# DEVICE MODELING AND CIRCUIT PERFORMANCE EVALUATION FOR NANOSCALE DEVICES: SILICON TECHNOLOGY BEYOND 45 nm NODE AND CARBON NANOTUBE FIELD EFFECT TRANSISTORS

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Jie Deng

June 2007

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

(H. -S. Philip Wong) Principal Advisor

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

(Boris Murmann) Co-Advisor

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

(Yoshio Nishi)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

(Subhasish Mitra)

Approved for the University Committee on Graduate Studies.

DEVICE MODELING AND CIRCUIT PERFORMANCE EVALUATION FOR NANOSCALE DEVICES: SILICON TECHNOLOGY BEYOND 45 nm NODE AND CARBON NANOTUBE FIELD EFFECT TRANSISTORS

Jie Deng, Ph.D.
Stanford University, 2007

Advisor: H. -S. Philip Wong

# Abstract

This thesis describes a body of work on modeling, understanding, and performance benchmarking for nanoscale devices and circuits, including both CMOS technology beyond the 45 nm node and carbon nanotube field effect transistors (CNFETs), with the aim of guiding nanoscale device and circuit design.

We propose a simple and accurate 4-point model for inverter effective drive current for nanoscale devices performance benchmarking. For CMOS technology beyond 65 nm node, it becomes more difficult to improve device performance by reducing the physical gate length. We propose improving the device and circuit performance by device footprint selective scaling and parasitic engineering, without reducing the physical gate length. The historic performance trend can continue for another 2 to 3 generations and the CMOS technology roadmap can be extended to 11 nm node with physical gate length no shorter than 10 nm.

Recognizing that the device structure has beening scaled from 3-D (bulk CMOS), quasi 2-D (partially depleted SOI), 2-D (fully depleted SOI), quasi 1-D (nanowire FET, FINFET, tri-gate FET), to 1-D (CNFET) for better channel electrostatics, it is important and necessary to model the behavior of 1-D device with the aim of guiding 1-D device and circuit design. Accurate (with less than 10% error) analytical models are presented to

calculate the electrostatic gate capacitance for 1-D and quasi 1-D FETs with high-*k* gate dielectric and multiple cylinder conducting channels.

Another goal of this thesis is to develop CNFET into a useful technology. Toward this goal, we develop and implement a universal circuit-compatible CNFET device model for device / circuit simulations, with the aim of evaluating and explaining device behavior and related physical phenomena, as well as obtaining the predictive performance metrics for guiding device and circuit design. Using this model, we found that the large device speed improvement (6× for nFET and 14× for pFET) of CNFET over CMOS technology at the device level is significantly degraded (by a factor of 5 to 8) by interconnect capacitance in a real circuit environment. Performance variations due to carbon nanotube (CNT) synthesis process induced imperfections further degrade circuit performance. In order to minimize performance variations at 32 nm node, less than 8% metallic CNTs and a density of at least 200 CNTs/μm are desired. Finally, the scalability of CNFET is examined using a simplified analytical model.

Approved for publication:

By: _____

For the Department of Electrical Engineering

# Acknowledgements

My five years experience at Stanford has been greatly enriched by my interactions with professors, friends, and colleagues. In particular, I am indebted to Professor H.-S Philip Wong for granting me the privilege of undertaking doctoral studies in his research group. He has allowed me tremendous freedom in defining and meeting my research objectives while providing invaluable guidance and insight. In addition to his technical expertise, I have learned a lot from his professionalism and his emphasis on clarity and precision in written and oral communication.

I greatly acknowledge the members of my dissertation reading committee: Professor Boris Murmann, Professor Yoshio Nishi, and Professor Subhasish Mitra. I greatly acknowledge the valuable independent study opportunity that Professor Boris Murmann granted me in his research group in 2004. Professor Subhasish Mitra has been of great help with technical discussions throughout this research work.

I wish to thank my colleagues, Nishant Patil, Lan Wei, Li-Wen Chang, Gordon Wan, and Albert Lin. Their great work has improved the quality of this work tremendously.

I thank Koungmin Ryu and Prof. Chongwu Zhou at USC for providing parallel-grown CNTs which are used to fabricate logic devices.

I wish to thank Amlani Islamshah and Lee King of Motorola for their providing the hardware data of the fabricated CNFET RF amplifier which make it possible to verify and improve our work.

I am especially grateful to Keunwoo Kim and Ching-Te Chuang of IBM for offering me the precious summer internship opportunity at the IBM T.J. Watson Research Center.

I wish to thank Dr. W. Haensch (IBM), Prof. M. Horowitz (Stanford), Dr. D. Singh (IBM), Dr. E.J. Nowak (IBM), Dr. J. Hutchby (SRC), Prof. D. Antoniadis (MIT), and my

group members, Arash Hazeghi, Gael Close, and Deji Akinwande, for the helpful technique discussions.

I greatly acknowledge Fely Barrera for supporting our group and doing an outstanding job at keeping our group functioning well.

The financial support of the SRC, MSD MARCO Focus Center Research Programs (FENA and C2S2), National Science Foundation, and Charles Powell Foundation are gratefully acknowledged.

I am deeply grateful to my wife, Yue Ri, whose love and understanding are great supports for me throughout my career at Stanford. I thank my parents for teaching me nearly everything I know. They have been sources of great inspiration to me.

Finally, I wish to thank my unborn little baby who makes me more than happy during the completion of this manuscript.

# TABLE OF CONTENTS

## Appendices

# LIST OF TABLES

# LIST OF FIGURES

# Chapter 1

## INTRODUCTION

Ever since the 0.35 μm node, the gate length of MOSFET has entered the deep-submicron region. 65 nm technology becomes the mainstream since 2006, and 45 nm technology has been announced in 2007. As CMOS continues to scale deeper into the nanoscale, various device non-idealities cause the I-V characteristics to be substantially different from well-tempered MOSFETs. It becomes more difficult to further improve device/circuit performance by reducing the physical gate length. The discrepancy between the fabricated physical gate length and the ITRS [1] projected gate length becomes larger as the technology advances, as shown in Figure 1.1. On the other hand, as the major driving force for the semiconductor industry, the device contacted gate pitch ($L_{pitch}$) is scaled down by a factor of 0.7 every technology node. Reasonable questions to ask are: Will the MOSFET scaling be stopped? Is there a way to extend the silicon technology roadmap? After silicon technology, or as a complement to silicon technology, is there any potential technology that may be used?

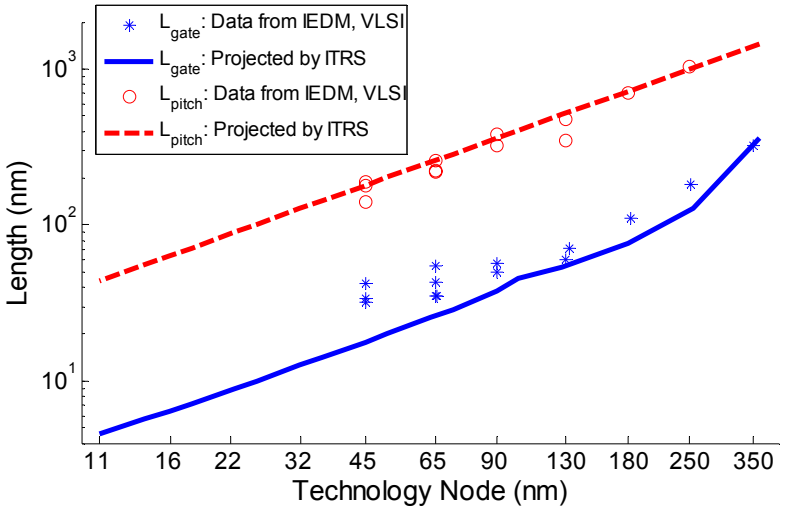



**Figure 1.1:** The physical gate length ($L_{gate}$) and the contacted gate pitch ($L_{pitch}$) of the fabricated devices (denoted by symbols) and projected by ITRS [1] (denoted by lines).

The last few years witnessed a dramatic increase in nanotechnology research, especially the nanoelectronics. These technologies vary in their maturity, as illustrated by Figure 1.2. The exciting opportunity is to design complex electronic circuits using the cutting-edge silicon technology and/or the novel nanometer-scale transistors in the future.

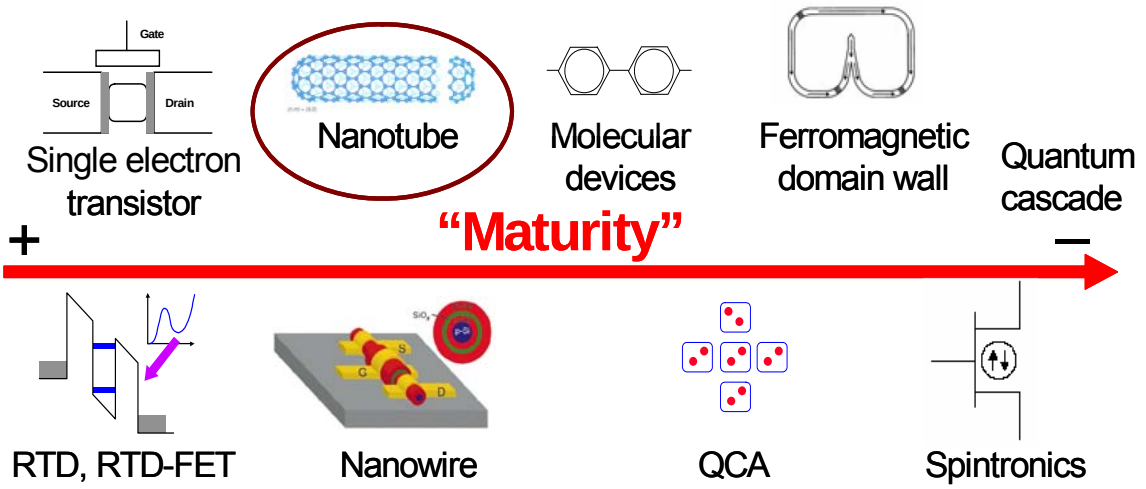



**Figure 1.2:** The novel nanoelectronic technologies in order of their maturity (presented at [2])

Carbon nanotubes (CNTs) are at the forefront of these new materials because of the unique mechanical and electronic properties. Carbon nanotube field effect transistor (CNFET) is the most promising technology to extend or complement traditional silicon technology due to three reasons: First, the operation principle and the device structure are similar to CMOS devices; we can reuse the established CMOS design infrastructure. Second, we can reuse CMOS fabrication process. And the most important reason is that CNFET has the best experimentally demonstrated device current carrying ability to date [3,4,5].

This thesis is a report on studies performed during the period 2005-2007. The objective of this work is modeling, understanding, and performance evaluation and prediction for nanoscale devices and circuits, including both silicon technology beyond the 45 nm CMOS technology node and carbon nanotube field effect transistors (CNFETs).

## 1.1 Nanoscale MOSFET

As CMOS continues to scale deeper into the nanoscale, various device non-idealities cause the I-V characteristics to be substantially different from well-tempered MOSFETs. For example, the source/drain series resistance is now a significant component of the total on-resistance. Proposals of metal contacted (Schottky) source/drain UTB SOI FET [6] also alter the I-V characteristics significantly. Novel non-Si devices such as the carbon nanotube FETs (CNFETs) operate with completely different device physics with quasi-ballistic transport in the channel [3] and Schottky barriers at the source/drain contacts [7]. In order to properly benchmark [8] future nanoscale Si FETs and novel FET such as CNFETs, it is necessary to develop a benchmarking metric that takes into account of the shape of the I-V characteristics with the device operated in a circuit environment. In Chapter 2, we propose a simple and accurate expression for inverter effective drive current for nanoscale Si and novel device (e.g. CNFET) performance benchmarking. This model includes the effects of both the nFET and the pFET of an inverter, and accurately captures the inverter delay performance over many CMOS technology nodes and in the presence of device non-idealities.

Technology boosters such as strain have helped the continuation of CMOS historic performance trend up to 45 nm node. As device physical gate length is reduced to below 25 nm at/beyond 65 nm technology node, various leakage currents and device parameter variation become the most important considerations for device optimization. In fact, it can be argued that reduction of gate length below 25 nm may not offer the same advantage as short-gate devices had provided historically in terms of power and performance at the system level [9]. The major detractors are: the lack of a thin equivalent gate oxide (with low leakage current) for effective short channel effect control, the increasing contribution of the fringing parasitic capacitance to the total gate capacitance, and the rising contribution of the source/drain resistance to the total device on-resistance. A reasonable question to ask is: if the gate length scaling stops, can we achieve performance gains through a different device scaling scenario?

The size of the active channel is only a small fraction of the total device footprint. Device footprint is mainly determined by patterning tolerances, spacer width, contact, and isolation sizes. In fact, the scaled device footprint has stayed fairly constant throughout many generations from 1 μm to 65 nm technology. Figure 1.3 shows historical data of device footprint for contacted pitch transistors and isolated transistors [10]. In [11] the role of device pitch was explored at the device level. In Chapter 3, we propose selective scaling of device footprint for 65 nm and beyond CMOS technologies. The benefits of selective scaling of device footprint are illustrated at both the device level and the circuit level using both ultra-thin body (UTB) fully-depleted SOI (FD-SOI) transistor and bulk MOSFET as examples.

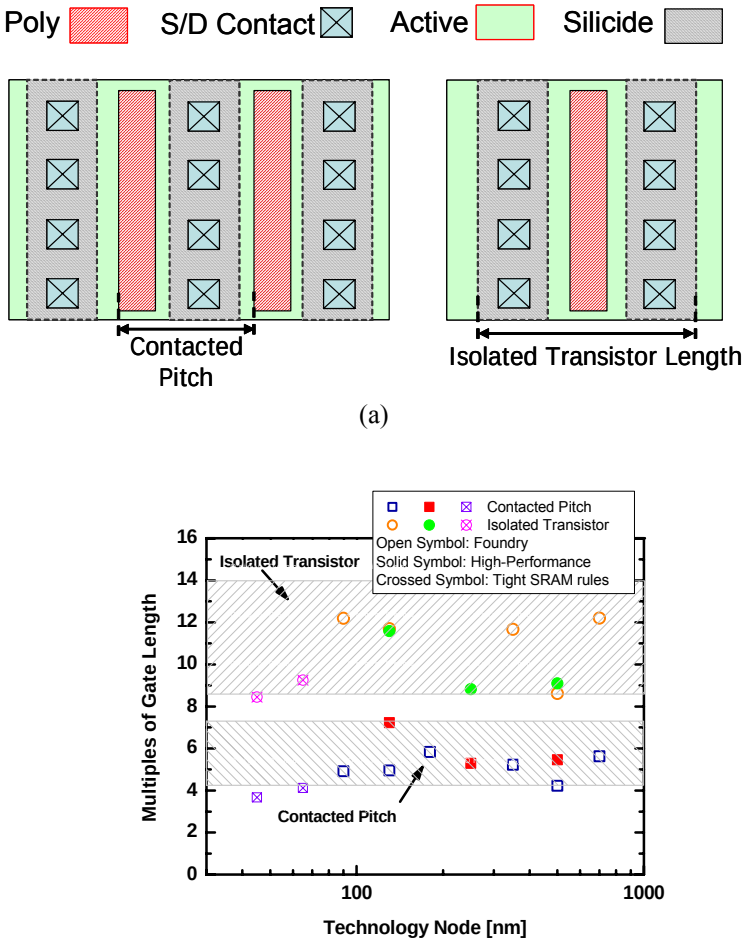



**Figure 1.3:** (a) The definitions for "Contacted pitch" and "Isolated transistor length" transistors. (b) The total device length vs. technology node. The ratio of the total device length over the gate length remains almost constant from 1 μm down to 65 nm technology node for both contacted pitch and isolated transistor. Adapted from a figure in Wong et al. [10] with additional data.

## 1.2 Carbon Nanotube Field-Effect Transistors

As one of the promising new devices, CNFET avoid most of the fundamental limitations for traditional silicon devices. All the carbon atoms in CNT are bonded to each other with $sp^2$ hybridization and there is no dangling bond which enables the integration with high-$k$ dielectric materials. In the next section, we will introduce the basic properties of CNFET, and will describe the problem to be modeled.

### 1.2.1 *Carbon Nanotube*

Carbon atom has an electron configuration of $1s^2 2s^2 2p^2$ in its ground state. In graphene, $sp^2$ hybridization occurs through covalent bonding of the two outermost shells. A carbon atom in graphene assembles in a single-sheet hexagonal lattice. The inter-carbon-atom distance within the hexagonal lattice ($d$) is approximately 1.44 Å, and the angle between carbon-carbon bonds (σ-bond) is 120 degrees. The lattice constant ($a$) is given by $\sqrt{3}d$ = 2.49 Å. The 2p electrons from all the atoms on a lattice form a delocalized π-orbital between the two adjacent sheets. The inter-layer spacing between the multiple sheets in graphene is about 3.35 Å. The weak electrostatic interactions between the sheets make it possible to assume the electrical characteristics of the graphite sheets are independent each other.

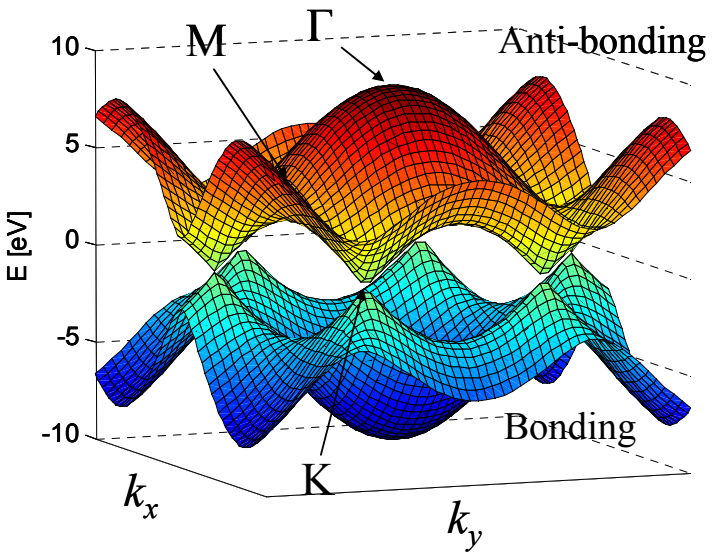



**Figure 1.4:** *E-k* dispersion relation for graphene, calculated using a nearest-neighbour tight-binding model. The three high-symmetry points are indicated by capital letters.

The dispersion relation for graphene, obtained by the Slater-Koster tight-binding scheme, considering only the π-orbital, is given by,

$$E_{g2D}(k_x, k_y) = \pm V_\pi \left\{ 1 + 4\cos\left(\frac{\sqrt{3}k_x a}{2}\right)\cos\left(\frac{k_y a}{2}\right) + 4\cos^2\left(\frac{k_y a}{2}\right) \right\}^{\frac{1}{2}} \qquad (1.1)$$

Where ($k_x$, $k_y$) are wavevectors, $V_\pi$ is the transfer integral (or the nearest-neighbour parameters). Figure 1.4 illustrates the band structure calculated with the above equation. The high-symmetry points are indicated by capital letters. *K*-points are degenerate, indicating the zero-bandgap (semi-metallic[+]) characteristic of graphene sheet.

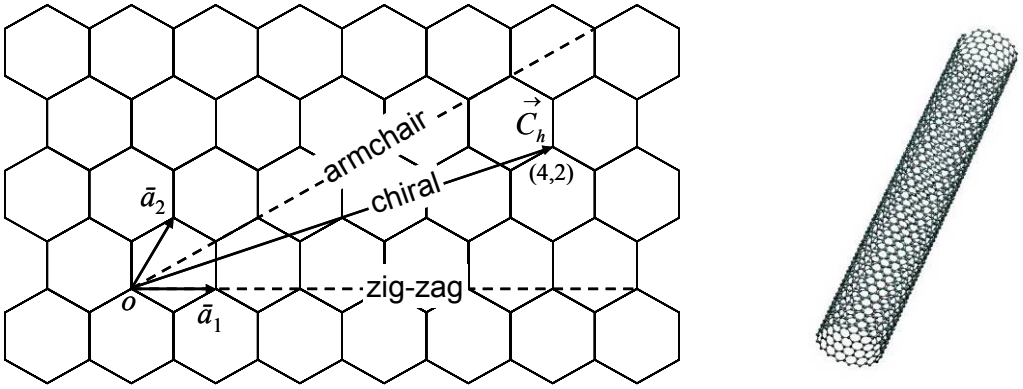



**Figure 1.5:** (a) Unrolled graphite sheet and (b) the rolled carbon nanotube lattice structure.

A single-walled carbon nanotube (SWCNT) can be visualized as a sheet of graphite which is rolled up and joined together along a wrapping vector $C_h = n_1.\bar{a}_1 + n_2.\bar{a}_2$, where $[\bar{a}_1, \bar{a}_2]$ are lattice unit vectors as shown by Figure 1.5, and the indices ($n_1$, $n_2$) are positive integers that specify the chirality of the tube [12]. The length of $C_h$ is thus the circumference of the CNT, which is given by,

$$C_h = a\sqrt{n_1^2 + n_1^2 + n_1 n_2} \qquad (1.2)$$

Single-walled CNTs are classified into one of there groups (Figure 1.5(a)), depends on the chiral number ($n_1$, $n_2$): (1) armchair ($n_1 = n_2$), (2) zigzag ($n_1 = 0$ or $n_2 = 0$), and (3) chiral (all other indices).

[+] Real graphite is a metal since the π-orbitals interaction between graphene sheets causes the bands to overlap by 40 mV [12].

The diameter of the CNT is given by the formula $D_{CNT} = C_h/\pi$. The typical diameters of CNTs are about several nanometers. Due to the small diameter of CNT, the quantization of wavevector in the circumferential direction occurs. A general analytic *E-k* dispersion relation for CNT is obtained by applying periodic boundary conditions in the circumferential direction to the 2D graphite sheet *E-k* dispersion relation. Pictorially, the dispersion of CNT is obtained by taking slices of the surface in Figure 1.4 with each cut determined by the circumferential quantization.

To be compatible with the quasi 1D structure of CNT, we convert the ($k_x$, $k_y$) coordinate to ($k_t$, $m$), where the wavevector $k_t$ is in the direction of transport, and $m$ is quantization number in the circumferential direction. Analytically, the dispersion of CNT is given by the formula [12],

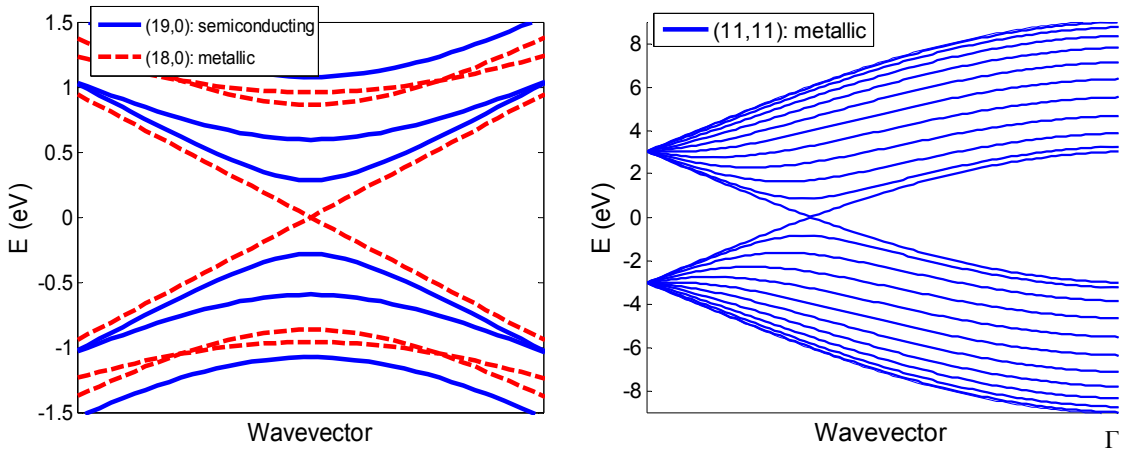
$$E_{CNT}(k_t, m) = \pm V_\pi \left\{1 + 4\cos t_1 \cos t_2 + 4\cos^2 t_2\right\}^{\frac{1}{2}} \tag{1.3}$$

Where the parameters are given by,

$$\begin{aligned}
t_1 &= \frac{\sqrt{3}a}{4}\frac{n_2 - n_1}{n}k_t + \frac{\pi}{2}\frac{n_1 + 3n_2}{n^2}m \\
t_2 &= \frac{\sqrt{3}a}{4}\frac{n_2 + n_1}{n}k_t + \frac{\pi}{2}\frac{3n_1 - n_2}{n^2}m \\
n^2 &= n_1^2 + n_2^2 + n_1 n_2 \\
-\frac{\pi \cdot d_R}{\sqrt{3}a \cdot n} &< k_t < \frac{\pi \cdot d_R}{\sqrt{3}a \cdot n} \\
m &= 0 : \frac{2n}{d_R} - 1 \\
d_R &= \gcd(2n_1 + n_2, 2n_2 + n_1)
\end{aligned} \tag{1.4}$$

In terms of the electrical conductivity, SWCNT is either metallic[+] (when $|n_1 - n_2|$ is a multiple of 3) with zero bandgap or semiconducting with finite bandgap. The band structure for both metallic CNTs and semiconducting CNTs are illustrated by Figure 1.6, using both armchair SWCNTs and zigzag SWCNTs as examples.

[+] To be accurate, achiral armchair SWCNTs ($n_1 = n_2$) and zigzag SWCNTs ($n_1 n_2 = 0$) are metallic, and the others ($|n_1 - n_2|$ is a multiple of 3 but $n_1 \neq n_2$ and $n_1 n_2 \neq 0$) are quasimetallic with small bandgap, calculated with Equiations (1.3, 1.4) [12].




(a) (b)

**Figure 1.6:** Energy dispersion relation for (a) (19,0), (18,0) zigzag CNTs, and (b) (11,11) armchair CNT.

The density of states (DOS) can be obtained as,

$$g(E) = \frac{2}{\pi}\sum_{m}\int \left|\frac{\partial E_{CNT}}{\partial k_t}\right|^{-1} dE_{CNT} \tag{1.5}$$

[13] proposes an universal expression to describe CNT band structure and the density of states, which are valid for the energy range $E_{CNT} << V_{\pi}$ . Figure 1.7 shows the DOS for (19,0) and (18,0) SWCNT with similar diameter (~ 1.5 nm).

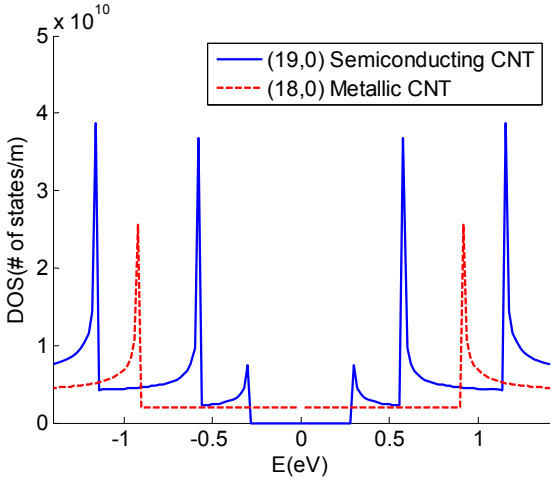



**Figure 1.7:** The density of states for (19,0) semiconducting and (18,0) metallic CNT, calculated using the universal DOS in [13].

The electrons in CNT are confined within the atomic plane of graphene. Due to the quasi-1D structure of CNT, the motion of the electrons in the nanotubes is strictly restricted.

Electrons may only move freely along the tube axis direction. As a result, all wide angle scatterings are prohibited. Only forward scattering and backscattering due to electron-phonon interactions are possible for the carriers in nanotubes. The experimentally observed ultra long elastic scattering mean-free-path (MFP) (~ 1 μm) [3,4,5,14] implies ballistic or near-ballistic carrier transport. High mobility, typical in the range of $10^3 \sim 10^4$ $cm^2/V{\cdot}s$ which are derived from conductance experiments in transistors, has been reported by a variety of studies [15,16]. Theoretical study also predicts a mobility of ~ $10^4$ $cm^2/V{\cdot}s$ for semiconduting CNTs [17]. The current carrying capacity of multi-walled CNTs are demonstrated to be more than $10^9$ A/ $cm^2$, about 3 orders higher than the maximum current carrying capacity of copper which is limited by the electron migration effect, without performance degradation during operation well above room temperature [18]. The superior carrier transport and conduction characteristic makes CNTs desirable for nanoelectronics applications, e.g. interconnect and nanoscale devices.

### 1.2.2 *CNFET*

The operation principle of carbon nanotube field-effect transistor (CNFET) is similar to that of traditional silicon devices. This three (or four) terminal device consists of a semiconduting nanotube, acting as conducting channel, bridging the source and drain contacts. The device is turned on or off electrostatically via the gate. The quasi-1D device structure provides better gate electrostatic control over the channel region than 3D device (e.g. bulk CMOS) and 2D device (e.g. fully depleted SOI) structures [19].

In terms of the device operation mechanism, CNFET can be categorized as either Schottky Barrier (SB) controlled FET (SB-CNFET) or MOSFET-like FET [3,4,20]. The conductivity of SB-CNFET is governed by the majority carriers tunneling through the SBs at the end contacts. The on-current and thereby device performance of SB-CNFET is determined by the contact resistance due to the presence of tunneling barriers at both or one of the source and drain contacts, instead of the channel conductance, as shown by Figure 1.8(a). The SBs at source/drain contacts are due to the Fermi-level alignment at the metal-semiconductor interface. Both the height and the width of the SBs, and therefore the conductivity, are modulated by the gate electrostatically. SB-CNFET shows

ambipolar transport behavior [27]. The work function induced barriers at the end contacts can be made to enhance either electron or hole transport. Thus both the device polarity (n-type FET or p-type FET) and the device bias point can be adjusted by choosing the appropriate work function of source/drain contacts [21]. On the other hand, MOSFET-like CNFET exhibits unipolar behavior by suppressing either electron (pFET) or hole (nFET) transport with heavily doped source/drain. The non-tunneling potential barrier in the channel region, and thereby the conductivity, is modulated by the gate-source bias (Figure 1.8(b)).

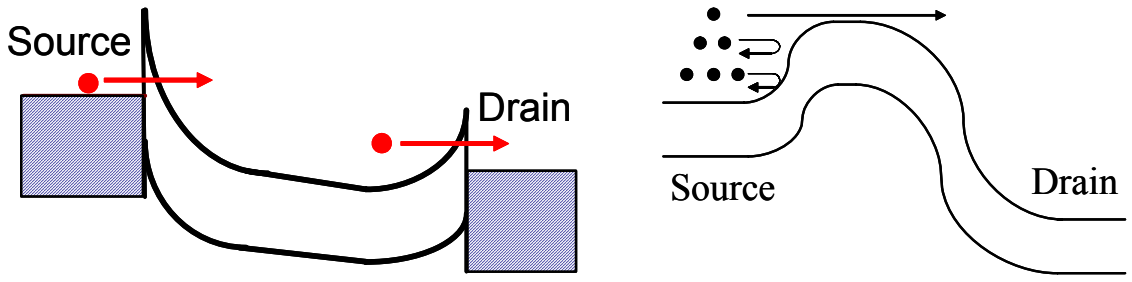



**Figure 1.8:** The energy band diagram for (a) SB-CNFET, and (b) MOSFET-like CNFET.

The first fabricated CNFET devices with Au or Pt source/drain metal contacts were reported in 1998 [22,23]. The gate dielectric material was a thick $SiO_2$ layer. A highly-doped Si back gate was used to control the conductivity. The $Al_2O_3$ gate dielectric was introduced to improve the gate controllability over the channel region [24]. The front-gate device structure, by placing the gate electrode over the thin gate oxide that covers CNT, was used to further improve the channel electrostatics [ 25 ]. Better gate electrostatics was achieved by using high-*k*, e.g. $HfO_2$, gate dielectric material [20,26]. The source/drain contacts using a variety of metals (Ti, Ni, Al, Pd, …) were fabricated to study the effect of the work function difference between the metal contacts and CNT on device conductivity. Ti source/drain metallization was reported to be efficient on reducing the contact resistance [27]. The device fabricated with Pd source/drain metal contact, Al gate electrode, and $HfO_2$ gate dielectric was reported to achieve excellent dc characteristics [3].

Logic circuits with field-effect transistors based on single carbon nanotubes have been demonstrated in the past few years. In 2001, [28] demonstrated one-, two-, and three-

transistor circuits that exhibit a range of digital logic operations, including an inverter, a logic NOR, a static random-access memory (SRAM) cell, and an three-stage ac ring oscillator operating at 5 Hz. A five-stage CMOS type nanotube ring oscillator using palladium p-type gates and aluminum n-type gates was reported in 2006 [29]. Owing to the compact device/circuit design, this ring oscillator works at a frequency of 72 MHz. Regarding RF analog application using CNFET, the first demonstration of ac gain in a single-walled carbon nanotube common-source amplifier was reported in 2006 [30]. The low frequency gain was ~ 11.3 dB, and the unity-gain frequency was about 560 kHz which was mostly limited by the parasitic load capacitance.

While the CNT synthesis / fabrication technique and the performance of CNFET devices and circuits have been significantly improved since the first fabricated device in 1998, CNFETs is still premature for very large scale integrated (VLSI) circuits design and commercial use. In order for CNFET to develop into a technology, first, we need tools to enable circuit design and performance benchmarking.

Efforts have been made in recent years on modeling semiconducting CNFET [31,32,33,34] for digital logic applications and CNT for interconnects [35,36] in order to evaluate the potential performance at the device level. This thesis will mostly focus on the device applications of CNT. A numerical model was reported in [37] to evaluate the dc current of SB-CNFET. The model reported in [38] predicts the dc performance of short channel SB-CNFET. Though good dc current can be achieved by SB-CNFET with the self-aligned structure [3,37,38], its ac performance is going to be poor due to the proximity of the gate electrode to the source/drain metal. The ambipolar behavior of SB-CNFET also makes it undesirable for complementary logic design. Considering both the fabrication feasibility [39] and superior device performance of MOSFET-like CNFET as compared to SB-CNFET, we will focus on MOSFET-like CNFETs. To evaluate the device/circuit performance as well as the performance dependence on device/geometry parameters, the requirements for a good device model include:

(1) Good scalability.
(2) Physics-based, or at least semi-physics based.

(3) Reasonable accuracy for both large signal and small signal analysis.

(4) Acceptable run time.

The reported compact models to date [31,32,33,34] used one or more lumped static gate capacitances and assumed an ideal ballistic transport channel. These simplifications make it questionable when evaluating the transient response and device dynamic performance. The integral function used in [31,32] requires intensive calculation efforts and thereby makes it difficult to implement in circuit simulators, e.g. HSPICE [40]. The model in [33] improves the run time significantly by using a polynomial fitting approach. This methodology dilutes the physical meaning of the device model and makes evaluating CNFET performance with different device parameters (e.g. CNT chiralities, gate oxide thickness) inconvenient. The reported models to date [31,32,34] used a simple coaxial or planer gate structure that differs from the typical realistic CNFET gate structure that consists of high-*k* gate oxide on top of $SiO_2$ insulating bulk [26]. For a CNFET with multiple parallel CNTs [3], these published models cannot examine the multiple CNT-to-CNT screening effect on both the driving current and the effective gate capacitance [41]. All the reported device models assumed CNFET devices with perfect and ideal CNT channel. Compared to the intrinsic performance of CNFET predicted by theoretical studies [31], the actual device and circuit level performance is mostly limited by various parasitics and process induced imperfections. The device parasitics and/or non-idealities include, but are not limited to: the channel length dependence of current drive, the finite scattering mean free path, the source/drain series resistance, the source/drain contacts (SBs) resistance, the geometry dependence of the gate to channel capacitance, and the interconnect wiring capacitance. To evaluate CNFET device/circuit performance with improved accuracy, a CNFET device model with a more complete circuit-compatible structure and also incorporating the typical device/circuit non-idealities is necessary. A good balance between the simulation run-time and accuracy is desired.

The circuit macro level performance is not only limited by the performance of one single device, but also limited by the device performance variations which are significant for nanometer scale devices [42]. There are a variety of device parameter variations and imperfections caused by today's CNT synthesis/fabrication technique: (1) CNT diameter

and chirality control [43]; (2) Doping level control [39]; (3) The probability of a CNT to be metallic [12,44]; (4) Directed-CNT-growth [45,46]. A reasonable question to ask is: considering these imperfections, what can be gained at the circuit-level using CNFET technology compared to cutting-edge Si CMOS?

Finally, in order for CNFET to develop into a competitive technology, it should have good scalability, i.e. the performance advantage of CNFET over MOSFET is expected to improve (or at least maintain the same) as the technology node advances.

## 1.3 Thesis Outline

The core chapters of this thesis, from Chapter 2 to Chapter 7, is a collection of manuscripts published in or submitted to journals or conference proceedings based on work performed at Stanford during the period 2005-2007. This thesis focuses on the performance prediction and evaluation of nanoscale devices and circuits including the realistic device structures, device/circuit non-idealities and process related imperfections. The first part of this thesis, Chapter 2 and Chapter 3, covers nanoscale CMOS technology. The latter part of this thesis, Chapter 4 to Chapter 7, discusses the modeling, the performance benchmarking, the opportunities, and the challenges for CNFET technology as an extension to or a complementary of the traditional CMOS technology.

Chapter 2 presents a simple and accurate expression for inverter effective drive current for nanoscale Si and novel device (e.g. CNFET) performance benchmarking.

Chapter 3 shows that both the device and circuit performance can be improved by selective scaling of device footprint for 65 nm and beyond CMOS technologies. Device selective footprint scaling and parasitic engineering are shown to be the efficient ways to extend the CMOS technology roadmap.

Chapter 4 presents accurate analytical models to calculate the electrostatic gate capacitance of 1-D field effect transistors (FETs) with multiple cylinder conducting channels.

Chapter 5 and Chapter 6 present a physics-based circuit-compatible compact model for MOSFET-like single-walled CNFET. Chapter 5 describes the compact model for the intrinsic channel region of CNFET, and Chapter 6 reports the complete CNFET model including the practical device/circuit non-idealities. This model is implemented in both HSPICE and Verilog-A[+] that enables both device performance evaluation and macro circuit level performance benchmarking with reasonable accuracy and run time.

Chapter 7 evaluates CNFET circuit performance, in the presence of process related non-idealities and imperfections in a real circuit environment, at the 32 nm technology node. We also demonstrate the design principles for designing CNFET-based logic circuits that are guaranteed to implement correct logic functions even in the presence of misaligned CNTs. This chapter also presents an analytical model describing the scalability of CNFET technology.

Chapter 8 summarizes the key findings and contributions of this thesis, and proposes some recommendations for future work.

A brief description for the key parameters in the implemented HSPICE model and the default parameter settings are summarized in Appendix A.

Appendix B shows an example of using this compact model to match both the dc and ac characteristics of the fabricated CNFET RF amplifier.

## 1.4 Bibliography

[1] Data from http://www.itrs.net/Links/2005ITRS/Home2005.htm

[2] J. Deng, N. Patil, K. Ryu, A. Badmaev, C. Zhou, S. Mitra, H.-S. P. Wong, "Carbon Nanotube Transistor Circuits: Circuit-Level Performance Benchmarking and Design Options for Living with Imperfections," *International Solid State Circuits Conference (ISSCC)*, pp. 70-71, San Francisco, CA, February, 2007.

[3] A. Javey, J. Guo, D. B. Farmer, Q. Wang, E. Yenilmez, R. G. Gordon, M. Lundstrom, H. Dai, "Self-Aligned Ballistic Molecular Transistors and Electrically Parallel Nanotube

[+] Available at https://www.stanford.edu/group/nanoelectronics/model_downloads.htm

Arrays," *Nano Letters*, vol. 4, pp. 1319-1322, 2004.

[4] A. Javey, R. Tu, D.B. Farmer, J. Guo, R.G. Gordon, H. Dai, "High Performance n-Type Carbon Nanotube Field-Effect Transistors with Chemically Doped Contacts," *Nano Letters*, vol. 5, pp. 345-348, 2005.

[5] D. Mann, A. Javey, J. Kong, Q. Wang, and H. Dai, "Ballistic Transport in Metallic Nanotubes with Reliable Pd Ohmic Contacts," *Nano Letters*, vol. 3, pp. 1541-1544, 2003.

[6] J. Kedzierski, P. Xuan, E.H. Anderson, J. Bokor, T-J. King, C. Hu, "Complementary silicide source/drain thin-body MOSFETs for the 20 nm gate length regime," *IEDM*, pp. 57-60, 2000.

[7] J. Appenzeller, J. Knoch, R. Martel, V. Derycke, S.J. Wind, Ph. Avouris, "Carbon Nanotube Electronics," *IEEE Trans. Nanotech.*, vol. 1, no. 4, pp. 184-189, 2002.

[8] R. Chau, S. Datta, M. Doczy, B. Doyle, B. Jin, J. Kavalieros, A. Majumdar, M. Metz, M. Radosavljevic, "Benchmarking Nanotechnology for High-Performance and Low-Power Logic Transistor Applications," *IEEE Trans. Nanotech.*, vol. 4, no. 2, pp. 153-158, 2005.

[9] M. Horowitz, E. Alon, D. Patil, "Scaling, Power, and the Future of CMOS," *IEDM Tech. Digest*, pp. 11-14, 2005.

[10] H.-S. P. Wong, G.S. Ditlow, P.M. Solomon, X. Wang, "Performance Estimation and Benchmarking for Carbon Nanotube FETs and Nanodiode Arrays," invited paper, *International Conferences on Solid State Devices and Materials (SSDM)*, Japan, September 16 – 18, pp. 802 – 803, 2003.

[11] J. W. Sleight, I. Lauer, O. Dokumaci, D. M. Fried, D. Guo, B. Haran, S. Narasimha, C. Sheraw, D. Singh, M. Steigerwalt, X. Wang, P. Oldiges, D. Sadana, C.Y. Sung, W. Haensch, M. Khare, "Challenges and Opportunities for High Performance 32 nm CMOS Technology, **"** *IEEE International Electron Devices Meeting (IEDM)*, pp. 697-700, San Francisco, CA, December 11-13**,** 2006.

[12] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, "Physical Properties of Carbon Nanotubes," *Imperial College Press, London*, 1998.

[ 13 ] J. W. Mintmire and C. T. White, "Universal Density of States for Carbon Nanotubes," *Physical Review Letters*, vol. 81, pp. 2506-2509, 1998.

[14] Z. Yao, C. L. Kane, and C. Dekker, "High-field Electrical Transport In Single-Wall Carbon Nanotubes," *Phys. Rev. Lett.*, vol. 84, pp. 2941–2944, 2000.

[15] M. S. Fuhrer, B. M. Kim, T. Dürkop, and T. Brintlinger, "High-Mobility Nanotube Transistor Memory," *Nano Letters*, vol. 2, pp. 755-759, 2002.

[16] Paul L. McEuen, Michael S. Fuhrer, and Hongkun Park, "Single-Walled Carbon Nanotube Electronics," *IEEE Trans. Nanotech.*, vol. 1, pp. 78-85, 2002.

[17] G. Pennington and N. Goldsman, "Semiclassical Transport and Phonon Scattering of Electrons in Semiconducting Carbon Nanotubes," *Phys. Rev. B.*, vol.68, pp. 045426, 2003.

[18] B. Q. Wei, R. Vajtai, and P. M. Ajayan, "Reliability and Current Carrying Capability of Carbon Nanotubes," *Appl. Phys. Lett.*, vol. 79, pp. 1172-1174, 2001.

[19] Y. Cui, Z. Zhong, D. Wang, W. U. Wang, and C. M. Lieber, "High Performance Silicon Nanowire Field Effect Transistors," *Nano Letters*, vol. 3, pp. 149-152, 2003.

[20] A. Javey, J. Guo, D. B. Farmer, Q. Wang, D. Wang, R. G. Gordon, M. Lundstrom, and H. Dai, "Carbon Nanotube Field-Effect Transistors with Integrated Ohmic Contacts and High-*k* Gate Dielectrics," *Nano Letters*, vol. 4, pp. 447-450, 2004.

[21] Z. Chen, J. Appenzeller, P. M. Solomon, Y.-M. Lin, P. Avouris, "Gate Work Function Engineering for Nanotube-Based Circuits," *International Solid State Circuits Conference (ISSCC)*, p. 68 – 69, 2007.

[22] R. Martel, T. Schmidt, H. R. Shea, T. Hertel, and Ph. Avouris, "Single- and Multi-wall Carbon Nanotube Field-Effect Transistors," *Appl. Phys. Lett.*, vol. 73, pp. 2447-2449, 1998.

[23] S. J. Tans, A. R. M. Verschueren, and C. Dekker, "Room-Temperature Transistor Based on a Single Carbon Nanotube," *Nature*, vol. 393, pp. 49-52, 1998.

[24] A. Bachtold, P. Hadley, T. Nakanishi, and C. Dekker, "Logic Circuits with Carbon Nanotube Transistors," *Science*, vol. 294, pp. 1317-1320, 2001.

[25] S. J. Wind, J. Appenzeller, R. Martel, V. Derycke, and Ph. Avouris, "Vertical Scaling of Carbon Nanotube Field-Effect Transistors using Top Gate Electrodes," *Appl. Phys. Lett.*, vol. 80, pp. 3817-3819, 2002.

[26] A. Javey, H. Kim, M. Brink, Q. Wang, A. Ural, J. Guo, P. McIntyre, P. McEuen, M. Lundstrom, and H. Dai, "High-*k* Dielectrics for Advanced Carbon-Nanotube Transistors and Logic Gates," *Nature Materials*, vol. 1, pp. 241-246, 2002.

[27] R. Martel, V. Derycke, C. Lavoir, J. Appenzeller, K. K. Chan, J. Tersoff, and Ph. Avouris, "Ambipolar Electrical Transport in Semiconducting Single-wall Carbon Nanotubes," *Phys, Rev. Lett.*, vol 87, pp. 256805, 2001.

[28] A. Bachtold, P. Hadley, T. Nakanishi, C. Dekker, "Logic Circuits with Carbon Nanotube Transistors," *Science*, Vol. 294, pp. 1317–1320, 2001.

[29] Z. Chen, J. Appenzeller, Y. M. Lin, J. Sippel-Oakley, A. G. Rinzler, J. Tang, S. J. Wind, P. M. Solomon, P. Avouris, "An Integrated Logic Circuit Assembled on a Single Carbon Nanotube," *Science*, Vol. 311, pp. 1735-1735, 2006.

[30] I. Amlani. J. Lewis, K. Lee, R. Zhang, J. Deng, H.-S. P. Wong, "First Demonstration of AC Gain From a Single-Walled Carbon Nanotube Common-Source Amplifier," *IEEE International Electron Devices Meeting (IEDM)*, pp. 559-562, San Francisco, CA, December 11-13, 2006.

[31] K. Natori, Y. Kimura, and T. Shimizu, "Characteristics of a Carbon Nanotube Field-Effect Transistor Analyzed as a Ballistic Nanowire Field-Effect Transistor," *Journal of Applied Physics*, vol. 97, pp. 034306, 2005.

[32] J. Guo, M. Lundstrom, and S. Datta, "Performance Projections for Ballistic Carbon

Nanotube Field-Effect Transistors," *Applied Physics Letters*, vol. 80, pp. 3192-3194, 2002.

[33] A. Raychowdhury, S. Mukhopadhyay, and K. Roy, "A Circuit-Compatible Model of Ballistic Carbon Nanotube Field-Effect Transistors," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 23, pp. 1411-1420, 2004.

[34] C. Dwyer, M. Cheung, and D. J. Sorin, "Semi-empirical SPICE models for carbon nanotube FET logic," *4th IEEE Conference on Nanotechnology*, pp. 386-388, 2004.

[35] P. J. Burke, "Lüttinger Liquid Theory as a Model of the Gigahertz Electrical Properties of Carbon Nanotubes," *IEEE Transactions on Nanotechnology*, vol. 1, pp. 129-144, 2002.

[36] A. Naeemi, R. Sarvari, and J. D. Meindl, "Performance Comparison between Carbon Nanotube and Copper Interconnects for Gigascale Integration (GSI)," *Electron Device Letters, IEEE*, vol. 26, pp. 84-86, 2005.

[37] D. L. John, L. C. Castro, J. Clifford, and D. L. Pulfrey, "Electrostatics of Coaxial Schottkt-Barrier Nanotube Field-Effect Transistors," *IEEE Trans. Nanotech.*, vol. 2, pp. 175-180, 2003.

[38] A. Hazeghi, T. Krishnamohan and H.-S. P. Wong, "Schottky-Barrier Carbon Nanotube Field-Effect Transistor Modeling," *IEEE Transactions on Electron Devices,* vol. 54, pp. 439-445, 2007.

[39] J. Chen, C. Klinke, A. Afzali, and P. Avouris, "Self-aligned carbon nanotube transistors with charge transfer doping," *Applied Physics Letters*, vol. 86, pp. 123108, 2005.

[40] *HSPICE*®, Version 2004.09, Synopsys Corp., CA

[41] J. Deng and H.-S P. Wong, "Modeling and Analysis of Planar Gate Capacitance for 1-D FET with Multiple Cylindrical Conducting Channels," *IEEE Transactions on Electron Devices*, to appear, 2007.

[42] J. Deng, K. Kim, C.-T. Chuang, H.-S P. Wong, "The Impact of Device Footprint Scaling on High Performance CMOS Logic Technology," *IEEE Transactions on Electron Devices,* vol. 54, pp. 1148-1155*,* 2007.

[43] Y-C Tseng, K. Phoa, D. Carlton, and J. Bokor, "Effect of Diameter Variation in a Large Set of Carbon Nanotube Transistors," *Nano Letters,* vol. 6, pp. 1364 – 1368, 2006.

[44] 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, H. Dai, "Preferential Growth of Semiconducting Single-Walled Carbon Nanotubes by a Plasma Enhanced CVD Method," *Nano Letters*, vol. 4, pp. 317-321, 2004.

[45] S. Han, X. Liu, C. Zhou, "Template-Free Directional Growth of Single-Walled Carbon Nanotubes on a- and r-Plane Sapphire," *J. Am. Chem. Soc.*, vol. 127, pp. 5294 – 5295, 2005.

[46] 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*, vol. 2, pp. 230-236, 2007.

# Chapter 2

# METRICS FOR PERFORMANCE BENCHMARKING OF NANOSCALE DEVICES



## 2.1 Introduction

A commonly used FET performance metric is the gate delay metric $C_{Load}V_{DD}/I$, where I is the saturation on-current ($I_{Dsat}=I_{DS}@V_{GS}=V_{DS}=V_{DD}$) [1]. Previous works [2,3] have shown that the "I" of the CV/I (the effective drive current) is better represented by a value other than $I_{Dsat}$ because $I_{Dsat}$ is never reached during switching [2,3].

As CMOS continues to scale deeper into the nanoscale, various device non-idealities cause the I-V characteristics to be substantially different from well-tempered MOSFETs (Figure 2.1(a)). For example, the source/drain series resistance is now a significant component of the total on-resistance. Proposals of metal contacted (Schottky) source/drain UTB SOI FET [4] also alter the I-V characteristics significantly. Novel non-Si devices such as the carbon nanotube FETs (CNFETs) operate with completely different device physics with quasi-ballistic transport in the channel [5] and Schottky barriers at the source/drain contacts [6]. The I-V characteristics of these novel devices, while they appear deceptively similar, are <u>not</u> scaled versions of the Si FET (Figure 2.1(b)). In order to properly benchmark [7] future nanoscale Si FETs and novel FET such as CNFETs, it is necessary to develop a benchmarking metric that takes into account of the shape of the I-V characteristics with the device operated in a circuit environment.

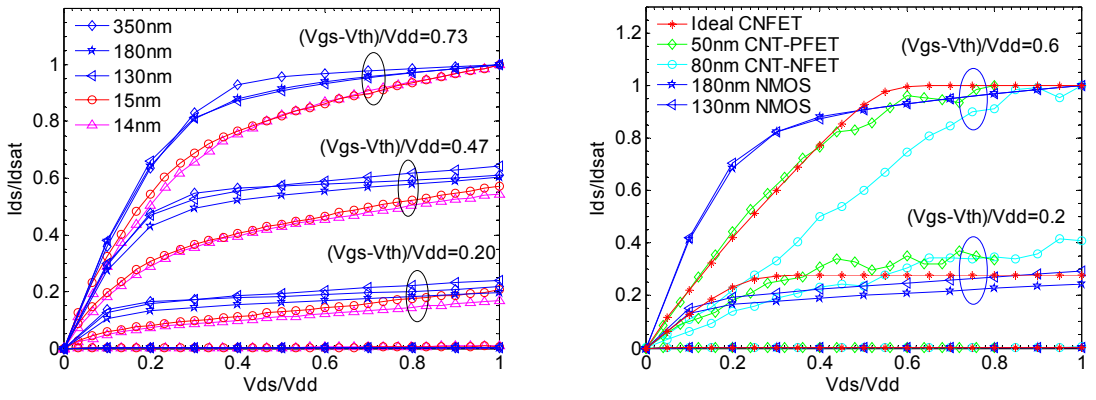


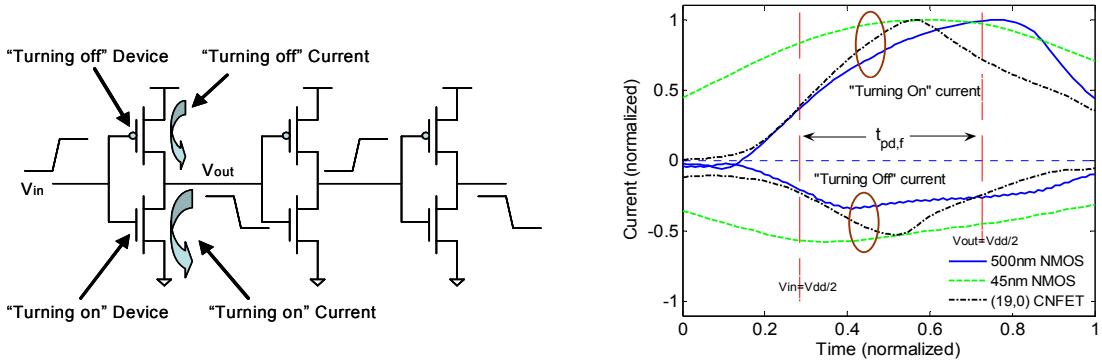

**Figure 2.1:** NMOS $I_{ds}/I_{Dsat}$ as a function of $V_{ds}/V_{dd}$ for (a) IBM 350nm, 180 nm and 130 nm CMOS technology, 15 nm [10], 14 nm [11] CMOS devices. (b) IBM 180 nm and 130 nm CMOS technology, 80 nm n-CNFET from [12], 50 nm p-CNFET from [5], and the ideal (19, 0) CNFET with appropriate source/drain resistance (simulated with [13]). The gate overdrive ratios are adjusted to the same values for all devices.

## 2.2 Models for Effective Drive Current

The nFET and pFET in an inverter are sized to give the same rising delay and falling delay so that only one transition edge needs to be considered. For historical reason and to be convenient, the nFET data and the rise-to-fall transition edge are usually chosen for analysis, but the fall-to-rise transition edge follows the same principle. Na et al. [3] showed that the choice of $I_{eff}=(I_H+I_L)/2$, where $I_L= I_{DS}$ ($V_{GS}=0.5V_{DD}$, $V_{DS}=V_{DD}$), $I_H=I_{DS}$ ($V_{GS}=V_{DD}$, $V_{DS}=0.5V_{DD}$), say *2-point model*, has better correlation with inverter gate delay than $I_{Dsat}$. This expression has two implicit assumptions: (1) The "turning on" current, which flows through the device that is being turned on (Figure 2.2(a)), is a linear function of time from the 50% to 50% transition point of the input/output waveform, and (2) the ratio of the "turning off" current, which flows through the device that is being turned off (Figure 2.2(a)), to the "turning on" current is constant from 50% to 50%, or the "turning off" current is relatively small compared to the "turning on" current.

For conventionally scaled MOSFETs, both assumptions are valid (Figure 2.2(b)). As CMOS scales deeper into the nanoscale, as well as some novel devices, the "turning on" current is no longer a linear function of time within 50% to 50% region (Figure 2.2(b))

due to the worse short channel effect and non-conventional scaling of the power supply and threshold voltage.




(a) (b)

**Figure 2.2:** Signal transition wave forms of an inverter in the ring oscillator, $t_{pd,f}$ and $t_{pd,r}$ are trimmed to be the same. (a) Definition of the "turning on" and "turning off" current of inverter (b) For IBM 500nm technology, the switching current through NMOS is almost a linear function of time within 50% to 50% region, while for ultra short channel device (Berkeley 45 nm NMOS) and novel device (45nm CNFET), the "turning on" current starts to saturate and then fall off in 50% to 50% region, and the "turning off" current become important compared to the "turning on" current.

The comparison of the normalized switching current trajectories of an inverter among 500nm, 45nm silicon technology and (19,0) CNFETs (Figure 2.3) illustrates that both assumptions made above are no longer valid for aggressively scaled devices and CNFETs.

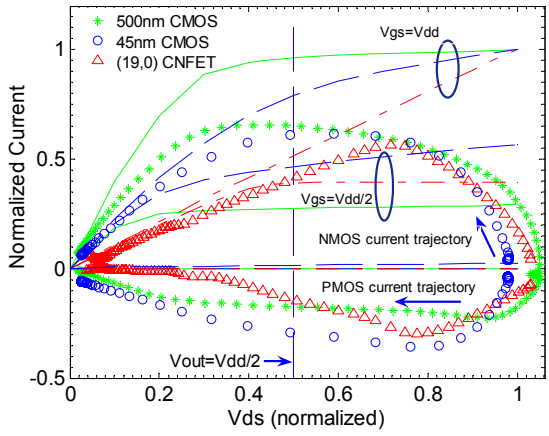



**Figure 2.3:** The switching current trajectories $I_{ds}(t)$ (symbols) of an inverter of the ring oscillator are mapped to NMOS static $I_{ds}/I_{Dsat}$ as a function of $V_{ds}/V_{dd}$ for IBM 500nm CMOS technology, Berkeley 45nm predictive BSIM4 model and (19,0) CNFET with 45nm channel length (simulated with [13]). The "turning on" current trajectories are bounded by the Ids vs. Vds curve.

The "turning on" current trajectory of CNFET which is bounded by the $I_{DS}$ vs. $V_{DS}$ curve does not follow that of silicon devices. The "turning off" current for short channel device gets more significant compared to the "turning on" current, due to 1) $V_T/V_{DD}$ is getting

smaller with continuous scaling which results in larger subthreshold current; 2) the circuit gets faster so that the output decreases into the saturation region, instead of triode region, of the pFET while the input still keeps pFET open, and 3) the shape of the I-V curve gets more non-ideal, or "resistive", i.e. the ratio of the current in the linear region of pFET over the complementary current in the saturation region of nFET is larger for a "resistive" I-V curve than for an ideally shaped I-V curve, given the same input/output voltage.

For novel FET such as CNFETs and metal contact (Schottky barrier) source/drain UTB SOI FETs, the gradually rising $I_{DS}$ current as a function of $V_{DS}$ (Figure 2.1 (b)) results in an inefficient "turning off" transition. It is necessary to consider both nFET and pFET currents through the 'turning on' and 'turning off' devices when benchmarking device performance in terms of inverter delay.

The inverter effective drive current ($I_{eff}=1/2*C_{eff}V_{dd}/\tau_{pd}$) is a strong function of the output load. A handy universal expression is not likely to be available. Here, we consider the fan-out one (FO1) inverter delay in a ring oscillator which is often used to benchmark the device/circuits performance. We investigated several models (Table 2.1) using the linear combination of some critical points data: (1) Saturation on-current model, $I_{Dsat}$; (2) *2-point model* [3]; (3) *3-point model*: To refine the first assumption listed earlier, one more point is added at the middle of the switching trajectory with the assumption that the "turning on" current is segmentable linear; (4) *4-point model*: To refine the second assumption, we include the effect of "turning off" currents. The 4-point model utilizes 3 pairs of data of nFET and pFET in principle, but requires only 4 data in calculation, since the other two pFET data, $I_{SD(P\text{-}FET)}@(V_{SG}=V_{DD,}\ V_{SD}=0)$, $I_{SD(P\text{-}FET)}@(V_{SG}=0,\ V_{SD}=V_{DD})$, are zero by definition. The currents $I_{NL}$, $I_{NM}$…etc. in Table 2.1 are defined in Table 2.2.

**Table 2.1:** The models to predict the Inverter effective drive current

| Model | $I_{DSAT}$ | 2-point [3] |
|---|---|---|
| $I_{eff}$ | $I_{DS}@V_{GS}=V_{DS}=V_{DD}$ | $\frac{I_{NL}+I_{LH}}{2}$ |
| Data | nFET | nFET |
| Model | 3-point | 4-point |
| $I_{eff}$ | $\frac{I_{NL}+I_{NM}+I_{NH}}{3}$ | $\frac{I_{NL}+I_{NM}+I_{NH}-I_{P}}{3}$ |
| Data | nFET | nFET, pFET |

**Table 2.2:** The definition of current components utilized by models

| **N-FET** | **$V_{DS}$@** | **$V_{GS}$@** |
|---|---|---|
| $I_{NL}$ | $V_{DD}$ | $0.50V_{DD}$ |
| $I_{NM}$ | $0.75V_{DD}$ | $0.75V_{DD}$ |
| $I_{NH}$ | $0.50V_{DD}$ | $V_{DD}$ |
| **P-FET** | **$V_{SD}$@** | **$V_{SG}$@** |
| $I_P$ | $0.25V_{DD}$ | $0.25V_{DD}$ |

## 2.3 Results and Discussion

To analyze the merits of the purposed models, the modeled inverter $I_{eff}$ is compared to the actual inverter effective drive current extracted with HSPICE [8]. These two values, modeled $I_{eff}$ and measured $I_{eff}$ should be linearly correlated. For the purpose of device performance comparison, the actual value of the modeled Ieff is not a big concern. Instead, we are interested with both the slope of the Least Mean Square (LMS) fitted line and the correlation relation between the modeled $I_{eff}$ and measured $I_{eff}$, which can be quantified by the norm of the residuals between the modeled $I_{eff}$ (normalized) and the LMS fitted line (normalized). However, while the fitted line slope can be calibrated for technology prediction, a 1:1 relationship is convenient and preferred in device performance prediction. The smaller norm of the residuals indicates better correlation between two sets of data. Both axes of $I_{eff}$ are normalized to the same scale to eliminate the possible error caused by different scaling. The Berkeley 90nm CMOS model [9] was chosen for analysis.

### 2.3.1 *Conventionally Scaled Devices*

When benchmarking the performance of the conventionally scaled devices (operated under nominal power supply), the device performance predicted by the 4-point model has the best correlation with the actual device performance as the fitted line slope is closest to 1, and the norm of the residuals is the smallest among all the models (Figure 2.4). Smaller differences are found among the normalized norms of the fitted line residuals (0.19 for $I_{DSAT}$-model, 0.16 for 2-point and 3-point model, 0.15 for 4-point model), after eliminating the effect of different axis scaling, which suggests that the saturation on-

current may continue to be a reasonable predictor of inverter gate delay for technologies down to 65 nm node.

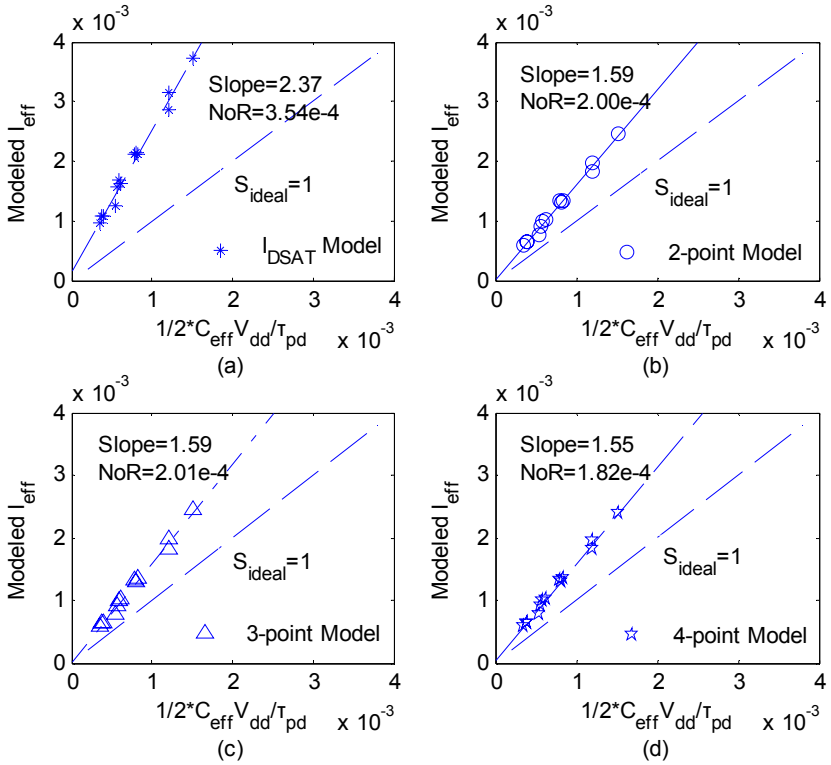



**Figure 2.4:** The modeled inverter $I_{eff}$ is compared to the actual $I_{eff}$ for various CMOS technologies from 500 nm to 65 nm nodes (original data). The LMS fitted lines are also plotted. (a) With saturation on-current model, the LMS fitted line slope is 2.37 and the norm of the residuals (NoR) is 3.54E-4. (b) With 2-point model, the LMS fitted line slope is 1.59 and the NoR is 2.00E-4. (c) With 3-point model, the LMS fitted line slope is 1.59 and the NoR is 2.01E-4. (d) With 4-point model, the LMS fitted line slope is 1.55 which is closest to the ideal slope 1, and the NoR is 1.82E-4 which is the smallest among all the models.

As device continues to scale deeper into the nanoscale, various device non-idealities become more significant. The non-idealities and scaling effects can be classified into three major categories: 1) source/drain series resistance $R_{sd}$, 2) $V_{TH}/V_{DD}$ ratio, and 3) short channel effect, such as drain-induced barrier lowering (DIBL). The proposed models were analyzed and compared by benchmarking device performance with one or more of these scaling effects.

### 2.3.2 *Novel Devices with Non-idealities*

For nanoscale devices, the source/drain series resistance is a significant component of the total on-resistance [1] due to in part the ultra shallow source/drain junction depth. The device performance was benchmarked with $R_{sd}/R_{ch\text{-}ideal}$ varying from 5% to 45%. All the

models track $R_{sd}$ variation fairly well, and all the multi point models track the effect of $R_{sd}$ variation on the device performance in a much more accurate way than the single point model, as quantified by the norm of the fitted line residuals (Figure 2.5(a)).

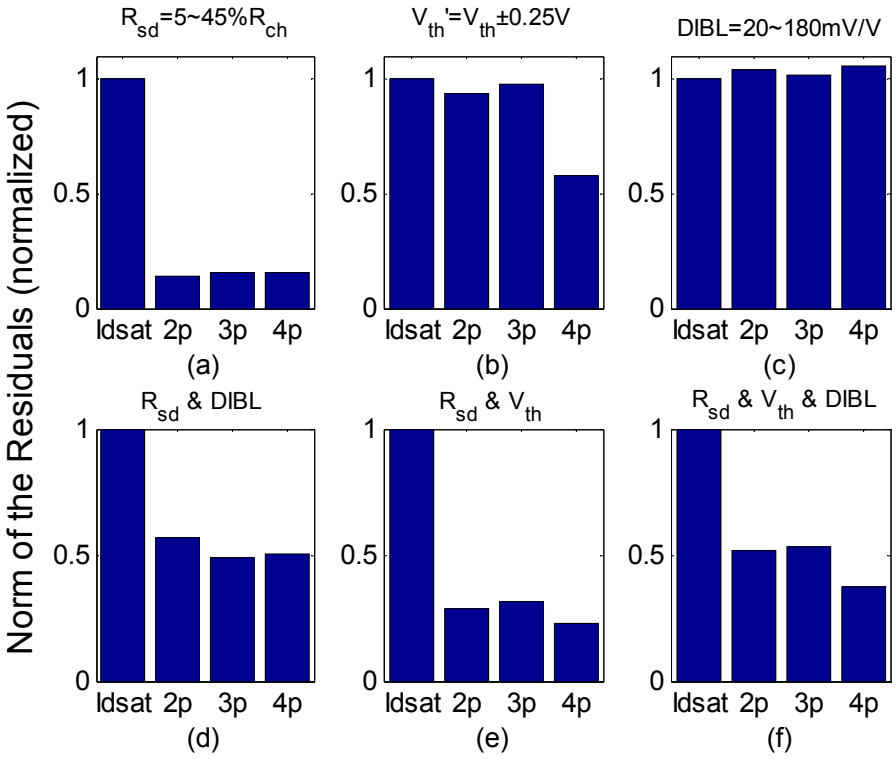



**Figure 2.5:** The norm of the residuals of the LMS fitted line of different models, with different scaling effect. (a) $R_{sd}/R_{ch}$ varies from 5% to 45%. (b) $V_{TH}$ changes within $\pm 0.25$V around the nominal value. (c) DIBL coefficient changes from 20mV/V to 180mV/V. (d) Combination of $R_{sd}$ and DIBL scaling. (e) Combination of $R_{sd}$ and $V_{TH}$ scaling. (f) Combination of $R_{sd}$, $V_{TH}$ and DIBL scaling.

Unlike devices in previous technology generations, the $V_{TH}/V_{DD}$ ratio is no longer a constant for aggressively scaled devices because $V_{DD}$ is hardly scaled after 130nm node. In circuit design, $V_{TH}/V_{DD}$ is often either intentionally decreased for higher dynamic performance or increased for better power performance. It is therefore important to study effects due to varying the $V_{TH}/V_{DD}$ ratio. The device performance was studied with $V_{TH}$ varying in a wide range, i.e. nominal $V_{TH} \pm 0.25$V.

The inverter made with devices with a small $V_{TH}/V_{DD}$ ratio has significant current flowing through the "turning off" device during 50% to 50% transition. The models using only nFET data failed to capture such current component and overestimated the effective drive current as $V_{TH}$ decreases, while the 4-point model successfully tracked the actual

device performance all along $V_{TH}$ variation range (Figure 2.5(b), Figure 2.6). This observation is particularly important for novel devices performance benchmarking, e.g. the experimental data of CNFET showed negative $V_{TH}$ for nFET and positive $V_{TH}$ for pFET. A performance comparison without considering the "turning off" current will result in an overestimated performance of CNFET, in a circuit environment.

The short channel effects include the non-ideality of I-V curve as well as $V_{TH}$ roll off. DIBL is a metric to qualify the non-ideality of the device performance caused by channel length modulation. The nominal DIBL coefficient for 90nm technology is about 90mV/V, the device performance was investigated with DIBL coefficient varying from 20mV/V to 180mV/V. All the models show similar results (within 10% error range) on benchmarking the devices performance with different DIBL effects (Figure 2.5(c)).

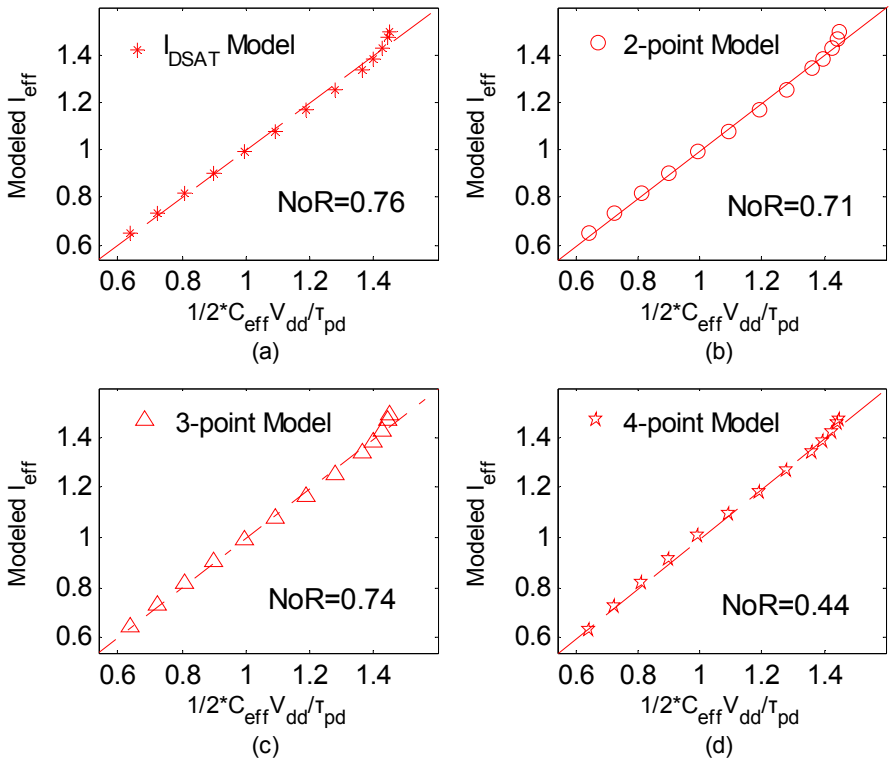



**Figure 2.6:** The modeled inverter $I_{eff}$ is compared to the actual $I_{eff}$ with $V_{TH}$ within $\pm 0.25$V around the nominal value. The LMS fitted lines are also plotted. (a) Saturation on-current model, the norm of the residuals (NoR) is 0.056. (b) 2-point model, the NoR is 0.076. (c) 3-point model, the NoR is 0.079. (d) 4-point model, the NoR is 0.043.

Practical device designs often result in more than one non-ideality in an uncorrelated fashion. To benchmark the performance of devices with different fabrication process over several technology nodes, more than one scaling effects need to be considered

simultaneously. The data distribution of the saturation on-current model is much more random and spans a much wider range than that of the other three cases where the non-idealities are varied independently (Figure 2.7). The 4-point model shows clearly the best correlation between the modeled $I_{eff}$ and the SPICE simulated $I_{eff}$ with the smallest predictive error for this set of uncorrelated device non-idealities (Figure 2.5(d~f)). The LMS fitted line slopes of the original data (2.30 for $I_{DSAT}$ model, 1.69 for 2-point and 3-point model, and 1.58 for 4-point model) also indicate that the device performance predicted by 4-point model is closest to the actual device performance.

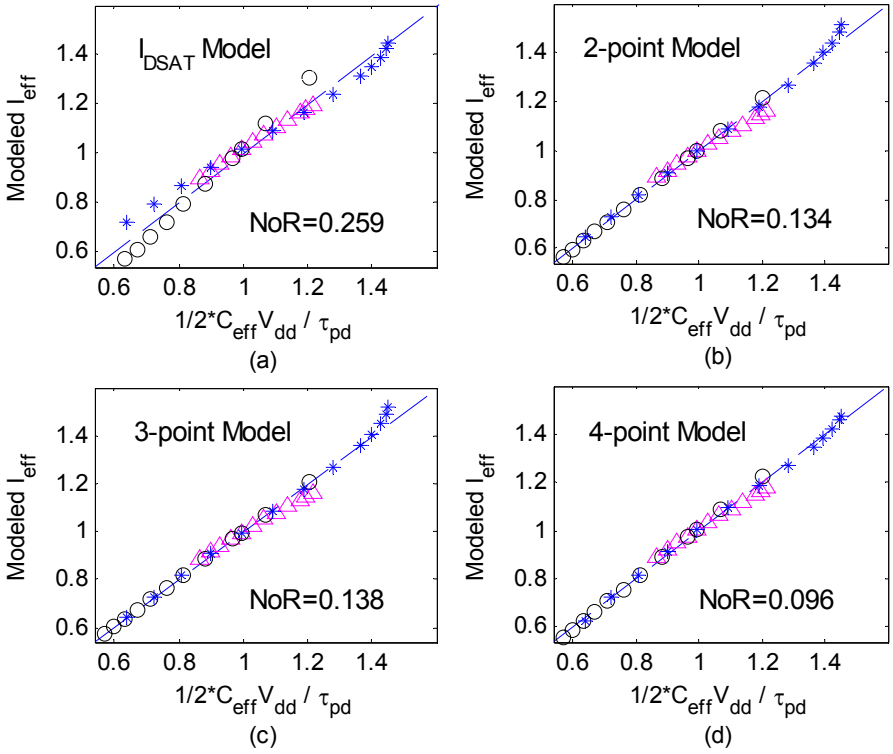



**Figure 2.7:** The modeled inverter $I_{eff}$ is compared to the actual $I_{eff}$ (normalized data) with $R_{sd}$ (circle symbol), DIBL (triangle symbol) and $V_{TH}$ (star symbol) non-idealities. One LMS fitted line of all data is also plotted. (a) Saturation on-current model. (b) 2-point model. (c) 3-point model. (d) 4-point model. (e) Comparison of the norm of the residuals between the modeled $I_{eff}$ and LMS fitted line among different models; the 4-point model has the smallest prediction error.

The analysis made above shows that the saturation on-current model and other models utilizing only nFET data are good metrics on benchmarking the performance of devices with close to ideal I-V characteristics. However, for novel devices, e.g. CNFET, with a non-ideal I-V behavior (e.g. large series/contact resistance, large $V_{Dsat}$, small $|V_{TH}|$), the 4-point model is needed. The simulation results also suggest that the accuracy of the model is not necessarily improved when the number of current points increases. The 3-

point model does not have an obvious advantage over the 2-point model (Figure 2.5). A balance exists among the simplicity, accuracy, and validity of the model. More simulations (not shown in this thesis) indicate that a simple linear model with more than 4 current points is not more accurate than the models studied here.

## 2.4 Model Application

The saturation on-current model, 2-point and 4-point model were used to predict and benchmark the device performance for various CMOS and CNFET technologies. For CMOS inverter, the nMOS and pMOS were sized to give the same rising delay and falling delay. For CNFET inverter, the current data for nFET (Lg=80nm) were taken from [12] and the current data for pFET (Lg=50nm) were taken from [5]. Since CNFET current with quasi-ballistic transport is a weak function of the channel length, the channel lengths of both CNFETs in an inverter are set to be the same without significantly affecting the drain current. A single 0.5V power supply was used for CNFET inverters. The effective gate capacitances, 8.2aF for 50nm CNFET and 13.4aF for 80nm CNFET, were calculated with a compact device model for the CNFET [13].

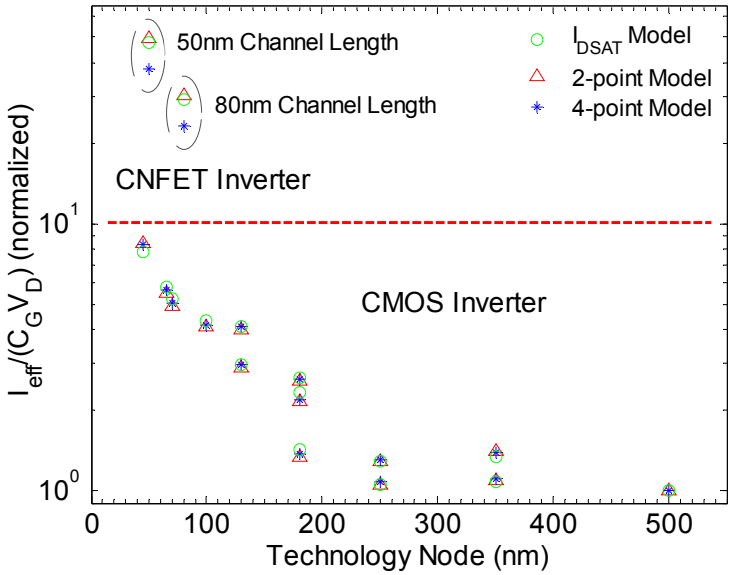



**Figure 2.8:** The modeled inverter effective drive current that is normalized to ($C_GV_D$) is plotted as a function of technology nodes for various CMOS and CNFET technologies. Single power supply is used for both CMOS and CNFET (operated under 0.5V) inverters for a fair comparison. For CNFET inverter, the current data for nFET (Lg=80nm) were taken from [12] and the current data for pFET (Lg=50nm) were taken from [5]. The channel lengths of both CNFETs in an inverter are set to be the same. The effective gate capacitances, 8.2aF for 50nm CNFET and 13.4aF for 80nm CNFET, were calculated with a compact device model for CNFET (4nm thick planar $HfO_2$ gate oxide was used in simulation) [13].

All the three models show similar results on benchmarking conventionally scaled CMOS inverters (Figure 2.8), as discussed earlier. CNFET inverters, with either 50nm channel length or 80nm channel length, show much smaller intrinsic delay than CMOS inverters with similar feather size, and the performance advantage of CNFETs over Si MOSFETs gets larger with shorter channel length because of the smaller effective gate capacitance while remaining essentially the same drain current. The saturation on-current model and 2-point model, considering only nFET current points, show the similar predictions for CNFET inverter. However, the proposed performance metric (4-point model), considering both nFET and pFET current points, indicates that the performance enhancement of CNFETs over Si MOSFETs is not as large as (around 20% smaller) that predicted by $I_{DSAT}$ and 2-point model in a circuit environment because of the non-ideal IV characteristics (Figure 2.8).

## 2.5 Summary

In summary, we found that $I_{Dsat}$ does not adequately describe inverter switching performance of nanoscale Si FETs as well as novel devices. We proposed a new metric which takes into account the following observations: (1) $I_{Dsat}$ is never reached in a switching event, (2) pFET data are also required (in addition to nFET data) to estimate the inverter speed, (3) different switching current trajectories for recent and future device generations. We proposed a 4-point model for future performance benchmarking of FETs. This new metric is very important for comparing conventional Si FETs with novel devices (e.g. CNFET and Schottky barrier metal source/drain FETs) since the I-V behavior are quite different and the non-idealities of nanoscale devices become more significant.

## 2.6 Bibliography

[1] Available at http://www.itrs.net/Links/2005ITRS/Home2005.htm

[2] K.K. Ng, C.S. Rafferty, H.I. Cong, "Effective On-Current of MOSFETs for Large-Signal Speed Considerations," *IEDM*, pp. 693-696, 2001.

[3] M.H. Na, E.J. Nowak, W. Haensch, J. Cai, "The Effective Drive Current in CMOS Inverters," *IEDM*, pp. 121-124, 2002.

[4] J. Kedzierski, P. Xuan, E.H. Anderson, J. Bokor, T-J. King, C. Hu, "Complementary silicide source/drain thin-body MOSFETs for the 20 nm gate length regime," *IEDM*, pp. 57-60, 2000.

[5] A. Javey, J. Guo, D. B. Farmer, Q. Wang, E. Yenilmez, R. G. Gordon, M. Lundstrom, H. Dai, "Self-Aligned Ballistic Molecular Transistors and Electrically Parallel Nanotube Arrays," *Nano Lett.*, vol. 4, no. 7, pp. 1319-1322, 2004.

[6] J. Appenzeller, J. Knoch, R. Martel, V. Derycke, S.J. Wind, Ph. Avouris, "Carbon Nanotube Electronics," *IEEE Trans. Nanotech.*, vol. 1, no. 4, pp. 184-189, 2002.

[7] R. Chau, S. Datta, M. Doczy, B. Doyle, B. Jin, J. Kavalieros, A. Majumdar, M. Metz, M. Radosavljevic, "Benchmarking Nanotechnology for High-Performance and Low-Power Logic Transistor Applications," *IEEE Trans. Nanotech.*, vol. 4, pp. 153-158, 2005.

[8] HSPICE® from Synopsys Corp.

[9] http://www.eas.asu.edu/~ptm/latest.html

[10] B. Yu, H. Wang, A. Joshi, Q. Xiang, E. Ibok, M-R. Lin, "15nm Gate Length Planar CMOS Transistor", *IEDM*, pp. 937-939, 2001.

[11] A. Hokazono, K. Ohuchi, M. Takayanagi, Y. Watanabe, S. Magoshi, Y. Kato, T. Shimizu, S. Mori, H. Oguma, T. Sasaki, H. Yoshimura, K. Miyano, N. Yasutake, H. Suto, K. Adachi, H. Fukui, T. Watanabe, N. Tamaoki, Y. Toyoshima and H. Ishiuchi, "14 nm Gate Length CMOSFETs Utilizing Low Thermal Budget Process with Poly-SiGe and Ni Salicide", *IEDM*, pp. 639-642, 2002.

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

[13] J. Deng and H.-S. P. Wong, "A Circuit Compatible SPICE model for Enhancement Mode Carbon Nanotube Field Effect Transistors," *2006 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD),* pp.166-169, Monterey, CA, September, 2006.

# Chapter 3

# SELECTIVE DEVICE FOOTPRINT SCALING AND PARASITIC ENGINEERING



## 3.1 Introduction

Technology boosters such as strain have helped the continuation of the historic performance trend up to the 45 nm node. As device physical gate length is reduced below 20 nm, gate length scaling becomes less effective because of the increasing contribution of parasitic capacitance [1, 2, 3]. Furthermore, the shorter gate lengths must be traded off against various leakage (subthreshold, gate, BTBT) currents. A reasonable question to ask is: if the gate length scaling stops, can we achieve performance gains through a different device scaling scenario?

In [2] the role of device pitch was explored at the device level. In this chapter, we show that even if gate length scaling slows down, we can achieve significant performance gains through aggressive scaling of the device footprint (contact sizes, overlay tolerances) selectively. The main reason is that scaling device footprint reduces the device parasitics ($C_{par}$ and $R_{ext}$) and circuit interconnect lengths, which improve speed and power efficiency at the ***circuit level***. We make the bold assumption that the sizes of various device structures (e.g. contact sizes, overlay spacings) can be arbitrarily reduced using yet-to-be developed process technologies [4, 5]. The purpose of this work is to illustrate the opportunity for a different device scaling scenario and provide the impetus for further research in process technologies to realize this new scaling scenario.

We perform device and circuit simulations for both ultra-think-body fully-depleted SOI (UTB FD-SOI) and bulk MOSFET to support our postulate. Device density, gate delay, and power consumption are analyzed for several device footprint scaling scenarios. Using a combination of 3D field solver, mixed-mode device/circuit simulations, and BSIM modeling based on actual circuit macro floor planning, we show that 25% performance improvement is obtained for a full custom 53-bit multiplier. Historic performance trend can continue for another 2 to 3 generations even without gate length scaling.

## 3.2 Contacted Gate Pitch Scaling

Contacted gate pitch ($L_{pitch}$) is the main driver for cost and performance. It has scaled along with general lithography from 1 μm through 65 nm node [6]. Here, we make the bold, yet plausible, proposal that contact sizes & overlay tolerances should be aggressively reduced (faster than general lithography) through process innovations. For example, aggressive reduction of contact sizes and overlay tolerance can potentially be achieved by self-assembly patterning techniques augmented by conventional photolithography. Block copolymer [7] (an organic material similar to photoresist) can self-organize into sub-20 nm holes that are self-registered to an existing 40 nm topography (Figure 3.1) after a 180 $^{o}$C anneal, blanket exposure to DUV, and dissolution of the PMMA polymer in a solvent, leaving a matrix of polystyrene as a soft mask [4]. The sizes of holes and $L_{gc}$ are engineered by modifying the molecular weight and composition of the polymer. For example, the PS : PMMA copolymer gives 20 nm holes and the PS : PEO : PMMA copolymer gives 10 – 15 nm holes [8].

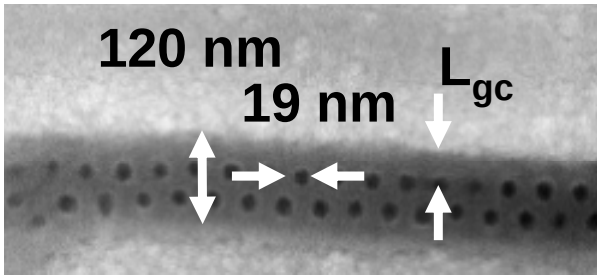



**Figure 3.1:** SEM photo of block copolymer self-assembled contact hole patterns. The holes are self-aligned to the edge of a topography 40 nm deep. Holes are 19 nm ± 1.8 nm with a pitch of 42 nm ± 2.7 nm. $L_{gc}$ is the offset of the holes from the edge. (Copyright: Liwen Chang, lwchang@stnford.edu)

## 3.3 Selective Footprint Scaling

### *(I) Reducing $L_{gc}$*

We start with a detailed analysis of the gate capacitance ($C_{gg}$) and S/D node capacitance ($C_{sd}$), including the 3D fringing capacitance (Figure 3.2) from the gate to contact plugs.

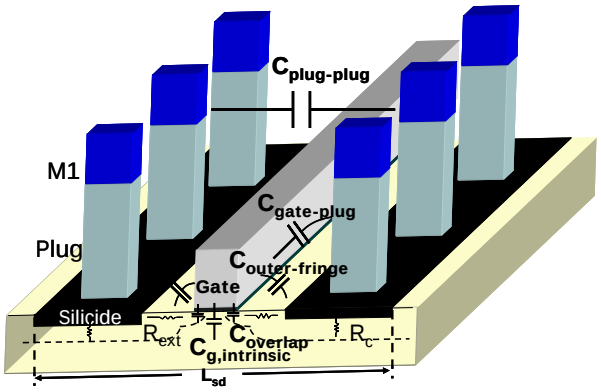



**Figure 3.2:** 3D structure of planar device with related parasitics.

Figure 3.3 shows the comparison between 2-D simulations and 3-D simulations. For long $L_{gc}$, 3-D simulation capacitance values converge with 2-D simulation values. For the shorter $L_{gc}$ (< 20 nm), due to the increased portion of the elliptical shape (the fringing portion) of the E-field, 3-D predicted capacitance values (solid lines) become much lower compared with 2-D counterparts (dashed lines). To evaluate the effect of the device scaling on device/circuit performance actually, we use 3-D simulations to calculate the capacitance in this work.

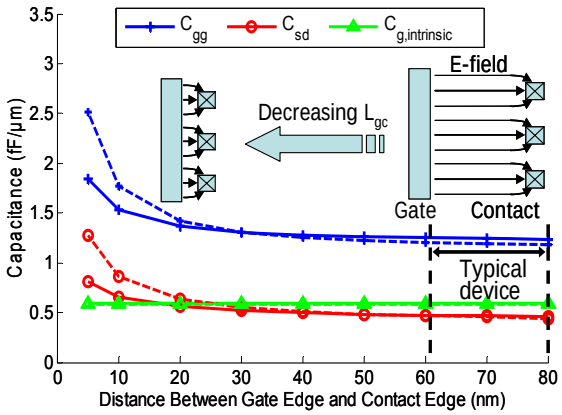



**Figure 3.3:** The capacitance comparison between 2-D simulations and 3-D simulations. Due to the increased portion of the elliptical shape of the E-field for the shorter $L_{gc}$ (<20nm), 3-D predicted capacitance values (solid lines) become much lower, compared with 2-D counterparts (dashed lines).

Both $C_{gg}$ and $C_{sd}$ increase as the gate to contact plug distance ($L_{gc}$) is reduced. The intrinsic gate capacitance is only ~50% of $C_{gg}$, the total gate capacitance. The $C_{outer\text{-}fringe}$, $C_{gate\text{-}plug}$ are responsible for ~40% (Figure 3.4).

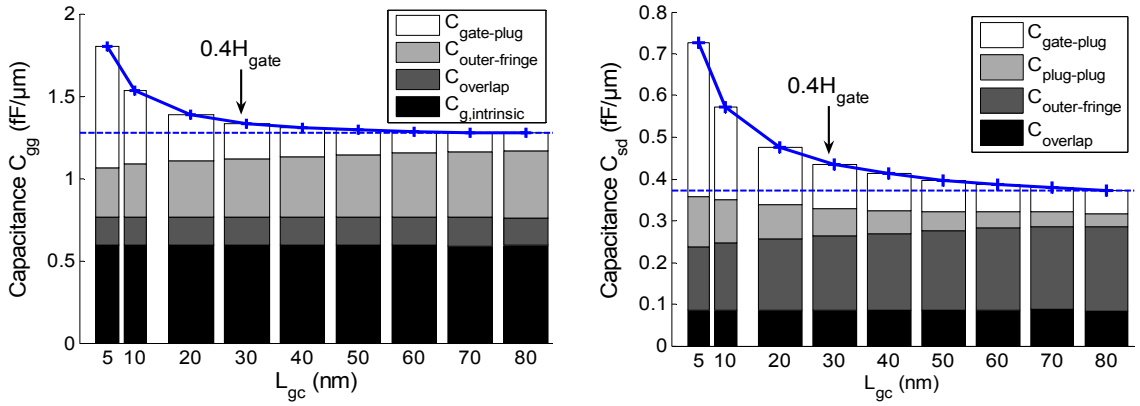
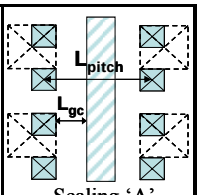
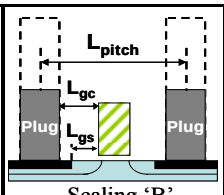
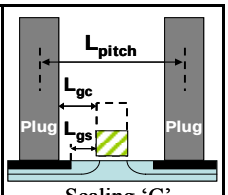



**Figure 3.4:** (a) Gate Capacitance, (b) S/D node capacitance, and their breakdown including Miller effect vs. gate to plug edge distance $L_{gc}$. The minimum $L_{gc}$ is about 0.4$H_{gate}$ to keep $C_{gg}$ ($C_{sd}$) within 8% (15%) of the nominal values for typical devices, respectively.

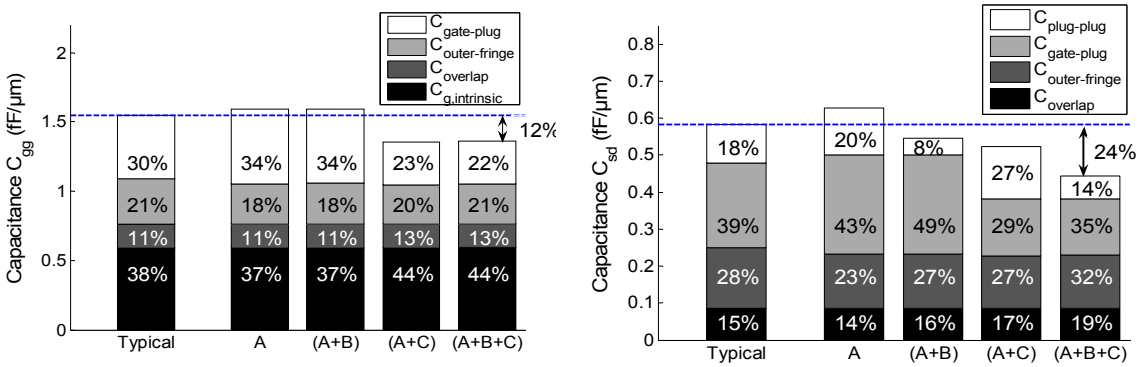
At small $L_{gc}$, the rapid increase of capacitance is due to $C_{gate\text{-}plug}$. A 2D slot approximation of the plug is pessimistic at small $L_{gc}$. To first order, the increase in $C_{gg}$ ($C_{sd}$) will be less than 8% (15%) of the nominal values for "typical" devices (designed with standard design rules) if $L_{gc}$ is larger than 0.4× gate height ($H_{gate}$). For $L_{gc} < 0.4H_{gate}$, there is significant performance loss due to increasing parasitic capacitance.

### *(II) Reducing contact size, plug height, and gate height*

**Table 3.1:** The effects of selective scaling scenarios on parasitics per unit gate width. Dashed lines denote the structures before scaling.

| |  Scaling 'A' |  Scaling 'B' |  Scaling 'C' |
|---|---|---|---|
| $C_{overlap}$ | no change | no change | no change |
| $C_{outer\text{-}finge}$ | ↓ | no change | ↓ |
| $C_{gate\text{-}plug}$ | ↑ | no change | ↓↓↓ |
| $C_{plug\text{-}plug}$ | ↑ | ↓↓ | ↑ |
| $R_c+R_{plug}$ | ↑ | ↓ | no change |
| $L_{pitch}$ | ↓ | no change | no change |

Selectively scaling the footprint (the contact size, plug height, and gate height) can further reduce the parasitic capacitance. We studied three 1-D scaling scenarios, as demonstrated by the insets in Table 3.1, and their combinations: *A) reducing the planar dimensions of S/D contacts by half, B) reducing the S/D contact plug height by half,* and *C) reducing the gate height by half*. The effect of any other 2-D or 3-D footprint scaling scenarios can be evaluated as a combination / superposition of the effects of these three 1-D scaling scenarios.




(a) (b)

**Figure 3.5:** (a) Gate capacitance, (b) S/D node capacitance, and their breakdown with different selective scaling scenarios. $L_{gc}$ is 10 nm. Scaling down the gate height is the most effective way of reducing device parasitic capacitance.

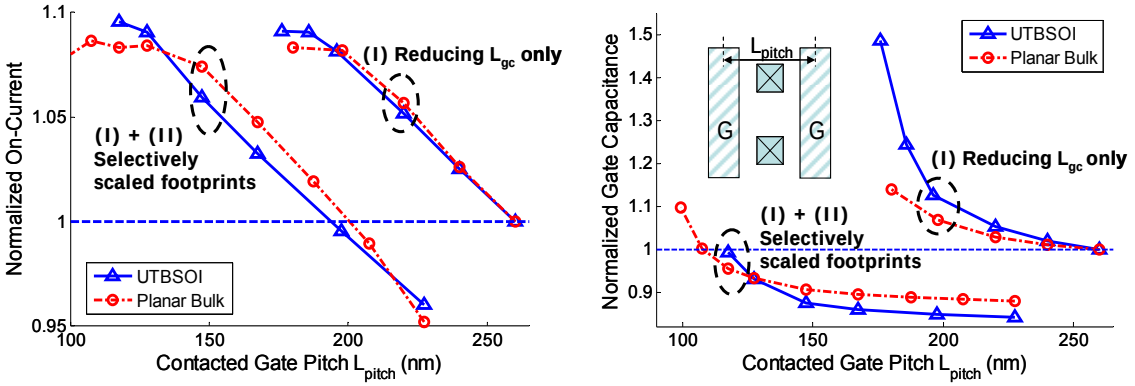
For an aggressively scaled device with $L_{gc}$ = 10 nm at the 65 nm node, Figure 3.5 shows device gate capacitance (Figure 3.5(a)), S/D node capacitance (Figure 3.5(b)), and their breakdown with different selective scaling scenarios. The effect of scaling scenarios *A*, (*A*+*B*), (*A*+*C*), and (*A*+*B*+*C*) are illustrated directly on Figure 3.5, and the effect of any other scaling scenarios can also be read out from Figure 3.5. For example, the effect of scaling scenario "C" can be read as the difference between the scaling scenario "*A*" and "*A*+*C*". As shown by Figure 3.5, scenario '*A*' does not reduce either $C_{gg}$ or $C_{sd}$, scenario '*B*' reduces $C_{sd}$ by reducing $C_{plug-plug}$, but does not reduce $C_{gg}$, reducing $H_{gate}$ (Scenario '*C*') is effective in reducing both $C_{gg}$ and $C_{sd}$.

The effects of the selective scaling scenarios on parasitics are summarized in Table 3.1. At the device level, the effectiveness of reducing $C_{par}$ is, in descending order: *(B+C) > (A+B+C) > C > (A+C) > B > (A+B) > NotScaled > A*. The 3-D device footprint scaling (*A*+*B*+*C*) is able to reduce the gate capacitance and S/D node capacitance by 12% and

24%, respectively (Figure 3.5). When interconnect capacitance at the circuit level is also considered, this order becomes: *(A+B+C) > (B+C), (A+C) > C > B > (A+B) > A > NotScaled* because a smaller device footprint helps reduce the interconnect capacitance. As a general rule, the most effective way to reduce the device parasitic capacitance is to reduce the height of the lowest components, e.g. the gate height for a planar bulk device, or the raised S/D and gate height for ultra-thin body SOI (UTBSOI). In the following sections, we show examples to illustrate how the selective scaling affects device and circuit level performance.

## 3.4 Inverter Delay Improvement

First, we analyze the behavior of the saturation on-current and the device capacitances as a function of the distance ($L_{sd}$) between the gate edge and the S/D contact stud for both planar bulk CMOS and UTBSOI in 65 nm node with the scaling scenarios: (I) reducing $L_{gc}$ only and (I) + (II) 3-D (*A+B+C*) selectively scaled footprints.




(a) (b)

**Figure 3.6:** (a) On-current, and (b) Gate capacitance vs. $L_{pitch}$ for both planar bulk CMOS and UTBSOI in 65 nm node with (I) reducing $L_{gc}$ only and (I) + (II) selectively scaled footprints.

The device on-current increases with reduced distance $L_{sd}$ between the gate edge and S/D contact stud mostly due to the smaller S/D extension resistance $R_{ex}$ (Figure 3.6(a)). The S/D contact spreading resistance has little dependence on the footprint scaling scenarios given the fact that the current transfer length is shorter than the contact length in the cases studied. The maximum on-current improvement of the scaled device over the device with the standard digital circuit design rule (termed as "typical device" in the following) is

about 10%. When evaluating dynamic performance (circuit speed), capacitance effects are also significant. 3-D device footprint scaling is able to reduce the gate capacitance by about 10% as compared to that of the typical device (Figure 3.6(b)).

A three-stage ring-oscillator (RO) and a four-stage inverter chain are used to examine the dynamic performance for both planar bulk CMOS and UTBSOI. Circuit simulations are performed using device/circuit mixed-mode numerical simulation in Taurus-Device [9]. The device width ratio between pFET and nFET is designed at 1.5 to balance the pull-up and the pull-down delay. The inverter delay ($\tau$) is evaluated between the 50% to 50% points. A trade off between the on-current and capacitances exists when sizing the dimensions of S/D region.

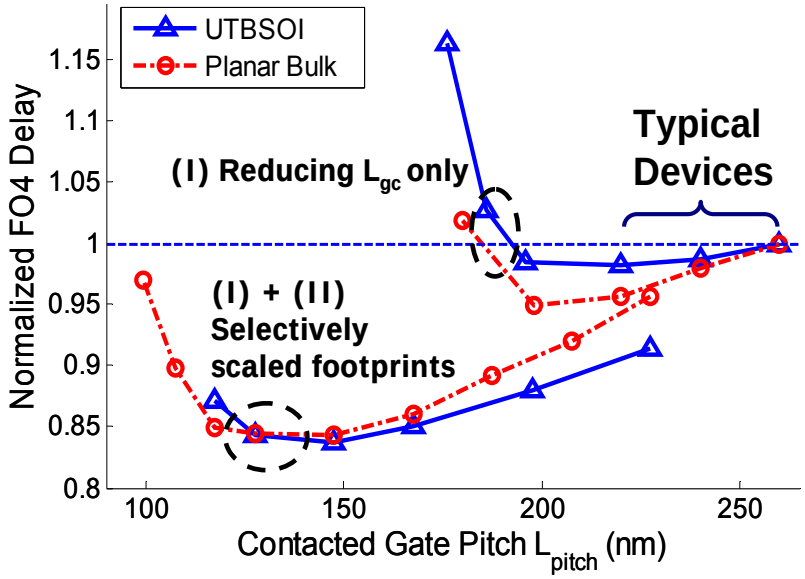



**Figure 3.7:** FO4 delay vs. $L_{pitch}$ for both planar bulk CMOS and UTBSOI with (I) reducing $L_{gc}$ only and (I) + (II) selectively scaled footprints.

Up to 5% higher speed can be obtained by simply pushing the gate to silicide distance $L_{gs}$ ($L_{gc}$=$L_{gs}$+$\Delta$) smaller (Figure 3.7). With 3-D selective footprint scaling (*A*+*B*+*C*), the speed is about 15% faster for both the planar bulk device and UTBSOI, compared with devices with standard layout rules ($L_{sd}$ = 12 $\lambda$, Figure 3.2). At the design point with minimum FO4 delay, the on-current improvement over typical device is about 7%, and the total device length is about $L_{sd0}$ = 6.6 $\lambda$ which is 45% smaller (isolated device) layout area than the typical device. In the following section, we use an optimal value of $L_{sd}$ = 7 $\lambda$

for the following reasons: (1) the delay is sensitive to $L_{sd}$ variation with $L_{sd}$ less than $L_{sd0}$, but is rather flat in the region with $L_{sd}$ slightly larger than $L_{sd0}$; (2) it is more convenient to set up the design rule to be a multiple number of λ; (3) a relaxed design with smaller parasitic capacitance helps to reduce dynamic power consumption with very minor speed/area penalty.

The reduced junction capacitance for bulk device with selective scaling has trivial impact (<3%) on speed because $C_{plug\text{-}plug}$, $C_{gate\text{-}plug}$, $C_{outer\text{-}fringe}$ and $C_{overlap}$ are the largest components of $C_{sd}$. Mobility degradation due to reduced stress of the smaller active area (~50MPa less stress by reducing $L_{gc}$ by 65 nm) is smaller than 4.5% [10, 11] which corresponds to less than 3% on-current degradation.

## 3.5 Circuit-Level Improvement

The previous section presented results for small circuits with very little wiring load. For larger circuit structures, the interconnect delay is comparable to the gate delay. The smaller device dimension permits the use of shorter interconnects, hence results in higher speed and lower power consumption at the circuit macro level. A fully custom designed 53-bit multiplier using a 7-metal layer technology is used to evaluate the benefits of device footprint scaling on complex circuits.

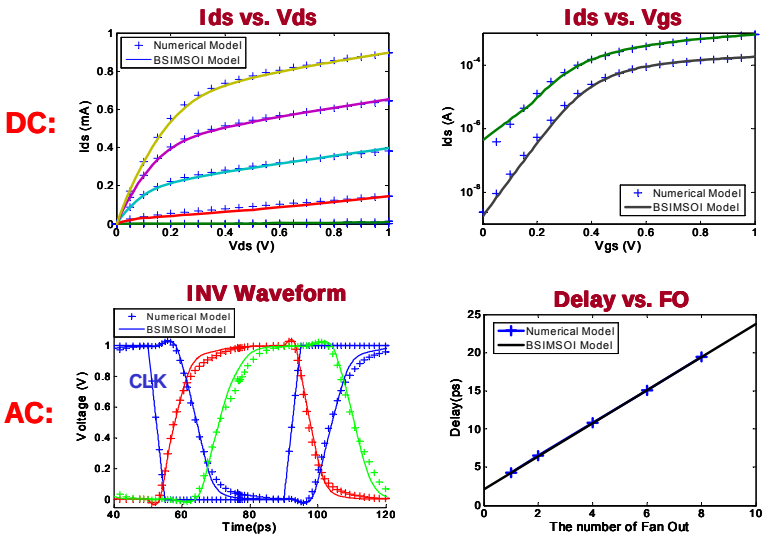



**Figure 3.8:** The simulation results comparison between the numeric device model and extracted BSIM3vSOI compact model in both dc (Ids vs. Vds, and Ids vs. Vgs) and ac (waveform, and delay vs. FO) characteristics.

Since it is impossible to perform mixed-mode simulation for such a large circuit, we first extract device parameters from the Taurus-Device simulator into the BSIM3vSOI compact model using Aurora [12]. Figure 3.8 shows good agreement between the numerical device-simulation results (Taurus) and the extracted BSIM3vSOI compact model for both dc and ac characteristics.

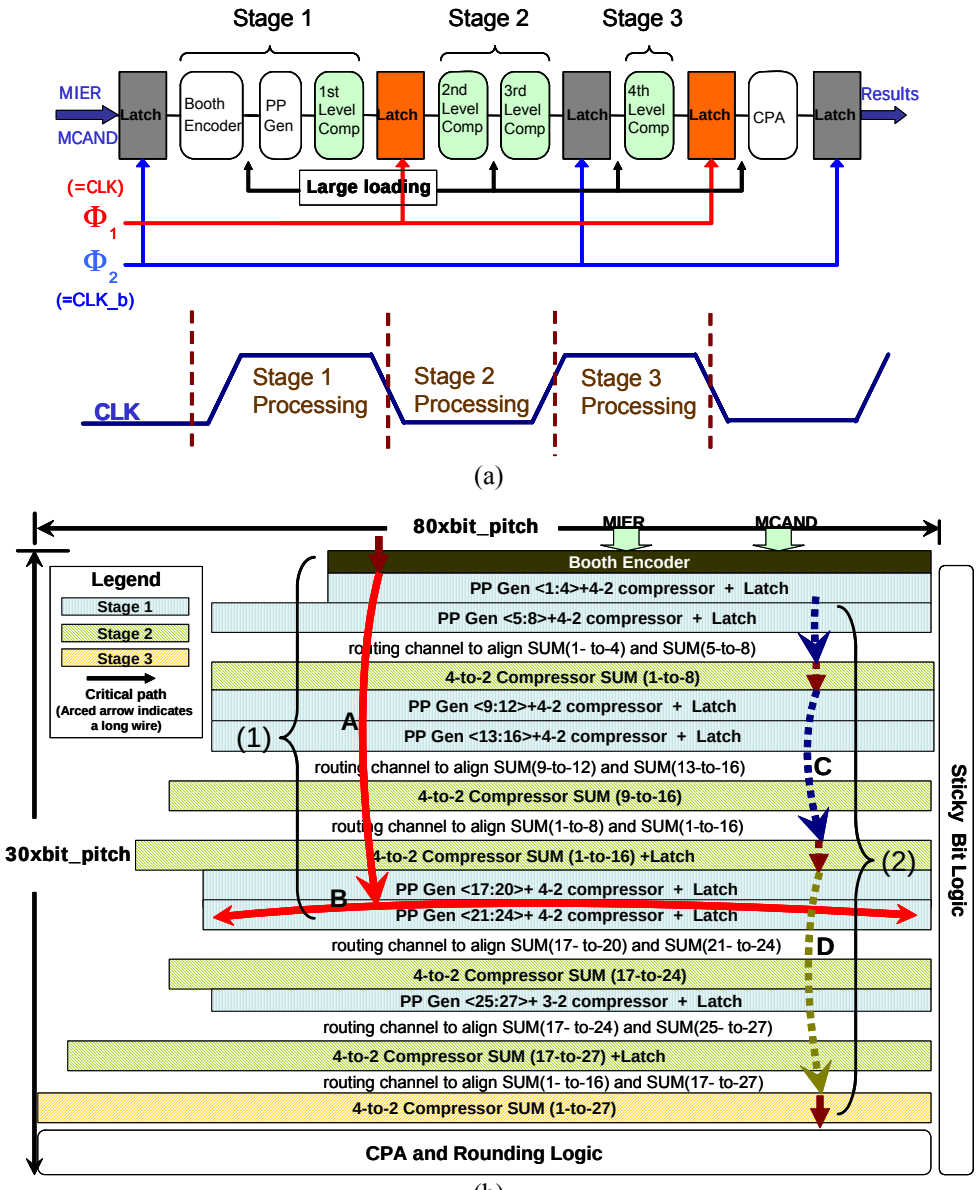



**Figure 3.9:** (a) The logic architecture, and (b) The illustration of chip floor plan for the fully custom designed 53-bit pipelined multiplier. There are two critical paths with long routing interconnects: (1) a long vertical wire (segment A) and a long horizontal wire (segment B) between the outputs of the Booth encoder and the inputs of the partial-product generators in the first stage, and (2) a long vertical wire (segment C) between the second level and the third level Wallace tree within the second stage, and a long vertical wire (segment D) between the latch outputs of the second stage and the inputs of the fourth level Wallace tree in the third stage.

Pipelined dual-rail dynamic domino logic is chosen for the circuit design. We use the Booth algorithm to calculate the partial product, and all the partial products are summed up with a three-stage pipelined Wallace tree (Figure 3.9(a)). The latency time is two clock cycles and the throughput is one output per cycle. With the folded chip floor plan, there are two critical paths (Figure 3.9(b)) with long routing interconnects: (1) (solid arced arrow in Figure 3.9(b)) a long vertical wire (segment A) and a long horizontal wire (segment B) between the outputs of the Booth encoder and the inputs of the partial-product generators in the first stage, and (2) (dashed arced arrow in Figure 3.9(b)) a long vertical wire (segment C) between the second level and the third level Wallace tree within the second stage, and a long vertical wire (segment D) between the latch outputs of the second stage and the inputs of the fourth level Wallace tree in the third stage. All the interconnect capacitance (about 0.2 fF/μm) and resistance extracted from the chip floor plan and the routing map are included in the SPICE simulation. A minimum positive setup time of 15 ps for latches is guaranteed in evaluating the maximum operation frequency. The circuit layout area can be expressed as the summation of the device layout area and the routing area, where single device layout area is given by $\alpha \cdot Area_{gate}$. The parameter α is the ratio of the total device length to the gate length (6 for the typical device and 3.5 for the selectively scaled device). The routing areas for the circuits built with both the typical device and the optimal device are approximately the same with the same interconnect technology and the same design rule for the metal layers (i.e. the same minimum width and minimum pitch for metal layers).

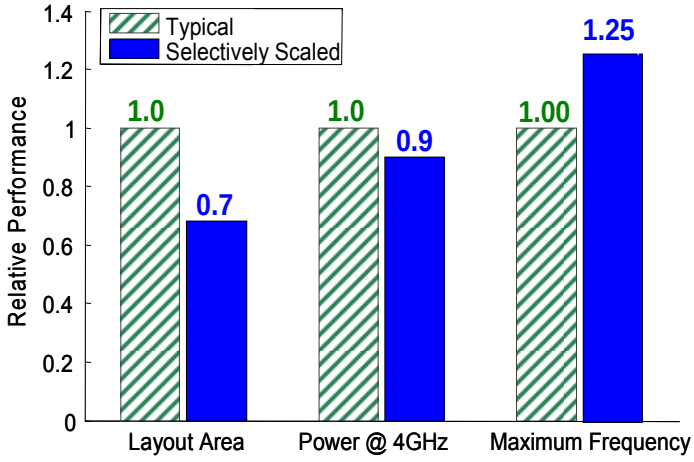



**Figure 3.10:** The performance comparison between the multipliers made with typical devices (shaded) and that made with selectively scaled devices (solid).

As shown by Figure 3.10, compared with the multiplier made with typical devices ($L_{sd}$=12λ), the multiplier built with the selectively scaled devices ($L_{sd}$=7λ) occupies 30% less layout area, operates at 25% higher speed (~15% is due to the faster device, another 10% comes from the shorter interconnects), and consumes 10% less dynamic power due to the smaller interconnect capacitance of the smaller circuit layout area.

## 3.6 Variability Analysis

One major reason that prohibits the further device scaling is the performance variation caused by the increased process variations beyond 65 nm technology node. Thus it is important to investigate the performance variability for the optimal device. We use UTBSOI as an example to study the effect of selective footprint scaling on device variability. We assume a super-halo profile [13], and the number of dopants in channel region is assumed to be constant. We use Monte-Carlo simulation to study the effects of three major process related variations, gate length ($L_{poly}$) variation, body thickness ($T_{body}$) variation, and contact size (both the contact length $L_{cnt}$ and the height of raised source/drain region $H_{cnt}$) variation, on both dc ($V_{th}$) and ac (FO1 delay) characteristics. These variations are treated as independent processes. The parameters and simulation results are summarized in Table 3.2.

**Table 3.2:** Vth and FO1 delay variability analysis for both typical device and optimal device with gate length ($L_{poly}$) variation, body thickness ($T_{body}$) variation, and contact size (both contact length $L_{cnt}$ and the height of raised S/D region $H_{cnt}$) variation (assuming 10% / 3σ variation). Both $L_{cnt}$ and $H_{cnt}$ scale in the same direction to represent the worse case for contact size variation.

| | | **Vth (mV)** | | | **FO1 Delay (ps)** | | |
|---|---|---|---|---|---|---|---|
| | | Nominal Value | Mean Value | 1σ Variation (% of mean value) | Nominal Value | Mean Value | 1σ Variation (% of mean value) |
| Typical Device | $L_{poly}$ | 174.2 | 173.7 | 4.7 (2.71%) | 4.10 | 4.11 | 0.093 (2.26%) |
| | $T_{body}$ | 174.2 | 173.8 | 11.3 (6.50%) | 4.10 | 4.13 | 0.087 (2.10%) |
| | $L_{cnt}/H_{cnt}$ | 174.2 | 174.2 | < 0.1 (0.06%) | 4.10 | 4.12 | 0.025 (0.61%) |
| | **Total** | **174.2** | **173.9** | **12.24 (7.04%)** | **4.10** | **4.12** | **0.130 (3.15%)** |
| Optimal Device | $L_{poly}$ | 175.1 | 174.6 | 4.7 (2.69%) | 3.70 | 3.68 | 0.097 (2.64%) |
| | $T_{body}$ | 175.1 | 174.8 | 11.4 (6.52%) | 3.70 | 3.71 | 0.072 (1.93%) |
| | $L_{cnt}/H_{cnt}$ | 175.1 | 176.0 | 1.2 (0.70%) | 3.70 | 3.69 | 0.031 (0.85%) |
| | **Total** | **175.1** | **175.1** | **12.39 (7.08%)** | **3.70** | **3.69** | **0.125 (3.38%)** |

Though the footprint scaled device is a little more sensitive (~0.2%) to the contact size variation because of the smaller contact size than the typical device, the effect of contact size variation is secondary compared to the effects of $L_{poly}$ variation and $T_{body}$ (for UTBSOI) variation. The scaled footprints do not increase the performance variation in terms of both dc and ac characteristics (Table 3.2). The average values are close to the nominal values. FO1 delay is less sensitive to $T_{body}$ variations than $V_{th}$. Using a first order approximation, the delay due to the inner fringing capacitance can be expressed as

$$\tau_{fr} \propto R_{on} \cdot C_{gsd,fr} \propto \frac{1}{T_{body}} \cdot T_{body} \sim \text{Constant}$$

The weak dependence of FO1 delay on body thickness is due in part to the cancellation of the gate to S/D inner channel fringing capacitance of the load devices and the on resistance of the driving device for FD-SOI.

## 3.7 Extending the Technology Roadmap

In addition to the limitations of parasitic capacitance on device/circuit performance, the parasitic series resistance also has significant impact on the device speed for future techniques. As illustrated in Figure 3.11, scaling the length of the S/D silicide ($L_{silicide}$) below the current transfer length causes rapid increase of contact resistance $R_c$. In order not to degrade the on-current for aggressively scaled devices, a relaxed silicide length in source/drain contact regions has to be used.

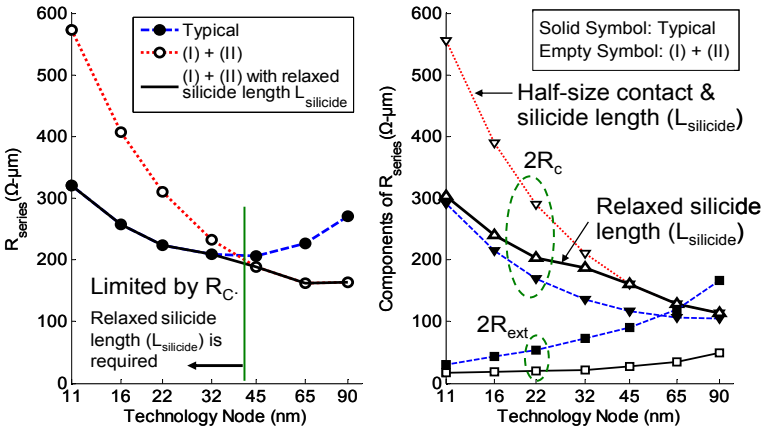



**Figure 3.11:** $R_{series}$ (left) and its components (right) as a function of technology nodes. $R_c$ increase quickly with half size footprint scaling. A relaxed contact silicide length ($L_{silicide}$) should be used for future technology nodes in order not to degrade on-current.

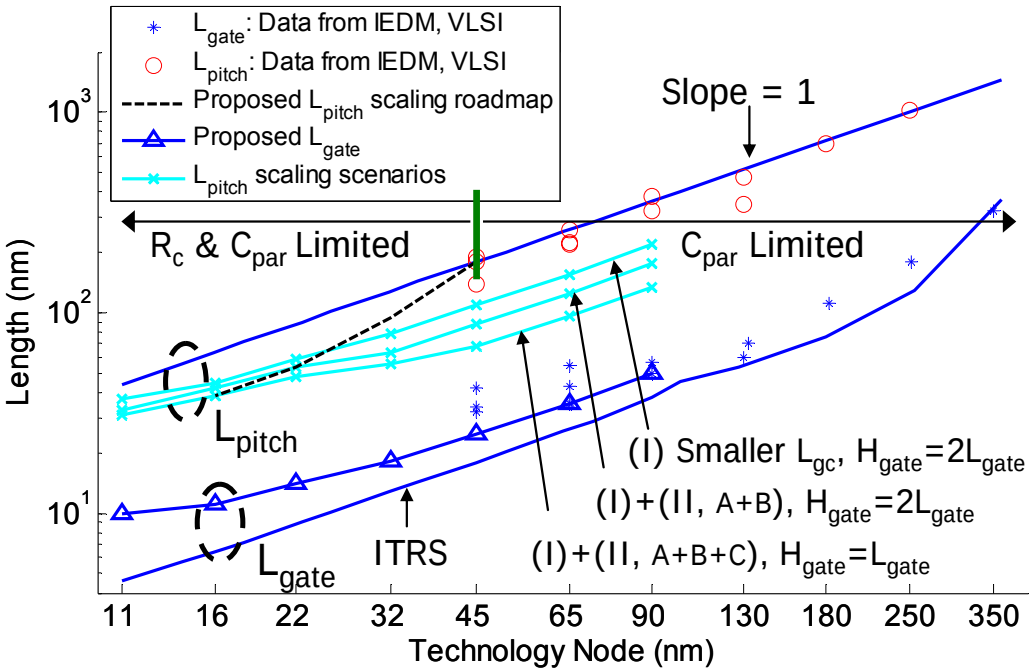



**Figure 3.12:** Contacted gate pitch ($L_{pitch}$) and physical gate length ($L_{gate}$) vs. technology node down to 11 nm node. The minimum $L_{pitch}$ (lines with '×' symbol) is bounded by either parasitic capacitance ($C_{par}$) or contact resistance ($R_c$) or both. By reducing gate height and selective footprint scaling, the historic performance trend can continue for another 2 to 3 generations even without gate length scaling. The suggested technology scaling path is denoted by dashed curve.

Figure 3.12 shows a scaling scenario in which aggressive $L_{pitch}$ scaling compensates for the slower than 0.7× per node $L_{gate}$ scaling. With selective footprint scaling (reducing contact size and $L_{gc}$) along with reduced gate and plug height, the technology roadmap can be extended to the 11 nm node with physical gate length no shorter than 10 nm. $L_{pitch}$ scaling is bounded by parasitic capacitance and contact resistance. Figure 3.12 marks the $L_{pitch}$ boundary within which the device on-current is equal to or larger than the values for "typical" devices with zero or trivial $C_{par}$ penalty. By reducing gate height and selective footprint scaling, the historic performance trend can continue for another 2 to 3 generations even without gate length scaling. The suggested technology scaling path is denoted by dashed curve for a smooth technology transition.

## 3.8 Conclusions

In this chapter, we propose a new device scaling scenario for sub-65 nm technology node high performance CMOS technology. We postulate that even with the gate length

remaining the same, selectively scaling the device footprint will provide significant circuit-level performance improvement from technology generation to technology generation. For small benchmark circuits such as ring oscillators and inverter chains, a selectively scaled device with reduced footprint achieves 45% smaller device layout area, 7% higher drive current, and 15% shorter delay than the typical (standard) device. For a fully custom designed 53-bit multiplier, the selectively scaled device with reduced footprint occupies 30% less layout area, operates at 25% higher frequency, and consumes 10% less dynamic power while keeping the static power dissipation the same.

Ever since the 0.35 μm node, the gate length has been selectively scaled to improve performance. Gate length selective scaling has been effective when the gate delay of unloaded (or lightly loaded) circuits were dominated by the intrinsic gate capacitance. Beyond the 65 nm technology node, the physical gate length is less than 1/10 of that at 0.35 μm node, but the parasitic gate to source/drain capacitance/resistance remains roughly the same. Therefore, further aggressive gate length scaling does not bring the same benefits as it did when the gate lengths were longer. At the same time, interconnect delay becomes a larger fraction of the total circuit delay. Therefore, a new selective scaling scenario where the parasitic capacitance/resistance and the interconnect delay are minimized is effective. The analyses of this chapter suggest that this new selective scaling scenario can achieve significant performance improvement. Historic performance trend can continue for another 2 to 3 generations even without gate length scaling.

The extended scaling path requires tight pitch patterning, tight overlay tolerances, and a short gate height processes – potential yield limiters. Novel nanofabrication techniques, such as self-assembly [4, 5] and low-*k* spacer, are needed to realize the substantial benefits offered by $L_{pitch}$ scaling and parasitics engineering. Additionally, packaging and/or architectural solutions are also required to mitigate the increased power density due to the smaller device footprint. Future work on this subject may further consider the effects of stress-dependent carrier mobility [14], low-*k* isolation dielectric (low-*k* STI), simultaneous optimization of footprint, and device width.

## 3.9 Bibliography

[1] S. E. Thompson and Srivatsan Parthasarathy, “Moore's law: the future of Si microelectronics,” *Materials Today*, vol. 9, pp. 20-25, 2006.

[2] J. W. Sleight, I. Lauer, O. Dokumaci, D. M. Fried, D. Guo, B. Haran, S. Narasimha, C. Sheraw, D. Singh, M. Steigerwalt, X. Wang, P. Oldiges, D. Sadana, C.Y. Sung, W. Haensch, M. Khare, “Challenges and Opportunities for High Performance 32 nm CMOS Technology, ” *IEEE International Electron Devices Meeting (IEDM)*, pp. 697-700, San Francisco, CA, December 11-13, 2006.

[3] Ali Khakifirooz, Dimitri A. Antoniadis, “Transistor Performance Scaling: The Role of Virtual Source Velocity and Its Mobility Dependence,” *IEEE International Electron Devices Meeting (IEDM)*, pp. 667-670, San Francisco, CA, December 11-13, 2006.

[4] L.-W. Chang, H.-S. P. Wong, “Diblock Copolymer Directed Self-Assembly for CMOS Device Fabrication,” *31st SPIE International Symposium on Microlithography, in Design and Process Integration for Microelectronic Manufacturing IV., edited by A. K. K Wong, V. K. Singh, Proceedings of the SPIE,* Vol. 6150, pp. 329-334, 2006.

[5] A. Hatzor, P. Weiss, “Molecular Rulers for Scaling Down Nanostructures,” *Science*, vol. 291, pp. 1019-1020, 2001.

[6] H.-S. P. Wong, G.S. Ditlow, P.M. Solomon, X. Wang, “Performance Estimation and Benchmarking for Carbon Nanotube FETs and Nanodiode Arrays,” invited paper, *International Conferences on Solid State Devices and Materials (SSDM)*, Japan, September 16 – 18, pp. 802 – 803, 2003.

[7] C. T. Black, K. W. Guarini, R. Ruiz, E. M. Sikorski, I. V. Babich, R. L. Sandstrom, Y. Zhang, “Polymer self assembly in semiconductor microelectronics,” *IEEE International Electron Devices Meeting (IEDM)*, pp. 439-442, San Francisco, CA, December 11-13, 2006.

[8] J. Bang, S. H. Kim, E. Drockenmuller, M. J. Misner, T. P. Russell, C. J. Hawker, “Defect-Free Nanoporous Thin Films from ABC Triblock Copolymers,” *J. Am. Chem. Soc.* Vol. 128, p.7622-7629, 2006.

[9] *Taurus-Device*®, Version 2005.10, Synopsys Corp., CA.

[10] S. Thompson, G. Sun, K. Wu, J. Lim, T. Nishida, “Key differences for process-induced uniaxial vs. substrate-induced biaxial stressed Si and Ge channel MOSFETs,” *IEEE International Electron Devices Meeting (IEDM)*, pp. 221-224, San Francisco, CA, December 13-15, 2004.

[11] N. Shah, M.S. Thesis, U. Florida, 2005.

[12] *Aurora*®, Version 2002.2.0, Avant! Corp. and TMA, Inc., CA.

[13] Y. Taur, C. Wann, D. J. Frank, “25 nm CMOS Design Considerations,” *IEDM Tech. Digest*, pp. 789 –792, 1998.

[14] V. Chan, R. Rengarajan, N. Rovedo, W. Jin, T. Hook, P. Nguyen, J. Chen, E. Nowak, X.–D. Chen, D. Lea, A. Chakravarti, V. Ku, S. Yang, A. Steegen, C. Baiocco, P. Shafer, H. Ng, S.-F. Huang, C. Wann, "High Speed 45nm Gate Length CMOSFETs Integrated Into a 90nm Bulk Technology Incorporating Strain Engineering," *IEDM Tech. Digest*, pp. 77-80, 2003.

# Chapter 4

# MODELING AND ANALYSIS OF 1-D FET PLANAR GATE CAPACITANCE

A version of this chapter has been submitted for publication. J. Deng and H.-S P. Wong, "Modeling and Analysis of Planar Gate Electrostatic Capacitance for 1-D FET with Multiple Cylindrical Conducting Channels," *IEEE Transactions on Electron Devices*, to appear, 2007.

## 4.1 Introduction

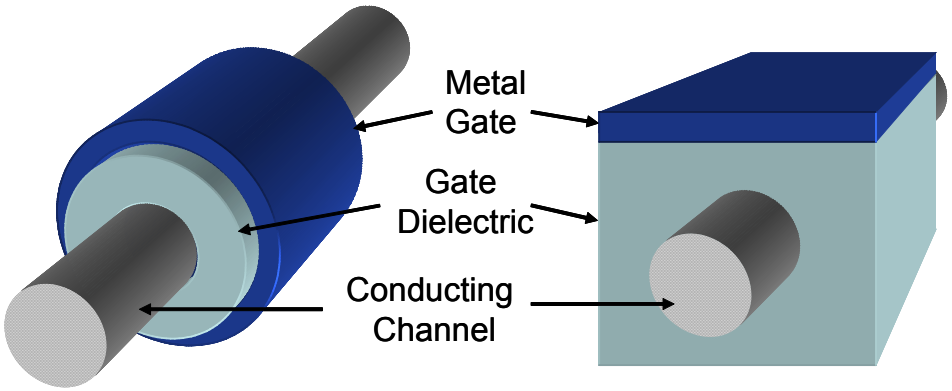



**Figure 4.1:** (a) Gate-All-Around (GAA) structure, and (b) Planar gate structure with single channel in a uniform gate dielectric material.

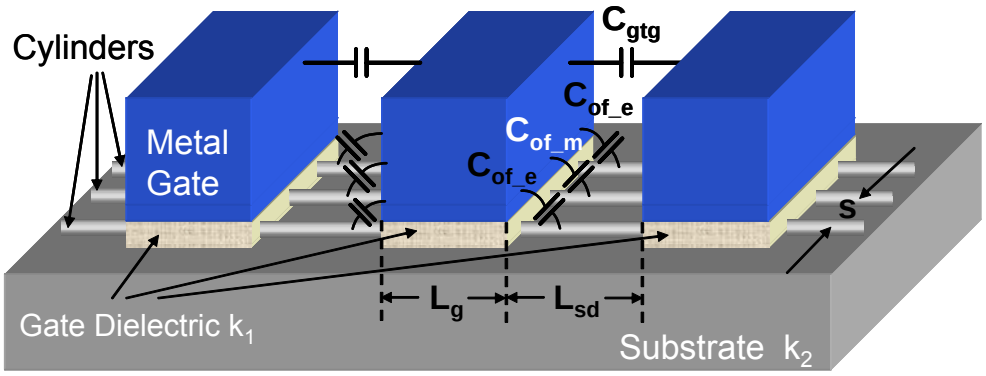
In recent years, 1-D field effect transistors (FETs) with cylinder conducting channels (e.g. semiconductor nanowire FET [1-3], carbon nanotube FET [4-8], and tri-gate transistors [9]) are proposed and reported due to the better electrostatic performance over planar devices (bulk CMOS) and SOI. Either a gate-all-around (GAA) gate structure or a planar gate structure with single conducting channel in a uniform dielectric material (Figure 4.1) was typically used to evaluate the gate capacitance and device performance [10-12].

However, for the devices with high-*k* gate dielectric whose permittivity is not the same as the substrate permittivity, 10% to 40% error will be incurred by these approximations. In addition, due to the small drive current that can be delivered by a single channel, multiple conducting channels per gate are usually required for 1-D FET to achieve competitive performance over the traditional silicon devices [6, 13]. For a device with multiple conducting channels, a planar gate structure with high-*k* gate dielectric material is a realistic structure. In the limit of ballistic or near-ballistic transport, the drive current of 1-D device highly depends the gate to channel capacitance ($C_{gc}$). The parallel conducting channels have screening / imaging effect on the actual potential profile in the gate region, and therefore affect the capacitance. Previous work treated the screening effect for calculating $C_{gc}$ using 2-D numerical modeling [14, 15]. However there have no reported analytical models available to offer insights for device design. In addition to $C_{gc}$, the device speed also strongly depends on the parasitic gate capacitance, including the outer-fringe gate capacitance ($C_{of}$) and the gate to gate (source/drain) coupling capacitance ($C_{gtg}$). It is very important to model the various components of total gate capacitance ($C_{gg}$) with reasonable accuracy in order to evaluate and predict the 1-D FET circuit performance (on-current, speed, and power). The contributions of this work are two fold: (1) we present an analytical model of the gate capacitance (including screening and fringing field effects) that can be incorporated in a compact model such as SPICE [16]; (2) using this simple analytical model for the gate capacitance, one can, for the first time, obtain a realistic estimate of circuit performance including the parasitic capacitance and screening effect from multiple parallel channels.

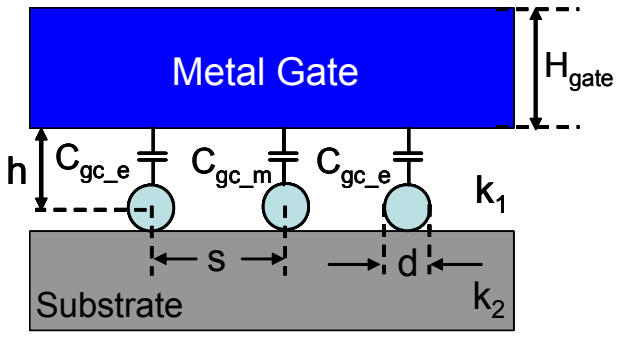
## 4.2 Gate Capacitance Modeling

In this work, we consider a planar gate structure with multiple cylindrical conducting channels and high-*k* gate dielectric material on a substrate with a different dielectric constant as illustrated in Figure 4.2. Multiple devices may be connected in series (e.g. in a NAND structure). The diameter of the cylinder is '*d*'. The normal distance between the gate and the cylinder center is denoted by '*h*', and the distance between the centers of the two adjacent parallel conductors is denoted by '*s*'. To be general, we start from a

description of the methodology to calculate the capacitance including the screening effect of the neighboring conductors.




(a)




(b)

**Figure 4.2:** (a) The 3-D structure of the devices with multiple channels and high-*k* gate dielectric material, and the related parasitic gate capacitances. (b) The 2-D view of the cross section in the channel region and the related gate to channel capacitance.

The important capacitances to be modeled are

(1) Gate to channel capacitance, $C_{gc}$

(2) Outer fringe capacitance, $C_{of}$

(3) Gate to gate, or gate to source/drain coupling capacitance, $C_{gtg}$

For $C_{gc}$, the image charge across the dielectric boundary should be taken into account. The outer fringe capacitance is a strong function of the device geometry. Both $C_{gc}$ and $C_{of}$ are strongly affected by screening of neighboring channels especially for closely spaced channels which provide large current drive per unit device width. This screening effect

must be properly accounted for in the model. $C_{gtg}$ cannot be simply modeled by a parallel plate approximation because the parallel plate dimension is comparable to the other 3-D geometrical dimensions. In the following sections, we derive simple analytical equations to model these three capacitances. We verify their accuracy by comparing the results with numerical solutions from a 3-D field solver [17].

### 4.2.1 *Capacitance Model for the Object in an Array*

Consider the structure in Figure 4.3(a). We represent the parallel conducting channels as $N$ ($N \geq 1$) identical objects in parallel, and the gate electrode as the planar electrode #0. In order to calculate the coupling capacitance $C_{01}$ between the electrode #0 and the object #1, the total effects of the other ($N$-1) objects around the object #1 on $C_{01}$ can be lumped and approximated as the two nearest objects, #2 and #3 (Figure 4.3(b)), because the objects with a large distance from object #1 have rather weak influence on the electric field distribution between electrode #0 and #1. Approximating the ($N$-*1*) objects with more than two objects may achieve better accuracy at the cost of a more complex equation for the general case.

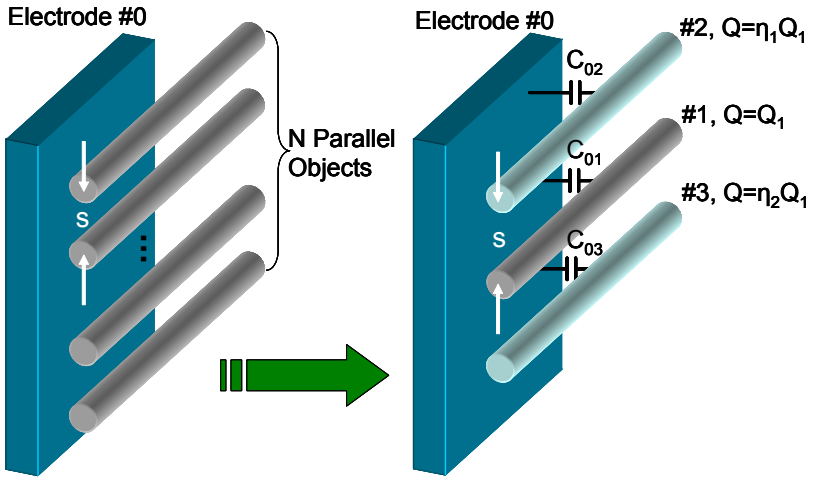



**Figure 4.3:** (a) There are $N$ identical objects in parallel in an array. (b) In order to calculate the coupling capacitance $C_{01}$, the effects of the ($N$-1) objects around the object #1 can be lumped into the two nearest objects #2 and #3. $C_{02}$ and $C_{03}$ are the equivalent capacitances assuming all the other ($N$-1) objects are lumped at the position of #2 and #3.

Applying the same voltage $V_1$ between the objects #1, #2, #3 and the electrode #0, an amount of charges $Q_1$, $\eta_1 Q_1$, and $\eta_2 Q_1$ are induced on the three objects #1, #2, and #3,

respectively, due to the different coupling capacitance $C_{01}$, $C_{02}$, and $C_{03}$. $C_{0i}$ is the equivalent coupling capacitance between the electrode #0 and the object #i. We can define $\eta_1$ and $\eta_2$ as the ratio of $C_{02}$ and $C_{03}$, respectively, over $C_{01}$, given by,

$$\eta_1 = \frac{C_{02}}{C_{01}} \qquad \eta_2 = \frac{C_{03}}{C_{01}} \tag{4.1}$$

The charges on objects #2 and #3 will affect the electric field and electrostatic potential profile between the electrode #0 and the object #1. For a 1-D device geometry, we ignore the charge redistribution in the circumferential direction. Using the superposition principle, the capacitance $C_{01}$ can be expressed as:

$$C_{01} = \frac{Q_1}{V_1} = \frac{Q_1}{V_o + V_{adj}} \tag{4.2}$$

$V_o$ and $V_{adj}$ are the potential differences between the electrode #0 and the object #1 caused by the charges on the object #1 and the adjacent objects (#2 and #3), respectively, acting as independent electrodes. Normalized to the same amount of charge $Q_1$, $V_{adj}$ is rewritten as:

$$V_{adj} = \eta_1 \cdot V_{adj,1} + \eta_2 \cdot V_{adj,2} \tag{4.3}$$

$V_{adj,1}$ and $V_{adj,2}$ are the potential differences between the electrode #0 and the object #1 caused by the objects #2 and #3, respectively, with the same amount of charge $Q_1$. With equations (4.2) and (4.3), we obtain,

$$C_{01} = \frac{1}{\frac{V_o}{Q_1} + \eta_1 \cdot \frac{V_{adj,1}}{Q_1} + \eta_2 \cdot \frac{V_{adj,2}}{Q_1}} = \frac{1}{\frac{1}{C_{\inf}} + \eta_1 \cdot \frac{1}{C_{sr,1}} + \eta_2 \cdot \frac{1}{C_{sr,2}}} \tag{4.4}$$

The capacitance $C_{01}$ is a series combination of the capacitance $C_{\inf}$, $C_{sr,1}/\eta_1$, and $C_{sr,2}/\eta_2$. $C_{\inf}$ is the capacitance between the electrode #0 and the object #1 without the screening of all other objects. $C_{sr,1}$ and $C_{sr,2}$ are the equivalent capacitances due to the screening effects of the objects #2 and #3, respectively. $\eta_1$, $\eta_2$ are functions of the geometry, the number of the objects of the array, and the position of the object in the array. In some particular cases, equation (4.4) can be simplified.

For the objects at the ends of an array, the screening objects are only at one side, therefore $\eta_2$=0. This capacitance, $C_e$, can be expressed as:

$$C_e = \frac{C_{sr,1} \cdot C_{\inf}}{C_{sr,1} + \eta_1 \cdot C_{\inf}} \tag{4.5}$$

For the objects around the middle of an array, because the geometry is symmetric around the object ($\eta_1 \cong \eta_2$ and $C_{sr,1} \cong C_{sr,2}$), the capacitance between the electrode and the object in the middle, $C_m$, can be approximated as

$$C_m = \frac{C_{sr,1} \cdot C_{\inf}}{C_{sr,1} + 2\eta_2 \cdot C_{\inf}} \tag{4.6}$$

With equations (4.5) and (4.6), we can eliminate $C_{sr,1}$ and express $C_m$ as a function of $C_e$ and $C_{\inf}$,

$$C_m = \frac{\eta_1 \cdot C_e \cdot C_{\inf}}{2\eta_2 \cdot C_{\inf} + (\eta_1 - 2\eta_2) \cdot C_e} \tag{4.7}$$

If we denote $\eta_2 = \alpha \cdot \frac{C_e}{C_m}$ (the reason will be discussed next), then $C_m$ is given by,

$$C_m = \frac{2\alpha}{\eta_1} \cdot C_e + (1 - \frac{2\alpha}{\eta_1}) \cdot C_{\inf} \tag{4.8}$$

The summation of the coefficients of $C_e$ and $C_{\inf}$ equals to 1. For a given geometry, the parameter $\eta_1$ is a function of the number of objects per array, and the parameter $\alpha$ is a function of both the number of objects per array and the position of the object in the array.

There are two special cases with which we can determine the parameters $\eta_1$ and/or $\alpha$ directly.

**Case 1:** For an array with only two objects, the two objects are identical and both objects are at the end. From equation (4.1), we know $\eta_1$=1. There is only one component of capacitance denoted by $C_e$, given by:

$$C_e = \frac{C_{sr,1} \cdot C_{\inf}}{C_{sr,1} + C_{\inf}} \tag{4.9}$$

**Case 2:** For an array with only three objects, there are two components of capacitances: the capacitance between #0 and the two objects at the ends ($C_e$) and the capacitance between #0 and the one object in the middle ($C_m$). When calculating $C_e$, we approximate the two objects #1 and #3 as one object in the position of object #1 as in Figure 4.3. By the definition of $\eta_2$ and $\alpha$, we know $\alpha$=1. Thus $C_e$ and $C_m$ are given by,

$$\begin{aligned} C_e &= \frac{C_{sr} \cdot C_{\inf}}{C_{sr} + \eta_1 \cdot C_{\inf}} \\ C_m &= \frac{2}{\eta_1} \cdot C_e + (1 - \frac{2}{\eta_1}) \cdot C_{\inf} \end{aligned} \tag{4.10}$$

We will be able to calculate $C_e$ and $C_m$ once we know the expressions of $\eta_1$, $C_{\inf}$ and $C_{sr}$. In the following sections, we derive $\eta_1$, $C_{\inf}$ and $C_{sr}$ for different gate regions along the channel length direction.

### 4.2.2 *Gate to Channel Capacitance*

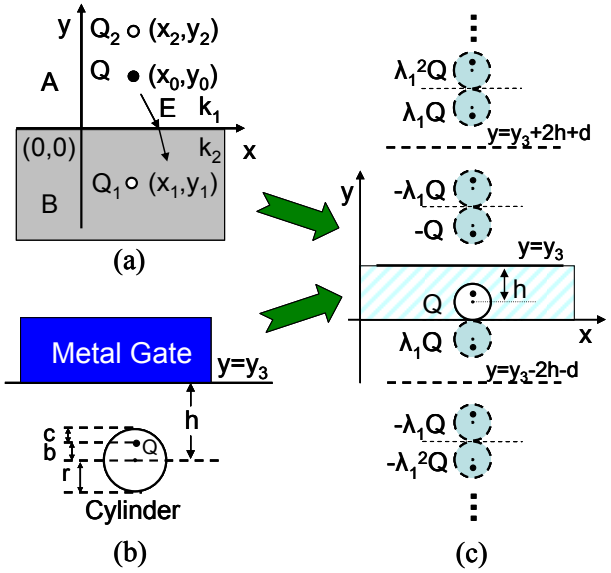



**Figure 4.4:** (a) One real charge $Q$ in dielectric material $k_1$, and the related image charges $Q_1$ and $Q_2$ due to $k_1 \neq k_2$. (b) The charge profile of one conducting cylinder under an infinite large metal gate in a uniform dielectric material $k_1$. (c) Combining the case (a) and (b), in order to calculate the potential profile in the shaded region. There is an infinite number of image charges (dashed circles) due to metal gate screening and the interface between $k_1$ and $k_2$.

In this section, we calculate the gate to channel capacitance per unit length ($C_{gc}$) for the planar gate structure with high-$k$ gate dielectric material and multiple parallel conducting cylindrical channels (Figure 4.2). We ignore the end effect in the axial direction, i.e. the length of the cylindrical channel is much longer than the diameter. We also assume that the gate width ($W_{gate}$) is much larger than the channel diameter.

$\boldsymbol{C_{gc,inf}}$ : First, we calculate $C_{gc,\mathrm{inf}}$, the capacitance between the gate and a single isolated cylinder with diameter '$d$'. First, consider the case where there is no metal gate on top of the cylinder. For an arbitrary charge $Q$ at ($x_0$, $y_0$), we consider two image charges $Q_1$ at ($x_1$, $y_1$), $Q_2$ at ($x_2$, $y_2$), as a first approximation (Figure 4.4(a)). The interface between the two dielectric materials is along the line $y$=0. Applying the following boundary conditions:

1) The tangential component of the electric field along the boundary is continuous, i.e. $E_{\|}(x, 0^+)= E_{\|}(x, 0^-)$

2) The normal component of the electrical displacement field across the boundary is continuous, i.e. $D\perp(x, 0^+)= D\perp(x, 0^-)$,

the solutions are given by,

$$\begin{aligned} &x_1 = x_2 = x_0 \qquad y_1 = -y_0 \qquad y_2 = y_0 \\ &Q_1 = \lambda_1 \cdot Q \qquad \lambda_1 = \frac{k_1 - k_2}{k_1 + k_2} \\ &Q_2 = \lambda_2 \cdot Q \qquad \lambda_2 = \frac{2k_2}{k_1 + k_2} \end{aligned} \tag{4.11}$$

$\lambda_1$ and $\lambda_2$ are the pre-factors that account for the interface due to $k_1 \neq k_2$. Thus the positions of the two image charges are symmetric across the interface $y$=0. $k_1$ and $k_2$ are the relative permittivity of the dielectric materials in region A and region B, respectively. To calculate the electrostatic properties in region A, only the image charge $Q_1$ in region B is required, we therefore ignore $Q_2$ in the following analysis.

Next, we consider the case of a metal gate on top of the cylinder along the interface $y=y_3$, and a uniform dielectric material ($k_1=k_2$) (Figure 4.4(b)). The relationships between the geometry parameters are,

$$b = h - \sqrt{h^2 - r^2} \qquad c = r - h + \sqrt{h^2 - r^2} \tag{4.12}$$

where $r$ is the radius of the cylinder.

Combining the above two geometries, there will be an infinite number of image charges in the whole space, as shown in Figure 4.4(c) due to the gate mirroring effect and the refection across the two interfaces $y=0$ and $y=y_3$. To simplify the analysis, we assume the charge distribution profile around the cylinder is not changed by the interface between the two dielectric materials. Mathematically, there are 4 image line charges in each group,

$$Q_{imag\_m,i} = (-1)^{i+1} \cdot \lambda_1^{\,m} \cdot Q \qquad i = 1,2,3,4 \qquad m = 1,2,3,... \tag{4.13}$$

The potential drop between the cylinder and the metal gate caused by the $m^{th}$ image line charges group is given by,

$$V_{imag\_m} = \sum_{i=1}^{4} V_{imag\_m,i} = \frac{(-1)^{m+1}}{2\pi k_1 \varepsilon_o} \cdot \lambda_1^{\,m} \cdot Q \cdot \ln\left( \frac{(2mh+md)^2}{(2mh+md)^2 - (2h-2b)^2} \right) \tag{4.14}$$

The gate to channel capacitance can be expressed as,

$$C_{gc\_\inf} = \frac{Q}{V} = \frac{Q}{V_o + \sum_{m=1}^{\infty} V_{imag\_m}} = \frac{1}{\frac{1}{C_{gco}} + \frac{1}{C_{gc\_imag}}} \tag{4.15}$$

Where $C_{gco}$ is the capacitance when $k_1=k_2$, and $C_{gc\_imag}$ is the equivalent series capacitance caused by the image charges when $k_1 \neq k_2$, given by,

$$C_{gco} = \frac{2\pi k_1 \varepsilon_o}{\cosh^{-1}\left(\frac{2h}{d}\right)} \tag{4.16}$$

$$C_{gc\_imag} = \frac{Q}{\sum_{m=1}^{\infty} V_{imag\_m}} = \frac{2\pi k_1 \varepsilon_o}{\sum_{m=1}^{\infty} (-1)^{m+1} \cdot \lambda_1^{\,m} \cdot \ln\left( \frac{(2mh+md)^2}{(2mh+md)^2 - (2h-2b)^2} \right)} \tag{4.17}$$

For a typical carbon nanotube field-effect-transistor (CNFET) device (1.5 nm in diameter ($d$=1.5nm) conducting channel, 4 nm thick ($h$=4nm) $HfO_2$ gate dielectric material ($k_1$=16) with $SiO_2$ substrate ($k_2$=3.9)), $C_{gco}$ is about 377 pF/μm and $C_{gc\_imag}$ is about 1380 pF/μm. Therefore $C_{gc\_inf}$ is 296 pF/μm using equation (4.15) which is within 2% of the numerical simulation result (~302 pF/μm), while the simple model $C_{gco}$ (Equ. 4.16) used in the literature [11, 12] overestimates the capacitance by 26%.

To simply the above equation, we approximate the effects of all the image line charges as one image line charge $\lambda_1 Q$ at (0, -$r$), the image capacitance is simplified as,

$$C_{gc\_imag} = \frac{2\pi k_1 \varepsilon_o}{\lambda_1 \cdot \ln\left(\frac{2h+2d}{3d}\right)} \tag{4.18}$$

The discrepancy between the capacitance $C_{gc\_inf}$ calculated by both analytical models (Equ. (4.15-4.18)) and the numerical simulation is within 2% (Figure 4.5). The results indicate that lumping the effects of all the image charges due to $k_1 \neq k_2$ as one image line charge $\lambda_1 Q$ at (0,-$r$) is a good approximation.

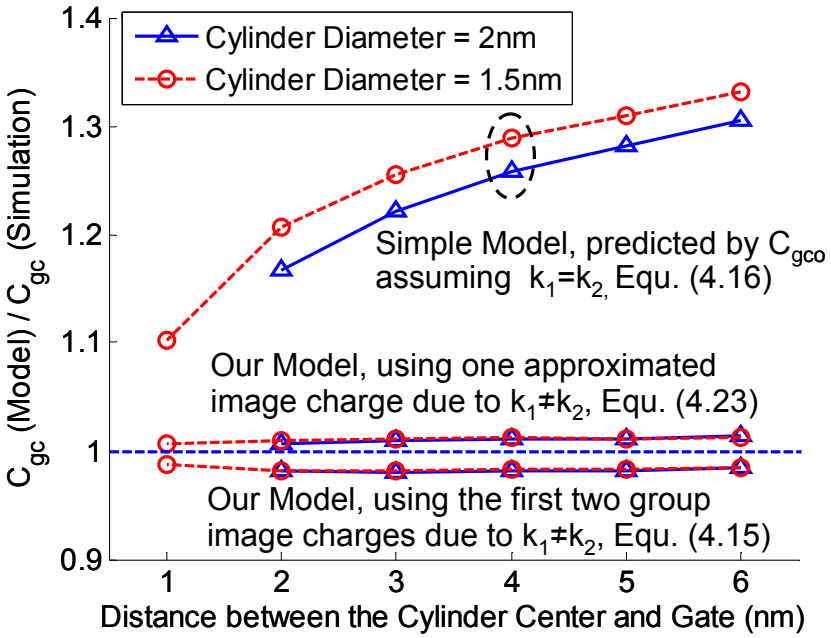



**Figure 4.5:** The comparison between the analytic models and the numerical simulations to calculate the gate to channel capacitance $C_{gc\_inf}$ with isolated cylinder channel. The simple model in Figure 4.1(b) overestimates $C_{gc\_inf}$ by about 10% to 35%. Our model using either one approximated image charge or the first two group (8) image charges due to $k_1 \neq k_2$ achieves 2% or better accuracy.

***$C_{gc\_e}$* and *$C_{gc\_m}$*:** Next, we consider the effects of the adjacent parallel cylinders on $C_{gc}$ for a more general gate structure and derive the expressions for $C_{gc\_e}$ and $C_{gc\_m}$. Consider a gate with two conducting channels, cylinder A and cylinder B, in parallel (Figure 4.6). Only two image line charges with each real line charge $Q$ is considered: the image line charge $-Q$ due to metal gate mirroring, and the image line charge $\lambda_1 Q$ due to $k_1 \neq k_2$ (Figure 4.6).

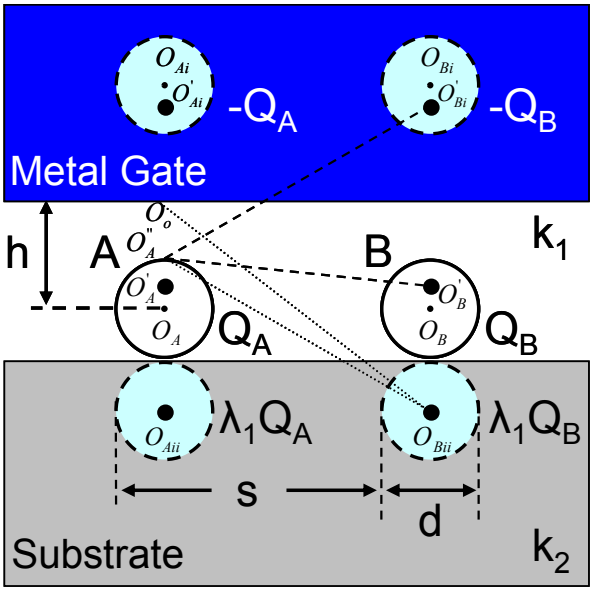



**Figure 4.6:** Two parallel cylinders A and B are under the same metal gate. The images charges are denoted by the dashed circles. The effects of the infinite number of image charges in Figure 4.5(c) due to the gate screening and the interface between $k_1$ and $k_2$ are approximated as one image charge $\lambda_1 Q_A$ ($\lambda_1 Q_B$). The charge distribution profile for $Q_A$($Q_B$) and -$Q_A$(-$Q_B$) are assumed not to be affected by the image line charge $\lambda_1 Q_A$ ($\lambda_1 Q_B$).

Due to the screening effect, in addition to the potential difference caused by the isolated cylinder A itself, there are additional potential drops between cylinder A and the gate caused the real and image line charges by cylinder B. The additional potential difference caused by the line charge $Q_B$ and the image line charge -$Q_B$ is given by,

$$V_{adj,1} = \frac{Q_B}{2\pi k_1 \varepsilon_o} \ln\left(\frac{\overline{O'_{Bi}O''_A}}{\overline{O'_B O''_A}}\right) = \frac{Q_B}{4\pi k_1 \varepsilon_o} \ln\left(\frac{s^2 + 2(h-r)\cdot[h+\sqrt{h^2-r^2}]}{s^2 + 2(h-r)\cdot[h-\sqrt{h^2-r^2}]}\right) \tag{4.19}$$

The additional potential difference caused by the image line charge $\lambda_1 Q_B$ is,

$$V_{adj,2} = f(h,r,s)\frac{\lambda_1 Q_B}{2\pi k_1 \varepsilon_o}\ln\left(\frac{\overline{O_{Bii}O_o}}{\overline{O_{Bii}O_A^{"}}}\right) = f(h,r,s)\frac{\lambda_1 Q_B}{4\pi k_1 \varepsilon_o}\ln\left(\frac{(h+d)^2+s^2}{9r^2+s^2}\right) \qquad (4.20)$$

The function $f(h, r, s)$ models the charge redistribution effects when the two cylinders are close enough. Due to the Coulomb interaction between adjacent charges, the displacement of the equivalent charge position from the center of the cylinder is of the form $\tanh(\mu_o)$ where $\mu_o$ is a function of the geometry [18]. Based on this observation, we empirically represent $f(h, r, s)$ as a tanh() function of the ratio of the vertical distance over the horizontal distance between these conductors,

$$f(h,r,s) = \tanh\left(\frac{h+r}{s-d}\right) \qquad (4.21)$$

When the same potential is applied to the two parallel cylinders, the charge distribution profiles in the regions around cylinder A and cylinder B should be identical. The equivalent series capacitance due to the adjacent channel screening is then given by,

$$C_{gc\_sr} = \frac{Q_B}{V_{adj,1} + V_{adj,2}} \qquad (4.22)$$

Since the electric field is well confined within the gate dielectric for the typical gate structure, we assume the cylinders which are more than '$s$' distance apart away have a minor effect on each other, i.e. the gate to channel capacitance does not depend on the number of the cylinders in the array. Applying $\eta_1$=1 (recall Equ. 4.10), with equations (4.10, 4.11, 4.15-4.22), the capacitances per unit length in the channel region are,

$$C_{gc\_e} = \frac{C_{gc\_\inf}\cdot C_{gc\_sr}}{C_{gc\_\inf} + C_{gc\_sr}} \qquad C_{gc\_m} = 2C_{gc\_e} - C_{gc\_\inf}$$

$$C_{gc\_sr} = \frac{4\pi k_1 \varepsilon_o}{\ln\left(\frac{s^2+2(h-r)\cdot[h+\sqrt{h^2-r^2}]}{s^2+2(h-r)\cdot[h-\sqrt{h^2-r^2}]}\right) + \lambda_1\cdot\ln\left(\frac{(h+d)^2+s^2}{9r^2+s^2}\right)\cdot\tanh\left(\frac{h+r}{s-d}\right)} \qquad (4.23)$$

$$C_{gc\_\inf} = \frac{2\pi k_1 \varepsilon_o}{\cosh^{-1}\left(\frac{2h}{d}\right) + \lambda_1\cdot\ln\left(\frac{2h+2d}{3d}\right)} \qquad \lambda_1 = \frac{k_1-k_2}{k_1+k_2}$$

$C_{gc\_e}$ is the unit capacitance of the gate to the cylinders at the two ends, and $C_{gc\_m}$ is the unit capacitance of the gate to the cylinders in the middle. The difference between the values calculated by the analytic model and the numeric 3-D field solver simulation is within 10% with various parameter settings and different number of channels per gate (Figure 4.7). The discrepancy comes from two sources: 1) we approximate all the image charges due to $k_1 \neq k_2$ with a single image line charge, 2) the gate capacitance is only classified into two groups, $C_{gc\_e}$ and $C_{gc\_m}$, which are assumed to be independent of the number of channels per gate. More accurate results can be obtained by carefully choosing the parameters $\eta_1$ and $\alpha$ for each channel under the gate.

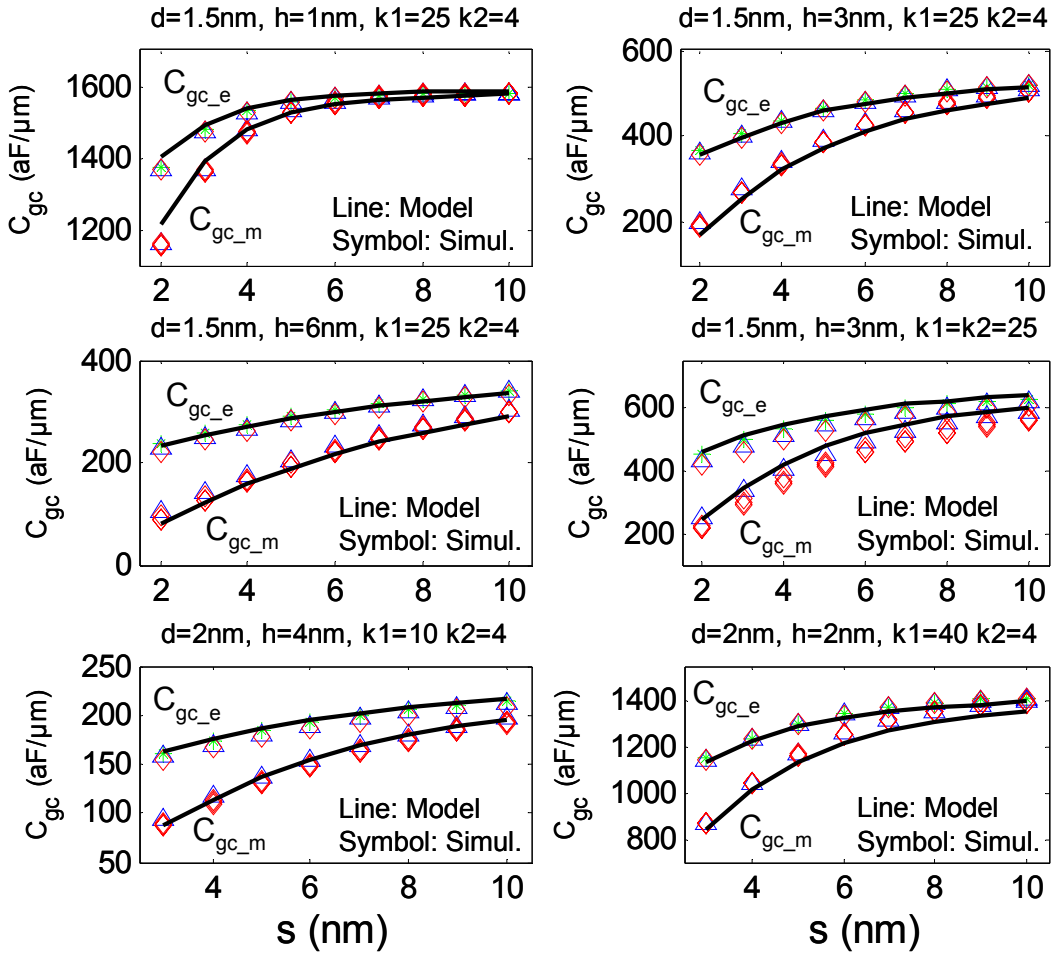



**Figure 4.7:** The comparison between the analytic model (solid lines) and the numerical simulation results (symbols) as a function of the distance between the adjacent channels ($s$). The gate to channel capacitance ($C_{gc}$) with different number of channels ($N$=2, 3, 7) per gate are plotted in the plot. $C_{gc}$ converges into two groups: the capacitance between the gate and the two channels at the ends ($C_{gc\_e}$) and the capacitance between the gate and the channels between the ends ($C_{gc\_m}$). 10% accuracy is obtained with various parameter settings.

### 4.2.3 *Gate Outer-Fringe Capacitance*

To evaluate the device speed accurately, it is necessary to include the parasitic gate capacitance, e.g. the outer-fringe capacitance ($C_{of}$), and the gate to gate (or gate to source/drain) coupling capacitance ($C_{gtg}$)[+] (Figure 4.2). Since the channel region is screened by the conducting cylinders, the inner fringe capacitance is ignored and uniform dielectric material with relative permittivity $k_2$ is assumed (Figure 4.8). For 1-D FET, the coupling capacitance between the gate and the source/drain due to the "sidewall fringing field" is more significant than the coupling capacitance due to the "normal fringing field", as illustrated by Figure 4.9 inset, because of the small diameter of 1-D channel compared to the gate width. For 1-D FET with $L_{sd}$ shorter than the gate interconnect length in the gate width direction, which is well satisfied for typical device design, $C_{of}$ is almost independent of the gate height and gate width, as shown by the 3-D numerical simulations (Figure 4.9). For the device with 32 nm long gate length ($L_g$), reducing the gate height ($H_{gate}$) from 64 nm to 10 nm results in less than 10% difference in $C_{of}$ in the range of $L_{sd} < W_{gate}$. Thus it is reasonable and convenient to assume $C_{of}$ is independent of $H_{gate}$ and $W_{gate}$ in order to benefit from the simplicity provided $L_{sd} < W_{gate}$.

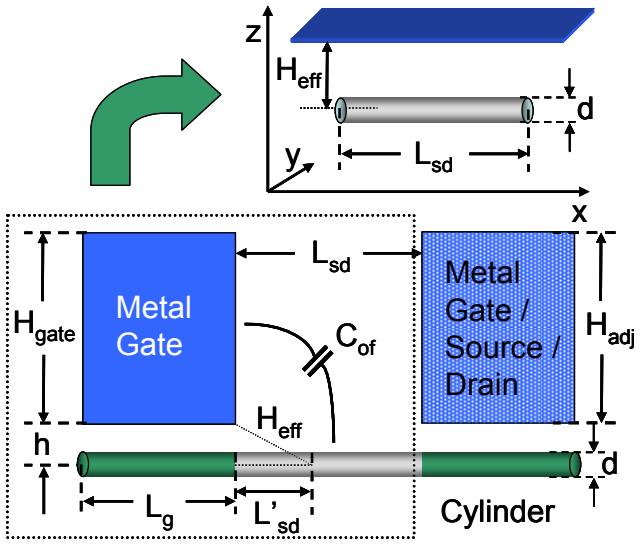



**Figure 4.8:** The structure to calculate the gate outer-fringe capacitance. The inset shows a parallel system with equivalent distance $H_{eff}$ between the cylinder and the plate electrode in order to calculate the outer-fringe capacitance $C_{of}$.

[+] To simplify the notation, we use $C_{gtg}$ even for the case of gate to source/drain capacitance. The two cases are identical from the electrostatics point of view.

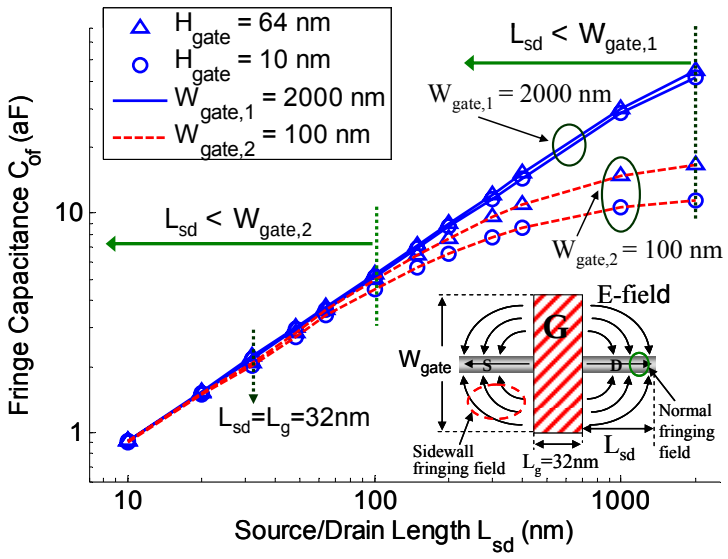



**Figure 4.9:** The outer-fringe capacitance ($C_{of}$) as a function of the source/drain length ($L_{sd}$) calculated with 3-D numerical simulations. Inset shows the top view of the simulated device structure and the related electric filed profile. For 1-D FET with $L_{sd} < W_{gate}$, $C_{of}$ is almost independent of the gate height ($H_{gate}$) and gate width ($W_{gate}$).

First we calculate $C_{of\_inf}$, the capacitance between the gate and the isolated S/D cylinder. It is convenient to calculate the capacitance if the cylinder is parallel with the sidewall of the gate, as shown by the inset in Figure 4.8. For a 2-D structure, it is possible to convert from an elliptical system to an equivalent parallel system using conformal mapping, while it is hard for a 3-D structure. We define an equivalent distance between the S/D cylinder and the sidewall of the gate as $H_{eff}$, in the form of,

$$H_{eff} = \sqrt{h^2 + (\gamma \cdot L_{sd})^2} \tag{4.24}$$

The parameter $\gamma$ is a fitting parameter which is function of the geometry. $\gamma$ is set to 0.28 empirically by matching the 3-D numerical simulation values. Thus $C_{of\_inf}$ is given by,

$$C_{of\_inf} = \alpha_{of\_sr} \cdot \frac{2\pi k_2 \varepsilon_o L_{sd}}{\cosh^{-1}\left(\frac{2H_{eff}}{d}\right)} = \alpha_{of\_sr} \cdot \frac{2\pi k_2 \varepsilon_o L_{sd}}{\cosh^{-1}\left(\frac{2\sqrt{h^2 + (0.28L_{sd})^2}}{d}\right)} \tag{4.25}$$

where $\alpha_{of\_sr}$ is the factor due to the screening of the adjacent gate/source/drain. $\alpha_{of\_sr} = 1$ if the height of the adjacent gate/source/drain $H_{adj} = 0$, and $\alpha_{of\_sr} = 0.5$ if $H_{adj} = H_{gate}$ which is the usual case (Figure 4.8). Equation (4.25) models the isolated outer-fringe

capacitance accurately for both cases provided $L_{sd} < W_{gate}$ (Figure 4.10). We choose $\alpha_{of\_sr}$ = 0.5 in the following assuming multiple devices are in series (as in a NAND stack).

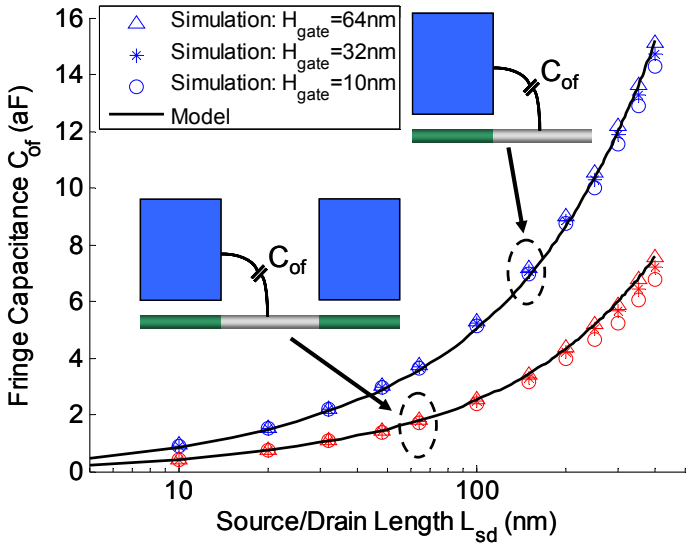



**Figure 4.10:** The comparison between the analytic model and 3-D simulation to calculate the fringe capacitance $C_{of}$ for the device with single isolated channel, with different gate height. The gate length ($L_g$) is 32 nm, $d$ = 1.5 nm, $h$ = 4 nm, and $k_2$ = 3.9.

With more than one channel per gate, there will be additional potential drop between the two electrodes of the capacitor $C_{of}$ caused by the adjacent cylinders (Figure 4.11).

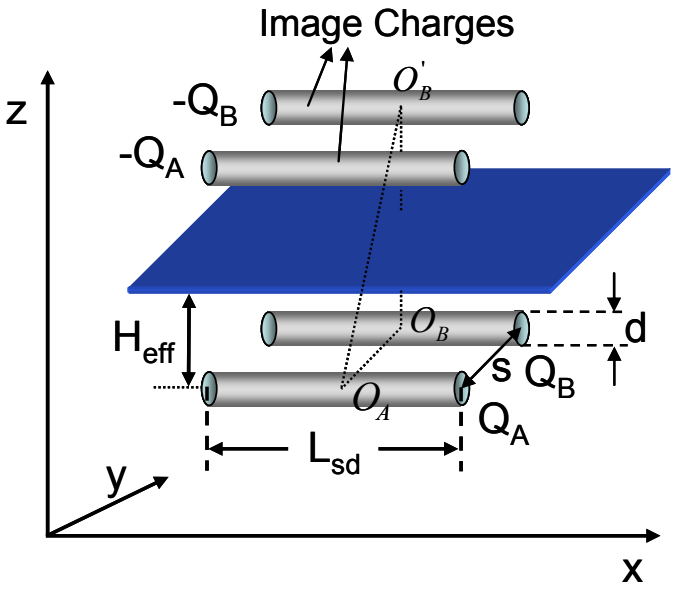



**Figure 4.11:** The equivalent parallel system in order to calculate the fringe capacitance $C_{of}$ with two parallel cylinders in the source/drain region. The actual potential difference between the cylinder A and the metal gate is affected by the adjacent line charge $Q_B$ and its image line charge $-Q_B$.

The equivalent series capacitance due to adjacent cylinder screening is,

$$C_{of\_sr} = \alpha_{of\_sr} \cdot \frac{2\pi k_2 \varepsilon_o L_{sd}}{\ln\left(\frac{\overline{O_B' Q_A}}{\overline{O_B O_A}}\right)} = \frac{\pi k_2 \varepsilon_o L_{sd}}{\ln\left(\frac{\sqrt{(2H_{eff})^2 + s^2}}{s}\right)} \tag{4.26}$$

The simulation results show that it is reasonable to group the fringe capacitances into two components: (a) the capacitance between gate and S/D at the ends ($C_{of\_e}$) and (b) the capacitance between gate and S/D in between ($C_{of\_m}$). Recall equations (4.5, 4.8), we have to determine the parameters $\eta_1$ and $\alpha$ to calculate $C_{of\_e}$ and $C_{of\_m}$. The outer-channel region is a more open structure than the inner channel region, thus we cannot use the simplified equation (4.10) for $N > 2$, where $N$ is the number of cylinders per gate, as in the channel region. Both $\eta_1$ and $\alpha$ are functions of $N$. $\eta_1$=1 for $N$=2, and $\alpha$=1 for $N$=3 as described before. For $N > 2$, because the additional potential drop imposed by line charges decreases logarithmically with the distance, we empirically represent $\eta_1$ and $\alpha$ as,

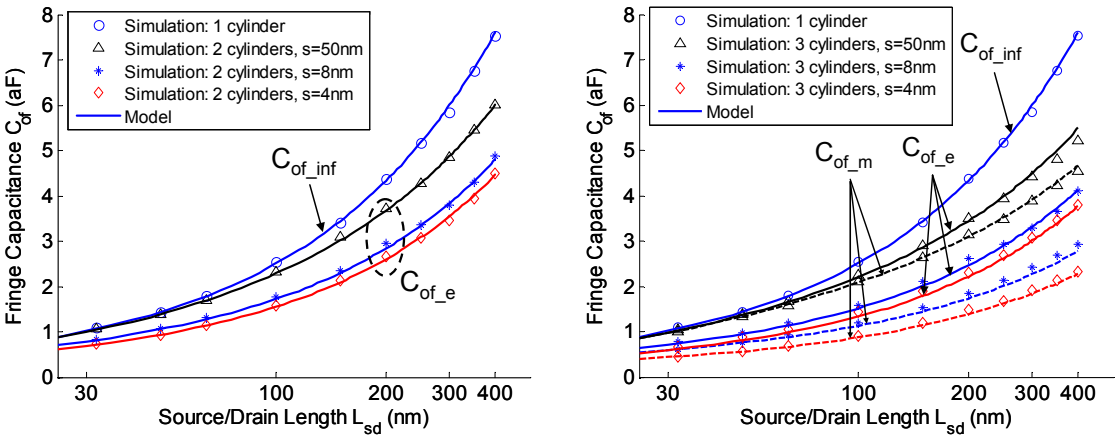
$$\begin{aligned} \eta_1 &= \exp\left(\frac{\sqrt{N^2 - 2N} + N - 2}{\tau_1 N}\right), \quad N \geq 2 \\ \alpha &= \exp\left(\frac{N-3}{\tau_2 N}\right), \quad N \geq 3 \end{aligned} \tag{4.27}$$

$\tau_1$ and $\tau_2$ are fitting parameters which describes how fast the electric flux of the adjacent cylinders decreases with increasing distance. With equations (4.5, 4.8, 4.25, 4.26), we get the fringe capacitances,

$$\begin{aligned} C_{of\_e} &= \frac{\pi k_2 \varepsilon_o L_{sd}}{\ln\left(\frac{\sqrt{(2h)^2 + (0.56L_{sd})^2 + s^2}}{s}\right) + \eta_1 \cdot \cosh^{-1}\left(\frac{\sqrt{(2h)^2 + (0.56L_{sd})^2}}{d}\right)} \\ C_{of\_m} &= \frac{2\alpha}{\eta_1} \cdot C_{of\_e} + \left(1 - \frac{2\alpha}{\eta_1}\right) \cdot \frac{\pi k_2 \varepsilon_o L_{sd}}{\cosh^{-1}\left(\frac{\sqrt{(2h)^2 + (0.56L_{sd})^2}}{d}\right)} \end{aligned} \tag{4.28}$$

where $\eta_1$ and $\alpha$ are given by equation (4.27), and $\tau_1$ and $\tau_2$ are empirically set as 2.5 and 2, respectively, to make $C_{of\_m}$ as the average value of the fringe capacitances for the cylinders in the middle of the array. The analytical models match the 3-D simulation

values very well with various number of cylinders per gate and various parameter settings provided $L_{sd} < W_{gate}$ (Figure 4.12, Figure 4.13).




(a) (b)

**Figure 4.12:** The comparison between the analytic model and 3-D simulation to calculate the fringe capacitance $C_{of\_e}$ and $C_{of\_m}$ for the device with (a) 2 parallel cylinders per gate, and (b) 3 parallel cylinders per gate, with different distance between the adjacent cylinders. $L_g$ = 32 nm, $H_{gate}$ = 64 nm, $d$ = 1.5 nm, $h$ = 4 nm, and $k_2$ = 3.9.

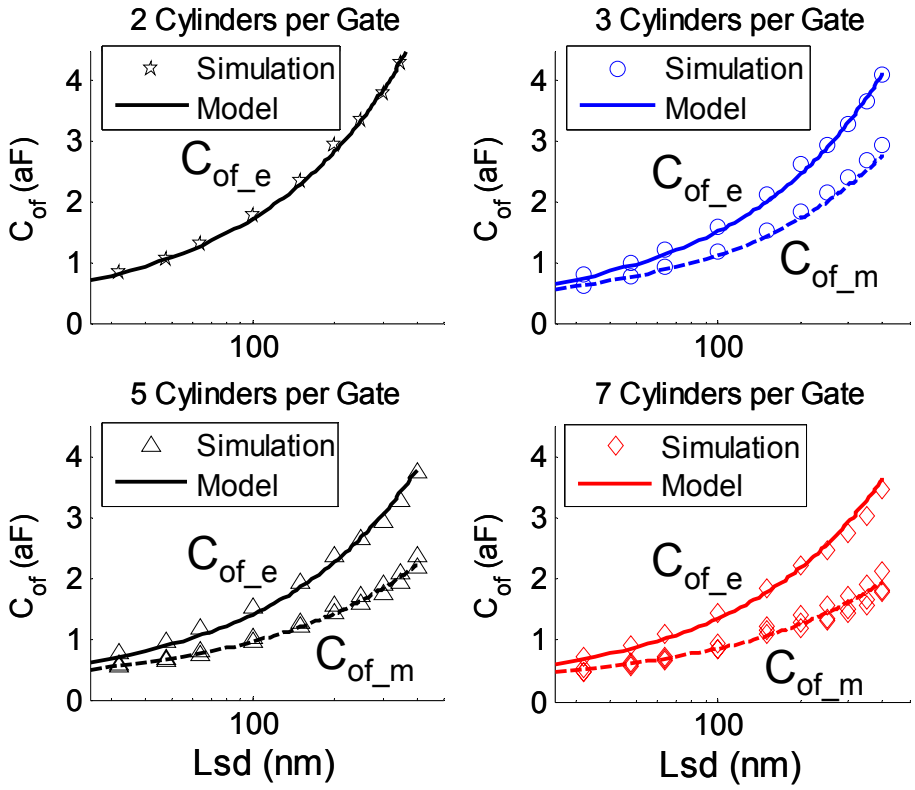



**Figure 4.13:** The comparison between the analytic model and 3-D simulation to calculate the fringe capacitance $C_{of\_e}$ and $C_{of\_m}$ for the device with different number of cylinders per gate. $L_g$ = 32 nm, $d$ = 1.5 nm, $h$ = 4 nm, $k_2$ = 3.9, and the distance between the adjacent cylinders ($s$) is 8 nm.

### 4.2.4 *Gate to Gate Capacitance*

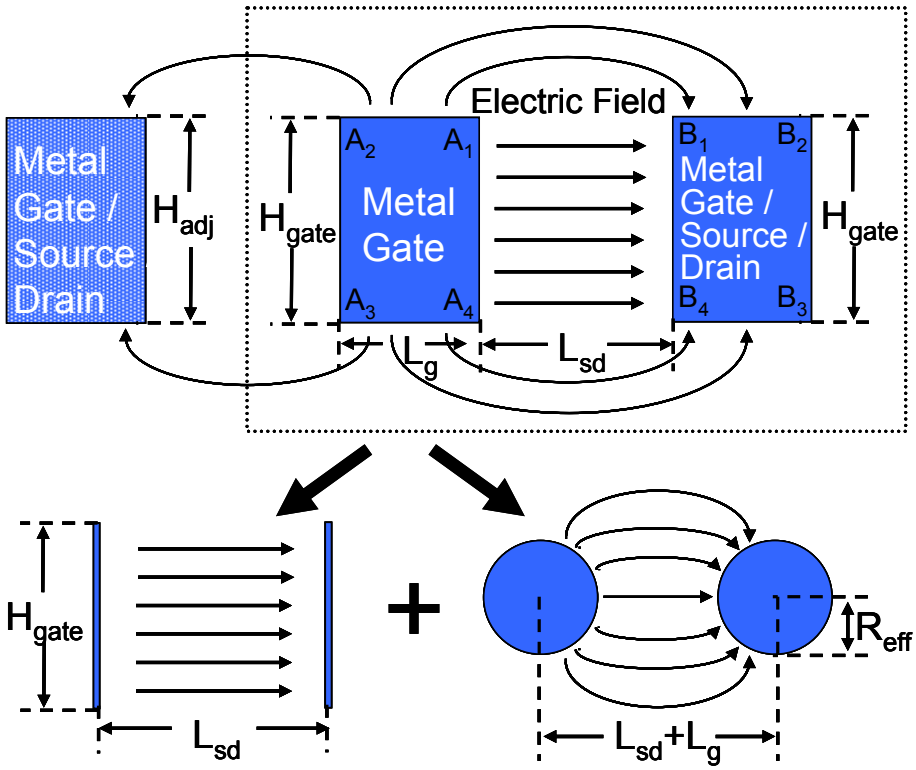



**Figure 4.14:** The gate to gate (source/drain) coupling capacitance ($C_{gtg}$) can be decomposed into two components: the parallel plate capacitance ($C_{gtg\_nr}$) due to the normal electric field between the gates, and the fringe capacitance between two cylinders ($C_{gtg\_fr}$) with equivalent radius $R_{eff}$ due to the fringing electric field between the gates.

The gate to gate capacitance ($C_{gtg}$) is another major component of the gate capacitance. We separate $C_{gtg}$ into two components (Figure 4.14): the gate to gate fringe capacitance per unit length ($C_{gtg\_fr}$) and the gate to gate plate capacitance per unit length ($C_{gtg\_nr}$). $C_{gtg\_nr}$ is due to the normal electrical filed between the two parallel plates,

$$C_{gtg\_nr} = \frac{k_2 \varepsilon_o H_{gate}}{L_{sd}} \tag{4.29}$$

Where $L_{sd}$ is the distance between the two parallel plates, and $H_{gate}$ is the gate height.

The fringe field caused by the top plates ($A_1A_2$, $B_1B_2$), bottom plates ($A_3A_4$, $B_3B_4$), and back plates ($A_2A_3$, $B_2B_3$) contribute to the fringe capacitance $C_{gtg\_fr}$. We approximate

$C_{gtg\_fr}$ as the capacitance between two parallel cylinders with equivalent radius $R_{eff}$ (Figure 4.14),

$$R_{eff} = \frac{2L_g + \tau_{bk} H_{gate}}{2\pi} \tag{4.30}$$

where $\tau_{bk}$ is the factor which accounts for the effects of the back plates ($A_2A_3$, $B_2B_3$) on $C_{gtg\_fr}$. Because the potential caused by fringe flux decreases logarithmically with the distance, $\tau_{bk} \to 0$ when $L_{sd} \to 0$, and $\tau_{bk} \to 1$ when $L_{sd} \to \infty$. We empirically approximate $\tau_{bk}$ as,

$$\tau_{bk} = \exp\left(2 - 2\sqrt{1 + \frac{2(H_{gate} + L_g)}{L_{sd}}}\right) \tag{4.31}$$

Thus $C_{gtg\_fr}$ is given by,

$$C_{gtg\_fr} = \alpha_{gtg\_sr} \cdot \frac{\pi k_2 \varepsilon_o}{\ln\left(\frac{L_{sd} + L_g}{R_{eff}}\right)} \tag{4.32}$$

The parameter $\alpha_{gtg\_sr}$ is the factor due to the screening of the adjacent gate/source/drain and interconnects. $\alpha_{gtg\_sr} = 1$ if the height of the adjacent gate/source/drain $H_{adj} = 0$. $\alpha_{gtg\_sr}$ is fitted to be 0.7 for the case $H_{adj} = H_{gate}$ (Figure 4.14). The total gate to gate capacitance per unit length is the summation of $C_{gtg\_fr}$ and $C_{gtg\_nr}$. With equations (4.29, 4.30, 4.32), we obtain,

$$C_{gtg} = \frac{k_2 \varepsilon_o H_{gate}}{L_{sd}} + \alpha_{gtg\_sr} \cdot \frac{\pi k_2 \varepsilon_o}{\ln\left(\frac{2\pi(L_{sd} + L_g)}{2L_g + \tau_{bk} H_{gate}}\right)} \tag{4.33}$$

$\tau_{bk}$ is given by equation (4.31). The above model accurately calculates the gate to gate capacitance with negligible mismatch with the numerical simulation results (Figure 4.15). We use $\alpha_{gtg\_sr} = 0.7$ in the following analysis.

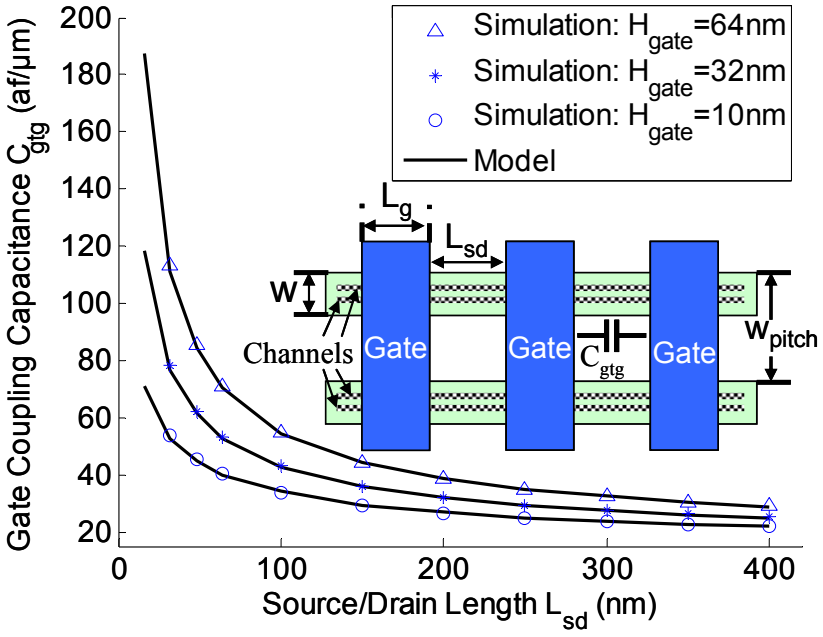



**Figure 4.15:** The comparison between the analytic model and simulation to calculate gate to gate capacitance ($C_{gtg}$) with different gate height. The inset shows the gate structure for simulation. $L_g$ = 32 nm, $d$ = 1.5 nm, $h$ = 4 nm, and $k_2$ = 3.9.

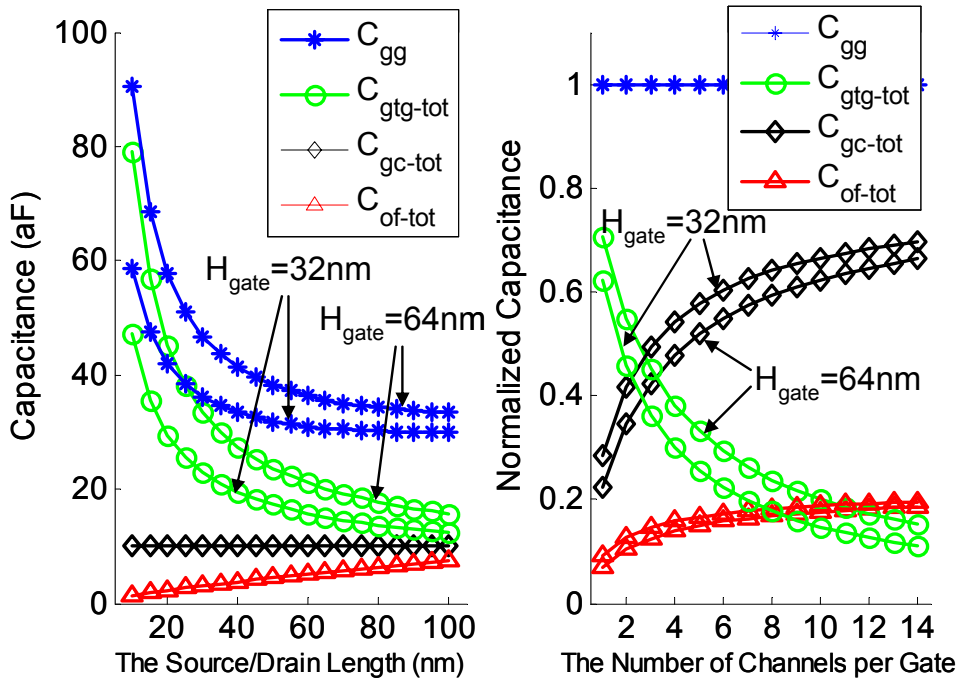
## 4.3 Gate Capacitance Analysis – Impact on Device Speed

In this section, we apply the above models to analyze the effects of gate capacitance ($C_{gg}$) on device performance. In order to emphasis the screening effect of multiple conducting channels on the electrostatic capacitance and thereby the device performance, we ignore the quantum capacitance ($C_Q$) of the conducting cylinders in this work by assuming $C_{gc} >> C_Q$. The total gate capacitance is expressed as,

$$\begin{aligned} C_{gg} &= C_{gc} \cdot L_g + f_{miller} \cdot 2(C_{of} + C_{gtg} W_{pitch}) \\ C_{gc} &= \min(N,2) \cdot C_{gc\_e} + \max(N-2,0) \cdot C_{gc\_m} \\ C_{of} &= \min(N,2) \cdot C_{of\_e} + \max(N-2,0) \cdot C_{of\_m} \end{aligned} \tag{4.34}$$

The parameter $f_{miller}$ is the Miller factor, set to be 1.5 for the switching device in inverter. $N$ is the number of channels per gate, $L_g$ is the physical gate length, and $W_{pitch}$ is the device pitch in the width direction (Figure 4.15). We evaluate the gate capacitance and device performance for a 32 nm node technology assuming metal-1 pitch=$L_g$=32 nm, $k_1$=16, $k_2$=3.9, channel diameter $d$=1.5 nm (e.g. (19,0) carbon nanotube) or $d$=5 nm (e.g.

silicon/germanium nanowire), and the gate oxide thickness (*h*-*r*) is 3 nm. For a minimum size device, the width of the channel region (*W*) is assumed to be the same as $L_g$, and the device pitch in the width direction ($W_{pitch}$) is about 3*W*.




(a) (b)

**Figure 4.16:** The gate capacitance $C_{gg}$ and its components as function of (a) the source/drain length, and (b) the number of channels per gate. The results with two gate heights (32 nm & 64 nm) are plotted. The width of gate region with high-*k* dielectric is 32 nm, the device pitch in the width direction ($W_{pitch}$) is 96 nm, $L_g$ = 32 nm, $d$ = 1.5 nm, $k_2$ = 3.9, and the gate oxide thickness (*h*-*r*) is 3 nm.

Figure 4.16 shows $C_{gg}$ and its components as a function of the source/drain length ($L_{sd}$), and the number of channels per gate (Figure 4.16). We define the total gate to channel capacitance as $C_{gc\_tot}=C_{gc}\cdot L_g$, the total outer fringe capacitance as $C_{of\_tot}=2f_{miller}\cdot C_{of}$, the total gate to gate (source/drain) capacitance as $C_{gtg\_tot}=2f_{miller}\cdot C_{gtg}\cdot W_{pitch}$, and the total parasitic gate capacitance as $C_{par}=C_{of\_tot}+C_{gtg\_tot}$. Unlike 3-D or 2-D devices, underlapped gate structure is not likely to improve the device speed for 1-D device because the fringe capacitance $C_{of\_tot}$ is not a major component of $C_{gg}$. For the device with a single channel per gate, $C_{gtg\_tot}$ is the largest component, far more than 50% of $C_{gg}$. There are three ways to improve the percentage of $C_{gc\_tot}$ out of $C_{gg}$ (Figure 4.16): 1) Increase the source/drain length. While $C_{of\_tot}$ increases with $L_{sd}$, $C_{gtg\_tot}$ decreases more quickly with increasing $L_{sd}$. This results in a smaller $C_{gg}$, at the cost of larger source/drain extension resistance ($R_{ext}$)

and lower device density. 2) Increase the number of channels per gate. In this case, $C_{gtg\_tot}$ remains almost the same. Because the parallel channels are screened by adjacent cylinders, $C_{of\_tot}$ increases slower than $C_{gc\_tot}$ as a function of $N$ due to the more open geometry (due to less gate shielding) of the region outside the gate than the inner channel region. With about 4 to 5 channels per gate, $C_{gc\_tot}$ is comparable to the parasitic capacitance ($C_{par}=C_{of\_tot}+C_{gtg\_tot}$). 3) Reduce the gate height. Both $C_{gc\_tot}$ and $C_{of\_tot}$ are almost independent of $H_{gate}$ while $C_{gtg\_tot}$ decreases with $H_{gate}$. Reducing $H_{gate}$ from 64 nm ($H_{gate}=2L_g$) to 32 nm ($H_{gate}=L_g$) results in 20% to 30% smaller $C_{gtg}$ with varying source/drain length (Figure 4.16).

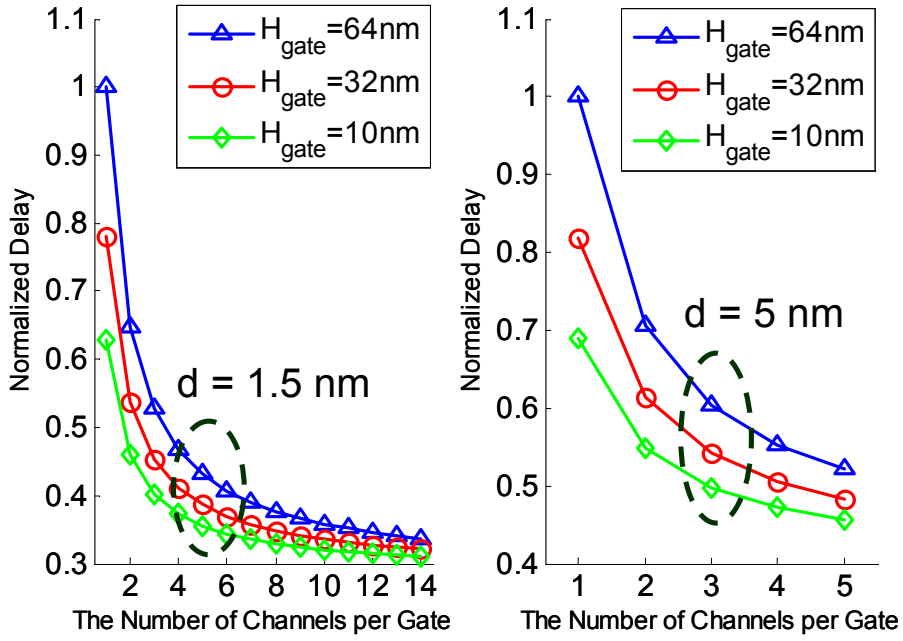



**Figure 4.17:** The device delay as a function of the number of channels per gate, with different gate height. The width of gate region with high-$k$ dielectric is 32 nm, the device pitch in the width direction ($W_{pitch}$) is 96 nm, $L_g$=32nm, $k_1$=16, $k_2$=3.9. The channel diameter $d$ = 1.5 nm for the left plot, and $d$ = 5 nm for the right plot. The gate oxide thickness ($h$-$r$) is 3 nm for both plots.

To evaluate the device speed, we assume the drive current is proportional to gate to channel capacitance per unit channel length ($C_{gc}$), i.e. $I_{on} \propto C_{gc}$, for a given $L_{sd}$[+]. The local interconnect series resistance between devices ($R_s$) is usually much smaller than the device intrinsic resistance ($R_{on}$) (including channel resistance $R_{ch}$, source/drain extension

[+] This assumption is valid in the ballistic transport region [12]

resistance $R_{ext}$, and source/drain contact resistance $R_c$), thus the device delay is proportional to $C_{gg}$,

$$\tau_{delay} \propto \frac{C_{gc} \cdot L_g + 3(C_{of} + C_{gtg} W_{pitch})}{C_{gc}} \tag{4.35}$$

For a device with a single channel, doubling the number of channels improves device speed by 35%, and halving the gate height makes the device 20% faster (Figure 4.17). Both increasing the number of channels per gate and reducing the gate height are effective ways in improving the device delay, while reducing the gate height is also efficient in reducing the dynamic power consumption due to the smaller parasitic capacitance.

## 4.4 Conclusions

This chapter presents accurate analytical models to calculate the gate capacitance of the device with high-*k* gate dielectric material and multiple cylindrical conducting channels including the screening effect. The accuracy of the analytic models is within 10% of the values simulated by 3-D numerical field solvers. For non-cylindrical conducting channels, 15% accuracy can be achieved for the devices with square cross-section channels by substituting the channel diameter in equations with the square width. For rectangular cross-section channels with an arbitrary aspect ratio, the capacitance equations can be derived analogously using the procedures described in this chapter. These models are suitable for incorporation into compact models for circuit simulations such as SPICE [13].

Using these simple analytical models, one can gain insights into device performance as a function of various device design parameters. Using these simple analytical formulae for the gate capacitance, one can obtain a realistic estimate of circuit performance (Equ. 4.35) including the parasitic capacitance. We show that the gate to gate capacitance ($C_{gtg}$) is the largest component of the total gate capacitance ($C_{gg}$) for typical 1-D device. Unlike 3-D or 2-D devices, underlapped gate structure is not likely to improve the device speed for 1-D device because the fringe capacitance ($C_{of}$) is not a major component of $C_{gg}$. Both increasing the number of channels per gate and reducing the gate height are effective in

improving the device speed. The analysis made in this work ignored quantum capacitance related with nanoscale device and used a simplified model to estimate $I_{on}$. A more accurate estimate of the dependence of device performance on device parameters (e.g. $N$, $L_{sd}$, $L_g$, $T_{ox}$) can be made by incorporating the capacitance models developed here with realistic device current models and quantum capacitance model [19]. This is outside the scope of this present work. On the other hand, the analytical capacitance model presented in this chapter is broadly applicable to all 1-D devices such as carbon nanotube transistors [4-8], 1-D semiconductor nanowire transistor [1-3], and scaled down tri-gate FET [9] or FINFET with a small aspect ratio. Our analysis of device speed shows the importance of accurately modeling the parasitic capacitance for performance evaluation of these 1-D FETs.

## 4.5 Bibliography


[1] G.F. Zheng, W. Liu, S. Jin, C.M. Lieber, "Synthesis and Fabrication of High-Performance n-Type Silicon Nanowire Transistros," *Advanced Materials*, vol. 16, pp.1890-1893, 2004.

[2] Z.Y. Fan and J. G. Lu, "Electrical properties of ZnO nanowire field effect transistors characterized with scanning probes," *Applied Physics Letters*, vol. 86, pp. 032111, 2005.

[3] F. Liu, M.Q. Bao, and K. L. Wang, "One-dimensional Transport of $In_2O_3$ Nanowires," *Applied Physics Letters*, vol. 86, pp. 213101, 2005.

[4] R. Martel, T. Schmidt, H. R. Shea, T. Hertel, and Ph. Avouris "Single- and Multi-wall Carbon Nanotube Field-effect Transistors," *Applied Physics Letter*s, vol. 73, pp. 2447-2449, 1998.

[5] H.-S. P. Wong, J. Appenzeller, V. Derycke, R. Martel, S. Wind, Ph. Avouris, "Carbon Nanotube Field Effect Transistors – Fabrication, Device Physics, and Circuit Implications", *Proc. Intl. Solid State Circuits Conference*, pp. 370 – 371, 2003.

[6] A. Javey, J. Guo, D. B. Farmer, Q. Wang, E. Yenilmez, R. G. Gordon, M. Lundstrom, H. Dai, "Self-Aligned Ballistic Molecular Transistors and Electrically Parallel Nanotube Arrays," *Nano Letters*, vol. 4, no. 7, pp. 1319-1322, 2004.

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

[8] I. Amlani, J. Lewis, K. Lee, R. Zhang, J. Deng, and H.-S. P. Wong, "First Demonstration of AC Gain From a Single-walled Carbon Nanotube Common-Source Amplifier," *Intl. Electron Devices Meeting,* pp. 559-562, December, 2006.

[9] S.-T. Chang, "Nanoscale Strained Si/SiGe Heterojunction Trigate Field Effect Transistors," *Japanese Journal of Applied Physics*, vol. 44, pp. 5304–5308, 2005.

[10] J. Guo, A. Javey, H. Dai, and M. Lundstrom, "Performance Analysis and Design Optimization of Near Ballistic Carbon Nanotube Field-Effect Transistors," *Proc. Intl. Electron Devices Meeting,* pp. 703 – 706, December, 2004.

[11] P. J. Burke, "Luttinger Liquid Theory as a Model of the Gigahertz Electrical Properties of Carbon Nanotubes," *Nanotechnology, IEEE Transactions on*, vol. 1, pp. 129-144, 2002.

[12] K. Natori, Y. Kimura, and T. Shimizu, "Characteristics of a Carbon Nanotube Field-effect Transistor Analyzed as a Ballistic Nanowire Field-effect Transistor," *Journal of Applied Physics*, vol. 97, pp. 034306, 2005.

[13] J. Deng, H.-S. P. Wong, "A Circuit-Compatible SPICE Model for Enhancement Mode Carbon Nanotube Field Effect Transistors," *Proc. Intl. Conf. Simulation of Semiconductor Processes and Devices,* pp. 166 – 169, September, 2006.

[14] X. Wang, H.-S. P. Wong, P. Oldiges, R.J. Miller, "Electrostatic Analysis of Carbon Nanotube Arrays," *Proc. Intl. Conf. Simulation of Semiconductor Processes and Devices,* pp. 163 – 167, September, 2003.

[15] X. Wang, H.-S. P. Wong, P. Oldiges, R.J. Miller, "Gate Capacitance Optimization for Arrays of Carbon Nanotube Transistors," *Device Research Conference,* pp. 87 – 88, 2003.

[16] L.W. Nagel, "SPICE2: A Computer Program to Simulate Semiconductor Circuits," *Technical Report ERL-M520, University of California, Berkeley*, 1975,

[17] Maxwell 3D®, Ansoft Corporation, PA.

[18] C. Snow, "Formulas for Computing Capacitance and Inductance," *Washington : U.S. Govt. Print. Off.*, 1954.

[19] J. Deng, N. Patil, K. Ryu, A. Badmaev, C. Zhou, S. Mitra, H.-S. P. Wong, "Carbon Nanotube Transistor Circuits: Circuit-Level Performance Benchmarking and Design Options for Living with Imperfections," *International Solid State Circuits Conference (ISSCC)*, pp. 70-71, San Francisco, CA, February, 2007.

# Chapter 5

# COMPACT MODEL FOR THE INTRINSIC CHANNEL REGION OF CNFET

A version of this chapter has been submitted for publication. J. Deng and H.-S P. Wong, "A Compact SPICE Model for Carbon Nanotube Field Effect Transistors Including Non-Idealities and Its Application — Part I: Model of the Intrinsic Channel Region," submitted to *IEEE Transactions on Electron Devices,* 2007.

## 5.1 Introduction

As one of the promising new transistors, Carbon Nanotube File Effect Transistor (CNFET) avoids most of the fundamental limitations for traditional silicon MOSFETs. With ultra long (~1μm) mean-free-path (MFP) for elastic scattering, ballistic or near-ballistic transport can be obtained with intrinsic CNT under low voltage bias to achieve the ultimate device performance [1,2,3,4]. The quasi-1D structure provides better electrostatic control over the channel region than 3D device (e.g. bulk CMOS) and 2D device (e.g. fully depleted SOI) structures [5].

Efforts have been made in recent years on modeling semiconducting CNFET [6,7,8,9] for digital logic applications and CNT for interconnects [10,11] in order to evaluate the potential performance at the device level. Very promising single device, dc performance over silicon CMOS has been demonstrated either by modeling or experimental data. The reported compact models to date [6,7,8,9] used one or more lumped static gate capacitances and an ideal ballistic transport model. These simplifications are questionable when evaluating the transient response and device dynamic performance. The integral function used in [6,7] requires intensive calculation efforts and thereby makes it difficult to implement in circuit simulators, e.g. HSPICE [12]. The polynomial fitting approach used in [8] improves the run time significantly, but it makes evaluating CNFET

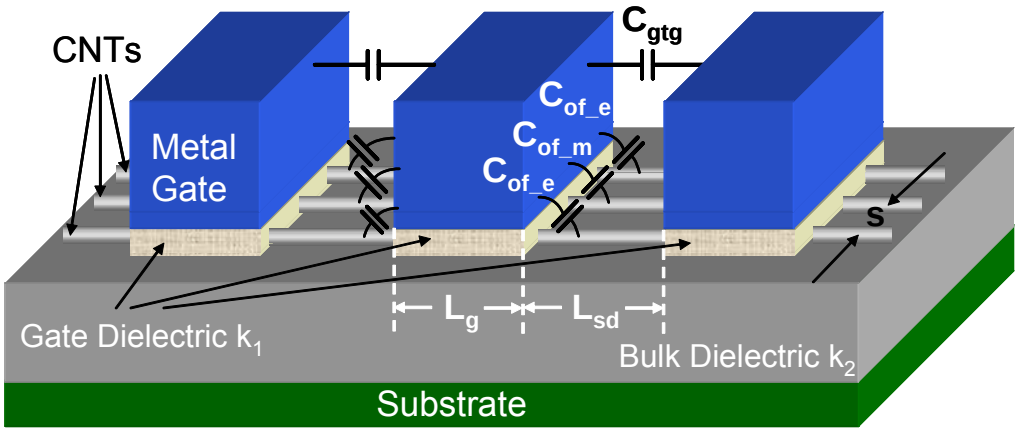
performance with different device parameters (e.g. CNT chiratities, gate oxide thickness) inconvenient. The simple coaxial or planer gate structures utilized in [6,7,9] differ from the typical realistic CNFET gate structure that consists of high-*k* gate oxide on top of $SiO_2$ insulating bulk. For a CNFET with multiple parallel CNTs [2], these published models cannot examine the multiple CNT-to-CNT screening effect on both the driving current and the effective gate capacitance. To evaluate CNFET circuit performance with improved accuracy, a CNFET device model with a more complete circuit-compatible structure and also incorporating including the typical device/circuit non-idealities is necessary.

In terms of the device operation mechanism, CNFET can be categorized as either Schottky Barrier (SB) controlled FET or MOSFET-like FET [1,2,13]. Though good dc current can be achieved by SB-controlled CNFET with the self-aligned structure [2], its ac performance is going to be poor due to the proximity of the gate electrode to the source/drain metal. The ambipolar behavior of SB-controlled CNFET also makes it undesirable for complementary logic design. Considering both the fabrication feasibility [14] and superior device performance of the MOSFET-like CNFET as compared to the SB-controlled FET, we choose to focus on MOSFET-like CNFETs in this work. This chapter models the intrinsic channel region of the CNFET which serves as the first level of the complete device model. This model includes the quantum confinement on both the circumferential and the axial directions, the acoustical/optical phonon scattering in the channel region, the screening effect by the parallel CNTs for CNFET with multiple CNTs, and the intrinsic ac behavior which is delivered by a dynamic gate capacitance network. The complete device model that includes the channel elastic scattering, the doped source/drain region, Schottky barrier (SB) resistance, multiple CNTs per device and other device/circuit non-idealities, and its applications are reported in Chapter 6[+]. The modeling approach and methodology described in this work is generally applicable to other 1-D device, e.g. silicon nanowire FET [5], provided that the appropriate equations for the band structure and/or density of states (DOS) are used.

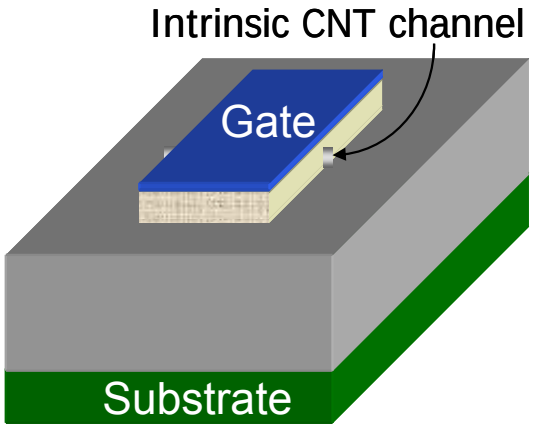
[+] The model is available at https://www.stanford.edu/group/nanoelectronics/model_downloads.htm

This chapter is organized as follows: First, we describe the device structure used for the modeling. Next, we show both the mathematical expressions and the circuit representations of each major component. Finally we will discuss the application of this model for a complete device model for circuit simulation.

## 5.2 Device Structure




(a)




(b)

**Figure 5.1**: (a) The 3-D device structure of CNFETs with multiple channels, high-*k* gate dielectric material, and the related parasitic gate capacitances. In this example, three CNFETs are fabricated along one single CNT. The channel region of CNTs is un-doped, and the other regions of CNTs are heavily doped. (b) The 3-D device structure of CNFET that is modeled in this chapter, with only the intrinsic channel region.

A typical layout of a MOSFET-like CNFET device is illustrated in Figure 5.1(a). One or multiple devices can be fabricated along a single CNT, and multiple CNTs may be placed under the same gate in order to improve the drive current. The CNT channel region is

undoped, and the other regions are heavily doped, acting as both the source/drain extension region and/or interconnects between two adjacent devices (un-contacted source-gate/gate-drain configurations). In order to account for the screening by the adjacent CNTs for the device with multiple CNTs, the nanotubes under the gate are grouped into (1) the two CNTs at the edges; (2) the other CNTs in the middle. The CNTs in each group are treated identically [15].

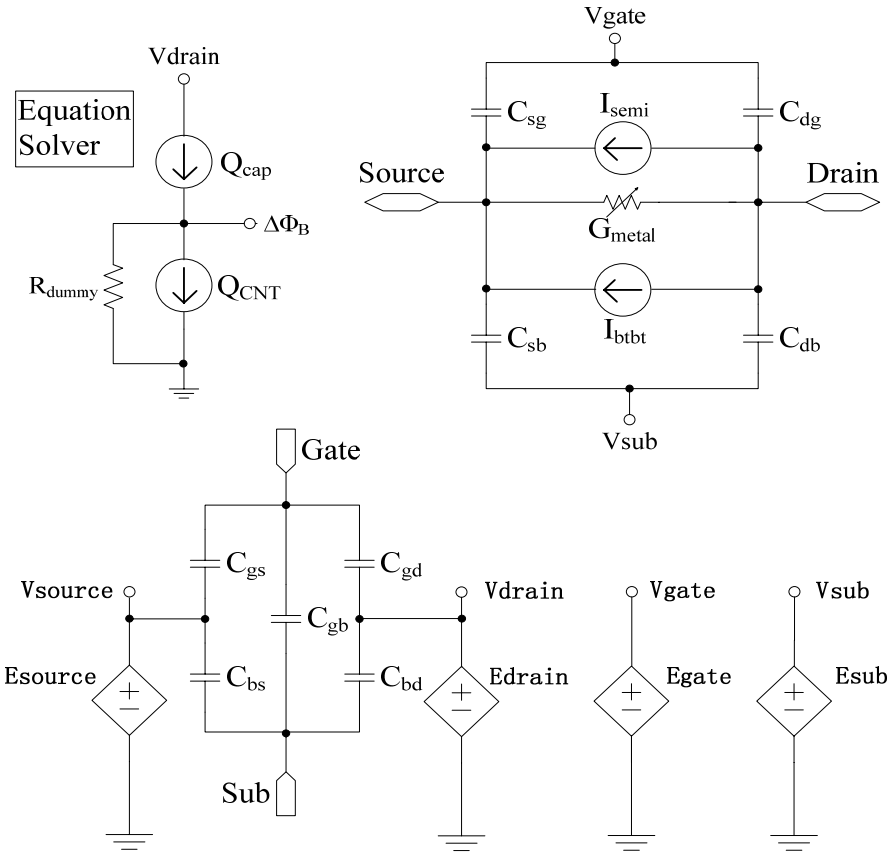
This chapter describes the modeling of one single intrinsic channel of CNFET, as shown in Figure 5.1(b), which is a starting point towards the complete device model reported in Chapter 6. For MOSFET-like CNFET, since pFET behavior is similar to nFET, we will only describe the equations for nFET in this chapter, though we implemented both nFET and pFET for the SPICE simulations.

From the circuit point of view, the input / output signals can be defined as either current or potential, and the outputs (current or potential) are just the responses to the inputs (current or potential). The potential can be either electrostatic potential or chemical potential (referred as the "Fermi level" in the following). For macro scale circuits/devices, potential usually means electrostatic potential because the Fermi level profile follows the electrostatic potential profile around the contacts which are usually the input/output ports. For a 1-D quantum wire with two contacts (source and drain), the source Fermi level $\mu_s$ and the drain Fermi level $\mu_d$ will be split apart with finite drain bias ($V_{ds}$) due to the finite density of states (DOS) [16]. Thus it becomes ambiguous to describe 1-D device behavior with electrostatic potential only, especially for the devices that are connected serially without an intermediate reservoir of electrons provided by a metal contact. In this work, we use both the Fermi level and the surface (electrostatic) potential to describe CNFET device behavior.

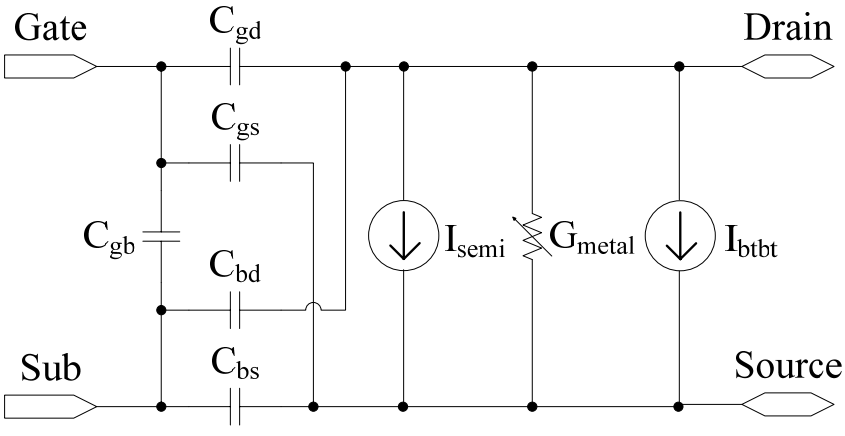
## 5.3 Model of the Intrinsic Channel Region

This part models the intrinsic channel region of CNFET with near-ballistic transport, and without any parasitic capacitance and parasitic resistance. The equivalent circuit model is shown as Figure 5.2. It consists of two major parts: the current sources and the trans-capacitance network to account for both dc and ac behavior. Figure 5.2(a) is the

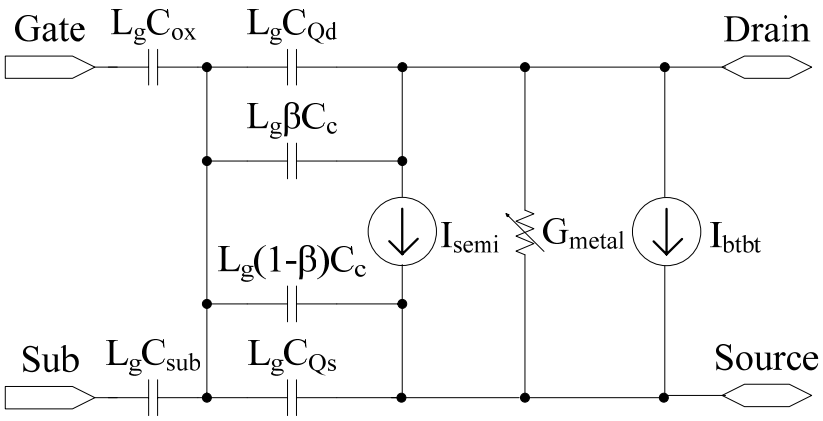
equivalent circuit implemented with HSPICE, and Figure 5.2(b, c) are the other two possible implementations for the trans-capacitance network which will be discussed in Section 5.3.2.




(a)




(b)




(c)

**Figure 5.2**: The equivalent circuit model for the intrinsic channel region of CNFET. (a) The 9-capacitor model assuming the carrier distribution along the channel is uniform. Exxx is the voltage controlled voltage source, and the potential of Vxxx equals to the controlling voltage source. $R_{dummy}$ is a large value (>1E15) resistor to keep the circuit stable. (b) The 5-capacitor model, and (c) the 6-capacitor model, assuming all the carriers from $+k$ branches are assigned to the source and all the carriers from $-k$ branches are assigned to the drain.

The Fermi level profiles and the energy band diagram in the channel region with ballistic transport are illustrated in Figure 5.3. There are four Fermi levels (both input and output Fermi levels for source and drain) for each device due to the quantum resistance.

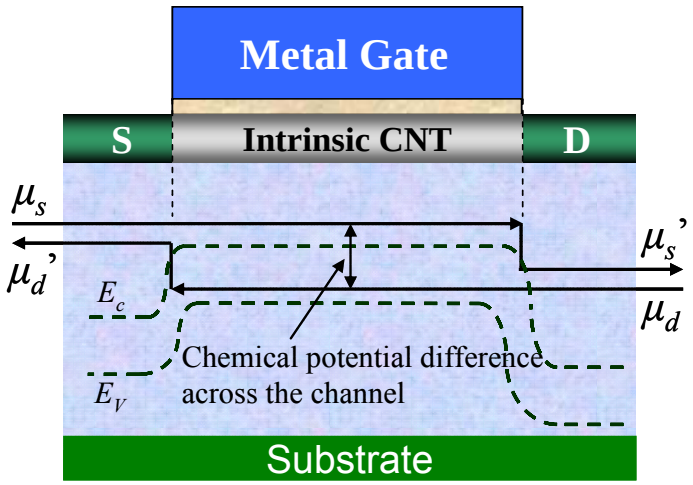



**Figure 5.3**: Ideal CNFET with ballistic (intrinsic) channel. Superposed are the Fermi level profiles (solid arrows) from source to drain and the energy band diagram (dashed lines) with bias $V_{DS} = (\mu_d - \mu_s)/e$.

Applying biases $V_G$, $V_S$, $V_D$, and $V_{sub}$ to the four terminals of the device, the potential difference $V_{DS}$ does not necessarily equal to $\mu_d$-$\mu_s$ with non-ballistic transport for 1-D quantum wire. The potential differences $\mu_s$-$\mu_s$' and $\mu_d$-$\mu_d$' are determined by the property

of the source/drain extension region. We will treat the non-ballistic transport and the potential drop at the source/drain extension region and the contacts in the complete device model. We assume near-ballistic transport and ideal (reflectionless) contacts in this chapter, i.e. $eV_{DS} \approx \mu_d - \mu_s$ so $\mu_s$ remains almost constant in the source-channel region and $\mu_d$ remains almost constant in the channel-drain region (Figure 5.3).

### 5.3.1 *Current Sources*

The single-walled carbon nanotube (SWCNT) is treated as quasi 1-D quantum wire in this work. For SWCNT with charities ($n_1$, $n_2$), the diameter ($D_{CNT}$) is given by ($a$ = 2.49 Å is the lattice constant) [17],

$$D_{CNT} = \frac{a\sqrt{n_1^2 + n_1 n_2 + n_2^2}}{\pi} \tag{5.1}$$

SWCNTs can be grouped as either metallic nanotubes if $n_1$-$n_2$ is an integer multiple of 3 (mod($n_1$-$n_2$,3)=0), or semiconducting nanotubes if $n_1$-$n_2$ is not an integer multiple of 3 (mod($n_1$-$n_2$,3)$\neq$ 0) [17]. For SWCNT with a finite length ($L_g$) and a finite diameter ($D_{CNT}$), there are quantization effects in both the circumferential direction and the axial (channel length) direction. Applying the Born-von Karman boundary condition on both directions, the $E$-$k$ dispersion relation is quantized into discrete sub-states. We denote ($m$,$l$) as the $l^{\text{th}}$ sub-state at the $m^{\text{th}}$ sub-band, $k_m$ as the wave-number of the $m^{\text{th}}$ sub-band in circumferential direction, and $k_l$ as the wave-number of the $l^{\text{th}}$ sub-state in current flow direction. We define the sub-bands with positive band gap as "semiconducting sub-bands", and the sub-bands with zero or negative band gap as "metallic sub-bands". Thus the band structure of metallic nanotubes can be treated as a summation of metallic sub-bands and semiconducting sub-bands.

The wave numbers related with semiconducting sub-bands are given by [16,17],

$$k_m = \frac{2\pi}{a\sqrt{n_1^2 + n_1 n_2 + n_2^2}} \cdot \lambda \tag{5.2a}$$

$$\lambda = \begin{cases} \dfrac{6m-3-(-1)^m}{12} & m = 1,2,... \quad ,\mathrm{mod}(n_1 - n_2,3) \neq 0 \\ m & m = 0,1,... \quad ,\mathrm{mod}(n_1 - n_2,3) = 0 \end{cases} \tag{5.2b}$$

$$k_l = \frac{2\pi}{L_g} l \quad ,l = 0,1,2,... \tag{5.2c}$$

$m$=0 is reserved for the metallic sub-band. $k_l$ approaches continuous values for large $L_g$. Around the Fermi point, or more explicitly for carrier energy $E_{m,l} << V_\pi$ (~3.033eV, the carbon $\pi$-$\pi$ bond energy in the tight bonding model), CNT $E$-$k$ dispersion relation can be approximated as [17],

$$E_{m,l} \approx \frac{\sqrt{3}}{2} a V_\pi \sqrt{k_m^2 + k_l^2} \tag{5.3}$$

$E_{m,l}$ is the carrier energy at the ($m,l$) sub-state above the intrinsic level $E_i$, and $E_{m,0}$ is the half band gap of the $m^{\text{th}}$ sub-band. The above equations are valid for both metallic and semiconducting nanotubes. As an example, for the (19,0) semiconducting SWCNT, $E_{1,0}$=0.29eV, $E_{2,0}$=0.58eV, and $E_{3,0}$=1.16eV. For the (18,0) metallic SWCNT, $E_{0,0}$=0, $E_{1,0}$=0.92eV, and $E_{2,0}$=1.83eV. The band gap of metallic sub-band ($m$=0) is zero.

We consider three current sources in the CNFET model: (1) the theriomic current contributed by the semiconduting sub-bands ($I_{semi}$) with the classical band theory, (2) the current contributed by the metallic sub-bands ($I_{metal}$), and (3) the leakage current ($I_{btbt}$) caused by the band to band tunneling mechanism through the semiconducting sub-bands.

***I*$_{semi}$**: For semiconducting sub-bands, we only consider the electron current for the nFET because the hole current is suppressed by the n-type heavily doped source/drain, and usually is negligible compared to the electron current. The current contributed by the sub-state ($m,l$) is given by,

$$J_{m,l}(V_{xs}, \Delta\Phi_B) = 2env_F \tag{5.4}$$

$V_{xs}$ is potential difference between node $x$ and source. The Fermi velocity $v_F = 1/\hbar \cdot \partial E/\partial k_l$. The factor of 2 is due to electron spin degeneracy, $e$ is the unit electronic charge, and $n$ is the number of electrons that occupy the sub-state ($m,l$), given by,

$$n = \frac{f_{FD}(E_{m,l} + eV_{xs} - \Delta\Phi_B)}{L_g} \tag{5.5a}$$

$$f_{FD}(E) = \frac{1}{1+e^{E/kT}} \tag{5.5b}$$

$f_{FD}(E)$ is the Fermi-Dirac distribution function. $k$ is the Boltzmann constant and $T$ is the temperature in Kelvin. $E_{m,l}$ is the carrier energy at the substate ($m$, $l$). $\Delta\Phi_B$ is the channel surface potential change with gate/drain bias.

With equations (5.4, 5.5), we obtain,

$$J_{m,l}(V_{xs}, \Delta\Phi_B) = \frac{2e}{h}\frac{\sqrt{3}a\pi V_\pi}{L_g}\frac{k_l}{\sqrt{k_m^2 + k_l^2}}\frac{1}{1+e^{(E_{m,l}+eV_{xs}-\Delta\Phi_B)/kT}} \tag{5.6}$$

The total current contributed by all sub-states is equal to the current flowing from the drain to the source (+$k$ branch) minus the current flowing from the source to the drain (-$k$ branch),

$$I_{semi}(V_{ch,DS}, V_{ch,GS}) = 2\sum_{\substack{k_m \\ m=1}}^{M}\sum_{\substack{k_l \\ l=1}}^{L}\left[T_{LR}J_{m,l}(0, \Delta\Phi_B)\big|_{+k} - T_{RL}J_{m,l}(V_{ch,DS}, \Delta\Phi_B)\big|_{-k}\right] \tag{5.7}$$

$V_{ch,DS}$ and $V_{ch,GS}$ denotes the Fermi potential differences near source side within the channel. The factor of 2 is due to the double-degeneracy of the sub-band. $M$ and $L$ are the number of sub-bands and the number of sub-states, respectively. For typical devices with appropriate diameter range ($D_{CNT}$ < 3nm) and short gate length ($L_g \leq$ 100nm), only the first 2 or 3 sub-bands and the first 10~15 sub-states have a significant impact on the current using a sub-1V power supply. Including more sub-bands should be done with more caution due to two limitations: (1) the band structure model used in this work requires $E_{m,l} << V_\pi$, (2) the complex phonon modes at high energy level. For long channel devices ($L_g >$ 100nm), one can either approximate the current for short device equation (5.7) by setting $L_g$=100nm in equations (5.2c, 5.5a, 5.6), or use the long channel model introduced in equation (5.14) below. $T_{LR}$ and $T_{RL}$ are the transmission probability of the carriers at the sub-state ($m$,$l$) in +$k$ branch and –$k$ branch, respectively. We consider three

typical scattering mechanisms in the channel region: (1) acoustic phonon scattering (near elastic process [4]), (2) optical phonon scattering (inelastic process [3]), and (3) elastic scattering. The elastic scattering probability is assumed to be independent of the carrier energy, and will be treated in the complete device modeling in Chapter 6. Both the acoustic phonon scattering and optical phonon scattering depend on the carrier energy. We assume there is no interaction between the sub-bands, thus only intra-band scatterings are considered in this work. Random angle scatterings are suppressed and only backscattering and forward scattering can occur in a 1-D quantum wire due to the Pauli's Exclusion principle and the confined *k*-space [18]. A scattering event from the sub-state $(m,l_1)$ in $+/-k$ branch to the sub-state $(m,l_2)$ in $-/+k$ branch can occur only if two conditions are satisfied: (1) the sub-state $(m,l_1)$ is filled with electrons. (2) the sub-state $(m,l_2)$ is empty so it can accept the scattered carrier from $(m,l_1)$. Assuming the optical phonon scattering mean free path (MFP) ($\lambda_{op} \sim 15$ nm [19]) and the acoustic phonon scattering MFP ($\lambda_{ap} \sim 500$ nm [20]) are constant if both conditions are met, we normalize the effective acoustic phonon scattering MFP ($l_{ap}$) and the effective optical phonon scattering MFP ($l_{op}$) of the semiconducting sub-bands to the available target empty states,

$$l_{ap}(V_{xs},m,l) = \frac{\lambda_{ap} D_o}{D(E_{m,l})[1 - f_{FD}(E_{m,l} - \Delta\Phi_B + eV_{xs})]} \tag{5.8a}$$

$$l_{op}(V_{xs},m,l) = \frac{\lambda_{op} D_o}{D(E_{m,l} - \hbar\Omega)[1 - f_{FD}(E_{m,l} - \hbar\Omega - \Delta\Phi_B + eV_{xs})]} \tag{5.8b}$$

$\hbar\Omega$ (~0.16eV [19]) is the optical phonon energy that a carrier attains before a optical phonon scattering can occur. Optical phonon scattering becomes more significant at high $V_{ch,DS}$ bias. $D_o$ is a constant $8/(3\pi V_\pi \cdot d)$ where $d$ is the carbon-carbon bond distance, about 0.144 nm. $D(E)$ is the CNT universal density of states (DOS) which is valid in the range $E_{m,l} << V_\pi$ [17],

$$D(E) = \begin{cases} D_0 \cdot E / \sqrt{E^2 - E_{m,0}^2} & E > E_{m,0} \\ 0 & E \le E_{m,0} \end{cases} \tag{5.9}$$

The effective phonon scattering MFP is in the form of,

$$\frac{1}{l_{eff}(V_{xs},m,l)} = \frac{1}{l_{ap}(V_{xs},m,l)} + \frac{1}{l_{op}(V_{xs},m,l)} \tag{5.10}$$

It is reasonable to assume that the phonon back-scattered carriers are not likely to be back-scattered again due to the energy loss and/or the occupied states. Thus the transmission probabilities in equation (5.7) are given by,

$$T_{LR} = \frac{l_{eff}(V_{ch,DS},m,l)}{l_{eff}(V_{ch,DS},m,l) + L_g} \tag{5.11a}$$

$$T_{RL} = \frac{l_{eff}(0,m,l)}{l_{eff}(0,m,l) + L_g} \tag{5.11b}$$

The key parameter for evaluating CNFET current is $\Delta\Phi_B$, the channel surface potential change in response to changes in gate and source/drain bias.

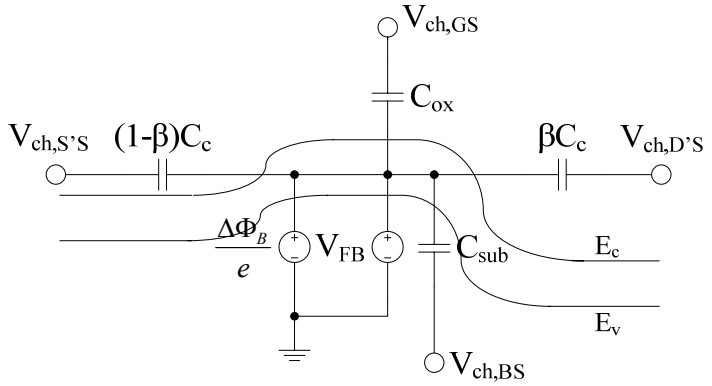



**Figure 5.4:** The electrostatic capacitor model used to calculate the channel surface potential change $\Delta\Phi_B$ before and after Gate/Source/Drain/Substrate bias. All the node potentials are referred to the input source Fermi level. Superposed is the energy band diagram (only the first sub-band shown) from the external source node S' to the external drain node D'.

We define both internal source (S) / drain (D) inside the channel region and external source (S') / drain (D') outside of the channel region for the purpose of introducing the channel parasitic coupling. As shown in Figure 5.4, there are three electrostatic coupling capacitors assuming the channel material is with infinite DOS: the capacitance ($C_{ox}$) between the gate and channel, the capacitance ($C_{sub}$) between channel and substrate, and the capacitance ($C_c$) between channel and external drain (D') / source (S'). $\Delta\Phi_B$ is dynamically affected by the drain bias. $\beta C_c$ is a fitting parameter that describes this effect due to two mechanisms: (1) the surface potential lowering due to the electrostatic

coupling between the channel region and the external drain electrode through fringing electric field; (2) the surface potential lowering due to non-uniform channel surface potential profile caused by DIBL effect. Operationally, the parameters $C_c$ and $\beta$ are chosen to fit the sub-threshold slope and the measured short channel effect. For a semiconducting channel with a finite DOS, the channel surface potential $\Delta\Phi_B$ changes with the gate bias at a rate $\Delta\Phi_B/\Delta V_{GS} < 1$, a phenomenon known as the effect of quantum capacitance. Here, we calculate $\Delta\Phi_B$ directly using the charge conservation equations (instead of using the quantum capacitance),

$$Q_{cap} = Q_{CNT} \tag{5.12a}$$

$$Q_{cap} = C_{ox}(V_{ch,GS} - V_{FB}) + C_{sub}V_{ch,BS} + \beta C_c V_{ch,D'S} + (1-\beta)C_c V_{ch,S'S} - (C_{ox} + C_{sub} + C_c)\frac{\Delta\Phi_B}{e} \tag{5.12b}$$

$$Q_{CNT} = \frac{4e}{L_g}\sum_{\substack{k_m \\ m=m0}}^{M}\sum_{\substack{k_l \\ l=0}}^{L}\left[\frac{1}{1+e^{(E_{m,l}-\Delta\Phi_B)/kT}} + \frac{1}{1+e^{(E_{m,l}-\Delta\Phi_B+eV_{DS})/kT}}\right] \tag{5.12c}$$

$$m0 = \begin{cases} 1 & ,\mathrm{mod}(n_1 - n_2, 3) \neq 0 \\ 0 & ,\mathrm{mod}(n_1 - n_2, 3) = 0 \end{cases} \tag{5.12d}$$

When implementing the model, S' and D' can either be defined as two external coupling nodes or be defined the same as S and D nodes if the Fermi level difference between the node S (D) and the node S' (D') is small which is valid for CNFET with ideal reflectionless source/drain contacts. For metallic nanotubes, the metallic subband ($m$=0) is included in equation (5.12). The factor of 4 includes both the spin degeneracy and the double-degeneracy of the sub-band. $V_{FB}$ is the flat band voltage, and $V_{BS}$ is the potential difference between substrate and source. $Q_{cap}$ is the charge induced by the electrodes, and $Q_{CNT}$ is the total charge induced on SWCNT surface. To solve equation (5.12), we use a construct available in HSPICE*. We make two current sources with the currents equals to $Q_{cap}$ and $Q_{CNT}$ respectively, and force the two currents to be equal to each other by connecting them in series. The value $\Delta\Phi_B$ is automatically calculated as the node voltage (Figure 5.2(a)).

* A similar construct is also available in other circuit simulators, e.g. Verilog-A.

The front gate capacitance $C_{ox}$ is modeled as a planar gate structure with high-$k$ gate dielectric on top of $SiO_2$ insulating layer (Figure 5.1). For the device with multiple SWCNTs in parallel, $C_{ox}$ is grouped into the capacitance between gate and SWCNT at the two ends ($C_{ox\_e}$), and the capacitance between gate and SWCNT in the middle ($C_{ox\_m}$). This approximation is accurate within 10% of the 3-D numerical model treating each CNT correctly in its electrostatic environment [15]. The equations are summarized as below,

$$C_{ox\_e} = \frac{C_{ox\_\inf} \cdot C_{ox\_sr}}{C_{ox\_\inf} + C_{ox\_sr}} \qquad C_{ox\_m} = 2C_{ox\_e} - C_{ox\_\inf}$$

$$C_{ox\_sr} = \frac{4\pi k_1 \varepsilon_o}{\ln\left(\frac{s^2 + 2(h_{ox} - r)\cdot[h_{ox} + \sqrt{h_{ox}^2 - r^2}]}{s^2 + 2(h_{ox} - r)\cdot[h_{ox} - \sqrt{h_{ox}^2 - r^2}]}\right) + \lambda_1 \cdot \ln\left(\frac{(h_{ox} + 2r)^2 + s^2}{9r^2 + s^2}\right) \cdot \tanh\left(\frac{h_{ox} + r}{s - 2r}\right)} \tag{5.13}$$

$$C_{ox\_\inf} = \frac{2\pi k_1 \varepsilon_o}{\cosh^{-1}\left(\frac{h_{ox}}{r}\right) + \lambda_1 \cdot \ln\left(\frac{2h_{ox} + 4r}{6r}\right)} \qquad \lambda_1 = \frac{k_1 - k_2}{k_1 + k_2}$$

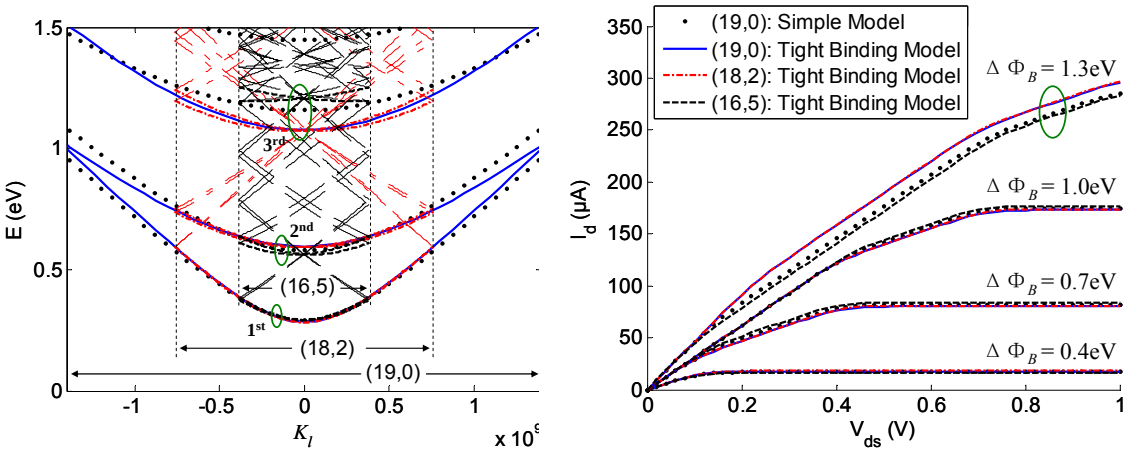
$C_{ox\_inf}$ is the coupling capacitance between the gate and one isolated SWCNT, and $C_{ox\_sr}$ is the equivalent series capacitance due to the imaging/screening effect of the adjacent nanotubes. $r$ is the radius of the SWCNT, $h_{ox}$ is the gate dielectric thickness between the SWCNT center and gate, $k_1$ and $k_2$ are the dielectric constants of the gate oxide and insulating bulk oxide respectively, and $s$ is the inter-CNT spacing. For SWCNT of 1.5 nm diameter with 4 nm thick $HfO_2$ ($k_1$=16) and 5 nm inter-CNT spacing, $C_{ox\_e}$ = 246 aF/μm and $C_{ox\_m}$ = 186 aF/μm. The substrate to gate capacitance $C_{sub}$ can either be calculated with equation (5.13) if a double gate device is desired, or be calculated with the simple equation, $C_{sub} = 2\pi k_2 \varepsilon_0 / \ln(2H_{sub}/r)$. $C_{sub}$ is approximately 21 aF/μm with 10 μm thick $SiO_2$ insulating layer. $C_c$ and $\beta$ are fitting parameters to approximate the sub-threshold behavior and drain bias induced current.

For a long channel device ($L_g$ >> 100 nm), the wave number $k_l$ can be represented as a continuous variable. By replacing the inner summation with the integral function and assuming $T_{LR}=T_{RL}=T_m$, equation (5.7) can be simplified as,

$$I_{semi}(V_{ch,DS},V_{ch,GS}) \approx \frac{4e^2}{h}\sum_{\substack{k_m \\ m=1}}^{M} T_m \cdot \left[ V_{ch,DS} + \frac{kT}{e}\ln\left( \frac{1+e^{(E_{m,0}-\Delta\Phi_B)/kT}}{1+e^{(E_{m,0}-\Delta\Phi_B+eV_{ch,DS})/kT}} \right) \right] \qquad (5.14)$$

For the limiting case of $T \to 0$, $V_{ch,DS} < \Delta\Phi_B - E_{m,0}$, and $T_m \to 1$, the conductance in equation (5.14) approaches the quantum conductance value, $4e^2/h$ per doubly-degenerated sub-band.

The above equations utilize the approximated SWCNT band structure (Equ. 2, 3) which is valid in the range $E_{m,l} << V_\pi$. This approach also eliminates the difference among SWCNTs with different chiralities but with the same diameter. A more accurate model can be obtained by replacing the simplified band-structure with the tight binding model [21] at the cost of more intensive calculations (~3x), or an exact analytical form valid only for achiral CNTs [22].




(a) (b)

**Figure 5.5:** The comparison of (a) the band structure, and (b) the drain current vs. drain-source bias with different channel surface potential ($\Delta\Phi_B$), calculated by the simple model used in this work (the dotted curves) and the corresponding results calculated using tight binding models with the same CNT diameter (1.5nm) for three different chiralities (the solid curve is for (19,0) CNT, the dot-dashed curve is for (18,2) CNT, and the dashed curve is for (16,5) CNT). For both the band structure and the drain current, the simple model matches well with the tight banding models for the first two subbands with $E_{m,l} < 1.0$eV, and significant discrepancies among the four models are found for the 3rd and higher subbands with $E_{m,l} > 1.0$eV.

Little difference is found for both the *E-k* relationship and the current drive between the model using simplified band-structure and the more complete tight binding model in low energy range (Figure 5.5). The chirality difference for SWCNTs with the same diameter can also be ignored for our purpose, in the range where the carrier energy is less than 1.0

eV (Fig. 5). Thus we use the simplified band-structure in this work for typical circuit applications using a sub-1V power supply.

***$I_{metal}$***: For metallic sub-bands of metallic nanotubes, the current includes both the electron current and the hole current,

$$I_{metal} = 2(1-m0)T_{metal}\sum_{\substack{k_l \\ l=1}}^{L}\left[J_{ele\_0,l} + J_{hole\_0,l}\right] \tag{5.15a}$$

$$J_{ele\_0,l} = \frac{2e}{h}\frac{\sqrt{3}a\pi V_{\pi}}{L_g}\left(f_{FD}(E_{0,l} - \Delta\Phi_B) - f_{FD}(E_{0,l} + eV_{ch,DS} - \Delta\Phi_B)\right) \tag{5.15b}$$

$$J_{hole\_0,l} = \frac{2e}{h}\frac{\sqrt{3}a\pi V_{\pi}}{L_g}\left(f_{FD}(-E_{0,l} - \Delta\Phi_B) - f_{FD}(-E_{0,l} + eV_{ch,DS} - \Delta\Phi_B)\right) \tag{5.15c}$$

The transmission probability $T_{metal}$ is given by,

$$T_{metal} = \frac{\lambda_{ap}\lambda_{op}}{\lambda_{ap}\lambda_{op} + (\lambda_{ap} + \lambda_{op})\cdot L_g} \tag{5.16}$$

If the summation function is replaced with an integral, equation (5.15) can be simplified to,

$$I_{metal} = (1-m0)\frac{4e^2}{h}T_{metal}V_{ch,DS} \tag{5.17}$$

Thus $I_{metal}$ is independent of the channel surface potential change $\Delta\Phi_B$ as expected because the DOS of metallic CNT is independent of the carrier energy. The metallic current is implemented with Equ. (5.17) as a controlled resistor ($G_{metal}$) with conductance $(4e^2/h)\cdot T_{metal}$. The current delivered by the metallic CNTs is a summation of $I_{metal}$ in Equ. (5.17) and $I_{semi}$ in Equ. (5.7). For metallic CNTs of less than 3 nm in diameter, the half band-gap of the first semiconducting subband is larger than 0.43 eV. Considering the large quantum capacitance of metallic CNT (~ 800 aF/μm for the metallic sub-band and the first semiconducting subband) and the typical gate electrostatic capacitance discussed in Section 5.4, the semiconducting subbands in a metallic CNT are not likely to be populated in and thereby contribute to the current with sub-1V power supply.

***$I_{btbt}$*:** In the sub-threshold region, especially with negative gate bias (nFET), the band-to-band tunneling (BTBT) current from drain to source becomes significant. We include a voltage controlled current source $I_{btbt}$ in the device model in order to evaluate the device sub-threshold behavior and the static power consumption.

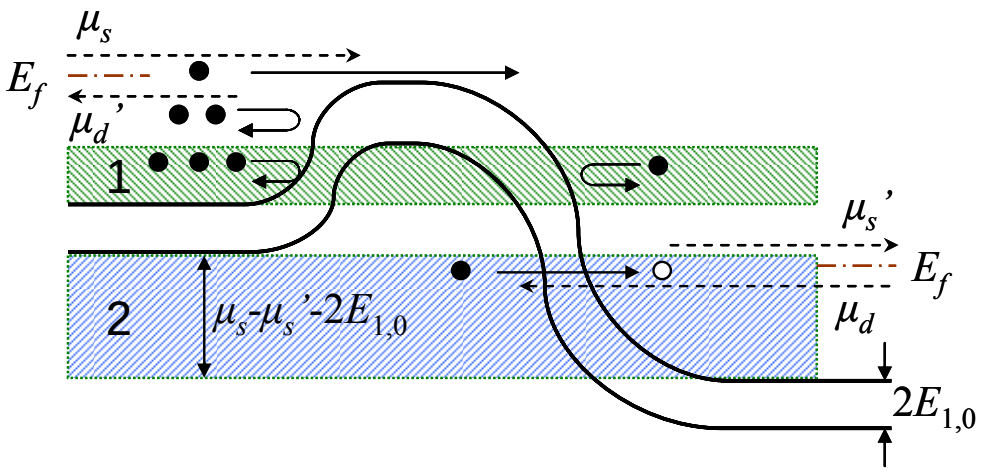



**Figure 5.6:** Energy band diagram (only the first sub-band is shown) and the associated Fermi levels at source/drain side for CNFET with moderate gate and drain bias. There are two possible tunneling regions: region 1 and region 2, which are shaded on the plot. We only consider the tunneling through region 2 in this work.

As shown in Figure 5.6, there are two possible tunneling regions: the "n" shape region 1 and the "L" shape region 2. With $V_{ch,DS} > E_{1,0}$, the tunneling through the drain junction in region 1 causes holes (electrons) pile up in the nFET (pFET) channel region because the source junction prohibits the holes (electrons) from escaping away. The hole (electron) pile up results in surface potential lowering and thereby a higher current and worse sub-threshold behavior [23]. This mechanism depends on the drain junction electrical field and thereby the doping profile. A gradual drain junction doping profile helps relieve this effect by spreading the potential drop over a longer distance. Little such effect is observed for well-tempered devices [14]. To simplify the modeling, we ignore this effect in this work. Because the tunneling through the source junction in region 1 is prohibited, we only consider the BTBT current through the drain junction in region 2. There are two prerequisites for tunneling to occur in region 2: First, the conduction band at the drain side is below the valance band at source side, i.e $V_{ch,DS} > 2E_{m,0}$ where $E_{m,0}$ is the half band gap of the $m^{th}$ sub-band; and second, there are enough empty states at the drain side to accept the carriers that tunnel from the source/channel region. Assuming ballistic

transport for the tunneling process, the BTBT current is approximated by the BTBT tunneling probability ($T_{btbt}$) times the maximum possible tunneling current integrating from the conduction band at drain side up to the valance band at source side,

$$I_{btbt} \approx \sum_{\substack{k_m \\ m=1}}^{M} T_{btbt} \frac{4e}{h} \int_{E_{m,0}-E_f}^{V_{ch,DS}-E_{m,0}-E_f} \left(1 - f_{FD}(E)\right) \cdot dE \tag{5.18}$$

$E_f$ is the Fermi level of the doped source/drain nanotube in units of eV. Applying the Fermi-Dirac function 5(b) in the above equation, we obtain,

$$I_{btbt} = \frac{4e}{h} kT \cdot \sum_{\substack{k_m \\ m=1}}^{M} \left[ T_{btbt} \ln\left( \frac{1 + e^{(eV_{ch,DS} - E_{m,0} - E_f)/kT}}{1 + e^{(E_{m,0} - E_f)/kT}} \right) \cdot \frac{\max(eV_{ch,DS} - 2E_{m,0}, 0)}{eV_{ch,DS} - 2E_{m,0}} \right] \tag{5.19}$$

Following the work of Kane [24,25], the WKB-like transmission coefficient is given by,

$$T_{btbt} \approx \frac{\pi^2}{9} \exp\left( - \frac{\pi m^{*(1/2)} (\eta_m 2E_{m,0})^{3/2}}{2^{3/2} e \cdot \hbar \cdot F} \right) \tag{5.20}$$

The "perpendicular energy" part in [25] is ignored in this case because the motion of the carriers in 1-D quantum wire is confined along the channel direction. $2E_{m,0}$ is the band gap, and $\eta_m$ is a fitting parameter, set to 0.5 in this work, which represents the band gap narrowing effect under high electrical field [26,27]. $m^*$ is the effective electron mass, defined as $\hbar^2/(\partial^2 E_{m,l}/\partial k_l^2)$. Applying equation (5.3) and approximating the effective mass near the bottom of the sub-bands, we obtain,

$$m^* \approx \frac{2\hbar^2 k_m}{\sqrt{3} \cdot a \cdot V_\pi} \tag{5.21}$$

$m^*$ is about $0.05m_o$ and $0.10m_o$ for the carriers in the 1st and the 2nd (semiconducting) sub-band, respectively, where $m_o$ is the electron rest mass. $F$ is the electrical field triggering the tunneling process near the drain side junction, we normalize the total potential drop across the channel-drain junction to a fitting parameter, $l_{relax}$, assuming the potential difference relaxes over the distance $l_{relax}$,

$$F = \frac{V_{ch,DS} + E_f - \Delta\Phi_B}{l_{relax}} \tag{5.22}$$

$l_{relax}$ affects both BTBT current slope and its magnitude. The default value is set to 40 nm to match the BTBT current slope vs. $V_{gs}$ of MOSFET-like CNFET in [1].

### 5.3.2 *Trans-Capacitance Network*

To model the intrinsic ac response of CNFET device, we use a controlled trans-capacitance array among the four electrodes (*G*, *S*, *D*, *B*) with the Meyer capacitor model [28]. The external source/drain nodes (*S'*, *D'*) are lumped with the internal source/drain nodes (*S*, *D*) in this part in order to simplify the analysis. $C_{ox}$, $C_{sub}$, $\beta C_c$, and $(1\text{-}\beta C_c)$ in Figure 5.4 are the electrostatic capacitance considered in this work, and $C_{IJ}$ is the mathematically derived trans-capacitance per unit gate length ($L_g$) between the node *i* and node *j*, defined as $\partial Q_I/\partial V_J$. $C_{IJ}$ has a negative value if $I \neq J$. We use absolute values, i.e. define $C_{IJ}=|\partial Q_I/\partial V_J|$, in the following in order to simply the mathematical expressions. The actual trans-capacitance in the channel region is $C_{ij} = C_{IJ}L_g$ (Figure 5.2(a)).

First, we consider the source/drain capacitance with respect to gate/substrate voltage variation. There are two methods to assign the charges in channel region to the source and the drain: (1) assuming near-ballistic transport in the channel, the carrier distribution along the channel should be almost uniform, i.e. $Q_{s,ch} \approx Q_{d,ch} = Q_{cap}/2 = Q_{CNT}/2$; (2) all the carriers from $+k$ branches are assigned to the source and all the carriers from $-k$ branches are assigned to the drain. The first approach is more reasonable in representing the physical meaning of the capacitor (a carrier reservoir which does not distinguish where the carriers come from), while it may result in $C_{ij} \neq C_{ji}$. For traditional 2-D/3-D drift-diffusion devices, these two approaches produce similar results. We first discuss the former (charge separation) approach which results in the equivalent circuit model in Figure 5.2(a). All the carriers in both the channel region and the source/drain nodes (Figure 5.4) come from the (external) source and drain electrodes, thus $Q_S = L_g \cdot (Q_{cap}/2 + (1-\beta)C_c \cdot \Delta\Phi_B)$ and $Q_D = L_g \cdot (Q_{cap}/2 + \beta C_c(\Delta\Phi_B - V_{DS}))$. We denote the total electrostatic coupling capacitance per unit length between channel and other electrodes as $C_{tot}=C_{ox}+C_{sub}+C_c$. Taking the partial dirivative of $Q_S$ and $Q_D$ over $V_G$, we obtain,

$$C_{sg} = \frac{L_g}{2}\left( C_{ox} - \frac{1}{e}\frac{C_{tot} - 2(1-\beta)C_c}{\partial V_G / \partial \Delta\Phi_B} \right) \tag{5.23a}$$

$$C_{dg} = \frac{L_g}{2}\left( C_{ox} - \frac{1}{e}\frac{C_{tot} - 2\beta C_c}{\partial V_G / \partial \Delta\Phi_B} \right) \tag{5.23b}$$

$\partial V_G/\partial\Delta\Phi_B$ can be calculated by equating $\partial Q_{cap}/\partial\Delta\Phi_B$ and $\partial Q_{CNT}/\partial\Delta\Phi_B$ with fixed $V_{ch,S}$, $V_{ch,D}$ and $V_B$ using Equ. (5.12b, 5.12c),

$$\frac{\partial V_G}{\partial \Delta\Phi_B} = \frac{1}{eC_{ox}}\left(C_{tot} + C_{Qs} + C_{Qd}\right) \tag{5.24a}$$

$$C_{Qs} = \frac{4e^2}{L_g \cdot kT}\sum_{\substack{k_m \\ m=m0}}^{M}\sum_{\substack{k_l \\ l=0}}^{L}\left[ \frac{e^{(E_{m,l}-\Delta\Phi_B)/kT}}{(1+e^{(E_{m,l}-\Delta\Phi_B)/kT})^2} \right] \tag{5.24b}$$

$$C_{Qd} = \frac{4e^2}{L_g \cdot kT}\sum_{\substack{k_m \\ m=m0}}^{M}\sum_{\substack{k_l \\ l=0}}^{L}\left[ \frac{e^{(E_{m,l}-\Delta\Phi_B+eV_{ch,DS})/kT}}{(1+e^{(E_{m,l}-\Delta\Phi_B+eV_{ch,DS})/kT})^2} \right] \tag{5.24c}$$

$m0$ is given by Equ. (5.12d). We define $C_{Qs}$ and $C_{Qd}$ as the quantum capacitance due to the carriers from source ($+k$ branch) and drain ($-k$ branch), respectively. With small gate bias ($E_{m,0} >> \Delta\Phi_B$), $\partial V_G/\partial\Delta\Phi_B \approx C_{tot}/(eC_{ox})$, thus the channel acts as a linear voltage divider which has little dependence on quantum capacitance. With large gate bias ($E_{m,0} < \Delta\Phi_B$), $\partial V_G/\partial\Delta\Phi_B > C_{tot}/(eC_{ox})$, thus the surface potential will be limited by the quantum capacitance. With this observation, it is possible to further simply/approximate the expressions for the quantum capacitance in each segmented gate bias region which is beyond the scope of this work. With equations (5.23, 5.24), we obtain,

$$C_{sg} = \frac{L_g C_{ox}}{2}\frac{C_{Qs} + C_{Qd} + 2(1-\beta)C_c}{C_{tot} + C_{Qs} + C_{Qd}} \tag{5.25a}$$

$$C_{dg} = \frac{L_g C_{ox}}{2}\frac{C_{Qs} + C_{Qd} + 2\beta C_c}{C_{tot} + C_{Qs} + C_{Qd}} \tag{5.25b}$$

We can follow a similar approach to calculate the capacitances $C_{sb}$ and $C_{db}$ as $C_{sb}=C_{sg}\cdot(C_{sub}/C_{ox})$ and $C_{db}=C_{dg}\cdot(C_{sub}/C_{ox})$, respectively.

The charges accumulated on the gate and substrate (back gate) electrodes are given by $Q_G=L_g\cdot C_{ox}\cdot(V_{GS}-V_{FB}-\Delta\Phi_B)$ and $Q_B= L_g\cdot C_{sub}\cdot(V_{BS}-\Delta\Phi_B)$, respectively. With a similar approach, the coupling capacitance between the gate and the substrate is derived as,

$$C_{bg} = C_{gb} = \frac{L_g C_{sub} C_{ox}}{C_{tot} + C_{Qs} + C_{Qd}} \tag{5.26}$$

$C_{sb}$ ($C_{db}$) and $C_{bg}$ ($C_{gb}$) are usually small values given that $C_{sub}$ is much smaller than $C_{ox}$ for typical devices (with thick $SiO_2$ bulk) and $C_{Qs}$ and $C_{Qd}$ are already larger than zero. They are usually much smaller than the direct coupling capacitance between gate and substrate ($C_{gsub}$ in Chapter 6), thus they can be ignored for most applications. But these components should be included for devices with the substrate acting as the back gate.

Next, we consider the gate/substrate capacitance due to source/drain voltage variation. With the similar approach as above, we obtain,

$$C_{gs} = \frac{L_g C_{ox}[C_{Qs} + (1-\beta)C_c]}{C_{tot} + C_{Qs} + C_{Qd}} \tag{5.27a}$$

$$C_{gd} = \frac{L_g C_{ox}(C_{Qd} + \beta C_c)}{C_{tot} + C_{Qs} + C_{Qd}} \tag{5.27b}$$

$$C_{bs} = C_{gs}\frac{C_{sub}}{C_{ox}} \tag{5.27c}$$

$$C_{bd} = C_{gd}\frac{C_{sub}}{C_{ox}} \tag{5.27d}$$

$C_{bs}$ and $C_{bd}$ are only important when the substrate is the driving (switching) gate.

The above equations give the values of the 9 capacitors in Figure 5.2(a). If we use the8 second channel charge separation approach ($+k$ carriers for source and $-k$ carriers for drain), reciprocality is guaranteed and $C_{sg}=C_{gs}$, $C_{dg}=C_{gd}$, $C_{sb}=C_{bs}$, and $C_{db}=C_{bd}$, thus the gate capacitance network can be simply represented by the 5-capacitor model in Figure 5.2(b), or by the 6-capacitor model, as shown in Figure 5.2(c), that shows explicitly the

electrostatic capacitance and the quantum capacitance with the same transfer function as the 5-capacitor model.

## 5.4 Discussion

Figure 5.7 shows the intrinsic channel current with incremental non-idealities. Assuming ballistic transport, there is little difference (< 3%) between the on-current for an infinitely long gate length (calculated with closed form approximation Equ. (5.14)) and the on-current for a 100 nm gate length device, thus it is reasonable to assume the ideal device current drive with gate length longer than 100 nm to be independent of the gate length. With 32 nm gate length, the on-current is about 90% of the long channel value. This slight ballistic current drop from long channel device to short channel device is due to the energy quantization ($k_l$ quantization) in the axial direction. Phonon scattering in the 32 nm long channel region further reduces the on-current by ~7%. BTBT current is only significant with high $V_{ds}$ bias, and the sub-threshold slope gets worse with larger electrostatic capacitance between the channel and the substrate (the inset in Figure 5.7).

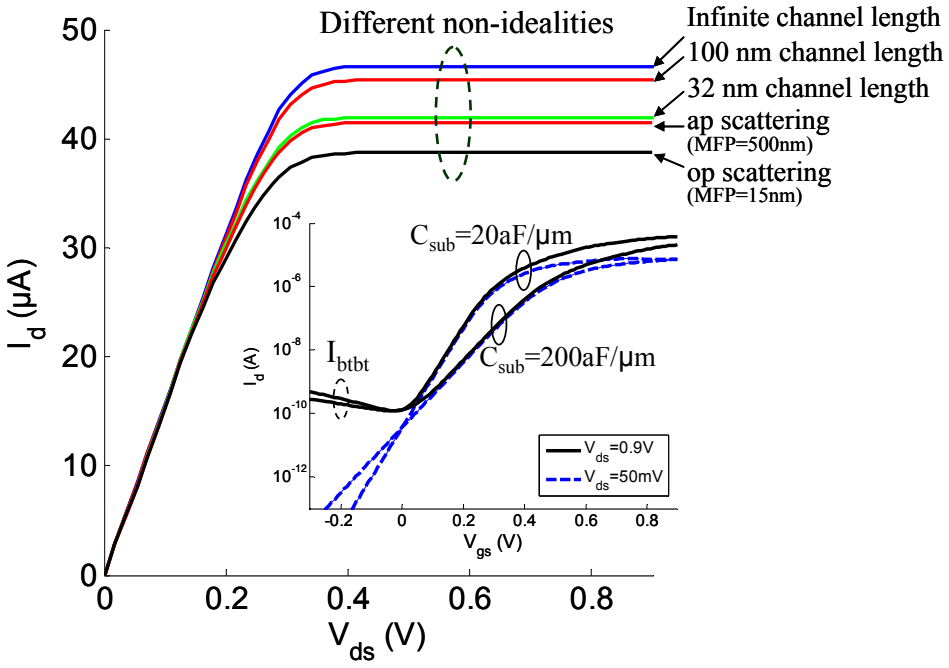



**Figure 5.7:** The drain current @ ($V_{gs}$=0.9V, $V_{FB}$=0V) for (19, 0) chirality CNFET with incremental non-idealities. The front gate dielectric material is 3 nm thick $HfO_2$ on top of 10 μm thick $SiO_2$ insulating layer. Inset plot shows the drain current as a function of $V_{gs}$ with different channel to substrate electrostatic capacitance.

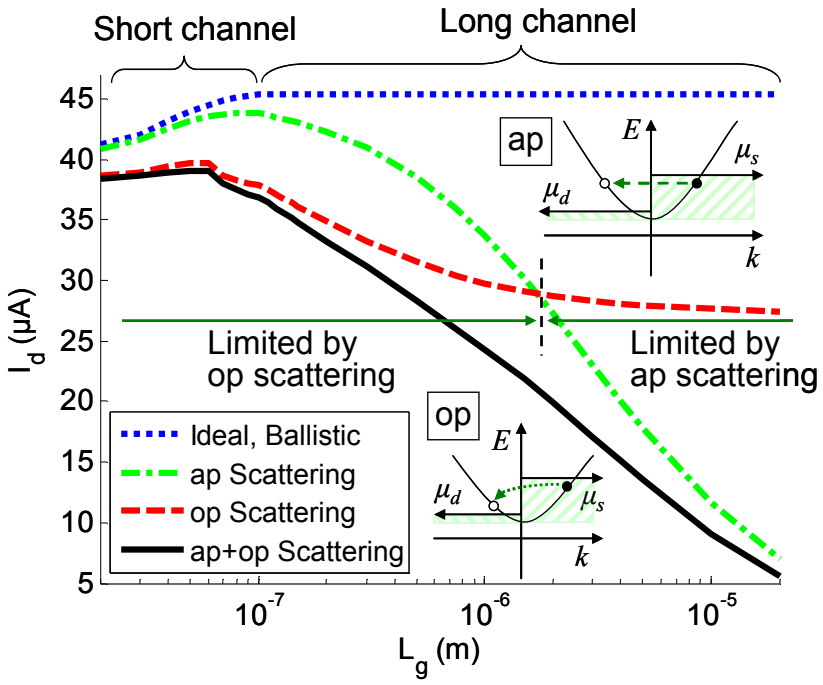



**Figure 5.8:** The on-current @ ($V_{gs}$=$V_{ds}$=0.9V) as a function of the gate length $L_g$. With ballistic transport, the on-current is almost constant for long gate ($L_g$ > 100 nm) CNFET, and there is a slight drop in on-current for short gate ($L_g$ < 100 nm) CNFET due to energy quantization in the axial direction. Optical phonon scattering is important for shorter gate lengths because of its short MFP (~15nm). Acoustic phonon scattering continues to be important as $L_g$ increases.

The intrinsic on-current ($I_{on}$@$V_{ds}$=$V_{gs}$=0.9V) dependence on the gate length is illustrated in Figure 5.8. With ideal ballistic transport, the on-current is almost constant with respect to the gate length except for a slight current drop for short gate lengths ($L_g$ < 100 nm) due to the energy quantization in the axial direction. This small current drop is likely to be smeared out by phonon scattering in practice. Optical phonon scattering depends on the carrier energy. A smaller current means a smaller number of high-energy carriers, and therefore there is less chance that optical phonon scattering can occur. As a result, the current reduction rate with only optical phonon scattering becomes smaller as $L_g$ increases because optical phonon scattering rate decreases as current decreases. Optical phonon scattering is important for short channel device due to its short MFP (~15nm). Acoustic phonon scattering with longer MFP (~500nm) continues to be important as $L_g$ increases as the acoustic phonon energy is small (assuming zero in the model), and therefore acoustic phonon scattering has a weak dependence on the carrier energy. Diffusive transport ($I_{on} \propto 1/L_g$) dominates as $L_g$ increases and acoustic phonon scattering increases.

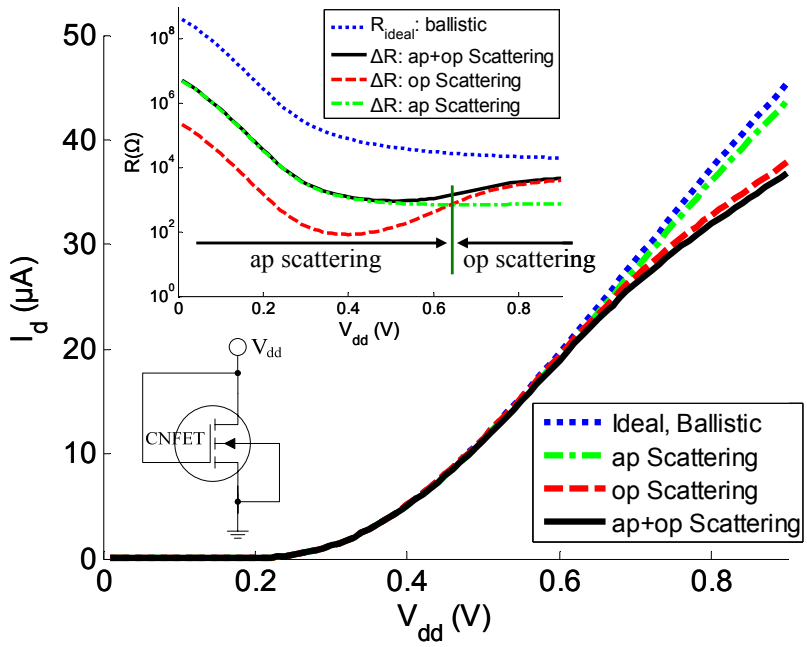



**Figure 5.9:** The on-current as a function of the bias $V_{dd}$ for diode-connected CNFET (the lower inset) with 100 nm gate length. The upper inset shows the corresponding dc resistance. The total channel resistance $R_{tot} = R_{ideal} + \Delta R$. In saturation region, optical phonon scattering increases with bias, thereby the resultant $\Delta R_{op}$ tends to saturate on-current in high bias region.

The intrinsic on-current also depends on the Gate/Drain bias. Figure 5.9 shows both the drain current and the resultant dc resistance (the inset) of a diode-connected CNFET with 100 nm gate length. The total channel dc resistance can be written as $R_{tot}=R_{ideal}+\Delta R$, where $R_{ideal}$ is the quantum resistance which is a function of the band-structure with ballistic transport, and $\Delta R$ denotes the additional resistance contributed by phonon scattering. $R_{ideal}$ is almost constant in the saturation region (the upper inset in Figure 5.9). Acoustic phonon scattering weakly depends on carrier energy, thus acoustic phonon scattering induced $\Delta R$ is almost constant in the saturation region. A constant linear channel resistance approximation is only valid with small bias where optical phonon scattering is not significant. Optical phonon scattering increases as bias increases due to more high-energy carriers. Therefore optical phonon scattering induced $\Delta R$ increases with bias $V_{dd}$. For a first order approximation, it can be approximated as $\Delta R_{op} \approx (V_{dd}-V_{th})/I_o$ which increases as the drain bias (and therefore current) increases [29]. As a result, optical phonon scattering tends to make the on-current saturate with large $V_{dd}$ bias which is consistent with the Monte-Carlo simulation results in [20].

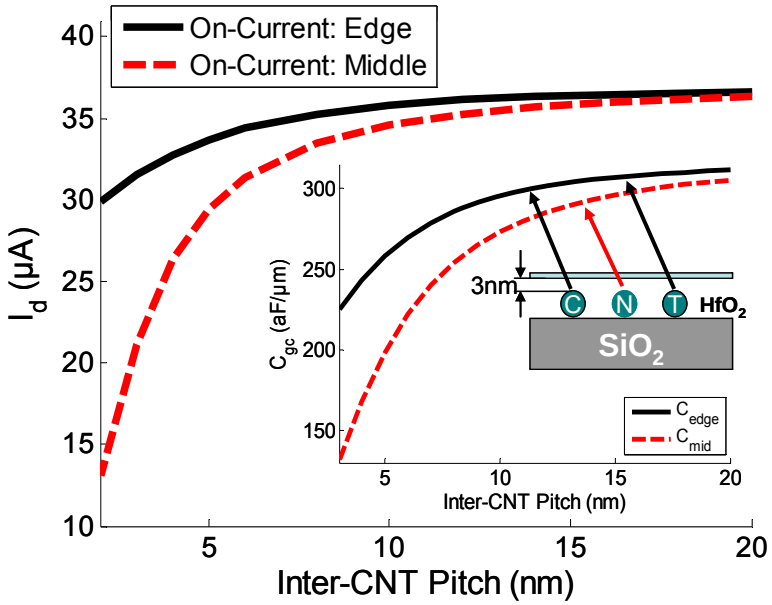
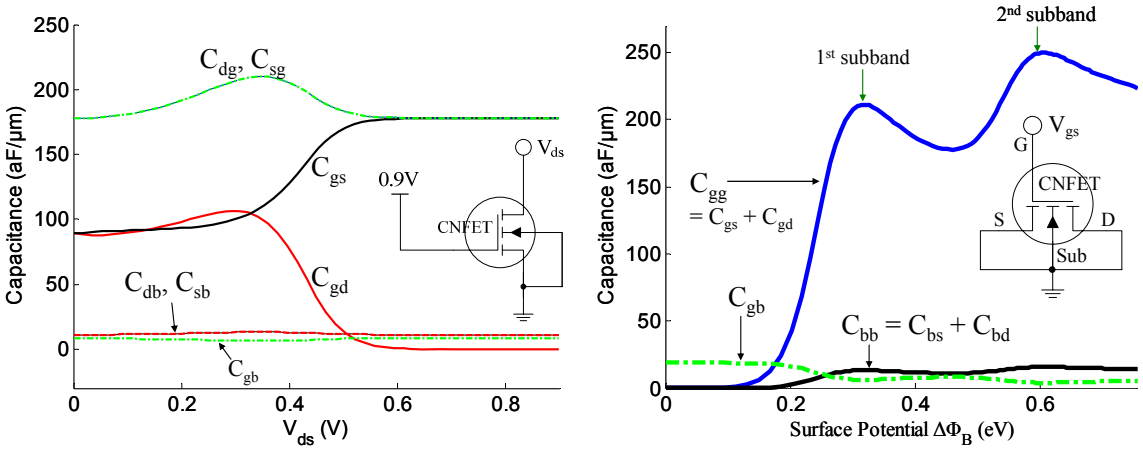



**Figure 5.10:** For CNFET with multiple parallel CNTs, the CNT to CNT screening reduces both the gate to channel electrostatic capacitance (inset) and the drain current. For a typical gate structure with 3 nm thick $HfO_2$ gate dielectric material, the screening effect is easily observable when the inter-CNT pitch is smaller than 20 nm.

For the device with multiple CNTs, screening by the parallel CNTs affects both the gate to channel electrostatic capacitance and the current drive delivered by individual CNTs under the same gate (Figure 5.10). With a typical realistic gate structure (3 nm thick $HfO_2$), the currents carried by the individual CNTs are almost identical if the inter-CNT pitch is larger than 20 nm. With smaller inter-CNT pitch, both the gate capacitance and the current of the CNTs in the middle are smaller than that of the CNTs at the edge due to screening by the adjacent CNTs. A factor of 2 reduction in current can be observed for dense CNT array (~ 2.5 nm inter-CNT pitch). The screening effect should be seriously taken into account when designing high-performance CNFET circuits to avoid overestimating the circuit performance.

Two types of CNFET device connections and the resultant trans-capacitances are illustrated by Figure 5.11. All the capacitances are non-linear components which depend on the bias due to the energy-dependent DOS. With MOS-CAP connection (Figure 5.11(b)), the two peaks correspond to the position of the first two sub-bands (~0.3eV for the 1$^{st}$ subband and ~0.6eV for the 2$^{nd}$ subband, for (19,0) CNT). This property can potentially be used to determine the nanotube diameter once the gate capacitance is

measured [30]. These non-linear properties are particularly important for analog small signal applications.




(a) (b)

**Figure 5.11:** (a) The trans-capacitances $C_{IJ}$ per unit length as a function of $V_{ds}$ @ ($V_{gs}$=0.9V), and (b) the gate and substrate node capacitances ($C_{gg}$ and $C_{bb}$) per unit length as a function of the channel surface potential $\Delta\Phi_B$, for (19,0) semiconducting CNFET at room temperature (T=300K). The flat band voltage is zero. The front gate dielectric is 3 nm thick $HfO_2$ on top of 10 μm thick $SiO_2$ insulting layer.

The gate electrostatic capacitance is about 316 aF/μm CNT length with 3 nm thick $HfO_2$ on top of 10 μm thick $SiO_2$ insulating layer. The total gate capacitance is about 180 aF/μm CNT length with only the first subband populated, thereby the CNT quantum capacitance is calculated to be 420 aF/μm which is larger than or comparable to the typical electrostatic gate capacitance (about 150~350 aF/μm). The CNFET effective gate capacitance for one CNT per gate is about 3.6 aF with 18 nm physical gate length, which is about 4% of the bulk CMOS gate capacitance (predictive BSIM model [31,32,33]) with minimum gate width (48 nm) at the 32 nm node. Considering the large drive current which can be delivered by a single CNT (~35μA@$V_{dd}$=0.9V, ~50% of the bulk n-type MOSFET on-current with 48 nm gate width), the CV/I improvement of intrinsic CNFET over bulk MOSFET device is about 13× better. This large improvement comes from both the much higher carrier velocity of CNT with ballistic transport (the Fermi velocity $v_{F,CNT}$ ≈ $8.0\times10^7$ cm/s, $v_{F,si}$ ≈ $2.5\times10^7$ cm/s), and the large parasitic gate capacitance of MOSFET device. We will show in Chapter 6 that this optimistic performance advantage is not achievable in a practical device structure and will be significantly degraded by the device/circuit non-idealities, including the series resistance of doped source/drain region,

the Schottky barrier (SB) resistance at the metal/CNT interface, the gate outer-fringe capacitance, and the interconnect wiring capacitance.

## 5.5 Summary

We present a circuit-compatible compact model of the intrinsic channel region of MOSFET-like single-walled Carbon Nanotube Field-Effect Transistors (CNFETs) including some channel region non-idealities. The modeled non-idealities include the quantum confinement effects on both circumferential and axial directions, the acoustical/optical phonon scattering in the channel region, and the screening effect by the parallel CNTs for CNFET with multiple CNTs. Comparison with a more accurate device model using the tight binding band structure model shows that this model is valid for CNFET with a wide range of chiralities and diameters. This model uses a sub-state summation approach, instead of the integral, to calculate the parameters. This approach makes the modeling methodology described in this work is generally applicable to other 1-D devices, e.g. silicon nanowire FET, and requires less computation efforts, thereby is more compatible with a circuit simulator. The complete dynamic gate capacitance network makes the model suitable for both small signal (analog) and large signal (digital) applications. Using this model, we project a 13× CV/I improvement of intrinsic CNFET with (19,0) CNT over bulk n-type MOSFET at the 32 nm node. This model serves as a start point towards the complete CNFET device model including device/circuit level non-idealities and multiple CNTs reported in Chapter 6.

## 5.6 Bibliography

[1] A. Javey, J. Guo, D. B. Farmer, Q. Wang, D. Wang, R. G. Gordon, M. Lundstrom, and H. Dai, “Carbon Nanotube Field-Effect Transistors with Integrated Ohmic Contacts and High-K Gate Dielectrics,” *Nano Letters*, vol. 4, pp. 447-450, 2004.

[2] A. Javey, J. Guo, D. B. Farmer, Q. Wang, E. Yenilmez, R. G. Gordon, M. Lundstrom, H. Dai, “Self-Aligned Ballistic Molecular Transistors and Electrically Parallel Nanotube Arrays,” *Nano Letters*, vol. 4, pp. 1319-1322, 2004.

[3] Z. Yao, C. L. Kane, and C. Dekker, “High-field Electrical Transport In Single-Wall Carbon Nanotubes,” *Phys. Rev. Lett.*, vol. 84, pp. 2941–2944, 2000.

[4] D. Mann, A. Javey, J. Kong, Q. Wang, and H. Dai, “Ballistic Transport in Metallic Nanotubes with Reliable Pd Ohmic Contacts,” *Nano Letters*, vol. 3, pp. 1541-1544, 2003.

[5] Y. Cui, Z. Zhong, D. Wang, W. U. Wang, and C. M. Lieber, “High Performance Silicon Nanowire Field Effect Transistors,” *Nano Letters*, vol. 3, pp. 149-152, 2003.

[6] K. Natori, Y. Kimura, and T. Shimizu, “Characteristics of a Carbon Nanotube Field-Effect Transistor Analyzed as a Ballistic Nanowire Field-Effect Transistor,” *Journal of Applied Physics*, vol. 97, pp. 034306, 2005.

[7] J. Guo, M. Lundstrom, and S. Datta, “Performance Projections for Ballistic Carbon Nanotube Field-Effect Transistors,” *Applied Physics Letters*, vol. 80, pp.3192-3194, 2002.

[8] A. Raychowdhury, S. Mukhopadhyay, and K. Roy, “A Circuit-Compatible Model of Ballistic Carbon Nanotube Field-Effect Transistors,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 23, pp. 1411-1420, 2004.

[9] C. Dwyer, M. Cheung, and D. J. Sorin, “Semi-empirical SPICE models for carbon nanotube FET logic,” *4th IEEE Conference on Nanotechnology*, pp. 386-388, 2004.

[10] P. J. Burke, "Lüttinger Liquid Theory as a Model of the Gigahertz Electrical Properties of Carbon Nanotubes," *IEEE Transactions on Nanotechnology*, vol. 1, pp. 129-144, 2002.

[11] A. Naeemi, R. Sarvari, and J. D. Meindl, “Performance Comparison between Carbon Nanotube and Copper Interconnects for Gigascale Integration (GSI),” *Electron Device Letters, IEEE*, vol. 26, pp. 84-86, 2005.

[12] *HSPICE®*, Version 2004.09, Synopsys Corp., CA

[13] A. Javey, R. Tu, D.B. Farmer, J. Guo, R.G. Gordon, H. Dai, “High Performance n-Type Carbon Nanotube Field-Effect Transistors with Chemically Doped Contacts,” *Nano Letters*, vol. 5, pp. 345-348, 2005.

[14] J. Chen, C. Klinke, A. Afzali, and P. Avouris, “Self-aligned carbon nanotube transistors with charge transfer doping,” *Applied Physics Letters*, vol. 86, pp. 123108, 2005.

[15] J. Deng and H.-S P. Wong, “Modeling and Analysis of Planar Gate Capacitance for 1-D FET with Multiple Cylindrical Conducting Channels,” *IEEE Transactions on Electron Devices*, to appear, 2007.

[16] Supriyo Datta, “Electronic Transport in Mesoscopic Systems,” *Cambridge University Press*, Chapter 2, 1995.

[17] J. W. Mintmire and C. T. White, "Universal Density of States for Carbon Nanotubes," *Physical Review Letters*, vol. 81, pp. 2506-2509, 1998.

[18] Heiko Stahl, “Electronic Transport in Ropes of Single Wall Carbon Nanotubes,” *Ph.D. Dissertation, Dept. of Mathematics, Informatics and Natural Sciences, Aachen University of Technology, Germany*, 2000.

[19] A. Javey, J. Guo, M. Paulsson, Q. Wang, D. Mann, M. Lundstrom, and H. Dai, "High-Field Quasiballistic Transport in Short Carbon Nanotubes," *Physical Review Letters*, vo. 92, pp. 106804, 2004.

[20] J. Guo and M. Lundstrom, "Role of Phonon Scattering In Carbon Nanotube Field-Effect Transistors," *Applied Physics Letters*, vol. 86, pp. 193103, 2005.

[21] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, "Physical Properties of Carbon Nanotubes," *Imperial College Press, London*, 1998.

[22] D. Akinwande and H.-S. P. Wong, "An Analytical Tight-Binding Density of States for Achiral Carbon Nanotubes," unpublished, 2007.

[23] J. Knoch, S. Mantl, J. Appenzeller, "Comparison of Transport Properties in Carbon Nanotube Field-Effect Transistors with Schottky Contacts and Doped Source/Drain Contacts," *Solid-State Electronics*, vol. 49, pp. 73-76, 2005.

[24] E. O. Kane, "Zener Tunneling In Semiconductors," *Journal of Physics and Chemistry of Solids*, vol. 12, pp. 181-188, 1959.

[25] E. O. Kane, "Theory of Tunneling," *Journal of Applied Physics*, vol. 32, pp. 83-91, 1961.

[26] J. Geist, J. R. Lowney, "Effect of Band-Gap Narrowing On the Build-In Electrical Field In n-type Silicon," *Journal of Applied Physics*, vol. 52, pp. 1121-1123, 1981.

[27] V. K. Arora, H. Sakaki, "High-Field-Induced Hot-Carrier Temperature, Bandgap Narrowing, and Carrier Multiplication In Bulk Semiconductors," *Proc. SPIE, Conf on Physics and Technology of Semiconductor Devices and Integrated Circuits*, Vol. 1523, p. 160-171, 1992.

[28] T. A. Fjeldly, T. Ytterdal, M. S. Shur, "Introduction to Device Modeling and Circuit Simulation," *Wiley-Interscience, New York*, 1998.

[29] A. Naeemi and J. D. Meindl, "Impact of Electron-Phonon Scattering On the Performance of Carbon Nanotube Interconnects for GSI," *Electron Device Letters, IEEE*, vol. 26, pp. 476-478, 2005.

[30] S. Ilani, L. A. K. Donev, M. Kindermann, and P. L. McEuen, "Measurement of the quantum capacitance of interacting electrons in carbon nanotubes," *Nature Physics*, vol.2 pp. 687-691, 2006.

[31] W. Zhao, Y. Cao, "New generation of Predictive Technology Model for sub-45nm design exploration," *IEEE International Symposium on Quality Electronic Design (ISQED)*, pp. 585-590, 2006.

[32] Y. Cao, T. Sato, D. Sylvester, M. Orshansky, C. Hu, "New paradigm of predictive MOSFET and interconnect modeling for early circuit design," *IEEE Custom Integrated Circuits Conference (CICC)*, pp. 201-204, 2000.

[33] Available at http://www.eas.asu.edu/~ptm/latest.html

# Chapter 6

# FULL DEVICE MODEL AND CIRCUIT PERFORMANCE BENCHMARKING

A version of this chapter has been submitted for publication. J. Deng and H.-S P. Wong, "A Compact SPICE Model for Carbon Nanotube Field Effect Transistors Including Non-Idealities and Its Application — Part II: Full Device Model and Circuit Performance Benchmarking," submitted to *IEEE Transactions on Electron Devices,* 2007.

## 6.1 Introduction

Ballistic or near-ballistic transport is observed with intrinsic CNT under low voltage bias because of the ultra long (~1μm) scattering mean-free-path (MFP) [1,2,3,4]. The quasi-1D structure provides better electrostatic control over the channel region than 3D device (e.g. bulk CMOS) and 2D device (e.g. fully depleted SOI) [5]. These properties make CNFET one of the promising new devices to extend or complement traditional silicon technology [6].

Various CNFET models at the device level have been reported in recent years [7,8,9,10]. Very promising single device, dc performance over silicon CMOS has been demonstrated either by modeling or experimental data. In Chapter 5, we report that there is ~13× CV/I improvement of intrinsic CNFET with (19,0) CNT over bulk n-type MOSFET at the 32 nm node. In order for CNFET to develop into a viable technology, the CNFET performance in a real circuit environment, instead of the intrinsic CNFET device performance, should be evaluated. The parasitic capacitance and resistance of the source/drain region and interconnect are likely to degrade the circuit performance. In addition, the performance degradation due to device and material parameter variations also needs to be considered [11]. Therefore a circuit-compatible CNFET device model

including the device/circuit non-idealities is necessary. In this chapter, we start from the model reported in Chapter 5, and present a complete device model that includes the channel elastic scattering, the doped source/drain region, Schottky barrier (SB) resistance, multiple CNTs per device and other device/circuit non-idealities, and its applications.

This chapter is organized as follows: first, we will describe the organization of this model from the circuit point of view. Next, we show both the mathematical expressions and the circuit representations of each major component. Finally to illustrate the application of this model, we report the circuit performance comparison between CNFET and CMOS circuits at the 32 nm node.

## 6.2 Circuit Topology

A MOSFET-like CNFET device structure (Figure 6.1) is used for the modeling because of both the fabrication feasibility and superior device performance of the MOSFET-like CNFET as compared to the SB-controlled FET. One or multiple devices can be fabricated along a single CNT and multiple CNTs may be placed under the same gate in order to improve the drive current. The CNT channel region is undoped, and the other regions are heavily doped, acting as both the source/drain extension region and/or interconnects between two adjacent devices (un-contacted source-gate/gate-drain configurations). The nanotubes under the gate are grouped into (1) the two CNTs at the edges; (2) the other CNTs in the middle [12].

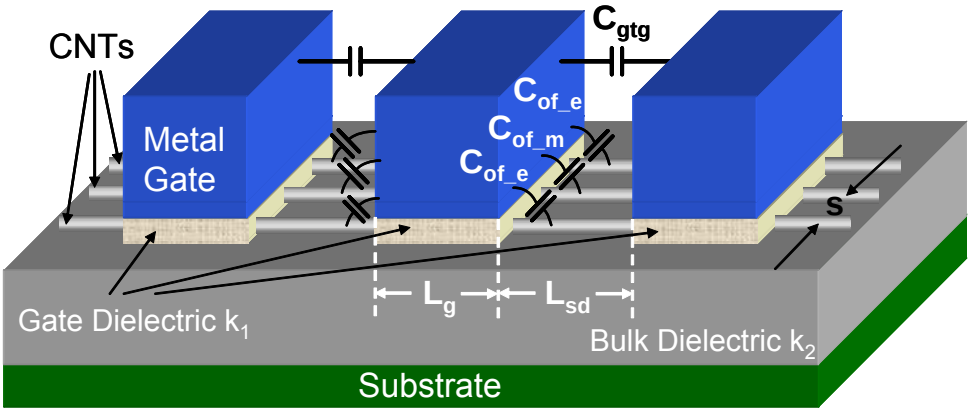



**Figure 6.1:** The 3-D structure of CNFETs with multiple channels, high-*k* gate dielectric material, and the related parasitic gate capacitances. Three CNFETs are fabricated along one single CNT. The channel region of CNTs is un-doped, and the other regions of CNTs are heavily doped.

The complete CNFET device model is implemented hierarchically in three levels (Figure 6.2). Device non-idealities are included hierarchically at each level. Level 1, denoted as CNFET_L1, models the intrinsic behavior of MOSFET-like CNFET. The model at this level is similar to the device-level models such as [7,8]. The second level, denoted as CNFET_L2, includes the device non-idealities: the capacitance and resistance of the doped S/D CNT region, as well as the possible Schottky Barrier (SB) resistances of S/D contacts. The first two levels deal with only one CNT under the gate. The top level, denoted as CNFET_L3, models the interface between CNFET device and CNFET circuits. This level deals with multiple CNTs per device, and includes the parasitic gate capacitance and screening due to adjacent CNTs. This chapter presents the second level and the third level model, as an extension to the first level model, CNFET_L1, in Chapter 5. Only the equations for nFET are described in this chapter, though both nFET model and pFET model are implemented with HSPICE [13]. We use both the Fermi level (chemical potential) and the surface (electrostatic) potential to describe the device behavior in this work, as in Chapter 5.

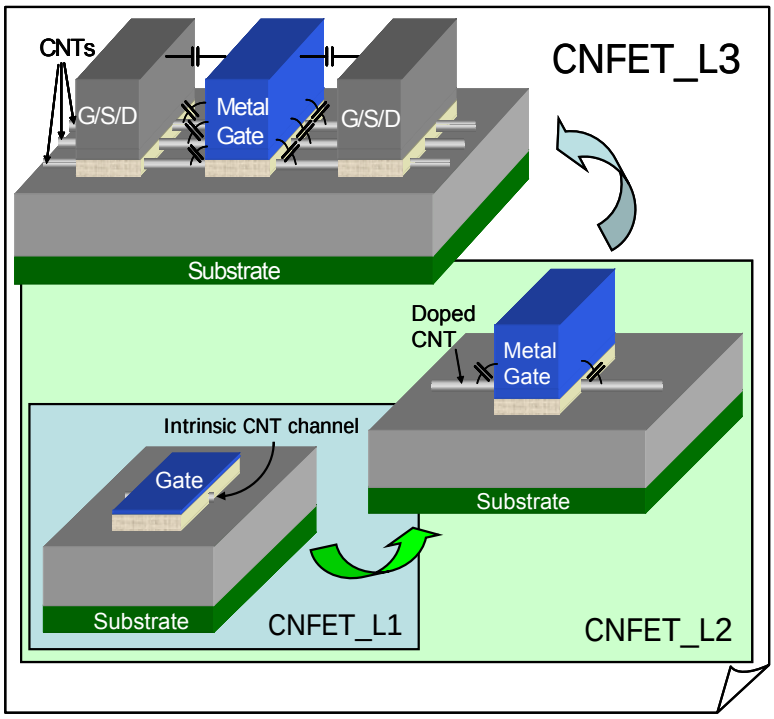



**Figure 6.2:** The complete CNFET device model is implemented with hierarchical three levels. Level 1, CNFET_L1, models the intrinsic behavior of CNFET. The second level, CNFET_L2, includes the device non-idealities. The first two levels deal with only one CNT under the gate. The top level, CNFET_L3, models the interface between CNFET device and CNFET circuits. This level deals with multiple CNTs per device, and includes the parasitic gate capacitance and screening due to adjacent CNTs.

First we estimate with first order approximation if inductance needs to be included in the CNFET device model. Since the magnetic inductance is about 4 orders smaller than the kinetic inductance [14], we ignore the magnetic inductance in this work. The simplest CNT model is a transmission line, as shown in Figure 6.3.

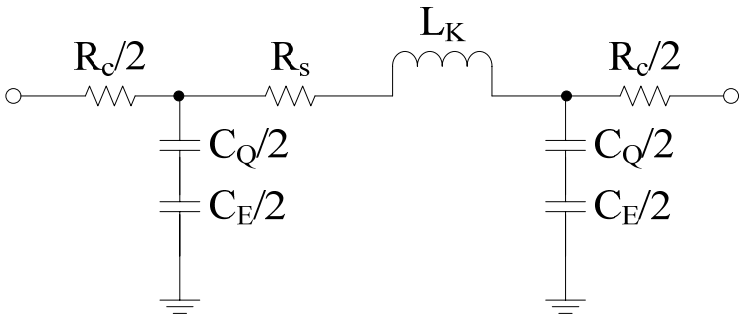



**Figure 6.3:** The simple transmission line model for CNT. $L_K$ is the kinetic inductance, $R_s$ is the series resistance, and $R_c$ is the contact resistance. $C_Q$ and $C_E$ are the quantum capacitance and the electrostatic capacitance, respectively.

With "*N*" modes contributing to the current flow, the quantum resistance is about 25.6/*N* kΩ [15]. Assuming ~1μm elastic scattering mean free path (MFP) for the optimistic intrinsic CNT and ~50nm for the doped CNT, the series resistance ($R_s$) is ~25.6/*N* kΩ/μm and ~512/*N* kΩ/μm, respectively. The kinetic inductance ($L_k$) is ~16/*N* nH/μm, and the quantum capacitance ($C_Q$) is ~100*N* aF/μm [14]. The shunt capacitance ($C_s$) is the series combination of the quantum capacitance ($C_Q$) and the electrostatic capacitance ($C_E$) due to the gate electrode, i.e. $C_s = C_Q C_E/(C_Q+C_E)$. The critical frequencies at which the conductance of the inductor and capacitor become comparable to the series resistance are $f_L=R_s/(2\pi L_k)$ and $f_c=1/(2\pi R_s C_s)$, respectively. $f_L$ is independent of the CNT length, and $f_c$ is a function of the inverse square of CNT length. Both $f_L$ and $f_c$ are independent of the number of modes. There are three typical cases: (1) In the intrinsic channel region: $R_s$ is typically much larger than 25.6/*N* kΩ/μm in the saturation region, say 80/*N* kΩ/μm in the optimal case (~50μA on-current per doubly degenerated subband with 1V power supply), and $C_s \approx C_Q$. Thus $f_L$ ~ 800 GHz, and $f_c$ ~ 20 GHz·μm$^2$. (2) In doped CNT source/drain region: $R_s$ is about 512/*N* kΩ/μm in the linear region and assuming $C_s \approx C_E \approx 30$ aF/μm $<< C_Q$, which gives $f_L$ ~ 5 THz and $f_c$ ~ 10.4*N* GHz·μm$^2$. (3) For intrinsic metallic CNT as an interconnect: $f_L$ ~ 255 GHz and $f_c$ ~ 207*N* GHz·μm$^2$. For CNFET, the inductance becomes significant when the signal frequency is higher than 800 GHz in the channel

region and 5 THz in the source/drain region, respectively, which are far beyond our interested frequency range for typical applications. For metallic CNT interconnect, the kinetic inductance becomes significant above 255 GHz and its effect is more significant than the effect of $C_Q$ with interconnect length shorter than 1.8 μm (assuming $N$=4). Local, device level metallic CNT bridge wiring is unlikely to be as long as 1.8 μm. Thus we ignore the inductance for CNFET device modeling. However, kinetic inductance must be included for metallic CNT and/or larger diameter CNT interconnect modeling.

## 6.3 Device Model – The 2nd Level

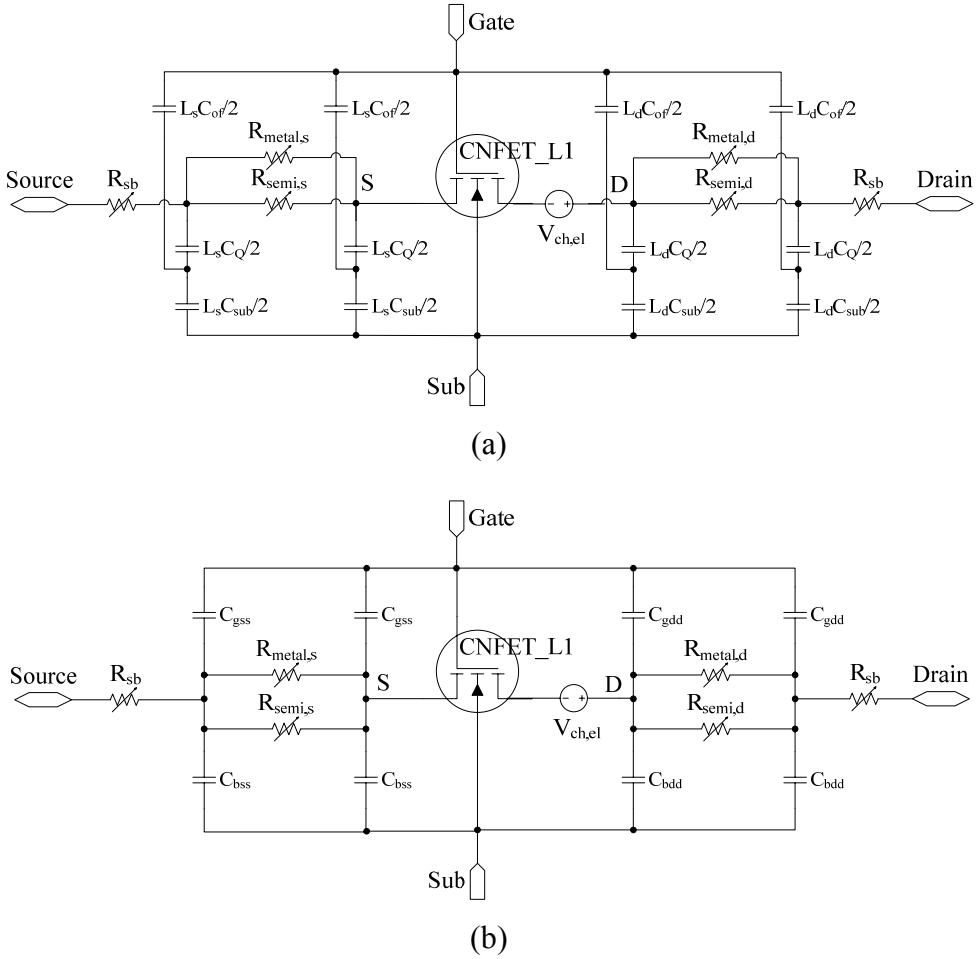



**Figure 6.4:** The 2nd level equivalent circuit model CNFET_L2 for CNFET. (a) The 6-capacitor model, and (b) the 4-capacitor model.

As an extension to the first level CNFET model, CNFET_L1, of the intrinsic channel region modeled in chapter 5, this level models the device non-idealities, including the

elastic scattering in the channel region, the quantum / series resistance and the parasitic capacitance of the doped source/drain region, as well as the Schottky barrier resistance at the interface between the doped CNT and the source/drain metal contacts. The equivalent circuit diagram is shown in Figure 6.4.

### 6.3.1 *Channel Resistance*

We consider three typical scattering mechanisms in the channel region: (1) acoustic phonon scattering (near elastic process [4]), (2) optical phonon scattering (inelastic process [3]), and (3) elastic scattering. Both the acoustic phonon scattering and optical phonon scattering depend on the carrier energy, and are treated in the first level device modeling [1]. The elastic scattering rate and thereby the MFP are assumed to be independent of the carrier energy. We include the elastic scattering in this work in a computationally-efficient way.

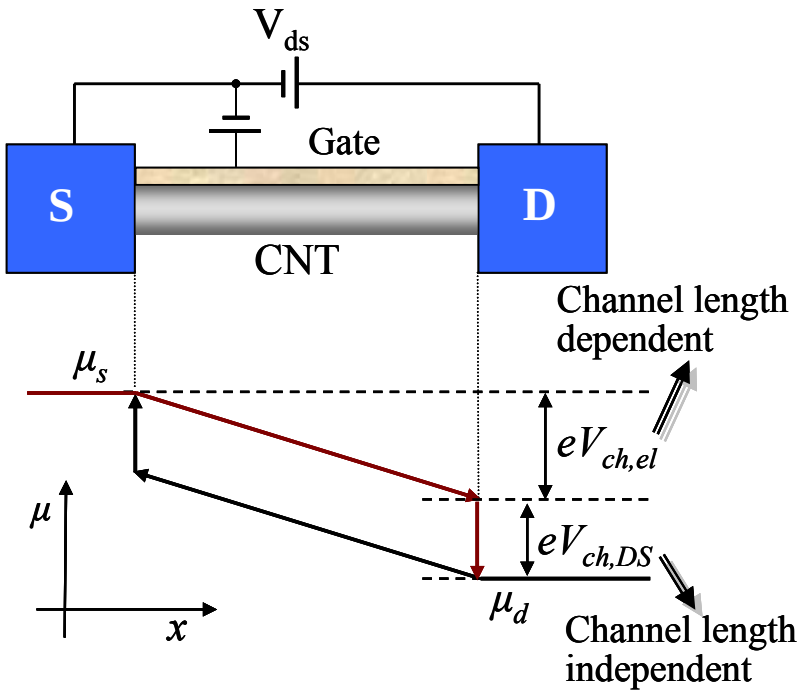



**Figure 6.5:** The Fermi level profile for 1-D device with series resistance (finite MFP along CNT). The Fermi level profile is approximated as a linear function of the position along the channel.

Though the elastic scattering MFP of intrinsic CNT can be longer than 1 μm [2], the fabricated CNTs often contain nonideal scattering centers (e.g. defects) which may degrade the MFP significantly and in turn cause additional potential drop along the

channel region. With a finite scattering MFP, the Fermi level profile along the channel region is illustrated in Figure 6.5. The total potential drop across the channel region is a summation of the potential drop ($V_{ch,DS}$) due to the channel quantum resistance $R_{ch,c}$ (the Fermi level difference near source side because of the finite DOS) and the potential drop ($V_{ch,el}$) over the channel resistance $R_{ch,el}$ due to the elastic scattering, i.e. $V_{DS} = V_{ch,DS} + V_{ch,el}$, $V_{ch,c} = I_{DS} R_{ch,c}$ and $V_{ch,el} = I_{DS} R_{ch,el}$. $R_{ch,el}$ equals to $(1-T_{ch})/T_{ch} \cdot R_{ch,c}$ [15], where $T_{ch}$ is the transmission probability in the channel region, $T_{ch} = l_{eff} / (L_g + l_{eff})$, $L_g$ and $l_{eff}$ are the channel length and the effective elastic scattering MFP, respectively. We further assume the MFP $l_{eff}$ is linearly proportional to the nanotube's diameter [16,17], i.e. $l_{eff} = D_{CNT}/(1.5\text{nm}) \cdot \lambda_{eff}$ where $D_{CNT}$ is the CNT diameter and $\lambda_{eff}$ (~200 nm [18]) is the elastic scattering MFP for 1.5 nm in diameter CNT. With the above equations, we can represent the potential drop over $R_{ch,el}$ as,

$$V_{ch,el} = \frac{L_g}{L_g + \frac{D_{CNT}}{1.5nm} \cdot \lambda_{eff}} V_{DS} \tag{6.1}$$

$V_{DS}$ is the potential drop across the entire channel region between the node S and the node D in Figure 6.4. Representing the effect of the channel resistance due to elastic scattering as a voltage-controlled voltage source $V_{ch,el}$ (Figure 6.4) can avoid calculating $R_{ch,el}$ directly, thereby simplifying the computation.

### 6.3.2 *Doped Source/Drain CNT*

The heavily doped nanotube regions of CNFET act as both the source/drain extension region and the local interconnect between two adjacent devices (Figure 6.1). We use a π model to represent the equivalent circuit of the doped source/drain region (Figure 6.4).

**Resistance**: First, we discuss the model for the resistance of the doped CNT region. Similar to the channel region, CNT can be either metallic or semiconducting. The source/drain resistance is modeled as two paralleled resistors: $R_{semi,s}$ ($R_{semi,d}$) due to the semiconducting sub-bands, and $R_{metal,s}$ ($R_{metal,d}$) due to the metallic sub-bands of metallic nanotubes. We consider two typical cases for device connectivity: (1) the drain of one CNFET is connected to the source of another CNFET, i.e. the doped CNT acts as

interconnect between two devices in series (Figure 6.6(a)) without a metal contact in between; (2) the drain/source of one CNFET is connected to the metal contact, e.g. at the output node (Figure 6.6(b)).

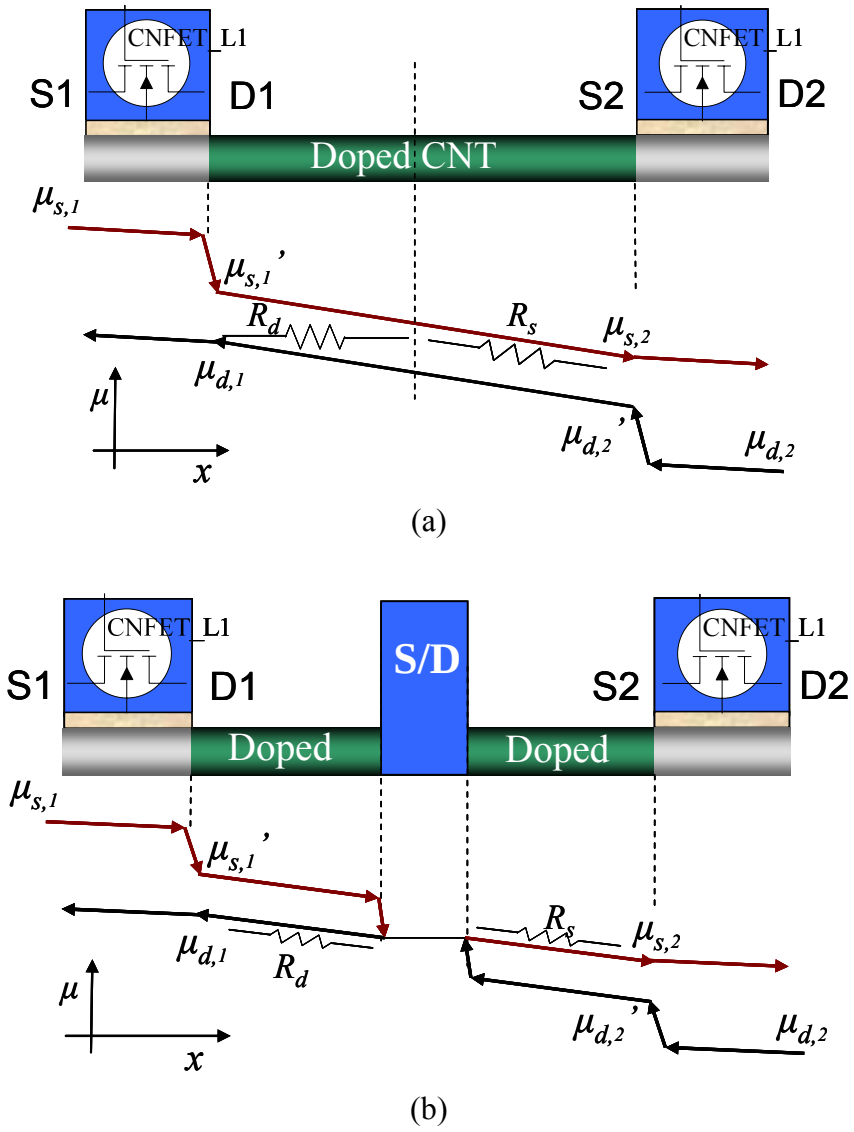



**Figure 6.6:** The related Fermi level profiles for (a) Two CNFETs are connected with doped CNT, and (b) two CNFETs are connected by an ideal metal contact (without considering the Schottky Barriers between CNT and metal interface). Superposed are the equivalent source/drain resistors.

For the first case, the two segment doped nanotubes should be the same as one continuous doped nanotube in the model, i.e. the potential profile along the two segments is continuous (Figure 6.6(a)). Furthermore, we describe the device intrinsic behavior with source/drain input Fermi levels ($\mu_s$, $\mu_d$) as above, while the input source Fermi level $\mu_s$ of

one device is connected to the output source Fermi level $\mu_s'$ of another device (Figure 6.6). Thus we need to convert the output port of the equivalent circuit model from $\mu_d$ to $\mu_s'$ for this case. For the second case, the ideal metal contact is an electron reservoir that has infinite DOS and acts as an infinite scattering center so that there is additional potential drop across the boundary between the doped CNT and the metal contact (called the quantum contact resistance [15]) due to mode mismatch (Figure 6.6(b)). The device model is able to handle both cases correctly as described below.

We define two parameters $S_{out}$ ($D_{out}$) representing the source/drain connectivity: equal to 0 if source (drain) is connected to doped CNT, otherwise equal to 1. Consider the Fermi level profiles for both cases in Figure 6.6, following a similar approach as in section 6.3.1, we obtain the total effective resistance of the doped source/drain region,

$$R_{x,s} = L_s /(\lambda_{sd} G_{x,c}) \tag{6.2a}$$

$$R_{x,d} = \eta_{deff} / G_{x,c} \tag{6.2b}$$

$$\eta_{deff} = \begin{cases} (L_d - \lambda_{sd}) / \lambda_{sd} & , D_{out} = 0 \\ L_d / \lambda_{sd} & , D_{out} = 1 \end{cases} \tag{6.2c}$$

The subscript $x$ denotes either "*semi*" or "*metal*". $L_s$ and $L_d$ are the lengths of the doped source and drain region, respectively. $\lambda_{sd}$ is the impurity scattering MFP, assumed a constant with a default value of 15 nm, a pessimistic estimation, for degenerately doped nanotubes. A longer MFP, 20 nm to 50 nm, can be derived from the work in [19] with charge transfer doping.

$G_{x,c}$ is the quantum conductance of doped CNT. $G_{x,c}$ depends on CNT diameter, the doping level ($E_f$), and the potential difference ($eV_c$=$|\mu_s - \mu_d|$) between the source and drain Fermi levels. For simplicity, we ignore the effect of the coupling capacitance between doped CNT and the other electrodes on doped CNT surface potential (this is valid because the electrostatic coupling capacitance of source/drain extension region is usually much smaller than the quantum capacitance). With applied bias at the two ends, the source and drain Fermi levels splits apart (the inset in Figure 6.7) by $V_c$ which results in carrier re-distribution between $+k$ states and $-k$ states while keeping the total number of carriers $Q_o$ the same. Denoting the surface potential changes by $\Delta\Phi_s$ referred to the

source Fermi level, the total carriers of semiconducting sub-bands per unit length are given by,

$$Q_o = \sum_{\substack{k_m \\ m=1}}^{M} \int_{E_{m,0}}^{E_{\max}} \frac{D(E)}{2} \cdot [f_{FD}(E - E_f - \Delta\Phi_s) + f_{FD}(E - E_f - \Delta\Phi_s + V_c)]dE \tag{6.3}$$

$f_{FD}(E)$ is the Fermi-Dirac distribution function. $D(E)$ is the CNT universal density of states (DOS) [20]. $E_{m,0}$ is the half band gap of the $m^{\text{th}}$ sub-band, and $k_m$ is the wave number due to the quantum confinement in the circumferential direction,

$$D(E) = \begin{cases} D_0 \cdot E / \sqrt{E^2 - E_{m,0}^2} & E > E_{m,0} \\ 0 & E \le E_{m,0} \end{cases} \tag{6.4a}$$

$$E_{m,0} = \frac{\sqrt{3}}{2} a V_\pi k_m \tag{6.4b}$$

$$k_m = \frac{2\pi}{a\sqrt{n_1^2 + n_1 n_2 + n_2^2}} \cdot \lambda \tag{6.4c}$$

$$\lambda = \begin{cases} \dfrac{6m - 3 - (-1)^m}{12} & m = 1,2,... \quad ,\mathrm{mod}(n_1 - n_2, 3) \neq 0 \\ m & m = 0,1,... \quad ,\mathrm{mod}(n_1 - n_2, 3) = 0 \end{cases} \tag{6.4d}$$

$D_o$ is a constant $8/(3\pi V_\pi \cdot d)$ where $d$ is the carbon-carbon bond distance, about 0.144 nm. For heavily doped nanotubes of typical CNFETs, we assume the doping level $E_f$ is above the first semiconducting subband, but does not exceed the third semiconducting subband. At room temperature, $f_{FD}()$ is quite steep, thus the normalized charge $Q_{Ef} = 2Q_o/D_o$ with small $V_c$ is approximately,

$$Q_{Ef} \approx \begin{cases} 2\sqrt{E_f^2 - E_{1,0}^2} & ,E_{1,0} < E_f < E_{2,0} \\ 2\left(\sqrt{E_f^2 - E_{1,0}^2} + \sqrt{E_f^2 - E_{2,0}^2}\right) & ,E_{2,0} \le E_f < E_{3,0} \end{cases} \tag{6.5}$$

At the point $V_c = E_f - E_{1,0} + \Delta\Phi_s$, $Q_{Ef}$ is given by,

$$Q_{Ef} \approx \begin{cases} \sqrt{(E_f + \Delta\Phi_{s,\max})^2 - E_{1,0}^2} & ,E_{1,0} < E_f + \Delta\Phi_{s,\max} < E_{2,0} \\ \sqrt{(E_f + \Delta\Phi_{s,\max})^2 - E_{1,0}^2} + \sqrt{(E_f + \Delta\Phi_{s,\max})^2 - E_{2,0}^2} & ,E_{2,0} \le E_f + \Delta\Phi_{s,\max} < E_{3,0} \end{cases} \tag{6.6}$$

Equating the equations (6.5) and (6.6), we obtain the maximum surface potential change,

$$\Delta\Phi_{s,\max} \approx \begin{cases} \sqrt{Q_{Ef}{}^2 + E_{1,0}^2} - E_f & , E_{1,0} < E_f + \Delta\Phi_{s,\max} < E_{2,0} \\ \dfrac{\sqrt{(E_{2,0}^2 - E_{1,0}^2)^2 + 2(E_{2,0}^2 + E_{1,0}^2)Q_{Ef}{}^2 + Q_{Ef}{}^4}}{2Q_{Ef}} - E_f & , E_{2,0} \le E_f + \Delta\Phi_{s,\max} < E_{3,0} \end{cases} \quad (6.7)$$

$Q_{Ef}$ is given by equation (6.5). We assume the ratio of ($\mu_s$–$E_f$) over ($E_f$–$\mu_d$) is constant respect to $V_c$ in the non-saturation region ($V_c < E_f$–$E_{1,0}$+$\Delta\Phi_{s,\max}$), thus the surface potential change is normalized as,

$$\Delta\Phi_s(V_c) = \Delta\Phi_{s,\max} \frac{\min(V_c, E_f - E_{1,0} + \Delta\Phi_{s,\max})}{E_f - E_{1,0} + \Delta\Phi_{s,\max}} \quad (6.8)$$

Following the similar approach in chapter 5, we obtain the quantum conductance of the semiconducting sub-bands,

$$G_{semi,c}(V_c) = \frac{4e^2}{h} \sum_{\substack{k_m \\ m=1}}^{2} \left[ 1 + \frac{kT}{eV_c} \ln\left( \frac{1 + e^{(E_{m,0} - E_f - \Delta\Phi_s)/kT}}{1 + e^{(E_{m,0} - E_f - \Delta\Phi_s + eV_c)/kT}} \right) \right] \quad (6.9)$$

The analytic model accurately describes the quantum conductance of doped semiconducting CNT as a function of bias provided $E_f > E_{1,0}$, compared to the more accurate numerical simulation results with the tight binding band structure [21] (Figure 6.7). With small drain bias ($V_c < E_f$–$E_{1,0}$+$\Delta\Phi_s$), $G_{semi,c}$ can be approximated as a constant, $G_{semi,c} \approx 4e^2/(\mathrm{m}\cdot h)$, where m is the number of sub-bands below $E_f$. With large drain bias ($V_c > E_f$–$E_{1,0}$+$\Delta\Phi_s$), e.g. in the saturation region, the maximum surface potential change referred to $\mu_s$ will be pinned at $\Delta\Phi_{s,\max}$, and the drain Fermi level will be pushed below the first sub-band which cause a rapid increase in resistance. Therefore the source/drain resistance increases with increasing current (drain bias). On the other hand, the quantum conductance of the metallic sub-bands is almost independent of bias because the DOS is constant, $G_{metal,c} = (1\text{-}m0)4e^2/h$, where $m0$=0 if mod($n_1$-$n_2$,3)=0 with ($n_1$, $n_2$) CNT, otherwise 1. The internal parameter $V_c$ is related to the circuit parameter $V_{series,s}$ or $V_{series,d}$ (the potential drop over the series resistor at the source side or the drain side, respectively) by equation $V_c = V_{series,d}/\eta_{deff} = V_{series,s}\lambda_{sd}/L_s$. With the above equations, we are ready to calculate the source/drain resistance as voltage controlled resistors ($R_{semi,s}$, $R_{semi,d}$, $R_{metal,s}$, $R_{metal,d}$) that are functions of the potential drop across the resistors.

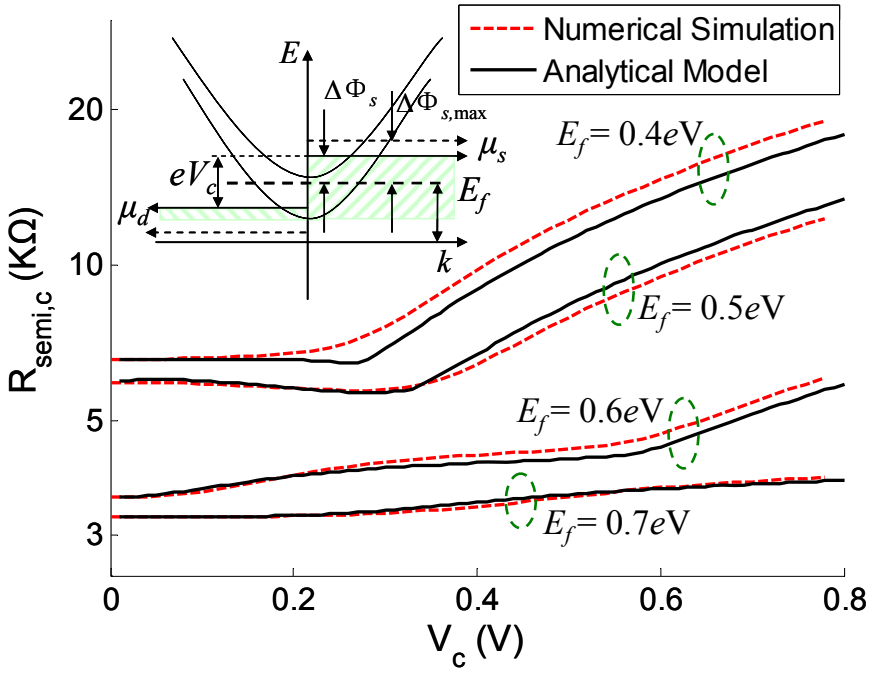



**Figure 6.7:** Comparison of the quantum resistance calculated with the analytical model and the more complete numerical simulations with tight binding band structure, for (19,0) semiconducting CNT with a doping level which varies from 0.4 eV to 0.7 eV. The inset shows the *E*-*k* desperation relationship and the Fermi levels (the solid arrows denote the Fermi levels with smaller $V_c$ and the dashed arrows denote the Fermi levels with larger $V_c$) for doped nanotube with (n-type) doping level $E_f$. The shaded regions are filled with electrons. With applied drain bias, the source Fermi level ($\mu_s$) and drain Fermi level ($\mu_d$) splits apart due to the finite carrier DOS. Referred to $\mu_s$, the surface potential of nanotube changes by $\Delta\Phi_s$. $\Delta\Phi_{s,\max}$ is the maximum surface potential change with applied bias $V_c$.

**Capacitance:** Similar to the channel region modeled in Chapter 5, there are two implementations for the extrinsic capacitor network: (1) the 6-capacitor model consists of the electrostatic capacitance and quantum capacitance (Figure 6.4(a)), or (2) the 4-capacitor model with 4 trans-capacitances (Figure 6.4(b)). The two implementations are equivalent in terms of ac response. The 4 equivalent capacitances can be expressed in terms of the six physical capacitors by,

$$C_{gzz} = \frac{L_z \cdot C_{of} C_Q}{2(C_Q + C_{of} + C_{sub})} \tag{6.10a}$$

$$C_{bzz} = \frac{L_z \cdot C_{sub} C_Q}{2(C_Q + C_{of} + C_{sub})} \tag{6.10b}$$

The subscript "*z*" denotes either "*s*" or "*d*", thus $L_z$ is the length of either source or drain. $C_Q$ is the quantum capacitance of the doped source/drain region. For the channel region,

we need to model $C_Q$ accurately because the gate capacitance ($C_{ox}$) is comparable to $C_Q$. However, for the heavily doped CNT region, $C_Q$ (~ 400 aF/μm per sub-band) is typically much larger than the electrostatic capacitance $C_E$ (typically << 100aF/μm), i.e. the AC performance of this region is limited by $C_E$, instead of $C_Q$. Thus we approximate $C_Q$ to first order as $C_Q = [\Theta(E_f - E_{1,0}) + \Theta(E_f - E_{2,0})] \times 400$ aF/μm, where $\Theta(x)$ is a step function that equals to 1 if $x > 0$, otherwise 0. Thus $C_Q$ of the source/drain region is assumed to be independent of bias. $C_{sub}$ is the coupling capacitance between CNT and the substrate, approximately the same as $C_{sub}$ in channel region. For multiple CNTs per gate, the gate outer-fringe capacitance ($C_{of}$), are grouped into the fringe capacitance between gate and S/D CNT at the two edges ($C_{of\_e}$), and the fringe capacitance between gate and S/D CNT in the middle ($C_{of\_m}$), analytically expressed as [12],

$$C_{of\_e} = \frac{2\pi k_2 \varepsilon_o}{\ln\left(\frac{\sqrt{(2h)^2 + (0.56L_z)^2 + s^2}}{s}\right) + \eta_1 \cdot \cosh^{-1}\left(\frac{\sqrt{(2h)^2 + (0.56L_z)^2}}{2r}\right)} \tag{6.11a}$$

$$C_{of\_m} = \frac{2\alpha}{\eta_1} \cdot C_{of\_e} + \left(1 - \frac{2\alpha}{\eta_1}\right) \cdot \frac{2\pi k_2 \varepsilon_o}{\cosh^{-1}\left(\frac{\sqrt{(2h)^2 + (0.56L_z)^2}}{2r}\right)} \tag{6.11b}$$

$$\eta_1 = \exp\left(\frac{\sqrt{N^2 - 2N} + N - 2}{2.5N}\right), \quad N \geq 2 \tag{6.11c}$$

$$\alpha = \exp\left(\frac{N-3}{2N}\right), \quad N \geq 3 \tag{6.11d}$$

$N$ is the number of CNTs per device, $r$ is the radius of SWCNT and $s$ is inter-CNT spacing. $L_z$ is the length of device source or drain, and should be set to half of the distance between two adjacent gates if two devices are connected in series. These capacitances depend only on the geometry and therefore need to be calculated only once before circuit simulation. For heavily doped CNTs with the first or two subbands populated, $C_Q$ (400 ~ 800 aF/μm) >> $C_{of}$ (~ 30 aF/μm) > $C_{sub}$ (~ 20 aF/μm), thus it is safe to ignore $C_Q$ for most applications.

### 6.3.3 *Schottky Barrier Resistance*

Schottky Barrier (SB) may exist at the interface between CNT and metal contact [22,23,24], and in carbon nanotube heterojunctions between metallic and semiconducting CNTs [25]. In this work, we use a simplified model to describe SB resistance between doped CNT and metallic electrode to include some signature effects of SB on device performance. We made the following simplifying assumptions in this model: (1) the doped CNT region is long enough so that there is no surface potential modification due to the quantum confinement within a short CNT; (2) dipole effects are ignored; (3) no pinning effects [26]; (4) the depletion profile is steep. Figure 6.8 shows an example of the potential profile of SB at the source side. The current flowing through the device is the net result of the carriers tunneling from the metal contact to the doped CNT, and the carriers tunneling from doped CNT to metal contact. In this simple model, we only consider tunneling through the first sub-band, assuming the carriers injecting from the metal contact can re-distribute over all the sub-bands near the contact region. The potential barriers seen by the carriers at the metal contact side ($\Phi_1$) and the doped CNT side ($\Phi_2$) are given by,

$$\Phi_1 = \Phi_M - \Phi_C + E_{1,0} \tag{6.12a}$$

$$\Phi_2 = \begin{cases} \Phi_1 + \left|V_{sb,s}\right| + \Delta\Phi_s & , Source - SB \\ \Phi_1 - \left|V_{sb,d}\right| - V_c + \Delta\Phi_s & , Drain - SB \end{cases} \tag{6.12b}$$

$V_{sb,s}$ and $V_{sb,d}$ are the potential drops over the equivalent SB resistor at the source side and drain side, respectively, from the equivalent circuit point of view. $\Phi_M$ and $\Phi_C$ are the metal work function and CNT work function, respectively. $\Delta\Phi_s$ is the surface potential change before and after bias of the doped CNT referring to the source Fermi level. $\Delta\Phi_s$ and $V_c$ are defined in Section 6.3.2. With positive bias $V_{DS}$ and $V_{GS}$, the carriers of the source side see a higher SB, while the carriers of drain side see a lower SB.

With doping level $E_f$, the normalized volume doping density is,

$$N_D = \frac{32E_{1,0}^2}{3\pi^2 V_\pi^3 d^3}\sqrt{E_f^2 - E_{1,0}^2} \tag{6.13}$$

The depletion length is then approximately,

$$W_d = \sqrt{\frac{2k_2\varepsilon_0}{eN_D}V_{bi}} \tag{6.14a}$$

$$V_{bi} = E_f - E_{1,0} + \Phi_2 \tag{6.14b}$$

$V_{bi}$ is the build in potential with applied bias. For low doping level (< $2\times10^{-4}$), the depletion width for CNT is microns or so, precluding a nanoscale device [24]. At high doping (> $10^{-3}$), the length scale becomes small enough so that the contact is essentially ohmic through tunneling. The conduction band potential profile is,

$$E_C(x) = \frac{eN_D(W_d - x)^2}{2k_2\varepsilon_0} + E_{1,0} - E_f - \Phi_2 + \Phi_1 \tag{6.15}$$

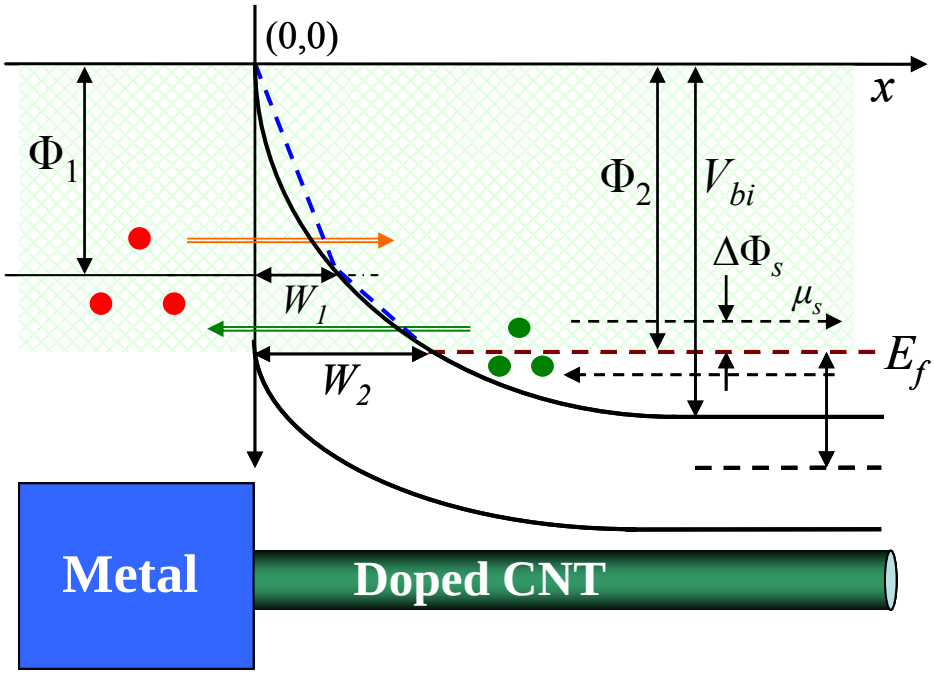



**Figure 6.8:** The energy band diagram for the contacted metal and doped CNT with bias. The potential barrier in the tunneling region (shaded area) is approximated as a triangle potential barrier.

We approximate the potential profile with a triangular potential profile with the same classical turning points, 0, $W_1$ (before bias), and $W_2$ (after bias) (Figure 6.8),

$$W_1 = \sqrt{\frac{2k_2\varepsilon_0}{eN_D}}\left(\sqrt{V_{bi}} - \sqrt{E_f - E_{1,0} + \Phi_2 - \Phi_1}\right) \tag{6.16a}$$

$$W_2 = \sqrt{\frac{2k_2\varepsilon_0}{eN_D}}\left(\sqrt{V_{bi}} - \sqrt{E_f - E_{1,0}}\right) \tag{6.16b}$$

Using the WKB approximation, we obtain the tunneling probability through a triangular potential barrier with height *v* and width *w*,

$$\ln T \cong -2\int_0^w |k(x)|dx = -\frac{4\sqrt{2m^*}w\sqrt{v}}{3\hbar} = -\frac{4w}{9r}\sqrt{\frac{2v}{E_{1,0}}} \tag{6.17}$$

The average transmission probability through the 1st sub-band is approximately,

$$T_{SB} \approx \frac{\int_{\Phi_1}^{\Phi_2} TdE}{\Phi_2 - \Phi_1} \approx \frac{2}{3(\Phi_2 - \Phi_1)\tau^{\frac{2}{3}}}\left[\Gamma(\frac{2}{3}, \tau\cdot\Phi_1^{\frac{3}{2}}) - \Gamma(\frac{2}{3}, \tau\cdot\Phi_2^{\frac{3}{2}})\right] \tag{6.18a}$$

$$\tau \approx \frac{4\sqrt{2}\cdot W_2}{9r\sqrt{E_{1,0}}\cdot\Phi_2} \tag{6.18b}$$

The potential barrier height $\Phi_2$ and thickness $W_2$ depends on the bias. With high degenerate doping which is satisfied in this work, to improve run-time, the above equation can be approximated with,

$$T_{SB} \approx 0.5\times\left(\exp(-\tau\cdot\Phi_1^{\frac{3}{2}}) + \exp(-\tau\cdot\Phi_2^{\frac{3}{2}})\right) \tag{6.19}$$

The equivalent SB series resistance is then given by,

$$R_{sb} = \frac{1}{G_{semi,c}}(\frac{1}{T_{SB}} - 1)(1 - X_{out}) \tag{6.20}$$

The symbol *"X"* denotes either "S" (source side SB) or "D" (drain side SB) which are defined as either 0 or 1 in Section 6.3.2. The SB resistance is modeled as a voltage-controlled resistor (Figure 6.4). The equations in this section are highly simplified and only valid for heavily doped CNT, i.e. $E_f > E_{1,0}$, otherwise the performance of CNFET will be heavily limited by source/drain resistance which is out of the scope of this work[+]. With this model, we can observe that the source (drain) SB resistance increases

[+] For a more accurate model of the short channel SB-controlled CNFET, see [23].

(decreases) with increased source-drain current (incresed $V_{DS}$ and $V_{GS}$) due to the increased (reduced) SB seen by the carriers tunneling from doped CNT to metal contact (Figure 6.9). SB resistance strongly depends on both the metal work function and the CNT doping level, as shown in Figure 6.9. With high doping (~0.8%) and $\Phi_M = \Phi_C = 4.5$ eV, SB resistance can be suppressed to a small value (<1 KΩ) compared to the typical device resistance (~40 KΩ), thus it can be ignored in most applications if this high level of doping can be achieved experimentally [27,28].

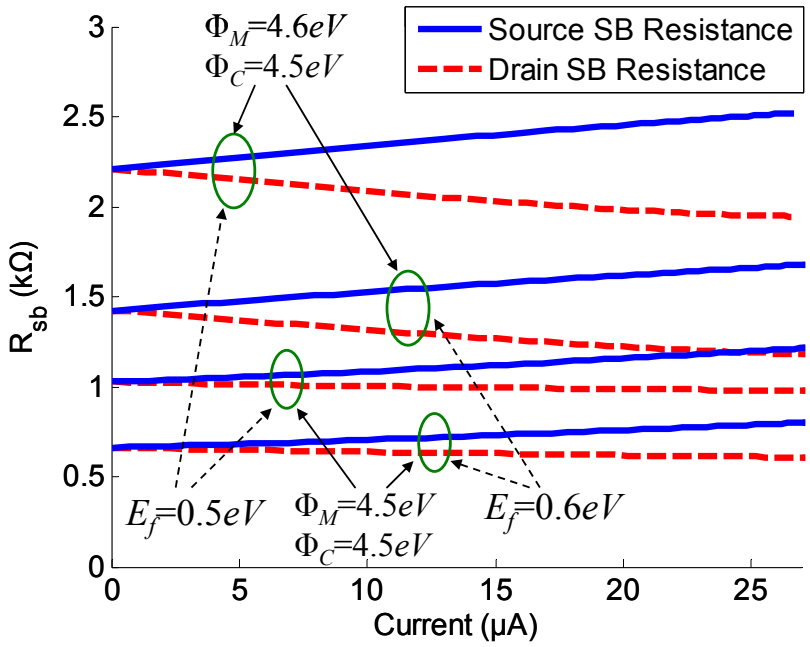



**Figure 6.9:** Schottky Barrier (SB) resistances as a function of the current, with different metal/CNT work functions and different CNT doping levels. Both smaller barrier height and higher CNT doping level help to reduce SB resistance significantly.

## 6.4 Device Model – The 3rd Level

This level is the top level of the device model, which allows for multiple CNTs for each device (Figure 6.10). Consider the case where there are N CNTs under the gate. The CNTs are grouped into (1) a number of min(N,2) CNTs at the two edges; (2) the other (N-min(N,2)) CNTs in the middle. All CNTs in each group are treated identically, and the CNTs in these two groups are connected in parallel in the composite device. By specifying the parameter "s" (the inter-CNT pitch) for each device, the screening effect of

parallel CNTs on both the current drive and the coupling capacitance are taken into account automatically by the device model.

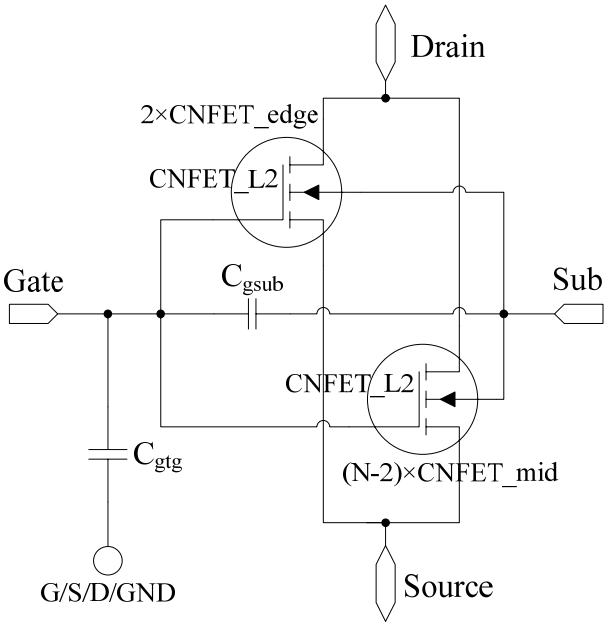



**Figure 6.10:** The 3rd level equivalent circuit model CNFET_L3 for CNFET. There are N nanotubes under the gate. These CNTs are grouped into (1) a number of min(N, 2) CNTs at the two edges; (2) the other N-min(N, 2) CNTs in the middle. All CNTs in each group are treated identically. $C_{gtg}$ is the gate parasitic coupling capacitance connected between the gate and the source/drain/ground or the gate of the adjacent devices, according to the device layout.

The direct coupling capacitance ($C_{gsub}$) between gate and the substrate is simply expressed as $C_{gsub} \approx 2\pi L_g k_2 \varepsilon_0 / \ln(4H_{sub}/H_{gate})$, where $H_{sub}$ is the insulating bulk thickness and $H_{gate}$ is the gate height. $C_{gsub}$ is about 1 aF, about 1/3 of the gate intrinsic capacitance, assuming a 10 μm thick $SiO_2$ bulk, a 64 nm gate height and a 32 nm channel length, and an infinitely large substrate.

## 6.5 Gate and Interconnect Parasitic Capacitance

To be compatible with CMOS process, we assume CNFET circuits use the same conventional metal interconnect technology (the feature size is defined by photolithography) as that for silicon technology (Figure 6.1). The local interconnect wiring resistance is ignored in this work because it is much smaller ($\rho$ = 2~10 μΩ-cm) than typical CNFET device resistance. Consider the layout in Figure 6.1, the gate

parasitic capacitance between gate and the adjacent gate/source/drain contacts per unit length is given by [12],

$$C_{gtg} = \frac{k_2 \varepsilon_o H_{gate} W_g}{L_{sd}} + \frac{0.7\pi k_2 \varepsilon_o W_g}{\ln\left(\frac{2\pi(L_{sd}+L_g)}{2L_g + \tau_{bk} H_{gate}}\right)} \tag{6.21a}$$

$$\tau_{bk} = \exp\left(2 - 2\sqrt{1 + \frac{2(H_{gate}+L_g)}{L_{sd}}}\right) \tag{6.21b}$$

$L_{sd}$ is the distance between the gate and the adjacent gate/source/drain contacts, $H_{gate}$ is the gate height, and $W_g$ is the gate width or the local interconnect length between devices. For devices at 32 nm node ($L_{sd}$=32nm, $L_g$=32nm, $H_{gate}$=64nm, $k_2$=3.9), $C_{gtg}$ is about 110 aF/μm for one side (each gate has two sides). Thus $C_{gtg}$ of such a device with 32 nm gate width is about 11 aF (including Miller effect). This value is more than 2× larger than the intrinsic gate capacitance (~ 4 aF per CNT channel). Therefore, it is very important to include the extrinsic parasitic capacitances for ac performance evaluation.

## 6.6 Device and Circuit Performance

In this section, we compare our model with experimental data and use the model to project circuit performance of CNFET circuits. The above model is implemented in HSPICE[+].

Both the dc and ac performance evaluated with the device model match well with the experimental data of the fabricated CNFET RF device [18]. The observable 10% mismatch between the simulation and the experimental data may due to two reasons: (1) this model uses simplified band structure, instead of the more complete tight binding model. This may introduce inaccuracy for the DOS of the second sub-band which will affect the transconductance. (2) We assume a linear potential profile in both the channel region and the source/drain region which may not be accurate enough for the fabricated device. In spite of these approximations, 10% accuracy should satisfy most analog and

[+] The same model is also implemented in Verilog-A and is available at https://www.stanford.edu/group/nanoelectronics/model_downloads.htm

digital applications, and should be adequate for circuit performance predictions at this point.

Next, we evaluate CNFET device and circuits performance at the 32 nm node with a 0.9 V power apply for high performance logic. All CNTs are assumed to be (19, 0) semiconducting carbon nanotubes with 1.5 nm diameter, and 0.6 eV (~0.8%) source and drain doping level[++], otherwise specified. The gate dielectric is 3 nm thick $HfO_2$ (dielectric constant $k_1$=16) on top of 10 μm thick $SiO_2$ insulating layer. The metal work function is assumed to be the same as the CNT work function (4.5 eV).

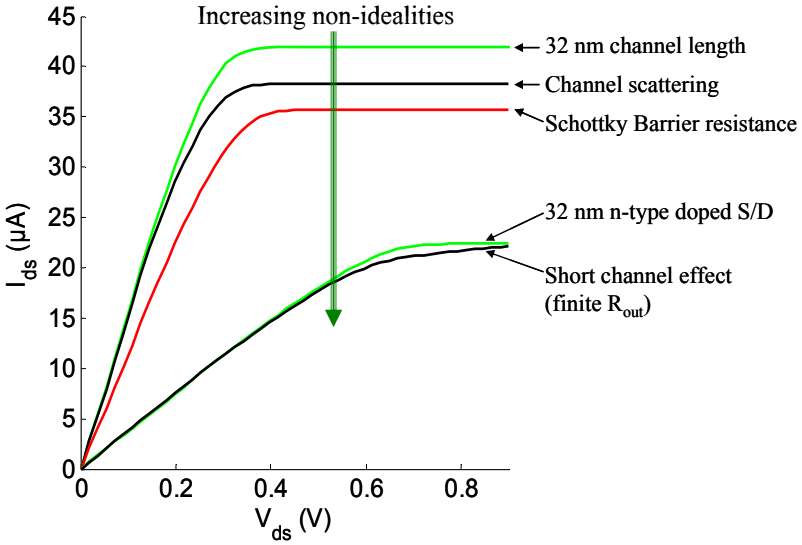



**Figure 6.11:** The drain current @ ($V_{gs}$=0.9V, $V_{FB}$=0V) for (19, 0) chirality CNFET with incremental device non-idealities. The front gate dielectric is 3 nm thick $HfO_2$, and insulating bulk is 10 μm thick $SiO_2$. The metal work function and CNT work function are 4.6eV and 4.5eV, respectively.

Figure 6.11 shows the device current in the presence of non-idealities. The ballistic current of a 32 nm gate length CNFET is about 42 μA. The elastic scattering, acoustic phonon scattering, and optical phonon scattering in the channel region decreases the on-current by about 10%. The SB resistance further reduces the on-current by another 5%. The largest current drive detractor is the source/drain series resistance due to the heavily doped CNTs, which reduces the on-current to 22 μA (about 40% reduction of the on-current including other non-idealities). Compared to silicon bulk CMOS technology (benchmarked with the BSIM4 predictive model [29,30,31]), CNFET shows better single

[++] The effect of a mixture of CNT chirality and doping variations have been reported in [11] using this model.

device performance based on intrinsic CV/I gate delay metric (6× for nFET and 14× for pFET) than MOSFET (where C is the intrinsic gate capacitance) in the 32 nm node, even with device non-idealities (Table 6.1).

**Table 6.1:** The gate effective capacitance and on-current comparison between CNFET and MOSFET at the 32 nm node. The off-currents per gate capacitance for CNFET and MOSFET are adjusted to the same value.

| $L_{Channel}$ =18nm | Gate $C_{eff}$ | $I_{off\ (nA/fF)}$ | $I_{on\ (mA/fF)}$ | $I_{on}/I_{off}$ | CNFET/MOS |
|---|---|---|---|---|---|
| nMOS | 1.1 fF/μm | 383 | 1.2 | $3.1\times10^3$ | N/A |
| nCNFET | 3.6 aF/FET | 383 | 7.2 | $1.9\times10^4$ | ~ 6 |
| pMOS | 1.1 fF/μm | 253 | 0.5 | $2.1\times10^3$ | N/A |
| pCNFET | 3.6 aF/FET | 253 | 7.1 | $2.8\times10^4$ | ~ 14 |

The device performance improvement (~ 6× for nFET) is smaller than the value (~13×) reported in Chapter 5, because the current is degraded by about a factor of two due to the source/drain extension resistance and SB resistance considering the actual device layout.

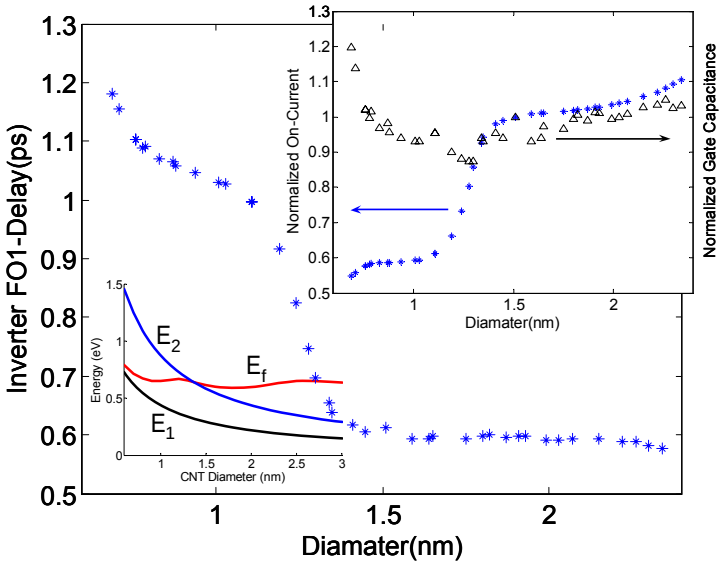



**Figure 6.12:** The FO1 delay vs. CNT diameter for CNFET inverter at the 32 nm node. Gate parasitic capacitance and interconnect capacitance are not included. The lower-left inset shows the first two subbands ($E_1$, $E_2$) and the Fermi level ($E_f$) with 0.8% doping level. The upper-right inset shows both the on-current (lest axes) and the effective gate capacitance (right axes) as a function of CNT diameter. The off-currents of CNFETs are trimmed to the same as that of MOSFET, as shown in Table 6.1.

The CNFET circuit performance depends on CNT diameter. Figure 6.12 shows CNFET inverter fan-out of one (FO1) delay as a function of the CNT diameter (simulated by choosing CNTs with different chiral number ($n_1$, $n_2$) randomly). The off current per unit gate capacitance is set to the same as that of MOSFET by setting the appropriate flat band

gate voltage ($V_{FB}$). The effective gate capacitance is almost independent of CNT diameter (Figure 6.12 inset). This is because only the first subband is likely to be populated in with sub-1V power supply (assuming the same off current). The CNT diameter dependence of the on-current comes from the source/drain resistance, instead of the channel resistance. The Fermi level of doped CNTs with different diameters is almost constant (~ 0.6 eV) for the same doping level (0.8%), as shown by the lower-left inset in Figure 6.12. As a result, for CNTs of less than 1.3 nm in diameter, only the first subband is degenerate which results in a higher source/drain resistance and a smaller current drive and, therefore, lower speed (Figure 6.12). This result illustrates the need to account for the band structure of the CNT correctly in the compact model. It further shows that in order to avoid large variation of circuit speed, the CNT diameter should be targeted away from the 1.25 nm range with 0.8% doping level.

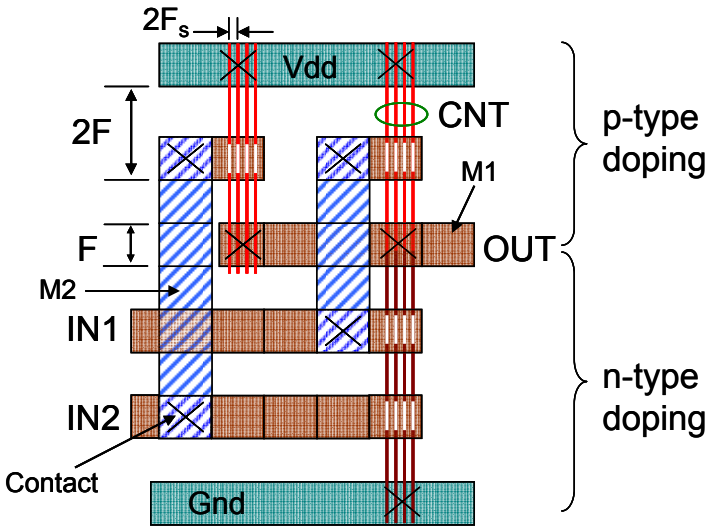



**Figure 6.13:** A sample layout for 2-input CNFET NAND gate with 4 nanotubes per device. The channel regions (white bars) of CNTs are un-doped, and the other regions of the nanotubes are either n-type doped or p-type doped. The coupling capacitances related with the layout pattern are included in the circuit simulation.

For the CNFET inverter with one CNT per device without including gate parasitics and interconnect capacitance, the FO1 delay (~ 0.6 ps) is about 10 times smaller than that of 32 nm bulk CMOS inverter. This result corroborates the performance improvement at the device level reported in Table 6.1. Though large improvement of CNFET over CMOS can be potentially achieved at the device level, the circuit performance is likely to be limited by the circuit parasitics, e.g. the interconnect capacitance. Next, we consider the

realistic layout pattern for CNFET circuits. All the related interconnect capacitances are extracted and included in the circuit simulation. Figure 6.13 shows one example layout for 2-input CNFET NAND gate.

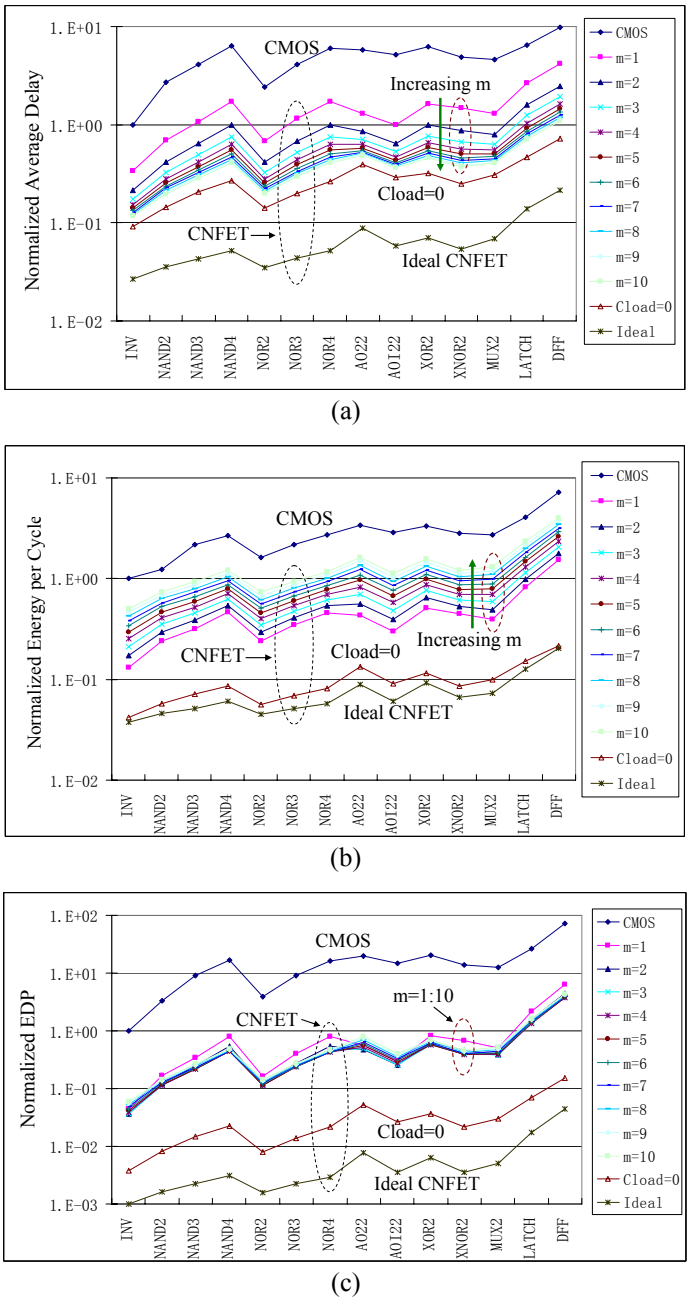



**Figure 6.14:** (a) The circuit speed, (b) dynamic energy per cycle, and (c) the energy-delay-product (EDP) comparison between CMOS circuits and CNFET circuits with realistic layout pattern and interconnect capacitance, in 32 nm technology node. All values are normalized to the performance of a CMOS FO1 inverter. The number of nanotubes per device (m) for CNFET circuits ranges from 1 to 10. Also shown are the relative performances of CNFET circuits (1) without interconnect capacitance (Cload=0), and (2) with ideal CNFET (assuming zero series resistance for doped CNTs, and zero parasitic capacitance).

We now compare CNFET circuit performance with CMOS circuits benchmarked with the standard digital library cells, including INV, NAND, NOR, ADDER, MUX, LATCH, DFF, and so on. The number of CNTs per device, m, is increased from 1 to 10 in order to evaluate the trade-off between speed and energy. The FO1 speed improvement of the ideal CNFET (CNFET_L1 intrinsic device without any parasitics) circuits over bulk CMOS circuits is about 20× to 40× (Figure 6.14(a)). This large number is degraded by a factor of 2 or so, represented by the curve with "Cload=0" in Figure 6.14(a), by the source/drain and SB resistance considering the device layout. After including the interconnect wiring capacitance extracted from the circuit layout pattern, the speed improvement of CNFET circuits with one CNT per device over bulk CMOS circuits is further degraded to ~ 2×, by another factor of 5 or so, because of the small CNT quantum-capacitance-limited effective gate capacitance. Increasing the number of CNTs per device illustrates the trade-off between the speed and energy consumption. With 1 to 10 CNTs per device, the FO1 speed of CNFET circuits is about 2× to 10× faster compared to CMOS circuits (Figure 6.14(a)), the energy consumption per cycle is about 7× to 2× lower (Figure 6.14(b)), and the energy-delay product is about 15× to 20× lower (Figure 6.14(c)). Large advantages of CNFET circuits over CMOS circuits can be potentially achieved even with the traditional (Cu) interconnect technology.

## 6.7 Summary

We present a circuit-compatible compact model for single-walled Carbon Nanotube Field-Effect Transistors (CNFETs), as an extension to the model in Chapter 5. A universal model including the practical device non-idealities is implemented with HSPICE. In addition to the non-idealities included in Chapter 5, this chapter includes the elastic scattering in the channel region, the resistive source/drain (S/D), the Schottky Barrier (SB) resistance, and the parasitic gate capacitances. More than one CNT per device is allowed, and the screening effect by the parallel channels is also included in the device model. Good agreement for both dc and ac characteristics between the device model and the experimental data has been verified with the fabricated CNFET RF device.

The source/drain resistance and SB resistance degrade CNFET on-current by a factor of 2 at the 32 nm node. Compared to silicon technology, CNFET shows better device performance (based on the intrinsic CV/I gate delay metric (6× for nFET and 14× for pFET)) than MOSFET device at the 32 nm node, with device non-idealities. This large speed improvement is significantly degraded (~5× degradation) by interconnect capacitance in a real circuit environment. Increasing the number of CNTs per device is the most effective way to improve the circuit speed. Compared to CMOS circuits, CNFET circuits with 1 to 10 CNTs per device is about 2× to 10× faster, the energy consumption per cycle is about 7× to 2× lower, and the energy-delay product (EDP) is about 15× to 20× lower, considering the realistic layout pattern and the interconnect capacitance.

Further improvements to the implemented device model may include the following: (1) This model utilizes a simplified band structure which restricts the use of this model for the applications that requires a high power supply and high CNT surface potential (>>1.0eV). A more complete tight binding model can alleviate this issue at the cost of longer run time. Separating the operation region into multiple sections and deriving approximated analytical equations in each section is another way to both enlarging the applicable range and improving the run time. (2) For a better sub-threshold behavior modeling, the surface potential lowering and consequent higher current caused by the holes (electrons) pile up in the nFET (pFET) channel region should be considered, especially in the high bias region ($V_{ch,DS} > E_{1,0}$) [32]. This effect is similar to floating body effect for PDSOI MOSFET and depends on the drain junction doping profile. A gradual doping profile can alleviate this effect by relaxing the potential drop over a longer distance and reducing band-to-band tunneling (BTBT) through the drain junction. (3) We ignored the diffusion capacitance due to the minor carriers at the source/drain junctions. Though this is a higher order effect, the diffusion capacitance may affect the ac response of small signal analog circuits. More accurate device modeling for analog circuit applications should be done and verified with experiment data. (4) The performance of CNFET circuits is severely degraded by the traditional interconnect technology. The resistance of heavily doped CNT also degrades the on-current. One way

to further improve CNFET circuit performance is to use metallic CNTs, multi-walled CNTs, or intrinsic large diameter CNTs as local interconnects because of the much higher current density and much smaller parasitic fringe capacitance. Thus a simple and universal interconnect model, similar to the model by [33], is necessary to evaluate all-CNT CNFET circuit performance. This interconnect model should be able to handle both metallic carbon nanotubes and semiconducting carbon nanotubes, as well as single-walled CNT and multiple-walled CNT. The kinetic inductance should also be included as discussed in Section 6.2. (5) A more accurate device model should also include the defect and device reliability analysis as a function of bias and operation time. Most of the carrier scattering and thermal relaxation processes occur around the contact/junction region due to the near-ballistics transport, thus defects are likely to accumulate along the nanotubes especially around the contact region for short gate CNFET. Circuit performance may suffer from this gradually degraded current drive.

## 6.8 Bibliography

[1] A. Javey, J. Guo, D. B. Farmer, Q. Wang, D. Wang, R. G. Gordon, M. Lundstrom, and H. Dai, "Carbon Nanotube Field-Effect Transistors with Integrated Ohmic Contacts and High-K Gate Dielectrics," *Nano Letters*, vol. 4, pp. 447-450, 2004.

[2] A. Javey, J. Guo, D. B. Farmer, Q. Wang, E. Yenilmez, R. G. Gordon, M. Lundstrom, H. Dai, "Self-Aligned Ballistic Molecular Transistors and Electrically Parallel Nanotube Arrays," *Nano Letters*, vol. 4, pp. 1319-1322, 2004.

[3] Z. Yao, C. L. Kane, and C. Dekker, "High-field Electrical Transport In Single-Wall Carbon Nanotubes," *Physical Review Letters*, vol. 84, pp. 2941–2944, 2000.

[4] D. Mann, A. Javey, J. Kong, Q. Wang, and H. Dai, "Ballistic Transport in Metallic Nanotubes with Reliable Pd Ohmic Contacts," *Nano Letters*, vol. 3, pp. 1541-1544, 2003.

[5] Y. Cui, Z. Zhong, D. Wang, W. U. Wang, and C. M. Lieber, "High Performance Silicon Nanowire Field Effect Transistors," *Nano Letters*, vol. 3, pp. 149-152, 2003.

[6] H.-S. P. Wong, J. Appenzeller, V. Derycke, R. Martel, S. Wind, Ph. Avouris, "Carbon Nanotube Field Effect Transistors – Fabrication, Device Physics, and Circuit Implications", *IEEE International Solid State Circuits Conference (ISSCC)*, p. 370 – 371, 2003.

[7] K. Natori, Y. Kimura, and T. Shimizu, "Characteristics of a Carbon Nanotube Field-Effect Transistor Analyzed as a Ballistic Nanowire Field-Effect Transistor," *Journal of*

*Applied Physics*, vol. 97, pp. 034306, 2005.

[8] J. Guo, M. Lundstrom, and S. Datta, "Performance Projections for Ballistic Carbon Nanotube Field-Effect Transistors," *Applied Physics Letters*, vol. 80, pp. 3192-3194, 2002.

[9] A. Raychowdhury, S. Mukhopadhyay, and K. Roy, "A Circuit-Compatible Model of Ballistic Carbon Nanotube Field-Effect Transistors," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 23, pp. 1411-1420, 2004.

[10] C. Dwyer, M. Cheung, and D. J. Sorin, "Semi-empirical SPICE models for carbon nanotube FET logic," *4th IEEE Conference on Nanotechnology*, pp. 386-388, 2004.

[11] J. Deng, N. Patil, K. Ryu, A. Badmaev, C. Zhou, S. Mitra, H.-S. P. Wong, "Carbon Nanotube Transistor Circuits: Circuit-Level Performance Benchmarking and Design Options for Living with Imperfections," *International Solid State Circuits Conference (ISSCC)*, p. 70 – 71, 2007.

[12] J. Deng and H.-S P. Wong, "Modeling and Analysis of Planar Gate Capacitance for 1-D FET with Multiple Cylindrical Conducting Channels," *IEEE Transactions on Electron Devices*, to appear, 2007.

[13] *HSPICE®*, Version 2004.09, Synopsys Corp., CA

[14] P. J. Burke, "Lüttinger Liquid Theory as a Model of the Gigahertz Electrical Properties of Carbon Nanotubes," *IEEE Transitions on Nanotechnologies*, vol. 1, pp. 129-144, 2002.

[15] Supriyo Datta, "Electronic Transport in Mesoscopic Systems," *Cambridge University Press*, Chapter 2, 1995.

[16] C. T. White and T. N. Todorov, "Carbon Nanotubes as Long Ballistic Conductors," *Nature*, vol. 393, pp. 240–242, 1998.

[17] J. Jiang, J. Dong, H. T. Yang, and D. Y. Xing, "Universal Expression for Localization Length in Metallic Carbon Nanotubes," *Physical Review B*, vol. 64, 045409, 2001.

[18] I. Amlani. J. Lewis, K. Lee, R. Zhang, J. Deng, H.-S. P. Wong, "First Demonstration of AC Gain From a Single-Walled Carbon Nanotube Common-Source Amplifier," *IEEE International Electron Devices Meeting (IEDM)*, pp. 559-562, San Francisco, CA, December 11-13, 2006.

[19] J. Chen, C. Klinke, A. Afzali, and P. Avouris, "Self-aligned carbon nanotube transistors with charge transfer doping," *Applied Physics Letters*, vol. 86, pp. 123108, 2005.

[20] J. W. Mintmire and C. T. White, "Universal Density of States for Carbon Nanotubes," *Physical Review Letters*, vol. 81, pp. 2506-2509, 1998.

[21] R. Saito, G. Dresselhau, and M. S. Dresselhaus, "Physical Properties of Carbon Nanotubes," *Imperial College Press, London*, 1998.

[22] T. Nakanishi and T. Ando, "Contact between Carbon Naonotube and Metallic Electrode," Journal of the Physical Society of Japan, vol. 69, pp. 2175-2181, 2000.

[23] A. Hazeghi, T. Krishnamohan and H.-S. P. Wong, "Schottky-Barrier Carbon Nanotube Field-Effect Transistor Modeling," *IEEE Transactions on Electron Devices,* vol. 54, pp. 439-445, 2007.

[24] F. Léonard and J. Tersoff, "Novel Length Scales in Nanotube Devices," *Physical Review Letters*, vol. 83, pp. 5174–5177, 1999.

[25] A. A. Odintsov, "Schottky Barriers in Carbon Nanotube Heterojunctions," Physical Review Letters, vol. 85, pp. 150-153, 2000.

[26] F. Léonard and J. Tersoff, "Role of Fermi-Level Pinning in Nanotube Schottky Diodes," vol. 84, *Physical Review Letters,* vol. 84, pp. 4693-4696, 2000.

[27] M. J. O'Connell, E. E. Eibergen and S. K. Doorn, "Chiral Selectivity in the Charge-Transfer Bleaching of Single-Walled Carbon-Nanotube Spectra," *Nature Materials*, vol. 4, pp. 412 – 418, 2005.

[28] T. Takenobu, T. Takano, M. Shiraishi, Y. Murakami, M. Ata, H. Kataura, Y. Achiba and Y. Iwasa, "Stable and Controlled Amphoteric Doping by Encapsulation of Organic Molecules Inside Carbon Nanotubes," *Nature Materials* vol. 2, pp. 683 – 688, 2003.

[29] W. Zhao, Y. Cao, "New generation of Predictive Technology Model for sub-45nm design exploration," *IEEE International Symposium on Quality Electronic Design (ISQED)*, pp. 585-590, 2006.

[30] Y. Cao, T. Sato, D. Sylvester, M. Orshansky, C. Hu, "New paradigm of predictive MOSFET and interconnect modeling for early circuit design," *IEEE Custom Integrated Circuits Conference (CICC)*, pp. 201-204, 2000.

[31] Available at http://www.eas.asu.edu/~ptm/latest.html

[32] J. Knoch, S. Mantl, J. Appenzeller, "Comparison of Transport Properties in Carbon Nanotube Field-Effect Transistors with Schottky Contacts and Doped Source/Drain Contacts," *Solid-State Electronics*, vol. 49, pp. 73-76, 2005.

[33] A. Naeemi and J. D. Meindl, "Compact Physical Models for Multiwall Carbon-Nanotube Interconnects," *IEEE Electron Device Letters*, vol. 27, pp. 338-340, 2006.

# Chapter 7

# CIRCUIT PERFORMANCE BENCHMARKING INCLUDING IMPERFECTIONS

 Part of this chapter has been published, reprinted with permission, from J. Deng, N. Patil, K. Ryu, A. Badmaev, C. Zhou, S. Mitra, H.-S. P. Wong, "Carbon Nanotube Transistor Circuits: Circuit-Level Performance Benchmarking and Design Options for Living with Imperfections," *International Solid State Circuits Conference (ISSCC)*, pp 70-71, San Francisco, CA, February, 2007.

The work in this chapter has been performed in close collaboration with Nishant Patil, a graduate student advised by Prof. Subhasish Mitra at Stanford. Part of this collaboration is reported in this thesis. Other parts of this collaboration will be reported in the Ph.D. thesis of Nishant Patil.

## 7.1 Introduction

Carbon Nanotubes Field Effect Transistors (CNFETs) are promising candidates as extensions to Si CMOS due to excellent CV/I device performance [1,2]. CNFETs can be categorized as either Schottky Barrier (SB) controlled FET or MOSFET-like FET [3,4,5] according to the operation mechanism. Considering both the fabrication feasibility [6] and superior device ac performance of the MOSFET-like CNFET, as discussed in previous chapters, as compared to the SB-controlled FET, MOSFET-like CNFETs are used for the performance evaluation in this work.

At the 32 nm node, 13× CV/I improvement of intrinsic CNFET with (19,0) CNT over bulk n-type MOSFET due to near-ballistic transport [3] is projected in Chapter 5. While similar numbers are often quoted in the literature [2,3], these numbers are optimistic because local interconnect capacitances and CNT imperfections are not included. Considering the realistic layout pattern, the related source/drain series resistance, and the interconnect wiring capacitance, CNFET circuits with 1 to 10 CNTs per device is expected to be about 2× to 10× faster, the energy consumption per cycle is about 7× to 2× lower, and the energy-delay product (EDP) is about 15× to 20× lower than the CMOS counterparts at the 32 nm node as reported in Chapter 6. Though large performance

advantages over MOSFET technology are projected, these analyses assume CNFETs that are made of homogeneous, identical CNTs, i.e. well-aligned parallel semiconducting CNTs that have the same chirality, the same doping level, and the same inter-CNT pitch. With practical CNT fabrication techniques, both the aligned-growth-CNT and the CNT chirality control, and thereby the CNT diameter control and the electrostatic property control, are big challenges [7]. These imperfections cause circuit performance variation which is important to nanoscale circuit design. The motivation of this research is: given the opportunities with CNFET devices – fine pitch, excellent CNFET device characteristics – what can be gained at the circuit-level compared to cutting-edge Si CMOS?

This chapter is organized as follows: First, we analyze the speed and power of CNFET circuits and their tolerance to the imperfections inherent to CNFET synthesis such as misaligned and metallic CNTs and CNT diameter and doping variations. Next, we demonstrate design principles for designing CNFET-based logic circuits that are guaranteed to implement correct logic functions even in the presence of misaligned CNTs. Finally, we evaluate the scalability of CNFET compared to MOSFET technology.

## 7.2 Circuit Performance Including Imperfections

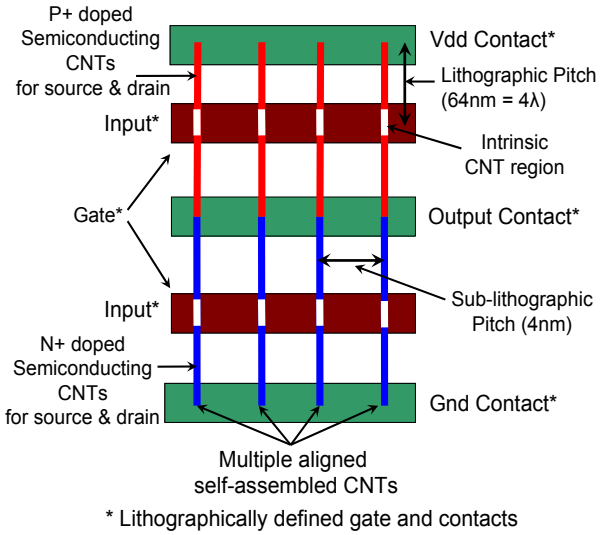
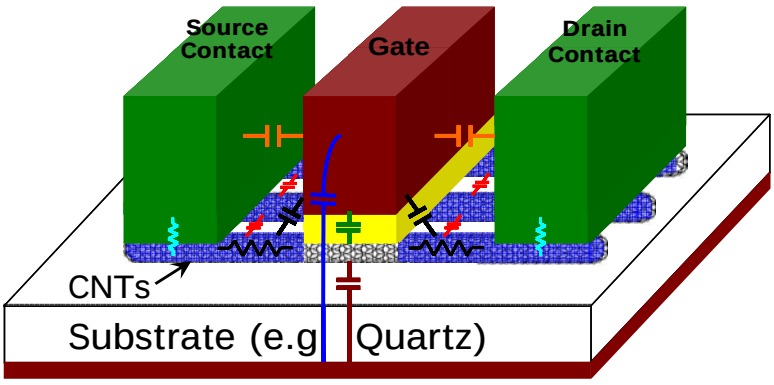



**Figure 7.1:** The top view of the layout of CNFET inverter with multiple CNTs per FET. The Gate/Source/Drain pitch (in the vertical direction) is defined by traditional optical lithography, and the inter-CNT pitch under the gate (in the horizontal direction) is defined by self-assembly sub-lithographic technology [7,11,12].

For circuit simulations in this work, we use a CNFET HSPICE model (presented in chapter 5 and chapter 6) implemented including practical device non-idealities [8,9]. We use 0.9V power supply [10] for both CNFET and CMOS circuits. The realistic circuit layout patterns and the related parasitic resistance and interconnect wiring capacitance are included in the simulations. Figure 7.1. shows the layout of a CNFET inverter with multiple CNTs. Gates and contacts are defined by lithography while the inter-CNT pitch within the same gate is defined by the sublithographic self-assembly process [9,11,12].




| | Description | CNFET | MOSFET |
|---|---|---|---|
| $C_{gc}$ | Intrinsic Gate Cap. | ~ 4 aF / CNT | ~ 35 aF |
| $C_{fr}$ | Outer-fringe Cap. | ~ 2 aF / CNT | ~ 19 aF |
| $C_j$, $C_{jsw}$ | Junction Cap. | — | ~ 14 aF |
| $C_{ovlap}$ | Overlap Cap. | — | ~ 7 aF |
| $C_{gsd}$ | Gate to S/D Contact | ~ 15 aF | ~ 15 aF |
| **$C_{par}$** | Gate Parasitic Cap. | ~ 17 aF | ~ 55 aF |
| **$C_{tot}$** | Total Gate Cap. | ~ 21 aF | ~ 90 aF |
| **$I_{on}$** | On Current | ~ 25 µA / CNT | ~ 65 µA |

**Figure 7.2:** The 3D view of a CNFET. The inset table shows a comparison of the device capacitance and on-current between CNFET and MOSFET at the 32 nm node. The Gate/Source/Drain lengths are 32 nm, the gate width is 48 nm, and the gate height is 64 nm for both CNFET and MOSFET.

Figure 7.2 shows the 3D view of a CNFET device and the related parasitics. The inset Table in Figure 7.2 shows a comparison of the device capacitance (including Miller effect) and the on-current between the CNFET and the bulk MOSFET with 48 nm (3λ) gate width, $L_g$ = 18 nm at the 32 nm node. We simulate a five-stage FO4 inverter chain consisting of CNFETs with no imperfections (CNT diameter = 1.5 nm, Doping level = 1%). The CNFET inverters are minimum sized with 32 nm gate width.

**Table 7.1:** Device parameters and process assumptions for simulations.

| **Variable Parameters** | |
|---|---|
| Source/Drain Doping Level | 0.59 eV - 0.75 eV (0.7% - 1.3%) Uniformly distributed* |
| CNT Diameter | 1.2 nm - 1.8 nm Uniformly distributed* |
| Probability of a CNT to be Metallic | 8% - 32% |
| **Fixed Parameters** | |
| Oxide Thickness ($T_{ox}$) | 4 nm |
| Gate Dielectric (Dielectric Constant:$K_{ox}$) | $HfO_2$ (16) |
| CNT Pitch | 4 nm |
| Power Supply | 0.9V |
| Mean Free Path: Intrinsic CNT | 200 nm |
| Mean Free Path: Doped CNT | 15 nm |
| Gate/Source/Drain Length (CNT) | 32 nm |
| Work Function: contact ($\Phi_M$) | 4.5 eV |
| Work Function: CNT ($\Phi_{CNT}$) | 4.5 eV |
| Interconnect Capacitance | 0.22 fF/μm |

* In the absence of large scale statistical data.

Our first simulation setup consists of inverters with single semiconducting CNTs. The 3rd CNFET inverter stage in the chain is ~ 2× faster and has ~ 6× lower switching energy per cycle compared to a 32 nm Si CMOS inverter using a similar setup (modeled using the BSIM4 predictive model [13,14,15]). Note that the delay advantage is smaller than the intrinsic CV/I delay advantage reported in Chapter 6. This is because, unlike in CMOS circuits, the local interconnect capacitance dominates the gate capacitance in CNFET circuits (Figure 7.2, Table 7.1). Next, we measure the delay through the 3rd stage while increasing the number of CNTs per inverter for all inverters without changing their gate widths. The FO4 delay through the 3rd stage initially improves because of the increased total drive current due to the increased CNT number (Figure 7.3). The delay subsequently worsens due to the reduced drive current per CNT as a result of increased inter-CNT charge screening. A 4 nm minimum CNT pitch (8 tubes per 32 nm of gate width) gives optimal delay (Figure 7.3) with typical FO4 load at the 32 nm node. A heavier interconnect wiring load tends to shift the curve in Figure 7.3 to the right side which means a tighter inter-CNT pitch.

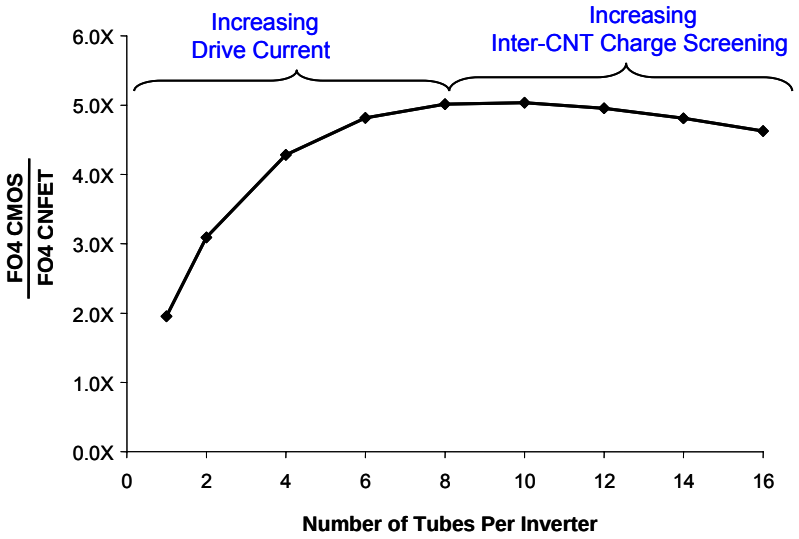



**Figure 7.3:** Mean FO4 delay improvement for CNFET inverter with no imperfections compared to 32 nm Si CMOS. A constant gate width 32 nm (varying CNT pitch) is assumed in this example.

The previous analysis assumed homogenous semiconducting CNTs that have the same diameter and the same doping levels. In reality, there are limitations of CNT fabrication technology over and above process variations introduced by lithography. These include:

1) Doping variations in CNFET source and drain regions that cause drive current variations [16];

2) CNT diameter variations that cause variations in the CNFET drive current [8,17,18];

3) Metallic CNTs that cannot be used to make CNFETs. The current of the metallic CNT can not be controlled by the gate voltage, therefore causes a short between the source and the drain of a CNFET. For a random distribution of CNT chirality, about 1/3 of the CNTs are metallic [17]. Current CNT synthesis techniques yield between 10% to 70% metallic CNTs [19]. One can, in principle, attain the target current drive in a CNFET inverter by starting with a larger number of CNTs (mixture of semiconducting and metallic) and then removing the metallic CNTs (e.g. by electrical burning [20] or chemical/plasma etching [21,22]). In this work, we consider the case where all metallic CNTs can be removed by a perfect removal process. Even in this optimistic scenario, the number of remaining semiconducting CNTs per inverter becomes stochastic variables, causing significant drive current variations.

4) Misaligned CNTs that can cause unintended shorts inside logic structures as reported in [23]. The best CNT synthesis techniques today produce misaligned CNTs like the ones shown in [7].

To quantify the effects of CNFET doping variations, CNT diameter variations, and variation caused by removal of metallic CNTs, we performed the following separate Monte-Carlo simulations. We use the same inverter chain described before assuming perfectly aligned and uniformly spaced tubes with a minimum pitch of 4 nm. The following parameters of test inverter (the 3rd stage in inverter chain) are individually varied (Table 7.1):

1. Source / drain doping levels
2. CNT diameters
3. Probability of a CNT to be metallic.

We compared the mean, 3σ, and 6σ values of the FO4 delay and energy per cycle of this test inverter to those of a 32 nm Si CMOS test inverter in a CMOS inverter chain (without any process variations), as shown by Figure 7.4 and Table 7.2.

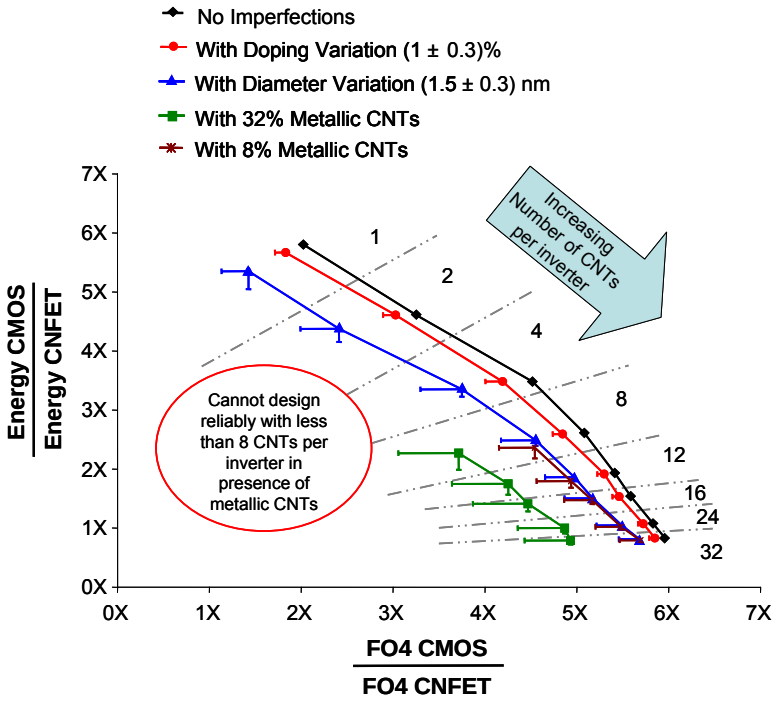



**Figure 7.4:** Energy per cycle and FO4 delay improvement for CNFET inverter at 3σ points (Minimum CNT pitch = 4 nm and Minimum gate width = 32 nm) compared to 32 nm CMOS FO4 inverter. The Error bars indicate 6σ points.

**Table 7.2:** FO4 delay and energy per cycle gain for 8 CNTs per inverter in the presence of imperfections.

| | | Sources of CNFET Imperfections | | | | |
|---|---|---|---|---|---|---|
| | | No Imperfection | Doping Variation (1 ± 0.3)% | Diameter Variation (1.5 ± 0.3) nm | 8% Metallic CNTs | 32% Metallic CNTs |
| Mean FO4 delay (ps) | | 2.95 ps | 3.03 ps | 3.00 ps | 3.00 ps | 3.17 ps |
| Std. Dev. of FO4 delay (ps) [% mean] | | – | 0.02 ps [0.5%] | 0.09 ps [3.3%] | 0.10 ps [3.5%] | 0.29 ps [9.2%] |
| FO4 delay advantage vs. 32 nm Si CMOS | Mean | 5.1× | 5.0× | 5.0× | 5.0× | 4.7× |
| | 3σ point | – | 4.9× | 4.6× | 4.5× | 3.7× |
| | 6σ point | – | 4.7× | 4.2× | 4.2× | 3.0× |
| Energy per cycle advantage vs. 32 nm Si CMOS | Mean | 2.6× | 2.6× | 2.6× | 2.6× | 2.6× |
| | 3σ point | – | 2.6× | 2.5× | 2.4× | 2.3× |
| | 6σ point | – | 2.6× | 2.4× | 2.2× | 2.0× |

The energy-delay-product (EDP) improvement of CNFET inverter without imperfections over 32 nm Si CMOS inverter is about 13×. The doping variation (1% ± 0.3%) of the source/drain region reduces the maximum EDP improvement (with 6σ variation) to 12×. For MOSFET-like CNFET, the doping source/drain CNT region acts as both the electron reservoir and interconnects between the devices. The total resistance of the doped source/drain region is a summation of the Schottky Barrier (SB) resistance ($R_{SB}$) at the metal/CNT interface and the series resistance ($R_{series,sd}$) due to scattering. A moderately high doping in source/drain region suppresses the SB resistance and the metal/CNT contact is essentially ohmic through tunneling. $R_{series.sd}$ only weakly depends on both the doping level and CNT diameter provided that the first two subbands are degenerate. As a result, the source/drain doping variation of CNFETs with high doping (1% ± 0.3%) has the minimal impact on both the circuit speed and dynamic energy consumption. CNT diameter affects not only the source/drain resistance, but also the conductivity of the channel region. Since the bandgap of semiconducting CNTs is reciprocally proportional to the diameter, the threshold voltage of the intrinsic CNT channel can be approximated to the first order as the half bandgap which is an inverse function of the diameter,

$$V_{th} \approx \frac{E_g}{2e} = \frac{\sqrt{3}}{3}\frac{aV_\pi}{eD_{CNT}} \tag{7.1}$$

The parameter $a$ (~ 2.49 Å) is the carbon to carbon atom distance, $V_\pi$ (~ 3.033 eV) is the carbon $\pi$-$\pi$ bond energy in the tight bonding model) [24], $e$ is the unit electron charge, and $D_{CNT}$ is the CNT diameter. Taking the partial derivative of $V_{th}$ over $D_{CNT}$, we obtain,

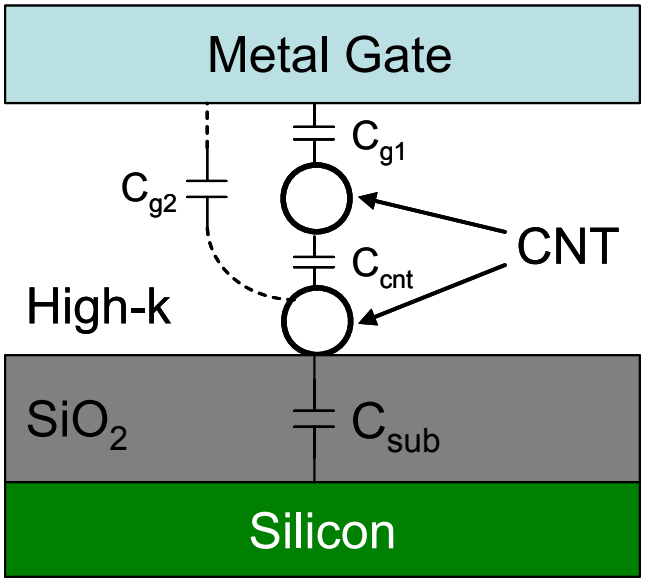
$$\frac{\partial V_{th}}{\partial D_{CNT}} = -\frac{\sqrt{3}}{3}\frac{aV_\pi}{eD_{CNT}^2} \approx -\frac{0.44}{D_{CNT}^2}\,\text{V·nm} \tag{7.2}$$

± 0.3 nm variation around 1.5 nm in diameter results in about −/+ 0.06 V threshold voltage variation which reduces the maximum EDP improvement (with 6σ variation) to 10×.

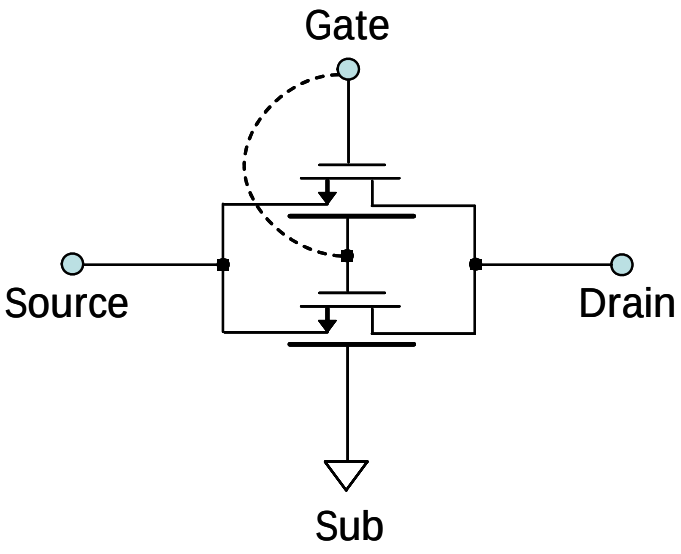
The impact of CNFET doping variation and CNT diameter variation on delay and energy per cycle is not significant compared to the impact of variation caused by metallic CNTs. For devices targeted to have 4 or fewer CNTs, there is a non-negligible probability that all CNTs in the device are metallic and therefore removed, causing an open circuit. Thus the probability of metallic CNTs sets a lower bound on the minimum number of CNTs required per logic gate in order to make viable circuits. With 8% metallic CNT, at least 8 CNTs per device is required in order to reduce the probability of all CNTs of one device to be metallic below $2\times10^{-9}$. The maximum EDP improvement (with 6σ variation) is reduced to 6× with 32% metallic tubes. With 8% metallic tubes the FO4 delay improvement and the maximum EDP improvement at 6σ point is better at 2.2× and 9× respectively which is comparable to that for diameter variation (Table 7.2).

Now we address the problem of misaligned CNTs. Misalignment may occur either within or beyond the gate region [23]. Misaligned CNTs within the gate degrades circuit performance because of inter-CNT pitch variation and the resultant drive current variation. Comparing two logic gates with the same number of CNTs per gate, the worse case of misalignment is that two CNTs overlaps each other so the lower CNT is shielded from the gate by the upper CNT (Figure 7.5(a)). As shown by Figure 7.5(b), the lower CNT is controlled by two front gates: the metal gate and the upper CNT, as well as one back gate. The electrostatic coupling capacitance among the metal gate, the two CNTs, and the substrate are calculated with 3-d field solver and used as the input parameters for SPICE simulations. The difference between the total current driven by the two crossed

tubes and the current driven by two identical parallel tubes is within ± 15%. Thus even in the worst case of misalignment, there is no significant reduction in drive current of the lower CNT, since the lower CNT is coupled to the top gate through fringing electric fields.




(a)




(b)

**Figure 7.5:** (a) Two CNTs overlap each other. This is the worst case of two crossed tubes in the channel region which gives the maximum current variation due to crossing. (b) The equivalent circuit model of two crossed tubes used for simulation. The bottom CNFET is controlled by three gates which include two front gates and one back gate.

The misaligned CNTs beyond the gate region can cause unintended shorts inside logic structures. In [23], we describe a technique to design CNFET-based logic structures that are guaranteed to implement correct logic functions even in the presence of large number of misaligned CNTs. A detailed description on automated design of misaligned-CNT-immune circuits is reported in [9].

## 7.3 Scalability of CNFET Technology

### 7.3.1 *CNFET Device Speed*

The above analyses are evaluated at the 32 nm node. Next, we examine the scalability of the performance advantage of CNFET technology over CMOS technology. The CNFET on-current can be approximated as,

$$I_{CNFET} = n \cdot g_{CNT}(V_{DD} - V_{SS'} - V_{th,CNT}) \tag{7.3}$$

The parameter $n$ is the number of CNTs per device, $V_{th,CNT}$ is the threshold voltage which is about 0.3 V for (19,0) semiconducting CNT, $g_{CNT}$ is the transconductance per CNT, and $V_{SS'}$ is the voltage drop between the inner source node S' and the external source node S (Figure 7.6 inset),

$$V_{SS'} = \frac{I_{CNFET} L_s \rho_s}{n} \tag{7.4}$$

$L_s$ is the source length (doped CNT region), and $\rho_s$ is the source resistance per unit length of doped CNT. The effect of the quantum contact resistance on device on-current is lumped into the transisconductance $g_{CNT}$ as described in Chapter 5 and Chapter 6.

With equation (7.3) and (7.4), we obtain,

$$I_{CNFET} = \frac{n \cdot g_{CNT}(V_{DD} - V_{th,CNT})}{1 + g_{CNT} L_s \rho_s} \tag{7.5}$$

The CNFET gate capacitance ($C_{gg,CNFET}$) consists of three components [25]: the gate to channel capacitance ($C_{gc,tot,CNT}$), the gate outer fringe capacitance ($C_{fr,tot,CNT}$), and the

coupling capacitance between the gate and the adjacent contacts ($C_{gtg,tot}$). These components are approximately,

$$\begin{aligned} C_{gc,tot,CNT} &= n \cdot C_{gc,CNT} L_g \\ C_{fr,tot,CNT} &\approx C_{fr,CNT} L_s \\ C_{gtg,tot} &= C_{gtg} W_g \end{aligned} \tag{7.6}$$

$C_{gc,CNT}$ and $C_{fr,CNT}$ are the capacitance per unit CNT length, and $C_{gtg}$ is the capacitance per unit gate width, including Miller effect. For a typical 1-D device, the outer fringe capacitance is usually quite small (~10% of $C_{gg,CNFET}$) compared to the gate to contact capacitance and gate to channel capacitance [25]. Thus the total gate capacitance is given by,

$$C_{gg,CNFET} \approx n \cdot C_{gc,CNT} L_g + C_{gtg} W_g \tag{7.7}$$

CNFET device delay CV/I can be derived with equations (7.5, 7.7),

$$\begin{aligned} \tau_{CNFET} &= \frac{C_{gg,CNFET} V_{DD}}{I_{CNFET}} \\ &= (1 + \frac{C_{gtg} W_g}{n \cdot C_{gc,CNT} L_g}) \cdot (1 + g_{CNT} L_s \rho_s) \cdot \frac{C_{gc,CNT} L_g V_{DD}}{g_{CNT} (V_{DD} - V_{th,CNT})} \\ &= \eta_{CNT,C} \cdot \eta_{CNT,R} \cdot \frac{C_{gc,CNT} L_g V_{DD}}{g_{CNT} (V_{DD} - V_{th,CNT})} \end{aligned} \tag{7.8}$$

Here, $\frac{C_{gc,CNT} L_g V_{DD}}{g_{CNT} (V_{DD} - V_{th,CNT})}$ can be interpreted as the intrinsic gate delay metric for a CNFET with a single CNT without source/drain series resistance and without considering parasitic capacitances.

We define the pre-factor $\eta_{CNT}$ as,

$$\begin{aligned} \eta_{CNT} &= \eta_{CNT,C} \cdot \eta_{CNT,R} \\ \eta_{CNT,C} &= (1 + \frac{C_{gtg} W_g}{n \cdot C_{gc,CNT} L_g}) \\ \eta_{CNT,R} &= (1 + g_{CNT} L_s \rho_s) \end{aligned} \tag{7.9}$$

Therefore the CNFET device intrinsic speed is degraded by both the pre-factor $\eta_{CNT,C}$ due to the gate parasitic capacitance and the pre-factor $\eta_{CNT,R}$ due to the extension series resistance. In the quantum capacitance limited regime for high performance CNFET device (i.e. $C_{ox} >> C_Q$), $C_{gc,CNT}$ is almost constant, both $C_{gtg}$ and the ratio of $W_g$ over $L_g$ are constant with traditional device scaling scenarios. Thus the pre-factor $\eta_{CNT,C}$ decreases with larger number of CNTs per device. On the other hand, the source series resistance per unit length of heavily doped CNT ($\rho_s$) is almost constant (limited by the maximum doping achievable), and the transconductance per CNT ($g_{CNT}$) is also constant because it weakly depends on the channel length with near-ballistic transport. Thus the pre-factor $\eta_{CNT,R}$ deceases with reduced source/drain length as technology scales down. As an overall effect, the value of the pre-factor $\eta_{CNT}$ for scaled device depends on the tradeoff between the number of CNTs per device and the source series resistance.

### 7.3.2 *MOSFET Device Speed*

We assume the MOSFET device on-current continues to be inversely proportional to the gate length as projected by ITRS [10], to the first order approximation,

$$I_{Si} = \frac{W_g g_{Si}(V_{DD} - V_{th,Si})}{L_g + L_s g_{si} \rho_{si,s}} \tag{7.10}$$

$g_{CNT}$ is the transconductance per square which is independent of the technology node, $\rho_{si,s}$ is the source serier resistance per square, and $V_{th,Si}$ is the threshold voltage of MOSFET.

For the MOSFET device, the gate capacitance ($C_{gg,Si}$) also consists of three major components: the gate to channel capacitance ($C_{gc,tot,Si}$), the gate outer fringe capacitance ($C_{fr,tot,Si}$), and the coupling capacitance between gate and the adjacent contacts ($C_{gtg,tot}$),

$$C_{gg,Si} = (C_{gc,Si} + C_{fr,Si} + C_{gtg}) \cdot W_g \tag{7.11}$$

$C_{gc,si}$, $C_{fr,si}$, and $C_{gtg}$ are the capacitance per unit gate width. The same (compatible) metallization / interconnect technology is assumed for both technologies. The MOSFET junction capacitance is ignored in this analysis which approximates the situation for an optimistic device structure such as the fully depleted SOI and FINFET.

With equations (7.10, 7.11), we obtain the MOSFET device delay CV/I,

$$\begin{aligned}\tau_{Si} &= \frac{C_{gg,Si}V_{DD}}{I_{Si}} \\ &= (1+\frac{C_{fr,Si}+C_{gtg}}{C_{gc,Si}})\cdot(1+\frac{L_s}{L_g}g_{si}\rho_{si,s})\cdot\frac{C_{gc,si}L_gV_{DD}}{g_{si}(V_{DD}-V_{th,Si})} \\ &= \eta_{si}\cdot\frac{C_{gc,si}L_gV_{DD}}{g_{si}(V_{DD}-V_{th,Si})}\end{aligned} \tag{7.12}$$

We define the pre-factor $\eta_{si}$ as,

$$\begin{aligned}\eta_{Si} &= (1+\frac{C_{fr,Si}+C_{gtg}}{C_{gc,Si}})\cdot(1+\frac{L_s}{L_g}g_{si}\rho_{si,s}) \\ &\approx 1+\frac{C_{fr,Si}+C_{gtg}}{C_{gc,Si}}\end{aligned} \tag{7.13}$$

The source/drain series resistance has much less significant effect on device on-current for silicon device than that for CNFET device because the resistivity of the heavily doped source/drain region of silicon devices is much smaller (< 10 %) than the resistivity of the channel region.

Thus the MOSFET device intrinsic speed ($\frac{C_{gc,si}L_gV_{DD}}{g_{si}(V_{DD}-V_{th,Si})}$) is mainly degraded by the gate parasitic capacitance. The source/drain series resistance and contact resistance will also affect the device speed, but their effects are the secondary compared to the effect of the gate parasitic capacitance on speed, provided the on-current can be scaled up with technology node. For the purpose of this work, $C_{gc,si}$, $C_{fr,si}$, and $C_{gtg}$ are assumed to be independent of the technology node with the traditional device scaling scenarios for simplicity[+]. Thus the pre-factor $\eta_{si}$ for MOSFET device is almost constant and technology node independent.

[+] For devices beyond 65 nm node, the gate to channel capacitance $C_{gc,si}$ is hard to be constant due to the slower scaling of physical gate length and oxide thickness, as well as the poly depletion effect [26].

### 7.3.3 *CNFET Scalability*

The scalability of CNFET technology is limited by: (1) the performance of the CNFET device without considering the imperfections and the performance variations; (2) the performance degradation due to the imperfections as discussed in Section 7.2. In order for CNFET to develop into a competitive technology, it should have good scalability, i.e. the performance advantage of CNFET over MOSFET is expected to improve (or at least maintain the same) as the technology node advances. We first discuss the relative speed improvement of CNFET over MOSFET as a function of the technology node, without worrying about the performance degradation due to the imperfections. Assuming $V_{th,CNT} = V_{th,Si}$, with equations (7.8, 7.9, 7.12, 7.13), the device speed improvement of CNFET over MOSFET can be expressed as,

$$\frac{\tau_{Si}}{\tau_{CNFET}} = \frac{\eta_{Si}}{(1+\frac{a}{n})(1+b \cdot L_s)} \cdot \beta_{intr} \tag{7.14a}$$

$$\beta_{intr} = \frac{C_{gc,Si} \cdot g_{CNT}}{C_{gc,CNT} \cdot g_{Si}} \tag{7.14b}$$

$$\eta_{Si} = 1 + \frac{C_{fr,Si} + C_{gtg}}{C_{gc,Si}} \tag{7.14c}$$

$$a = \frac{C_{gtg} W_g}{C_{gc,CNT} L_g} \tag{7.14d}$$

$$b = g_{CNT} \rho_s \tag{7.14e}$$

$\beta_{intr}$ is the intrinsic speed improvement of CNFET over CMOS device, $\eta_{si}$, $a$ and $b$ are technology node independent assuming traditional device scaling scenario. From equation (7.14), we observe that there are multiple ways to increase the device speed improvement of CNFET over MOSFET: 1) Increase the number of CNTs per device in order to increase the total current drive. 2) Use the minimum sized device, i.e. reduce $W_g$ to reduce the gate parasitic capacitance of the CNFET. 3) Reduce the resistivity of the doped CNTs. A potential solution is to replace the doped S/D CNTs with metallic CNTs.

**Table 7.3:** The device parameters of CNFET and MOSFET used for CNFET scalability analysis.

| | | | |
|---|---|---|---|
| $C_{gtg}$ (aF/μm gate width) | 300 | $g_{si}$ (μs /□) $^{+}$ | 75 |
| $C_{fr,si}$ (aF/μm gate width) | 400 | $g_{CNT}$ (μs) | 80 |
| $C_{gc,si}$ (aF/μm gate width) | 700 | $\rho_s$ (kΩ/μm) $^{*}$ | 80 |
| $C_{gc,CNT}$ (aF/μm gate length) | 140 | | |
| $\beta_{intr}$ | 5.3 | $\eta_{Si}$ | 2.0 |
| $a$ ($W_g/L_g$) | 2.1 | $b$ ($\mu m^{-1}$) | 6.4 |

$^{+}$ Assuming the gate length $L_g$ equals the feature size $2\lambda$.
$^{*}$ Assuming the mean free path of doped CNT is 40 nm.

The parameters for typical CNFET and MOSFET devices are summarized in Table 7.3. For high performance CNFET, we assume the first two subbands of doped S/D are degenerated, while only the first subband of the intrinsic channel is populated in at on-state, provided the fact that it is hard to populate the second subband with typical gate to channel capacitance ($C_{QC}$ < 500 aF/μm [25]) and sub-1V power supply. The parameters for CNFET are benchmarked with the compact model reported in Chapter 6. The capacitance values and transconductance for MOSFET are obtained from [26] and ITRS projection [27], respectively. The parameters for silicon MOSFET are different from the parameters predicted by the predictive BSIM model [15] in order to avoid underestimating MOSFET performance. The junction capacitance of the MOSFET is not included in this analysis. The fringing/interconnect capacitance values include Miller effect. The physical gate length is assumed to be 2λ for simplicity. The impurity scattering MFP of degenerately doped CNT with charge transfer doping is assumed to be 40 nm [14]. The intrinsic device speed improvement of CNFET over n-type MOSFET ($\beta_{intr}$) is about 5.3× which is almost independent of the technology node. Considering CMOS gate parasitic capacitance, the speed improvement is increased to ($\eta_{si}\cdot\beta_{intr}$) 10.6×. The speed advantage of CNFET over MOSFET technology is sensitive to the gate parasitic capacitance, the source/drain series resistance, and the number of CNTs per device. To the first order approximation, for the minimum sized device with $W_g$=1.5$L_g$ and $L_s$=$L_g$=32 nm, CNFET with 1, 2, 10 CNTs per device ($\tau_{si}/\tau_{CNFET}$) is about 2.5×, 4.1×, 8.1× faster than MOSFET at the 32 nm node, respectively. The results are promising, but

are with two pre-conditions: (1) the minimum sized device is guaranteed to have the specified number of semiconducting CNTs per device; (2) there is no imperfection. The maximum achievable CNT density with today's directed-CNT-growth synthesis technology is about 4 CNTs/μm [7,28] which requires a minimum gate width of 250 nm in a circuit environment to connect two adjacent devices together. Thus CNFET device speed also depends on CNT density.

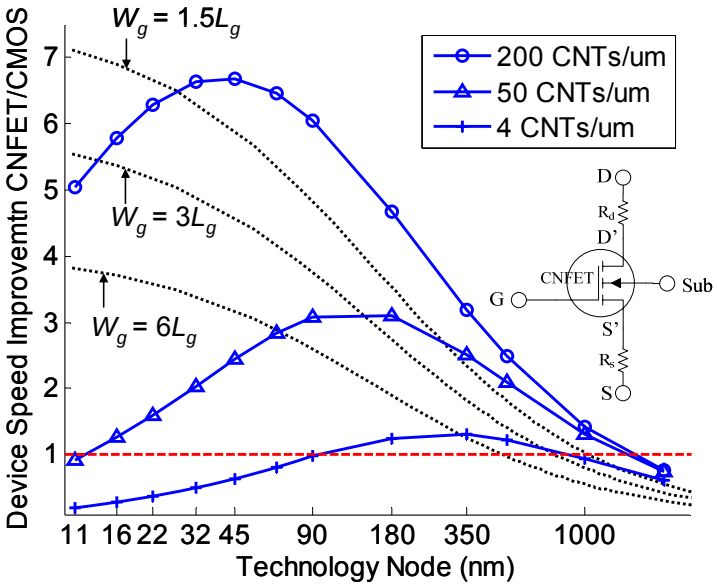



**Figure 7.6:** CNFET device speed advantage over MOSFET vs. technology node, with different CNT density. The dotted three curves denote the technology nodes at which the CNFET with the gate width $W_g$ = $1.5L_g$, $3L_g$, $6L_g$, respectively, can hold 8 CNTs per device. Significant performance degradation is expected for the more advanced technologies beyond these curves (to the left hand side of the curves) due to the smaller number of CNTs per device.

Now we consider the effect of CNT density on CNFET device speed and the scalability. We consider the case where the inter-CNT pitch (or CNT density) is a constant for all the technology nodes for simplicity. The total current drive of CNFET device depends on both the number of CNTs per device and the doped source/drain series resistance. A narrower gate at the more advanced technology has less number of CNTs per device while the reduced source/drain length increases the current per CNT. As shown by Figure 7.6, the speed improvement of CNFET (with the fixed CNT density) over MOSFET increases first, and then decrease as the technology node advances. The increase in CNFET speed advantage is due to the shorter source/drain length and consequently the smaller source/drain series resistance. The benefit from the shorter source/drain length

becomes smaller, i.e, the parameter $b \cdot L_s$ in equation (7.14a) is less significant on device speed, as technology advances. On the other hand, the total current drive per device decreases and the parameter $a/n$ in equation (7.14a) becomes significant as the technology advances because of lesser number of CNTs per device. Assuming a minimum device gate width $W_{g,min}$, an inter-CNT pitch larger than $W_{g,min}$ requires a wider gate width which increases the gate parasitic capacitance and therefore reduces the device speed. As a trade-off among the current drive per device, the number of CNTs per device, and the gate parasitic capacitance, an optimal point with the maximum speed improvement over MOSFET exists for CNFET technology for a given CNT density. For CNFET with CNT density of 4, 50, and 200 CNTs/μm, the optimal technology is around 350 nm, 90 nm, and 45 nm node, respectively. With 4 CNTs/μm, CNFET loses the speed advantage over MOSFET technology beyond the 90 nm node. Higher CNT density helps improve the device speed and achieve the maximum speed improvement at the more advanced technology node. Though the screening of parallel CNTs may cause current reduction and speed degradation for CNFET with high CNT density [25], the CNFET scaling trend described by equation (7.14) and Figure 7.6 is still valid and accurate to the first order approximation.

The above analysis assumes CNFET device without imperfections. As discussed in Section 7.2, various device imperfections degrade the circuit performance. With 8% metallic CNT growth rate and the typical diameter, doping level variation, at least 8 CNTs per device is required to minimum the performance variations (Table 7.2). The dotted three curves in Figure 7.6 indicates the most advanced technology node that can accommodate at least 8 CNTs per device with the minimum gate width ($W_{g,min}$) equals to $1.5L_g$, $3L_g$, and $6L_g$, respectively. Significant performance degradation is expected for the more advanced technologies beyond these curves (to the left hand side of the curves) due to the smaller number of CNTs per device. Larger gate width relaxes the CNT density requirement at the cost of larger device area and higher power consumption. In order to minimum performance variation at the 32 nm node for the minimum sized device with $W_{g,min}=1.5L_g$, at least 200 CNTs/μm (an inter-CNT pitch smaller than 5 nm) is desired (This CNT density is 50× higher compared to today's CNT synthesis technology).

## 7.4 Conclusions

This chapter evaluates CNFET circuit performance, in the presence of process related non-idealities and imperfections in a real circuit environment, at the 32 nm technology node using the compact CNFET SPICE model. This work uses Monte-Carlo simulation to investigate the effect of three major process related imperfections on circuit level performance: (1) Doping variations in the CNFET source and drain regions; (2) carbon nanotube (CNT) diameter variation; and (3) the variation caused by the removal of metallic CNTs.

In conclusion, we have a lot to gain from CNFET-based circuits (2.5× energy per cycle advantage and 4.6× FO4 delay advantage at 3σ points after addressing diameter and doping variation) over 32 nm Si CMOS as long as we can build circuits and architectures immune to misaligned CNTs and metallic CNTs. As described in [9], it is possible to design robust standard cells with guaranteed correct functionality in the presence of misaligned CNTs. We require an interdisciplinary approach to overcome the problem of metallic CNTs. CNT synthesis techniques that can significantly and reliably reduce the probability of metallic CNTs (e.g., to fewer than 8% metallic CNTs) are extremely important. We need to combine novel CNT synthesis techniques, reliable ways of removing metallic CNTs after CNT growth, together with new circuit and architectural techniques to fully overcome the challenge of metallic CNTs. This work describes a methodology to correlate circuit-level performance with CNT material and fabrication specifications required to achieve the required circuit performance. An analogical approach can be for other novel device such as semiconductor nanowire transistors, MEMS switches, and spintronics.

We also showed that CNT density affect CNFET device speed significantly. CNT density (4 CNTs/μm) provided by current CNT synthesis technology is not sufficient for CNFET to achieve competitive performance over MOSFET beyond the 90 nm node. CNFET technology has good scalability if CNT density can be significantly improved. To maximize circuit performance at the 32 nm node, a minimum CNT density of 200 CNTs per μm gate width is desired.

## 7.5 Bibliography

[1] H.-S. P. Wong, J. Appenzeller, V. Derycke, R. Martel, S. Wind, Ph. Avouris, "Carbon Nanotube Field Effect Transistors – Fabrication, Device Physics, and Circuit Implications", *Proc. Intl. Solid State Circuits Conference*, pp. 370 – 371, 2003.

[2] J. Guo, A. Javey, H. Dai, and M. Lundstrom, "Performance Analysis and Design Optimization of Near Ballistic Carbon Nanotube Field-Effect Transistors," *Proc. Intl. Electron Devices Meeting,* pp. 703 – 706, December, 2004.

[3] A. Javey, J. Guo, D. B. Farmer, Q. Wang, D. Wang, R. G. Gordon, M. Lundstrom, and H. Dai, "Carbon Nanotube Field-Effect Transistors with Integrated Ohmic Contacts and High-K Gate Dielectrics," *Nano Letters*, vol. 4, pp. 447-450, 2004.

[4] A. Javey, J. Guo, D. B. Farmer, Q. Wang, E. Yenilmez, R. G. Gordon, M. Lundstrom, H. Dai, "Self-Aligned Ballistic Molecular Transistors and Electrically Parallel Nanotube Arrays," *Nano Letters*, vol. 4, pp. 1319-1322 (2004).

[5] A. Javey, R. Tu, D.B. Farmer, J. Guo, R.G. Gordon, H. Dai, "High Performance n-Type Carbon Nanotube Field-Effect Transistors with Chemically Doped Contacts," *Nano Letters*, vol. 5, pp. 345-348, 2005.

[6] J. Chen, C. Klinke, A. Afzali, and P. Avouris, "Self-aligned carbon nanotube transistors with charge transfer doping," *Applied Physics Letters*, vol. 86, pp. 123108, 2005.

[7] S. Han, X. Liu, C. Zhou, "Template-Free Directional Growth of Single-Walled Carbon Nanotubes on a- and r-Plane Sapphire," *J. Am. Chem. Soc.*, vol. 127, pp. 5294 – 5295, 2005.

[8] J. Deng, H.-S. P. Wong, "A Circuit-Compatible SPICE model for Enhancement Mode Carbon Nanotube Field Effect Transistors," *Proc. Intl. Conf. Simulation of Semiconductor Processes and Devices,* pp. 166 – 169, September, 2006.

[9] N. Patil, J. Deng, H.-S. P. Wong, and S. Mitra, "Automated Design of Misaligned-Carbon-Nanotube-Immune Circuits," *44th Design Automation Conference (DAC)*, San Diego, CA, June 4-8, 2007.

[10] Data from http://www.itrs.net/Links/2005ITRS/Home2005.htm

[11] C.T. Black, O. Bezencenet, "Nanometer-scale pattern registration and alignment by directed diblock copolymer self-assembly," *IEEE Trans. Naonotechnology*, pp. 412-415, 2004.

[12] L.-W. Chang, H.-S. P. Wong, "Diblock Copolymer Directed Self-Assembly for CMOS Device Fabrication," *SPIE 31st International Symposium on Microlithography*, paper 6156-66, Feb 19 – 24, 2006.

[13] W. Zhao, Y. Cao, "New generation of Predictive Technology Model for sub-45nm design exploration," IEEE International Symposium on Quality Electronic Design (ISQED), pp. 585-590, 2006.

[14] Y. Cao, T. Sato, D. Sylvester, M. Orshansky, C. Hu, "New paradigm of predictive MOSFET and interconnect modeling for early circuit design," IEEE Custom Integrated Circuits Conference (CICC), pp. 201-204, 2000.

[15] Available at http://www.eas.asu.edu/~ptm/latest.html

[16] D. Kang, N. Park, J,H. Ko, E. Bae, and W. Park, "Oxygen-Induced P-type Doping of A Long Individual Single-walled Carbon Nanotube," *Nanotechnology*, vol. 16, pp. 1048 – 1052, 2005.

[17] Y-C Tseng, K. Phoa, D. Carlton, and J. Bokor, "Effect of Diameter Variation in a Large Set of Carbon Nanotube Transistors," *Nano Letters,* vol. 6, pp. 1364 – 1368, 2006.

[18] R. Saito, G. Dresselhaus and M. Dresselhaus, "Physical Properties of Carbon Nanotubes," *Imperial College Press*, 1998.

[19] 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, H. Dai, "Preferential Growth of Semiconducting Single-Walled Carbon Nanotubes by a Plasma Enhanced CVD Method," *Nano Letters*, vol. 4, pp. 317-321, 2004.

[20] P. G. Collins, M. S. Arnold, P. Avouris, "Engineering carbon nanotubes and nanotube circuits using electrical breakdown," *Science*, vol. 292, pp. 706 – 709, 2001.

[21] A. Hassanien, M. Tokumoto, P. Umek, D. Vrbanic, M. Mozetic, D. Mihailovic, P. Venturini, S. Pejovnik, "Selective etching of metallic single-wall carbon nanotubes with hydrogen plasma," *Nanotechnology*, vol. 16, pp. 278 – 281, 2005.

[22] G. Zhang, P. Qi, X. Wang, Y. Lu, X. Li, R. Tu, S. Bangsaruntip, D. Mann, L. Zhang, H. Dai**,** "Selective Etching of Metallic Carbon Nanotubes by Gas-Phase Reaction," *Science*, Vol. 314, pp. 974 – 977, 2006.

[23] J. Deng, N. Patil, K. Ryu, A. Badmaev, C. Zhou, S. Mitra, H.-S. P. Wong, "Carbon Nanotube Transistor Circuits: Circuit-Level Performance Benchmarking and Design Options for Living with Imperfections," *International Solid State Circuits Conference (ISSCC)*, pp. 70-71, San Francisco, CA, February, 2007.

[24] J. W. Mintmire and C. T. White, "Universal Density of States for Carbon Nanotubes," *Physical Review Letters*, vol. 81, pp. 2506-2509, 1998.

[25] J. Deng and H.-S P. Wong, "Modeling and Analysis of Planar Gate Electrostatic Capacitance for 1-D FET with Multiple Cylindrical Conducting Channels," *IEEE Transactions on Electron Devices*, to appear, 2007.

[26] J. Deng, L. Wei, L.-W. Chang, K. Kim, C.-T. Chuang, H.-S P. Wong, "Extending the Technology Roadmap by Selective Device Footprint Scaling and Parasitics Engineering," Submitted to *2007 ICCAD*, San Jose, CA, 2007.

[27] Data from http://www.itrs.net/Links/2005ITRS/Home2005.htm

[28] 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*, vol. 2, pp. 230-236, 2007.

# Chapter 8

# CONCLUSIONS

This thesis describes a body of work on modeling, understanding, and performance evaluation and prediction for nanoscale devices and circuits, including both CMOS technology beyond the 45 nm node and carbon nanotube field effect transistors (CNFETs), with the aim of guiding nanoscale device and circuit design. This thesis covers: device performance metric definition, extending the silicon CMOS technology roadmap by selective footprint scaling, 1-D FET gate capacitance modeling, CNFET device modeling, CNFET device / circuit / system performance prediction, performance comparison between CNFET and CMOS technology, and CNFET technology scalability analysis. The core part of this thesis consists of the manuscripts published in or submitted to journals or conference proceedings based on work performed at Stanford.

## 8.1 Summary of this Thesis

The I-V behavior of nanoscale silicon devices and the novel devices (e.g. CNFET and Schottky barrier metal source/drain FETs) are quite different from that of well-tempered silicon devices. The non-idealities of nanoscale devices become more significant than before. It is very important to develop a new metric, other than the conventional metric $I_{Dsat}$, for comparing conventional Si FETs with novel devices. We proposed a simple and accurate 4-point model for inverter effective drive current for nanoscale Si MOSFET and novel devices performance benchmarking (Chapter 2). This model accounts for the following observations: (1) $I_{Dsat}$ is never reached in a switching event, (2) pFET data are also required (in addition to nFET data) to estimate the inverter speed, (3) different switching current trajectories for recent and future device generations. This model accurately captures the inverter delay performance over many CMOS technology nodes and in the presence of device non-idealities. About 20% slower inverter speed is

predicted by using this model, as compared to the performance prediction using the metric $I_{Dsat}$.

Since 1960's, CMOS device have been scaled down for close to half century. 65 nm technology becomes the mainstream since 2006. It is evident that it becomes more difficult to further push device / circuit performance by reducing device physical gate length beyond 65 nm node. For nanoscale devices, the parasitics and non-idealities are more significant. In Chapter 3, we propose improving the device and circuit performance by device footprint selective scaling and parasitic engineering, without reducing the physical gate length. This principle, device footprint selective scaling, is verified and supported by device and circuit simulations using both 3-D device (bulk CMOS) and 2-D device (UTB FD-SOI) as examples. With selective footprint scaling along with reduced gate and plug height, the historic performance improvement trend can continue for another 2 to 3 generations and the CMOS technology roadmap can be extended to the 11 nm node with physical gate length no shorter than 10 nm.

Recognizing that the device structure has beening scaled from 3-D (bulk CMOS), quasi 2-D (partially depleted SOI), 2-D (fully depleted SOI), quasi 1-D (nanowire FET, FINFET, tri-gate FET), to 1-D (CNFET) for better channel electrostatics, it is important and necessary to model the behavior of 1-D device with the aim of guiding 1-D device and circuit design. Gate to channel capacitance ($C_{gc}$) is important for determining both the dc and ac characteristics of devices. Chapter 4 presents accurate analytical models to calculate the electrostatic gate capacitance of 1-D FETs with high-$k$ gate dielectric and multiple cylinder conducting channels. This model is suitable for both 1-D devices and quasi 1-D devices with error within 10% of the values simulated by 3-D numerical field solvers. Analysis of device speed using this model shows the importance of accurately modeling the parasitic capacitance for performance evaluation of these 1-D (quasi 1-D) FETs.

While significant advances on CNT synthesis and CNFET device/circuit have been made in the past decade, CNFETs is still premature for macro level circuits design and commercial use. In order for CNFET to develop into a technology, a device model that enables circuit design and performance benchmarking is necessary. A universal circuit-

compatible CNFET device model is derived and implemented in Chapter 5 and Chapter 6, with the aim of evaluating and explaining device behavior and related physical phenomena, as well as obtaining the predictive performance metrics for guiding device and circuit design. This model achieves reasonable accuracy (less than 10% mismatch compared with experimental data) as shown in Appendix B. Using this model, CNFET device and circuit performance are predicted and compared with the cutting-edge CMOS technology at the 32 nm node. It is found that the large speed improvement (6× for nFET and 14× for pFET) of CNFET over CMOS technology at the device level including device parasitics is significantly degraded (by a factor of 5 to 8) by interconnect capacitance in a real circuit environment.

At present, carbon nanotube material synthesis entail process induced variations and imperfections for CNTs and CNFETs. This study, presented in Chapter 7, uses Monte-Carlo simulations to investigate the effect of three major process related imperfections on circuit level performance: (1) Doping variations in the CNFET source and drain regions; (2) CNT diameter variation; and (3) the variation caused by the removal of metallic CNTs. The simulation results indicate that the metallic CNTs are most detrimental. Less than 8% metallic CNT growth rate is desired in order to reduce the circuit performance variation. This work also presents an analytical model for the scalability analysis for CNFET technology. A minimum CNT density of 200 CNTs (about 50× higher than today's CNT density) per μm gate width is required to minimize performance variation at the 32 nm node.

This thesis emphasizes that the device parasitics and non-idealities, in addition to the device/material intrinsic performance, are important for assessing the circuit level performance. One perfect device by itself is not enough to make a successful technology. More attention and effort should be placed on reducing the parasitics, process related variations, and imperfections, in addition to maximizing the device/material intrinsic performance.

## 8.2 Recommendations for Future Work

### 8.2.1 *Device Footprint Scaling*

We make the bold assumption, in Chapter 3, that the sizes of various device structures (e.g. contact sizes, overlay spacings) can be arbitrarily reduced using yet-to-be developed process technologies. The extended scaling path requires tight pitch patterning, tight overlay tolerances, and a short gate height processes. All these are potential yield limiters. More studies on the effect of these assumptions on yield should be carried out to verify the concept of device footprint scaling from the manufacturability point of view. Novel nanofabrication techniques, such as self-assembly and low-*k* spacer, are needed to realize the substantial benefits offered by $L_{pitch}$ scaling and parasitics engineering. Additionally, packaging and/or architectural solutions are also required to mitigate the increased power density due to the smaller device footprint. Future work on device/circuit performance optimization may further consider the effects of stress-dependent carrier mobility, low-*k* isolation dielectric (low-*k* STI), simultaneous optimization of footprint, and device width.

### 8.2.2 *CNFET Modeling*

As a compact model, good agreement with the fabricated device on both dc and ac characteristics have been demonstrated, as shown in Appendix B. Further improvements to this implemented model, in order to improve the accuracy and reduce the run time, may include:

(1) This model utilizes a simplified band structure which restricts the use of this model for the applications that requires a high power supply and high CNT surface potential (>>1.0eV). Using a more accurate band structure model (e.g. tight binding model) can alleviate this issue at the cost of longer run time.

(2) The implemented model uses a circuit self-consistent construct to calculate the channel surface potential recursively. One efficient way to improve the run time (by about one order) is to derive a closed form expression for the channel surface

potential. One way to accomplish this is to separate the operation region into multiple sections and deriving approximated analytical equations in each section.

(3) For a better sub-threshold behavior modeling, the surface potential lowering and consequent higher current caused by the holes (electrons) pile up in the nFET (pFET) channel region should be considered, especially in the high bias region ($V_{ch,DS} > E_{1,0}$). This effect is similar to floating body effect for PDSOI MOSFET and depends on the drain junction doping profile. A gradual doping profile can alleviate this effect by relaxing the potential drop over a longer distance and reducing band-to-band tunneling (BTBT) across the drain junction.

(4) The diffusion capacitances due to the minority carriers at the source/drain junctions are ignored. Though this is a higher order effect, the diffusion capacitance may affect the ac response of high frequency small signal analog circuits. More accurate device modeling for analog circuit applications should be done and verified with experiment data.

(5) To predict and evaluate the ultimate performance of all-CNT circuits, the modeling of CNT interconnects, in additional to the CNFET device model, is important. The performance of CNFET circuits is severely degraded by the traditional interconnect technology. The resistance of heavily doped CNT S/D extension also degrades the on-current. One way to further improve CNFET circuit performance is to use metallic CNTs, multi-walled CNTs, or intrinsic large diameter CNTs as local interconnects because of the much higher current density and much smaller parasitic fringe capacitance. Thus a simple and universal interconnect model is necessary to evaluate all-CNT CNFET circuit performance. This interconnect model should be able to handle both metallic nanotubes and semiconducting nanotubes, as well as single-walled CNT and multiple-walled CNT.

(6) The reported CNT bundle interconnect models treated all CNTs in a bundle identically. In reality, not all CNTs in a bundle are identical. The different electrostatic coupling capacitance between the CNT, depending on the position of

the CNT in a bundle, and the external electrodes, as well as the different resistance which depends on CNT parameters, makes the signal delay different for the CNTs in a bundle. This may result in a skewed signal at the output which is similar to the dispersion effect. The skewed signal is not likely to be recovered due to the randomness of CNTs in a bundle. This effect may affect the validity of using CNT bundle interconnects for RF analogy applications, and must been addressed.

(7) Because of the pre-assumed frequency range in Chapter 6 which is mostly limited by the parasitic resistance and parasitic capacitance, the kinetic inductance and the magnetic inductance are ignored in the model. The high-frequency models employed in this thesis are quasi-static in nature. These issues must be addressed if CNFET are expected to operate at THz frequency range.

(8) The effect of CNT defect on device performance and device reliability analysis as a function of bias and operation time should be addressed. Most of the carrier scattering and thermal relaxation processes occur around the contact/junction region due to the near-ballistics transport, thus defects are likely to accumulate along the nanotubes especially around the contact region for short gate CNFET. Circuit performance may suffer from this gradually degraded current drive.

(9) The number of subband and the number of sub-states in the current implemented model are constant which do not depend on the device parameter, the bias, and operating conditions. More subbands and sub-states used in simulations give a more accurate result at the cost of longer run time. To satisfy the run time and accuracy requirements, the next implementation of this model may set up the number of subband and sub-states dynamically with respect to the device parameters (e.g. CNT diameter, gate to channel capacitance) and operating conditions.

(10) More experiments must be taken to extract the accurate and/or high frequency device parameters, e.g. kinetic inductance, the temperature and bias dependence

of CNT band structure. The possible $sp^3$ hybridization for CNTs with small diameter (< 1nm), and its effect on the band structure also deserve investigation.

(11) A method of parameter extraction to determine the CNFET model parameters from experimental measurements must be developed.

After nearly half a century of evolution since 1960's, the 45 nm CMOS technology has been announced in 2007. Silicon technology is believed to be able to solve its impending scalability problems and extend the roadmap beyond 22 nm node, based on the efforts of the semiconductor industry. As a new technology first conceived in 1998, CNT-based devices have demonstrated attractive material and device level performance compared to conventional CMOS technology in a short time. Though today's CNT-based device and circuit performances cannot compete with the well-developed CMOS technology, CNT still remains an attractive material for nanoelectronics devices, interconnects, owing to the near-ballistic transport, high current capability, nanometer scale, and low-cost synthesis technique. It should not be surprising that CNFET technology and CMOS technology will cross and/or join together at some point in the future.

# Appendix A.

# A BRIEF DESCRIPTION OF CNFET MODEL

## A.1 Capability of this Model

The current version is implemented with HSPICE (ver. 2004.09) Marco model. This model is designed for unipolar behavior CMOS-like CNFET device. The minimum channel length (~10 nm) is restricted by the complex quantum mechanisms which are not implemented in this model. In principle, this model has no limitation on the maximum gate length of CNFET. For gate length longer than 100 nm, the device is treated as long-channel device. The transition from the short channel model (10 nm < Lg < 100 nm) to the long channel model (Lg > 100 nm) is continuous and is automatically handled by the model.

Schottky Barrier (SB) effects can be modeled and observed with this model. The simple SB model employed in this model requires that the doping level in doped Source/Drain extension region be degenerate, with the Fermi level above the first conduction band of carbon nanotube.

## A.2 Function Calling and Local Parameters

The usage of this model is similar to that of the Si CMOS model. The local parameters can be set up for every single device. The syntax to call device models is as below:

**XDevice** ***Drain Gate Source Sub*** **NFET** < Lch=L_channel Lgeff=Lceff Lss=L_sd Ldd=L_sd Efi=Efo Kgate=Kox Tox=4e-9 Csub=20e-12 Csd='Ccsd' CcdBeta='CoupleRatio' Vfbn=0 Dout=0 Sout=0 Pitch=20e-9 Wgate=sub_pitch {*CNPOS=1*} n1=19 n2=0 Mul=1 >

**XDevice** ***Drain Gate Source Sub*** **PFET** < Lch=L_channel Lgeff=Lceff Lss=L_sd Ldd=L_sd Efi=Efo Kgate=Kox Tox=4e-9 Csub=20e-12 Csd='Ccsd' CcdBeta='CoupleRatio' Vfbp=0 Dout=0 Sout=0 Pitch=20e-9 Wgate=sub_pitch {*CNPOS=1*} n1=19 n2=0 Mul=1

The ports definitions (*Drain, Gate, Source, Sub*) for CNFET are the same as that for CMOS device. The ports "*Drain*" and "*Source*" are not swappable in the model though the two ports are physically symmetric in a real device. Since CNFET sits on an insulator, the port "*Sub*" can also act as back gate in some applications such as double gate CNFET. But the driving gate (switching gate) should be the "*Gate*" port because of the simplifications made in substrate region in order to improve the run time.

The parameters between symbol < > are **optional**. The definitions and default values of these parameters are summarized in Table A.1.

**Table A.1.** The definition and default values of local parameters

| Local Parameter | Description | Default Value |
|---|---|---|
| Lch | Physical channel length. This model may not be valid for channel length below 10 nm where other quantum mechanical effects may need to be considered. | 32 nm (Set by global parameter *L_channel*) |
| Lgeff | The mean free path in the (intrinsic) channel region due to non-ideal elastic scattering. | 200 nm (Set by global parameter *Lceff*) |
| Lss | The length of doped source side extension region (CNT) | 32 nm (Set by global parameter *L_sd*) |
| Ldd | The length of doped drain side extension region (CNT) | 32 nm (Set by global parameter *L_sd*) |
| Efi | The Fermi level of the doped S/D tube | 0.6 eV (Set by global parameter *Efo*) |
| Kgate | The dielectric constant of high-*k* front gate dielectric material (planer gate) | 16 (Set by global parameter *Kox*) |
| Tox | The thickness of high-*k* front gate dielectric material (planer gate) | 4 nm |
| Csub | The coupling capacitance between channel region and substrate (back gate) | 20 pF/m (assume 10 μm thick $SiO_2$) |

| | | |
|---|---|---|
| Csd | The coupling capacitance between channel region and source/drain region | 0 pF/m (Set by global parameter *Ccsd*) |
| CcdBeta | The percentage of coupling capacitance between channel and drain region out of Ccsd | 0 pF/m (Set by global parameter *CoupleRatio*) |
| Vfbn, Vfbp | Flat band voltage for nFET, pFET, respectively | 0 eV |
| Dout | Describe the property of the drain side output:<br>1: the drain output is connected to another CNFET directly<br>0: the drain output is connected to metal contact | 0 |
| Sout | Describe the property of the source side output:<br>1: the source output is connected to another CNFET directly<br>0: the source output is connected to metal contact | 0 |
| Pitch | The distance between the center of two adjacent tubes under the same gate. This parameter is used to include the screening effects. It is also useful to change the Gate-Tube coupling capacitance in case *Tox* and *Kgate* are fixed. | 20 nm |
| Wgate | The width of metal gate. This parameter is used to include interconnect capacitance, approximated as 0.213 fF/μm * | 6.4 nm (set by global parameter *sub_pitch*) |
| CNPOS | The position of CNT under the gate (unique for model file *'CNFET_nonBalistic_single_CNT.lib'*:<br>1: the tube is at the two ends<br>0: the tube is in the middle | 1 |
| (n1, n2) | The chirality of tube. n1 and n2 are symmetric. | (19, 0) |
| Mul | The number of tubes under the same gate. | 1 |

* Assuming the height of G/S/D metal contacts is 64nm, contact spacing is 32nm

## A.3 Global Parameters

In addition to the local parameters which can be set for every single device, there are some global parameters in “PARAMETER.lib” file that can affect all the devices. The definition and default values of the major global parameters are summarized in Table A.2.

**Table A.2.** The definition and default values of major global parameters

| Global Parameters | Description | Default Value |
|---|---|---|
| Klowk | The dielectric constant of low-*k* oxide material | 2 |
| Ksub | The dielectric constant of back gate (substrate) dielectric material | 4 |
| Kox | The dielectric constant of high-*k* gate oxide material | 16 |
| Ld_par | Fitting parameter. The length of the drain CNT, to calculate parasitic diffusion capacitance at drain side junction. | 15 nm |
| Efo | The Fermi level of n+/p+ doped source/drain tube. This parameter is internally limited to be above the first conduction band. | 0.6 eV (~0.8% doping level) |
| Lambda_op | The Optical Phonon backscattering mean-free-path in Matallic CNT | 15 nm |
| Lambda_ap | The Acoustic Phonon backscattering mean-free-path in Matallic CNT | 500 nm |
| Photon | The optical phonon energy | 0.16 eV |
| L_channel | Physical gate length | 32 nm |
| L_sd | The length of doped source/drain extension tube. | 32 nm |
| L_relax | Fitting parameter. Carrier relaxation range at drain side, used to match BTBT current. | 40 nm |
| Sub_pitch | Sublithography full pitch | 6.4 nm |
| Ccsd | The coupling capacitance between channel region and source/drain region. | 0 pF/m |
| CoupleRatio | The percentage of coupling capacitance between channel and drain region out of Ccsd. | 0 |
| Lceff | The mean free path in intrinsic CNT. | 200 nm |
| Leff | The mean free path in p+/n+ doped CNT. | 15 nm |
| fai_M | The work function of Source/Drain metal contact | 4.6 eV |
| fai_S | CNT work function | 4.5 eV |

# Appendix B.

## MATCHING EXPERIMENTAL DATA



## B.1 Device Structure

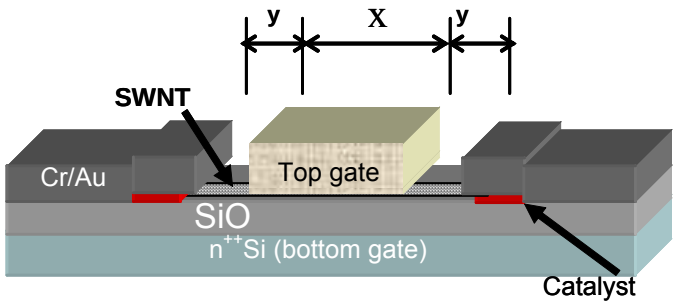



**Figure B.1:** 3-dimensional cross-sectional layout of CNFET; x = 0.75 μm, y = 0.375 μm.

Top-gated CNFETs were fabricated on an oxidized Si substrate where highly doped Si acted as a back gate. The device structure was laid out in ground-signal-ground configuration as shown in Figure B.1. This device is operation in depletion mode. A fixed negative bias is applied to the backgate to electrostatically dope the un-gated regions of the nanotube on either side of the channel. The maximum transconductance ($g_m$) is ~ 10 μS and the observed high saturation current approaches 20 μA. The CNFET was configured as a common-source amplifier and dc offsets were applied to the gate and drain to bias it in saturation near maximum $g_m$.

## B.2 Modeling Results

The CNFET model described in Chapter 5 and Chapter 6 was used to model both dc and ac characteristics of this fabricated CNFET. There are essentially two ways to model the

SWNT-FET as shown in Fig. B.2. One way is to model the top gated section of the nanotube as the FET and the un-gated sections on either side of the top gate as resistors (Figure B.2(a)). In this approach, the effect of electrostatic doping achieved by the bias applied to the back gate is modeled by changing the doping concentration of the un-gated sections. The limitation of this approach, however, is that the source and drain resistances do not depend on the back gate bias. The second and more accurate approach is to model the device as three separate FETs connected in series (Figure B.2(b)) with the nanotube at the two ends with only the back gate control and the middle section with both the top and the back gate control. Despite being computationally more intensive, the second approach, 3-FET model is used to fit the measured data due to its better accuracy.

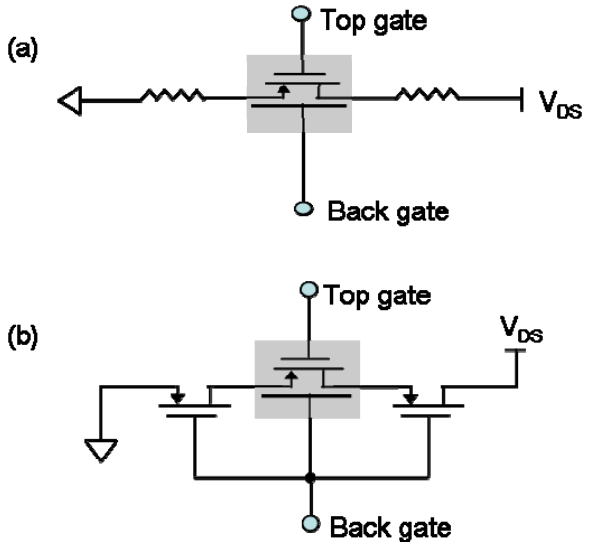



**Figure B.2:** Two approaches to model the CNFET (a) a FET with source and drain series resistors and (b) three FETs in series.

The parameters used for dc and ac modeling are listed in Table B.1 and Table B.2, respectively.

**Table B.1:** The parameters used in dc modeling.

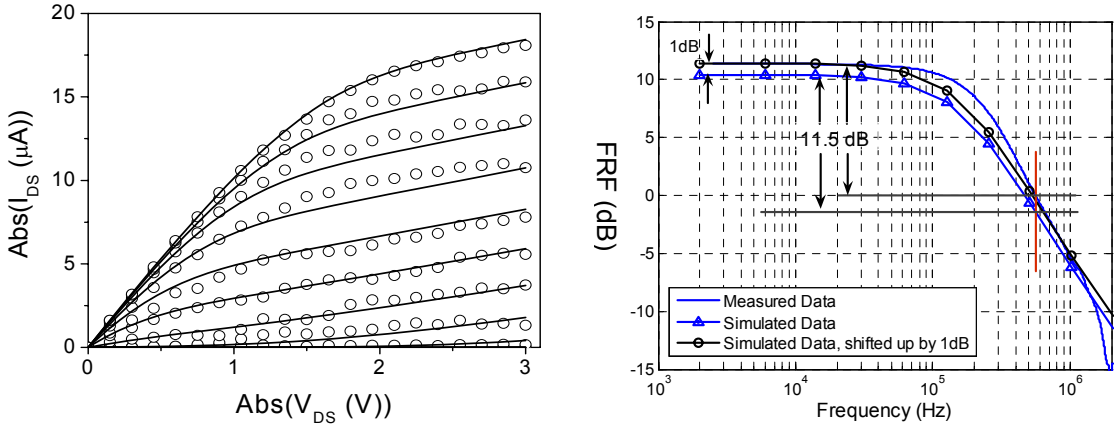
| | | | |
|---|---|---|---|
| Tox ($Al_2O_3$) | 10nm | CNT diameter | 2nm |
| Kox ($Al_2O_3$) | 9 | Channel length | 700nm |
| Tsub ($SiO_2$) | 130nm | Source/Drain CNT length | 400nm |
| MFP* (channel) | 100nm | S/D work function ($\Phi_s$) | 4.6eV |
| MFP* (S/D CNT) | 100nm | CNT work function ($\Phi_{cnt}$) | 4.5eV |
| $V_{sub}$ | -5V | Gate capacitance ($C_{ox}$) | 70aF/μm |
| β | 0.13 | Parasitic capacitance ($C_c$) | 98aF/μm |

*: Scattering Mean Free Path

**Table B.2:** The parameters used in ac modeling.

| | | | |
|---|---|---|---|
| $V_{DS}$ | -10V | $R_D$ | 620K |
| $I_{dd}$ (dc) | 11.75μA | $R_{probe}$ | 1M |
| $I_{ss}$ (dc) | 9.027μA | $C_L$ | 2.8pF |
| $V_{in}$ (dc) | -0.3V | $V_{out}$ (dc) | -2.718V |

The solid line in Figure B.3(a) shows a fit to the dc transport characteristics using the physical parameters shown in Table B.1. In order to get the best fit, a coupling capacitance between the channel region and the drain is fitted as 12.7 aF/μm. Due to the presence of the back gate and other parasitic couplings, the effective top gate to the channel coupling capacitance is only about 50% of the theoretical value with no back-gate. Figure B.3(b) shows a comparison of the measured and simulated voltage gain using the parameters listed in Table B.2. The agreement between the simulations and measurement are quite reasonable with less than 10% error for both dc and ac characteristics.




(a) (b)

**Figure B.3:** (a) Measured (symbols) and simulated (lines) dc transport characteristics of the CNFET as a function of top gate. Both horizontal and vertical scales show absolute values of the drain current and the drain voltage, respectively. A fixed back gate bias of -5 V is applied during the measurement. $V_G$ varies between 1.7 V to -1.3 V in increments of 300 mV. (b) Solid line shows the measured frequency response function (FRF) of the CNFET. A gain of ~11.5 dB with a unity voltage gain frequency of 560 kHz is observed. The device bandwidth is constrained due to the parasitic RC roll-off. The load capacitance $C_L$ can be reduced significantly by fabricating CNFET on a low-loss substrate such as quartz and using a high-Z probe with lower capacitance in fF range. Solid line with "Δ" shows the simulation result. As can be seen, there is a 1 dB difference (~10%) between the measurement and simulation. Better match can be achieved by shifting the simulation data up by 1 dB as shown by the solid line with "o".

The reasonably good agreement of modeling with both the dc and ac experimental data is obtained with the same set of physical parameters. This gives us confidence in projecting the CNFET ac performance under more ideal conditions, as shown by Figure B.4. If the load capacitance ($C_L$) is zero and the probe resistance is infinite, then the intrinsic low frequency gain of the circuit is about 15 dB and the voltage unity-gain frequency is up to 29 GHz (Figure B.4). The device transconductance is degraded by the back gate, which also controls the CNFET channel potential in addition to the top-gate. If back gate coupling and the parasitic coupling are removed, the low frequency gain will be improved to 23 dB, and the voltage unity-gain frequency will increase to 220 GHz. The predicted cut-off frequency for this device based on extracted parameters is ~ 50 GHz. Further improvement in ac performance requires a smaller channel length and a doped S/D.

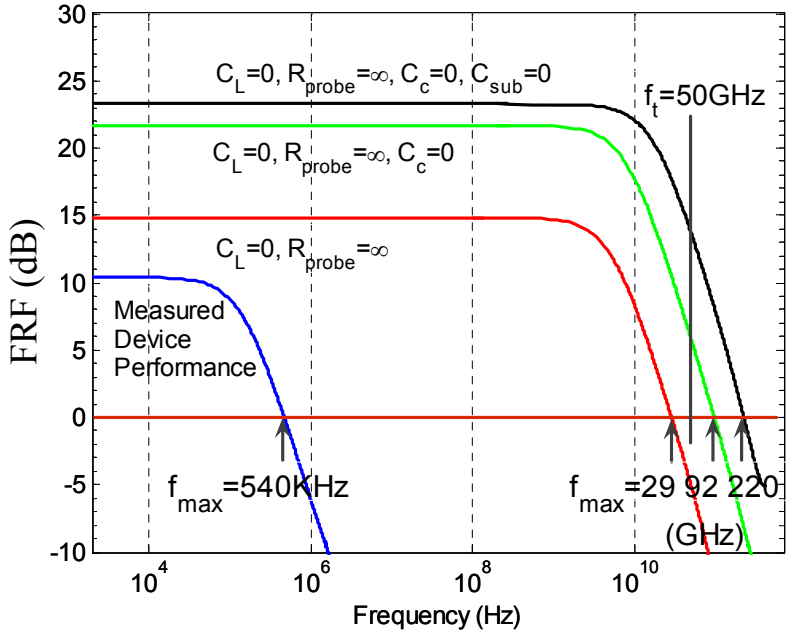



**Figure B.4:** The predicted frequency response function (FRF) of the CNFET circuit as a function of frequency under more ideal conditions. Symbol "Δ" shows the well-fitted voltage gain for the measurement circuit (with a 620 KΩ load resistor). Symbol "o" predicts the actual circuit performance without the interference from the measurement, i.e. the load capacitance is zero and the probe resistance is infinite. The curve with symbol "*" predicts the circuit performance assuming there is no parasitic coupling capacitance $C_c$. Symbol "◊" predicts the circuit performance assuming the function of the back gate is substituted with chemical doping. With a 620 KΩ load resistor, the voltage gain increases to 23 dB and the unity gain frequency approaches 220 GHz.

# Appendix C.

## LIST OF PUBLICATIONS

### JOURNAL PAPERS

[1] **J. Deng** and H.-S P. Wong, "A Compact SPICE Model for Carbon Nanotube Field Effect Transistors Including Non-Idealities and Its Application — Part I: Model of the Intrinsic Channel Region," Submitted to *IEEE Transactions on Electron Devices*, 2007.

[2] **J. Deng** and H.-S P. Wong, "A Compact SPICE Model for Carbon Nanotube Field Effect Transistors Including Non-Idealities and Its Application — Part II: Full Device Model and Circuit Performance Benchmarking," Submitted to *IEEE Transactions on Electron Devices*, 2007.

[3] N. Patil, **J. Deng**, K. Ryu, A. Badmaev, C. Zhou, H.-S. P. Wong, and S. Mitra, "Carbon Nanotube Transistor Circuits: Circuit-Level Performance Benchmarking and the Scalability Analysis," Submitted to *IEEE Journal of Solid-State Circuits*, 2007.

[4] **J. Deng**, Kunal Ghosh, and H.-S. Philip Wong, "Modeling Carbon Nanotube Sensors," Submitted to *IEEE Sensors Letters*, 2007.

[5] **J. Deng** and H.-S P. Wong, "Modeling and Analysis of Planar Gate Electrostatic Capacitance for 1-D FET with Multiple Cylindrical Conducting Channels," *IEEE Transactions on Electron Devices*, to appear, 2007.

[6] **J. Deng**, K. Kim, C.-T. Chuang, and H.-S P. Wong, "The Impact of Device Footprint Scaling on High Performance CMOS Logic Technology," *IEEE Transactions on Electron Devices,* vol. 54, pp. 1148-1155*,* 2007.

[7] **J. Deng** and H.-S. P. Wong, "Metrics for performance benchmarking of nanoscale Si and Carbon Nanotube FETs including device non-idealities," *IEEE Transactions on Electron Devices,* vol. 53, pp.1317-1322, June, 2006.

### CONFERENCE PAPERS

[8] A. K. Friesz, A. C. Parker, C. W. Zhou, K. M. Ryu, J. M. Sanders, **J. Deng**, and H. -S. P. Wong, "A Biomimetic Carbon Nanotube Synapse Circuit," Submitted to *2007 BMES Annual Fall Meeting*, Los Angeles,CA, September 26-29, 2007.

[9] **J. Deng**, L. Wei, L.-W. Chang, K. Kim, C.-T. Chuang, H.-S P. Wong, "Extending the Technology Roadmap by Selective Device Footprint Scaling and Parasitics Engineering," Submitted to *2007 ICCAD*, San Jose, CA, 2007.

[10] N. Patil, **J. Deng**, S. Mitra, and H.-S. P. Wong, "Automated Design of Misaligned-Carbon-Nanotube-Immune Circuits," *44th Design Automation Conference (DAC)*, San Diego, CA, June 4-8, 2007.

[11] **J. Deng**, G. C. Wan, and H.-S. P. Wong, "Carbon Nanotube Transistor Compact Model," *invited talk, Workshop on Compact Modeling (WCM),* Santa Clara, CA, May 20 – 24, 2007.

[12] **J. Deng**, K. Kim, C-T. Chuang, and H.-S P. Wong, "Device Footprint Scaling for Ultra Thin Body Fully Depleted SOI," *IEEE International Symposium on Quality Electronic Design (ISQED)*, San Jose, CA, March 26-28, 2007.

[13] N. Patil, **J. Deng**, S. Mitra, and H.-S. P. Wong, "Design of Imperfection-Immune Carbon Nanotube Field Effect Transistor Circuits," *GOMACTech-07 Conference*, Lake Buena Vista, FL*,* March 20, 2007.

[14] **J. Deng**, N. Patil, K. Ryu, A. Badmaev, C. Zhou, S. Mitra, and H.-S. P. Wong, "Carbon Nanotube Transistor Circuits: Circuit-Level Performance Benchmarking and Design Options for Living with Imperfections," *International Solid State Circuits Conference (ISSCC)*, pp. 70-71, San Francisco, CA, February, 2007.

[15] I. Amlani. J. Lewis, K. Lee, R. Zhang, **J. Deng**, and H.-S. P. Wong, "First Demonstration of AC Gain From a Single-Walled Carbon Nanotube Common-Source Amplifier," *IEEE International Electron Devices Meeting (IEDM)*, pp. 559-562, San Francisco, CA, December 11-13, 2006.

[16] H.-S. P. Wong, **J. Deng**, A. Hazeghi, T. Krishnamohan, and G. C. Wan, "Carbon Nanotube Transistor Circuits – Models and Tools for Design and Performance Optimization," *2006 International Conference on Computer-Aided Design (ICCAD)*, San Jose, CA, November 7, 2006.

[17] **J. Deng** and H.-S. P. Wong, "A Circuit Compatible SPICE model for Enhancement Mode Carbon Nanotube Field Effect Transistors," *2006 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD),* pp.166-169, Monterey, CA, September, 2006.

[18] **J. Deng**, N. Patil, S. Mitra, and H.-S. P. Wong, "Designing Circuits with Carbon Nanotubes: Open Questions and Some Possible Directions," invited paper, *IEEE Nano, Special Session on Nanocircuits and Architectures: Manufacturing Issues and Reliability*, Cincinnati, Ohio, July 17 – 20, 2006.

[19] **J. Deng** and H.-S. P. Wong, "Circuit analysis of sublithographic nanodevice logic array," *Symposium on Nanoscale Materials, Process and Devices,* Hawaii, November, 2005.

[20] **J. Deng** and H.-S. P. Wong, "Metrics for performance benchmarking of nanoscale Si and carbon nanotube FET," *Symposium on Nanoscale Materials, Process and Devices,* Hawaii, November, 2005.