Portland State University

PDXScholar

Dissertations and Theses

Dissertations and Theses

Winter 3-15-2019

Efficient Methods for Robust Circuit Design and Performance Optimization for Carbon Nanotube Field Effect Transistors

Muhammad Ali
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

Part of the [Electrical and Computer Engineering Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Ali, Muhammad, "Efficient Methods for Robust Circuit Design and Performance Optimization for Carbon Nanotube Field Effect Transistors" (2019). *Dissertations and Theses*. Paper 4833.

<https://doi.org/10.15760/etd.6709>

This Dissertation is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Efficient Methods for Robust Circuit Design and Performance Optimization
for Carbon Nanotube Field Effect Transistors

by
Muhammad Ali

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
James Morris
Marek Perkowski
John Acken
Rolf Koenenkamp

Portland State University
2019

ABSTRACT

Carbon nanotube field-effect transistors (CNFETs) are considered to be promising candidate beyond the conventional CMOSFET due to their higher current drive capability, ballistic transport, lesser power delay product and higher thermal stability. CNFETs show great potential to build digital systems on advanced technology nodes with big benefits in terms of power, performance and area (PPA). Hence, there is a great need to develop proven models and CAD tools for performance evaluation of CNFET-based circuits. CNFETs specific parameters, such as number of tubes, pitch (spacing between the tubes) and diameter of CNTs determine current driving capability, speed, power consumption and area of circuits and play a significant role in accurate PPA evaluation. Furthermore, count and density variations in carbon nanotubes (CNTs) due to manufacturing limitations, like the presence of metallic tubes in the CNFET channel, pose major obstacles to robust and energy-efficient CNFET digital circuit designs and degrade the anticipated PPA benefits. CNFET-based circuits can suffer from large performance variations and reduction in functional yield due to these variations in CNFETs. Moreover, modeling the CNFET parameters, CNT variations and etching techniques for CNTs create additional complexity during performance optimization. Hence, for realistic optimization of CNFET circuit's performance, it is imperative to incorporate the impact of these parameters and variations.

We present a capacitance-based Logical Effort (LE) framework to investigate design is-

sues of high-speed and low-power circuit designs implemented by considering specific requirements and challenges of the CNFET technology. The LE technique is widely recognized as a pedagogical method to quickly estimate and optimize the propagation delay and transition time in CMOS circuits equivalently without performing transient simulations and detailed delay calculations. In this thesis, we propose novel delay models [Pitch-Aware Logical Effort (PALE) and Position-Aware Pitch Factor (PAPF)] for fast and accurate performance evaluation by including the impact due to CNFET-specific parameters and CNT variations.

1. Ideal case (CNTs variations are not considered):

During our research on CNFET-based circuits, we analyzed the impact of CNFET specific parameters, such as CNTs count, diameter and spacing between tubes, on the performance of CNFET-based circuits. The screening effect is critical to take into account for accurate performance evaluation. Hence, PALE model is developed by extending LE formulation to include influence of CNFET specific parameters.

2. Realistic case (CNTs variations are considered):

We have studied CNFET-based logic gates and circuits in the presence of major CNTs variations using Monte Carlo simulations. The removal of the initially present unwanted metallic tubes, by the known processing techniques, causes non-uniformity of CNT density in the channel. Such variations in the number of CNTs impact circuit performance and functional yield. We develop variation-aware model (PAPF) based on LE technique to include impact of CNTs variations on the delay of large CNFET-based circuits.

Our developed models are correlated with SPICE simulations using different types of

gates and circuits with an average error of 3% and 5% for ideal and realistic cases respectively. Our framework is capable of estimating performance more than 100x faster as compared to SPICE simulations methods.

Furthermore, using our models (PALE and PAPF), we present an optimization tool to minimize the area and delay product (ADP) of CNFET circuits. We deploy circuit-level techniques (CLT) prior to the optimizing the tubes (CNTs) in the logic gates to achieve highly optimized solution with global approach. For better optimization of the circuits, the impact of wire parasitic in estimating the delay of the individual gates is included as well. Our optimization tool results in maximum and average delay improvement by 27% and 17% respectively, and 2.5X reduction in area for standard ISCAS and OpenSPARC benchmark circuits. Fast and fairly accurate delay computation in our optimization framework offers great runtime benefits as compared to state-of-the-art SPICE simulation and statistical-based methods.

Finally, we propose more accurate probabilistic model for yield estimation which incorporates the impact of screening effect on the functional yield after the removal of metallic tubes.

Overall, the objective of this thesis is to develop comprehensive LE-based framework and optimization tool and methodology which comprehend CNFET specific parameters for accurate performance evaluation as well as estimation of delay, power, functional yield and do ADP optimization in presence of CNTs variations. Our models are easily scalable to future technology nodes.

DEDICATION

To My Parents

ACKNOWLEDGEMENTS

It is with great admiration that I express deep appreciation to my academic advisor, Professor Malgorzata Chrzanowska-Jeske for her support and guidance during my PhD. I have indeed learned many things from her. I highly appreciate her for accommodating my meetings during late hours due to my full time job. I would also like to thank my PhD dissertation committee members, Dr. James Morris, Dr. Marek Perkowski, Dr. John Acken and Dr. Rolf Koenenkamp for their valuable feedback.

I would like to extend my appreciation to Dr. Ivan Sutherland at Portland State University (inventor of standard Logical Effort model) for providing guidance to enhance LE framework for CNFET technology and Dr. Prof. H.S. Philip Wong and his team at Stanford University for the many useful discussions related to CNFET models.

I am thankful to my colleagues Rehman Ashraf, Mohammad A. Ahmed and Sarita T. Sanchez for many technical discussions during my stay at Portland State University. Their help and feedback greatly improved the quality of this work.

My most heartfelt gratitude is for my parents and family for encouragement, love and affection during my PhD. I would like to acknowledge the sacrifice and efforts of my parents to enable me to pursue good education.

Finally, special thanks to my wife Hina, for her encouragement and support by taking

care of family and allowing me to focus on my studies after work-hours and over week-ends. Indeed, her help and care provided me great motivation. My deepest love must go to my kids Mahnoor, Aleena and Zakariyya for their understanding and patience during my PhD.

TABLE OF CONTENTS

Abstract	i
Dedication	iv
Acknowledgements	v
List of Tables	x
List of Figures	xii
Chapter 1: Introduction	1
1.1 Beyond CMOS	1
1.2 Challenges	2
1.3 Carbon nanotube FETs (CNFETs)	5
1.4 Research Contributions	12
1.4.1 Extending Logical Effort model for CNFET-based circuits (Ideal case)	12
1.4.2 Logical Effort based variation aware model (realistic case)	13
1.4.3 Area and Delay Optimization of CNFET-based circuits	14
1.4.4 CNT Position-Aware Yield Estimation model	14
1.5 Outline	15
Chapter 2: CNFET Technology, Advantages and Dis-Advantages	16
2.1 Advantages of CNFETs	16
2.2 CNFET specific and technology parameters	17
2.2.1 CNTs array	17
2.2.2 Pitch	18
2.2.3 Diameter	19
2.2.4 Width of CNFET gate	19
2.2.5 The technology parameters	20
2.3 CNFET Gate Capacitance Model	20
2.4 Charge Screening Effect	21
2.5 The variations and challenges in CNT technology	24

2.5.1	CNT type (m-CNT or s-CNT) variations	24
2.5.2	CNT diameter variations	24
2.5.3	CNT density and Spacing variations	25
2.5.4	Misalignment of CNTs	27
2.5.5	CNT doping variations	27
2.6	Summary and Conclusion	30
Chapter 3: Pitch-Aware Logical Effort Model		32
3.1	The standard Logical Effort Model	34
3.1.1	Evaluation of Logical Effort model for CNFET-based circuits . . .	36
3.1.2	Comparison of CMOS and CNFET Logical Efforts	38
3.2	Limitations of standard LE model for CNFET	39
3.3	Pitch-Aware Logical Effort (PALE) Model	41
3.3.1	Derivation of Pitch-Factor (PF) to model the Screening Effect . . .	41
3.3.2	Derivation of Pitch-Aware Logical Effort (PALE) and Estimation of Normalized Delay	43
3.3.3	A Reference Inverter	46
3.3.4	Estimation of τ for Absolute Delay calculation in CNFET-based circuits	46
3.4	Experiment Results	48
3.4.1	CNFET Gate Level Results	49
3.4.2	CNFET Circuit Level Results	53
3.5	Summary and Conclusion	58
Chapter 4: Position-Aware Pitch-Factor Model		59
4.1	Evenly-spaced LE (ESLE) Model	61
4.2	Position-Aware Pitch-Factor (PAPF) Model	62
4.2.1	Statistical Analysis of Capacitance in the presence of Metallic Tubes	65
4.2.2	PAPF Model Closed-form Expression	67
4.3	Experiment Results for PAPF LE Model	70
4.3.1	Comparison of PAPF and ESLE	71
4.3.2	PAPF correlation with SPICE Simulation Methods	71
4.3.3	Impact of C_{gtg} on the Variance in Statistical Delay	73
4.3.4	Decoder circuit-level Analysis	73
4.3.5	Runtime Analysis	75
4.4	Summary and Conclusion	76
Chapter 5: Optimization of CNFET Circuits		78
5.1	Optimization Methodology	79
5.1.1	Optimization Flow	79
5.1.2	Cost Function	81

5.1.3	Algorithm	82
5.2	Parameters Estimation	83
5.2.1	Delay Estimation	83
5.2.2	Impact of Wire Load and Estimation	85
5.2.3	Area Estimation	87
5.3	Experiment Results	88
5.4	Summary and Conclusion	92
Chapter 6:	CNT position-aware Yield estimation	93
6.1	Functional Yield of Inverter	96
6.2	Functional Yield of 3-stage Inverter:	98
6.3	Experimental Results	98
6.3.1	Analysis of single Inverter Gate	99
6.3.2	Analysis of 2-stage Inverter Gate	101
6.4	Summary and Conclusion	103
Chapter 7:	Contributions, Conclusions & Future Work	104
7.1	Contributions and Conclusions	105
7.1.1	Conclusions	107
7.2	Future Work	108
Bibliography		110
Appendix: Glossary		121
A.1	Nomenclature	121
A.2	Acronyms	123

LIST OF TABLES

Table 2.1	Description of technology parameters	20
Table 3.1	Logical Effort, using actual input gate capacitance, of CNFET-based gates with $N_{tur}=8$ in each transistor.	37
Table 3.2	Logical Effort of CNFET versus CMOS gates	38
Table 3.3	Look-up-table (LUT): Pitch factor (PF) values for different spacing between CNTs	43
Table 3.4	Logical Effort of different CNFET inverter configuration for same width	49
Table 3.5	Gate level comparison of Delay computation from PALE and SPICE simulations with different FO	50
Table 3.6	Circuit level comparison of Delay computation from model and SPICE simulation for a) multi-stage circuits b) multi-branch circuits for FO4	56
Table 3.7	Test-cases description	56
Table 3.8	Comparison of Delay from LE model and SPICE simulations for 3-stage Decoder, with B = 8 and H = 9.2	57
Table 4.1	Mean and Variance of 1000 inverter instances in the presence of different percentage of metallic tubes and their removal technique	69
Table 4.2	PAPF for VMR and SCE removal techniques for different logic gates at different percentage of P_m	70
Table 4.3	Runtime comparison between SPICE simulations and CNFET LE model.	76
Table 5.1	Total delay and area of ISCAS and OpenSPARC benchmark circuits with tube optimization and with tube + CLT both including runtime comparison of our Algorithm with linear model (LM) and non-linear model (NLM) on a single 2.6GHz processor with no parallelization	89
Table 6.1	Comparison of number of CNTs and pitch desired to meet Delay limits	97

Table 6.2	Comparison of desired tubes and pitch for functional inverter using existing and proposed models for CNT pitch 4nm	100
Table 6.3	Comparison of desired tubes and pitch for functional inverter using existing and proposed models for CNT pitch 2nm	100
Table 6.4	Comparison of accuracy and runtime using existing and proposed models	103

LIST OF FIGURES

Figure 1.1	Relationship of More Moore, Beyond CMOS, and Novel Computing Paradigms and Applications [1].	2
Figure 1.2	Options for Emerging Memory Devices [1].	4
Figure 1.3	Options for Emerging Logic Devices [1].	5
Figure 1.4	(a) Side view of a CNFET layout (b) Top view of CNFET layout with array of six CNTs.	6
Figure 1.5	Molybdenum (Mo) end-contacted SWCNT or SWNT transistors. Figures are showing the conversion from a side-bonded contact (a), where the SWCNT is partially covered by Mo, to end-bonded contact (b), where the SWCNT is attached to the bulk Mo electrode through carbide bonds while the C atoms from originally covered portion of the SWCNT uniformly diffuse out into the Mo electrode [11].	7
Figure 1.6	CNT Contact resistance with channel length [25].	7
Figure 1.7	CNFET cross section with 9nm channel length and I_{on}/I_{off} for CNFET with channel length 9nm [27].	8
Figure 1.8	Carbon nanotube transistor with ideal gate-all-around geometry. (a) Cross-sectional schematic of the device illustrating how the GAA-CNT channel is suspended across the Si trench and contacted on either side by Pd source/drain. Cross-sectional Transmission Electron Microscope (TEM) images of (b) an array of CNTs with GAA and (c) a higher magnification of a GAA with the CNT visible in the center [26].	9
Figure 1.9	The fabrication flow for CNFET-based computer. The steps 1–4 prepare the final substrate for circuit fabrication, steps 5–8 transfer the CNTs from the quartz wafer (where highly aligned CNTs are grown) to the final SiO ₂ substrate and steps 9–11 continue final device fabrication on the final substrate [57].	9
Figure 1.10	(a) Schematic of SWCNT FET architecture (not to scale). (b) Schematic of PFO-BPy-wrapped SWCNT arrays. (c) Top-down scanning electron micrograph of SWCNTs spanning Pd electrodes of a 240 nm L_{ch} SWCNT FET (scale bar = 100 nm) [10].	10

Figure 1.11	Aligned growth of CNTs. (a) Optical image of aligned CNTs grown on quartz and (b) a higher resolution AFM image. (c) SEM image of CNT-FETs fabricated from aligned CNTs and (d) electrical characteristics of an inverter fabricated with aligned CNTs [23] [52] [64].	12
Figure 2.1	(a) Three-dimensional structure of the devices with multiple channels and high-k gate dielectric material, and the related parasitic gate capacitances. (b) Cross section of the channel region and the related gate-to-channel capacitances.	18
Figure 2.2	The structure of a carbon nanotube with chirality vector $C_h(n, m)$ [59].	19
Figure 2.3	There are identical CNTs in parallel in an array. The coupling capacitance $C_{inf,01}$ is calculated, by considering the effects of the CNTs around middle tube #1, which can be lumped into the two nearest CNTs 2 and 3 on both edges. $C_{inf,02}$ and $C_{inf,03}$ are the equivalent capacitances assuming all the other neighboring CNTs are lumped at the position of 2 and 3.	22
Figure 2.4	Impact of charge screening on gate-to-channel capacitance for middle (C_{gc_m}) and edge (C_{gc_e}) CNTs.	23
Figure 2.5	For CNFET with multiple parallel CNTs, the CNT to CNT screening reduces both the gate to channel electrostatic capacitance (inset) and the drain current [21].	26
Figure 2.6	CNT alignment variation [45].	28
Figure 2.7	Some CNTs become metallic and cannot be switched ON and OFF. Metallic tubes removal and misaligned of CNTs [2].	28
Figure 2.8	S/D doping [70].	29
Figure 2.9	$\sigma(I_{ON})/I_{ON}$ of minimum-width n-type CNFET caused by different variation sources [70].	30
Figure 2.10	VLSI Compatible Metallic CNT Removal (VMR).	30
Figure 2.11	Selective Chemical Etching (SCE).	31
Figure 3.1	Logical Effort for INV, NAND and NOR CMOS gates	35
Figure 3.2	Electrical Effort	35
Figure 3.3	Logical Effort per input and sizing of CNFET-based logic gates	37
Figure 3.4	Gate to Channel capacitance (C_{gc}) for an inverter (a) number of tubes in the transistors (b) pitch.	40
Figure 3.5	Input capacitance of an inverter with given pitch normalized with reference inverter pitch (4nm).	44

Figure 3.6	(a) Reference inverter with number of tubes $N_{ref}=8$, Pitch= $4nm$, Logical Effort ($g_{ref}=1$) (b) 2-inputs NAND gate with $N_{tur}=6$, Pitch= $5nm$ (c) Inverter gate with different CNTs and Pitch, $N_{tur}=4$, Pitch= $8nm$. The configurations in (b) and (c) need to be normalized using reference inverter.	47
Figure 3.7	Technology parameter Tau calculation using different Fanout.	48
Figure 3.8	Correlation of inverter delay from the model with SPICE simulations for various tubes/pitch arrangements; 2/16 and 16/2.	51
Figure 3.9	Correlation of inverter delay from the model with SPICE simulations for various tubes/pitch arrangements; 8/4 and 4/8.	51
Figure 3.10	Maximum and average error in INV gate delay computation from original and developed LE models with SPICE simulation	52
Figure 3.11	Gate level Delay results for FO4	52
Figure 3.12	Average error in delay computation between SPICE simulations and developed model of the test-circuits.	55
Figure 3.13	Circuit level Delay results for FO4.	55
Figure 4.1	(a) The m-CNT (blue) and s-CNT (black) CNFET in CNFET channel. (b) The CNFET channel after removal of m-CNTs. (c) The CNFET channel with s-CNTs with evenly spaced assumption.	63
Figure 4.2	(a) CNFETs randomly placed on aligned s- and m-CNTs. N_{m-CNT} represents the number of m-CNT and N_{s-CNT} expresses the number of s-CNT in a CNFET. CNFET ₂ and CNFET ₃ are functional where CNFET ₁ and CNFET ₄ have short and open defects, respectively. (b) The classification of CNFETs defects based on the number of m- and s-CNTs placed in the active region. (c) Impact of metallic tube removal resulting in non-uniform pitch.	66
Figure 4.3	Influence of position of CNT removal on the gate capacitance, 1 CNT removed per transistor, 2 CNTs removed per transistor.	67
Figure 4.4	Mean capacitance of inverter gate at given P_m (VMR) normalized to $P_m=0\%$ for different N_{tur} .	69
Figure 4.5	Comparison between inverter gate capacitance for position-aware and evenly-spaced tubes. The values represent the mean of 1000 instances. (a) VMR (b) SCE	72
Figure 4.6	Statistical analysis of 1000 instances of two stage inverter using Stanford Spice and CNFET LE model for CNT pitch at (a) 2nm. (b) 6nm. (c) 10nm	74
Figure 4.7	Impact of C_{gtg} on the variance in statistical delay of a CNFET circuit in the presence of metallic tubes.	75
Figure 4.8	Delay distribution of 100 decoder circuits for $P_m=5\%$ and with 5% error amount.	76

Figure 5.1	Flow diagram of optimization framework for CNFET-based circuits.	80
Figure 5.2	Impact of wires on the (a) delay of c17 circuit, (b) optimal gate sizes of c17.	85
Figure 5.3	Optimized gate sizes in 10k circuit with and without CLT.	86
Figure 5.4	The runtime and normalized area and delay product (ADP) after tube optimization for different size of OpenSPARC circuits.	88
Figure 5.5	Delay Distribution of Gates in 10k circuit.	91
Figure 5.6	Delay of 100 critical gates for different benchmark circuits with and without CLT.	92
Figure 6.1	Comparison of existing yield model and proposed model with MC simulations.	102

CHAPTER 1

INTRODUCTION

1.1 BEYOND CMOS

Over the past few decades, Silicon MOSFET circuits have experienced remarkable improvements in terms of both integration density and operating frequency. The CMOS dimensional and functional scaling has enabled broadening spectrum of new applications and technology realm through increased performance and complexity. But now, dimensional scaling of CMOS is experiencing saturation and these devices are approaching physical and technological limits due to variation in process parameters, short channel effects and reliability challenges, resulting in an exponential increase in leakage power, large deviations in the performance of the circuits and other technical limitations. Several new information processing devices and microarchitectures for both existing and new functions are being explored to continue the pace of historical integrated circuit scaling. Hence, it is becoming critical to come up with novel devices to address needs for new paradigms for system architecture, technologies where multiple functions are intended to be integrated, information processing and memory. The technological maturity, long term potential to address these new challenges and other scientific risks related to these emerging devices and novel architecture may delay their acceptance and future development by semiconducting industry. It has been discussed in IRDS,

these challenges have some potential solutions and can be addressed by focusing on two focus areas: 1) Novel architecture development and extending functionality in CMOS, referred as “More Moore”, and 2) Exploration of new devices to address needs for new information paradigm, referred as “Beyond CMOS”. Figure 1.1 demonstrates the relationship between these focus areas. The trend shows that it is hard for saturating More Moore technology to meet demands for new market segments and to address higher performance and efficiency requirements for novel computing paradigm and applications pulls (e.g., big data, IoT, artificial intelligence, autonomous systems, high-speed computing), there is absolute need to focus on Beyond CMOS with emerging material, devices/process and architecture [1].

Figure 1.1: Relationship of More Moore, Beyond CMOS, and Novel Computing Paradigms and Applications [1].

1.2 CHALLENGES

The major bottlenecks to widely acceptance of beyond-CMOS devices are associated with challenges in memory technologies, power and performance requirements of logic

devices, and heterogeneous integration of multi-domain components.

The requirements of the emerging memory devices is to combine the superior features of present memory in manufacturing technology compatible with the CMOS process flow, and is also scalable beyond the limits of current SRAM and FLASH technology. This would laid the foundation for memory devices that can be manufactured for either standalone or embedded applications. The mere scaling of devices doesn't provide efficient solution to enhance the ability of an MPU to run programs, as its limited by the interaction between memory and processor. Increasing the memory of MPU cache is the most feasible solution, but it increases the SRAM area on a MPU chip. However, increase in the area would impact the net delay impacting the overall throughput. The other external storage media with slower operation such as magnetic hard drives, optical CD also need attention. Development of high speed and high density non-volatile memory would create a revolution in the computer architecture.

Second challenge is maintain the scalability of CMOS logic technology below 10nm. As the miniaturization of strained silicon MOSFET channel reaches saturation, there is a need for alternative materials for continuing performance gains. The promising materials to substitute CMOS devices include strained Ge, SiGe, III-V compound semiconductors, and carbon materials. However, the fabrication of non-silicon materials introduces very difficult challenges such as fabrication of defect free channel and source/drain regions, minimizing the tunneling in narrow bandgap channel materials, and compatibility with high-k gate materials. Moreover, continuing the desired reduction in leakage current and power dissipation with these nano-scaled CMOS can be challenging. This can be achieved only with the new materials while minimizing the variations in the critical parameters that can impact the power and performance of these devices.

The exploration of new devices beyond silicon transistors can serve as novel logic

switches as a replacement for silicon devices. However, the potential devices need to meet following requirements:

Figure 1.2: Options for Emerging Memory Devices [1].

1. High device density and reduction in cost, which is not achievable by further scaling of CMOS counterparts;
2. Increased switching speed, either through improvement in drive current or reduction in switched capacitance;
3. Larger reduction in dynamic or leakage power consumption compared to CMOS devices;
4. The methods to process information that cannot be achieved by existing CMOS devices. The Figure 1.3 demonstrates options for potential systematic transition

from CMOS to devices which have very different structure, materials, or operation as compared to CMOS [1].

Figure 1.3: Options for Emerging Logic Devices [1].

1.3 CARBON NANOTUBE FETS (CNFETS)

As one of the promising emerging devices, CNFETs address most of the fundamental limitations for traditional silicon devices. CNFET uses a Single-walled Carbon Nanotube (SWCNT) as channel material. The operation of CNFET like Schottky barrier transistors with nearly transparent barriers to carrier injection, which is demonstrated for both N and P type transport. There is no inversion layers of carriers required to allow current flow, since these are intrinsic semiconductors and do not need to be doped in the traditional way. The carriers are injected from metal contacts due to lowering of energy

barrier in CNFET channel by the gate field. The control electrode (gate) is placed above the conduction channel and which is separated from it by a thin layer of dielectric (gate oxide). Figure 1.4 shows the side and top view of a CNFET where an array of six single-walled CNTs is used as a channel material.

Figure 1.4: (a) Side view of a CNFET layout (b) Top view of CNFET layout with array of six CNTs.

Over the past decade, significant research has been made to understand and enhance device performance in CNTFETs. These key points of the research are discussed as follows [1].

1. The demonstration of an effective contact length of 0 nm with reasonable performance by realizing end-bonded contacts [11] as shown in Figure 1.5.
2. The impact of contact scalability in CNFETs, which is discussed in reference [25] and shown in Figure 1.6.
3. The performance is maintained as the channel length is scaled down to 9 nm with no short channel effects [27], as shown in Figure 1.7.

Figure 1.5: Molybdenum (Mo) end-contacted SWCNT or SWNT transistors. Figures are showing the conversion from a side-bonded contact (a), where the SWCNT is partially covered by Mo, to end-bonded contact (b), where the SWCNT is attached to the bulk Mo electrode through carbide bonds while the C atoms from originally covered portion of the SWCNT uniformly diffuse out into the Mo electrode [11].

Figure 1.6: CNT Contact resistance with channel length [25].

Figure 1.7: CNFET cross section with 9nm channel length and I_{on}/I_{off} for CNFET with channel length 9nm [27].

4. The fabrication of complementary gate-all-around (GAA) FETs [26], shown in Figure 1.8.
5. The realization of an CNFET, showing radio-frequency performance with intrinsic cut-off (f_T) frequency of 153 GHz [58].
6. The fabrication of CMOS inverters and pass-transistor logic using non-doped CNTs with operating voltage of 0.4 V [24].
7. The very first realization of a Carbon nanotube computer composed of 178 FETs [57] as shown in Figure 1.9.
8. The reduction in variability in CNFETs is shared in [12].
9. The origins of hysteresis in CNFETs are studied in [44].
10. The demonstration of CNFETs with ON-current of 0.5 mA/ μ m [10] as shown in Figure 1.10.

Figure 1.8: Carbon nanotube transistor with ideal gate-all-around geometry. (a) Cross-sectional schematic of the device illustrating how the GAA-CNT channel is suspended across the Si trench and contacted on either side by Pd source/drain. Cross-sectional Transmission Electron Microscope (TEM) images of (b) an array of CNTs with GAA and (c) a higher magnification of a GAA with the CNT visible in the center [26].

Figure 1.9: The fabrication flow for CNFET-based computer. The steps 1–4 prepare the final substrate for circuit fabrication, steps 5–8 transfer the CNTs from the quartz wafer (where highly aligned CNTs are grown) to the final SiO₂ substrate and steps 9–11 continue final device fabrication on the final substrate [57].

Figure 1.10: (a) Schematic of SWCNT FET architecture (not to scale). (b) Schematic of PFO-BPy-wrapped SWCNT arrays. (c) Top-down scanning electron micrograph of SWCNTs spanning Pd electrodes of a 240 nm L_{ch} SWCNT FET (scale bar = 100 nm) [10].

There is good progress has been made towards overcoming the challenges related to CNT material and other characteristics [64], which includes the need to achieve purified and sorted semiconducting CNTs growth with a relatively uniform diameter distribution and then position the CNTs into aligned fashion, the arrays of CNTs with consistent pitch and closely packed as shown in Figure 1.11. There is more work needed to achieve a target purity of 99.9999% semiconducting CNTs and placement density of >125 CNTs/ μm (<8 nm pitch) at large scale. Furthermore, there is need for further research toward improving other device-level aspects, including further reduction of contact resistance effects at small contact lengths, including reduction in variability, improved control of gate dielectric interfaces and properties. The experimental study of devices and circuits fabrication using the most scaled and relevant device structures and materials is required to be conducted [1]. In summary, CNFETs have exhibited some of the most substantial potential in high performance, low-voltage, sub-10 nm scaled transistor applications but there is still more work needed to address some of the challenges.

The effective usage of CNFETs in logic gates and circuits requires further research and development in performance models and CAD tools development. CNFET technology needs efficient and accurate industry standard methods for robust circuit design and PPA (Power, Performance and Area) optimization in early design stage, which need to comprehend, 1) CNFET-specific parameters 2) CNT variations, and 3) Fast and accurate shift-left approach to help with timely design convergence.

Figure 1.11: Aligned growth of CNTs. (a) Optical image of aligned CNTs grown on quartz and (b) a higher resolution AFM image. (c) SEM image of CNT-FETs fabricated from aligned CNTs and (d) electrical characteristics of an inverter fabricated with aligned CNTs [23] [52] [64].

1.4 RESEARCH CONTRIBUTIONS

The key contributions of my research are as follows:

1.4.1 Extending Logical Effort model for CNFET-based circuits (Ideal case)

CNFET-specific additional parameters such as number of tubes, pitch (spacing between tubes), tube position and diameter in array of tubes play a significant role in accurate Power Performance and Area (PPA) evaluation. The charge screening effect at smaller pitch values degrades the performance of CNFET-based circuits significantly. In early design phase, when physical design is not available, there is need for models to predict delay of CNFET circuits to meet certain performance requirements. Standard Logical Effort (LE) calculates delay for CMOS-based designs and we extend LE model to do delay

analysis for CNFET-based circuits by considering CNT count (number of tubes), spacing and how it translates into delay calculations. First, the standard CMOS Logical Effort is evaluated to calculate delay in CNFET circuits. Existing standard LE model cannot comprehend charge screening and number of CNTs in CNFET channel and gives an average error in delay up to 55% in comparison with SPICE simulation. Further, the impact of different number of CNTs and spacing between them (with charge screening) on CNFET capacitance is studied for wide range of CNT pitch and count. The Pitch-Aware Logical Effort (PALE) model is developed by incorporating the effect of CNFET-specific parameters, including of the development of reference inverter and Technology parameter for CNFET. Our developed empirical model (PALE) is correlated well with SPICE simulations with an average error of 3% for representative CNFET gates and circuits. Our model is significant faster (25X) as compared to SPICE simulation methods.

1.4.2 Logical Effort based variation aware model (realistic case)

The CNT variations from m-CNT removal and density or spacing between CNTs (CNT count variations) impact performance, power and yield of CNFET-based circuits. We develop closed-form Position-Aware Pitch-Factor (PAPF) model for performance evaluation of large CNFET-based circuits in the presence of imperfection. Monte Carlo simulation methods are used to study the impact of these variations on CNFET capacitance and delay for samples of gates and circuits for different CNT removal techniques. Our developed PAPF model is validated using SPICE simulations methods with average delta around 5% for standard benchmarks circuits. Our PAPF model provides significant runtime benefit as compared to SPICE and Monte Carlo simulation methods without much impact on accuracy.

1.4.3 Area and Delay Optimization of CNFET-based circuits

Furthermore, we develop an optimization tool using PALE and PAPF models based on LSGS algorithm by incorporating CNFET-specific parameters and CNTs count variations, to minimize the area and delay product (ADP) of CNFET circuits. We integrate standard circuit-level techniques (CLT); FO optimization and sizing of gates, used prior to optimizing the CNTs under delay constraint, in the logic gates to achieve highly optimized solution with global approach. For better optimization of the circuits, we also include the impact of wire parasitic using Rent's Rule method in estimating the delay of the individual gates. For standard ISCAS and OpenSPARC benchmark circuits, we have investigated that our optimization tool results in maximum and average delay improvement by 27% and 17% respectively, and 2.5X reduction in area. Fast and fairly accurate delay computation in our optimization framework provides great runtime benefits as compared to state of-the-art SPICE simulation and statistical-based methods.

1.4.4 CNT Position-Aware Yield Estimation model

The existing yield models have limitations and do not include the screening effect and actual position of the CNTs in the channel during yield estimation. Therefore, the predicted functional yield may not be accurate. In this work, we propose more accurate probabilistic yield model based on conditional probability, which uses the enhanced delay models to account for the screening effect and position of CNTs in the presence of CNT density and count variations at smaller CNT pitch. To estimate yield for CNT pitch greater than 6nm, probabilistic yield model proposed by [7] is used. Our proposed model is correlated within 1% with Monte Carlo simulations at small pitch values for CNFET gates and circuits and offers significant runtime benefit.

1.5 OUTLINE

The rest of the thesis is organized as follow:

Chapter 2 introduces CNFET specific parameters, CNFET gate capacitance model, charge screening effect, advantages and challenges faced by CNFET and techniques to remove metallic tubes.

Chapter 3 describes the Logical Effort (LE) background, its benefits and limitations in CMOS LE models. Then it focuses on the development of LE models for CNFET technology with charge screening effect considering all semiconducting tubes.

Chapter 4 provides the details on the development of Logical Effort based framework to evaluate the performance of CNFET-based circuits in the presence of metallic tubes and density variations.

Chapter 5 discusses the development of early design optimization framework to minimize area and delay product (ADP) for CNFET-based circuits, by optimizing the number of tubes and performing circuit-level techniques (CLT).

Chapter 6 focuses on the analysis of the yield of CNET-based circuits and propose methodology for accurate yield estimation with screening effect by considering the actual position of tubes into account.

Chapter 7 concludes the thesis with summary of work and also suggesting future work.

CHAPTER 2

CNFET TECHNOLOGY, ADVANTAGES AND DIS-ADVANTAGES

In this chapter, we discuss technology specific parameters, CNFET capacitances, Advantages and limitations about CNFETs.

2.1 ADVANTAGES OF CNFETS

Silicon MOSFET circuits have experienced tremendous improvements in terms of both performance and integration density over the past few decades. However, now, conventional silicon-CMOS devices are approaching their physical and technological limits due to variation in process parameters and other short channel effects. Also, transistor scaling faces significant challenges especially with continuously increasing energy/power dissipation [39] [67]. Carbon-based materials such as Carbon Nanotubes (CNTs) have drawn considerable attention due to superior electrical, thermal, and mechanical properties [21]. Hence, Carbon Nanotube Field-Effect Transistor (CNFET) has been considered as one of new promising devices for post silicon era. CNFET is characterized by ultra-long mean-free-path (MFP) for elastic scattering similar for electrons and holes, high Fermi velocity, easy integration of high-k dielectric material, and other excellent device characteristics [53]. Stanford engineers built first basic computer chip successfully and demonstrated advantages of CNFET technology [57]. It is observed by Deng *et al.* [20] [66] that electrical properties of CNFET are not exactly same as of CMOS due to

unique CNFET device parameters and characteristics (numbers and diameter of CNTs, position and spacing between two CNTs) of CNFET on top of conventional CMOS specific parameters like node voltage, threshold voltage and gate width which effects I-V characteristics of a transistor. The most prominent advantages of CNFETs over other options for aggressively scaled devices are the room temperature ballistic transport of charge carriers, the reasonable energy gap, the demonstrated potential to yield high performance at low operating voltage, and scalability to sub-10nm dimensions with minimal short channel effects [1].

2.2 CNFET SPECIFIC AND TECHNOLOGY PARAMETERS

The delay, power and area of CNFET-based circuits are determined by CNFETs specific parameters which are explained below. Here, we have included the description of technology parameters for 32nm, which are used in our experiments.

2.2.1 CNTs array

CNTs are hollow cylindrical nano structures made up of carbon atoms. CNTs with a single shell of carbon atoms are called Single-Walled carbon nanotubes (SWCNTs) or simply CNTs. A structure of a CNFET is mostly similar to MOS transistors with a major difference being a channel that in CNFET is built with an array of parallel semiconducting carbon nanotubes (CNTs), referred as N_{tur} in this thesis. The region of the carbon nanotube used as a channel is undoped and the regions used as source and drain are heavily doped.

2.2.2 Pitch

The distance between the centers of two CNTs is referred as CNT Pitch as shown in Figure 2.1(b). When spacing of CNTs on channel is too narrow then the gate capacitance is not proportional to number of the carbon nanotubes and it impacts the gate capacitance. This effect is called screening effect and details are discussed in Section 2.4.

Figure 2.1: (a) Three-dimensional structure of the devices with multiple channels and high-k gate dielectric material, and the related parasitic gate capacitances. (b) Cross section of the channel region and the related gate-to-channel capacitances.

2.2.3 Diameter

The diameter d_{CN} of a CNT is specified by the chirality vector $C_h(n, m)$ and is shown in Figure 2.2. Both n and m are positive integers. CNT is metallic if $|n - m|$ is an integer multiple of 3. Alternatively, if $|n - m|$ is not an integer multiple of 3, SWCNT is a semiconductor [59]. We have considered diameter as 1.5nm, which is optimum diameter value corresponding to chirality vector (19, 0). Sayed *et al.* [54] shared simulation results to show that 1.5nm is optimum value to achieve minimum PDP. Imran *et al.* [32] discussed using optimum diameter value as 1.5nm, to achieve better trade-off between bandwidth and port impedance.

Figure 2.2: The structure of a carbon nanotube with chirality vector $C_h(n, m)$ [59].

2.2.4 Width of CNFET gate

In CNFET, the total area of a channel is determined by the physical gate width (W_g) as shown in Figure 2.1(b). W_g is determined by the inter-tube pitch, the number of CNTs

(N_{tur}), the diameter of CNTs (d_{CN}), and the gate extension beyond the carbon nanotubes at the two ends of the channel [59]. Due to multiple parallel CNTs in the channel, the physical channel width can be expressed as product of N_{tur} and $Pitch$, which is considered in our work.

2.2.5 The technology parameters

The description of technology parameters for 32nm which are used in our experiments is shown in the Table 2.1.

Parameters	Value/Range
Supply Voltage	0.9v
Threshold Voltage	0.2v
Gate length	32nm
Gate height	64nm
Pitch	[2nm-32nm]
Diameter	1.5nm
Percentage of metallic tubes	[0%-25%]
R	8 $\Omega/\mu\text{m}$
C	0.2 fF/ μm

Table 2.1: Description of technology parameters

2.3 CNFET GATE CAPACITANCE MODEL

The CNFET 3-D structure is shown in Figure 2.1(a). The capacitance models for CNFET are taken from Deng *et al.* [21]. The capacitances associated with CNFET are shown in Figure 2.1. The different components of C_{input_gate} are shown in (2.1) and discussed below.

$$C_{input_gate} = C_{gtg} + C_{gc} + C_{of} \quad (2.1)$$

where C_{gtg} is the gate-to-gate, or gate-to-source/drain coupling capacitance, C_{gc} is

the gate-to-channel capacitance, and C_{of} , outer-fringe capacitance.

The C_{gtg} , is major component of gate capacitance and is calculated using the formula in [21]. It is separated into two components, the first component of C_{gtg} is due to normal electrical field between two parallel plates and the second component represents fringe capacitance between two cylinders. As shown in Figure 2.1(a), this capacitance is between two metal gates and any impact due to CNTs is very negligible on C_{gtg} . C_{gc} is also referred as intrinsic gate capacitance and different components are shown in Figure 2.1(b). There are two types of C_{gc} ; the C_{gc_e} is the capacitance of the CNT located in the edge of the CNFET device and the C_{gc_m} is the capacitance of the CNT located in the middle of the CNFET device. The gate capacitance of multi-channel CNFETs is calculated by considering the coupling capacitance between the gate and one isolated CNT (C_{gc_inf}) and the equivalent capacitance (C_{gc_sr}) due to charge screening from the adjacent tubes as given in [21]. The charge screening is discussed in detail in Section 2.4.

For 32nm source/drain length and 64nm gate height, the contribution of C_{of} , is almost negligible as compared to C_{gtg} and C_{gc} , and which can be ignored [21] [20].

2.4 CHARGE SCREENING EFFECT

In an array of CNTs that creates a channel for a CNFET, electric field lines can be screened due to the proximity of tubes, and that would effect the distribution of charges among the nanotubes [59]. This effect is called *screening effect*. It is explained by Deng *et al.* in [21] with the help of (2.2).

$$C_{inf,01} = \frac{1}{\frac{1}{C_{inf}} + \eta_1 \cdot \frac{1}{C_{sr,1}} + \eta_2 \cdot \frac{1}{C_{sr,2}}} \quad (2.2)$$

Figure 2.3: There are identical CNTs in parallel in an array. The coupling capacitance $C_{inf,01}$ is calculated, by considering the effects of the CNTs around middle tube #1, which can be lumped into the two nearest CNTs 2 and 3 on both edges. $C_{inf,02}$ and $C_{inf,03}$ are the equivalent capacitances assuming all the other neighboring CNTs are lumped at the position of 2 and 3.

where η_1 and η_2 are defined as ratio of $C_{inf,02}$ and $C_{inf,03}$, respectively, over $C_{inf,01}$, these are functions of the geometry, the number of the objects of the array, and the position of the object in the array. C_{inf} is the capacitance between electrode or gate and middle tube #1 without considering charge screening of all neighboring tubes. $C_{sr,1}$ and $C_{sr,2}$ are the equivalent capacitances due to the screening effects of edge tubes #2 and #3, respectively. The Figure 2.3(a) shows N identical tubes in parallel in an array. Figure 2.3(b) shows a simple representation to explain screening effect and corresponding coupling capacitances. The charge distribution due to screening effect influences the intrinsic capacitance of a CNFET. At smaller pitch, the screening effect becomes significant and the gate capacitance is not anymore directly proportional to the number of the CNTs in the channel. Essentially, the charge screening reduces the effective width of the channel, affects gate to channel electrostatic capacitance, thereby degrading the

device current. It means the spacing between the adjacent parallel tubes in a CNFET channel array impacts the drive strength of parallel tube transistors due to the screening of charge from the adjacent tubes [59]. It is known that tubes present on the edges of an array have screening from only one side, and tubes in the middle get charge screening from adjacent CNTs on both sides as shown in Figure 2.3(b). It is observed that the capacitance of edge tubes is twice that of the middle tubes at smaller inter-CNT pitch as shown in Fig 2.4. This indicates that screening effect is dominant when tubes are separated by small distance which degrades both gate capacitance and current in CNFETs. To avoid overestimating the circuit performance for high-speed CNFET circuits, the screening effect needs to be taken into account [21]. Therefore, we have accounted for screening effect in our proposed models and discussed in detail in section 2.4.

Figure 2.4: Impact of charge screening on gate-to-channel capacitance for middle (C_{gc_m}) and edge (C_{gc_e}) CNTs.

2.5 THE VARIATIONS AND CHALLENGES IN CNT TECHNOLOGY

It has been discussed that CNFETs are less sensitive to conventional CMOS variations such as channel length, oxide thickness and threshold voltage variations [49]. We will discuss CNT-specific variations and consider some of them in our research.

2.5.1 CNT type (m-CNT or s-CNT) variations

SWCNTs can be metallic or semiconducting depending on chirality (different ways carbon atoms are arranged in CNTs [53]) as shown in Figure 2.2. It is difficult to control chirality of CNT during growth, so there is no CNT synthesis technique that will guarantee to grow 100% s-CNTs (semiconducting tubes). Today's CNT growth techniques typically produce 4% to 50% m-CNTs [22] [41] [60] [57]. CNT type variations can lead to CNFET circuit variations because: a) the electrical properties of m-CNTs and s-CNTs are very different; b) when m-CNTs removal techniques are applied, the number of surviving CNTs can vary significantly. The m-CNT causes short between the source and drain of the CNFET, which cannot be controlled by gate terminal anymore. Therefore, delay, static current, functional yield and power of complementary logic gates and circuits with pull-up and pull-down networks are greatly impacted by these shorts caused by presence of m-CNTs. There are techniques proposed [17] [48] [69], to remove m-CNTs but unfortunately it causes the density variations in the remaining majority of s-CNTs.

2.5.2 CNT diameter variations

The variations in CNT chirality cause CNT diameter variations, because the diameter of a CNT is a function of its chirality [53]. The fabrication of CNTs results in the varia-

tion in the diameter of the tubes where normally fabricated CNTs have diameters within 1nm to 2nm. The diameter variations of CNTs have a Gaussian distribution as shown in experimental results by [42]. The CNT diameter determines the bandgap of a CNT. Thus, CNT diameter variations can cause threshold voltage variations. According to Zhang *et al.* [72], Hills *et al.* [31] and Raychowdhury *et al.* [50], while the on-current (I_{ON}) of a CNFET with only a single CNT as its channel is highly sensitive to CNT diameter variations. CNFETs in practical VLSI circuits consist of multiple CNTs to provide sufficient drain-to-source drive current in a CNFET (I_{ON}), which is one of the main assumptions in our work. Thus, the impact of diameter variations is reduced due to statistical averaging as compared to CNT count and density variations. It is also discussed that variations in I_{ON} and I_{OFF} due to diameter are within the limits of tolerance even in the state-of-the-art Si process and CNFET structure can tolerate significant variations in the diameter [50] as shown in Figure 2.9. Hence, we focus on the CNT count and density variations and have not considered the impact of diameter variation in this work.

2.5.3 CNT density and Spacing variations

An array of densely packed CNTs in parallel is considered to drive large current requirements which cannot be met with single or fewer CNTs. Chemical synthesis techniques which are used to grow CNTs, do not provide precise control of the individual tube locations in the array CNTs. The density variations of grown CNTs represent the variations of inter-CNT spacing during growth and the resulting variations in the CNT count inside CNFETs. This variation has been explored by [36] and [48]. The drive current of a parallel tube CNFET depends upon the gate to channel capacitance as shown in figure 9. 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 [21]. 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 and it has been discussed in detail in chapter 3. 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 change in the charge screening because of change in the spacing and in the density of CNTs, influence the drive current of CNFETs greatly. The screening effect becomes prominent for pitch values less than 10nm as it can be seen in Figure 2.5.

Figure 2.5: For CNFET with multiple parallel CNTs, the CNT to CNT screening reduces both the gate to channel electrostatic capacitance (inset) and the drain current [21].

2.5.4 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 (i.e., having a non-zero angle with respect to the alignment direction) as shown in Figures 2.6 and 2.7. Significant progress has been made in the fabrication of aligned CNTs [38] [47]], and less than 0.5% of CNTs fabricated on the single-crystal quartz substrate are misaligned [36]. The misaligned tubes can cause either a short between the output and the supply rail, or an incorrect logic function [45] as shown in Figures 2.6 and 2.7. Such functional failures can be dealt with using special layout techniques, as previously reported in [45]. The most of the CNTs are aligned to a single direction in the CNT arrays produced by CNT growth techniques [38] [47]. However, a small fraction of the CNTs can be misaligned (i.e., having a non-zero angle with respect to the alignment direction). Such alignment variation causes changes in actual CNT length in the CNFET channel and also introduces CNT-to-CNT junctions. In the extreme cases, poor alignment of CNTs can cause functional failures of logic gates [45]. Such functional failures can be dealt with using special layout techniques, as previously reported in [9] [45].

2.5.5 CNT doping variations

These are variation of doping concentration in the source / drain extension regions of a CNFET (i.e., the regions of exposed CNTs between the gate and source/drain contacts in a CNFET).

As shown in Figur 2.9, m-CNTs and CNT density variations are dominant contributors to the on-current variations. This is because they directly contribute to the number of conducting CNTs (channels) in a given CNFET. Variations in CNT diameters can result

Figure 2.6: CNT alignment variation [45].

Figure 2.7: Some CNTs become metallic and cannot be switched ON and OFF. Metallic tubes removal and misaligned of CNTs [2].

Figure 2.8: S/D doping [70].

in a significant change in the on-current of a single CNT. However, statistical averaging effects resulting from the multi-CNT CNFET structure substantially suppress the overall variations caused by CNT diameter variations. Due to these variations specific to carbon nanotubes (CNTs), there are challenges to build energy-efficient and robust CNFET digital VLSI. As shown in Figure 2.9, one of the major contributors of variations is the unwanted growth of metallic tubes. The current known CNT fabrication technologies do not result in 100% semiconducting CNTs in the channel. The presence of metallic tubes creates a short between the drain and source terminals of the transistor and impacts performance and functional yield of CNT based gates. CNT density variation can potentially cause complete failure of CNFETs in cases when there is no s-CNT between drain and source. Therefore, modeling, understanding and mitigating the impact of these variations are critical. To address fabrication issues related to unwanted m-CNTs in CNFETs, two techniques were proposed, 1) Removal of unwanted tubes by Selective Chemical Etching (SCE) [69], and 2) VLSI Compatible Metallic Carbon Nanotube Removal (VMR) [48]. Both tube removal techniques (SCE) and (VMR), will remove close to 100%, of metallic tubes, but unfortunately, will also remove some of needed semiconducting tubes. In this work, we consider removal of unwanted tubes

Figure 2.9: $\sigma(I_{ON})/I_{ON}$ of minimum-width n-type CNFET caused by different variation sources [70].

using both SCE and VMR. The large performance variations are observed after the removal of tubes in the CNFET-based circuits and sometimes it can create open-circuit gates (worst-case) or short-circuit, when the removal process cannot remove all metallic tubes.

Figure 2.10: VLSI Compatible Metallic CNT Removal (VMR).

2.6 SUMMARY AND CONCLUSION

Carbon nanotube field-effect transistors (CNFETs) show great potential to build digital systems on advanced technology nodes with big benefits in terms of power, performance

Figure 2.11: Selective Chemical Etching (SCE).

and area (PPA). However, CNFETs specific additional features such as number of tubes, pitch (spacing between tubes), tube position and diameter in array of tubes play a significant role in accurate PPA evaluation. Furthermore, count and density variations in Carbon nanotubes (CNTs) due to manufacturing limitations, like presence of metallic tubes in the CNFET channel, degrades the anticipated PPA benefits. Moreover, modeling the CNFET parameters, CNT variations and etching techniques for CNTs create additional complexity during performance optimization. Hence, for realistic optimization of CNFET circuit's performance, it is imperative to incorporate the impact of these parameters and variations. As part of our work, we have presented models, methods and techniques, which can address some of the challenges faced by CNFET circuits with reasonable accuracy.

CHAPTER 3

PITCH-AWARE LOGICAL EFFORT MODEL

Part of this chapter has been published in:

- M. Ali, M. A. Ahmed and M. Chrzanowska-Jeske, "Logical Effort Framework for CNFET-Based VLSI Circuits for Delay and Area Optimization," in IEEE Transactions on Very Large Scale Integration (VLSI) Systems, March 2019, pp. 573-586.
- M. Ali, M. Ahmed, and M. Chrzanowska-Jeske, "Fast and Accurate Evaluation of delay in CNFET Circuits," in 2016 IEEE 16th International Conference on Nanotechnology (IEEE-NANO), Aug 2016, pp. 659–662.
- M. Ali, M. Ahmed, and M. Chrzanowska-Jeske, "Stochastic Analysis of CNFET circuits using Enhanced Logical Effort Model in the Presence of Metallic Tubes," in 2014 21st IEEE International Conference on Electronics, Circuits and Systems (ICECS), Dec 2014, pp. 774–777.
- M. Ali, M. Ahmed, and M. Chrzanowska-Jeske, "Logical Effort model for CNFET-based circuits," in 14th IEEE International Conference on Nanotechnology, Aug 2014, pp. 460–465.

The Logical Effort (LE) technique was first proposed by Sutherland *et al.* [61] [62] and it is used in several industry standard computer-aided-design (CAD) tools and other

applications because of its elegant and simple nature. Due to aggressive time to market (*TTM*) requirements for most modern digital CMOS designs, industry is focusing more and more on shift left techniques, which means there is a great demand to estimate delays of logic gates and data-paths, even during the primitive design phase, and to size the logic gates in order to meet the timing requirements for better convergence and timely sign-off. The sizing of logic gates in order to meet a delay constraints, with minimal power dissipation has been a key requirement in most digital designs. Hence, there has been great interest in LE model and significant research has been done [34] [35] [43] [51] to improve LE model for MOSFET applications. Kim *et al.* [37] discussed very basic design methodology using CMOS-based LE model in CNFETs without considering CNFET specific parameters and CNT variations.

The focus of most of the models and methods developed for performance evaluation of CNFET-based circuits, is SPICE like transient and MC-based simulations, which are very computation intensive for large circuits and result in very long simulation runtime. Hence, there is a great need to develop robust models to quickly estimate and optimize delay, without performing long SPICE simulations and detailed delay calculations. There have been several methods developed in CMOS technology for quick and early estimations of performance and power. There has been research conducted to improve Logical Effort model in CMOS technology. In this chapter, we discuss the background of industry standard Logical Effort (LE) model, its advantages, and limitations in CNFET technology. The design methodology is discussed using Logical Effort in CNFET without considering CNT specific parameters in [37] and CNFET specific parameters are not modeled in LE framework which impact the performance significantly. These parameters are discussed in section 2.2. We discuss development of proposed LE models for CNFET technology for ideal cases (considering screening effect with no CNT variations) and realistic cases

(with CNT density and metallic tube variations). We also present correlation of developed models using SPICE simulation methods.

3.1 THE STANDARD LOGICAL EFFORT MODEL

The Logical Effort (LE) technique is widely recognized as a pedagogical method to quickly estimate and optimize, without performing long SPICE simulations and detailed delay calculations, single paths by modelling equivalently propagation delay and transition time in CMOS circuits. In order to meet a delay constraint by sizing gates with minimal power dissipation has been a key requirement in most digital designs. Logical Effort model provides designers an elegant and intuitive solution in estimating gate delay in early design phase. First, we discuss standard LE model for CMOS. The propagation delay of a gate is represented as:

$$t_p = (g.h + p).\tau_0 \quad (3.1)$$

$$t_p = d.\tau_0 \quad (3.2)$$

where τ_0 is the intrinsic delay of a reference inverter without parasitic. t_p is propagation delay (also referred as absolute delay) and d is normalized delay of a gate and is given as:

$$d_{norm} = (g.h + p) \quad (3.3)$$

where g is the Logical Effort of the gate, h is the electrical effort, C_{out}/C_{in} , and p is the parasitic (or self-loading) delay. The parameter g represents the input capacitance required for a gate to have the equivalent drive strength as an inverter. The electrical effort is shown in the Figure 3.2. Logical Effort technique in general calculates a gate delay in

terms of the basic inverter delay and input loading of the gate represented as a multiple of the minimum-size transistor's gate capacitance. We use capacitance-based *LE* model, which is defined as *the ratio of its input capacitance to that of an inverter that delivers equal output current*.

Figure 3.1: Logical Effort for INV, NAND and NOR CMOS gates

Figure 3.2: Electrical Effort

Following is summary of steps to develop the PALE model:

- (i) To evaluate well-known CMOS-based LE model for CNFET-based circuit applications.
- (ii) Developed Pitch-Aware LE (PALE) model to include the impact of the screening effect.
- (iii) Correlate our new model (PALE) with SPICE simulation-based approaches for different technology using gate delay as criteria.

3.1.1 Evaluation of Logical Effort model for CNFET-based circuits

In this section, we evaluate the LE model for CNFET technology. The input gate and parasitic capacitances for CNFET gates are calculated by using (2.1) and shown in Table 3.1 for 32nm technology node. In CNFET, the p-type and n-type carbon nanotubes have almost the same carrier mobility, so the equal numbers of tubes are considered in PU (pull-up) and PD (pull-down) networks. For the sake of simplicity, we ignore the screening effect that is due to inter-CNT electrostatic discharge, which is being discussed in detail in chapter 2 and taken into account in our proposed models. The LE for CNFET-based combinational gates (Inverter, n-input NAND and NOR) are calculated using (3.4) as shown in Table 3.1, which shows the LE for 2-input NAND and NOR are the ratio of C_{input_gate} of these gates to the C_{input_gate} of inverter used as reference [61] [62]. C_{input_gate} is calculated using (2.1).

$$g = \frac{C_{in_gate}}{C_{in_inv}} \quad (3.4)$$

Figure 3.3 shows how CNFET-based gates are sized and how LE is calculated for these gates using inverter as reference. These numbers match LE values which are calculated using specific capacitance values shown in Table 3.1. The LE for these gates is compared with the LE of Si-CMOS logic gates at 32nm technology node as presented in Table 3.2.

Figure 3.3: Logical Effort per input and sizing of CNFET-based logic gates

Table 3.1: Logical Effort, using actual input gate capacitance, of CNFET-based gates with $N_{tur}=8$ in each transistor.

Gate Type	$C_{gtg}(aF)$	$C_{gc}(aF)$	$C_{input_gate}(aF)$	Logical Effort
Inverter	14.8	64.0	78.8	1.0
2-input NAND	22.2	96.0	118.2	1.5
2-input NOR	22.2	96.0	118.2	1.5

For CNFET large circuits, path delay is given by: $D = (N * F)^{\frac{1}{N}} + P$, with path effort $F = G * B * H$, where G is LE of all gates in the path, H is equal to $\frac{C_{out_load}}{C_{in_driver}}$, and C_{out} and C_{in} are directly influenced by the number of tubes in the driver and load gates and B represents branching effort.

Table 3.2: Logical Effort of CNFET versus CMOS gates

No. of inputs	Logical Efforts			
	NAND (CNFET)	NAND (CMOS)	NOR (CNFET)	NOR (CMOS)
2	3/2	4/3	3/2	5/3
3	4/2	5/3	4/2	7/3
4	5/2	6/3	5/2	9/3
n	$\frac{(N+1)}{2}$	$\frac{(N+2)}{3}$	$\frac{(N+1)}{2}$	$\frac{(2N+1)}{3}$

3.1.2 Comparison of CMOS and CNFET Logical Efforts

Following observations are made based on Table 3.2 [5].

- (i) The LE for both CNFET-based NAND and NOR gates are the same, because mobility of electrons and holes are almost the same and therefore the number of tubes in the PU and PD networks of the reference inverter are the same.
- (ii) The LE of CNFET-based NAND gate is larger than CMOS-based NAND. CNFET-based NAND has to exert more effort to deliver the same amount of current as inverter as compared to CMOS-based NAND because PU network in CNFET-based inverter is stronger.

- (iii) The LE of CNFET-based NOR gate is smaller than CMOS-based NOR, since PD network of the reference inverter is the same in both CNFET and CMOS.
- (iv) The gate capacitance for CNFETs is greatly influenced by the number of tubes (CNT density) and technology (Pitch). Two CNFET gates with similar width and topology but different number of tubes and pitch have different driving capabilities. The LE model for CMOS cannot comprehend these parameters and needs to be modified.

The Logical Effort technique shows high potential for being very efficient in evaluation and optimization of stochastic CNFET-based designs. The models which we are planning to develop based on Logical Effort framework are inherently path-based. According to [62]: *Synthesis tools make some effort to explore topologies, but still cannot match experienced designers on critical paths.*

3.2 LIMITATIONS OF STANDARD LE MODEL FOR CNFET

As discussed, Logical Effort represents the driving capability of a gate and depends on the gate topology for CMOS devices. It models the amount of current generated by a gate when a certain charge is provided at the input with constant width. The current generated by a CMOS inverter is constant for a certain input voltage applied at the gate of inverter. However, for CNFET gates, driving strength of a CNFET depends on the number of tubes in the channel and spacing between them, which we referred as pitch. Therefore, a CNFET gate with constant width outputs different current for the same input voltage. The LE framework models the driving strength of the gates based on the various capacitances associated with the gate as shown in (2.1). Therefore, we use the CNFET capacitance model to examine the behavior of CNFET input capacitance in the

presence of charge screening. We observe the impact of the number of CNTs and pitch on the input gate capacitance of CNFET. In Figure 3.4(a), we show the behavior of gate capacitance with the number of tubes in the gate. It can be seen that gate capacitance of CNFET gates varies linearly with the number of tubes. However, Figure 3.4(b) shows that at smaller pitch values, e.g., 6nm and below, the impact of charge screening on C_{gc} is more visible. It varies with CNTs spacing and has non-linear impact on the gate input capacitance. The slopes of C_{gc} suggest that the driving capability of CNFET varies differently with number of CNTs for various pitch. Therefore, it is critical to model the impact of screening effect on the input gate capacitance for accurate delay evaluation of CNFET-based circuits. We also observe that the CNFET gate with a given width may have different possible configurations depending on the number of tubes and pitch. Hence, the long SPICE simulations to evaluate the circuits with different configuration may be a cumbersome process. In such scenarios, our developed PALE model can be used for quick evaluation of such CNFET circuits.

Figure 3.4: Gate to Channel capacitance (C_{gc}) for an inverter (a) number of tubes in the transistors (b) pitch.

3.3 PITCH-AWARE LOGICAL EFFORT (PALE) MODEL

We extend LE framework first time for CNFET era, which involves evaluation of LE model for CNFET-based gates and circuits, and expanding Logical Effort model for CNT specific parameters i.e. screening effect due to narrow pitch between tubes and its impact on CNFET capacitances due to different number of tubes in channel for fixed width.

3.3.1 Derivation of Pitch-Factor (PF) to model the Screening Effect

In order to develop the LE model for CNFET gates, we use 32nm technology node for our model development. As discussed in [61] LE model evaluates the delay of logic gates normalized to reference inverter. The details about reference inverter are discussed in next section. The reference inverter in CNFET has minimum gate width of 32nm. As we discussed in section 3.2, there can be multiple combinations of number of CNTs and pitch for the same width of inverter. We decide to choose our reference inverter for CNFETs based on the desired CNT density ($250/\mu\text{m}$) to match the performance of CMOS circuits [46]. The desired density requires at least 8 CNTs for minimum width (32nm) inverter. Hence, our reference inverter has 8 tubes separated by 4nm spacing. The delay of inverters and other logic gates with different number of tubes and pitch combinations is normalized with respect to the reference inverter.

To evaluate the CNFET speed, it is shown by Deng *et al.* [21] that the drive current is proportional to gate-to-channel capacitance per unit channel length C_{gc} , and local interconnect series resistance is usually much smaller than the CNFET overall intrinsic resistance; thus CNFET delay is proportional to input gate capacitance and represented as:

$$t_p \propto \frac{C_{gc} \cdot L_g + 3(C_{of} + C_{gtg} \cdot W_{pitch})}{C_{gc}} \quad (3.5)$$

Where, L_g is the physical gate length, and W_{pitch} is the device pitch in the width direction, both of these terms are constant.

In (3.1), t_p is represented in the form of LE model expression and is proportional to normalized delay of the gate and delay of reference inverter without parasitic (C_{of} and C_{gtg} of inverter are ignored). Therefore, to model the screening effect for LE framework, C_{gc} of inverter with certain CNT spacing is normalized with respect to reference inverter CNT pitch (4nm). The outputs from multiple SPICE simulations are summarized in Figure 3.5 for different N_{tur} and pitch values. It can be seen that impact of screening on normalized gate capacitance is much larger at smaller CNT spacing. Also, the normalized gate capacitance varies from 0.5 to 2.2 times of reference inverter. We choose the value of normalized gate capacitance, referred as pitch factor (PF) for each pitch value where the curve for inverter gate capacitance reaches saturation. In first order approximation, PF captures the non-uniform behavior of normalized C_{gc} in presence of charge screening for wide range of pitch values including due to both edge and middle CNTs in the CNFET channel, which is needed in simple expression form for LE framework. Table 3.3, which is referred as look-up-table (LUT), shows the values of PF for wide range of CNT pitch. Note that the reference pitch of 4nm has the pitch factor (PF) of 1.

The PALE model is derived based on experiments and used in the form of normalized expression. The LE of reference inverter, normalized to reference CNT count and pitch (N_{ref}, P_{ref}) is represented as “ g ”. Whereas, the Logical Effort of another inverter with the same gate width, but different CNT count and pitch (N_{tur}, P_{tur}) is represented as “ g' ” and can be computed as follows.

$$g' = g * (\text{normalized screening effect}) * (\text{normalized CNT count})$$

Pitch (nm)	2	4	6	8	10	12	14	16	18	20
PF	0.5	1.0	1.3	1.4	1.5	1.57	1.62	1.67	1.7	1.73

Table 3.3: Look-up-table (LUT): Pitch factor (PF) values for different spacing between CNTs

Normalized screening effect is the screening effect in another inverter or gate normalized to reference inverter. Since the screening effect has inverse impact on the capacitance and driving capability of the gate, the normalized screening effect is given by expression below and it is modeled as $1/PF$ in (3.6).

$$\text{Normalized Screening Effect} = 1/PF$$

Similarly, the inverse linear effect of CNT count on Logical Effort as shown in Figure 3.4(a) can be captured as the ratio of tubes in a certain gate to the reference inverter ($\frac{N_{tur}}{N_{ref}}$) and shown in the expression below:

$$\text{Normalized CNT count} = \frac{1}{N_{tur}/N_{ref}}$$

Thus, this pitch factor (PF) and ratio of tubes are included in the LE framework to model the driving strength of the inverter at different CNT spacing and count.

3.3.2 Derivation of Pitch-Aware Logical Effort (PALE) and Estimation of Normalized Delay

The new developed LE model is referred as Pitch-Aware Logical Effort (PALE) and shown in (3.6). The screening effect is modeled by Pitch-Factor (PF). N_{ref} and N_{tur} represent the number of tubes in the reference and a given gate, respectively. g_{GATE} of any logic gate is

Figure 3.5: Input capacitance of an inverter with given pitch normalized with reference inverter pitch (4nm).

equal to its original LE (g_{REF_GATE}) when gate has same number of tubes and pitch values as reference INV as shown in (3.7). The larger number of tubes and pitch lead to lower effort required by the gate to drive the load. The electrical effort “h” is defined as ratio of output to input capacitance as shown in (3.8). The parasitic effort “p” depends on the number of inputs as shown in (3.9). The normalized delay for CNFET-based circuits in the presence of screening effect for single stage is given by (3.10). The normalized delay (D_{norm}) for CNFET circuits with multiple stages and branches is given by (3.11) and steps to derive normalized and absolute delay based on PALE model g' are shown below.

$$g' = g \cdot \left(\frac{1}{PF}\right) \cdot \left(\frac{N_{ref}}{N_{tur}}\right) \quad (3.6)$$

$$g_{GATE} = g' * g_{REF_GATE} \quad (3.7)$$

$$h = \frac{C_{out}}{C_{input_gate}} \quad (3.8)$$

$$p = 1, 2, 3, \dots, n \quad (3.9)$$

The normalized delay in the presence of screening effect is presented as:

$$d' = g' \cdot h + p \quad (3.10)$$

$$D_{norm} = N * F^{1/N} + P \quad (3.11)$$

$F = G * B * H$, where F is path effort, B represents number of branches and N is the number of stages.

$$G = g_1 * g_2 \dots g_n, H = C_{out}/C_{in}$$

Equation (3.2) for CNFET-based circuits is represented as below:

$$t_{p,CNFET} = d' \cdot \tau_0 = (g \cdot (\frac{1}{PF}) \cdot (\frac{N_{ref}}{N_{tur}}) \cdot h + p) \cdot \tau_0 \quad (3.12)$$

It is observed that larger N_{tur} and pitch values lead to lower effort required by the gate to drive the load. As the CNT spacing increases, the impact of screening effect reduces and the value of pitch factor increases. The increase in PF reduces the g' showing the better driving strength of the inverter, hence improved delay, i.e., $t_p \propto \frac{1}{PF}$.

To represent PF in overall LE framework in the form of expression, it is shown in (3.13) in simplified form instead of LUT.

$$PF = \frac{(2 \cdot P_{abs} - 2)}{(P_{abs} + 2)} \quad (3.13)$$

Where, P_{abs} represents absolute value of CNT pitch and normalized to (CNT pitch(nm))

/1nm). Hence, (3.12) is updated as follow to calculate the propagation delay of any logic gate.

$$t_{p,CNFET} = (g \cdot (\frac{(P_{abs} + 2)}{(2 \cdot P_{abs} - 2)}) \cdot (\frac{N_{ref}}{N_{tur}}) \cdot h + p) \cdot \tau_0 \quad (3.14)$$

3.3.3 A Reference Inverter

It is discussed earlier that LE model evaluates the delay of logic gates normalized to reference inverter. The reference inverter in CMOS has minimum width equal to the length of the gate. Hence, the reference inverter in CNFET also has minimum gate width of 32nm. Since, the performance of CNFETs depends on the number of tubes and pitch, so there can be many possible reference inverters in CNFETs. However, our choice of the reference inverter is based on the desired CNT density (250/ μm) to match the performance of CMOS circuits. The desired density requires at least 8 tubes for minimum width (32nm) inverter. Hence, our reference inverter has 8 tubes separated by 4nm spacing. The normalized parameters like Pitch Factor and tubes (N_{tur}/N_{ref}) are always 1 for a reference inverter. The delay of inverters and other logic gates with different number of tubes and pitch combinations is normalized with respect to the reference inverter. Figure 3.6 shows CNTs and Pitch configurations for reference inverter (INV) template and some other gates. The inveter and other gates with different numbers of CNTs and pitch values need to be normalized using reference inverter and delay is calculated using equations (3.12)-(3.14). Hence, Logical Effort and propagation delay for CNFET-based gates and circuits can be calculated using equations (3.6) to (3.14).

3.3.4 Estimation of τ for Absolute Delay calculation in CNFET-based circuits

In this section, we show the calculation of technology dependent CNFET parameter Tau (τ_0), which is obtained using slope-intercept method as described in [61] and shown in

(a)

(b)

(c)

Figure 3.6: (a) Reference inverter with number of tubes $N_{ref}=8$, Pitch= $4nm$, Logical Effort ($g_{ref}=1$) (b) 2-inputs NAND gate with $N_{tur}=6$, Pitch= $5nm$ (c) Inverter gate with different CNTs and Pitch, $N_{tur}=4$, Pitch= $8nm$. The configurations in (b) and (c) need to be normalized using reference inverter.

Figure 3.7 using SPICE simulations. For 32nm technology the value of τ is calculated as 0.45ps. The absolute delay in the circuits is estimated as shown in (3.14).

Figure 3.7: Technology parameter Tau calculation using different Fanout.

3.4 EXPERIMENT RESULTS

In this section, we compare the delay computed using our PALE model with HSPICE simulations for CNFET logic gates and circuits. We assume that the reference inverter has 8 tubes separated by 4nm pitch. For these set of experiments, we keep the gate width constant, and vary the number of tubes and spacing between the tubes. Since, the gate width and other physical parameters remains the same, the C_{gtg} does not change with different tube and pitch configuration. For circuit-level run, we assume all the gates have same number of tubes and pitch. The fanout for gate and circuit level simulations is considered between 2X – 8X. The correlation between delay estimated from PALE and SPICE simulation is done with same tubes and spacing in the channel. The SPICE model and physical parameters are used from Stanford [20] for simulation purposes. The

Table 3.4: Logical Effort of different CNFET inverter configuration for same width

N_{tur} <i>Pitch</i>	8tubes 4nm	4tubes 8nm	16tubes 2nm	2tubes 16nm
PALE (g)	1	1.39	1.12	2.13

absolute delay of the gates and circuits is calculated using technology delay parameter τ (τ). We only consider ideal CNFET transistors with all semiconducting tubes. The accuracy of the delay model is validated for the impact of two CNFET parameters, 1) number of tubes (CNTs), and 2) pitch. The experiments were performed on 4x Dual Core Intel CPUs at 2.0 GHz and 128 GB RAM.

3.4.1 CNFET Gate Level Results

To validate our model, we started our experiments with small gates like inverter and then extended to include other gates. We simulated inverter with constant width (32nm). Table 3.4 shows the Pitch-Aware Logical Effort (PALE), obtained from the model. It can be observed that for the same gate width (32nm), the driving capability (represented by PALE) of inverter depends on the number of tubes and pitch.

Further, we have used CNFET INV (inverter), NAND2, NAND3, NOR2 logic gates with variable fanout (FO) between 2 and 8 as shown in Table 3.5 . The physical width of the gates is kept the same for different configuration of N_{tur} and pitch ($N_{tur}/Pitch$). The standard LE model results in the same delay for these gates with the same channel width for all $N_{tur}/Pitch$ combinations. However, it can be seen in the results as shown in Figures 3.8 and 3.9, the delay for gates with the same channel width significantly changes depending on the number of CNTs and spacing between them. A little deviation from SPICE simulation is observed for higher fanout, same behavior is also observed in stan-

Table 3.5: Gate level comparison of Delay computation from PALE and SPICE simulations with different FO

FO	INVERTER											
	8tubes/4nm			4tubes/8nm			16tubes/2nm			2tubes/16nm		
	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)
2	1.265	1.26	0.4	1.514	1.60	6.0	1.506	1.41	6.3	2.402	2.44	1.4
4	2.089	2.10	0.5	2.657	2.82	6.13	2.399	2.35	2.1	4.394	4.45	1.3
6	2.842	2.94	3.4	3.764	3.99	6.10	3.160	3.29	4.0	6.283	6.47	2.95
8	3.616	3.78	4.5	4.729	5.1	7.8	3.989	4.23	6.1	8.230	8.48	3.08
FO	NAND2											
	8tubes/4nm			4tubes/8nm			16tubes/2nm			2tubes/16nm		
	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)
2	1.921	2.10	7.9	2.678	2.60	2.8	2.247	2.10	7.0	3.854	3.86	0.2
4	3.463	3.36	3.0	4.554	4.44	2.6	3.337	3.36	0.7	6.913	6.89	0.3
6	4.540	4.62	1.8	6.133	6.24	1.8	4.560	4.62	1.3	10.28	9.91	3.7
8	5.837	5.88	0.7	7.808	8.04	3.0	5.609	5.88	4.7	13.09	12.94	1.2
FO	NOR2											
	8tubes/4nm			4tubes/8nm			16tubes/2nm			2tubes/16nm		
	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)
2	2.089	2.10	0.5	2.615	2.60	0.5	2.066	2.10	1.7	3.761	3.86	2.6
4	3.397	3.36	1.1	4.318	4.44	2.7	3.199	3.36	4.8	7.175	6.89	4.0
6	4.534	4.62	1.9	6.091	6.24	2.4	4.420	4.62	4.4	10.19	9.91	2.8
8	5.647	5.88	4.0	7.863	8.04	2.3	6.108	5.88	4.2	13.59	12.94	5.0
FO	NAND3											
	8tubes/4nm			4tubes/8nm			16tubes/2nm			2tubes/16nm		
	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)
2	2.892	2.94	1.7	3.527	3.66	3.8	2.956	2.94	0.5	5.255	5.29	0.7
4	4.342	4.62	6.5	5.755	6.06	5.2	4.395	4.62	5.0	8.937	9.12	2.0
6	5.559	5.56	0	7.876	8.56	7.3	5.947	6.30	5.9	12.95	13.36	3.0
8	7.271	7.98	9	10.24	10.86	5.8	7.426	7.98	7.0	17.03	17.38	2.0

standard LE model for CMOS circuits. The absence of PF results in the maximum and average error of 55% and 23%, respectively. The newly developed PALE model with the added PF reduces the average error to approximately 3.0% on average. It is observed that the developed model shows very good correlation with SPICE simulations for FO4 metric. Figure 3.10 compares maximum and average error in delay estimation from original and PALE model for two inverters with same width (32nm).

The PALE model is capable of doing quick and accurate analysis without running long HSPICE simulations for different type of gates to make certain decision for design and CNFET parameter choices at gate level. As shown in Figure 3.11, estimated propa-

Figure 3.8: Correlation of inverter delay from the model with SPICE simulations for various tubes/pitch arrangements; 2/16 and 16/2.

Figure 3.9: Correlation of inverter delay from the model with SPICE simulations for various tubes/pitch arrangements; 8/4 and 4/8.

Figure 3.10: Maximum and average error in INV gate delay computation from original and developed LE models with SPICE simulation

gation delay for NAND and NOR is same. By increasing the number of tubes for fixed width results in smaller number of tubes and causes significant increase in the delay. Also, by increasing the number of tubes for fixed width may not improve delay due to higher screening effect.

Figure 3.11: Gate level Delay results for FO4

3.4.2 CNFET Circuit Level Results

For circuit level PALE model validation, we use 3-stage decoder with INV-NAND4-INV topology. The number of branches (B) is 8 and the fanout (H) for the given topology is 9.2 as shown in Table 3.8. The delay predicted from PALE model is compared with SPICE simulation results for a given technology node and using CNFETs with different gate widths. The gate width in CNFET circuits represents the product of the CNTs and the pitch between them. It is observed that PALE model saves significant computational time keeping the average error in delay with SPICE model around 2.15%. The small runtime of developed model, scalable to different technology node, makes it suitable for early delay analysis and exploration. The delay in the decoder at 2nm pitch and 16 tubes is 13.37ps, which reduces to 12ps when the number of tubes increase to 30 with gate width of 60nm. This shows that almost doubling the tubes, reduces the delay in the decoder merely by 10.27%. However, at 4nm pitch, the delay in the decoder reduces to 10.96 ps (reduction of 18%), with only 8 tubes in the transistor (50% smaller tubes than at 2nm pitch). Moreover, when the transistor has 15 tubes in the channel, the decoder delay at pitch 3nm is around 10.42ps, which reduces by 12% to 9.17ps at 4nm. Hence, we can conclude, that at pitch smaller than 10nm, due to charge significant screening effect, increasing CNTs in the channel have limited influence on the driving capability of the transistor. Whereas, at larger pitch, the number of CNTs have more linear impact on the delay.

Next, we validate the model for a multiple-stage circuit and multiple-branch circuits. We have assumed the fanout of 4 for all the circuits with the same gate width (32nm). The delay computed from the developed model (PALE) and SPICE simulations is reported in Table 3.6. The average error in estimated delay for multi-stage circuits is 5.2%, while for

multi-branches circuits is around 3.5%.

Figure 3.12 shows average error in delay computation between SPICE simulations and developed model of the test circuits with description shown in Table 3.7. It was observed that keeping the same channel width and increasing the number of tubes does not result in better performance due to significantly larger screening effect due to reduced pitch. It is interesting to observe that the gates with 16tubes/2nm configuration have twice the number of tubes as compared to 8tubes/4nm. But due to smaller spacing between tubes, the delay is not better than the gate with 4nm pitch. This influence is successfully captured at both gate and circuit level by the developed model and the average error in the given range of pitch is less than 5%. The developed model allows accurate prediction of the delay in CNFET circuits for wide range of pitch and numbers of tubes in much shorter time as compared to SPICE simulations. The CNFETs in logic gates with sufficiently large CNT pitch can be represented by edge tubes, e.g., the circuits with 2tubes/16nm pitch in CNFETs have minimal screening effect and tubes can be represented as edge tubes. The minimal error is observed in estimated delay due to a smaller screening effect. The delay in this configuration is driven by number of tubes, which have linear behavior and hence, easy to capture.

As we discussed, PALE model is useful for quick and accurate estimation of delay at circuit level and making important decisions at early design phase, such analysis at circuit level is very important for fast design convergence with minimum design iterations. As shown in Figure 3.13, for pitch less than 6nm, screening effect dominates and for pitch higher than 6nm, the effect of number of tubes dominates. As discussed, increasing the number of tubes in a fixed width transistor may not improve driving capability always. For 16nm pitch, the only 2 tubes results in 45% larger delay on average. The circuits with 4 CNTs and 8nm pitch have 15% larger delay on average compared to reference circuit.

Figure 3.12: Average error in delay computation between SPICE simulations and developed model of the test-circuits.

Figure 3.13: Circuit level Delay results for FO4.

Table 3.6: Circuit level comparison of Delay computation from model and SPICE simulation for a) multi-stage circuits b) multi-branch circuits for FO4

CNFET circuits	MULTI-STAGE CIRCUITS											
	8tubes/4nm			4tubes/8nm			16tubes/2nm			2tubes/16nm		
	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)	SPICE	PALE	Δ (%)
INV-NAND2	5.65	5.38	5%	6.52	6.18	5%	5.383	5.38	0%	8.052	7.64	5%
INV-NOR2	5.69	5.38	5%	6.57	6.18	5%	5.298	5.38	-1%	7.972	7.64	4%
INV-INV-INV	4.66	4.55	2.3%	4.91	5.51	-10%	5.45	5.00	-9%	6.67	6.92	-4%
INV-NAND2-NOR2	9.78	8.70	9%	10.62	10.45	2%	8.278	8.70	-9%	13.52	13.94	-3%
INV-NAND2-NAND3	10.7	9.37	12%	11.95	11.29	5%	10.10	9.37	7%	15.89	15.13	5%
	MULTI-BRANCH CIRCUITS											
2:4 Decoder	11.01	11.21	-1.8%	8.78	8.47	3.7%	7.58	7.42	2.3%	7.95	8.26	-4.8%
3:8 Decoder	15.46	14.87	4.0%	12.31	12.66	-3.9%	10.62	11.07	-4.1%	11.13	11.54	-3.6%

Test-cases	Description
Test I	INV-NAND2
Test II	INV-NOR2
Test III	INV-INV-INV
Test IV	INV-NAND2-NOR2
Test V	INV-NAND2-NAND3

Table 3.7: Test-cases description

Table 3.8: Comparison of Delay from LE model and SPICE simulations for 3-stage Decoder, with B = 8 and H = 9.2

Gate Width	CNT Pitch	No. of Tubes per Transistor	Logical Effort (g)	Normalized Delay (d_{norm})	Absolute Delay (d_{abs}) (ps)	Model runtime (sec)	Simulations (d_{sim}) (ps)	Simulations runtime (sec)	Δ Delay(%)
32nm	2nm	16	5	27.49	12.92	0.117	13.377	270.27	3.38
	4nm	8	2.5	23.06	10.83	0.117	10.956	333.52	1.06
	8nm	4	3.57	25.21	11.85	0.117	12.192	352.3	2.78
	16nm	2	6	28.84	13.55	0.117	13.438	336.01	0.88
45nm	3nm	15	2.08	22.05	10.36	0.117	10.417	279.81	0.48
	5nm	9	1.70	21.00	9.87	0.117	9.823	303.17	0.51
	9nm	5	1.89	21.54	10.12	0.117	10.165	318.92	0.37
	15nm	3	2.45	22.96	10.79	0.117	10.417	291.7	3.61
60nm	2nm	30	2.66	23.43	11.21	0.117	12.003	281.01	6.57
	3nm	20	1.56	20.58	9.67	0.117	9.769	315.72	0.94
	4nm	15	1.33	19.83	9.32	0.117	9.169	313.77	1.68
	5nm	12	1.27	19.63	9.22	0.117	9.026	312.98	2.24
	6nm	10	1.28	19.65	9.23	0.117	9.097	320.83	1.52
	10nm	6	1.48	20.33	9.55	0.117	10.14	310.94	5.82
	12nm	5	2.59	23.26	10.93	0.117	10.79	309.73	1.31
	15nm	4	3.6	25.26	9.87	0.117	9.767	298.69	0.92
	20nm	3	5.95	28.79	13.53	0.117	13.365	314.86	1.25
	30nm	2	5.51	28.21	13.26	0.117	13.731	327.62	3.42

3.5 SUMMARY AND CONCLUSION

Our experiments show clearly that the CNFET specific parameters, such as tube diameter and pitch impact the performance of gates and circuits. Using Logical Effort model that was developed for CMOS gates would introduce significant errors to delay evaluations of CNFET-based circuits. The PALE model presented in this chapter models one of the most critical parameter which is charge screening effect. Delay evaluations performed using the proposed pitch-aware Logical Effort model generated results are on average within 3.7% for gates and 5.2% for circuits as compare to SPICE simulation values for fanout ranging from 2 to 8. It was observed that keeping the same channel width and increasing the number of tubes does not result in better performance due to significantly larger screening effect due to reduced pitch. The developed model enables accurate delay prediction in CNFET circuits for wide range of pitch and numbers of tubes in much shorten time that could be done with SPICE simulations. The CNFETs in logic gates with sufficiently large CNT pitch can be represented by edge tubes. The least impact of screening in these gates results in more consistent delay correlation with SPICE simulations.

CHAPTER 4

POSITION-AWARE PITCH-FACTOR MODEL

Part of this chapter has been published in:

- M. Ali, M. A. Ahmed and M. Chrzanowska-Jeske, "Logical Effort Framework for CNFET-Based VLSI Circuits for Delay and Area Optimization," in IEEE Transactions on Very Large Scale Integration (VLSI) System, March 2019, pp. 573-586.
- M. Ali, M. Ahmed, and M. Chrzanowska-Jeske, "Fast and Accurate Evaluation of delay in CNFET Circuits," in 2016 IEEE 16th International Conference on Nanotechnology (IEEE-NANO), Aug 2016, pp. 659–662.
- M. Ali, M. Ahmed, and M. Chrzanowska-Jeske, "Stochastic Analysis of CNFET circuits using Enhanced Logical Effort Model in the Presence of Metallic Tubes," in 2014 21st IEEE International Conference on Electronics, Circuits and Systems (ICECS), Dec 2014, pp. 774–777.
- M. Ali, M. Ahmed, M. Chrzanowska-Jeske, and J. Morris, "Logical Effort model for CNFET circuits with CNTs variations," in 2015 IEEE 15th International Conference on Nanotechnology (IEEE-NANO), July 2015, pp. 1218–1221.

The statistical nature of variations using some analytical modeling are explored in [7], [14] and [71], which are mainly helpful to predict performance at gate level using currents obtained from SPICE simulations. Our focus is to predict performance at circuit

level using capacitance based Logical Effort framework. Authors in [19] and [56] discussed that statistical variations are easy to be modeled in process dependent parameter τ_0 . As discussed in [56], physical parameters variability in CMOS can be translated in output loads using gate delay models. The authors in [8] proposed Stochastic Logical Effort (SLE) to capture the effect of statistical parameter variations on the delay for CMOS.

Recently, some approaches have been reported in the technical literature to develop models to evaluate performance and yield of CNFET. Cho *et al.* [13] talked about MATLAB-based model to analyze performance of CNFETs for variable CNT spacing at gate level only without presenting any closed form model for circuit level. The SPICE simulation-based linear programming (LP) approach is presented in [15] at circuit level, it is observed that non-linear simulations for bigger circuits are very time consuming and equivalence relationship approach proposed in this paper to calculate parameters imposes difficult requirements to determine parameters required for SPICE simulation, which adds up to complexity, error and run time. The proposed approach was not demonstrated for bigger and complex circuits. Hills *et al.* [31], Ghavami *et al.* [29] [28] and Ashraf *et al.* [7] discussed about CNFET variations and impact on CNFET-based gates and methods to improve yield.

In this chapter, we focus to develop Position-Aware Pitch-Factor (PAPF) LE model to account for variations in CNTs due to density and metallic tube removal. Our research for PAPF development consists of two parts. We first developed simple model based on certain assumption using PALE which is referred as Evenly-Spaced LE (ESLE) model, then more accurate realistic model (PAPF). We have discussed both models below including comparison results for both.

4.1 EVENLY-SPACED LE (ESLE) MODEL

As discussed previously, the delay estimation using conventional LE model cannot comprehend screening effect and it needs to be enhanced to account for CNFET-specific parameters for accurate performance evaluation. We incorporated the developed PALE model for performance evaluation of CNFET-based circuits in presence of CNTs density and count variations. The development of ESLE model is intermediate step between PALE and PAPF, and can be referred as PALE-based variation-aware model. We make certain assumptions in the beginning for ESLE model development, some important assumptions are addressed as part of PAPF model. We evaluate the impact of metallic tube variations using our PALE model in Monte Carlo-based analysis in beginning. The steps and assumptions for delay estimation in the presence of metallic tubes using Monte Carlo analysis are discussed below.

- **Number of CNTs in CNFET:** It is decided based on the number of desired tubes in a transistor, number and type of gates. We have assumed that in an ideal case the transistor has 8 semiconducting tubes. Hence, a tube array consists of 8 tubes. The tubes have diameter varying between 1nm to 2nm.
- **Percentage of Metallic Tubes:** The percentage of metallic tubes in the given sample depends on the growing technique. In our analysis we have assumed percentages of metallic tubes between 1% and 25%.
- **Removal of Metallic Tubes:** The two most preferred techniques for tubes removal are used. Based on published data [12], we can assume that VMR technique results in 100% of metallic and 0% of semiconducting tubes being removed. As discussed in [69], we assume SCE technique removes 100% metallic tubes but also

some percentage of semiconducting tubes. In SCE, the tube removal is based on the tube diameter. To generate the number of semiconducting tubes removed, we used diameter threshold of 1.4nm as suggested in [69]. We also assume that no transistor has less than two tubes after the removal process. The more details about SCE and VMR removal techniques are discussed in chapter 2.

- **Number of Tubes and Pitch:** The transistors are constructed based on the remaining tubes in the CNT array. We assume an even distribution of remaining s-CNT tubes in the channel. The pitch between the tubes is then easily calculated for each transistor. The removal of m-CNTs and remaining evenly spaced s-CNTs are shown in Figure 4.1.
- **Delay Estimation:** The pitch-aware LE model is used to predict the delay for each instance of the circuit. The delay of the instances may be different depending on the number of remaining tubes and estimated pitch. As it can be seen, the mean delay correlates well with the SPICE simulations. Since PALE model cannot model position of remaining CNTs in the channel after removal of metallic CNTs (m-CNTs) so it is assumed that remaining tubes are spaced evenly with same Pitch value for both PALE and SPICE simulations.

4.2 POSITION-AWARE PITCH-FACTOR (PAPF) MODEL

After further research, it is observed that the CNT density variations and the presence of metallic tubes and their removal may result in non-uniform pitch distribution. We consider non-uniform pitch distribution based on density variations, which is significant enough to impact delay. The impact on delay due to small density variations is assumed

Figure 4.1: (a) The m-CNT (blue) and s-CNT (black) CNFET in CNFET channel. (b) The CNFET channel after removal of m-CNTs. (c) The CNFET channel with s-CNTs with evenly spaced assumption.

within limits of tolerance due to statistical averaging similar to diameter variations. The standard LE model for CMOS circuits does not comprehend these variations and needs to be enhanced for accurate delay modeling. ESLE model discussed in section 4.1 is not accurate for all scenarios as shown in the Experiment section of this chapter. So we present the development of accurate LE model (PAPF) in this section.

As discussed in [30], there are four different scenarios applicable during CNFETs placement over aligned CNTs as shown in Figure 4.2(a). CNTs (s-CNT or m-CNT) are assigned under CNFET active region in random number (N_{m-CNT} for m-CNT and N_{s-CNT} for s-CNT). It is shown, CNFET₁ has two s-CNTs and one m-CNT. CNFET₂ and CNFET₃ have one and two s-CNT in the active region, respectively. There is no CNT under CNFET₄ and it is referred as a “void CNFET” such CNFET configuration results in an *open defect* (CNFET₄). m-CNT in the active region of a CNFET causes source-drain to short and referred as *short defect* (CNFET₁). A CNFET can be a “functional” if it encounters neither open nor short defects (CNFET₂ and CNFET₃). Fig 4.2(b) shows the classification of a CNFET based on above description. If the functionality of a CNFET changes due to either an open or a short defect, it is considered as *defective*.

As discussed previously that the two methods, SCE and VMR, are used to remove unwanted metallic tubes. A given percentage of metallic tubes is identified by P_m factor and the the total percentage of tube removed is given by P_r . The SCE removes all the metallic tubes, but may also remove some of the semiconducting tubes. The percentage of semiconducting tubes removed depends on the diameter of the tubes. The VMR is a highly efficient technique, removing all the metallic tubes only. Both of these metallic tubes removal techniques cause CNTs density variations in the CNFET channel, which results in non-uniform pitch as shown in Figure 4.2(c). To account for non-uniform pitch effect, the analytical model for the gate-to-channel capacitance (C_{gc}) can be rewritten

for tubes in the middle [16] and is given by (4.1), which includes the influence of non-uniform pitch between tubes on the charge screening effect since the pitch on either side of a middle tube may be different if adjacent tube is removed.

$$C_{gc_m} = \frac{C_{gc_inf}C_{gc_sr}(P1).C_{gc_sr}(P2)}{C_{gc_sr}(P1).C_{gc_sr}(P2)+C_{gc_sr}(P2)C_{gc_inf}+C_{gc_sr}(P1)C_{gc_inf}} \quad (4.1)$$

$$C_{gc_e} = \frac{1}{C_{gc_inf}} + \frac{1}{C_{gc_m}} \quad (4.2)$$

Where, $P1$ and $P2$ are pitch on the either side of the middle tube. As shown in Figure 4.2(c), $P1$ and $P2$ values for a tube depend on the charge screening of tubes on either side. The neighboring metallic tubes and their removal may increase either or both $P1$ and $P2$. However, the edge tube has only one pitch value which depends on screening effect of adjacent tube as given by (4.2).

4.2.1 Statistical Analysis of Capacitance in the presence of Metallic Tubes

Further, Monte Carlo simulations are used to generate sample population of logic gates, using sample tubes with given diameter ranges from 1 to 2nm, which is required to construct the given number of gates for statistical analysis. The random removal of tubes using SCE and VMR results in non-uniform pitch distribution in transistors. In our analysis we have assumed percentage of metallic tubes present is between 0% and 25%. We use 1.4nm as diameter threshold for removal of semiconducting tubes as suggested in [69] for SCE. We also assume that no transistor has less than two tubes after the removal process. The gate-to-channel capacitance for each tube is estimated based on its position and neighboring tubes using (4.1). The influence of non-uniform pitch is also considered for the edge tubes. The capacitance of each tube in the tube array for each

Figure 4.2: (a) CNFETs randomly placed on aligned s- and m-CNTs. N_{m-CNT} represents the number of m-CNT and N_{s-CNT} expresses the number of s-CNT in a CNFET. $CNFET_2$ and $CNFET_3$ are functional where $CNFET_1$ and $CNFET_4$ have short and open defects, respectively. (b) The classification of CNFETs defects based on the number of m- and s-CNTs placed in the active region. (c) Impact of metallic tube removal resulting in non-uniform pitch.

transistor array is summed up to obtain the total capacitance of each gate. The impact of positions of removed tubes on the gate capacitance of a basic inverter is shown in Figure 4.3. The inverter gate capacitance for given number and position of the tubes removed are normalized to ideal case with no metallic tubes. We have conducted detailed analysis to study impact of number of tube removed from a certain position on normalized gate capacitance which is published in [4] along with initial PAPF implementation. It is evident from the experiments that the initial assumptions of uniform pitch after removal of metallic tube [65] introduces significant error in the delay estimation. We share the comparison results between ESLE and PAPF in experiment section below.

Figure 4.3: Influence of position of CNT removal on the gate capacitance, 1 CNT removed per transistor, 2 CNTs removed per transistor.

4.2.2 PAPF Model Closed-form Expression

The mean capacitance of the given population of inverter gates for different values of P_m is calculated. The reference inverter with $P_m = 0\%$ is used to normalize the mean capacitance of inverter gates value at each P_m value. The PAPF in (4.3) models the behavior of normalized capacitance for given P_m , P_r and applied removal technique. The fitness factor α represents the σ/μ at $P_m = 0$ and captures the effect of different ranges of transistor tubes (N_{tur}) on capacitance.

$$PAPF = (1 - \alpha.(x.P_m + (1 - x).P_r)) \quad (4.3)$$

where $\alpha = (\sigma/\mu)_{P_m=0} = 0.005$, $x=1$ for VMR and $x=0$ for SCE. Now, LE model can be represented as follow:

$$g' = g.(\frac{1}{PAPF}) \quad (4.4)$$

Normalized delay can be calculated by (4.5).

$$D_{norm} = g'.h + p \quad (4.5)$$

The normalized capacitance values at different P_m value for CNFET with different tube array sizes is shown in Figure 4.4. It is observed that the capacitance of inverter gate reduces linearly with increase in P_m which is expected. As the number of tubes removed P_r increases with increase in number of tubes, the mean capacitance at higher P_m also decreases significantly. The average slope of the capacitance is captured in the form of α which also represents standard deviation to mean ratio at $P_m = 0\%$. As discussed, the use of SCE technique also removes the semiconducting tubes. Hence, the total number of tubes removed (P_r) factor is used instead of percentage of metallic tubes (P_m).

To demonstrate the impact of removing techniques, we measured mean and variance of gate capacitance for different removal techniques at different percentage of P_m and P_r . It is observed as shown in Table 4.1, the mean capacitance at $P_m=5\%$ for VMR is around 21% more than the SCE technique. A smaller capacitance variance in VMR compared to SCE for similar P_m signifies large percentage of inverter instances has gate capacitance around the mean value. It is due to the simple fact that the variance depends really on the percentage of removed tubes (P_r) and not on P_m . The CNFET circuits with VMR

Table 4.1: Mean and Variance of 1000 inverter instances in the presence of different percentage of metallic tubes and their removal technique

Removal Technique	P_m (%)	P_r (%)	Mean Cap (aF)	Variance Cap (aF)
SCE	0.0	0.0	77.16	0.511
	5.0	31.0	62.57	5.379
	10.0	35.0	60.05	6.408
	15.0	37.8	57.57	7.287
	20.0	40.8	55.76	7.769
VMR	0.0	0.0	77.16	0.511
	5.0	5.0	75.84	2.375
	10.0	10.0	74.01	3.731
	15.0	15.0	72.01	4.658
	20.0	20.0	69.94	5.765

Figure 4.4: Mean capacitance of inverter gate at given P_m (VMR) normalized to $P_m = 0\%$ for different N_{tur} .

technique result in superior delay performance compared to SCE technique because SCE also removes a significant percentage of semiconducting tubes in addition to removing metallic tubes. In the SCE technique, the ratio between percentages of metallic tubes removed to all tubes removed increases with the increase in the percentage of the metallic

Table 4.2: PAPF for VMR and SCE removal techniques for different logic gates at different percentage of P_m

VMR (x=1)					SCE (x=0)				
$P_m(\%)$	PAPF	Logical Effort (g')			$P_r(\%)$	PAPF	Logical Effort (g')		
		INV	NAND	NOR			INV	NAND	NOR
0	1.00	1.0	1.5	1.5	0	1.0	1	1.5	1.5
5	0.975	1.025	1.538	1.538	31	0.845	1.183	1.775	1.775
10	0.95	1.052	1.578	1.578	35	0.825	1.212	1.818	1.818
15	0.925	1.081	1.622	1.622	38	0.81	1.234	1.852	1.852
20	0.9	1.11	1.667	1.667	41	0.795	1.257	1.887	1.887
25	0.875	1.142	1.714	1.714	44	0.78	1.282	1.823	1.823

tubes present initially in the tube array. Hence, it is important to include the impact of removal technique in the model for quick and accurate estimation of delay as shown in PAPF model in (4.3).

Table 4.2 shows the PAPF and Logical Effort (g') calculation for VMR and SCR removal techniques at different percentage of P_m . PAPF model calculates LE values for CNFET gates in quick and accurate way.

4.3 EXPERIMENT RESULTS FOR PAPF LE MODEL

We have assumed SCE and VMR techniques for tube removal. A sample size of tubes is generated to create N instances of a gate or circuit using MATLAB. The metallic tubes are randomly selected within the sample size based on the user define percentage of metallic tubes. For VMR, we assume only metallic tubes are removed, whereas, SCE also removes some of the semiconducting tubes with diameter below defined threshold (1.4nm). After random tube removal from the distribution, we construct the transistors with an array of 8 tubes. The position of tubes removed for each transistor is tracked

and same data is exported to SPICE simulation for comparison.

4.3.1 Comparison of PAPF and ESLE

As discussed in ESLE section, after the removal of metallic tubes and some of semiconducting using CNT removal techniques; the transistors were built with the remaining tubes. The average pitch for the transistors for modeling purpose was estimated based on the remaining tubes in the transistor array. The tubes were assumed to be separate by uniform pitch. However in reality, the pitch between the tubes is nonuniform and depends on the position of the tubes removed. The tube removal and increase in the pitch only impacts the adjacent tubes. The Figure 4.5 shows the comparison of gate-to-channel capacitance for position-aware and position-unaware scenarios. The difference between the capacitance values of position-aware and position-unaware increases with increase in P_m . As more and more tubes are removed, the impact of tube position on capacitance becomes significant. It can be also observed that the difference between the two is significant for SCE techniques even at lower P_m value. Since, the semiconducting tubes are also removed in SCE technique resulting in large P_r . The capacitance value represents the mean of 1000 inverter instances.

4.3.2 PAPF correlation with SPICE Simulation Methods

For correlation with the simulation methods, we have conducted statistical analysis for different circuits sizes like INV-INV and decoders. We measure mean and variance delay using PAPF LE model and compare with the SPICE simulation results. The SPICE simulations are also performed by considering the impact of P_m on the CNTs for right comparison. In our first experiment, 1000 instances of a given test circuit are created using sample population of tubes with initial presence of metallic tube in range of 0% to

(a)

(b)

Figure 4.5: Comparison between inverter gate capacitance for position-aware and evenly-spaced tubes. The values represent the mean of 1000 instances. (a) VMR (b) SCE

20%. The metallic tubes are randomly identified from the given tubes population, and removed using VMR technique, which is most efficient technique for removing metallic tubes. The results of statistical analysis for two stage inverters using PAPF LE model and Stanford SPICE model are shown in Figure 4.6. The results confirm that mean and variance of delay distribution from LE model correlates well with SPICE model, with average error less than 3%. It can be observed that for 20% metallic tubes (P_m), the delay in the circuit increases on an average by 15%. Moreover, the variance in delay is much higher and mean delay is worst at smaller pitch due to screening effect as shown in Figure 4.6(a). It suggests that the delay in CNFET circuits show less variation with tube removal at larger CNT pitch.

4.3.3 Impact of C_{gtg} on the Variance in Statistical Delay

In some cases, by ignoring the C_{gtg} capacitance had negligible impact on the mean delay. However, the variance in the delay distribution compared to SPICE simulation is significantly large as shown in Figure 4.7. The average variance in the delay without C_{gtg} is 71%, which reduces to around 2% when C_{gtg} is included in the capacitance formula. Hence, we have included C_{gtg} in our analysis and models development.

4.3.4 Decoder circuit-level Analysis

Next, we do the statistical delay analysis for more complex circuits, such as decoder, the results are shown in Figure 4.8. In this experiment, 100 decoders with $P_m=5\%$ are simulated using Stanford SPICE model as well as PAPF LE model. The mean delay from the SPICE simulations and the model is 11.83ps and 11.9ps respectively, with an error of 0.53%. The variance of delay distribution using the model has an error of 2.64%. Apart from very few corner cases, the model correlates closely with the SPICE simulations.

(a)

(b)

(c)

Figure 4.6: Statistical analysis of 1000 instances of two stage inverter using Stanford Spice and CNFET LE model for CNT pitch at (a) 2nm. (b) 6nm. (c) 10nm

Figure 4.7: Impact of C_{gtg} on the variance in statistical delay of a CNFET circuit in the presence of metallic tubes.

It is observed that for 5% metallic tubes, the circuit delay may vary up to 14%. Our experiments suggest that the removal of tubes is more significant for larger CNT pitch. The increase in delay due to removal of tubes at smaller pitch is mitigated by the reduced screening effect. The runtime of statistical delay analysis for decoder circuits using PAPF based model is 0.120 sec as compared to 32200 sec from SPICE simulation, which can be highly desirable for early fast and accurate delay analysis. We have conducted detailed runtime analysis for different circuit types by PAPF LE model, this is presented in [6].

4.3.5 Runtime Analysis

Table 4.3 shows the runtime comparison of CNFET LE model with SPICE simulation results for different stages, branching and circuit complexity, and number of instances. The circuit with single instance represents the ideal case with no metallic tubes present. The circuits with multiple instances represent the statistical analysis in the presence of metallic tubes. The percentage of metallic tubes has no impact on the SPICE simulations

Figure 4.8: Delay distribution of 100 decoder circuits for $P_m = 5\%$ and with 5% error amount.

Table 4.3: Runtime comparison between SPICE simulations and CNFET LE model.

Circuit (instances)	Simulation (sec)	Our Model (sec)
INV-INV (1)	18.8	0.052
NAND-NAND (1)	21.2	0.075
NAND-NOR-INV (1)	62.3	0.094
2-stage decoder (1)	46.8	0.087
3-stage decoder (1)	310.5	0.117
3-stage decoder (100)	32200	0.120
INV-INV (200)	3980	0.118

runtime. The runtime of simulations depends greatly on the circuit topology.

4.4 SUMMARY AND CONCLUSION

We have analyzed the impact of metallic tube removal and resulting non-uniform pitch distribution on the gate capacitance and delay in the circuits. The position-aware gate capacitance shows large difference from the position-unaware capacitance at higher P_m factor. The difference between the capacitances calculated considering removed tube

positions and calculated using averaged pitch values, increases with increasing percentage of metallic tubes initially present, P_m . The variance of position-aware capacitance distribution increases with P_m , as the influence of position becomes significant. The delay estimated from position-aware Logical Effort model correlates well with the SPICE simulations. The difference in delay calculation for CNFET-based gates and circuits estimated from PAPF model is within 2% to 5% range as compared to SPICE simulation. The delay estimated from the developed LE model works well for P_m from 0% to 25%.

CHAPTER 5

OPTIMIZATION OF CNFET CIRCUITS

Part of this chapter has been published in:

- M. Ali, M. A. Ahmed and M. Chrzanowska-Jeske, "Logical Effort Framework for CNFET-Based VLSI Circuits for Delay and Area Optimization," in IEEE Transactions on Very Large Scale Integration (VLSI) Systems, March 2019, pp. 573-586.

In this chapter, we present a comprehensive methodology and steps for optimizing the performance and area of CNFET circuits [3]. As discussed in the introduction section, most of the proposed methods and techniques may help to improve the performance of the gates or to ensure that the gates are functional. However, the impact of critical circuit level parameters is ignored such as wire parasitic, number of stages, fan-out (FO) and branching effort and so on in the CNFET circuits. Hence, these methods have limited solution space which may result in the local optima. In this section, we propose comprehensive performance and area optimization framework for CNFET circuits based on developed PALE and PAPF LE models and LSGS (Large-scale Gate Sizing) algorithm from Joshi *et al.* [63]. The existing work [33] [63] modeled the combinational circuits as geometrical programming (GP) problem and presented the method for fast optimization of CMOS gate sizes. The key features developed for CNFET-based circuits as part of LSGS algorithm are discussed as follow.

1. Implementation of circuit level techniques (CLT), such as optimizing the number of stages, the fan-out and branching effort, to achieve global optima as well as optimization of number of CNTs or gate sizing in a CNFET.
2. Incorporating the impact of wire parasitic on the circuit performance for more realistic and accurate optimization using Rent's Rule method to obtain the wire-length distribution of each circuit.
3. Implementation of PALE and PAPF LE models to include the impact of CNFET screening and CNT density variations during optimization to meet the desired performance specifications accurately.

5.1 OPTIMIZATION METHODOLOGY

In this section, we present a comprehensive methodology and steps for optimizing the performance and area of CNFET circuits.

5.1.1 Optimization Flow

The flow diagram of the developed optimization framework is shown in Figure 5.1. The technology parameters such as CNT pitch, initial N_{tur} , percentage of metallic tube (P_m), percentage of removed tubes (P_r), etching techniques such as VMR or SCE, gate netlist and circuit topology are provided as input to the optimization framework. We write the equations for delay and arrival time for each gate based on PALE and PAPF-based LE models for ideal and realistic cases with metallic tubes variations, respectively.

Figure 5.1: Flow diagram of optimization framework for CNFET-based circuits.

5.1.2 Cost Function

The objective function of the problem is the optimization of gate sizes to achieve minimum area meeting the desired timing specifications, as shown in (5.1).

$$\begin{aligned} \min_{(area) \in \mathbb{R}} C(x) &= area \\ \text{subject to } \sum_{i=0}^n d_i^{path} &\leq d_{max} \end{aligned} \quad (5.1)$$

where d_i^{path} represents the delay of the i_{th} gate in a certain path and d_{max} is the upper bound on the gate delay.

The LE framework suggests that the number of stages can be critical in determining the overall delay of the circuits. In our algorithm, it is achieved by randomly selecting the nets and adding/removing even number of buffers from the circuit without altering the functionality. The FO of the critical nets is optimized before doing the sizing of all the gates. It helps to start the optimization phase with a better synthesized netlist with considering the screening effect and density variations, which useful for the algorithm to converge with global optimum solution. The cost function is evaluated at each iteration until stopping criteria are met. Our algorithm terminates if the certain number of iterations defined by the user are reached or desired timing specifications are met. The effectiveness of our approach in improving the area and performance of the benchmark circuits is discussed in the results section. The optimization algorithm and steps are discussed as follow.

Algorithm 1 Optimization Algorithm

```
1: x=initial number of tubes (gate size);
2:  $d_{min}$ =minimum delay of each gate;
3: k = wire load of each gate;
4: FO= fan-out;
5: area= initial area;
6: delay = initial delay based on  $d_{min}$ , k and FO;
7: T=timing specification vector;
8: while  $itr \leq length(T)$  do

9:     Apply CLT
10:     $a_{min}$  = minimize area s.t.  $d_{max} \leq T$ 
11:    new_x = x + (apply CLT+ $a_{min}$ )
12:    new_area = area + (apply CLT+ $a_{min}$ )
13:    new_delay = delay + (apply CLT+ $a_{min}$ )
14: end
15: return (new_x, new_area, new_delay)
```

5.1.3 Algorithm

1. **Initialization:** In the first step, all the gates are assigned initial number of tubes to minimize delay, represented as a vector x . The minimum delay of each gate is computed using PALE, and represented as d_{min} vector. The wire load for each gate is computed using Rent's Rule, and represented as k vector. The primary outputs are assigned maximum load value. The gate connectivity is represented as fan-out (FO) matrix, which defines fan-out for each gate. The initial area and delay of each gate are computed based on the x , d_{min} , k , FO.
2. **Timing Specification:** We define the timing specification vector (T) that each circuit is required to meet. The length of the vector T defines the number of iterations.
3. **Apply CLT:** We apply certain CLT to optimize the performance of the circuit

prior to tube optimization. These techniques are categorized as: (1) optimizing the number of stages (buffer insertion); (2) optimization of branching effort; (3) FO optimization and (4) optimizing the gate sizes. These techniques are applied to individual nets at each iteration prior to tube optimization. The CLT (1) and (2) are used for limited circuits with additional steps.

4. **Tube Optimization:** After performing CLT, we optimize the number of tubes in each gate to minimize the total area such that the maximum delay of each path does not exceed the timing specified in step 2. This step is performed using LSGS algorithm.
5. **Stopping Criteria:** The algorithm terminates after a certain number of iterations. The number of iterations is defined by timing specification. The final output of the flow is the optimized gate sizes, delay of each gate and area of the circuit.

5.2 PARAMETERS ESTIMATION

In this section, we describe the methods used for estimation of parameters defined in the cost function.

5.2.1 Delay Estimation

The delay of each gate is represented in (5.2), and depends on its size and size of load it drives.

$$D_i = d_i^{min} + \frac{g_i + \sum_{j \in FO(i)} F_{ij} x_j}{x_i}, \text{ where } i = 1, 2, ..n \quad (5.2)$$

Alternatively, the delay model in (5.2) can be represented as RC delay model of gates and wires as shown in (5.3). Here, r_i , c_i^{int} and c_i^{in} are driving resistance, internal capacitance

and the input capacitance of a minimum-size inverter of gate i , respectively. c_i^{wire} is the wire load capacitance for gate i . It is assumed that the internal capacitance and input capacitance of gate scales linearly with the scale factor, and therefore, it is represented as c_j^{in} .

$$D_i = r_i x_i \left(c_i^{int} x_i + c_i^{wire} + \sum_{j \in FO(i)} c_j^{in} x_j \right) \quad (5.3)$$

The equivalence between timing models in (5.2) and (5.3) can be represented as below.

$$d_i^{min} = r_i c_i^{int}, g_i = r_i c_i^{wire}, \quad i = 1, 2, \dots, n,$$

$$F_{ij} = \begin{cases} r_i c_j^{in}, & j \in FO(i) \\ 0, & otherwise \end{cases}$$

However, in our implementation d_i^{min} for CNFET gates is estimated based on PALE. If there are no metallic tubes in the circuit, ($P_m = 0\%$), internal delay of the gates is obtained using pitch factor (PF) from (3.6) and (3.10), which depends on CNT pitch and their count. On the other hand, for P_m greater than 0%, the delay of the gates also depend on the position of metallic tubes removed, and it is estimated using PAPF LE model from (4.3) and (4.4).

$$d_i^{min} = \begin{cases} \frac{r_i c_i^{int}}{PF}, & P_m = 0\% \\ \frac{r_i c_i^{int}}{PAPF}, & otherwise \end{cases}$$

5.2.2 Impact of Wire Load and Estimation

The wire load for output gates is computed based on the wire-length distribution obtained for each benchmark circuit using the Rent's Rule. We have used longest wire to compute the value of constant load driven by all the output gates. The inclusion of wire parasitic is necessary for realistic optimization of tubes and gate sizes for CNFET circuits. We take an example of small ISCAS benchmark circuit (c17) to show that the impact of wire load. The gate #5 and #6 are the output gates and drive constant wire load defined by minimum, average and maximum (worst) wire-length. Figure 5.2(a) shows that the delay of output gates vary significantly under different wire load and cannot be ignored during the optimization. If we ignore the impact of wires during optimization, we may get optimum gate sizes as for minimum wire in Figure 5.2(b). However, the optimal gate sizes needed for worst size are double of minimum wire. Figure 5.3 shows the optimal gate sizes in 10k circuit with and without applying the CLT. After applying the CLT, the gate sizes of some critical gate is larger than their size without any CLT. Hence, our optimization algorithm takes the impact of wire load into account and optimize the gate sizes better for overall minimization of the area and delay.

Figure 5.2: Impact of wires on the (a) delay of c17 circuit, (b) optimal gate sizes of c17.

Figure 5.3: Optimized gate sizes in 10k circuit with and without CLT.

In addition, c_i^{wire} in our algorithm is computed using the Rent's Rule [18] given by (5.4). The longest wire obtained from the wire-length distribution is used to calculate the total wire capacitance. The capacitance per unit length for $32\mu\text{m}$ ($0.2fF/\mu\text{m}$) is multiplied by the longest wire in the circuit to obtain wire capacitance which acts as load for all output gates. The delay estimation and optimization are done, considering a constant wire load, which results in better optimization of circuit performance. Since, the number of buffers and their positions are not known until placement and routing stages, so we have ignored their impact on the wire delay estimation during optimization.

Region I: $1 \leq \ell \leq \sqrt{N}$

$$i(\ell) = \frac{\alpha k}{2} \Gamma\left(\frac{\ell^3}{3} - 2\sqrt{N}\ell^2 + 2N\ell\right) \ell^{2p-4}$$

Region II: $\sqrt{N} \leq \ell \leq 2\sqrt{N}$

$$i(\ell) = \frac{\alpha k}{6} \Gamma \left(2\sqrt{N} - \ell \right)^3 \ell^{2p-4} \quad (5.4)$$

Where ℓ is the length of interconnect in units of gate pitches, N is the number of logic gates, p is the Rent's exponent, α is the fraction of the on-chip terminals acting as a sink and depends on average fanout given by (5.5), and Γ is given by (5.6).

$$\alpha = \frac{FO}{FO + 1} \quad (5.5)$$

$$\Gamma = \frac{2N(1 - N^{p-1})}{\left(-N^p \frac{1+2p-2^{2p-1}}{p(2p-1)(p-1)(2p-3)} - \frac{1}{6p} + \frac{2\sqrt{N}}{2p-1} - \frac{N}{p-1} \right)} \quad (5.6)$$

5.2.3 Area Estimation

The total area of CNFET circuits given by (5.7), where a_i is the area of the minimum-size i^{th} gate, and x_i represents the optimized size of i^{th} gate at each iteration.

$$area = \sum_{i=0}^n a_i x_i \quad (5.7)$$

The effective width of each gate is the product of the assigned number of tubes to the gate and given CNT spacing. The product of calculated width and given length (technology) of the channel is defined as the area occupied by each transistor in the gate. The total circuit area is obtained by the summation of area occupied by each gate depending on the connectivity of each gate in the circuit.

Figure 5.4: The runtime and normalized area and delay product (ADP) after tube optimization for different size of OpenSPARC circuits.

5.3 EXPERIMENT RESULTS

In this section, we present the delay and the area optimization of CNFET circuits using developed optimization tool. We have used ISCAS-85 and OpenSPARC benchmark circuits for these experiments. The gate-level verilog netlist from ISCAS and OpenSPARC benchmark circuits are used for the validation of optimization tool and it is translated into MATLAB format. The MATLAB database stores the area and initial delay of individual gates, connectivity matrix required by LSGS optimization algorithm. We consider initial size is the maximum area of the gates to achieve minimum delay. For simplistic comparison with linear and non-linear models, we have only considered the gate count and hardware specifications. The intrinsic delay of each gate is estimated using our PALE and PAPF models. The wire-length for each circuit is computed using Rent's Rule. The Rent's coefficient, k , is a scaling constant, which is found to be experimentally cor-

Table 5.1: Total delay and area of ISCAS and OpenSPARC benchmark circuits with tube optimization and with tube + CLT both including runtime comparison of our Algorithm with linear model (LM) and non-linear model (NLM) on a single 2.6GHz processor with no parallelization

ISCAS											
Circuit	Area		Total Delay (ns)						Runtime (sec)		
			Ideal case (P=4nm)			Metallic tube variation					
	Actual (μm)	Optimal (μm)	w/o CLT	with CLT	Δ (%)	w/o CLT	with CLT	Δ (%)	Our Algorithm	LM	NLM
c17	86.0	48.224	0.061	0.058	- 4.2	0.063	0.059	-3.1	0.2	201.5	1.63e+4
c1355	2.97e+04	1.49e+04	17.51	13.29	-24.2	17.96	13.72	-23.7	12.2	530.7	5.64e+4
c3540	7.62e+04	3.81e+04	61.77	48.18	-22.1	63.06	49.63	-21.3	57.1	1068.2	1.22e+5
c5315	1.05e+05	5.72e+04	96.98	73.65	-24.1	99.57	75.93	-23.8	156	1504.9	1.75e+5
c7552	1.63e+05	8.20e+04	145.1	116.0	-22.1	149.8	119.2	-20.5	369	2055.2	2.42e+5
10k	2.15e+05	1.08e+05	220.4	178.0	-19.3	226.4	183.2	-22.4	522	2657.4	3.15e+5
OpenSPARC											
des3_area	8.12e+04	2.04e+04	96.54	93.14	-3.6	102.26	97.82	-4.4	185	1.44E+03	1.66E+05
aes_core	3.76e+05	9.40e+04	396.34	374.98	-5.5	409.78	379.3	-7.5	4510	4.99E+03	6.45E+05
wb_conmax	4.83e+05	1.21e+05	1076.12	879.59	-18.3	1152.26	888.89	-22.9	8380	1.15E+04	9.00E+05
ethernet	7.47e+05	2.87e+05	4027.5	2939.9	-27	4135.7	3205.6	-22.5	18300	2.60E+04	1.81E+06

responding to the average number of pins per module and typically varies from 3.5-4.7. The Rent's exponent, p , is the feature parameter of the circuit. In realistic 2D circuits, it ranges from 0.5 for highly-regular circuits (such as SRAM) to 0.75 for random logic. Since, our experiments are done on logic circuits, we chose $p=0.75$ as more appropriate value. We chose $FO = 2.8$. It is assumed that each output gate is driving a constant wire load determined based on the wire-length distribution of the circuit.

Table 5.1 presents the area and total delay of ISCAS and OpenSPARC benchmark circuits for ideal case (semiconducting tubes only) and with metallic tubes variations. The delay of gates includes its intrinsic delay, as well as the delay due to driving other gates or constant wire load. The introduction of CLT prior to tube optimization improves the delay in the circuits by up to 27% for both ideal and realistic scenarios. The increase in total delay due to the presence of 25% metallic tubes is around 3% on average. This suggests that the overall impact of tube removal on the performance of CNFET circuits

averages out due to removal of tubes from input and fan-out gates. Our algorithm also reduces the total area by more than 2X for all benchmark circuits. The significant reduction in total delay is due to substantial decline in the worst delay which can be attributed to unoptimized FO or stages in the critical paths.

We have run experiments using our optimization tool with PALE model for large OpenSPARC benchmark circuits with high gate count, more complexity and higher re-convergent fanout to demonstrate the scalability of the proposed algorithm as shown in Figure 5.4. It shows the runtime and normalized ADP with tube optimization for different size of OpenSPARC modules. The normalized ADP is the ratio of ADP with tube optimization to the ADP without any optimization. It is observed that the average normalized ADP of tested modules is around 0.27. Figure 5.5 shows the delay distribution of individual gates in 10k circuit with and without the CLT. The encircled region shows the delay of the critical gates exceeding 100ps and significantly impacting the worst and total delay in the circuit. Figure 5.6 shows the delay of 100 critical gates in the different benchmark circuits. It can be observed that the delay of critical gates reduced significantly with the CLT prior to tube optimization, which plays critical role in decreasing the total delay significantly. It can be observed that our algorithm helps to minimize the delay of these critical gates, which helps in reducing the overall delay. The results suggest that prior circuit optimization helps to achieve global optima, as oppose to tube optimization which may give only local optimum solution. It can be concluded that our optimization algorithm helps reducing the delay in critical gates, thereby improving the area and the performance of the circuit.

We compare the runtime of our algorithm with the runtime of optimization algorithm discussed in [31] using linear model (LM) and non-linear timing model (NLM) for CNFET circuit delay computation. The reported runtime in Table 5.1 for NLM and

Figure 5.5: Delay Distribution of Gates in 10k circuit.

LM for ISCAS and OpenSPARC benchmark circuits is estimated based on reported gate count, comparable optimization options and machine specification, which is reasonable assumption for runtime comparison. The delay of CNFET circuits in optimization framework [31] is estimated using Monte Carlo statistical timing analysis approach as described in [72]. However, Hills *et al.* [31] linearized the nonlinear timing model in [72] to analyze the impact of CNT variations on CNFET circuit delay variations. In our optimization algorithm, we use PAPF model for quick and fairly accurate estimation of delay in the presence of CNT variations. This reduces the overall runtime of our algorithm significantly compared to linearized and non-linear timing model. It must be noted that the objective function in [31] is to minimize energy under delay constraints. Whereas, in our cost function, the area of CNFET circuits is minimized to meet specific timing constraint. Our optimization tool shows significant runtime improvement over linear and non-linear timing models as shown in Table 5.1.

Figure 5.6: Delay of 100 critical gates for different benchmark circuits with and without CLT.

5.4 SUMMARY AND CONCLUSION

Applying circuit level techniques (CLT) prior to tube optimization in CNFET circuits help to reduce the delay in the benchmark circuits by up to 27% and area by 2.5X. The inclusion of wire parasitic load is essential for achieving optimal number of tubes in the CNFET gates, and for the true optimization of area and delay in the circuits. The use of fast and more accurate LE-based models for the estimation of delay in CNFET circuits helps in reducing the runtime of optimization algorithm significantly as compared to existing SPICE simulation and statistical methods. The developed framework can be successfully deployed in early design cycle and pre-layout stages of the design phase for more realistic evaluation and optimization of the CNFET-based VLSI circuits.

CHAPTER 6

CNT POSITION-AWARE YIELD ESTIMATION

The authors in [7], [55] and [71] developed current-based models and methodologies to predict functional yield for the CNFET-based designs. The functional yield of the gate depends on the number of tubes in the gates in the presence of the metallic tubes, count and density variations after metallic tubes removal. The approaches for yield estimation for CNFET-based standard cells and SRAM are discussed in [68] and [40] respectively. However, most of this work relies on the probability of the number of tubes required for a gate to be functional and the required number of tubes to meet certain delay without considering the impact of pitch and charge screening effect. Most of these approaches predicts yields for gate-level, and the performance of the circuits is not modeled. However, it is known that performance of the circuit depends on the factors like fanout, wire parasitics and load capacitance which are not modeled at gate-level. Hence, these approaches have some limitations. We develop probabilistic model to estimate functional Yield in presence of CNT variations. Our experiments suggest that the performance of a gate is greatly influenced by screening effect at smaller CNT Pitch (in particular it depends on the position of the tube in the CNTs array), which needs to be considered for accurate yield estimation. Also, the SPICE simulation-based techniques which depends on calculating I_{ON} and I_{OFF} currents for estimation of functional yield, have large runtime overhead, and may not be suitable for large CNFET circuits. As

we discussed in previous chapters, our LE based models are very fast in nature and address runtime bottlenecks at circuit level. The existing yield model based on regular probability is taken from [7] for comparison with our model which is extended to include conditional probability.

As we discussed previously, the delay and power consumption in CNFET-based gates and circuits increase significantly due to the presence of metallic tubes in a CNFET. The rise in the CNFET delay and power consumption depend on the percentage of metallic tube and may result in functional failures. The functional yield of CNFET logic gates depends on the gate topology and number of tubes (CNTs) remaining in the transistor after removal of metallic tubes. The functional yield of a logic gate is defined as the ratio of a number of functional gates to the total number of gates. The gates are functional, if the delay in gates is below the defined delay threshold in the presence of metallic tubes.

In this work, we propose more accurate probabilistic yield models which incorporates the screening effect due to removal of the metallic tubes with significant runtime benefit. The steps for analytical model development is discussed below:

- **Estimation of number of CNTs and minimum Pitch to meet the Delay Constraint:** The developed CNFET LE models are used to determine the desire tube and pitch configuration in the transistor to keep the delay below the delay threshold. This step generates a list of possible configuration with number of tubes and minimum desired pitch to meet the delay limit, for e.g., $(N_{t1}, P_{t1}) (N_{t2}, P_{t2}).....(N_m, P_m)$.
- **Probability of Functional Gates:** The probability of the functional gates is determined using the configurations obtained at the end of previous step. The probability of the transistors to be functional in pull-up and pull-down positions of a

logic gate while meeting the delay threshold is represented as P_{f_d}, P_{f_u} . The conditional probabilities depend on the tube and pitch configuration and type of gate. For certain number of tubes in the transistor, the position of the tubes removed is critical for gate to be functional. Hence, the conditional probability is necessary to account for the desired tube configuration.

$$P_{f_d} = P_{f_u} = \sum_{i=0}^{N_n} \left(\frac{Pr_i}{Pr_{i-1}} + \frac{Pr_i}{Pr_{i+1}} \right)^i \left(1 - \left(\frac{Pr_i}{Pr_{i-1}} + \frac{Pr_i}{Pr_{i+1}} \right) \right)^{N_{tur}-i} \frac{N_{tur}}{i}$$

- **Functional Yield of CNFET Logic Gates:** A functional gate requires both pull-up and pull-down network to be functional. The yield of the gate can be defined as the product of the probabilities of pull-up and pull-down network being functional.

$$Y_f = P_{f_d} * P_{f_u}$$

- **Functional Yield of CNFET-based Circuits:** The functional yield of a connection using CNFETs requires all the gates in the connection to be functional. Hence, the functional yield of a connection is represented as the product of yield of individual gates. The pseudo code of functional yield estimation algorithm is shown in table (Algorithm 2). Where, N_{path} represents path in a circuit in which probability of each gate instance needs to be calculated and N_{gates} represents gates with pull-up and pull-down network in a specific path of the circuit.

$$Y = \prod_{i=0}^{N_{gates}} \alpha_i Y_{fi}$$

Algorithm 2 Yield Estimation using developed Logical Effort models

```
1: Begin;
2: Define  $N_{CNFET}, P_{CNFET}, P_m$ 
3: Define  $N_{path}$ 
4: for  $i \leftarrow 1$  to  $N_{path}$  do
5:   for  $j \leftarrow 1$  to  $N_{gates}$  do
6:     Using developed LE Model to calculate desired tubes ( $N_{tur}$ ) and pitch ( $P_{cnt}$ )
7:     Calculate Probability:
8:      $P_{f\_d} = P_{f\_u} = \sum_{i=0}^{N_n} \left( \frac{Pr_i}{Pr_{i-1}} + \frac{Pr_i}{Pr_{i+1}} \right)^i \left( 1 - \left( \frac{Pr_i}{Pr_{i-1}} + \frac{Pr_i}{Pr_{i+1}} \right) \right)^{N_{tur}-i} N_{tur}$ 
9:     Calculate Yield:
10:     $Y_{gate} = \prod_{i=0}^{N_{gates}} Pr_{gate}$ 
11:     $j++$ ;
12:   end;
13:   Calcualte Circuit Yield:
14:    $Y_{circuit} = \prod_{i=0}^{N_{gates}} Y_{gate}$ 
15: end;
16: end;
```

6.1 FUNCTIONAL YIELD OF INVERTER

In this section, the functional yield of inverter gate is estimated using proposed probabilistic yield model. The analysis is based on the following assumptions:

- The ideal CNFET has 8 tubes with 4nm pitch. The screening effect is dominant at smaller pitch.
- The removal process only removes metallic tube and no semiconducting tubes are removed.
- The delay limit is defined as 1.2x of the ideal gate delay.

Below are two main steps for estimating overall yield.

Step 1: Estimating the number of CNTs and minimum desired Pitch: The developed CNFET

Logical Effort models are used to determine the desired tube and pitch configuration to meet the delay threshold. Table 6.1 reports the desired number of tubes and minimum pitch using existing and proposed models. It can be clearly seen, that proposed model suggests more number of configurations passing delay criteria. Hence, for 6 tubes or above, yield is independent of the pitch and associated screening effect. However, for number of tubes below 6, the screening effect goes down due to removal of tubes. Due to lower screening effect, the delay of transistor remains below threshold in spite of only 4 tubes.

Step 2: Probabilities of Pull-up and Pull-down Network to be Functional: The probability of pull-up and pull-down network to be functional is same and computed using proposed configurations in previous step. As it can be seen in Table 6.1 that the for at least 6 tubes at 4nm technology, the probabilities of gate to be functional can be calculated using existing models. However, for the number of tubes below 6, the conditional probability is more suitable to account for pitch and position of tubes.

Table 6.1: Comparison of number of CNTs and pitch desired to meet Delay limits

Existing model		Proposed model	
N_{tl}	P_{tl}	N_{tl}	P_{tl}
8	4nm	8	4nm
7	4nm	7	4nm
6	4nm	6	4nm
-	-	5	6nm
-	-	4	8nm

The probability of 6 tubes or more in transistor is given by:

$$P_{f_d} = P_{f_u} = \sum_{i=6}^8 (Pr_i)^i (1 - Pr_i)^{8-i}$$

The probability of tubes below 6 and desired pitch is given by:

$$P_{f_d} = P_{f_u} = \sum_{i=0}^6 \left(\frac{Pr_i}{Pr_{i-1}} + \frac{Pr_i}{Pr_{i+1}} \right)^i \left(1 - \left(\frac{Pr_i}{Pr_{i-1}} + \frac{Pr_i}{Pr_{i+1}} \right) \right)^{8-i}$$

Step 3: Yield of the CNFET gate: The yield of the gate is defined by the product of pull-down and pull-up probabilities.

$$Y_f = P_{f_d} * P_{f_u}$$

6.2 FUNCTIONAL YIELD OF 3-STAGE INVERTER:

The functional yield of a circuit depends on the yield of individual gates in the circuit. The yield of a 3-stage inverter requires each gate to be functional in order to achieve a functional circuit. Once the yield of a logic gate for ideal number of tubes and pitch is estimated using steps discussed in section 6.1, the yield of 3-stage inverter can be calculated using following Eq.

$$Y_f = Y_{f1} * Y_{f2} * Y_{f3}$$

6.3 EXPERIMENTAL RESULTS

The reference data from Monte Carlo simulations for validating the probabilistic model is generated similar to section 4.3. We have used the probabilistic model proposed by

[7] for estimating the yield for pitch greater than 6nm. At smaller pitch, where the screening effect is significant, we have developed the conditional probabilistic model. The functional yield of logic gates and circuits using proposed Pitch-Aware LE (PALE) model is discussed here. The probabilistic model is validated using Monte Carlo (MC) simulations on sample size of 1000. The functional yield for different CNT pitch values is discussed below. Since, the screening effect is dominant at smaller CNT pitch, the experiments are performed for pitch below 6nm.

6.3.1 Analysis of single Inverter Gate

We assume the ideal transistor has 8 tubes. The percentage of metallic tubes (P_m) present in the circuit is 10%. The delay threshold required for the gate to be functional is 1.2X of the ideal delay. The functional yield of the inverter for given tubes, pitch and percentage of metallic tube is estimated using following steps.

Step 1: Estimation of Tubes and Pitch: We use PALE model to estimate the number of tubes and pitch to meet the delay threshold. Tables 6.2 and 6.3 show the number of tubes and desired pitch for 4nm and 2nm respectively. As shown in these tables, the screening effect based on actual positions of CNTs helps to qualify more configurations as functional gate.

Step 2: Probabilities of Pull-up and Pull-down Network to be Functional: The probability for networks being functional using existing and functional models is shown below:

Table 6.2: Comparison of desired tubes and pitch for functional inverter using existing and proposed models for CNT pitch 4nm

Existing model		Proposed model	
N_{tl}	P_{tl}	N_{tl}	P_{tl}
8	4nm	8	4nm
7	4nm	7	4nm
6	4nm	6	4nm
-	-	5	6nm
-	-	4	8nm

Table 6.3: Comparison of desired tubes and pitch for functional inverter using existing and proposed models for CNT pitch 2nm

Existing model		Proposed model	
N_{tl}	P_{tl}	N_{tl}	P_{tl}
8	2nm	8	2nm
7	2nm	7	2nm
6	2nm	6	2nm
-	-	5	3nm
-	-	4	4nm
-	-	3	5nm
-	-	2	8nm

$$P_{f_d}(Existing_model) = 0.962$$

$$P_{f_d}(Proposed_model) = 0.978$$

Step 3: Yield of the CNFET gate: The yield of the inverter gate is obtained by the product of the pull-up and pull-down probabilities for 4nm pitch.

$$Y_f(Existing_model) = 92.5\%$$

$$Y_f(Proposed_model) = 95.6\%$$

For 2nm pitch, the probabilities of the pull-up and pull-down networks are different based on the proposed model. However, the existing model gives the same yield as at 4nm pitch. The probability of gate being functional and yield at 2nm pitch using configuration in Table 6.4 using existing and proposed models is given below.

$$P_{f_d}(Proposed_model) = 0.989$$

$$Y_f(Proposed_model) = 98.1\%$$

6.3.2 Analysis of 2-stage Inverter Gate

The yield of the 2-stage inverter is obtained by the product of functional yield of individual inverter. Hence, the functional yield of 2-stage inverter chain using existing and proposed models at 4nm pitch is given below.

$$Y_{f_circuit}(Existing_model) = 85.6\%$$

$$Y_{f_circuit}(Proposed_model) = 91.3\%$$

Figure 6.1: Comparison of existing yield model and proposed model with MC simulations.

The yield of 2-stage inverter at 2nm CNT pitch is given by Eq. (9)

$$Y_f(Proposed_model) = 96.2\%$$

The comparison of the proposed yield model with MC simulations is shown in Figure 6.1 for different gates and circuits. The existing yield model shows a difference of 6% and 11% with MC simulations at gate-level and circuit level respectively. The proposed model predicts yield within 1% difference with MC simulations.

The accuracy and runtime of predicting yield using proposed model is compared with existing model and MC simulations in Table 6.4. The error in yield prediction using existing model is up to 10%, while the proposed model predicts yield with error less than 1%. Also, the proposed model has runtime 1000x better than MC simulations.

Table 6.4: Comparison of accuracy and runtime using existing and proposed models

Circuit	Error (%)		Runtime (sec)		
	Existing	Proposed	Existing	Proposed	Monte Carlo
	Model	Model	Model	Model	Simulation
INV(4nm)	3.334	0.094	1.5	1.6	716
INV(2nm)	6.377	0.709	1.5	1.6	744
2-stage INV (4nm)	5.582	0.705	2.0	2.1	3980
2-stage INV (2nm)	10.686	0.455	2.0	2.1	3980
2-stage decoder (4nm)	15.42	1.07	1.2	1.2	1750

6.4 SUMMARY AND CONCLUSION

Our experiments show that CNT position to account for accurate charge screening at smaller pitch is important for yield estimation. Our proposed model is correlated within 1% at small pitch values for CNFET gate and circuit level with Monte Carlo simulation methods and it is 1000X better in runtime.

CHAPTER 7

CONTRIBUTIONS, CONCLUSIONS & FUTURE WORK

Scaling transistors and following Moore's law have served the Silicon industry well for over five decades in providing integrated circuits that are denser, cheaper, higher performance, and lower power. However, Si based integrated circuits technology is approaching physical and manufacturing limits. The reliability challenges, process variations and short channel effects are also becoming prominent.

On top of that, Novel computing paradigms and applications introduce higher performance and efficiency requirements. The academic and industry are focusing more on Beyond-CMOS technologies, which may provide the required devices, processes, and architectures for the new era of computing. CNFETs have been identified as a potential candidate for future integrated circuits due to superior electrostatic properties, ballistic transport of charge carriers and other key advantages. The imperfections in CNT growth (cause of CNT variations) and CNFET specific parameters impact performance, power and yield of CNFET-based circuits. In this work, we have analyzed the impact of CNFET specific parameters and CNT variations on the drive strength and delay for different CNFET gates and circuits configurations. Moreover, models are developed to do accurate performance evaluation, optimization and yield estimation for large CNFET-based circuits.

7.1 CONTRIBUTIONS AND CONCLUSIONS

This thesis has demonstrated the development of accurate models and a framework for delay and area optimization of CNFET-based VLSI circuits considering the impact of CNFET specific parameters and CNT variations. These parameters like CNT pitch, tubes, CNTs density and metallic tube variations are critical in determining the performance, power and area (PPA) of CNFET-based VLSI circuits. Therefore, these parameters and variations must not be ignored for true performance optimization of CNFET circuits. Applying circuit level techniques (CLT) prior to tube optimization in CNFET circuits, help in achieving global optimum solution and reducing the delay significantly in the benchmark circuits.

In chapter 3, the impact of charge screening effect on the performance of CNFET-based circuits has been studied and CMOS based Logical Effort model is enhanced to account for CNFET specific parameters. The developed PALE model results in fast and accurate performance evaluation of CNFET gates and circuits with average error of 5% as compared to SPICE simulations.

We have analyzed the impact of CNT variations and metallic tube removal on the gate capacitance and delay in the CNFET gates and circuits for different CNT removal techniques using Monte Carlo simulations in chapter 4. The closed-form Position-Aware Pitch-Factor (PAPF) model for accurate performance evaluation of large CNFET-based circuits in the presence of imperfection is developed. The difference in delay calculation for CNFET-based gates and circuits estimated from PAPF model is within 2% to 5% range as compared to SPICE simulation and it works well for P_m from 0% to 25%.

In chapter 5, we have demonstrated an optimization tool using PALE and PAPF models based on LSGS algorithm by incorporating CNFET-specific parameters and CNTs

count variations, to minimize the area and delay product (ADP) of CNFET circuits. The proposed circuit-level techniques (CLT); FO optimization and sizing of gates, help to reduce the delay in the benchmark circuits by up to 27% and area by 2.5X. The addition of wire load, PALE and PAPF models to account for CNFET specific features and variations improve the overall accuracy of optimization framework.

In chapter 6, we propose more accurate probabilistic yield model which incorporates the screening effect due to removal of the metallic tubes with significant runtime benefit. The existing yield estimation approach relies on the probability of the number of tubes required for a gate to be functional and the required number of tubes to meet certain delay without considering the impact of pitch and charge screening effect. Our proposed yield model is validated with Monte Carlo simulations with 1% error at small pitch values for CNFET gates and circuits and it shows 1000X runtime improvement.

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

- Analysis of the impact of CNFET specific parameters and CNT variations on the drive strength and delay for different CNFET configurations; (width, number of CNTs and spacing between them), the effects due to these parameters and variations are captured accurately in simple empirical closed-form developed models which enable to do performance evaluation, optimization and yield estimation for large CNFET-based circuits
- For accurate and fast delay estimation for ideal case (without any CNT variations), Pitch-Aware Logical Effort (PALE) model is developed by incorporating the impact of CNFET specific parameters for CNFET-based circuits, including the development of the reference inverter and the technology parameter for CNFET.
- Development of closed-form Position-Aware Logical Effort (PAPF) model for per-

formance evaluation of large CNFET-based circuits in the presence of CNT count and density variations (different number of tubes with non-uniform spacing and position in CNFET channel).

- ADP (Area and Delay Product) optimization tool is developed, by incorporating circuit level techniques (FO optimization and gate sizing), wire load, PALE and PAPF developed LE models to incorporate CNFET-specific parameters and CNT variations in ADP optimization.
- The probabilistic model for better functional yield estimation is presented for CNFET gates and circuits by comprehending the impact due to statistical removal of CNTs including influence of position of removal tubes and its effect on charge screening at smaller pitch.

7.1.1 Conclusions

- CNFET gates with same width, can have significantly different delay and functional yield depending on the number of CNTs and spacing between them. The increase in CNT count at smaller pitch may not always help to improve CNFET performance due to increase in charge screening between CNTs.
- We showed that the removal of tubes due to the presence of metallic tubes may not impact the performance of the circuit. The tube removal from specific positions for smaller CNT pitch (6nm and below) reduces the screening effect significantly that compensates for the lower tube count.
- Area and Delay Product (ADP) results show that the CNFET-specific parameters and CNT variations impact the ADP optimization of CNFET circuits significantly

in early design phase and cannot be ignored for accurate optimization.

- Our experiments show that CNT position to account for accurate charge screening at smaller pitch is important for functional yield estimation.

7.2 FUTURE WORK

CNT technology and CNFET-based circuits are still evolving and there is room to improve and develop new techniques and models, which is important to realize this technology for real commercial products in future. Below is list of proposed future works that can be critical for CNFET circuits developments:

- The power analysis is very important in overall circuit level designs. Hence, the addition of power component to optimization framework in conjunction with performance and area for more comprehensive analysis will be very helpful for circuit designers in early design stage.
- For industry EDA tools, there is need to have technology file for CNFETs. There is need to develop accurate technology file for EDA tools by including the impact of CNFET specific parameters and CNT variations. Our developed PALE and PAPF models can be used to develop accurate technology files for EDA tools usage.
- The long simulation runtime for large circuits is always a concern. The parallelism features can be incorporated using distributed machine processing approaches to further improve runtime of our tool.
- Due to limitations in the SPICE models beyond 32nm, the validation of PALE and PAPF models is done for 32nm technology node only. Once accurate SPICE mod-

els are available, experimental validation of the optimization tool can be done to ensure compatibility with different technology nodes beyond 32nm.

Bibliography

- [1] [Online] Available, International Roadmap for Devices and Systems. <https://irds.ieee.org>, [Online]. [Accessed: 11-April-2018].
- [2] [Online] Available, Stanford engineers build first computer based on carbon nanotube technology; imperfections-immune design. <https://www.greencarcongress.com/2013/09/20130926-cnt.html>, [Online]. [Accessed: 26-September-2013].
- [3] M. Ali, M. Abrar Ahmed, and M. Chrzanowska-Jeske. Logical Effort Framework for CNFET-Based VLSI Circuits for Delay and Area Optimization. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 27(3):573–586, March 2019.
- [4] M. Ali, M. Ahmed, M. Chrzanowska-Jeske, and J. Morris. Logical Effort model for CNFET circuits with CNTs variations. In *2015 IEEE 15th International Conference on Nanotechnology (IEEE-NANO)*, pages 1218–1221, July 2015.
- [5] M. Ali, R. Ashraf, and M. Chrzanowska-Jeske. Logical Effort of CNFET-based Circuits in the Presence of Metallic Tubes. In *2012 12th IEEE International Conference on Nanotechnology (IEEE-NANO)*, pages 1–6, Aug 2012.
- [6] M. Ali, M. Ahmed, and M. Chrzanowska-Jeske. Fast and Accurate Evaluation of

- delay in CNFET Circuits. In *2016 IEEE 16th International Conference on Nanotechnology (IEEE-NANO)*, pages 659–662, Aug 2016.
- [7] R. Ashraf, M. Chrzanowska-Jeske, and S. G. Narendra. Functional Yield Estimation of Carbon Nanotube-Based Logic Gates in the Presence of Defects. *IEEE Transactions on Nanotechnology*, 9(6):687–700, Nov 2010.
 - [8] A. A. Bayrakci. Stochastic Logical Effort As a Variation Aware Delay Model to Estimate Timing Yield. *Integr. VLSI J.*, 48(C):101–108, Jan. 2015.
 - [9] S. Bobba, J. Zhang, A. Pullini, D. Atienza, and G. D. Micheli. Design of compact imperfection-immune CNFET layouts for standard-cell-based logic synthesis. In *2009 Design, Automation Test in Europe Conference Exhibition*, pages 616–621, April 2009.
 - [10] G. J. Brady, Y. Joo, M.-Y. Wu, M. J. Shea, P. Gopalan, and M. S. Arnold. Polyfluorene-Sorted, Carbon Nanotube Array Field-Effect Transistors with Increased Current Density and High On/Off Ratio. *ACS Nano*, 8(11):11614–11621, 2014. PMID: 25383880.
 - [11] Q. Cao, S.-J. Han, J. Tersoff, A. D. Franklin, Y. Zhu, Z. Zhang, G. S. Tulevski, J. Tang, and W. Haensch. End-bonded contacts for carbon nanotube transistors with low, size-independent resistance. *Science*, 350(6256):68–72, 2015.
 - [12] Q. Cao, J. Tersoff, S.-J. Han, and A. V. Penumatcha. Scaling of Device Variability and Subthreshold Swing in Ballistic Carbon Nanotube Transistors. *Phys. Rev. Applied*, 4:024022, Aug 2015.

- [13] G. Cho, Y. B. Kim, and F. Lombardi. Modeling Undeposited CNTs for CNTFET Operation. *IEEE Transactions on Device and Materials Reliability*, 11(2):263–272, June 2011.
- [14] G. Cho and F. Lombardi. On the Delay of a CNTFET with Undeposited CNTs by Gate Width Adjustment. *J. Electron. Test.*, 29(3):261–273, June 2013.
- [15] G. Cho and F. Lombardi. Circuit-Level Simulation of a CNTFET With Unevenly Positioned CNTs by Linear Programming. *IEEE Transactions on Device and Materials Reliability*, 14(1):234–244, March 2014.
- [16] M. Chrzanowska-Jeske, R. Ashraf, R. K. Nain, and S. G. Narendra. Performance Analysis of CNFET based Circuits in the Presence of Fabrication Imperfections. In *2012 IEEE International Symposium on Circuits and Systems*, pages 1363–1366, May 2012.
- [17] P. G. Collins, M. S. Arnold, and P. Avouris. Engineering Carbon Nanotubes and Nanotube Circuits Using Electrical Breakdown. *Science*, 292(5517):706–709, 2001.
- [18] J. A. Davis, V. K. De, and J. D. Meindl. A stochastic wire-length distribution for gigascale integration (GSI). I. Derivation and validation. *IEEE Transactions on Electron Devices*, 45(3):580–589, Mar 1998.
- [19] A. Demir and S. Tasiran. Stochastic Logical Effort: Designing for Timing Yield on the Back of an Envelope. *ACM/IEEE International Workshop (TAU)*, 2006.
- [20] J. Deng and H. S. P. 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*, 54(12):3186–3194, Dec 2007.
- [21] J. Deng and H. S. P. Wong. Modeling and Analysis of Planar-Gate Electrostatic Capacitance of 1-D FET With Multiple Cylindrical Conducting Channels. *IEEE Transactions on Electron Devices*, 54(9):2377–2385, Sept 2007.
- [22] L. Ding, A. Tselev, J. Wang, D. Yuan, H. Chu, T. P. McNicholas, Y. Li, and J. Liu. Selective Growth of Well-Aligned Semiconducting Single-Walled Carbon Nanotubes. *Nano Letters*, 9(2):800–805, 2009. PMID: 19159186.
- [23] L. Ding, D. Yuan, and J. Liu. Growth of High-Density Parallel Arrays of Long Single-Walled Carbon Nanotubes on Quartz Substrates. *Journal of the American Chemical Society*, 130(16):5428–5429, 2008. PMID: 18376825.
- [24] L. Ding, Z. Zhang, S. Liang, T. Pei, S. Wang, Y. Li, W. Zhou, J. Liu, and L.-M. Peng. CMOS-based carbon nanotube pass-transistor logic integrated circuits. *Nat. Commun*, 3(677), 2012.
- [25] A. D. Franklin, D. B. Farmer, and W. Haensch. Defining and Overcoming the Contact Resistance Challenge in Scaled Carbon Nanotube Transistors. *ACS Nano*, 8(7):7333–7339, 2014. PMID: 24999536.
- [26] A. D. Franklin, S. O. Koswatta, D. B. Farmer, J. T. Smith, L. Gignac, C. M. Breslin, S.-J. Han, G. S. Tulevski, H. Miyazoe, W. Haensch, and J. Tersoff. Carbon Nanotube Complementary Wrap-Gate Transistors. *Nano Letters*, 13(6):2490–2495, 2013. PMID: 23638708.

- [27] A. D. Franklin, M. Luisier, S.-J. Han, G. Tulevski, C. M. Breslin, L. Gignac, M. S. Lundstrom, and W. Haensch. Sub-10 nm Carbon Nanotube Transistor. *Nano Letters*, 12(2):758–762, 2012. PMID: 22260387.
- [28] B. Ghavami and M. Raji. Failure Characterization of Carbon Nanotube FETs Under Process Variations: Technology Scaling Issues. *IEEE Transactions on Device and Materials Reliability*, 16(2):164–171, June 2016.
- [29] B. Ghavami, M. Raji, H. Pedram, and M. Pedram. Statistical Functional Yield Estimation and Enhancement of CNFET-Based VLSI Circuits. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 21(5):887–900, May 2013.
- [30] B. Ghavami, M. Raji, H. Pedram, and M. Pedram. Statistical Functional Yield Estimation and Enhancement of CNFET-Based VLSI Circuits. *IEEE Trans. VLSI Syst.*, 21(5):887–900, 2013.
- [31] G. Hills, J. Zhang, M. M. Shulaker, H. Wei, C. S. Lee, A. Balasingam, H. S. P. Wong, and S. Mitra. Rapid Co-Optimization of Processing and Circuit Design to Overcome Carbon Nanotube Variations. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 34(7):1082–1095, July 2015.
- [32] A. Imran, M. Hasan, A. Islam, and S. A. Abbasi. Optimized Design of a 32-nm CNFET-Based Low-Power Ultrawideband CCII. *IEEE Transactions on Nanotechnology*, 11(6):1100–1109, Nov 2012.
- [33] S. Joshi and S. Boyd. An Efficient Method for Large-Scale Gate Sizing. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 55(9):2760–2773, Oct 2008.

- [34] A. Kabbani. Logical effort based dynamic power estimation and optimization of static CMOS circuits. *Integration, the VLSI Journal*, 43:279–288, 06 2010.
- [35] A. Kabbani, D. Al-Khalili, and A. J. Al-Khalili. Delay analysis of CMOS gates using modified logical effort model. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 24(6):937–947, June 2005.
- [36] S. J. Kang, C. Kocabas, T. Ozel, M. Shim, N. Pimparkar, M. A. Alam, S. V. Rotkin, and J. A. Rogers. High-performance electronics using dense, perfectly aligned arrays of single-walled carbon nanotubes. *Nature Nanotechnology*, 2:230–236, 2007.
- [37] Y. B. Kim, Y. B. Kim, and F. Lombardi. A novel design methodology to optimize the speed and power of the CNTFET circuits. In *2009 52nd IEEE International Midwest Symposium on Circuits and Systems*, pages 1130–1133, Aug 2009.
- [38] C. Kocabas, S. J. Kang, T. Ozel, M. Shim, and J. A. Rogers. Improved Synthesis of Aligned Arrays of Single-Walled Carbon Nanotubes and Their Implementation in Thin Film Type Transistors. *The Journal of Physical Chemistry C*, 111(48):17879–17886, 2007.
- [39] K. J. Kuhn. Moore’s Law Past 32nm: Future Challenges in Device Scaling. In *2009 13th International Workshop on Computational Electronics*, pages 1–6, May 2009.
- [40] T. Li, L. Jiang, X. Liang, Q. Xu, and K. Chakrabarty. Defect tolerance for CNFET-based SRAMs. In *2016 IEEE International Test Conference (ITC)*, pages 1–9, Nov 2016.
- [41] Y. Li, D. Mann, M. Rolandi, W. Kim, A. Ural, S. Hung, A. Javey, J. Cao, D. Wang, E. Yenilmez, Q. Wang, J. F. Gibbons, Y. Nishi, and H. Dai. Preferential Growth

- of Semiconducting Single-Walled Carbon Nanotubes by a Plasma Enhanced CVD Method. *Nano Letters*, 4(2):317–321, 2004.
- [42] X. Liu, T. Pichler, M. Knupfer, M. Golden, J. Fink, H. Kataura, and Y. Achiba. Detailed analysis of the mean diameter and diameter distribution of single wall carbon nanotubes from their optical response. *Physical Review B*, 66(4):45411, 2002.
- [43] A. Morgenshtein, E. G. Friedman, R. Ginosar, and A. Kolodny. Unified Logical Effort - A Method for Delay Evaluation and Minimization in Logic Paths With RC Interconnect. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 18(5):689–696, May 2010.
- [44] R. S. Park, M. M. Shulaker, G. Hills, L. Suriyasena Liyanage, S. Lee, A. Tang, S. Mitra, and H.-S. P. Wong. Hysteresis in Carbon Nanotube Transistors: Measurement and Analysis of Trap Density, Energy Level, and Spatial Distribution. *ACS Nano*, 10(4):4599–4608, 2016. PMID: 27002483.
- [45] N. Patil, J. Deng, A. Lin, H. S. P. 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*, 27(10):1725–1736, Oct 2008.
- [46] N. Patil, J. Deng, S. Mitra, and H. S. P. Wong. Circuit-level performance benchmarking and scalability analysis of carbon nanotube transistor circuits. *IEEE Transactions on Nanotechnology*, 8(1):37–45, Jan 2009.
- [47] N. Patil, A. Lin, E. R. Myers, K. Ryu, A. Badmaev, C. Zhou, H. . P. Wong, and S. Mitra. Wafer-Scale Growth and Transfer of Aligned Single-Walled Carbon Nanotubes. *IEEE Transactions on Nanotechnology*, 8(4):498–504, July 2009.

- [48] N. Patil, A. Lin, J. Zhang, H. Wei, K. Anderson, H. S. P. Wong, and S. Mitra. VMR: VLSI-compatible metallic carbon nanotube removal for imperfection-immune cascaded multi-stage digital logic circuits using Carbon Nanotube FETs. In *2009 IEEE International Electron Devices Meeting (IEDM)*, pages 1–4, Dec 2009.
- [49] B. C. Paul, S. Fujita, M. Okajima, T. Lee, H. S. P. Wong, and Y. Nishi. Impact of Process Variation on Nanowire and Nanotube Device Performance. In *2007 65th Annual Device Research Conference*, pages 269–270, June 2007.
- [50] A. Raychowdhury, V. K. De, J. Kurtin, S. Y. Borkar, K. Roy, and A. Keshavarzi. Variation Tolerance in a Multichannel Carbon-Nanotube Transistor for High-Speed Digital Circuits. *IEEE Transactions on Electron Devices*, 56(3):383–392, March 2009.
- [51] P. Rezvani and M. Pedram. A fanout optimization algorithm based on the effort delay model. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 22(12):1671–1678, Dec 2003.
- [52] K. Ryu, A. Badmaev, C. Wang, A. Lin, N. Patil, L. Gomez, A. Kumar, S. Mitra, H.-S. P. Wong, and C. Zhou. CMOS-Analogous Wafer-Scale Nanotube-on-Insulator Approach for Submicrometer Devices and Integrated Circuits Using Aligned Nanotubes. *Nano Letters*, 9(1):189–197, 2009. PMID: 19086836.
- [53] R. Saito, G. Dresselhaus, and M. Dresselhaus. Physical Properties of Carbon Nanotubes. *Imperial C. Press*, 1998.
- [54] S. I. Sayed, M. M. Abutaleb, and Z. B. Nossair. Optimization of CNFET Parameters for High Performance Digital Circuits. *Advances in Materials Science and Engineering*, 2016.

- [55] K. Sheikh, S. Han, and L. Wei. CNFET With Process Imperfection: Impact on Circuit-Level Yield and Device Optimization. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 63(12):2209–2221, Dec 2016.
- [56] K. Shinkai, M. Hashimoto, A. Kurokawa, and T. Onoye. A Gate Delay Model Focusing on Current Fluctuation over Wide-Range of Process and Environmental Variability. In *2006 IEEE/ACM International Conference on Computer Aided Design*, pages 47–53, Nov 2006.
- [57] M. M. Shulaker, G. Hills, N. Patil, H. Wei, H.-Y. Chen, H.-S. P. Wong, and S. Mitra. Carbon nanotube computer. *Nature*, 501:526–530, 2013.
- [58] M. Steiner, M. Engel, Y. M. Lin, K. J. Y. Q. Wu, D. B. Farmer, J. J. Humes, N. L. Yoder, J. W. T. Seo, A. A. Green, M. C. Hersam, R. Krupke, and P. Avouris. High-frequency performance of scaled carbon nanotube array field-effect transistors. *Appl. Phys. Lett*, 101:053123, 2012.
- [59] Y. Sun and V. Kursun. N-Type Carbon-Nanotube MOSFET Device Profile Optimization for Very Large Scale Integration. *Transactions on Electrical and Electronic Materials*, 2(2), Apr 2011.
- [60] L. Suriyasena Liyanage, X. Xu, G. Pitner, Z. Bao, and H.-S. P. Wong. VLSI-Compatible Carbon Nanotube Doping Technique with Low Work-Function Metal Oxides. *Nano Letters*, 14(4):1884–1890, 2014. PMID: 24628497.
- [61] I. Sutherland, B. Sproull, and D. Harris. Logical Effort: Designing Fast CMOS Circuits. *Morgan Kaufmann*, 1999.

- [62] I. E. Sutherland and R. F. Sproull. Logical effort: Designing for speed on the back of an envelope. *in Proc. Univ. California/Santa Cruz Conf. ARVLSI*, pages 1–16, 1991.
- [63] S.Joshi and S. Boyd. lsgs large-scale gatesizing MATLAB toolbox, version 0.25. March 2007.
- [64] G. S. Tulevski, A. D. Franklin, D. Frank, J. M. Lobe, Q. Cao, H. Park, A. Afzali, S.-J. Han, J. B. Hannon, and W. Haensch. Toward high-performance digital logic technology with carbon nanotubes. *ACS Nano*, 8(9):8730–8745, 2014. PMID: 25144443.
- [65] M. Ali, M. Ahmed, and M. Chrzanowska-Jeske. Stochastic Analysis of CNFET circuits using Enhanced Logical Effort Model in the Presence of Metallic Tubes. In *2014 21st IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, pages 774–777, Dec 2014.
- [66] J. Deng and H. S. P. 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 Transactions on Electron Devices*, 54(12):3195–3205, Dec 2007.
- [67] K. J. Kuhn. Reducing Variation in Advanced Logic Technologies: Approaches to Process and Design for Manufacturability of Nanoscale CMOS. In *2007 IEEE International Electron Devices Meeting*, pages 471–474, Dec 2007.
- [68] C. Wang, L. Jiang, S. Hu, T. Li, X. Liang, N. Jing, and W. Qian. Timing-driven placement for carbon nanotube circuits. In *2015 28th IEEE International System-on-Chip Conference (SOCC)*, pages 362–367, Sept 2015.

- [69] G. Zhang, P. Qi, X. Wang, Y. Lu, X. Li, R. Tu, S. Bangsaruntip, D. Mann, L. Zhang, and H. Dai. Selective Etching of Metallic Carbon Nanotubes by Gas-Phase Reaction. *Science*, 314(5801):974–977, 2006.
- [70] J. Zhang, A. Lin, N. Patil, H. Wei, L. Wei, H. . P. Wong, and S. Mitra. Carbon Nanotube Robust Digital VLSI. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 31(4):453–471, April 2012.
- [71] J. Zhang, N. Patil, A. Lin, H.-S. P. Wong, and S. Mitra. Carbon Nanotube Circuits: Living with Imperfections and Variations. In *Proceedings of the Conference on Design, Automation and Test in Europe*, DATE '10, pages 1159–1164, 3001 Leuven, Belgium, Belgium, 2010. European Design and Automation Association.
- [72] J. Zhang, N. Patil, H. S. P. Wong, and S. Mitra. Overcoming carbon nanotube variations through co-optimized technology and circuit design. In *2011 International Electron Devices Meeting*, pages 4.6.1–4.6.4, Dec 2011.

APPENDIX

GLOSSARY

A.1 NOMENCLATURE

η	Ratio of gate-to-channel capacitance between edge and middle tubes
a_{min}	Minimum area
B	Branching Effort
C_{gc}	Gate-to-channel capacitance
C_{gc_e}	Gate-to-channel capacitance of edge tubes
C_{gc_m}	Gate-to-channel capacitance of middle tubes
C_{gtg}	Gate-to-gate capacitance
C_{inf}	Coupling capacitance between electrode and tube
CMOS	Complementary Metal Oxide Semiconductor
C_{of}	Outer-fringe capacitance
C_{sr}	Equivalent capacitance due to Screening Effect
c_i^{int}	Internal capacitance
c_i^{in}	Input capacitance
D	Path Delay
d_{cn}	Diameter of CNT
d_{norm}	Normalized Delay

d_{min}	Minimum Delay
F	Path Effort
G	Logical Effort of all the gates in the path
g	Logical Effort of the gate
g_{ref}	Logical Effort of Reference Inverter
Ge	Germanium
H	Electrical Effort of the path
h	Electrical Effort
I_{ON}	On current
I_{OFF}	Off current
k	Wire load of each gate
L_{ch}	Percentage of tubes removal
$m - CNT$	Metallic CNT
N_{CNEFT}	Number of CNTs for yield estimation
N_{gates}	Represents gates in a specific path of the circuit
N_{path}	Represents path in a circuit
N_{ref}	Number of tubes in the reference gate
N_{tur}	Number of tubes in the actual gate
p	Parasitic Delay
P_{cnt}	Pitch of CNTs
P_{CNEFT}	Minimum desired pitch
P_{f_u}	Probability of the transistors in pull-up network
P_{f_d}	Probability of the transistors in pull-down network
P_m	Percentage of metallic tubes
P_r	Percentage of tubes removal

r_i	Driving resistance of gate i
$s - CNT$	Semiconducting CNT
Si	Silicon
$SiGe$	Silicon Germanium
T	Timing specification vector
t_p	Propagation Delay
τ_0	Intrinsic delay of a Reference Inverter
x	Initial number of tubes
$Y_{circuit}$	Functional Yield of the circuit
Y_f	Functional Yield
Y_{gates}	Functional Yield of the gate

A.2 ACRONYMS

ADP	Area-Delay Product
AFM	Atomic Force Microscope
CNT	Carbon Nanotube
$CNFET$	Carbon Nanotube Field-Effect Transistor
$ESLE$	Evenly-Spaced Logical Effort
FET	Field-Effect Transistors
FO	Fanout
$FO4$	Fanout of 4
GAA	Gate-all-around
$IRDS$	International Roadmap for Devices and Systems

<i>ISCAS</i>	International Symposium on Circuits and Systems
<i>LE</i>	Logical Effort
<i>LM</i>	Linear Model
<i>LSGS</i>	Large-Scale Gate Sizing
<i>LUT</i>	Look-up-table
<i>MC</i>	Monte Carlo
<i>NLM</i>	Non-linear Model
<i>PALE</i>	Pitch-Aware Logical Effort
<i>PAPF</i>	Position-Aware Pitch Factor
<i>PF</i>	Pitch Factor
<i>PPA</i>	Power, Performance and Area
<i>SCE</i>	Selective Chemical Etching
<i>SEM</i>	Scanning Electron Microscope
<i>SLE</i>	Stochastic Logical Effort
<i>SRAM</i>	Static random access memory
<i>SWCNT</i>	Single-walled CNT
<i>TEM</i>	Transmission Electron Microscope
<i>VLSI</i>	Very Large-Scale Integrated circuits
<i>VMR</i>	VLSI-compatible Metallic CNT Removal