% when more than 30% of the tubes are metallic. The detrimental impact of the presence of metallic tubes on the performance of an inverter makes it impossible to build circuits with acceptable performance and functional yield, when the percentage of metallic tubes is larger than 10%. In this work, the maximum percentage of metallic tubes considered for Monte Carlo simulations is 10%. It will be shown later that for more

complex gates like NAND gate, the presence of metallic tubes has a more adverse impact on the functional yield of gates, with even 10% of metallic tubes seeming too high to build robust circuits with acceptable power, performance and functional yield.

5.3.1 Inverter

The schematic and layout of an inverter consisting of *PT* CNFETs is shown in Figure 5-4. Each transistor has four channels (tubes). To analyze the functional yield of an inverter with Monte Carlo simulations we utilize the methodology given in Chapter 3.
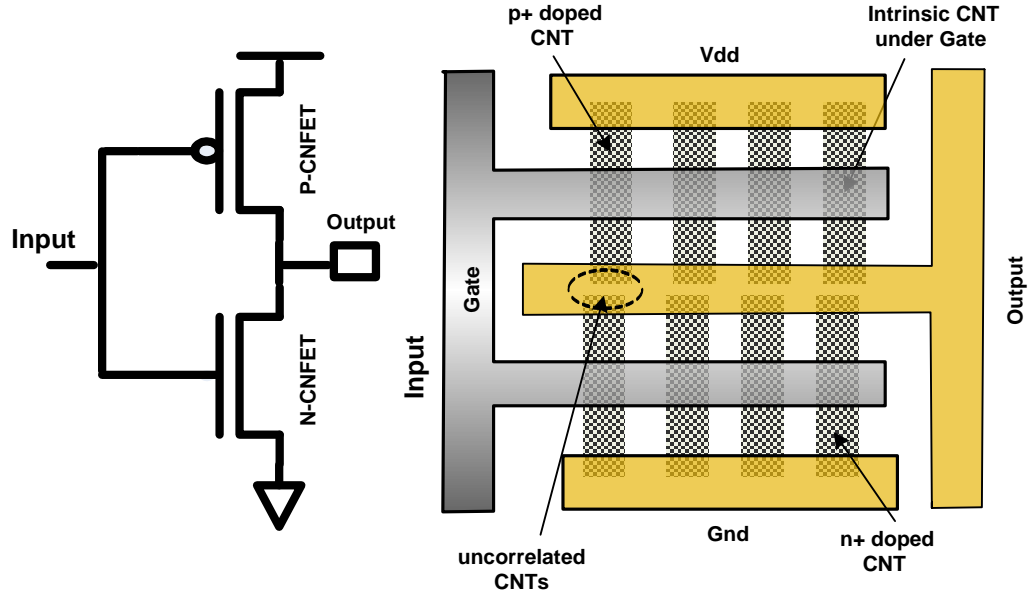
Figure 5-4: Schematic and layout of *PT* inverter containing an array of four CNTs in P-CNFET and N-CNFET.

Figure 5-5 shows normalized delay vs. normalized static power for *PT* inverter configuration generated through Monte Carlo simulation without and with a different percentage of metallic tubes present. The number of tubes in the gate (N_{tug}) was 16. The data points shown are for inverters that *(a)* do not exceed 1.3X delay of the fastest inverter

71

under absence of metallic tubes and *(b)* do not exceed 200X static power of the lowest static power in an inverter under the absence of metallic tubes. These variations in delay and static power are common in nanoscale CMOS technologies [105]. Sample size (*n*) of 10,000 was used for all Monte Carlo simulations. Figure 5-5(a) shows the impact of diameter variation of CNTs on delay and static power of the gate. In Figure 5-5(a) a reference case of 0% metallic tubes resulting in 100% functional yield is shown. In Figure 5-5(b) when 4% of the tubes are metallic, we will see two distributions one with gates having all the tubes being semiconducting, and the other distribution with one of the tubes being metallic but still not violating the maximum delay and power constraint. Results from Figure 5-5 clearly shows that as the metallic content is increased from 0% to 4% and to 10% the number of inverters that have no metallic tubes drop from 100% to 53% and to 19%, respectively. The presence of metallic tubes lowers the overall functional yield from 100% to 93% and to 67%. To increase the functional yield of gates in the presence of metallic tubes, we are using CNFET stacking configurations as discussed in Section 5.2[99].
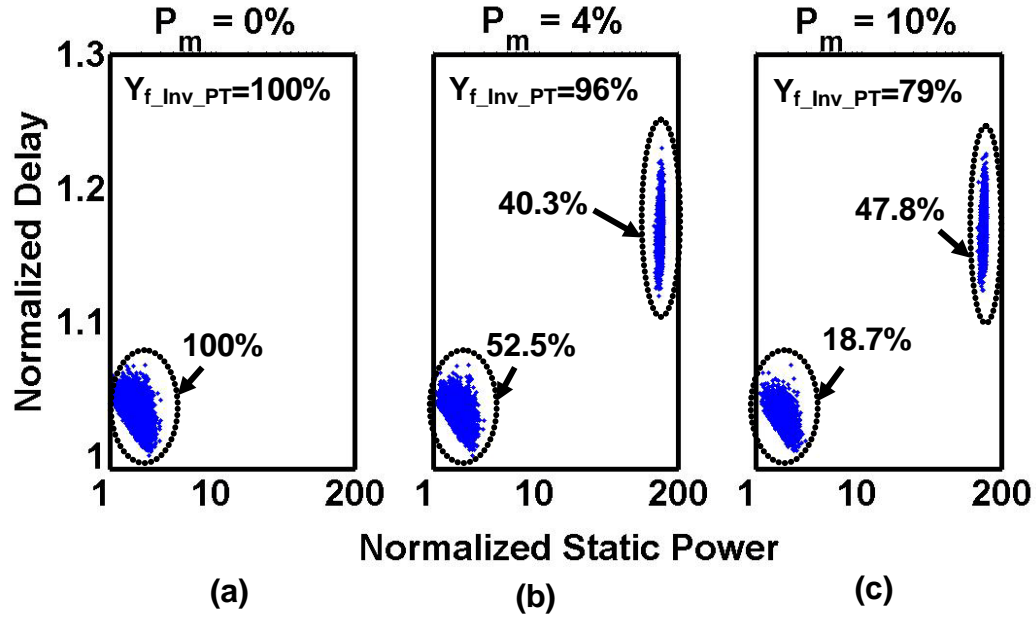
Figure 5-5: Monte Carlo Simulation for Parallel Tube (*PT*) inverters with N_{tug}=16, showing normalized delay vs. static power for (a) absence of metallic tubes – P_m=0%, (b) P_m=4% metallic tubes and (c) P_m=10% metallic tubes with delay constraint of 1.3X and static power constraint of 200X.

In Figure 5-6 Monte Carlo simulations of the functional yield for inverters implemented with *TuS* configuration for different percentages of metallic tubes are shown. We assume that the two arrays of un-correlated tubes used for the *TuS* configuration are perfectly aligned or their misalignment is negligible. As expected, the stacking configuration improves the functional yield as it reduces the statistical probability of a short circuit between the power and the output at the expense of an increase in delay, but with the reduction in static power on the positive side. The stacking configuration also helps to reduce the variation in delay and static power as it can be seen in Figure 5-6. For all the gates which are functional the maximum static power variation is within 10X of the minimum static power and delay variation is within 10% of the minimum delay. This order of magnitude reduction in static power and less variation in the delay help to implement low power circuits with reduced variations.
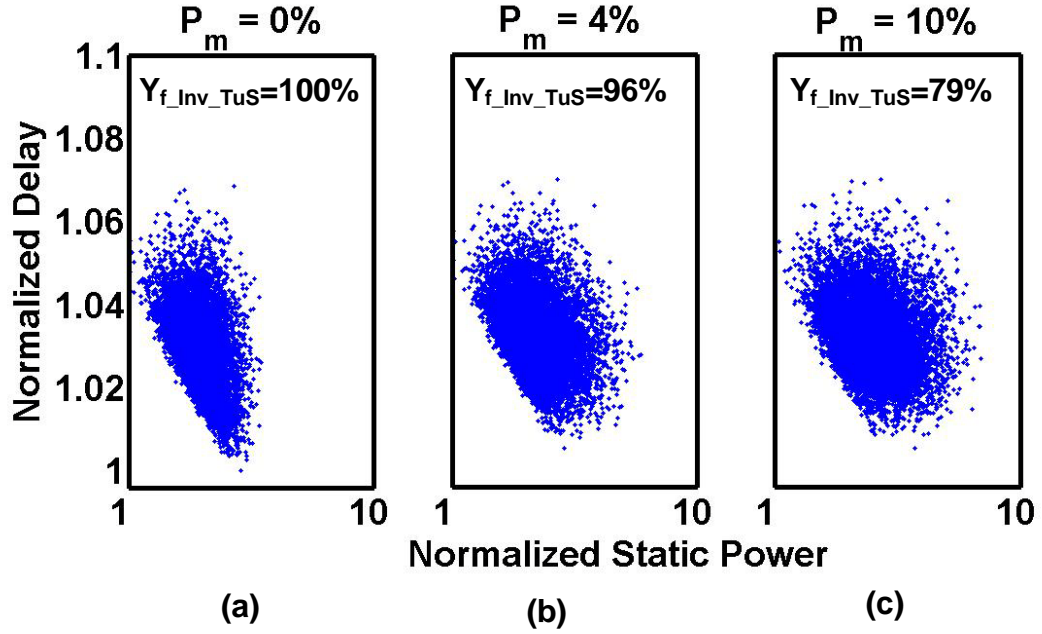
Figure 5-6: Monte Carlo simulation for Tube Stacking (*TuS*) inverters with N_{tug}=16, showing normalized delay vs. static power for (a) absence of metallic tubes – P_m=0%, (b) P_m=4% metallic tubes and (c) P_m=10% metallic tubes with delay constraint of 1.3X and static power constraint of 200X. Scale of 1.1X and 10X is used for normalized delay and static power because there are no gates with delay and static power between 1.1X-1.3X and 10X-200X respectively.

If two un-correlated tube arrays that create *TuS* configuration are not perfectly aligned, some of the stacked channels might be open. We use Monte Carlo simulations to analyze yield losses due to un-contacted tubes in the presence of metallic tubes. Depending upon the percentage of metallic tubes and the percentage of un-contacted tubes, the yield obtained from the *TuS* configuration may be less than that obtained from the *TrS* configuration.

Table 5-1 compares the functional yield of an inverter implemented with *TrS,* and with *TuS* configurations with various percentages of un-contacted tubes, from 0% to 5%. The highlighted numbers represent good design, optimal choices for various percentages of metallic tubes and un-contacted tubes, as *TuS* configurations result in better yields as compared to transistors implemented with *TrS* configurations. From Table 5-1 it can be

observed that for up to 4%, metallic tubes *TuS* gives better yield in the presence of metallic tubes only if the tubes can be precisely aligned. For 4% to 10% of metallic tubes, *TuS* gives better yield if the percentage of un-contacted tubes is not larger than 1%.

Table 5-1: Functional yield for an inverter with Transistor Stacking (*TrS*) and Tube Stacking (*TuS*) configurations for 4% and 10% of metallic tubes and three drive strengths of the inverter. The percentage of un-contacted tubes (*P_{UC}*) in *TuS* configuration varies from 0%-5%.

P_m	N_{tug}	Y_{f_TrS}	Y_{f_TuS}					
			P_{uc}=0%	P_{uc}=1%	P_{uc}=2%	P_{uc}=3%	P_{uc}=4%	P_{uc}=5%
4%	16	95.5	98.7	84.1	71.5	60.6	51.4	43.5
	32	99.7	100.0	97.6	92.1	84.5	75.8	66.6
	64	100.0	100.0	99.9	99.3	97.3	93.3	87.2
10%	16	77.7	92.3	78.6	66.8	56.7	48.0	40.6
	32	93.1	99.5	95.7	89.2	81.1	72.1	63.0
	64	99.1	100.0	99.8	98.7	95.9	91.1	84.2

5.3.2 NAND

A typical standard cell library used to design integrated circuits contains other complex gates like NAND, NOR, AND and OR. In this work, we start by analyzing the functional yield of 2-input NAND gates designed with *PT* configuration in the presence of imperfections such as variation in the diameter of the tubes and presence of metallic tubes. In the case of a 2-input NAND gate, two P-CNFETs are connected in parallel in the pull-up network and two N-CNFETs are connected in series in the pull-down network as shown in Figure 5-7. To obtain an almost equal worst case delay for both high-to-low (D_{HL}) and low-to-high (D_{LH}) transitions, the number of CNTs used in the transistors of the pull-down network is twice the number of CNTs used in the transistors of pull-up network.
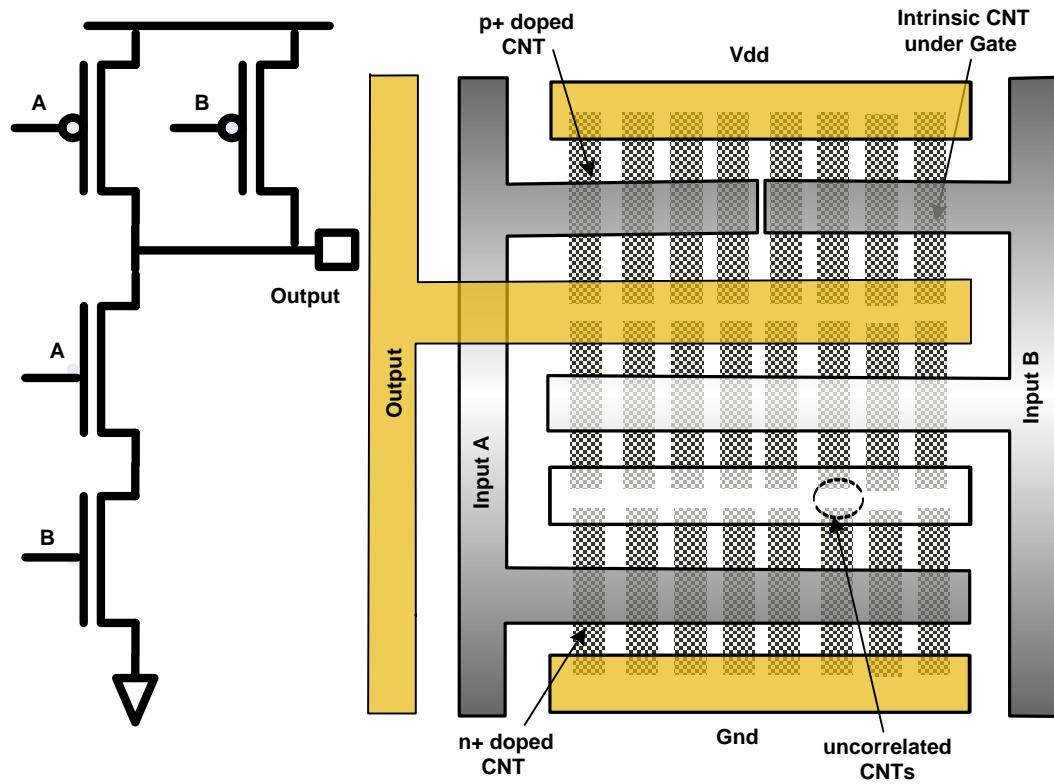
Figure 5-7: CNT based schematic and layout of 2-input NAND gate containing an array of four CNTs in P-CNFETs and an array of eight CNTs in N-CNFETs. The number of tubes in the N-CNFET is twice the number of tubes in the P-CNFET, to make the worst case rise and fall delays equal.

The normalized delay vs. normalized static power for PT NAND gate generated by Monte Carlo simulation is shown in Figure 5-8. Results indicate that the increase in the metallic content from 0% to 4% and to 10% drops the functional yield from 100% to 66% and to 14%, respectively. Please notice in Figure 5-8(a), that the variation in the diameter of the tubes does not impact the functional yield of NAND gates.
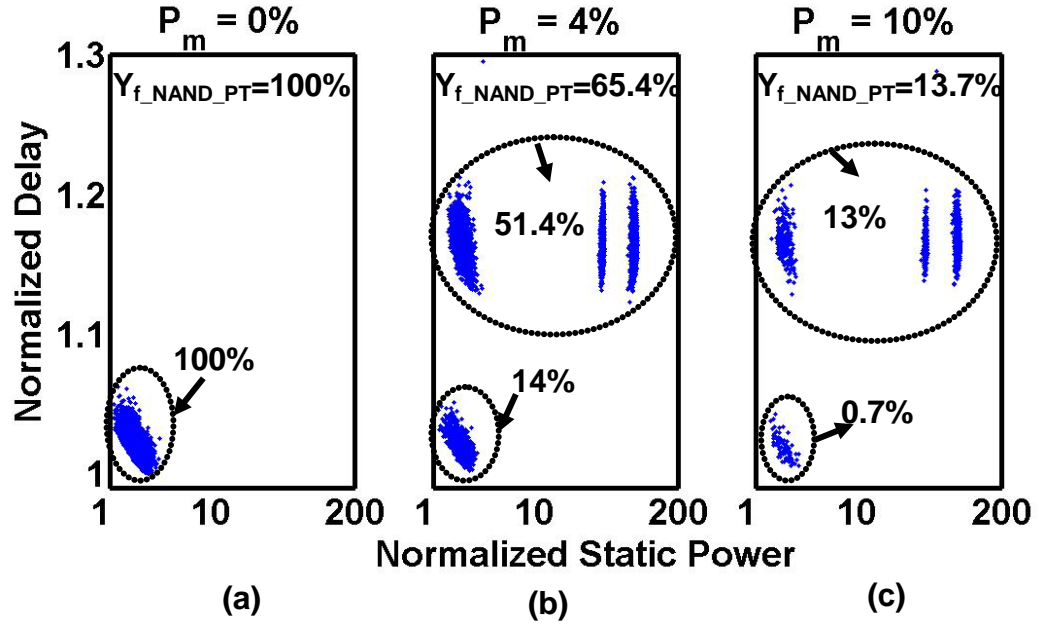
Figure 5-8: Monte Carlo simulation for Parallel Tube(*PT*) NAND gate with N_{tug}=48, showing normalized delay vs. static power for (a) absence of metallic tubes – P_m=0%, (b) P_m=4% metallic tubes and (c) P_m=10% metallic tubes with delay constraint of 1.3X and static power constraint of 200X.

Figure 5-9 shows the functional yield of a NAND gate when Tube Stacking configuration is used. It is observed that the stacking configuration increased the functional yield of NAND gate as expected. For example, for 4% metallic tubes the yield of *TuS* NAND gate is 96%, as compared to 66% when *PT* configuration is used.
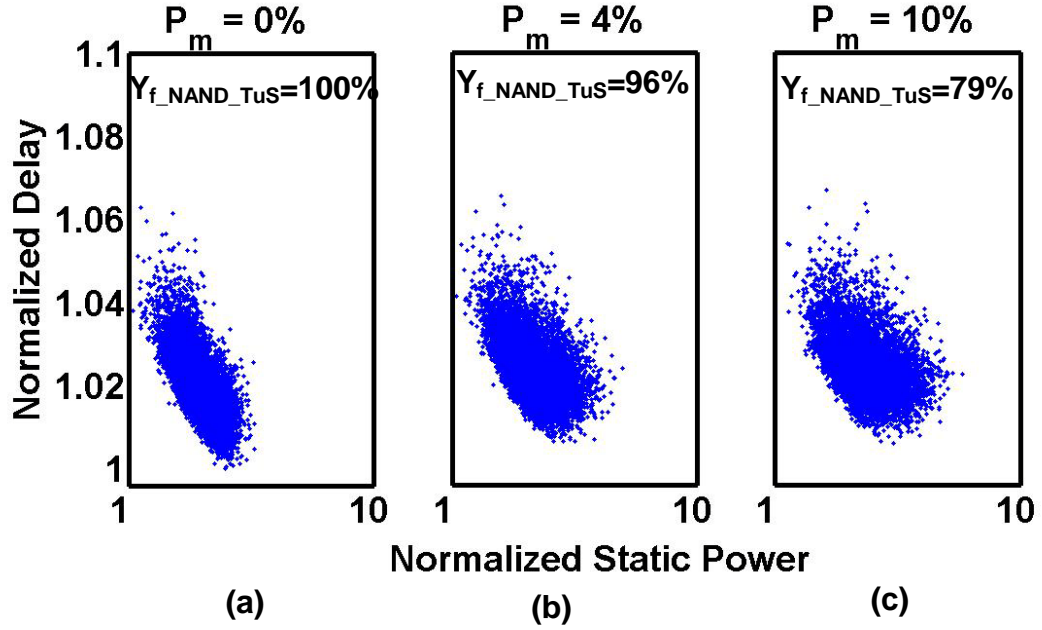
Figure 5-9: Monte Carlo simulation for Tube Stacking (*TuS*) NAND gate with N_{tug}=48, showing normalized delay vs. static power for (a) absence of metallic tubes – P_m=0%, (b) P_m=4% metallic tubes and (c) P_m=10% metallic tubes with delay constraint of 1.3X and static power constraint of 200X. Scale of 1.1X and 10X is used for normalized delay and static power because there are no gates with delay and static power between 1.1X-1.3X and 10X-200X, respectively.

5.3.3 Yield Comparison between Logic Gates

Monte Carlo simulation results for inverter and NAND gate reveals that variation in the diameter of tubes does not produce substantial variation in delay and static power consumption, therefore, not impacting the functional yield of the inverter and NAND gates. However, when we compare the functional yield of NAND gate with that of an inverter in the presence of metallic tubes, the metallic tubes have a more adverse impact on the functional yield of NAND gate than on the inverter. For example, for 10% metallic tubes and for the same drive strength of NAND gate and inverter, the yield of NAND gate is only 14% as compared to 67% for the inverter. On the other hand stacking configurations are more helpful in increasing the functional yield of complex gates like a NAND gate than an inverter. For example, for 10% metallic tubes the functional yield of 2-

78

input NAND gate *TuS* configuration increases by 5.5X of *PT* configuration as compared to 1.4X for inverter.

5.4   General Analytical Model for Yield

Since Monte Carlo simulations are computationally intensive, we have developed analytical models to quickly analyze the functional yield behavior of logic gates. As it is observed from Monte Carlo simulation results that for a delay constraint of 1.3X, all the gates that are functional have static power less than the maximum allowable static power constraint of 200X. Therefore, our analytical model derivation is based explicitly on the delay constraint and the power constraint is implied.

The analytical models compute the functional yield of gates on the basis of drive strength of a gate, number of parallel tubes in a gate, and percentage of metallic tubes. Here again the assumption is that all of the transistors are implemented with un-correlated CNTs. Please refer to the Appendix A for the symbols along with their definitions used in the derivation of analytical models.

If there are a finite number of metallic tubes, statistically, there will be a finite delay penalty compared to a gate with no metallic tubes due to contention current coming from the *OFF* network. A number of parallel tubes (N_{tur}) in a transistor is a parameter used in all models.

All analytical models are derived using the following procedure:

Step 1: Maximum number of metallic tubes tolerated in a network (N_m): Given a number of parallel tubes in a transistor and a value of the maximum acceptable delay ratio, X_{max}, we derive the expression for the maximum number of metallic tubes, N_m, that can be tolerated without

violating the acceptable delay penalty represented by the maximum delay ratio, X_{max}. It is assumed that for a semiconducting tube, the *ON* current is much larger than the *OFF* current, i.e. $I_{ons} >> I_{offs}$. For a metallic tube, the *ON* current is equal to the *OFF* current and both are equal to the *ON* current of a semiconducting tube, i.e. $I_{onm} = I_{offm} = I_{ons}$.

Step 2: Probability of PU/PD network being functional (Pr_{PU}/ Pr_{PD}): Given N_m, one can calculate the probability of pull-up, Pr_{PU}, and pull-down, Pr_{PD}, networks (pull-up/pull-down, Pr_{PU}/ Pr_{PD}) being functional by meeting the delay constraints. These probabilities depend on the type of a gate and on a tube configuration. Pr_{PU} and Pr_{PD} are functions of N_m, the maximum number of metallic tubes to be tolerated, N_{tur}, the number of tubes in a transistor and Pr_m, the probability of a tube being metallic. Probabilities of pull-up and pull-down networks to be functional are calculated by adding probabilities of a network being functional with a tolerable number of metallic tubes from zero to N_m as shown in equation(5.1).

$$\Pr{}_{PU/PD} = \sum_{i=0}^{N_m} (1 - \Pr{}_m)^{(N_{tur} - i)} \quad \Pr{}_m^i \quad {}^{N_{tur}}C_i \tag{5.1}$$

Where Pr_m^i is the probability of *i* out of N_{tur} tubes being metallic and $(1-Pr_m)^{(Ntur-i)}$ is the probability of $(N_{tur}-i)$ tubes being semiconducting. ${}^{N_{tur}}C_i$ $\binom{N_{tur}}{i}$ is the number of possible ways of *i* metallic tubes, being present among N_{tur} tubes.

Step 3: Functional yield of a gate (Y_f): A gate is considered functional if both the pull-up and pull-down networks are functional. The functional yield of a gate, Y_f, can be expressed as a product of the probabilities of both networks being functional as shown in equation(5.2).

$$Y_f = Pr_{PU} \times Pr_{PD} \tag{5.2}$$

5.5 Analytical Yield Model for Inverter

As the derivation procedure and final expressions for functional probabilities of PU (Pr_{PU}) and PD (Pr_{PD}) networks for an inverter are the same, $Pr_{PU}=Pr_{PD}$, therefore we present derivation for the functional probability of PU network only. We first derive the analytical model for the functional yield of PT inverter, and later extend it for two additional tube configurations, which we proposed in [99], TrS and TuS.

5.5.1 Parallel Tube

Step 1: Maximum number of metallic tubes that can be tolerated in PU/PD network of inverter (N_{m_Inv}): In case of an inverter, the pull-up and pull-down networks each consist of a single transistor. Therefore, the maximum number of metallic tubes that can be tolerated without violating the acceptable delay penalty can be obtained as shown in equation(5.3).

$$N_{m\_Inv} = \left\lfloor N_{tur}\left(1-\frac{1}{X_{max}}\right)\right\rfloor \tag{5.3}$$

Step 2: Probability of PU network of PT inverter being functional ($Pr_{PU\_Inv\_PT}$): Given N_{m_Inv}, the probability Pr_{PU_Inv} of the pull-up network being functional is calculated using equation (5.1) with N_m equal to N_{m_Inv} and N_{tur} equal to the actual number of tubes in a transistor as shown in equation(5.4).

$$Pr_{PU\_Inv\_PT} = \sum_{i=0}^{N_{m\_Inv}} (1-Pr_m)^{(N_{tur}-i)} \; Pr_m^i \;\; ^{N_{tur}}C_i \tag{5.4}$$

Step 3: Functional yield of an inverter with PT transistors ($Y_{f\_Inv\_PT}$): is obtained by substituting (5.4) into (5.2) as shown in equation(5.5). Each transistor in the inverter has N_{tur} tubes so the total number of tubes in the gate, N_{tug}, is $2N_{tur}$.

$$Y_{f\_Inv\_PT} = \left[\sum_{i=0}^{N_{m\_Inv}} (1 - Pr_m)^{(0.5N_{tug} - i)} \quad Pr_m^i \quad {}^{0.5N_{tug}}C_i \right]^2 \tag{5.5}$$

5.5.2 Transistor Stacking

The *TrS* configuration was proposed to reduce the probability of ohmic short between the power and the output of a transistor in the presence of metallic tubes. In stacking configurations, each transistor in both, the pull-up and pull-down networks are replaced with a stack of two transistors. Therefore, the functional probability of *PU* network depends upon the contention current coming from the *PD* network that consists of two stacked transistors N_1 and N_2, as shown in Figure 5-10. The *PU* network will be functional when *(a)* either *OFF* current of N_1 transistor is smaller than the maximum *OFF* current, I_{off_max} or *(b)* the *OFF* current of N_2 transistor is smaller than I_{off_max} or *(c)* the *OFF* current of both N_1 and N_2 is smaller than I_{off_max}. Where I_{off_max} is the maximum allowable current coming from the *PD* network, for which the *PU* network does not violate the maximum allowable delay constraint and remains functional.
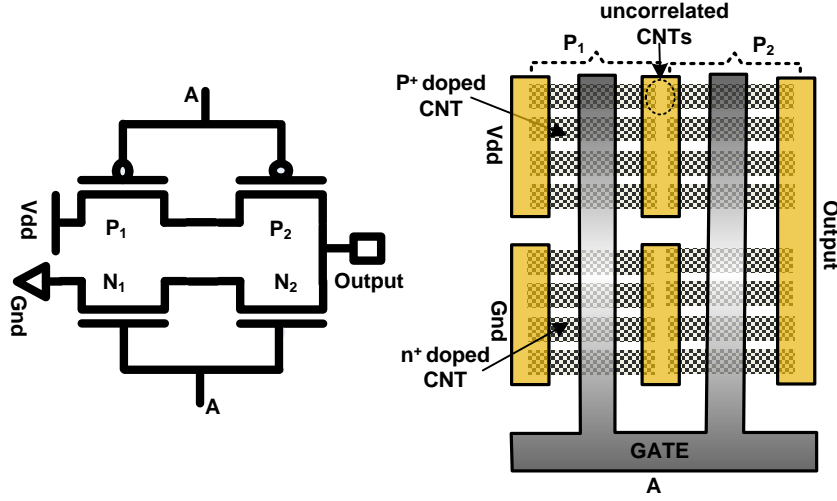
Figure 5-10: CNT based schematic and layout of inverter in which transistors in the pull-up and pull-down network are replaced by a stack of transistors.

In other words, the probability of *PU* network being functional is equal to the probability of N_1 transistor being functional plus the probability of N_2 transistor being functional minus the joint probability of both of transistors N_1 and N_2 being functional, as shown in (5.6). We assume the same functional probabilities for both N_1 and N_2 transistors.

$$Pr_{PU\_Inv\_TrS} = 2\,Pr_{PU\_Inv\_PT} - Pr_{PU\_Inv\_PT}^2 \qquad (5.6)$$

The functional probability of *PD* network depends upon the contention current coming from *PU* network that consists of two stacked transistors, P_1 and P_2, as shown in Figure 5-10. Since in case of inverter both the *PU* and *PD* are symmetrical, and the probability of *PD* network being functional can be obtained from (5.6) by substituting the functional probabilities of N_1 and N_2 transistors with functional probabilities of P_1 and P_2 transistors. We assume that a probability of an n-type transistor being functional is the same as the probability of a p-type transistor. The functional yield ($Y_{f\_Inv\_TrS}$) of *TrS* inverter can be

obtained by substituting the pull-up and pull-down network functional probabilities of *TrS* inverter gate into yield expression given by(5.2).

5.5.3 Tube Stacking

In *TuS* configuration, each parallel tube in a transistor is replaced with a stack of two tubes, called a double-stacked tube, as shown in Figure 5-2(b). For a double-stacked tube an ohmic short between source and drain contacts of a transistor can only happen when both tubes in a double-stacked tube are metallic. Therefore, the probability of a double-stacked tube to be metallic, Pr_{ms}, can be expressed as a product of a probability of one tube being metallic and the second tube being metallic as given in

$$\Pr_{ms} = \left(\Pr_m\right)^2$$

(5.7)

The maximum number of tubes, in a double-stacked configuration, that can be tolerated to be metallic without violating the acceptable delay penalty is obtained by replacing N_{tur} by N_{tusr} in(5.3). Here N_{tusr} is the total number of double-stacked parallel tubes in the transistor. The functional yield ($Y_{f\_Inv\_TuS}$) of an inverter designed with double-stacked tubes is obtained by replacing Pr_m with Pr_{ms} in(5.5).

$$Y_{f\_Inv\_TuS} = \left[\sum_{i=0}^{N_{m\_Inv}} (1 - \Pr_{ms})^{(0.5N_{tug} - i)} \quad \Pr_{ms}^{i} \quad ^{0.5N_{tug}}C_i \right]^2$$

(5.8)

5.5.4 Comparison between Monte Carlo and Analytical Model for Inverter

Figure 5-11 shows the comparison between the functional yield generated by Monte Carlo simulations and by using analytical models for an inverter with three discussed tube configurations. Analytical model results are shown with lines, while Monte Carlo results are

shown with symbols. Please observe the oscillatory nature of functional yield with respect to the number of tubes. The reason for the oscillatory nature is that by increasing the number of tubes in a transistor/gate, the probability of the presence of metallic tubes in the transistor/gate also increases, but the number of metallic tubes that can be tolerated can only increase in fixed intervals.
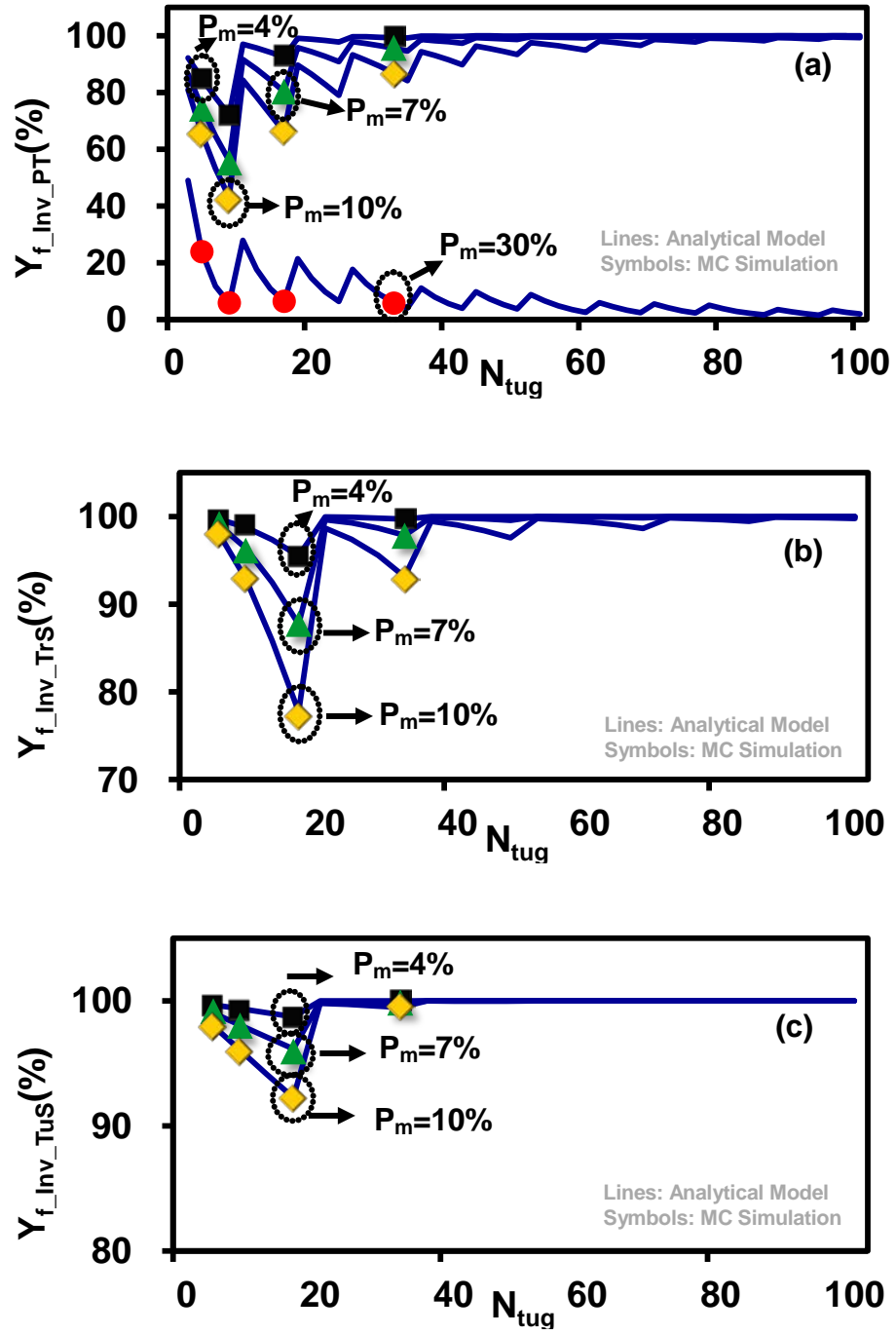
Figure 5-11: Functional yield, Y_{f_Inv} for (a) Parallel Tube (PT) (b) Transistor Stacking (TrS) (c) Tube Stacking (TuS), inverter as predicted by analytical model and Monte Carlo simulation for different drive strengths as measured by number of tubes in the inverter (N_{tug}) and for different percentage amount of metallic tubes (4%,7% and 10%) for allowed delay penalty of 1.3X.

For example, in the case of *PT* configuration with four tubes in a transistor no metallic tubes can be tolerated for the transistor/gate to be functional. Increasing the number of tubes from 4 to 8 will still not allow for any metallic tubes to be tolerated. A further increase in the number of tubes from 8 to 12, however, will allow for one metallic tube to be tolerated and consequently the functional yield will increase. For *PT* configuration with 4% to 10% of metallic tubes, the functional yield finally converges to almost 100% with an increase in the number of tubes. If 30% of metallic tubes are present, however, the functional yield asymptotically approaches 0% by increasing the number of the tubes as shown in Figure 5-11(a).

Table 5-2 shows absolute differences in functional yield magnitudes between Monte Carlo simulations and analytical models for different percentage of metallic tubes and different number of tubes in the inverter. In our experiments the range of absolute difference in functional yield magnitudes is between 0% to 0.9%, and absolute maximum error in functional yield is 0.9%, and it was recorded for *PT* inverter with $N_{twg}=8$ and 10% of metallic tubes. This small difference shows that our analytical model estimates the functional yield with excellent accuracy without going through computationally extensive Monte Carlo simulations.

Table 5-2: Absolute difference in functional yield magnitudes between Monte Carlo simulations and analytical model for different percentage of metallic tubes and different drive strengths of inverter. All numbers are in %. Maximum yield difference is 0.9% and minimum yield difference is 0%.

N_{tug}	Parallel Tube			Transistor Stacking			Tube Stacking		
	$P_m=4\%$	$P_m=7\%$	$P_m=10\%$	$P_m=4\%$	$P_m=7\%$	$P_m=10\%$	$P_m=4\%$	$P_m=7\%$	$P_m=10\%$
4	0.1	0.6	0.4	0.1	0.2	0.1	0.1	0.2	0.1
8	0.1	0.7	0.9	0.2	0.2	0.1	0.1	0.0	0.2
16	0.3	0.1	0.1	0.0	0.1	0.5	0.1	0.2	0.0
32	0.1	0.1	0.7	0.0	0.1	0.4	0.0	0.0	0.0

5.6 Analytical Yield Model for NAND Gate

The procedure for analytically finding the functional yield of NAND gate for different configurations of tubes in CNT transistors requires separate analysis of the pull-up and pull-down network for two reasons:

1. In the pull-up network, transistors are arranged in parallel and in the pull-down network transistors are arranged in series.

2. To make the worst case rise and fall delays equal, the number of tubes in the transistors in the pull-down network, N_{turN}, is twice the number of tubes in the transistors of the pull-up network, N_{turP}.

We follow the approach used for the inverter by first deriving the analytical model for the functional yield of *PT* NAND gate and later modifying the analytical model of *PT* NAND gate to develop models for *TrS* and *TuS* NAND gates.

5.6.1 Parallel Tube

PD network:

Step 1: Maximum number of metallic tubes that can be tolerated in PU network for the PD network to be functional in NAND gate (N_{mPU_NAND}): In a NAND gate, the functionality of the pull-down network depends upon the contention current coming from the pull-up network due to the presence of metallic tubes. We are assuming that two P-CNFETs, connected in parallel in the pull-up network, are equivalent to a single equivalent P-CNFET with twice the number of tubes of an individual P-CNFET, $2N_{turP}$. Therefore the maximum number of metallic tubes that can be tolerated in the pull-up network for the pull-down network to be functional (N_{mPU_NAND}) can be approximated by substituting N_{tur} with N_{turP} in (5.3) as shown in (5.9)

$$N_{mPU\_NAND} = \left\lfloor N_{turP}\left(1 - \frac{1}{X_{\max}}\right)\right\rfloor \tag{5.9}$$

Step 2: Probability of PD network of PT NAND being functional ($Pr_{PD\_NAND\_PT}$): In the *PT* NAND configuration the probability of the pull-down network to be functional can be calculated by substituting N_m with N_{mPU_NAND} and N_{tur} with $2N_{turP}$, in (5.1) as shown in (5.10)

$$Pr_{PD\_NAND\_PT} = \sum_{i=0}^{N_{mPU\_NAND}} (1 - Pr_m)^{(2N_{turP}-i)}\ \ Pr_m^i\ \ \ ^{N_{turP}}C_i \tag{5.10}$$

PU network:

Step 1: Maximum number of metallic tubes that can be tolerated in PD network for the PU network to be functional in NAND gate (N_{mPD_NAND}): The functionality of the pull-up network depends upon the contention current coming from the pull-down network that consists of two N-

CNFETs transistors connected in series, each with N_{turN} tubes. We consider the worst-case situations in which only one P-CNFET in the pull-up network is *ON* to pull the output node high, therefore only one N-CNFET in the pull-down network is *OFF*. Consequently, the maximum number of metallic tubes that can be tolerated in the pull-down network, N_{mPD_NAND}, for the pull-up network to be functional is calculated by substituting N_{tur} with $0.5N_{turN}$ in (5.3) as shown in (5.11)

$$N_{mPD\_NAND} = \left\lfloor \frac{N_{turN}}{2}\left(1 - \frac{1}{X_{max}}\right) \right\rfloor \tag{5.11}$$

Step 2: Probability of PU network of PT NAND being functional ($Pr_{PU\_NAND\_PT}$): The probability of the pull-up network being functional (Pr_{UP}) needs to be developed differently. We consider the worst case of low-to-high transition in which one P-CNFET is *ON* and other is *OFF*, in the pull-up network, and similarly one N-CNFET is *ON* and other is *OFF* in the pull-down network. Therefore, we need to consider two cases in the pull-down network. Either top N_1 transistor is *OFF* and bottom N_2 transistor is *ON* represented by P_{PU1_NAND}, or top N_1 transistor is *ON* and bottom N_2 transistor is *OFF* represented by P_{PU2_NAND}. If we assume that both cases are equally possible then the probability of the pull-up network being functional can be expressed as in (5.12)

$$Pr_{PU\_NAND} = Pr_{PU1\_NAND} \times Pr_{PU2\_NAND} \tag{5.12}$$

We can further assume that the worst case probabilities of the pull-up network being functional (P_{UP1_NAND}) and (P_{UP2_NAND}) are the same and both can be calculated using (5.1) by substituting N_m with N_{mPD_NAND} and N_{tur} with N_{turN}, as shown in

$$
\begin{aligned}
Pr_{PU1\_NAND\_PT} &= Pr_{PU2\_NAND\_PT} \\
&= \sum_{i=0}^{N_{mPD\_NAND}} (1 - Pr_m)^{(N_{turN} - i)} \quad Pr_m^i \quad {}^{N_{turN}}C_i
\end{aligned}
\tag{5.13}
$$

The overall expression for the pull-up network being functional is given by substituting (5.13) into (5.12) as shown in (5.14)

$$
Pr_{PU\_NAND\_PT} = \left[\sum_{i=0}^{N_{mPD\_NAND}} (1 - Pr_m)^{(N_{turN} - i)} \quad Pr_m^i \quad {}^{N_{turN}}C_i \right]^2
\tag{5.14}
$$

Step 3: Functional yield of NAND gate with PT transistors ($Y_{f\_NAND\_PT}$): The functional yield $Y_{f\_NAND\_PT}$ of PT NAND gate, shown in (5.15), is obtained by substituting (5.10), $Pr_{PD\_NAND\_PT}$, and (5.14),$Pr_{PU\_NAND\_PT}$, into yield expression given in (5.2). Since each transistor in the pull-up network of NAND gate has N_{turP} tubes and each transistor in the pull-down network has N_{turN} tubes, the total number of tubes in the NAND gate, N_{tug} will be $2N_{turP} + 2N_{turN}$.

$$
\begin{aligned}
Y_{f\_NAND\_PT} &= \left[\sum_{i=0}^{N_{mPD\_NAND}} \begin{array}{c} (1 - Pr_m)^{(N_{tug}/3 - i)} \\ \times Pr_m^i \quad {}^{(N_{tug}/3)}C_i \end{array} \right]^2 \\
&\times \left[\sum_{i=0}^{N_{mPU\_NAND}} \begin{array}{c} (1 - Pr_m)^{(N_{tug}/3 - i)} \\ \times Pr_m^i \quad {}^{(N_{tug}/3)}C_i \end{array} \right]
\end{aligned}
\tag{5.15}
$$

91

5.6.2 Transistor Stacking

The functional yield of *TrS* NAND gate is derived using the same reasoning as explained for inverter with *TrS* configuration. The functional probability of *PU* network depends upon the contention current coming from the pull-down network in which all the transistors are replaced with the stacked transistors. We again consider three cases of network being functional; (a) either the first stacked transistor *OFF* current is smaller than the maximum *OFF* current, I_{off_max} or (b) the *OFF* current of the second stacked transistor is smaller than, I_{off_max} or (c) the *OFF* current of both stacked transistors is less than I_{off_max}. Since the functional probabilities of both stacked transistors in the network are non-exclusive, the functional probability of the pull-up network of *TrS* NAND gate is given by (5.16)

$$\text{Pr}_{PU\_NAND\_TrS} = 2\,\text{Pr}_{PU\_NAND\_PT} - \text{Pr}^2_{PU\_NAND\_PT} \tag{5.16}$$

The functional probability of *PD* network of *TrS* NAND gate is obtained using (5.16) by substituting the functional probability of *PT* pull-down network of NAND gate given in(5.10). To calculate the functional yield of *TrS* NAND gate we substitute the pull-up and pull-down network functional probabilities of *TrS* NAND gate into yield expression given by(5.2).

5.6.3 Tube Stacking

As in the case of inverter with *TuS* configuration, the probability of double-stacked tubes to be metallic can be obtained from(5.7). The maximum number of double-stacked metallic tubes that can be tolerated in the pull-up and pull-down networks can be obtained from (5.9)and (5.11)by replacing N_{turP} with N_{tusrP} and N_{turN} with N_{tusrN}, respectively. Where N_{tusrP} is

the number of stacked tubes in P-CNFET of NAND gate and N_{tusrN} is the number of stacked tubes in N-CNFET of NAND gate. The functional yield of *TuS* NAND gate ($Y_{f\_NAND\_TuS}$) is obtained by replacing Pr_m with Pr_{ms} in(5.15).

$$Y_{f\_NAND\_TuS} = \left[\sum_{i=0}^{N_{mPD\_NAND}} (1-\mathrm{Pr}_{ms})^{(N_{tug}/3 - i)} \times \mathrm{Pr}_{ms}^{i} \ {}^{(N_{tug}/3)}C_i \right]^2$$
$$\times \left[\sum_{i=0}^{N_{mPU\_NAND}} (1-\mathrm{Pr}_{ms})^{(N_{tug}/3 - i)} \times \mathrm{Pr}_{ms}^{i} \ {}^{(N_{tug}/3)}C_i \right]$$

(5.17)

The analytical model derivation procedure for 2-input NOR gate is exactly the same as that of 2-input NAND gate as the NOR gate is a dual of NAND gate. The expressions for the functional yield of the pull-up and pull-down networks are switched, but the functional yield of the NOR gate is the same as that of NAND gate. Analytical models for the functional yield of logic gates with larger fan-in and for other complex logic gates can be derived in a similar manner.

5.6.4   *Comparison between Monte Carlo and Analytical Model for NAND Gate*

Figure 5-12 shows the functional yield comparison between Monte Carlo simulations and the analytical model results for 2-input NAND gate. Results from analytical models are shown with lines and Monte Carlo simulation results are shown with symbols. It can be observed that for *PT* NAND gate with 10% metallic tubes the functional yield saturates at around 35% which is very low and not good enough for robust CNT based circuits. For 4% metallic tubes the functional yield of *PT* NAND asymptotically approaches 90% by

increasing the number of tubes in the gate. *TuS* configuration proves very helpful in increasing the functional yield of NAND gate in the presence of metallic tubes as shown in Figure 5-12(c). For 4% and 10% metallic tubes the functional yield approaches almost 100% with the increase in drive strength of the gate represented by increasing the number of tubes in the gate.
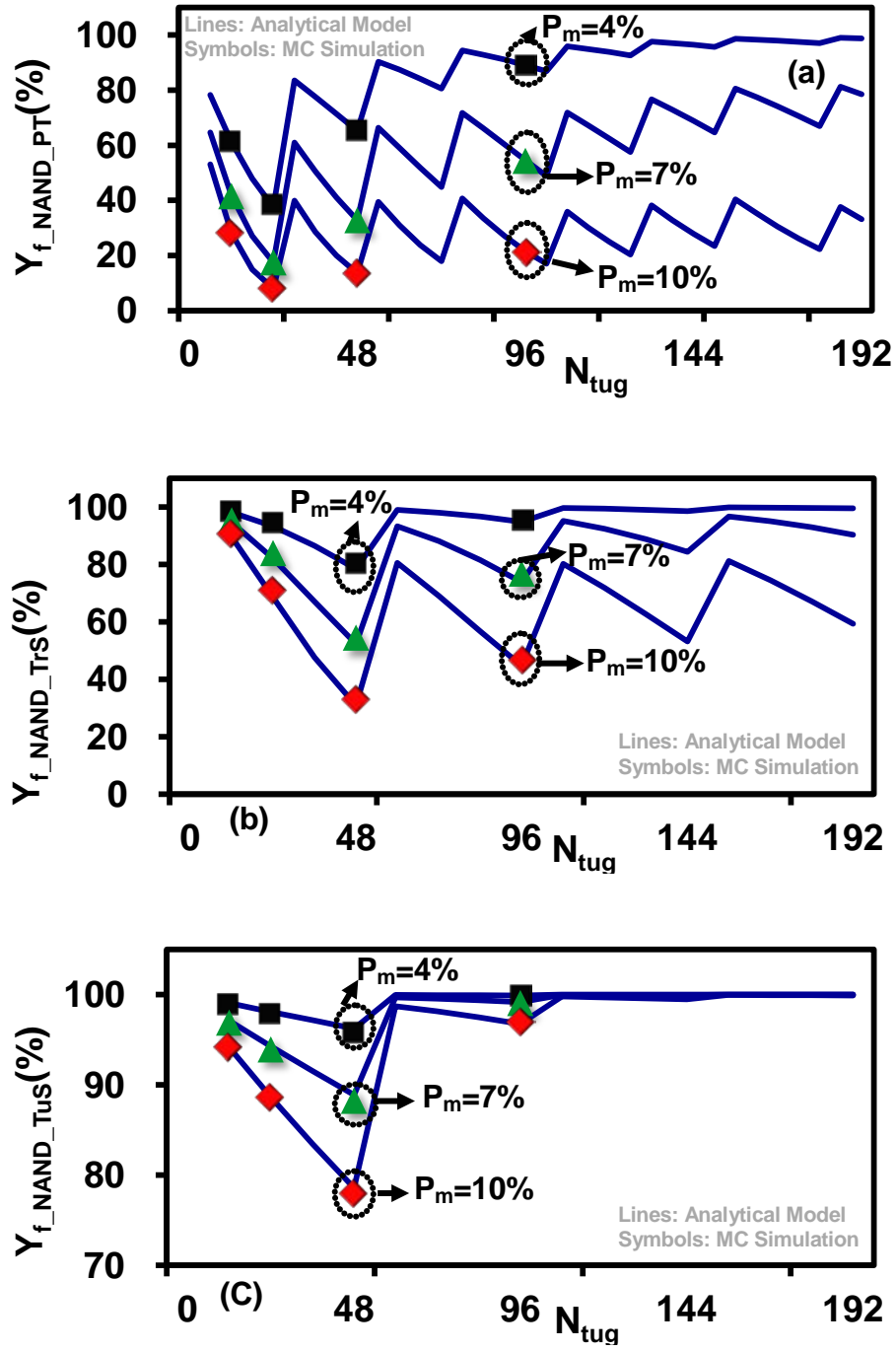
Figure 5-12: Functional yield, Y_{f_NAND}, for (a) Parallel Tube (PT) (b) Transistor Stacking (TrS) (c) Tube Stacking (TuS) NAND gate as predicted by analytical model and Monte Carlo simulation for different drive strengths as measured by number of tubes in the NAND gate (N_{tug}) and for different percentage amount of metallic tubes (4%,7% and 10%) for allowed delay penalty of 1.3X.

Table 5-3 shows the absolute differences in functional yields between data obtained from Monte Carlo simulations and the analytical models. Results are reported for different percentage of metallic tubes and different numbers of tubes in the NAND gate. In the table, the range of absolute error in functional yield is between 0% and 2.5%, and the maximum error in functional yield for NAND gate is 2.5%. The maximum error is observed for *TrS* configuration NAND gate with N_{tug}=48 and for 10% metallic tubes. Very small differences in yield numbers show that we can accurately predict the functional yield of NAND gate analytically without going through computationally expensive Monte Carlo simulations. The maximum error of 2.5% obtained for NAND gate as compared to 0.9% for the inverter is because of the complexity of the NAND gate as compared to the inverter.

Table 5-3: Absolute difference in functional yield magnitudes between Monte Carlo simulations and analytical model for different percentage of metallic tubes and different drive strengths of 2-input NAND gate .All numbers are in %. Maximum yield difference is 2.5% and minimum yield difference is 0%.

N_{tug}	Parallel Tube			Transistor Stacking			Tube Stacking		
	P_m=4%	P_m=7%	P_m=10%	P_m=4%	P_m=7%	P_m=10%	P_m=4%	P_m=7%	P_m=10%
12	0.1	0.3	0.3	0.3	0.8	1.0	0.1	0.2	0.0
24	1.0	0.0	0.3	1.2	1.9	2.4	0.2	0.3	0.0
48	0.1	0.1	0.0	2.2	2.2	2.5	0.3	0.6	0.5
96	0.3	0.3	0.6	0.7	1.8	2.3	0.0	0.1	0.1

5.7   Configuration Comparison Summary

Table 5-4 and Table 5-5 summarize the normalized mean delay, static power and yield results for inverter and NAND gate obtained from Monte Carlo simulations for different transistor configurations and different gate delays as measured by N_{tug}, the number of tubes in the gate. The percentage of metallic tubes, P_m, is considered to be 4%, the minimum percentage of unwanted metallic tubes reported by [92]. It is worth noting that the same trend in functional yield is observed for different configurations of the inverter when 4% or 10% tubes are metallic. D_{μ_Inv} and D_{μ_NAND} captures the mean delay of the inverter and NAND gate and SP_{μ_inv} and SP_{μ_NAND} captures the average static power (also a measure of total static power) of the inverter and NAND gate.

From Table 5-4 and Table 5-5 it is clear that there is no single favorite configuration to be chosen for the best of delay, static power and yield. Either *ST* or *PT* configurations should be chosen when delay is the primary objective – this would apply to critical paths. Either *TrS* or *TuS* configurations should be chosen when yield and static power are the primary objectives – this would apply to non-critical paths. The choice between *TrS* and *TuS* depends on how precisely the technology allows alignment of the tubes in the *TuS* configuration. If the advancement in the technology allows alignment of the tubes with nano-scale precision then *TuS* will give us a marginally higher yield than *TrS* configuration of transistors in the presence of metallic tubes. On the other hand if the technology does not allow the precise alignment of tubes, then the *TrS* configuration will be the optimal choice to handle the metallic tubes. The better choices for delay, static power and yield for various N_{tug} are highlighted in Table 5-4 and Table 5-5. Please note that for a given value of N_{tug} all configurations have iso-input capacitance. An architecture that utilizes an

appropriate combination of proposed configurations and possible hybrids between them is required to enable a better trade-off between delay, power, and yield when the percentage of metallic tubes is small.

Table 5-4: Normalized mean delay, normalized static power and yield comparisons for inverter. Optimal choices for delay, static power and yield for a given value of N_{tug} are highlighted.

P_m	Configuration	Inverter				
		N_{tug}	D_{μ_Inv}	SP_{μ_Inv}	$Y_{f\_Inv}(\%)$	
4%	**Shared Tube**	4	**1.1**	**7.3**	**92.3**	←
	Parallel Tube		1.1	3.7	81.7	
	Transistor Stacking		4.2	0.9	99.4	
	Tube Stacking		**4.2**	**0.9**	**99.4**	←
	Shared Tube	8	**1.1**	**7.3**	**92.3**	←
	Parallel Tube		1.0	2.7	72.3	
	Transistor Stacking		4.2	0.7	99.0	
	Tube Stacking		**4.2**	**0.6**	**99.2**	←
	Shared Tube	32	1.1	7.3	92.3	
	Parallel Tube		**1.1**	**60.7**	**99.4**	←
	Transistor Stacking		4.2	5.1	99.7	
	Tube Stacking		**4.1**	**1.2**	**100.0**	←

Table 5-5: Normalized mean delay, normalized static power and yield comparisons for 2-input NAND gate. Optimal choices for delay, static power and yield for a given value of N_{tug} are highlighted.

P_m	Configuration	NAND			
		N_{tug}	D_{μ_NAND}	SP_{μ_NAND}	$Y_{f\_NAND}(\%)$
4%	**Shared Tube**	12	**1.1**	**4.9**	**84.8** ←
	Parallel Tube		1.0	2.6	61.4
	Transistor Stacking		4.2	0.6	98.5
	Tube Stacking		**4.2**	**0.6**	**98.9** ←
	Shared Tube	24	**1.1**	**4.9**	**84.8** ←
	Parallel Tube		1.0	2.3	38.6
	Transistor Stacking		4.1	0.6	94.6
	Tube Stacking		**4.1**	**0.5**	**97.9** ←
	Shared Tube	96	1.1	4.9	84.8
	Parallel Tube		**1.2**	**43.1**	**88.8** ←
	Transistor Stacking		4.4	3.5	95.5
	Tube Stacking		**4.1**	**0.8**	**99.9** ←

For multiple-stage logic networks a certain level of statistical averaging in delay and power variation can be observed depending on the logic depth. Therefore, even if performance and power of some individual gates do not meet the specification, the circuit can still function properly due to statistical averaging. Here we assume that based on the defined limits of delay and static power the degradation of noise margin of logic gates in the presence of metallic tubes will be tolerable and will not result in logic failure as the signal can be restored in traversing through multistage logic networks. We will show the statistical averaging impact later in this chapter by analyzing the yield of full-adder [114] and 3-input functions implemented with regular logic bricks [114].

5.8   Architecture Solution

5.8.1   Design of Full Adder

A full adder is implemented using inverters, 2-input NANDs and 2:1 MUX as shown in

Figure 5-13. The 2:1 MUXs are implemented with an inverter and 2-input NAND gates.

Here MUX (M_1) and inverter (I_1) implement propagate function $P = (a \oplus b)$ and MUX

(M_2) and inverter (I_2) implement the sum function. Similarly, NAND (N_1) and inverter (I_3)

implement the generate function $G = ab$. The MUX (M_3) implements the carry function

$c_0 = G + Pc_i$. Please notice that generate and propagate are only functions of inputs a and

b, and are independent of input c_i. For multi-bit adders only gate M_3 is on the critical path.

As it was mentioned before, gates implemented with *TrS* configuration are 4X slower than

corresponding gates implemented with *PT* configuration. Therefore, using *TrS*

configuration gates in the adder to increase the yield significantly increases the adder delay.

However, since only gate M_3 is on the critical path, the yield of the adder can be increased

by implementing the adder with *TrS* configuration and using parallelism in the critical path

of the adder to mitigate the impact on the delay.
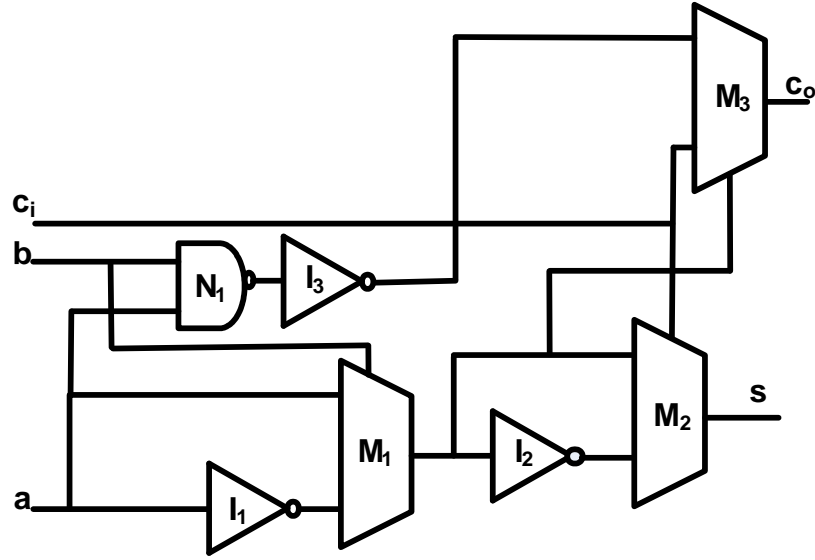
Figure 5-13: Schematic diagram of full adder implemented using inverters, 2-input NAND and 2:1 MUX.

5.8.2 Adder using TrS Configuration and Parallelism

We substitute MUX (M_3) with four parallel instances of M_3 to keep the delay of adder the same as the delay of the *PT* configuration adder. The advantage of using this design is a higher yield and much lower static power. The trade-offs are extra area because of three additional *M3* gates and additional dynamic power dissipation.

To verify advantages of the *TrS* adder with parallelism in the critical path, we performed Monte Carlo simulation of n=1,000 full-adders with a different percentage of metallic tubes and variation of tube diameters from 1nm to 2nm. Table 5-6 shows the yield results of adders implemented with *PT* and *TrS* configuration with parallelism in the critical path. Significant improvements in yield are observed. For example, for 7% metallic tubes the yield for *TrS* parallelized configuration of adder is 98% as compared to 54% for the *PT* adder.

Table 5-6: Functional yield of full adder implemented with Parallel Tube (*PT*) and Transistor Stacking (*TrS*) configurations.

Configuration	Pm=0%	Pm=4%	Pm=7%	Pm=10%
Parallel Tube	100.0	90.3	53.9	18.5
Transistor Stacking	100.0	100.0	98.1	67.5

Figure 5-14 shows the comparison between normalized delays for adders implemented with *PT* and parallelized *TrS* configuration. Please observe that for P_m=0% the same delay is obtained from both configurations, however, when the percentage of metallic tubes increases, the increase in the mean delay is less for parallelized *TrS* configuration as compared to *PT*. For example for 7% metallic tubes the increase in the mean delay for *PT* configuration is 30% as compared to 10% for adders with parallelized *TrS* configuration.
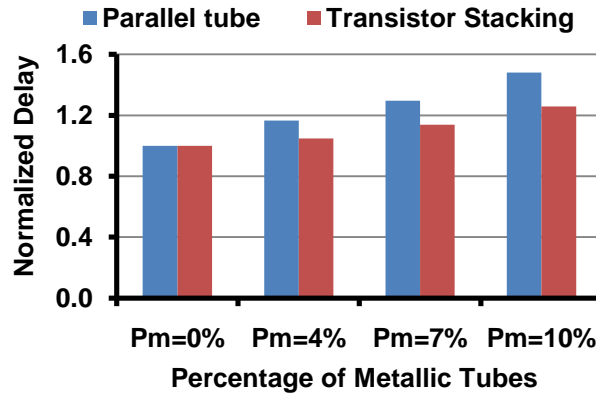
Figure 5-14: Comparison of normalized delay of full adders using *PT* and parallelized *TrS* configurations for varying percentage of metallic tubes.

The static power evaluation is based on the total current flowing in the *OFF* network of logic gates. This current has two components; the sub-threshold leakage current, and the current flowing because of the presence of metallic tubes. It can be easily observed that the contribution of the current due to the presence of metallic CNTs is much higher than the

subthreshold leakage current. To evaluate the static power of the adder we are considering the worst-case scenario by taking into account that the network with the higher static power is *OFF*. This assumption is to capture the total *OFF* current due to metallic tubes. The total static power is then obtained by adding the static power of all the gates in the adder as given by

$$SP_{add} = \sum_{i=1}^{3} SP_{Ii} + \sum_{i=1}^{3} SP_{Mi} + SP_{N1} \tag{5.18}$$
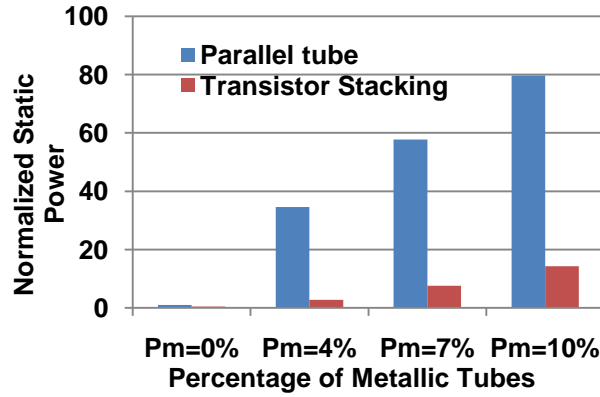
Figure 5-15: Comparison of static power consumption of full adders using *PT* and parallelized *TrS* configurations for varying percentage of metallic tubes.

Figure 5-15 shows the comparison between static power consumption for adders implemented with *PT* and parallelized *TrS* configurations. Similarly, a smaller increase in power consumption for an increased percentage of the metallic tubes is observed when the parallelized *TrS* configuration is used. For the case of 7% metallic tubes, the increase in static power is only 8% as compared to 58% for *PT* configuration, constituting almost 8X improvement.

5.8.3   *Via Configurable Logic Bricks*

Circuits fabricated using CNFETs have lithographic related variations in addition to the imperfections specific to the CNT technology. Therefore, for designs specific to ASIC

implementations we are proposing via-configurable regular logic blocks approach [107]. There are two main advantages of implementing the designs with regular logic blocks; *(1)* a reduction in the systematic process related variations in nanoscale technologies, and *(2)* acceptable levels of yield can be obtained by using redundant logic blocks for larger percentage of metallic tubes. The first advantage allows designers to focus only on challenges associated particularly with the CNT based technology, and the second allows for replacement of non-functional blocks that fail to meet the delay and power constraints. Among all the logic primitives used in [107] they are able to implement all three-input functions by using NAND, 2:1 MUXs, inverters and buffers. A finite small set of via-configurable logic blocks can be well tuned for manufacturability and performance. Figure 5-16 shows the schematic diagram of one of the five unique bricks used in [107] to implement 80 unique 3-input functions. In Figure 5-16, the top input of MUX (M_2) will be $(ab)'$ or $(a'b)'$ and similarly the lower input of MUX (M_2) will be $(ab' + a'b)$ or $(ab' + a'b)'$ based on which via is configured. The possible four 3-input functions implemented with the brick are shown below.

$$O_1 = \begin{cases} b + \overline{\overline{a}b}c & ; \\ ab\overline{c} + c(a \oplus b) & ; \\ \overline{ab} + \overline{a}\overline{b}c & ; \\ \overline{a}\overline{b}\overline{c} + c\left(\overline{a \oplus b}\right) & ; \end{cases} \tag{5.19}$$
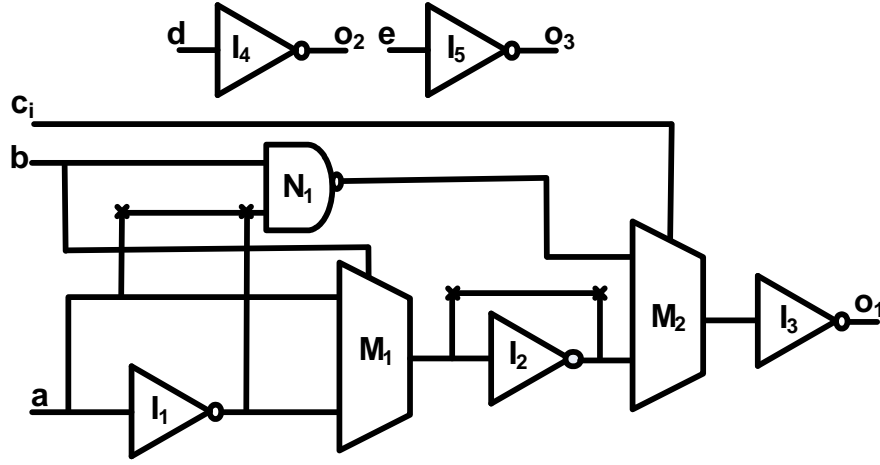
Figure 5-16: Sample brick to implement 3-input function [107].

Inverters I_4 and I_5 are used in the brick to make the foot print of this brick identical to that of D flip-flop with scan. These inverters can also be used for buffering of local and global signals. The inverters are of the minimum size and NAND gates are sized to have the same delay as that of inverter. The flowchart of Monte Carlo simulation setup used to obtain the delay, static power and functional yield of brick is provided in Section 10.3.1 of Appendix B. There are two possible delay paths for the brick shown in Figure 5-16. A path delay of the logic brick is obtained by adding the delays of all logic gates on the path as given in(5.20). The static power of the brick is obtained in the same way as the static power of the full adder in Section 5.8.2.

$$D_{P1} = \sum_{i=1}^{3} D_{Ii} + \sum_{i=1}^{2} D_{Mi}$$

$$D_{P2} = \sum_{i=1, i \neq 2}^{3} D_{Ii} + D_{Ni} + \sum_{i=2}^{2} D_{Mi}$$

(5.20)

The bricks are considered functional if their delays in the presence of metallic tubes are less than 1.3X of the delay of the fastest brick and the static power does not exceed 100X of the

lowest static power of a brick. The fastest brick with the lowest static power is a brick with all tubes being semiconducting.

5.8.4   Monte Carlo Simulation of Bricks

Functional yields of bricks are obtained through Monte Carlo simulation for n=1,000 bricks build of 2-input NAND, 2:1 MUXs, inverters and buffers. The gates in the bricks are implemented with *PT*, *TrS* and hybrid tube configurations. The hybrid configuration is composed of gates with *PT* gates in delay critical paths, *TrS* gates where delay is less important than power and yield contribution. For simulation purposes a diameter variation of 1nm to 2nm is considered [75] and the percentage of metallic tubes is varied between 0% (all semiconducting) to 10%.

Initially, all setups consist of homogeneous bricks implemented with CNFETs using only *PT* and *TrS* configurations of transistors used in all the logic gates inside the brick. Functional yield, delay and static power of a brick are obtained as a function of different percentages of the metallic tubes. The number of CNTs (represented by N_{tur}) used in N-CNFET and P-CNFET of the inverter in the Parallel Tube configuration is 8.

Table 5-7 shows the functional yield results obtained for bricks implemented with CNFETs using *PT*, *TrS,* and *PT-TrS* hybrid configurations and the percentage of metallic tubes varied from 0% to10%. It can be observed that the yield of bricks implemented with *PT* configuration drops significantly when the percentage of metallic tubes is 7% and gets to an extremely low level of 22% for 10% of metallic tubes. Results also show that *TrS* configuration improves the functional yield of bricks in the presence of metallic tubes. For 10% metallic tubes the functional yield obtained from *TrS* configuration is 81% as compared to 22% when the brick is implemented with *PT* configuration.

Table 5-7: Functional yield of bricks implemented with parallel tube (*PT*), transistor stacking (*TrS*) and hybrid configurations for different percentage of metallic tubes.

Brick Configuration	Pm=0%	Pm=4%	Pm=7%	Pm=10%
Parallel Tube	100.0	99.6	77.5	21.6
Transistor Stacking	100	100.0	97.7	80.5
Hybrid	100	99.6	87.5	53.5

In hybrid configurations some of the gates are implemented with *PT* transistors and others with *TrS* configuration. In the brick shown in Figure 5-16, we consider two critical paths $I_1 \rightarrow M_1 \rightarrow I_2 \rightarrow M_2 \rightarrow I_3$, and $I_1 \rightarrow N_1 \rightarrow M_2 \rightarrow I_3$. The larger of the two delays will define the delay of the brick, therefore, our hybrid-brick design strategy is to improve yield and power of the brick by using *TrS* transistors everywhere where the delay of the brick will not increase or will constitute a desirable trade-off. Please notice that gates I_1, M_2 and I_3 are common to both paths, so, to increase the yield we will implement these three gates with *TrS* transistors. Also the delay of $M_1 \rightarrow I_2$ in the first path is much larger than the delay of N_1 in the second path, hence we can allow for an increase in N_1 delay and it is implemented with *TrS* configuration. The hybrid configuration provides a compromise between the two configurations by obtaining yield, which is higher than for *PT* and lower than that for *TrS* configurations. The hybrid-brick delay is higher than obtained from *PT* and lower than that obtained from *TrS* configuration.

Figure 5-17 shows the delay of different brick configurations normalized with respect to the delay of brick implemented with *PT* configuration in the presence of varying percentage of metallic tubes. It can be observed that when comparing to *PT* bricks, that have the lowest yield and smallest delay, the hybrid configuration, showing significantly improved yield, also

reduces the delay increase to almost 2.2X as compared to 4X for homogenous *TrS* implementation.
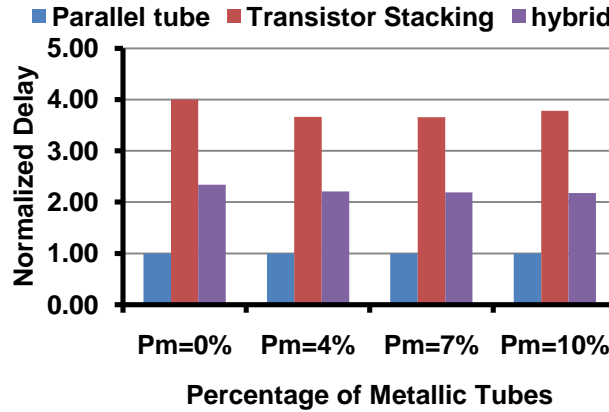
Figure 5-17: Delay of different configurations of bricks normalized to delay of a brick implemented with parallel tube configurations when percentage of metallic tubes are 0%,4%,7% and 10%.

Figure 5-18 shows the Static Power consumption of different configurations of bricks normalized to the static power of a brick implemented with *TrS* Stacking configurations when percentage of metallic tubes are 0%,4%,7% and 10%. From the figure it can be observed that for bricks implemented with *PT* configuration, the static power will increase from 40X up to 55X for 4% and 10% metallic tubes respectively. In the case of hybrid approach the static power is increased by 18X to 22X as compared to bricks implemented with *TrS* configurations but it is three orders of magnitude less when the bricks are implemented with *PT* configuration.
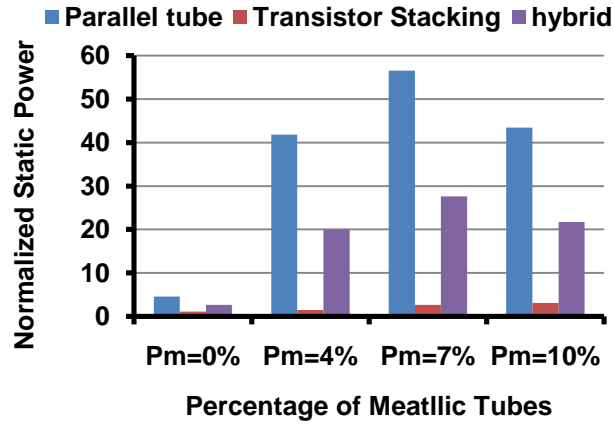
Figure 5-18: Static power consumption of different configurations of bricks normalized to the static power of brick implemented with Tube Stacking configurations when percentage of metallic tubes are 0%,4%,7% and 10%.

5.9   Conclusion

The undesired presence of metallic tubes is one of the major technological barriers faced by CNT technology that hinders the development of robust CNT based circuits for real applications. Both proposed CNFET stacking configurations are helpful in increasing the yield of CNT-based circuits in the presence of the metallic tubes, but the trade-off is an almost 4X increase in the delay. Moreover, the development of analytical models is helpful in quick analysis of finding the impact of different percentage of metallic tubes for different drive strength of logic gates. In this chapter we showed that by using innovative design methodology, we can leverage advantages of CNFET transistor stacking configurations to design high yield systems with low power dissipation and with delays comparable to most efficient Parallel Tube configuration. The performance degradation of high-yield achieving stacking configurations can also be minimized by exploiting architecture level techniques like parallelism or implementation of circuits with hybrid configuration of CNFETs.

6   Processing Techniques for Metallic Tubes

Part of this chapter has been published in Rehman Ashraf, Malgorzata Chrzanowska, Siva G. Narendra," Yield Enhancement by Tube Redundancy in CNFET-based Circuits", ICECS,2010

and part of this chapter is ready for submission in Rehman Ashraf, Malgorzata Chrzanowska, Siva G. Narendra," Yield Enhancement techniques for CNFET based Circuits in the Presence of Imperfections" , *IEEE Trans. Nanotechnology, 2011*

In the previous chapter, we assumed that all metallic tubes are present and we analyzed the impact of the presence of metallic tubes on the functional yield of gates and circuits. It was observed that the presence of metallic tubes has a detrimental impact on both the delay and static power consumption of the gates. Extra processing techniques must be used as described in Chapter 2, if the percentage of metallic tubes is larger than 7%. In this chapter we focus on yield enhancement of CNFET based gates and circuits in the presence of metallic tube removal by post processing techniques such as SCE or VMR.

6.1   Impact of Tube Removal Process

The removal of tubes by these extra processing steps increases the delay of CNFET based gates, and results in large variability in the performance and power of CNFET based devices. Furthermore, in the worst case, all tubes from a transistor can be removed and an open-circuit gate is created. The probability of open circuit CNFET based devices has been analyzed in [115]. In this chapter, we present the impact of extra processing techniques on *delay*, *power,* and *functional yield* of complementary CNFET based circuits.

6.2 Impact of Tube Correlation on the Functional Yield

It has been shown that in the presence of metallic tubes and due to their removal, the correlation among CNTs of different transistors, and pull-up and pull-down networks of logic gates has a strong impact on the functional yield of gates [99], [115]. It has been shown in [116] that when metallic tubes are present the use of highly un-correlated (different) tubes among different transistors reduces the probability of ohmic short, and increases the functional yield of logic gates.

Now if the extra processing techniques such as SCE and VMR are used to remove the metallic tubes, the Monte Carlo results show that both techniques remove more than 99.9% of metallic tubes. The trade-off of using these techniques is large performance variation due to removal of metallic and semiconducting tubes. It is observed that when the tubes are removed, the use of highly un-correlated tubes (different tubes) among different CNFETs results in large variation in performance, and hence low functional yield. On the other hand, when the tube removal process removes almost all the metallic tubes, then use of highly correlated tubes results in less variation in performance and high functional yields of CNFET based gates. Figure 6-1 shows the schematic and layout of a CNFET based inverter in which the same set of tubes have been used for pull-up and pull-down network. Thus, these CNTs are highly correlated in the pull-up and pull-down network.
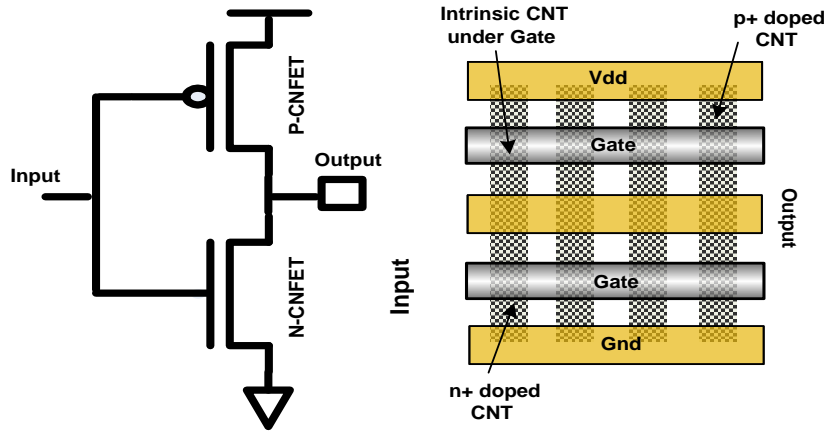
Figure 6-1: Schematic and layout of an inverter containing array of four CNTs in P-CNFET and N-CNFET. Highly correlated tubes are used in the pull-up and pull-down networks.

However, for complex logic gates like NAND and NOR, if highly correlated tubes are used in both pull-up and pull-down networks of the gate, then it results in irregular layout and increases the area of the gates which in turn makes the gates slower. In this work, we assume that for complex gates highly correlated tubes are used within the transistors of the individual pull-up and pull-down networks. But CNTs used in the pull-up and pull-down networks are un-correlated, *i.e. pull-up and pull down networks do not share tubes with each other.*

Figure 6-2 shows the schematic and layout of a 2-input NAND gate. It can be observed from the layout of Figure 6-2, that transistors within the pull-up network share the same tube (*tubes connected with the* V_{dd} *rail*). Thus tubes used by the transistors of the pull-up network are highly correlated. The same observation can be made for the transistors in the pull-down network. However the pull-up and pull down networks do not have any common tubes between them. Thus pull-up and pull down networks have highly un-correlated tubes.
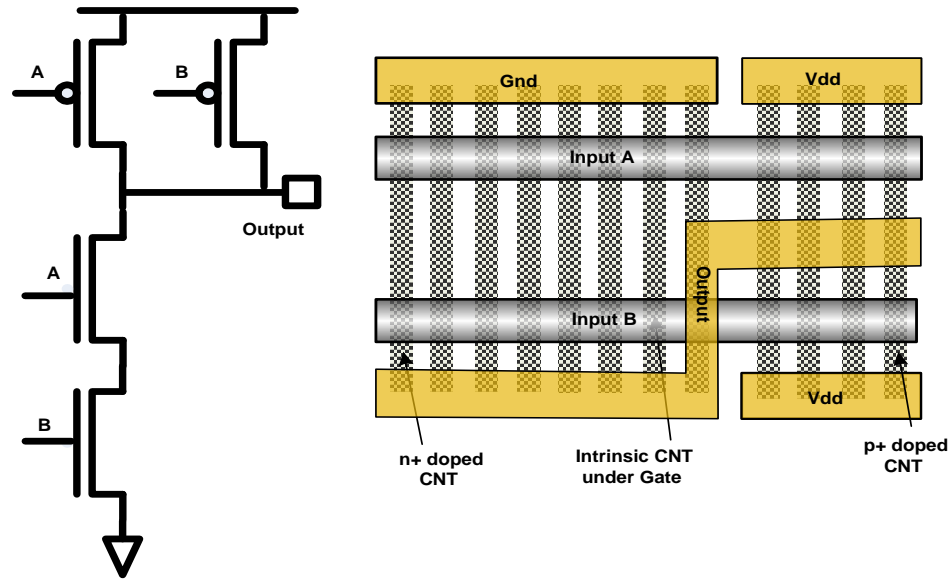
Figure 6-2: CNT based schematic and layout of 2-input NAND gate containing an array of four CNTs in P-CNFETs and an array of eight CNTs in N-CNFETs. Pull-up and pull-down networks are implemented with un-correlated tubes and transistors within pull-up and pull-down networks are implemented with highly correlated tubes.

During the analysis presented in [117], the authors assumed fanout of the logic gates to be constant. However, this scenario is only applicable while driving internal or external interconnect buses. In most cases, the gates will be driving other gates through local interconnects. In this case, tubes removed from both the driving gates and fanout gates will impact the performance of gates. If we consider the fanout to be constant then it will result in the underestimation of the functional yield of logic gates. Therefore, in this work we study the effect of fanout on the functional yield of logic gates.

6.3 Monte Carlo Simulation Results

Monte Carlo simulations are used to generate functional yield for an inverter and NAND gate built with CNFETs. The flowchart of Monte Carlo simulation setup used to obtain the delay, static power and energy of logic gates is provided in Section 10.4.2 of Appendix B. The yield results obtained from Monte Carlo simulations will be used to validate the yield results obtained from the analytical models developed in the subsequent sections of this chapter.

Figure 6-3 shows the Monte Carlo simulation yield summary for inverters *(with delay constraint of 1.3X and , static power constraint of 10X)* as a function of number of tubes in the gate (N_{tug}),the percentage of metallic tubes(P_m) prior to the application of the tube removal process, and the percentage of metallic and semiconducting tubes removed *(P_r)*. Here FO_1 load is considered, and tubes are removed from both the driving gate and fanout *(realistic scenario as discussed in the previous paragraph)*.Sample size (*n*) of 10,000 was used for all Monte Carlo simulations. The inset graph of Figure 6-3 shows the Monte Carlo simulation yield for inverters when it is assumed that no tubes are removed from the fanout. From the figure, it can be observed that extremely low yields are obtained when we consider the fanout to be constant.
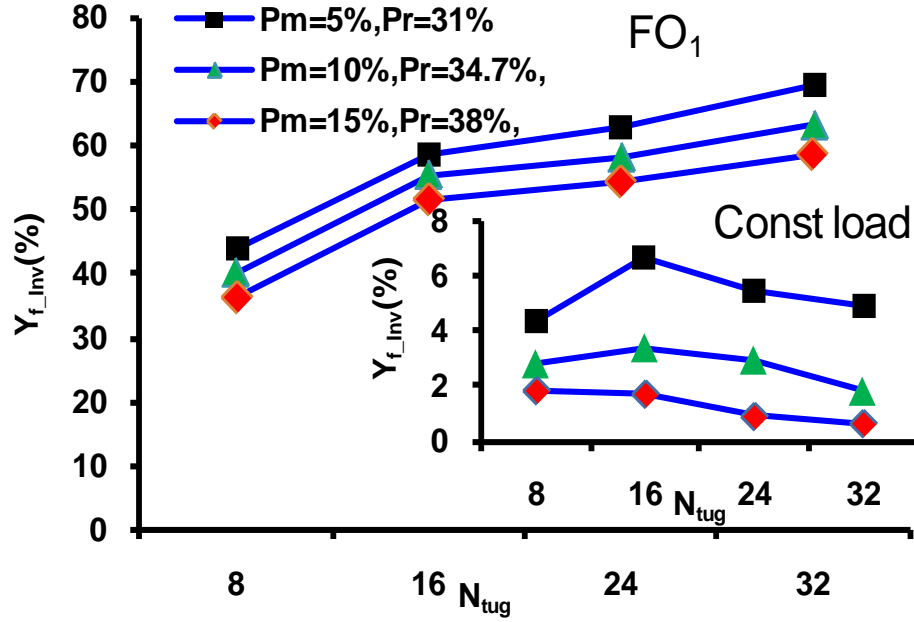
Figure 6-3: Monte Carlo simulation yield summary for inverters with delay constraint of 1.3X, static power constraint of 10X, as a function of N_{tug}, P_m and P_r when FO_1 is considered and tubes are removed from both driving and fanout gates (realistic scenario). The inset graph shows the MC simulation yield for inverters with delay and static power constraints of 1.3X and 10X as a function of N_{tug}, P_m and P_r when it is assumed that the load is constant.

Next we consider the impact of different fanout when a finite number of tubes are removed from the gates. Figure 6-4 shows the Functional yield of 2-input NAND gates obtained from the Monte Carlo simulation, with different drive strengths, and with FO_1 and FO_4. From the figure, two parallel lines for FO_1 and FO_4 can be observed for a constant P_m. The change in fanout shown by the parallel lines is within 8% difference in the functional yield of the gates. The reason is the probabilistic nature of tube removal from both the drive and fanout gates.
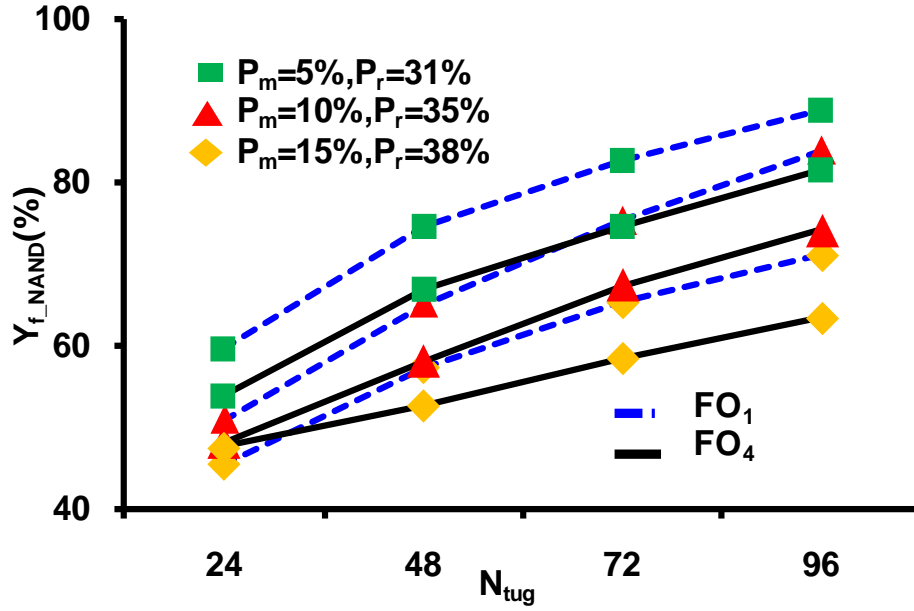
Figure 6-4: Monte Carlo simulation yield summary for NAND gate with FO$_1$ and FO$_4$. Yield is obtained as a function of N_{tug} , P_m and P_r for a delay constraint of 1.3X and static power constraint of 10X.

Figure 6-5 shows the impact of removal of tubes on the yield of NAND gates, assuming that almost all the metallic tubes are removed but no semiconducting tubes are removed (ideal case). We use the ideal case result as a baseline for our analysis. From the figure it can be observed that for up to 10% metallic tubes reasonable yield can be obtained at the gate level, and the circuit level performance variation can be within acceptable limits due to statistical averaging among gates. However, both of the tube removal processes, i.e. SCE and VMR, are not perfect as they remove metallic as well as semiconducting tubes, therefore, these techniques need to be improved.
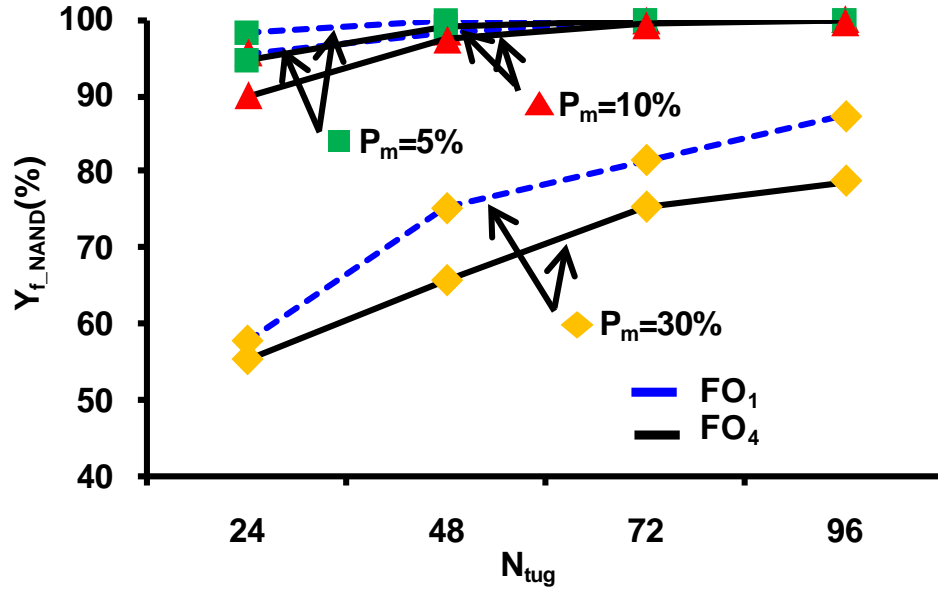
Figure 6-5: Monte Carlo simulation yield summary for NAND gate with FO$_1$ and FO$_4$. Yield is obtained as a function of N_{tug}, P_m and P_r for a delay constraint of 1.3X and static power constraint of 10X.

6.4 Analytical Yield Model

Since Monte Carlo simulations are computationally intensive, we have developed analytical models to quickly analyze the functional yield behavior of logic gates in the presence of fabrication imperfections. It is observed from Monte Carlo simulations that both SCE and VMR techniques are almost perfect, in terms of removal of metallic tubes, removing almost 99.99% of metallic tubes. We derive our analytical model explicitly on the delay constraint, and the power constraint implied.

The derived analytical models compute the functional yield of logic gates as a function of *a)* drive strength of a gate, *b)* the percentage of metallic tubes present prior to the application of tube removal process, and *c)* the percentage of metallic and semiconducting tubes removed after the application of the tube removal process.

117

6.4.1 General Analytical Model for Yield

If a finite number of tubes are removed from the driving gate, statistically, the logic gate will have a finite delay penalty due to the reduction in the drive strength of the gate, compared to the case when all the tubes are present in the gate. Similarly, if a finite number of tubes are removed from the fanout gate(s), there will be reduction in the delay of the logic gate because of the reduction in the total load capacitance that will be driven by a gate. We define the maximum acceptable increase in the delay, X_{max}, due to the finite number of tubes being removed from the drive and fanout gates as compared to the delay of the gate when all the tubes are present and semiconducting, and no tube removal process is applied. Here the objective is to obtain the number of different possible combinations of tubes removed from the drive and fanout gate(s) (N_C) that can be tolerated without violating the maximum allowable delay constraint.

The delay of the gates under an ideal scenario when all the tubes are semiconducting and no tube removal process is applied, is given by

$$D_{g\_ideal} = \frac{C_{L\_ideal}V_{DD}}{N_{tu\_dr}I_{on} - N_{tu\_off}I_{off}}$$ (6.1)

It is assumed that for a semiconducting tube the *ON* current is much larger than the *OFF* current, i.e. $I_{on} \gg I_{off}$ and no metallic tubes are remained after the application of the tube removal process. Similarly, the total load capacitance of the gate when all tubes are semiconducting and no tube process is applied is given by

$$C_{L\_ideal} = C_{p\_dr} + C_{p\_fo} + N_{tu\_fo}C_{g\_fo}$$ (6.2)

Now the delay of the gates when a finite percentage of tubes are removed due to the application of the tube removal process is given by

$$D_{g\_turm} = \frac{C_{L\_turm}V_{DD}}{\left(N_{tu\_dr} - N_{turm\_dr}\right)I_{on}}$$

(6.3)

Likewise, the load capacitance of the gate when a finite number of tubes are removed from the drive and fanout gate(s) is given by

$$C_{L\_turm} = C_{p\_dr} + C_{p\_fo} + \left(N_{tu\_fo} - N_{turm\_fo}\right)C_{g\_fo}$$

(6.4)

6.4.2 Allowed Combination of Tubes Removed from Drive and Fanout Gate(s) (N_C):

For a given number of parallel CNTs in the drive and fanout gates, and the maximum acceptable delay ratio X_{max}, we obtain the different combinations of tubes removed from the driving gate as given by

$$\sum_{i=1}^{N_C} N_{turm\_dr}\left(i\right)$$

(6.5)

and the combinations of tubes removed from fanout gate(s) as given by

$$\sum_{i=1}^{N_C} N_{turm\_fo}\left(i\right)$$

(6.6)

The delay of the gates which is within the acceptable limit after the tube removal process is given by

$$D_{g\_turm} \leq X_{max}D_{g\_ideal}$$

(6.7)

119

6.4.3 Probability of PU/PD Network Being Functional (Pr_{PU}/ Pr_{PD}):

After obtaining the different possible combination of tubes removed from the drive and fanout gates (*for which the delay is within acceptable limits*), we calculate the probability of the pull-up network, Pr_{PU}, and the pull-down network, Pr_{PD}, being functional by meeting the delay constraints.

$$\Pr_{PU} = \Pr_{PD} =$$
$$\sum_{i=0}^{N_C} (1 - P_r)^{\left(N_{tu\_dr} - N_{turm\_dr}(i)\right)} P_r^{N_{turm\_dr}(i)} \, {}^{\left(N_{tu\_dr}\right)}C_{N_{turm\_dr}(i)} \tag{6.8}$$

6.4.4 Functional Yield of a Gate (Y_f):

For a gate to be functional, the worst case delay of both the pull-up and the pull-down networks has to be less than the maximum allowable limit when a finite number of tubes are removed. Therefore the functional yield of a gate, Y_f, can be expressed as the *joint probability* of both *PU* and *PD* networks being simultaneously functional as shown in

$$Y_f = Pr_{PU} \times Pr_{PD} \tag{6.9}$$

6.4.5 Analytical Model of an Inverter

For the inverter, we first obtain the total number of possible combinations of tubes removed from the drive and fanout gates from equation(6.1),(6.3), and (6.7) . As we implemented the pull-up and the pull-down networks of inverters with the same tubes, the functional probabilities of pull-up and pull-down networks are the same. The computation of Pr_{PU} is sufficient to compute the functional probability given by equation(6.8). In this

120

case, the functional yield of the inverter will be equal to the functional probability of either the pull-up or the pull-down network.

6.4.6   Analytical Model Derivation of a NAND Gate

In the case of a NAND gate we are assuming that transistors within the pull-up (or pull-down) network are implemented with the same tubes, as shown in Figure 6-2. Furthermore, the pull-up and pull-down networks are implemented with un-correlated tubes, i.e. these two networks do not share any tubes between them (please see Figure 6-2). To make the worst case rise and fall delays equal, the number of tubes in the transistors in the pull-down network, N_{turN}, is twice the number of tubes in the transistors of pull-up network, N_{turP}. We can obtain the functional probability of *PU* and *PD* networks of NAND gate by inserting the number tubes in the transistors of pull-up and pull-down networks in equation (6.1) ,(6.3) and (6.7). Finally we can obtain the functional yield of NAND gate by substituting the functional probability of pull-up and pull-down networks of a NAND gate obtained from equation (6.8) in equation(6.9).

The analytical model for a 2-input NOR gate can be derived similarly as the NOR gate is the dual of a NAND gate. The expressions for the functional yield of pull-up and pull-down networks are switched, but the functional yield remains the same as that of a NAND gate. Analytical models for complex gates, and gates with larger fan-in can be derived similarly.

6.4.7   Comparison between Monte Carlo and Analytical Model for Inverter

Figure 6-6 shows the comparison between the functional yield obtained from Monte Carlo simulations, and our analytical model for an inverter. Analytical model results are shown

with lines and Monte Carlo results are shown with symbols. The upper inset table shows the absolute difference in functional yield magnitudes between Monte Carlo simulations and analytical models for *a)* different percentage of metallic tubes, *b)* different percentage of tubes removed, and *c)* different drive strength of the inverter. In our experiments the range of absolute difference in functional yield magnitudes is between 0% to 2.6% for inverter with N_{tug}= 48 and 15% of metallic tubes. This small difference shows that our analytical model estimates the functional yield with excellent accuracy without going through computationally extensive Monte Carlo simulations.

N_{tug}	Pm=5%	Pm=10%	Pm=15%
22	1.2	2.1	2.2
30	1.5	1.6	2.4
38	0.0	0.0	0.0
46	0.4	0.7	2.6
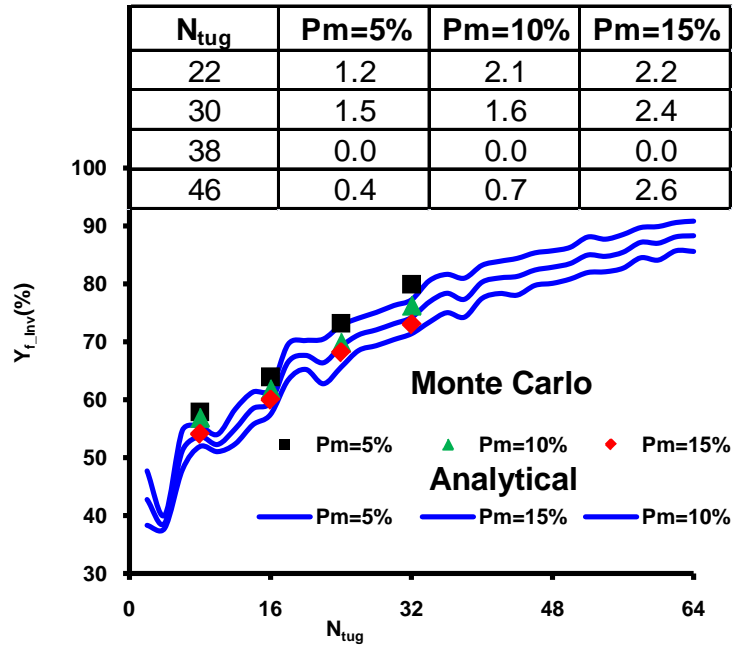
Figure 6-6: Functional yield, Y_{f_Inv}, for multi-channel CNT based inverter gate as predicted by the analytical model and Monte Carlo simulation for different drive strengths as measured by the number of tubes in the inverter (N_{tug}) gate and for a different percentage amount of metallic tubes (5%, 10% and 15%) for allowed delay constraint of 1.3X, and static power constraint of 10X. The upper inset table shows the absolute difference in functional yield between MC simulations and the analytical model for a different percentage of metallic tubes and different drive strengths of inverter gate.

Figure 6-7 shows the functional yield comparison between Monte Carlo simulations and the analytical model results for 2-input NAND gates. Results from the analytical model are shown with lines and Monte Carlo simulation results are shown with symbols. The upper inset table shows the absolute difference in functional yields between data obtained from Monte Carlo simulations and from the analytical model. Results are reported for different percentage of metallic tubes and different numbers of tubes in the NAND gate. In the table shown within Figure 6-7, the range of absolute error in functional yield is between 0% and 3.4%, The maximum error is observed for NAND gate with N_{tug}=96 and for 5% metallic tubes. Very small difference in yield numbers show that we can accurately predict the

functional yield of NAND gate analytically in constant runtime. The maximum error of 3.5% obtained for NAND gate as compared to 2.6% for an inverter is because of the increased complexity of the NAND gate compared to an inverter.

N_{tug}	P_m=5%	P_m=10%	P_m=15%
24	1.8	1.6	1.8
48	1.1	0.7	0.8
72	0.2	1.6	0.8
96	3.4	1.2	1.3
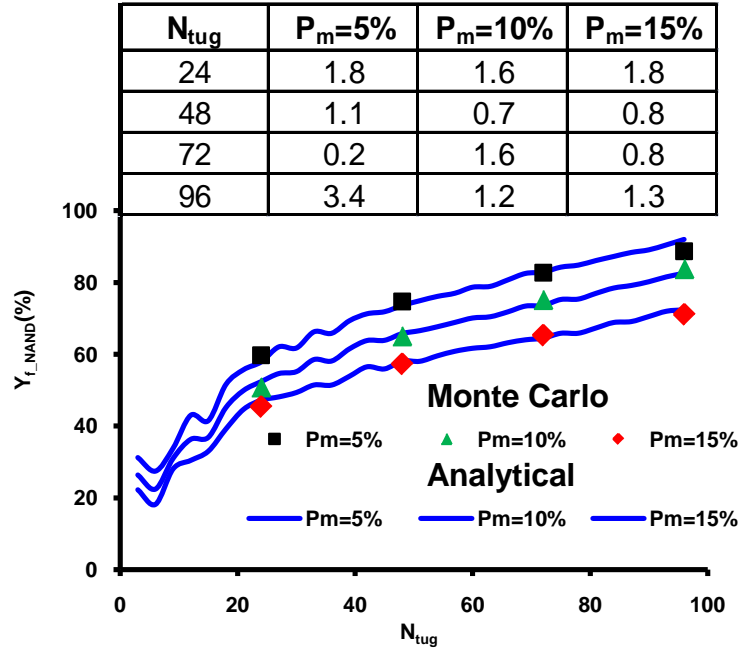
Figure 6-7: Functional yield, Y_{f_NAND}, for multi-channel CNT based NAND gate as predicted by analytical model and Monte Carlo simulation for different drive strengths as measured by the number of tubes in the NAND (N_{tug}) gate and for a different percentage amount of metallic tubes (5%,10% and 15%) for allowed delay constraint of 1.3X and static power constraint of 10X. The upper inset table shows the difference in functional yield between MC simulations and the analytical model for a different percentage of metallic tubes and different drive strengths of NAND gate.

6.5   Tube Level Redundancy (TLR)

From the data presented in the previous section, it can be concluded that removal of tubes creates two main problems *1)* open-circuit transistors/gates when all tubes are removed from a transistor and *2)* low functional yields because of a finite number of tubes removed from the gates. The flowchart of Monte Carlo simulation setup used is provided in Section 10.4.3 of Appendix B. To tackle these problems, we propose tube level redundancy (TLR)

to reduce the probability of open-circuit gates, and improve the functional yield of gates. *Our objective is to find the minimum number of tubes (N_{turmin}) required in a transistor prior to tube removal process for less than 0.001% probability of open-circuit CNFETs. N_{turmin}* can be calculated as shown in (6.10).

$$N_{turmin} = \log\left(\frac{10^{-5}}{P_r}\right)$$

(6.10)

Table 6-1 shows the N_{turmin}, required for a negligible probability of open circuit transistors for different percentage of metallic tubes. Numbers in Table 6-1 are calculated by assuming that the SCE technique is applied to remove metallic tubes and the cutoff diameters for metallic and semiconducting tubes are D_{CS}=1.4nm and D_{CM}=2nm. The same methodology can also be applied to VMR technique.

Table 6-1: Minimum number of CNTs required in a CNFET to produce 0.001% probability of open circuit transistors.

P_m	0%	5%	10%	15%	20%	25%	30%
N_{turmin}	1	8	9	10	11	12	13

Figure 6-8 shows the impact of addition of redundant tubes on the functional yield of 2-input NAND gates. From the figure it can be observed that the functional yield of a NAND gate is 60%, when P_m=5% and P_r=31%. Now by increasing the number of the tubes in the gate the functional yield of the gate increases. However to obtain the acceptable yield almost 6X increase in the number of tubes are required. Figure 6-8(b) and Figure 6-8(c) show the impact of the addition of redundant tubes on the area as well as average energy consumption of the gates. From the figure it can be observed that acceptable level of yields are obtained at the expense of almost 4.4X increase in the area

and almost 4.7X increase in the energy of the gates. This redundancy will be very expensive in terms of the area and energy requirements and diminishes the advantages of CNFET over silicon CMOS technology.
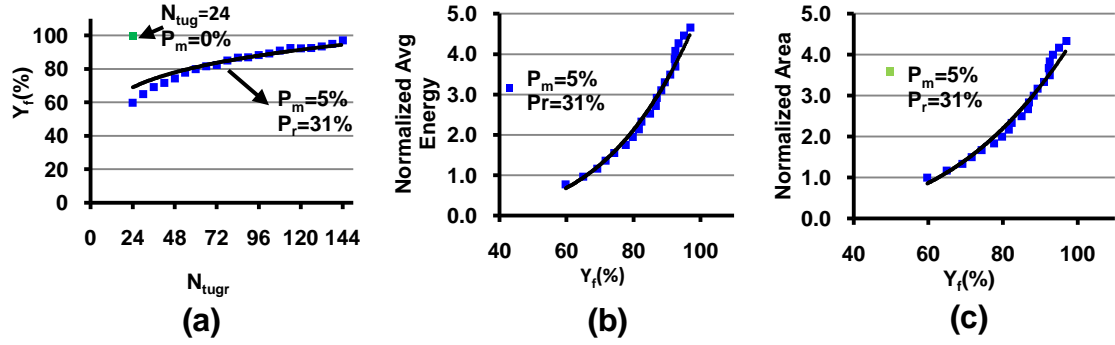
Figure 6-8: (a) Impact of redundancy on the functional yield of NAND gates, (b) Increase in mean energy of gates by increasing the functional yield due to tube level redundancy, (c) increase in area of gates by increasing the functional yield due to tube level redundancy on Area. Results are obtained from Monte Carlo simulation when P_m=5% and P_r=31% and N_{tug}=24. Sample size of 10,000 gates is used for Monte Carlo simulations.

An efficient TLR technique is proposed in this work which allows us to obtain acceptable levels of yield without sacrificing too much area and power. *Here we add the redundant tubes with the objective to obtain the same mean number of tubes in the CNFET after tube removal as required by the design prior to tube removal process.* Table 6-2 shows the efficient redundancy estimation technique to increase the functional yield of gates (*when a finite number of tubes are removed*) with minimal impact on the area and energy. Here we add the redundant tubes with the objective to obtain the same mean number of tubes in the CNFET after tube removal as required by the design prior to the tube removal process. For example, in Table 6-2 if the number of tubes required in a CNFET prior to tube removal are 8 and P_m=10%. Then after the tube removal process the mean tubes remaining are 5. However if put 13 tubes in the CNFET prior to tube removal process in the CNFET, then mean number of tubes remained after the tube removal process will be 8.

126

Table 6-2: Original CNTs in a CNFET before tube removal (*BTR*), and after tube removal (*ATR*), for different percentage of metallic tubes. The number of CNTs required in a CNFET *BTR* that will produce the same mean CNTs in a CNFET after tube removal as are initially required by design *BTR*.

P_m(%)	Original		Redundancy	
	CNTs BTR	μ CNTs ATR	CNTs BTR	μ CNTs ATR
5	8	5	12	8
10		5	13	
15		5	13	
5	16	11	24	16
10		10	25	
15		9	26	
5	32	22	47	32
10		20	49	
15		19	52	

Figure 6-9 shows the Monte Carlo simulation results for NAND gates for N_{tug}=48. Figure 6-9(b) shows the functional yield when P_m=10% and P_r=35% of the tubes are removed by the tube removal process. Here only 58% of the gates are functional. Now by adding the redundancy based on the methodology developed in the previous paragraph, N_{tugr}=72 are required in the gate that will yield mean tubes of 48 after the tube removal process. This added redundancy of 50% increased the functional yield from 58%to 66%. By using this redundancy technique the increase in area is 50% as compared to 6X required in the previous case. Similarly, as shown in Table 6-3 the mean energy of the gate increases by 17% compared to 4.7X in the previous technique. Here the yield is much less than 100% but it is anticipated that in large circuits, multiple gates are cascaded to form multi-stage logic network. Depending on the logic depth of logic network, certain amount of statistical averaging in delay variation is observed. Thus less than 100% yield at the *gate level* is sufficient to obtain acceptable yield at the *system level*, as explained in the next section.
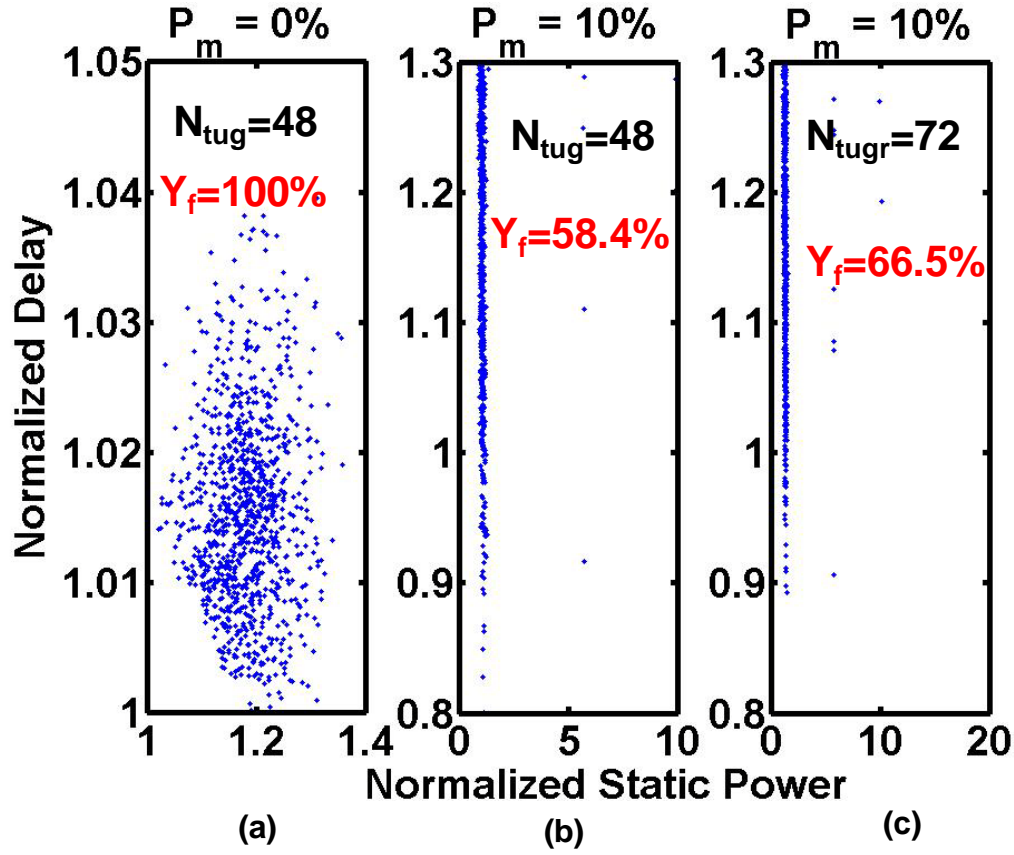
Figure 6-9: Monte Carlo simulation for NAND gates showing normalized delay vs. static power for (a) N_{tug}=48, 0% metallic tubes and no tube removal, (b) N_{tug}=48, 10% metallic tubes and 35% tubes are removed, (c) N_{tugr}=72, 10% metallic tubes and 35% tubes are removed. The yield is 100%, 58% and 66%, respectively.

Table 6-3: Impact of tube removal on the mean (μ) and (σ/ μ) energy without applying the redundancy (N_{tug}=48) and when redundancy is applied (N_{tugr}=72).

Energy	ideal	After TR	After TR & Redunancy
mean	1.00	0.78	1.17
sigma/mean	0.00	0.08	0.05

6.6   Critical Path Analysis

In a complex digital system like CPU the speed of the system is determined by the delay of

the critical path. In this work we represent a typical CPU pipeline stage with the critical path

composed of 9 levels of NAND gates [118].

Figure 6-10 shows the impact of the path depth (d_{path}) on the Functional yield of NAND gates for different drive strengths as measured by the number of tubes in the gate. From the figure it can be observed that when $N_{tug}=12$, the functional yield of NAND gates decreases by increasing the depth of logic path. The reason is that by increasing the path depth there will be higher probability of paths consisting of open circuit gates. However, when the number of tubes in the NAND gate is 24 or higher, the probability of open circuit gates is negligible and there we can see that the functional yield of the path increases by increasing the logic depth of the path. Figure 6-11 shows the σ/μ variation in the delay as a function of the path depth. Again it can be observed that by increasing the number of tubes in the gate the variation in delay decreases due to statistical averaging among the gates.

From Figure 6-11 it can be observed that for a typical path depth of 9 NAND gates the functional yield is still less than 100%. This deficiency in yield can be compensated by either adding the efficient redundancy technique at the gate level as described previously, or by further increasing the path depth.
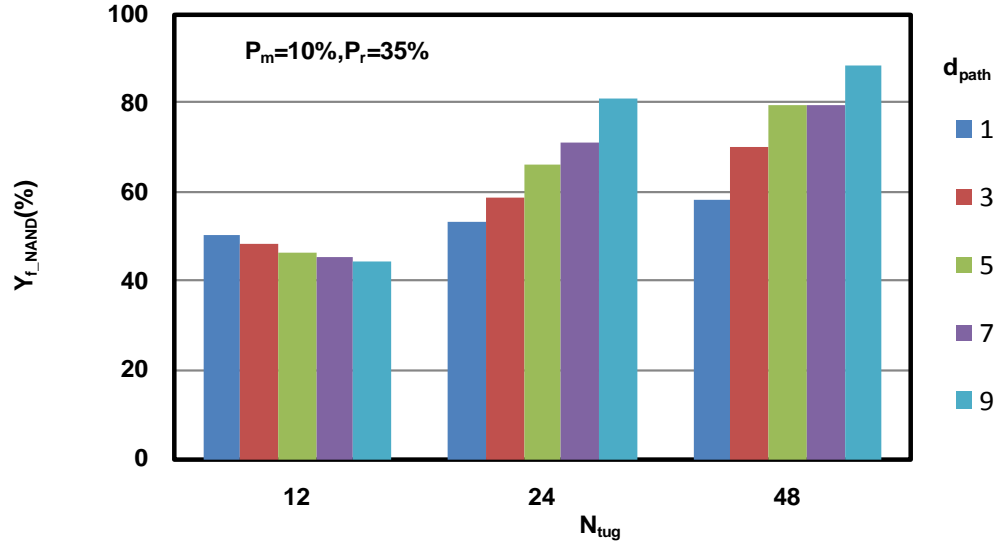
Figure 6-10: Impact of path depth(d_{path}) on the functional yield of NAND gate for different drive strengths of NAND gates, as measured by the number of tubes in the gate(N_{tug}) when P_m=10% and P_r=35%.
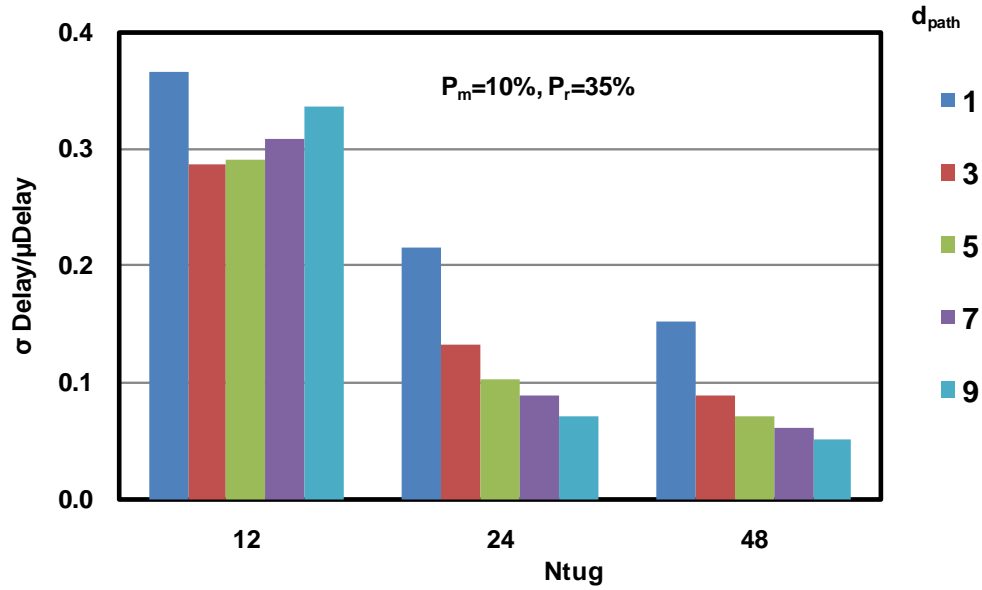
Figure 6-11: Impact of path depth(d_{path}) on (σ Delay/μ Delay) of NAND gate for different drive strengths of NAND gates as measured by the number of tubes in the gate(N_{tug}) when P_m=10% and P_r=35%.

Figure 6-12 shows the functional yield of a critical path consisting of 9 NAND gates obtained from Monte Carlo simulation when the number of tubes required in the design prior to tube removal process is 48. However, by adding the tube level redundancy as described previously for a typical path depth of 9 gates, almost 100% functional yield is obtained as shown in Figure 6-12(c). Table 6-4 shows the impact of tube removal on the mean and σ/μ energy for path depth of 9 NAND gates when no redundancy is applied (N_{tug}=48) and when redundancy is applied (N_{tugr}=72). From the table it can be observed that addition of 50% more tubes results in 12% increase in the average energy, and 5% increase in the variation in the energy as given in Table 6-4. The addition of redundant tubes results in 50% increase in the overall area of the circuit.
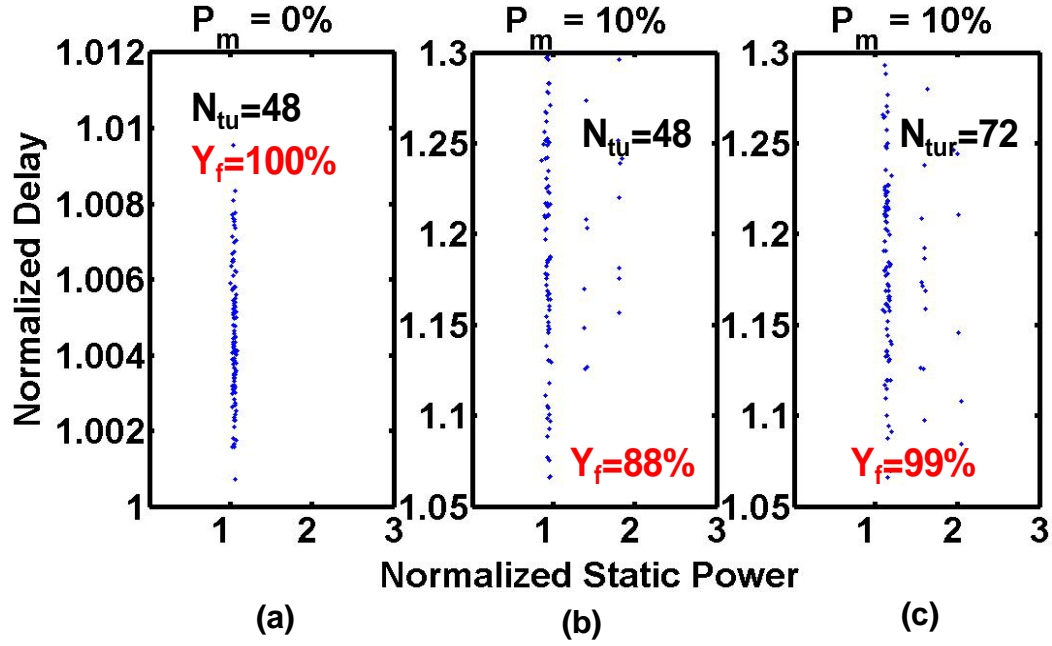
Figure 6-12: Monte Carlo simulation for NAND gates showing normalized delay vs. static power for (a) N_{tug}=48, 0% metallic tubes, d_{path}=9 and no tube removal (b) N_{tug}=48, 10% metallic tubes, d_{path}=9, 35% tubes are removed (c) N_{tugr}=72, 10% metallic, d_{path}=9 and 35% tubes are removed. The yield is 100%, 88% and 99% respectively.

Table 6-4: Impact of tube removal on the mean (μ) and (σ/μ) energy without applying the redundancy (N_{tug}=48) and when redundancy is applied (N_{tugr}=72). For d_{path}=9 gates.

Energy	ideal	After TR	After TR & Redunancy
mean	1.00	0.75	1.12
sigma/mean	0.00	0.06	0.05

We can also increase the functional yield of the gates in a path by increasing the depth of the path instead of using the tube level redundancy. The trade-off in this case will be reduction in performance. Figure 6-13 shows the impact of path depth on the functional yield of NAND gates when (N_{tug}=48) obtained from Monte Carlo simulation. It can be observed that almost 100% yield can be obtained when the path depth is increased to 17 gates which is almost twice the typical path depth of 9 NAND gates in digital systems. The trade-off of increasing the path depth is almost 2X reduction in performance. However,

same throughput can be achieved as that of the original design by using parallelism as in the case of datapaths. In this situation, the trade-off will be ~2X increase in area, 7.5% increase in average energy, and 1% increase in the variation in energy.
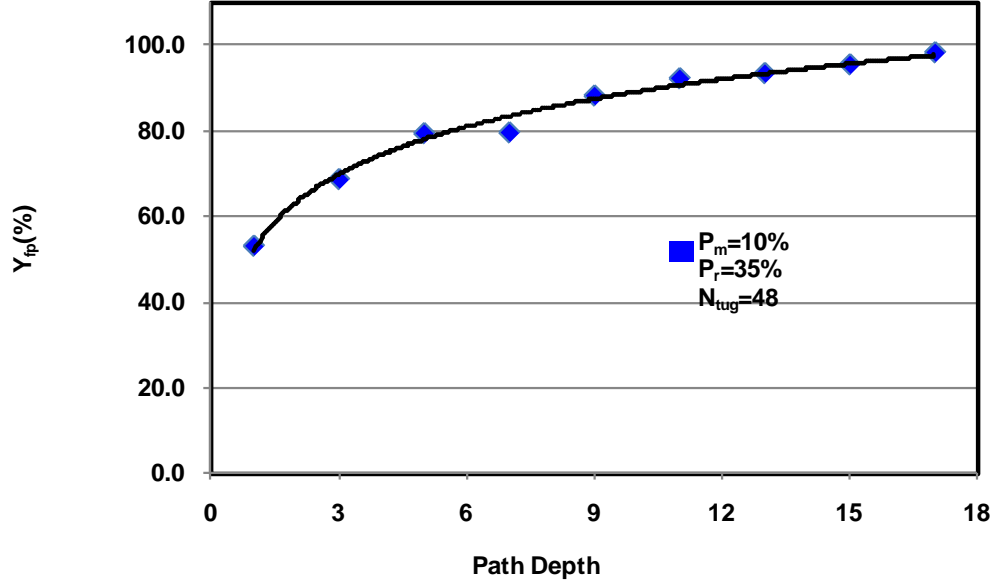
Figure 6-13: Impact of path depth (d_{path}) on the functional yield when (N_{tug}=48) P$_m$=10% and P$_r$=35%. Almost 100% functional yield is obtained when increasing the path depth to 17.

6.7 Conclusion

The removal of metallic tubes results in large performance variations and reduces the functional yield of CNFET based circuits. The analysis presented in this paper considers the impact of stochastic removal of tubes removed from the driving gates as well as from the fanout gates. We present analytical models for the yield estimation of gates which is extremely useful in predicting the impact of yield loss due to removal of tubes for various percentages of metallic tubes, percentage of tubes removed, drive strengths of the gates, and fanout of gates. An efficient tube level redundancy technique is proposed which helps

to increase the functional yield of CNFET based circuits with minimum impact in terms of area and energy. Analysis shows that the yield loss of CNFET based circuits due to tube removal can also be compensated by increasing the logic path depth. The performance loss due to increase in path depth can be compensated by architecture level parallelism.

7 Contributions, Conclusions and Future Work

Silicon based Integrated Circuit technology has witnessed aggressive scaling over the last four decades but now it is approaching its physical limits. Research has started in earnest for new materials in sub-10nm technology node. The superior electrostatic properties of CNFETs make them a potential candidate for future integrated circuits. However, because carbon nanotubes are grown by chemical synthesis, it is very difficult to obtain the precise control on the exact positioning and chirality of CNTs during their growth. These CNT growth imperfections lead to a misalignment of tubes, and the unwanted growth of metallic tubes. In this work, we have analyzed the impact of the unwanted growth of metallic tubes on the performance, power and yield of CNFET based circuits. Moreover solutions are proposed which help to build robust CNFET based circuits with reduced variability in the performance and power in the presence of fabrication imperfections.

7.1 Contributions and Conclusions

This thesis focused on the impact of fabrication imperfections on the performance, power, and yield of CNFET based integrated circuits. In Chapter 4 we analyzed the impact of variation in the diameter of CNTs, and spacing between adjacent CNTs on the drive strength of parallel tube CNFETs. The results showed that both the variations in the tube diameter and inter-tube spacing can be tolerated to a certain extent, because of statistical averaging among tubes in multi-channel CNFETs.

In Chapter 5 we showed, with the help of Monte Carlo simulations, that the unwanted growth of metallic tubes has a detrimental impact on the performance, power, and functional yield of CNFET based circuits. We proposed two new CNFET transistor

configurations: Transistor Stacking(*TrS*) and Tube Stacking(*TuS*), which increased the functional yield of CNFET based gates by reducing the statistical probability of an ohmic short between the drain and source terminals of parallel tube CNFETs. Furthermore, accurate analytical models are developed, estimating the functional yield of logic gates for different percentages of metallic tubes, and different drive strengths of logic gates implemented with different configurations of the CNFETs. It is observed that, although stacking configurations increased the functional yield significantly and reduced the static power by an order of magnitude, the trade-off of the stacking configurations is in terms of an almost 4X delay penalty.

The analysis shows that the delay penalty associated with the proposed stacking configurations can be compensated by using parallelism in the critical path of circuits. The implementation of circuits with the proposed stacking configurations, and parallelism in the critical path results in the same performance as obtained from parallel tube configurations of the transistors, but with 4X improvement in functional yield, and 6X reduction in the static power. As CNTs are grown using chemical self assembly, in addition to process variations observed in conventional CMOS fabrication, CNFET based circuits are subjected to sources of imperfections that are unique to CNTs. To handle the process related variations we proposed architecture based on regular logic bricks which are designed using hybrid configurations of transistors. There are two main advantages of implementing the designs with regular logic blocks: (1) reduction in the systematic process related variations in nanoscale technologies, and (2) for larger percentage of metallic tubes, acceptable levels of yield can be obtained by using redundant logic blocks. Our analysis

showed that for up to 10% metallic tubes, logic bricks implemented with hybrid configurations of CNFETs can help to reduce the performance impact by 2X, as compared to homogenous bricks implemented with only *TrS* CNFETs. In comparison to homogenous bricks realized with PT CNFETs, the static power can be reduced by 2X and yield can be increased by 2.5X.

The proposed circuit level techniques can handle the metallic tubes if we can reliably and grow CNTs with less than 5% metallic tubes. For a large percentage of metallic tubes, extra processing techniques are required to remove the unwanted metallic tubes. Significant progress has been made by researchers, and different techniques have been developed, selectively removing the metallic tubes from an ensemble of metallic and semiconducting CNTs. The trade-off with these extra processing techniques is that they also remove the finite number of semiconducting tubes. The removed metallic and semiconducting tubes result in density variations in the CNFETs, causing a large variability in the performance and power of CNFET based circuits, and in the worst case open circuit gates if all the tubes from the CNFET are removed.

To analyze the impact of a removal of tubes by these extra processing techniques, we have developed a Monte Carlo simulation engine in Chapter 6. The Monte Carlo simulation analyzes the impact of removing of metallic and semiconducting tubes, for different drive strengths of logic gates and for different percentages of metallic tubes before the application of extra processing techniques, and percentages of metallic and semiconducting tubes removed after the application of extra processing techniques. Furthermore, analytical models are developed to allow the designers to quickly analyze the impact of tube removal

from the driving and fan-out gates on the yield of CNFET based gates without going through the computationally intensive Monte Carlo simulation.

The efficient Tube Level Redundancy (*TLR*) technique is proposed, allowing for an increase in the functional yield of CNFET based circuits to acceptable levels when large fraction of tubes are removed. The trade-off of TLR is an almost 50% increase in area and an almost 12% increase in the average power, which is much less than that associated with the conventional redundancy techniques. Another architecture level solution is proposed where the functional yield is increased, and variability in the CNFET circuit parameters is reduced due to a removal of tubes by increasing the logic depth of logic gates. The trade-off of this approach is the reduction in the performance of CNFET based circuits. However, for applications where throughput is more important than latency, we can increase the performance to the same level as obtained from the ideal scenario where all the tubes are semiconducting, and no tubes are removed by using the parallelism. The trade-off of this approach is 2X penalty in terms of area and 8% increase in the average power.

We showed that even the unwanted growth of metallic tubes has a detrimental impact on the performance, power and yield of CNFET based circuits. Possible solutions to build robust CNFET based circuits with acceptable performance and reasonable functional yield can be achieved.

In summary below is the list of specific contributions of this work

- Analysis of the impact of variation in the diameter of CNTs, and spacing between adjacent CNTs on the drive strength of parallel tube

- For small percentage of metallic tubes i.e. < 5%, two new CNFET transistor configurations: Transistor Stacking(*TrS*) and Tube Stacking(*TuS*) are proposed, which increased the functional yield of CNFET based gates by reducing the statistical probability of an ohmic short between the drain and source terminals of parallel tube CNFETs.

- Accurate analytical models are developed, estimating the functional yield of logic gates for different percentages of metallic tubes, and different drive strengths of logic gates implemented with different configurations of the CNFETs.

- Presented a methodology for yield-aware carbon nanotube based circuit design in the presence of metallic tubes using different CNFET transistor configurations.

- Architecture level techniques such as parallelism and implementation of circuits with regular logic blocks are proposed to obtain better trade-off between delay, power and yield parameters.

- When the percentage of metallic tubes is large i.e. > 5%, we analyzed the impact of variability in the performance and static power due to removal of tubes with the help of Monte Carlo simulations.

- Analytical models are developed to allow the designers to quickly analyze the impact of tube removal from the driving and fan-out gates on the functional yield of CNFET based gates without going through the computationally intensive Monte Carlo simulation.

- The efficient Tube Level Redundancy (*TLR*) technique is proposed, allowing for an increase in the functional yield of CNFET based circuits to acceptable levels when large fraction of tubes are removed.

- Another architecture level solution is proposed where the functional yield is increased, and variability in the CNFET circuit parameters is reduced due to a removal of tubes by increasing the logic depth of logic gates.

7.1.1   Conclusions

- Diameter and spacing variations are issue at the tube level but not a big challenge to parallel tube CNFET based circuits. Both the tube diameter and inter-tube spacing can be tolerated to a certain extent, because of statistical averaging among tubes in parallel tube CNFETs.

- Unwanted growth of metallic tubes is one of the major technological barriers faced by the CNT technology that hinders the development of CNFET based circuits for real applications.

- We showed that even the unwanted growth of metallic tubes has a detrimental impact on the performance, power and yield of CNFET based circuits. Possible

solutions such as stacking configurations or tube level redundancy can be used to build robust CNFET based circuits with acceptable performance and reasonable functional yield.

7.2   Suggestions for Future Work

This thesis has proposed a few techniques to handle the unwanted growth of metallic tubes in CNFET based circuits. Below is the list of proposed topics for possible future works which can be instrumental in analyzing and overcoming the challenges faced by the commercialization of CNT based technology.

- Develop integrated CNFET based development tools to allow designers to consider the impact of different CNT fabrication imperfections, and to apply the solutions proposed. This will allow the designers to estimate the impact of these different sources of fabrication imperfections on the circuit parameters like area, performance, and power.

- Similarly, circuits fabricated using CNFETs have some of the lithographic related variations in addition to the imperfections specific to the CNT technology. The CNFET design methodology can be enhanced to incorporate the impact of lithographic related variations.

- Development of new methodology, giving the designers a fully integrated approach to implement complete integrated circuits with semiconducting CNTs used as channel material, metallic CNTs used as interconnects, decoupling capacitors, and inductors. This will allow to the semiconductor industry to estimate the full

advantage that can be obtained by using CNT technology to implement integrated circuits, as compared to integrated circuits implemented with current process technology.

8   Bibliography

[1]   Tze-Chiang Chen, "Overcoming research challenges for CMOS scaling: industry directions," in *International Conference on Solid-State and Integrated Circuit Technology*, pp. 4-7, 2006.

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

[3]   N. Weste and D. Harris, *CMOS VLSI Design: A Circuits and Systems Perspective*, 4th ed. Addison Wesley, 2010.

[4]   S. Tyagi et al., "An advanced low power, high performance, strained channel 65nm technology," in *Proc. Int. Electron Devices Meet*, pp. 245-247, 2005.

[5]   K. Mistry et al., "A 45nm Logic Technology with High-k+Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-free Packaging," in *Proc. Int. Electron Devices Meet*, pp. 247-250, 2008.

[6]   D. Hisamoto et al., "A folded-channel MOSFET for deep-sub-tenth micron era," in *Proc. Int. Electron Devices Meet*, pp. 1032-1034, 1998.

[7]   J. Kavalieros et al., "Tri-Gate Transistor Architecture with High-k Gate Dielectrics, Metal Gates and Strain Engineering," in *Proc. IEEE Int. Symp. VLSI Technol.*, pp. 50-51, 2006.

[8]   [Online] Available, "http://www.itrs.net/Links/2009ITRS/Home2009.htm/," *Emerging Research Devices*. [Online]. [Accessed: 24-Feb-2010].

[9]   Y. Cui, Z. Zhong, D. Wang, W. U. Wang, and C. M. Lieber, "High Performance Silicon Nanowire Field Effect Transistors," *Nano Lett.*, vol. 3, no. 2, pp. 149-152,

2003.

[10]   T. Ashley et al., "Novel InSb-based quantum well transistors for ultra-high speed, low power logic applications," in *International Conference on Solid-State and Integrated Circuits Technology*, vol. 3, pp. 2253-2256, 2004.

[11]   S. Datta et al., "85nm gate length enhancement and depletion mode InSb quantum well transistors for ultra high speed and very low power digital logic applications," in *Proc. Int. Electron Devices Meet*, pp. 763-766, 2005.

[12]   Chien-I Kuo et al., "A 40-nm-Gate InAs/In0.7Ga0.3As Composite-Channel HEMT with 2200 mS/mm and 500-GHz fT," in *IEEE International Conference on Indium Phosphide & Related Materials*, pp. 128-131, 2009.

[13]   T. Ashley, A. Dean, C. Elliott, R. Jefferies, F. Khaleque, and T. Phillips, "High-speed, low-power InSb transistors," in *Proc. Int. Electron Devices Meet*, pp. 751-754, 1997.

[14]   A. Javey, J. Guo, Q. Wang, M. Lundstrom, and H. Dai, "Ballistic Carbon Nanotube Field-Effect Transistors," *Nature*, vol. 424, no. 6949, pp. 654-657, 2003.

[15]   "EETimes.com - Junctionless transistor is ready for 20-nm node, says researcher." [Online].   Available: http://www.eetimes.com/news/latest/showArticle.jhtml?articleID=223100343. [Accessed: 07-Mar-2010].

[16]   S. D. Suk et al., "High-Performance Twin Silicon Nanowire MOSFET (TSNWFET) on Bulk Si Wafer," *IEEE Transactions on Nanotechnology*, vol. 7, no. 2, pp. 181-184, 2008.

[17]   S. Koo, M. D. Edelstein, Q. Li, C. A. Richter, and E. M. Vogel, "Silicon nanowires

as enhancement-mode Schottky barrier field-effect transistors," *Nanotechnology*, vol. 16, no. 9, pp. 1482-1485, 2005.

[18] R. Chau et al., "Benchmarking nanotechnology for high-performance and low-power logic transistor applications," *IEEE Trans. Nanotechnol.*, vol. 4, no. 2, pp. 153-158, 2005.

[19] "EETimes.com - Intel's Gargini pushes III-V-on-silicon as 2015 transistor option." [Online]. Available: http://www.eetimes.com/news/semi/rss/showArticle.jhtml?articleID=222900175 &cid=. [Accessed: 07-Mar-2010].

[20] N. Goel et al., "Addressing the gate stack challenge for high mobility InxGa1-xAs channels for NFETs," in *Proc. Int. Electron Devices Meet*, pp. 1-4, 2008.

[21] Y. Sun et al., "High mobility III-V channel MOSFETs for post-Si CMOS applications," in *IEEE International Conference on IC Design and Technology*, pp. 161-164, 2009.

[22] X. Wang, Y. Ouyang, X. Li, H. Wang, J. Guo, and H. Dai, "Room Temperature All Semiconducting sub-10nm Graphene Nanoribbon Field-Effect Transistors," *Phys. Rev. Lett.*, vol. 100, 2008.

[23] A. K. Geim and K. S. Novoselov, "The rise of graphene," *Nat Mater*, vol. 6, no. 3, pp. 183-191, Mar. 2007.

[24] "EETimes.com - Graphene wafers ready to fab carbon chips." [Online]. Available: http://www.eetimes.com/showArticle.jhtml;jsessionid=CZZP5MJHUZBGDQE1 GHRSKHWATMY32JVN?articleID=222600668. [Accessed: 07-Mar-2010].

[25] M. Lemme, T. Echtermeyer, M. Baus, and H. Kurz, "A Graphene Field-Effect

Device," *IEEE Electron Device Letters.*, vol. 28, no. 4, pp. 282-284, 2007.

[26]   S. Kim et al., "Realization of a high mobility dual-gated graphene field-effect transistor with Al2O3 dielectric," *Applied Physics Letters*, 2009. .

[27]   "EETimes.com - IBM demos 100-GHz graphene transistor." [Online]. Available: http://www.eetimes.com/showArticle.jhtml;jsessionid=CZZP5MJHUZBGDQE1 GHRSKHWATMY32JVN?articleID=222601227. [Accessed: 07-Mar-2010].

[28]   J. Kedzierski et al., "Epitaxial Graphene Transistors on SiC Substrates," *IEEE trans. Electron Devices*, vol. 55, no. 8, pp. 2078-2085, 2008.

[29]   "Technology Review: Graphene Transistors." [Online]. Available: http://www.technologyreview.com/computing/20119/page1/. [Accessed: 08-Mar-2010].

[30]   Z. Yao, C. L. Kane, and C. Dekker, "High-Field Electrical Transport in Single-Wall Carbon Nanotubes," *Phys. Rev. Lett.*, vol. 84, no. 13, pp. 2941-2944, Mar. 2000.

[31]   T. Durkop, S. A. Getty, E. Cobas, and M. S. Fuhrer, "Extraordinary Mobility in Semiconducting Carbon Nanotubes," *Nano Lett.*, vol. 4, no. 1, pp. 35-39, 2004.

[32]   A. Javey et al., "Carbon Nanotube Field-Effect Transistors with Integrated Ohmic Contacts and High-k Gate Dielectrics," *Nano Lett.*, vol. 4, no. 3, pp. 447-450, Mar. 2004.

[33]   S. J. Tans, A. R. M. Verschueren, and C. Dekker, "Room-temperature transistor based on a single carbon nanotube," *Nature*, vol. 393, no. 6680, pp. 49-52, May. 1998.

[34]   R. Martel, T. Schmidt, H. R. Shea, T. Hertel, and P. Avouris, "Single- and multi-wall carbon nanotube field-effect transistors," *Applied Physics Letters*, vol. 73, no. 17, p.

2447, 1998.

[35] V. Derycke, R. Martel, J. Appenzeller, and P. Avouris, "Carbon Nanotube Inter- and Intramolecular Logic Gates," *Nano Letters*, vol. 1, no. 9, pp. 453-456, 2001.

[36] Z. Chen et al., "An Integrated Logic Circuit Assembled on a Single Carbon Nanotube," *Science*, vol. 311, no. 5768, p. 1735, 2006.

[37] A. Bachtold, P. Hadley, T. Nakanishi, and C. Dekker, "Logic Circuits with Carbon Nanotube Transistors," *Science,*, vol. 294, no. 5545, pp. 1317-1320, Nov. 2001.

[38] "IBM builds first IC around a single carbon nanotube." [Online]. Available: http://www.physorg.com/news12109.html. [Accessed: 07-Mar-2010].

[39] Q. Cao et al., "Medium-scale carbon nanotube thin-film integrated circuits on flexible plastic substrates," *Nature*, vol. 454, no. 7203, pp. 495-500, Jul. 2008.

[40] R. Chau, B. Doyle, S. Datta, J. Kavalieros, and K. Zhang, "Integrated nanoelectronics for the future," *Nat Mater*, vol. 6, no. 11, pp. 810-812, Nov. 2007.

[41] P. McEuen, M. Fuhrer, and P. Hongkun, "Single-Walled Carbon Nanotube Electronics," *IEEE Trans. Nanotechnol.*, vol. 1, no. 1, pp. 78-85, Mar. 2002.

[42] D. S. Bethune et al., "Cobalt-catalysed growth of carbon nanotubes with single-atomic-layer walls," *Nature*, vol. 363, no. 6430, pp. 605-607, Jun. 1993.

[43] S. Iijima and T. Ichihashi, "Single-shell carbon nanotubes of 1-nm diameter," *Nature*, vol. 363, no. 6430, pp. 603-605, Jun. 1993.

[44] M. Ouyang, J. Huang, C. L. Cheung, and C. M. Lieber, "Energy Gaps in "Metallic" Single-Walled Carbon Nanotubes," *Science*, vol. 292, no. 5517, pp. 702 -705, Apr. 2001.

[45] M. S. Dresselhaus, G. Dresselhaus, and R. Saito, "Carbon fibers based on C_{60}

and their symmetry," *Physical Review B*, vol. 45, no. 11, p. 6234, Mar. 1992.

[46]   B. Q. Wei, R. Vajtai, and P. M. Ajayan, "Reliability and current carrying capacity of carbon nanotubes," *Applied Physics Letters*, 2001. .

[47]   G. F. Close, S. Yasuda, B. Paul, S. Fujita, and H. P. Wong, "A 1 GHz Integrated Circuit with Carbon Nanotube Interconnects and Silicon Transistors," *Nano Letters*, vol. 8, no. 2, pp. 706-709, Feb. 2008.

[48]   A. Nieuwoudt and Y. Massoud, "Evaluating the impact of resistance in carbon nanotube bundles for VLSI interconnect using diameter-dependent modeling techniques," *IEEE Transactions on Electron Devices*, vol. 53, no. 10, pp. 2460-2466, 2006.

[49]   M. Budnik, A. Raychowclhury, A. Bansal, and K. Roy, "A high density, carbon nanotube capacitor for decoupling applications," in *Proc. ACM/IEEE Design Automat. Conf.*, pp. 935-938, 2006.

[50]   A. Nieuwoudt and Y. Massoud, "High density integrated capacitors using multi-walled carbon nanotubes," in *IEEE Conference on Nanotechnology*, pp. 387-390, 2007.

[51]   K. Tsubaki, H. Shioya, J. Ono, Y. Nakajima, T. Hanajiri, and H. Yamaguchi, "Large magnetic field induced by carbon nanotube current -proposal of carbon nanotube inductors," in *Device Research Conference Digest*, vol. 1, pp. 119-120, 2005.

[52]   K. Tsubaki, Y. Nakajima, T. Hanajiri, and H. Yamaguchi, "Proposal of Carbon Nanotube Inductors," *Journal of Physics: Conference Series*, vol. 38, pp. 49-52, 2006.

[53]   A. Nieuwoudt and Y. Massoud, "Carbon nanotube bundle-based low loss integrated inductors," in *IEEE Conference on Nanotechnology*, pp. 714-718, 2007.

[54]   N. Chimot, V. Derycke, M. F. Goffman, J. P. Bourgoin, H. Happy, and G.

Dambrine, "Gigahertz frequency flexible carbon nanotube transistors," *Applied Physics Letters*, vol. 91, no. 15, pp. 153111-3, Oct. 2007.

[55] Q. Cao et al., "Medium-scale carbon nanotube thin-film integrated circuits on flexible plastic substrates," *Nature*, vol. 454, no. 7203, pp. 495-500, Jul. 2008.

[56] S. J. Tans et al., "Individual single-wall carbon nanotubes as quantum wires," *Nature*, vol. 386, no. 6624, pp. 474-477, Apr. 1997.

[57] C. Kocabas, M. Shim, and J. A. Rogers, "Spatially Selective Guided Growth of High-Coverage Arrays and Random Networks of Single-Walled Carbon Nanotubes and Their Integration into Electronic Devices," *Journal of the American Chemical Society*, vol. 128, no. 14, pp. 4540-4541, Apr. 2006.

[58] J. Kong, H. T. Soh, A. M. Cassell, C. F. Quate, and H. Dai, "Synthesis of individual single-walled carbon nanotubes on patterned silicon wafers," *Nature*, vol. 395, no. 6705, pp. 878-881, Oct. 1998.

[59] A. Javey and J. Kong, *Carbon nanotube electronics*. 2009.

[60] S. Heinze, J. Tersoff, R. Martel, V. Derycke, J. Appenzeller, and P. Avouris, "Carbon Nanotubes as Schottky Barrier Transistors," *Phys. Rev. Lett.*, vol. 89, no. 10, pp. 1068011-1068014, 2002.

[61] A. Javey, R. Tu, D. B. Farmer, J. Guo, R. G. Gordon, and H. Dai, "High Performance n-Type Carbon Nanotube Field-Effect Transistors with Chemically Doped Contacts," *Nano Lett.*, vol. 5, no. 2, pp. 345-348, 2005.

[62] J. Chen, C. Klinke, A. Afzali, and P. Avouris, "Self-aligned carbon nanotube transistors with charge transfer doping," *Appl. Phys. Lett.*, vol. 86, no. 12, p. 123108, 2005.

[63] Yu-Ming Lin, J. Appenzeller, and P. Avouris, "Novel carbon nanotube FET design with tunable polarity," in *Proc. Int. Electron Devices Meet*, pp. 687-690, 2004.

[64] M. J. Biercuk, D. J. Monsma, C. M. Marcus, J. S. Becker, and R. G. Gordon, "Low-temperature atomic-layer-deposition lift-off method for microelectronic and nanoelectronic applications," *Applied Physics Letters*, 2003. .

[65] A. Javey et al., "Self-Aligned Ballistic Molecular Transistors and Electrically Parallel Nanotube Arrays," *Nano Lett.*, vol. 4, no. 7, pp. 1319-1322, Jul. 2004.

[66] A. Raychowdhury, A. Keshavarzi, J. Kurtin, V. De, and K. Roy, "Carbon Nanotube Field-Effect Transistors for High-Performance Digital Circuits-DC Analysis and Modeling Toward Optimum Transistor Structure," *IEEE trans. Electron Devices*, vol. 53, no. 11, pp. 2711-2717, Nov. 2006.

[67] A. Javey et al., "High-K dielectrics for advanced carbon-nanotube transistors and logic gates," *Nat Mater*, vol. 1, no. 4, pp. 241-246, Dec. 2002.

[68] J. Guo et al., "Assessment of Silicon MOS and Carbon Nanotube FET Performance Limits Using a General Theory of Ballistic Transistors," in *Proc. Int. Electron Devices Meet*, pp. 711-714, 2002.

[69] J. Guo, A. Javey, H. Dai, and M. Lundstrom, "Performance Analysis and Design Optimization of Near Ballistic Carbon Nanotube Field-Effect Transistors," in *Proc. Int. Electron Devices Meet*, pp. 703-706, 2004.

[70] J. Deng et al., "Carbon Nanotube Transistor Circuits: Circuit-Level Performance Benchmarking and Design Options for Living with Imperfections," in *Proc. Int. Solid State Circuits Conf.*, pp. 70-588, 2007.

[71] R. Martel, T. Schmidt, H. R. Shea, T. Hertel, and P. Avouris, "Single- and multi-wall

carbon nanotube field-effect transistors," *Appl. Phys. Lett.*, vol. 73, no. 17, p. 2447, 1998.

[72]   Y. C. Tseng, K. Phoa, D. Carlton, and J. Bokor, "Effect of Diameter Variation in a Large Set of Carbon Nanotube Transistors," *Nano Lett.*, vol. 6, no. 7, pp. 1364-1368, Jul. 2006.

[73]   A. Raychowdhury, V. De, J. Kurtin, S. Borkar, K. Roy, and A. Keshavarzi, "Variation Tolerance in a Multichannel Carbon-Nanotube Transistor for High-Speed Digital Circuits," *IEEE Transactions on Electron Devices*, vol. 56, no. 3, pp. 383-392, Mar. 2009.

[74]   S. Han, X. Liu, and C. Zhou, "Template-Free Directional Growth of Single-Walled Carbon Nanotubes on a- and r-Plane Sapphire," *Journal of the American Chemical Society*, vol. 127, no. 15, pp. 5294-5295, Apr. 2005.

[75]   P. Avouris, J. Appenzeller, R. Martel, and S. J. Wind, "Carbon Nanotube Electronics," *Proc. IEEE*, vol. 91, no. 11, pp. 1772-1783, Nov. 2003.

[76]   X. Liu et al., "Detailed Analysis of the Mean Diameter and Diameter Distribution of Single-Wall Carbon Nanotubes from their Optical Response," *Physical Review B*, vol. 66, no. 4, p. 45411, Jul. 2002.

[77]   J. Deng, *Device Modeling and Circuit Performance Evaluation for Nanoscale Devices*. 2007.

[78]   N. Patil, Jie Deng, S. Mitra, and H. Wong, "Circuit-Level Performance Benchmarking and Scalability Analysis of Carbon Nanotube Transistor Circuits," *IEEE Trans. Nanotechnol*, vol. 8, no. 1, pp. 37-45, Jan. 2009.

[79]   [Online]   Available:, "http://www.itrs.net/Links/2007ITRS/2007_Chapters/2007_ERD.pdf," *Emerging*

Research Devices. [Online]. [Accessed: 24-Feb-2010].

[80] Y. Lin, J. Appenzeller, J. Knoch, and P. Avouris, "High-Performance Carbon Nanotube Field-Effect Transistor With Tunable Polarities," *IEEE Trans. Nanotechnol.*, vol. 4, no. 5, pp. 481-489, Sep. 2005.

[81] S. J. Kang et al., "High-Performance Electronics Using Dense, Perfectly Aligned Arrays of Single-Walled Carbon Nanotubes," *Nat. Nanotechnol.*, vol. 2, no. 4, pp. 230-236, Apr. 2007.

[82] N. Patil et al., "Wafer-Scale Growth and Transfer of Aligned Single-Walled Carbon Nanotubes," *IEEE Trans. Nanotechnol.*, vol. 8, no. 4, pp. 498-504, Jul. 2009.

[83] Jie Deng and H. Wong, "A Compact SPICE Model for Carbon-Nanotube Field-Effect Transistors Including Nonidealities and Its Application—Part I: Model of the Intrinsic Channel Region," *IEEE Transactions on Electron Devices*, vol. 54, no. 12, pp. 3186-3194, 2007.

[84] Jie Deng and H. Wong, "Modeling and Analysis of Planar-Gate Electrostatic Capacitance of 1-D FET With Multiple Cylindrical Conducting Channels," *IEEE Transactions on Electron Devices*, vol. 54, no. 9, pp. 2377-2385, 2007.

[85] N. Patil, J. Deng, H. P. Wong, and S. Mitra, "Automated Design of Misaligned-Carbon-Nanotube-Immune Circuits," in *Proc. ACM/IEEE Design Automat. Conf.*, pp. 958-961, 2007.

[86] C. Zhou, J. Kong, and H. Dai, "Electrical measurements of individual semiconducting single-walled carbon nanotubes of various diameters," *Appl. Phys. Lett.*, vol. 76, no. 12, p. 1597, 2000.

[87] J. Appenzeller, J. Knoch, M. Radosavljević, and P. Avouris, "Multimode Transport

in Schottky-Barrier Carbon-Nanotube Field-Effect Transistors," *Phys. Rev. Lett.*, vol. 92, no. 22, 2004.

[88]   W. Kim, A. Javey, R. Tu, J. Cao, Q. Wang, and H. Dai, "Electrical contacts to carbon nanotubes down to 1 nm in diameter," *Applied Physics Letters*, vol. 87, no. 17, p. 173101, 2005.

[89]   M. Radosavljevic, J. Appenzeller, P. Avouris, and J. Knoch, "High performance of potassium n-doped carbon nanotube field-effect transistors," *Applied Physics Letters*, 2004. .

[90]   P. Nikolaev et al., "Gas-phase catalytic growth of single-walled carbon nanotubes from carbon monoxide," *Chemical Physics Letters*, vol. 313, no. 1, pp. 91-97, Nov. 1999.

[91]   Y. Li et al., "Preferential Growth of Semiconducting Single-Walled Carbon Nanotubes by a Plasma Enhanced CVD Method," *Nano Lett.*, vol. 4, no. 2, pp. 317-321, 2004.

[92]   L. Qu, F. Du, and L. Dai, "Preferential Syntheses of Semiconducting Vertically Aligned Single-Walled Carbon Nanotubes for Direct Use in FETs," *Nano Lett.*, vol. 8, no. 9, pp. 2682-2687, Sep. 2008.

[93]   P. G. Collins, M. S. Arnold, and P. Avouris, "Engineering Carbon Nanotubes and Nanotube Circuits Using Electrical Breakdown," *Science*, vol. 292, no. 5517, pp. 706-709, Apr. 2001.

[94]   G. Zhang et al., "Selective Etching of Metallic Carbon Nanotubes by Gas-Phase Reaction," *Science*, vol. 314, no. 5801, pp. 974-977, Nov. 2006.

[95]   N. Patil et al., "VMR: VLSI-compatible metallic carbon nanotube removal for

imperfection-immune cascaded multi-stage digital logic circuits using Carbon Nanotube FETs," in *Proc. Int. Electron Devices Meet*, pp. 1-4, 2009.

[96] C. Kocabas et al., "High-Frequency Performance of Submicrometer Transistors That Use Aligned Arrays of Single-Walled Carbon Nanotubes," *Nano Lett.*, vol. 9, no. 5, pp. 1937-1943, May. 2009.

[97] C. Kocabas et al., "Radio frequency analog electronics based on carbon nanotube transistors," *Proceedings of the National Academy of Sciences*, vol. 105, no. 5, pp. 1405-1409, Feb. 2008.

[98] Z. Y. Zhang et al., "High-Performance N-Type Carbon Nanotube Field-Effect Transistors with Estimated Sub-10-ps Gate Delay," *Applied Physics Letters*, vol. 92, no. 13, p. 133117, Mar. 2008.

[99] R. Ashraf, M. Chrzanowska-Jeske, and S. Narendra, "Carbon Nanotube Circuit Design Choices in the Presence of Metallic Tubes," in *Proc. IEEE Int. Symp. Circuits Syst.*, pp. 177-180, 2008.

[100] N. Patil, Jie Deng, A. Lin, H. Wong, and S. Mitra, "Design Methods for Misaligned and Mispositioned Carbon-Nanotube Immune Circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 27, no. 10, pp. 1725-1736, 2008.

[101] Jie Deng and H. Wong, "A Compact SPICE Model for Carbon-Nanotube Field-Effect Transistors Including Nonidealities and Its Application—Part II: Full Device Model and Circuit Performance Benchmarking," *IEEE trans. Electron Devices*, vol. 54, no. 12, pp. 3195-3205, 2007.

[102] J. Deng and H. Wong, "A Circuit-Compatible SPICE model for Enhancement

Mode Carbon Nanotube Field Effect Transistors," in *International Conference on Simulation of Semiconductor Processes and Devices, SISPAD*, pp. 166-169, 2006.

[103] I. Amlani, J. Lewis, K. Lee, R. Zhang, J. Deng, and H. P. Wong, "First Demonstration of AC Gain From a Single-walled Carbon Nanotube Common-Source Amplifier," in *Proc. Int. Electron Devices Meet*, pp. 1-4, 2006.

[104] R. Ashraf, C. Malgorzata, and N. Siva G., "(Submitted) Yield Enhancement Techniques for CNFET based Circuit in the Presence of Imperfections," *IEEE Trans. Nanotechnol.*

[105] S. Shukla, *Nano, quantum, and molecular computing implications to high level design and validation.* Boston: Kluwer Academic Publishers, 2004.

[106] J. Zhang, N. Patil, A. Hazeghi, and S. Mitra, "Carbon nanotube circuits in the presence of carbon nanotube density variations," in *Proc. ACM/IEEE Design Automat. Conf.*, pp. 71-76, 2009.

[107] V. Kheterpal et al., "Design methodology for IC manufacturability based on regular logic-bricks," in *Proc. ACM/IEEE Design Automat. Conf*, pp. 353-358, 2005.

[108] A. Lin, N. Patil, W. Hai, S. Mitra, and H. Wong, "A Metallic-CNT-Tolerant Carbon Nanotube Technology Using Asymmetrically-Correlated CNTs (ACCNT)," in *Proc. IEEE Int. Symp. VLSI Technol.*, pp. 182-183, 2009.

[109] Jie Zhang, N. Patil, and S. Mitra, "Probabilistic Analysis and Design of Metallic-Carbon-Nanotube-Tolerant Digital Logic Circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,*, vol. 28, no. 9, pp. 1307-1320, 2009.

[110] S. Narendra, S. Borkar, V. De, D. Antoniadis, and A. Chandrakasan, "Scaling of Stack Effect and its Application for Leakage Reduction," in *Intl. Symp. Low Power*

Electronics and Design, pp. 195-200, 2001.

[111] Y. Ye, S. Borkar, and V. De, "A new technique for standby leakage reduction in high-performance circuits," in *Symp. of VLSI circuits*, pp. 40-41, 1998.

[112] J. Halter and F. Najm, "A gate-level leakage power reduction method for ultra-low-power CMOS circuits," in *Custom Integrated Circuits Conf.*, pp. 475-478, 1997.

[113] Zhanping Chen, M. Johnson, Liqiong Wei, and W. Roy, "Estimation of standby leakage power in CMOS circuit considering accurate modeling of transistor stacks," in *Intl. Symp. Low Power Electronics and Design*, pp. 239-244, 1998.

[114] R. Ashraf, R. Nain, M. Chrzanowska-Jeske, and S. Narendra, "Design methodology for Carbon Nanotube based circuits in the presence of metallic tubes," in *IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH)* ., pp. 71-76, 2010.

[115] J. Zhang et al., "Carbon Nanotube Correlation Promising Opportunity for CNFET Circuit Yield Enhancement," in *Proc. ACM/IEEE Design Automat. Conf*, 2010.

[116] R. Ashraf, M. Chrzanowska-Jeske, and S. G. Narendra, "Functional Yield Estimation of Carbon Nanotube based Logic Gates in the Presence of Defects," *IEEE Trans. Nanotechnol*, vol. 9, no. 6, pp. 687-700, Nov. 2010.

[117] R. Ashraf, N. Rajeev, C. Malgorzata, and N. Siva G., "Yield Enhancement by Tube Redundancy in CNFET-Based Circuits," in *IEEE International Conference on Electronics, Circuits, and Systems.*, 2010.

[118] K. A. Bowman, S. G. Duvall, and J. D. Meindl, "Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution for gigascale integration," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 2, pp. 183-190, 2002.

Appendix A

A.1 Nomenclature

ρ_S	Resistance per unit length of source region
a	Carbon to carbon atom distance
C_{g_fo}	Gate capacitance of the fanout gate(s)
C_{L_ideal}	Load capacitance of the gate when all the tubes are semiconducting
C_{L_turn}	Total load capacitance after the removal of tubes from both drive and fanout gate(s)
C_{p_dr}	Parasitic capacitance of the driving gate
C_{p_fo}	Parasitic capacitance of the fanout gate(s)
d	Diameter of CNT
D_{μ_Inv}	Mean delay of inverter
D_{μ_NAND}	Mean delay of NAND gate
D_g	Delay of a logic gate
Dg_ideal	Delay of the gate under an ideal scenario
$D_{HL/LH}$	Worst case delay for high-to-low/low-to-high transition
$D_{PU/PD}$	Delay of pull-up/pull-down network
D_{g_turn}	Number of tubes removed from the ON network of driving gate
e	Charge on electron
g_{CNT}	Transconductance of CNFET
I_{on}	ON current in a single CNT
I_{ons}	ON current of a semiconducting tube
I_{onm}	ON current of a metallic tube
I_{offs}	OFF current of a semiconducting tube
I_{offm}	OFF current of a metallic tube
I_{onm}	Mean value of ON current of a semiconducting
$I_{ON\_PU/PD}$	ON current of pull-up/pull-down network
$I_{OFF\_PU/PD}$	OFF current of pull-up/pull-down network

L_S	Length of doped CNT which acts as a source
n	Sample size
N_m	Maximum number of metallic tubes tolerated in a network
N_C:	Allowed combinations of tubes removed from the drive and fanout gate(s)
N_{m_Inv}	Maximum number of metallic tubes that can be tolerated in PU/PD network of inverter
N_{mPU_NAND}	Maximum number of metallic tubes that can be tolerated in PU network for the PD network to be functional in NAND gate
N_{mPD_NAND}	Maximum number of metallic tubes that can be tolerated in PD network for the PU network to be functional in NAND gate
N_{tu_dr}	Number of tubes in the ON network of driving gate
N_{tu_fo}	Total number of tubes in the fanout gate(s)
N_{tur}	Number of tubes in a transistor
$Nturm\_dr$	Number of tubes removed from the ON network of driving gate
N_{turm_fo}	Number of tubes removed from the fanout gate(s)
N_{tusr}	Number of stacked tubes in a transistor
N_{turP}	Number of tubes in P-CNFET of NAND gate
N_{tusrP}	Number of stacked tubes in P-CNFET of NAND gate
N_{turN}	Number of tubes in N-CNFET of NAND gate
N_{tusrN}	Number of stacked tubes in N-CNFET of NAND gate
N_{tug}	Number of tubes in a gate
P_m	Percentage of metallic tubes
P_{UC}	Percentage of un-contacted tubes
Pr_m	Probability of a tube to be metallic
Pr_{ms}	Probability of a stacked tube to be metallic
Pr_{PU}/Pr_{PD}	Probability of PU/PD network being functional
$Pr_{PU\_Inv\_PT}$	Probability of PU network of PT inverter being functional
$Pr_{PD\_Inv\_PT}$	Probability of PD network of PT inverter being functional
$Pr_{PU\_Inv\_TrS}$	Probability of PU network of TrS inverter being functional
$Pr_{PD\_Inv\_TrS}$	Probability of PD network of TrS inverter being functional

$Pr_{PU\_NAND\_PT}$	Probability of *PU* network of *PT* NAND gate being functional
$Pr_{PD\_NAND\_PT}$	Probability of *PD* network of *PT* NAND gate being functional
$Pr_{PU\_NAND\_TrS}$	Probability of *PU* network of *TrS* NAND gate being functional
$Pr_{PD\_NAND\_TrS}$	Probability of *PD* network of *TrS* NAND gate being functional
r	Fitting parameter
S	Subthreshold slope
SP_{μ_Inv}	Average static power of inverter
SP_{μ_NAND}	Average static power of NAND gate
V_{th}	Threshold voltage
V_{π}	Carbon $\pi-\pi$ bonding energy
x_m	Ratio of *ON* current of metallic to semiconducting tube
X_{max}	Ratio of the average delay when some tubes are removed to the average delay when all the tubes are present and semiconducting
Y_f	Functional yield of a gate
Y_{f_inv}	Functional yield of inverter
$Y_{f\_Inv\_PT}$	Functional yield of an inverter with *PT* transistors
$Y_{f\_Inv\_TrS}$	Functional yield of an inverter with *TrS* transistors
$Y_{f\_Inv\_TuS}$	Functional yield of an inverter with *TuS* transistors
Y_{f_NAND}	Functional yield of NAND gate
$Y_{f\_NAND\_PT}$	Functional yield of NAND gate with *PT* transistors
$Y_{f\_NAND\_TrS}$	Functional yield of NAND gate with *TrS* transistors
$Y_{f\_NAND\_TuS}$	Functional yield of NAND gate with *TuS* transistors

A.2   Acronyms

CNT	Carbon Nanotube
CNFET	Carbon nanotube field-effect transistor
N-CNFET	N-type CNFET
P-CNFET	P-type CNFET
PT	Parallel tube
ST	Shared tube

| *TrS* | Transistor stacking |
| *TuS* | Tube stacking |

Flowcharts of Monte Carlo Simulations

B.1   MC Flow when Metallic Tubes are Present (MTP)

B.1.1   Flow MTP 1

Generating delay, static power distributions and calculating functional yield of logic gates when metallic tubes are present.

```
        ┌─────────────┐
        │    Start    │
        └──────┬──────┘
               │
               ▼
     ╱───────────────────╲
    ╱  N_tug, P_m, d_μ, dσ ╲
    ╲───────────────────────╱
               │
               ▼
      ┌──────────────────┐
      │ Generate CNT     │
      │ diameter         │
      │ distribution     │
      └────────┬─────────┘
               │
               ▼
```

Start

N_{tug}, P_m, d_{μ}, $d\sigma$

Generate CNT diameter distribution

Calculate currents of individual CNTs
http://nano.stanford.edu/models.php

Randomly assign tubes to be metallic and recalculate their OFF currents

Distribution of currents for individual CNTs

Calculate currents for CNFETs with different numbers of CNTs

Calculate Delay and Static Power of different CNFET configurations and different logic gates

1) Distribution
    Delay
    Static Power
2) Functional Yield

End

B.1.2 Results obtained from Monte Carlo simulation using Flow MTP 1
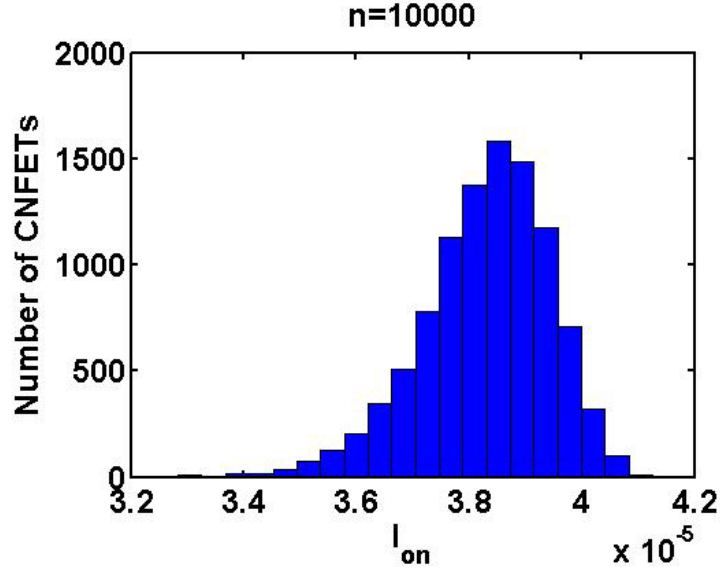
Figure B-1: Distribution of *ON* current for CNT diameter distribution with μ=1.5nm and 3σ=0.5nm.

B.1.3 Output Distributions from Monte Carlo Flow MTP 1

Distributions of delay and static Power of parallel tube inverter with N_{tug}=16
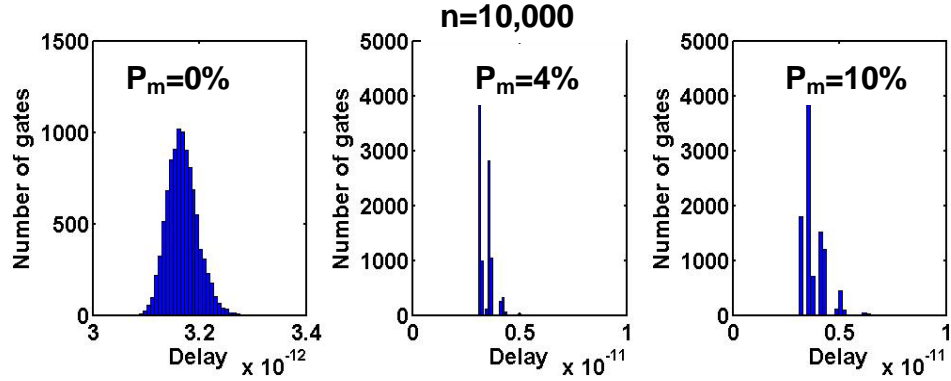
Figure B-2: Monte Carlo simulation for Parallel Tube (*PT*) inverters with N_{tug}=16, showing actual delay distribution for (a) absence of metallic tubes − P_m=0%, (b) P_m=4% metallic tubes and (c) P_m=10% metallic tubes.
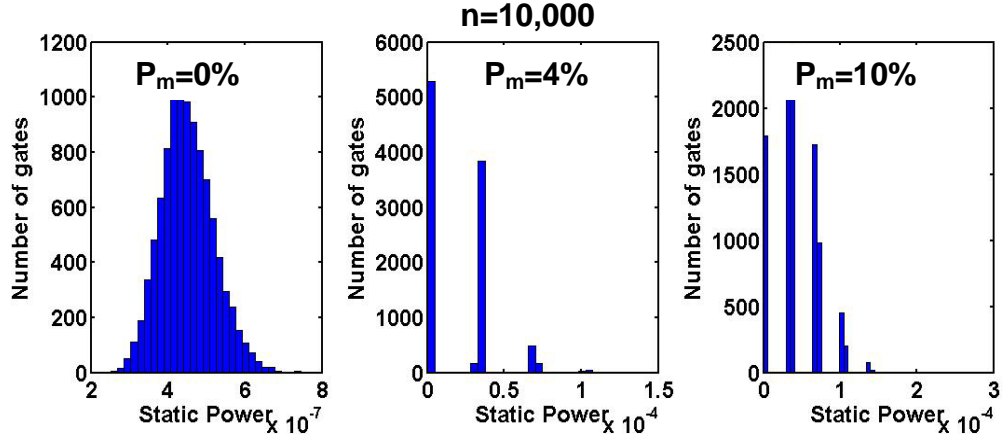
Figure B-3: Monte Carlo simulation for Parallel Tube (*PT*) inverters with N_{tug}=16, showing actual static power distribution for (a) absence of metallic tubes – P_m=0%, (b) P_m=4% metallic tubes and (c) P_m=10% metallic tubes.

From Figure B-2 and Figure B-3 it can be observed that the presence of metallic tubes results in increase in the delay and static power consumption of gates. Based on the delay and static power constraints defined in Chapter 3 we calculate the number of gates whose delay and static power are less than the maximum defined constraints. In the dissertation we normalized the delay and static power because we want to see the impact of metallic tubes on the delay and static power as compared to the case when all the tubes are semiconducting.

B.1.4   MC simulation results showing both functional and non-functional gates

Figure B-4 and Figure B-5 shows the distributions of functional as well as non-functional gates in the presence of metallic tubes.
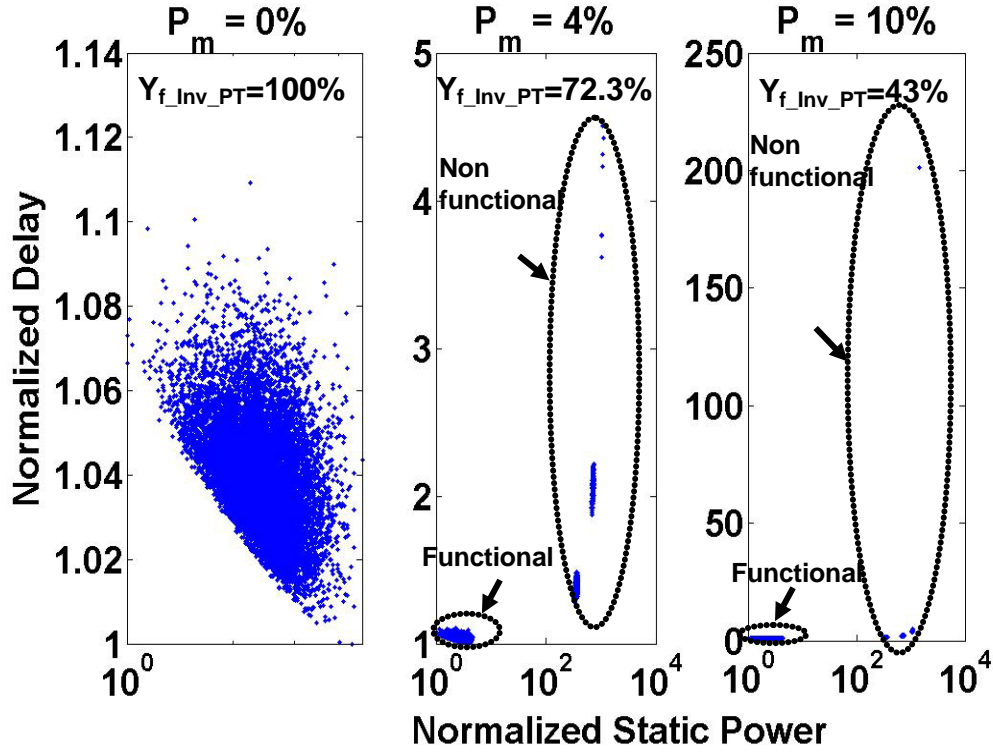
Figure B-4: Monte Carlo simulation for Parallel Tube (*PT*) inverters with N_{tug}=8, showing normalized delay vs. static power for (a) absence of metallic tubes – P_m=0%, (b) P_m=4% metallic tubes and (c) P_m=10% metallic tubes with delay constraint of 1.3X and static power constraint of 200X for functional yield calculation.
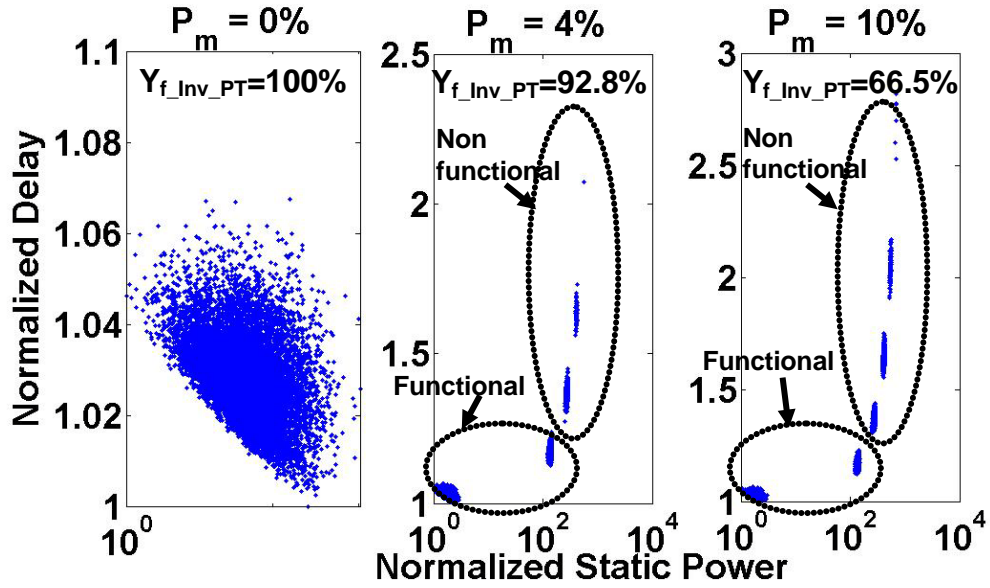
Figure B-5: Monte Carlo simulation for Parallel Tube (*PT*) inverters with N_{tug}=16, showing normalized delay vs. static power for (a) absence of metallic tubes – P_m=0%, (b) P_m=4% metallic tubes and (c) P_m=10% metallic tubes with delay constraint of 1.3X and static power constraint of 200X for functional yield calculation.

B.1.5 MC simulation results showing cut-off values of delay slicing the density

Figure B-6 and Figure B-7 shows that the defined delay constraint of 1.3X, slicing some distribution of gates.
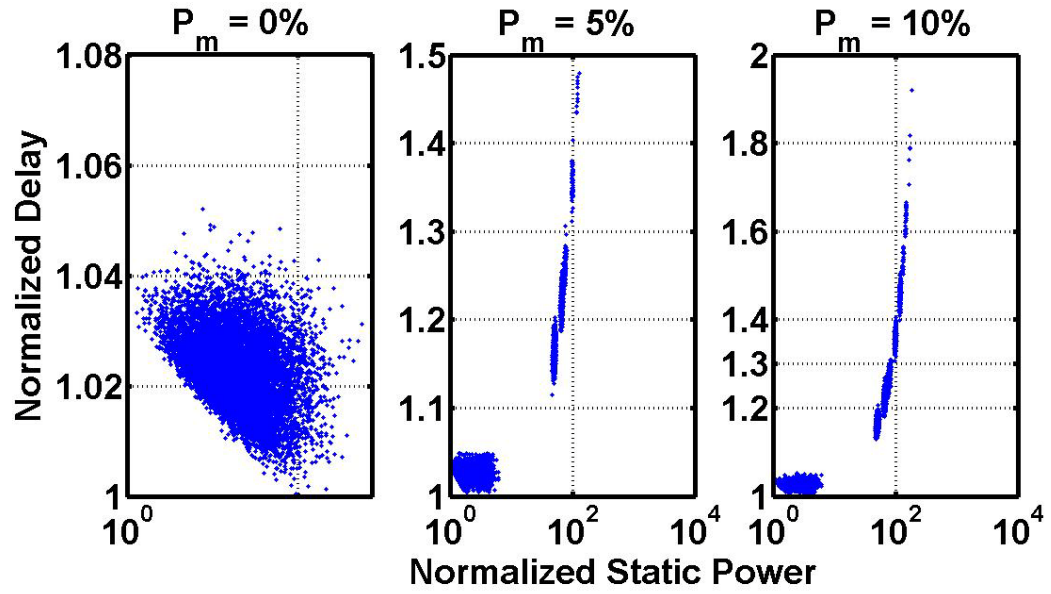
Figure B-6: Monte Carlo simulation for Parallel Tube (*PT*) inverters with N_{tug}=32, showing normalized delay vs. static power for (a) absence of metallic tubes – P_m=0%, (b) P_m=5% metallic tubes and (c) P_m=10% metallic tubes with delay constraint of 1.3X and static power constraint of 200X for functional yield calculation.
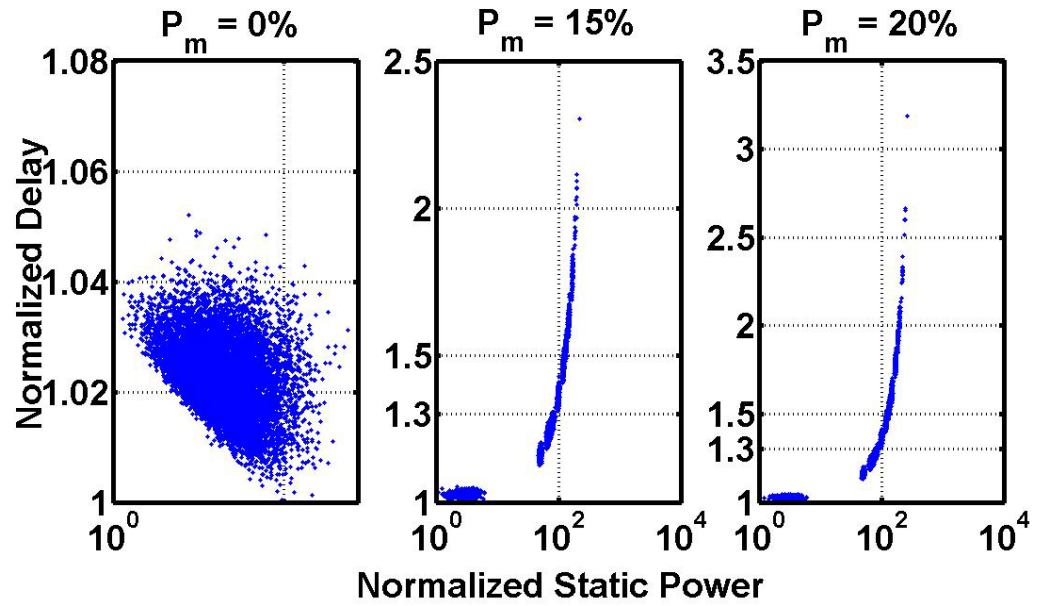
Figure B-7: Monte Carlo simulation for Parallel Tube (*PT*) inverters with *N*_{tug}=32, showing normalized delay vs. static power for (a) absence of metallic tubes – *P*_m=0%, (b) *P*_m=15% metallic tubes and (c) *P*_m=20% metallic tubes with delay constraint of 1.3X and static power constraint of 200X for functional yield calculation.

B.2   Scalability Analysis of Functional Yield when MTP

Figure B-8 and Figure B-9 shows the functional yield of inverters and NAND gates for different percentage of metallic tubes.
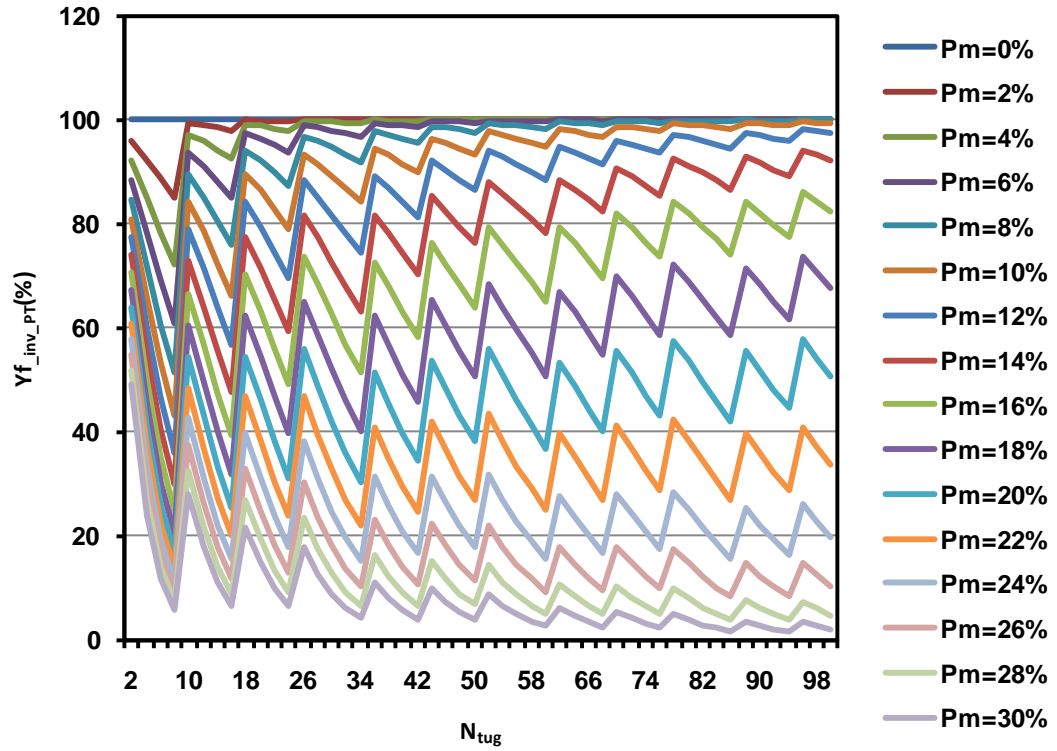
Figure B-8: Functional Yield of Inverter for different number of tubes in the gates(N_{tug}) and different percentage of metallic tubes(P_m)

The functional yield finally approaches acceptable level for up to 10% metallic tubes if we have sufficient number of tubes in the gate. The main problem will be large increase in the area of the gates and dynamic power consumption.
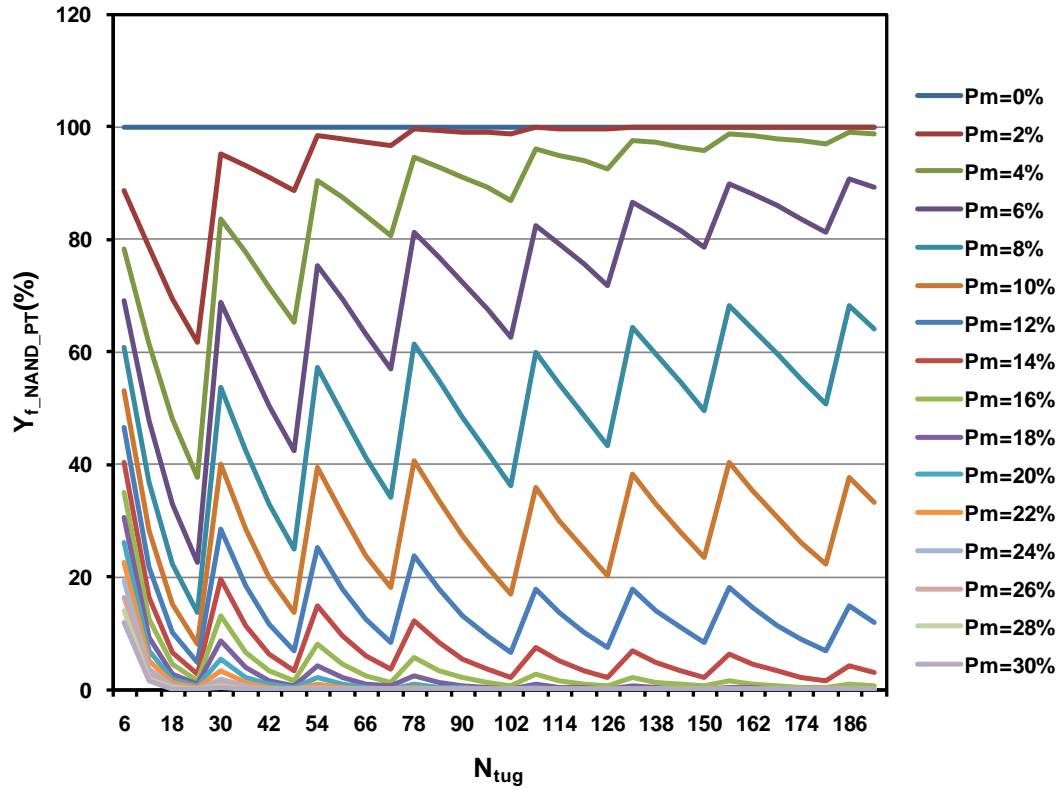
Figure B-9: Functional Yield of NAND gate for different number of tubes in the gates(N_{tug}) and different percentage of metallic tubes(P_m)

The functional yield finally approaches acceptable level for up to 4% metallic tubes if we put sufficient number of tubes in the gate. The trade-off will be large increase in the area of the gates and dynamic power consumption.

B.3　　　　　Building CNFET -based circuit Architecture

Adders are implemented with NAND gates and inverters. Based on the delay and power of inverters and NAND gates, we obtain the delay and power of different paths of adder and finally obtain functional yield of adders, (1) when gates in the adder are implemented with Parallel Tube (*PT*) configurations of transistors, and (2) when gates in the adder are implemented with Transistor Stacking (*TrS*) configurations of gates and Parallelism in the critical path.

B.3.1 Flow MTP 2

Yield of Bricks for various transistor configurations
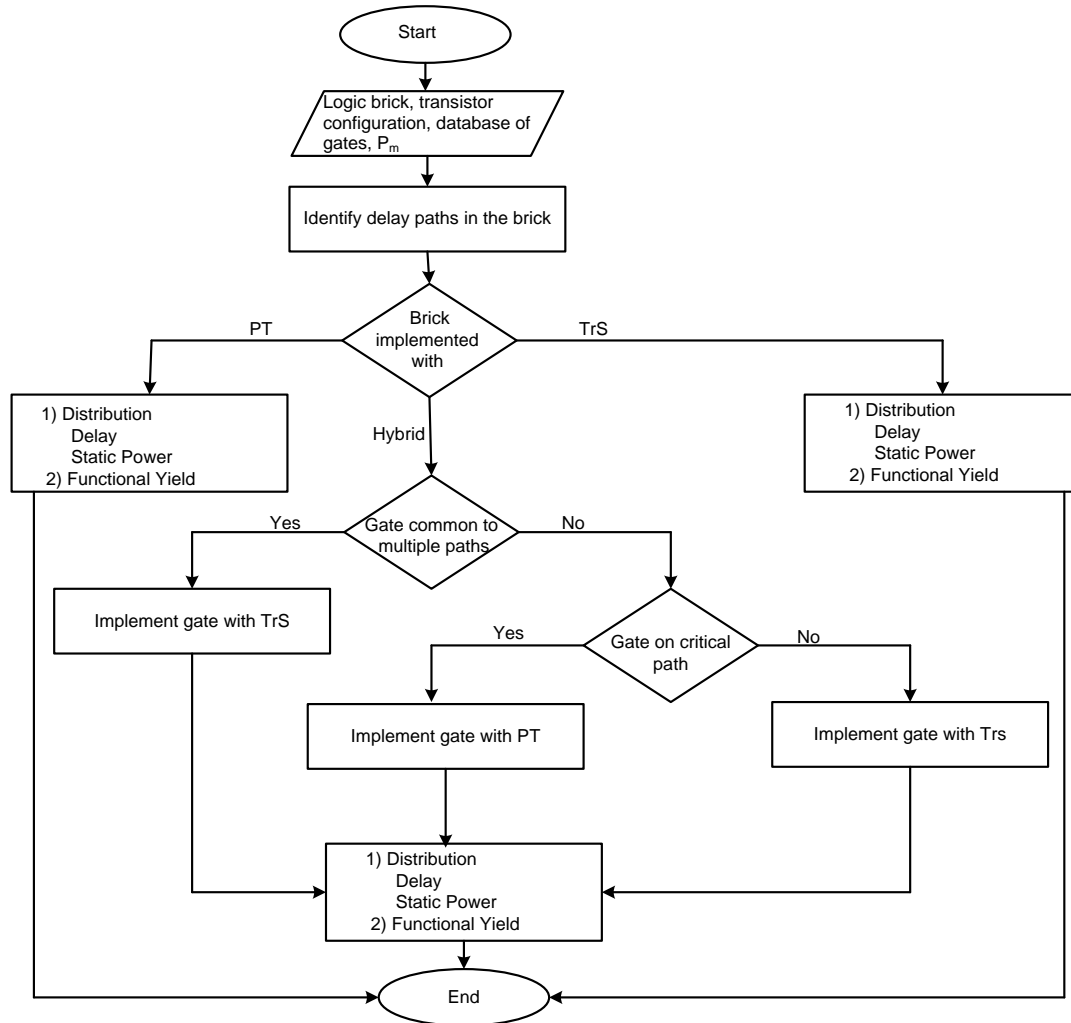
B.4 MC flow for CNFETs when Metallic Tubes are Removed (MTR)

B.4.1 Flow MTR 1

Distribution of CNFETs when MTR

```
                    ┌─────────────┐
                   (    Start     )
                    └─────────────┘
                          │
                          ▼
                  ╱───────────────────╲
                 ╱ N_tur, P_m,P_r,D_CS,D_CM╲
                ╱  d_μ,d_σ,S_μ, S_σ         ╲
               ╱─────────────────────────────╲
                          │
                          ▼
                  ┌─────────────────┐
                  │ Generate CNT    │
                  │ diameter        │
                  │ distribution    │
                  └─────────────────┘
                          │
                          ▼
                  ┌──────────────────────────┐
                  │ Calculate currents of    │
                  │ individual CNTs          │
                  │ http://nano.stanford.edu/models.php │
                  └──────────────────────────┘
                          │
                          ▼
                  ┌──────────────────────────┐
                  │ Calculate μ and σ of I_ON│
                  │ and I_OFF currents of CNFETs│
                  └──────────────────────────┘
                          │
                          ▼
                  ┌──────────────────────────┐
                  │ Calculate μ and σ of I_ON│
                  │ and I_OFF currents of    │
                  │ CNFETs with charge       │
                  │ screening effect (spacing│
                  │ between tubes)           │
                  └──────────────────────────┘
                          │
                          ▼
                  ┌──────────────────────────┐
                  │ Calculate μ and σ of I_ON│
                  │ and I_OFF currents of    │
                  │ CNFETs in the presence of│
                  │ pitch(diameter & spacing)│
                  │ variation                │
                  └──────────────────────────┘
                          │
                          ▼
                  ┌──────────────────────────┐
                  │ Randomly assign tubes to │
                  │ be metallic and recalculate│
                  │ their OFF currents       │
                  └──────────────────────────┘
                          │
                          ▼
                  ┌──────────────────────────┐
                  │ Randomly choose tubes to │
                  │ be removed by SCE process│
                  │ and recalculate ON and OFF│
                  │ currents of effected tubes│
                  └──────────────────────────┘
                          │
                          ▼
                  ┌──────────────────────────┐
                  │ Calculate currents of    │
                  │ CNFETs in the presence of│
                  │ pitch variation          │
                  └──────────────────────────┘
                          │
                          ▼
                  ┌──────────────────────────┐
                  │ Create a Databse of CNFETs│
                  │ with their I_ON and I_OFF currents│
                  └──────────────────────────┘
                          │
                          ▼
                    ┌─────────────┐
                   (     End      )
                    └─────────────┘
```

B.4.2 Flow MTR 2

Monte Carlo Simulation for distribution of delay, static power, energy and functional yield of gates when MTR
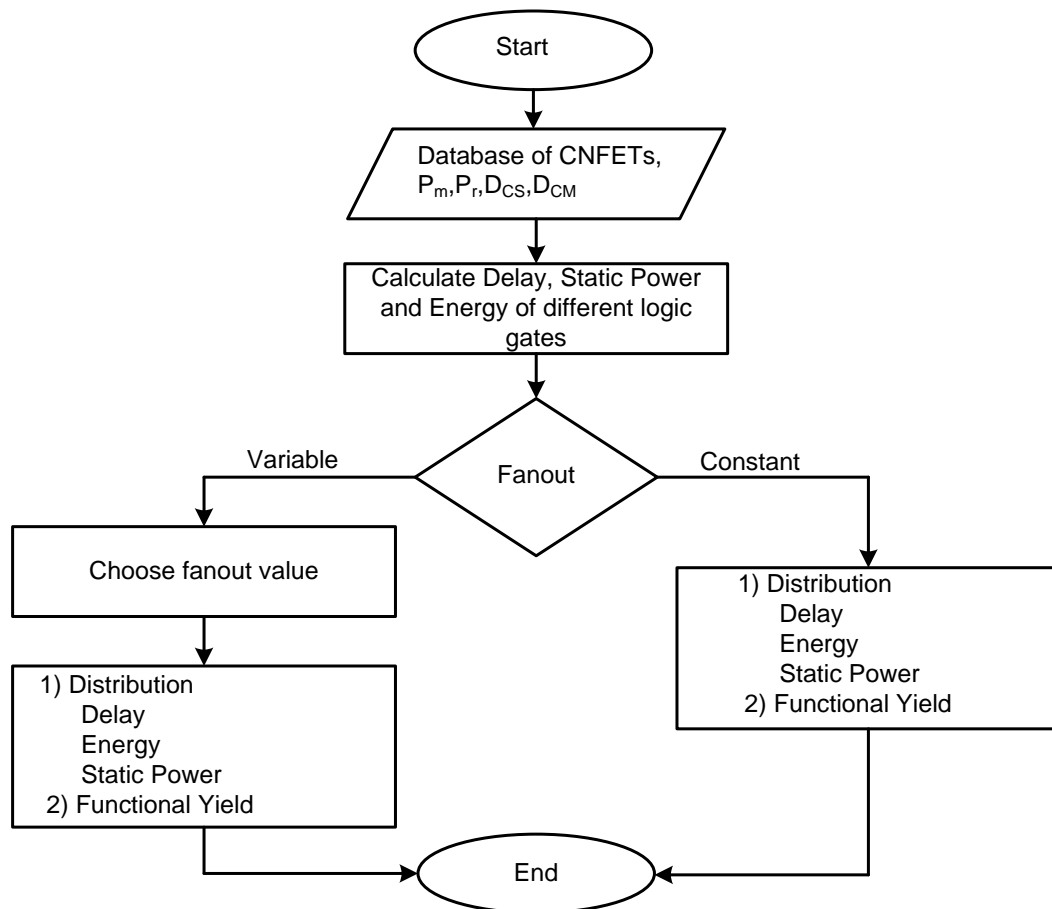
B.4.3   Flow MTR 3

Monte Carlo Simulation for distribution of delay, static power, energy and functional yield for logic paths when MTR

```
                    ┌─────────────┐
                    │    Start    │
                    └──────┬──────┘
                           │
                    ╱──────┴──────╲
                   │ N_tug, P_m, P_r, D_CS, │
                   │   D_CM, d_path         │
                    ╲──────┬──────╱
                           │
              ┌────────────┴────────────┐
              │ Build a LUT of minimum number │
              │ of tubes required in CNFETs to │
              │ avoid open circuit gates after │
              │            SCE            │
              └────────────┬────────────┘
                           │
                    ◇──────┴──────◇
          No       ◇     TLR     ◇       Yes
       ┌───────────◇            ◇───────────┐
       │            ◇──────┬──────◇          │
       │                                     │
┌──────┴──────┐              ┌───────────────┴───────────────┐
│ Calculate path depth │     │ Generate a LUT with the        │
│ required to obtain   │     │ number of redundant tubes      │
│ acceptable yield     │     │ required to obtain μ number of │
└──────┬──────┘        │     │ tubes in a CNFET, after SCE    │
       │                     │ process, what is equal to the  │
       │                     │ mean number of tubes required  │
┌──────┴──────┐              │          before SCE"           │
│ 1) Distribution │          └───────────────┬───────────────┘
│    Delay        │                          │
│    Energy       │          ┌───────────────┴───────────────┐
│    Static Power │          │ Calculate Delay, Static power  │
│ 2) Functional Yield │      │ and Energy of different logic  │
│ 3) Area overhead │         │            gates               │
└──────┬──────┘              │ Calculate Area impact due to   │
       │                     │            TLR                 │
       │                     └───────────────┬───────────────┘
       │                                     │
       │                     ┌───────────────┴───────────────┐
       │                     │ Calculate Delay, Static power  │
       │                     │ and Energy for different logic │
       │                     │           depths               │
       │                     │ Calculate Area impact due to   │
       │                     │ TLR for different logic depths │
       │                     └───────────────┬───────────────┘
       │                                     │
       │                     ┌───────────────┴───────────────┐
       │                     │ 1) Distribution                │
       │                     │    Delay                       │
       │                     │    Energy                      │
       │                     │    Static Power                │
       │                     │ 2) Functional Yield            │
       │                     │ 3) Area overhead               │
       │                     └───────────────┬───────────────┘
       │                                     │
       │        ┌─────────────┐              │
       └────────┤     End     ├──────────────┘
                └─────────────┘
```

176