CARBON NANOTUBE ELECTRONICS: MODELING, PHYSICS, AND APPLICATIONS

A Thesis

Submitted to the Faculty

of

Purdue University

by

Jing Guo

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August, 2004

ACKNOWLEDGMENTS

I would like to express my deep gratefulness to my thesis advisor, Prof. Mark Lundstrom who made the whole work possible. My experience of working as a student of Prof. Lundstrom is an invaluable treasure, which will benefit my whole life. He teaches me how to approach difficult research topics with simple, neat ways; he creates every opportunity to help me to connect, learn, and benefit from top researchers in the field. Prof. Lundstrom's contribution to this work is more than what can be described by words.

It is a great pleasure to thank Prof. Supriyo Datta for providing insightful suggestions and devoting a lot of his precious time to the work. I am also deeply indebted to Ali Javey and Prof. Hongjie Dai in Stanford University for extensive discussions and collaborations. I want to thank Profs. Kaushik Roy and Ron Reifenberger for serving on my committee.

It is a great joy to work in the Purdue computational electronics group, with generous help from Prof. Muhammad Alam, Dr. Zhibin Ren, Dr. Ramesh Venugopal, Dr. Jung-Hoon Rhew, Dr. Mani Vaidyananthan, Dr. Diego Kienle, Anisur Rahaman, Sayed Hasan, Jing Wang, Neophytos Neophytou, Dr. Avik Ghosh, Titash Rakshit, and Geng-Chiau Liang.

Finally I want to thank my parents, my brother, and my fiancée, Rachel Y. Zhang, for their enormous sacrifice to support my work.

This work was supported by the National Science Foundation under Grant No. EEC-0228390, and the MARCO Focus Center on Materials, Structures and Devices.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Figure Page

ABSTRACT

Jing Guo, Ph. D., Purdue University, August, 2004. Carbon Nanotube Electronics: Modeling, Physics, and Applications. Major Professor: Mark Lundstrom.

In recent years, significant progress in understanding the physics of carbon nanotube electronic devices and in identifying potential applications has occurred. In a nanotube, low bias transport can be nearly ballistic across distances of several hundred nanometers. Deposition of high-κ gate insulators does not degrade the carrier mobility. The conduction and valence bands are symmetric, which is advantageous for complementary applications. The bandstructure is direct, which enables optical emission. Because of these attractive features, carbon nanotubes are receiving much attention. In this work, simulation approaches are developed and applied to understand carbon nanotube device physics, and to explore device engineering issues for better transistor performance.

Carbon nanotube field-effect transistors (CNTFETs) provide a concrete context for exploring device physics and developing a simulation capability. We have developed an empirical (p_z orbital) atomistic, quantum simulator for nanotube transistors. This simulator uses the non-equilibrium Green's function (NEGF) formalism to treat ballistic transport in the presence of self-consistent electrostatics. We also separately developed a coupled Monte-Carlo/quantum injection simulator to understand carrier scattering in CNTFETs.

Numerical simulations are used to understand device physics and to explore device engineering issues. In chapter 4, we did a comprehensive study of the scaling behaviors for ballistic SB CNTFETs. In chapter 5, we analyzed a short-channel, high-performance CNTFET, to understand what controls and how to further improve the transistor performance. In chapter 6, we explored the interesting role of phonon scattering in CNTFETs.

1. INTRODUCTION

1.1 Overview

Since the discovery of carbon nanotubes (CNTs) by Iijima in 1991[1], significant progress has been achieved for both understanding the fundamental properties and exploring possible engineering applications [2]. The possible application for nanoelectronic devices has been extensively explored since the demonstration of the first carbon nanotube transistors (CNTFETs) [3, 4]. Carbon nanotubes are attractive for nanoelectronic applications due to its excellent electric properties. In a nanotube, low bias transport can be nearly ballistic across distances of several hundred nanometers. Deposition of high-κ gate insulators does not degrade the carrier mobility because the topological structure results in an absence of dangling bonds. Fermi level pining at the metal-nanotube interface is weak, so a range of Schottky barrier heights can be achieved by using different contact metals. The conduction and valence bands are symmetric, which is advantageous for complementary applications. The bandstructure is direct, which enables optical emission, and finally, CNTs are highly resistant to electromigration. Significant efforts have devoted to understand how a carbon nanotube transistor operates and to improve the transistor performance [5, 6]. It has been demonstrated that most CNTFETs to date operates like non-conventional Schottky barrier transistors [7, 8], which results in quite different device and scaling behaviors from the MOSFET-like transistors [9, 10]. Important techniques for significantly improving the transistor performance, including the aggressively scaling of the nanotube channel, integration of thin high-κ gate dielectric insulator [11, 12], use of excellent source/drain metal contacts [13], and demonstration of the self-align techniques, have been successfully developed. Very recently, a nanotube transistor, which integrates ultra-short channel, thin high-κ top gate insulator, excellent Pd source/drain contacts is demonstrated using a self-align technique [14]. Promising transistor performance exceeding the state-of-the-art Si

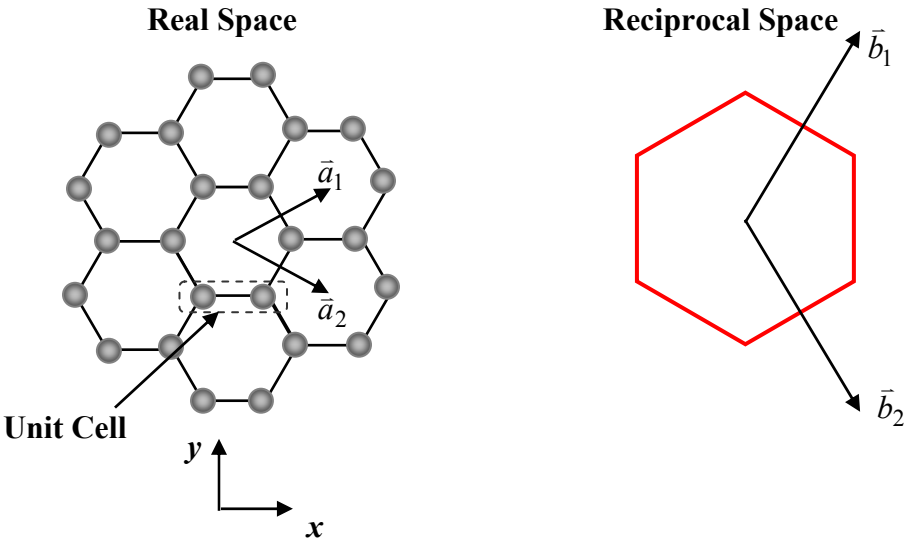
MOSFETs is achieved. The transistor has a near-ballistic source-drain conductance of $\sim 0.5 \times 4e^2 / h$ and delivers a current of $\sim 20 \mu A$ at $|V_G\text{-}V_T| \sim 1V$.

In this work, numerical simulations are developed to explain experiments, to understand how the transistor operates and what controls the performance, and to explore the approaches to improve the transistor performance. New simulation approaches are necessary for a carbon nanotube transistor because it operates quite different from Si transistors. The carbon nanotube channel is a quasi-one-dimensional conductor, which has fundamentally different carrier transport properties from the Si MOSFET channel. It has been demonstrated that treating the Schottky barriers at the metal/CNT interface and near-ballistic transport in the channel are important for correctly modeling the transistor. The CNT channel is a cylindrical semiconductor with a ~1nm diameter, which means the electrostatic behavior of the transistor is quite different from Si MOSFETs with a 2D electron gas. All carbon bonds are well satisfied at the carbon nanotube surface, which results in a different semiconductor/oxide interface. Furthermore, the phonon vibration modes and carrier scattering mechanisms are quite different in carbon nanotubes, which results in different roles of phonon scattering in CNTFETs. In this work, we developed physical simulation approaches to treat CNTFETs. We will show that our understanding of the carrier transport, electrostatics, and interracial properties seem to be sufficient to describe the behavior of the recently demonstrated short-channel CNTFETs [14].

1.2 Carbon Nanotube Basics

1.2.1 Graphene sheet

The nanotube can be conceptually viewed as a rolled-up graphene sheet [6, 15]. A simple way to calculate the one-dimensional E-k relation of carbon nanotube, which governs its electronic property, is to quantize the two-dimensional E-k of the graphene sheet along the circumfencial direction of the nanotube. Thus the first step to calculate the nanotube E-k is to calculate the band structure of the graphene sheet.



(a) (b)

Fig. 1.1 (a) The graphene lattice in real space with the basis vectors $\vec{a}_1$ and $\vec{a}_2$. (b) The first Brillouin zone of the reciprocal lattice with the basis vectors $\vec{b}_1$ and $\vec{b}_2$.

The two-dimensional graphene lattice in real space can be created by translating one unit cell by the vectors $\vec{T} = n\vec{a}_1 + m\vec{a}_2$ with integer combinations (*n,m*), where $\vec{a}_1$ and $\vec{a}_2$ are basis vectors (as shown in Fig. 1.1),

$$\vec{a}_1 = a_0(\frac{\sqrt{3}}{2}\hat{x} + \frac{1}{2}\hat{y})$$

$$\vec{a}_1 = a_0(\frac{\sqrt{3}}{2}\hat{x} - \frac{1}{2}\hat{y}) \qquad (1.1),$$

$a_0 = \sqrt{3}a_{cc}$ is the length of the basis vector, and $a_{cc} \approx 1.42\,\overset{o}{A}$ is the nearest neighbor C-C bonding distance.

A tight binding model, which includes one p_Z orbital per carbon atom and the nearest neighbor interaction, is used to calculate the graphene band structure. More detailed calculations including multiple orbitals and more levels of neighboring atoms show that the one-obital, tight-binding approximation works well at the energy range near the Fermi point of the graphene sheet, which is the region of interest for electronic transport [16]. Because the E-k relation describes the eigen-energies of the plane wave state (with wave vector $\vec{k}$) in a periodic crystal lattice, we write down the wave vector-dependent Hamiltonian for one unit cell, which treats the C-C bonding within the unit cell itself and the bonding with neighboring unit cells.

$$H(\vec{k}) = t \cdot \begin{bmatrix} 0 & 1 + e^{i\vec{k}\cdot\vec{a}_1} + e^{i\vec{k}\cdot\vec{a}_2} + e^{i\vec{k}\cdot\vec{a}_3} \\ 1 + e^{-i\vec{k}\cdot\vec{a}_1} + e^{-i\vec{k}\cdot\vec{a}_2} + e^{-i\vec{k}\cdot\vec{a}_3} & 0 \end{bmatrix} \tag{1.2}$$

where $t \approx -3.0eV$ is the C-C bonding energy and $\vec{a}_3 = \vec{a}_1 - \vec{a}_2$.

The E-k relation of the graphene sheet is then calculated by solving the eigen-energies of the Hamiltonian matrix in eqn. 1.2,

$$E(\vec{k}) = \pm |t| \cdot \sqrt{3 + 2\cos(\vec{k}\cdot\vec{a}_1) + 2\cos(\vec{k}\cdot\vec{a}_2) + 2\cos(\vec{k}\cdot\vec{a}_3)} \ . \tag{1.3}$$

where the positive sign is for the conduction band and the negative one for the valence band. In contrast to Si, which is an indirect band gap semiconductor and has asymmetric bandstructures for electrons and holes, graphene has symmetric conduction and valence bands.

We next show that the energy valleys are located at the corners of the Brillouin zones, which are usually referred as the Fermi points. The basis vectors in the reciprocal lattice $\vec{b}_j$, as shown in Fig. 1.1 (b), satisfies

$$\vec{a}_i \cdot \vec{b}_j = 2\pi\delta_{ij}, \tag{1.4}$$

where $\vec{a}_i$ are the basis vectors of the real space lattice expressed as eqn. (1.1) and $\vec{b}_j$ is computed as,

$$\vec{b}_1 = b_0(\frac{1}{2}\hat{x} + \frac{\sqrt{3}}{2}\hat{y})$$

$$\vec{b}_2 = b_0(\frac{1}{2}\hat{x} - \frac{\sqrt{3}}{2}\hat{y}), \tag{1.5}$$

where $b_0 = \frac{4\pi}{\sqrt{3}a_0}$ is the length of the basis vector in the reciprocal space. The wave vectors at the six corners of the Brillouin zone can be expressed in terms of b_1 and b_2 as

$$\vec{k}_F = (u \pm \frac{1}{3})\vec{b}_1 + (v \mp \frac{1}{3})\vec{b}_2, \tag{1.6}$$

where *u* and *v* are integers. Among the six valleys in the first Brillouin zone, only two of them are independent.

By substituting $\vec{k}_F$ to eqn. (1.3), we can show that the energy at the Fermi points of the Brillouin zones is zero,

$$E(\vec{k}) = \pm |t| \cdot \sqrt{3 + 2\cos(\vec{k}_F \cdot \vec{a}_1) + 2\cos(\vec{k}_F \cdot \vec{a}_2) + 2\cos(\vec{k}_F \cdot \vec{a}_3)}$$

$$= \pm |t| \cdot \sqrt{3 + 2\cos(\pm\frac{2}{3}\pi) + 2\cos(\mp\frac{2}{3}\pi) + 2\cos(\pm\frac{4}{3}\pi)} = 0. \tag{1.7}$$

Equation (1.3), which gives an analytical expression for the E-k relation, can be further simplified by Taylor expansion of the cosine function near the Fermi point. The

simplified E-k is isotropic around the Fermi point and indicates a linear dispersion relation,

$$E(\vec{k}) = \frac{3a_{cc} \mid t \mid}{2} \mid \vec{k} - \vec{k}_F \mid, \tag{1.8}$$

which indicates the E-k relation near the Fermi point is linear and isotropic. This linear E-k approximation agrees with the E-k in Eq. (1.3) within the energy range ~1eV near the Fermi point. Due to its mathematical simplicity, Eq. (1.8) is useful for deriving analytical forms of other electronic properties, such as density-of-states [17].

1.2.2 Carbon nanotubes

A carbon nanotube can be viewed as a rolled graphene sheet along its circumferential direction, $\vec{c} = n\vec{a}_1 + m\vec{a}_2$, where $\vec{a}_1$ and $\vec{a}_2$ are the basis vectors of the graphene sheet (in Fig. 1.1). Two special kinds of CNTs are defined as 1) the zigzag CNT when $m = 0$, and 2) the armchair CNT when $n = m$. CNTs other than these two special kinds are generally referred as chiral nanotubes.

Next we calculate the E-k relation of CNTs by discritizing the linear E-k relation of the graphene sheet in eqn. (1.8)]. The periodic boundary condition imposed along the circumference direction restricted the wave vectors to

$$\vec{k} \cdot \hat{c} = 2\pi q, \tag{1.9}$$

where $\vec{k}$ is an allowed wave vector and q is the quantum number.

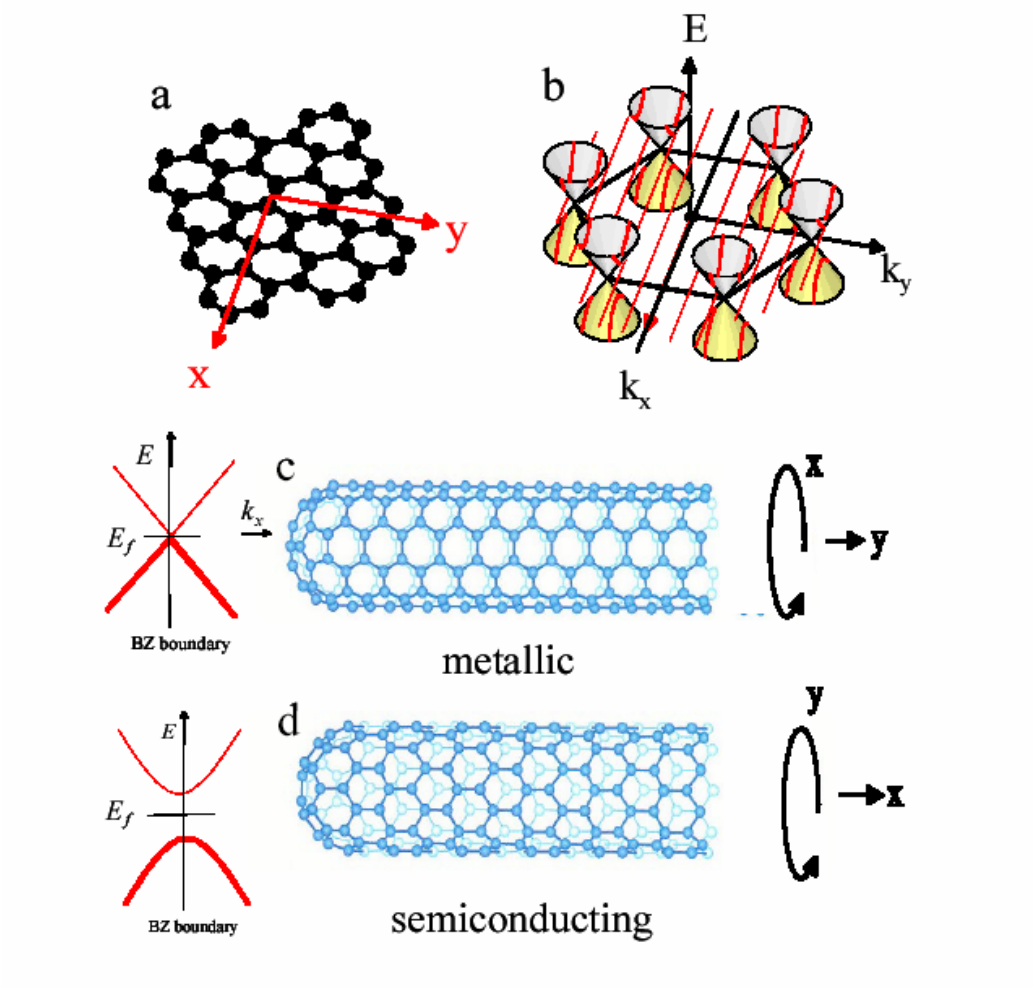


Fig. 1.2 Carbon nanotubes can be viewed as a rolled graphene sheet. The periodic boundary condition only allows quantized wave vectors around the circumferential direction, which generates one-dimensional bands for carbon nanotubes [6].

The E-k near the Fermi-points is the most interesting. We choose one Fermi-point, $\vec{k}_F = \frac{1}{3}\vec{b}_1 - \frac{1}{3}\vec{b}_2$, and compute its component along the circumferential direction,

$$\vec{k}_F \cdot \hat{c} = \frac{n-m}{3} \cdot 2\pi \,. \tag{1.10}$$

If the origin of the reciprocal lattice is reset to the Fermi point, the wave vector in the new coordinate system is

$$\vec{k}' = \vec{k} - \vec{k}_F = k_c'\hat{c} + k_t'\hat{t} \,, \tag{1.11}$$

where k_c' is the component along the circumference direction, which is quantized by the periodic boundary condition

$$k_{c,q}' = (\vec{k} - \vec{k}_F) \cdot \hat{c} = \frac{\vec{k} \cdot \vec{c} - k_F \cdot \vec{c}}{|\vec{c}|} = \frac{1}{3d}[3q - (n-m)] \tag{1.12}$$

and *d* is the diameter of the nanotube.

Based on eqn. (1.8), the linear E-k approximation for the graphene sheet, the E-k relation of the CNT is

$$E(\vec{k}) = \frac{3a_{cc}\,|t|}{2}|\vec{k}'| = \frac{3a_{cc}\,|t|}{2}\sqrt{k_{c,q}'^{\,2} + k_t'^{\,2}} \tag{1.13}$$

The lowest subband of the CNT is determined by the minimum value of $|k_{c,q}|$. The nanotube can be either metallic or semiconducting, depending on whether *(n-m)* is the multiple of 3.

1) If $(n-m) \bmod 3 = 0$, the CNT is metallic.

The minimum $k'_{c,q} = 0$ at $q = (n-m)/3$. The one-dimensional E-k relation of the nanotube is

$$E = \pm \frac{3a_{CC} \mid t \mid}{2} k'_t, \tag{1.14}$$

which is a one-dimensional linear dispersion relation independent of (*n,m*), as shown in Fig. 1.3. The Fermi level is located at $E = 0$, and this type of nanotube is referred to as semi-metallic. Note that the bandgap is zero. The 1D density of states contributed by the lowest subband of the metallic CNT is constant,

$$D(E) = 2 \times 2 \times \frac{1}{L} \sum_{\Delta k_t} \delta[E - E'(\Delta k_t)] = \frac{8}{3\pi a_{cc} \mid t \mid}. \tag{1.15}$$

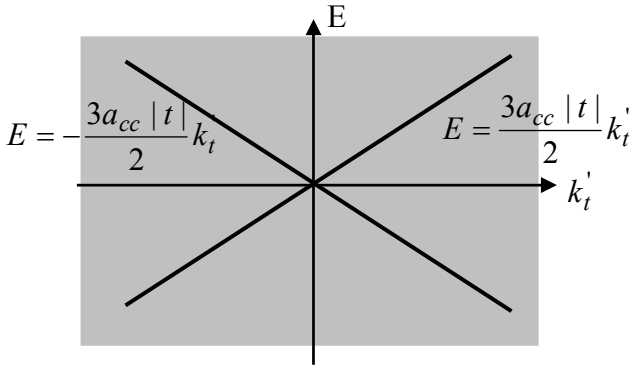


Fig. 1.3 The E-k relation of a CNT metallic band.

2) If $(n-m) \bmod 3 \neq 0$, the CNT is semiconducting.

The E-k relation for the lowest subband is determined by the minimum value of

$$k_{c,q} = \frac{2}{3d}, \tag{1.16}$$

where d is the diameter of the CNT.

By substituting Eq. (1.16) into the linear E-k approximation for graphene as shown in eqn. (1.8), we get

$$E(k_t') = \pm \frac{3a_{CC} \mid t \mid}{2} \sqrt{k_t'^2 + (2/3d)^2} \tag{1.17}.$$

The band gap is

$$E_G = \frac{2a_{cc} \mid t \mid}{d} \approx \frac{0.8eV}{d}, \tag{1.18}$$

where the units of d are nm. Based on this simple derivation, the E(k) relation and the bandgap are functions of the CNT diameter alone.

The one-dimensional density of states for one semiconducting band is,

$$D(E) = 2 \times 2 \times \frac{1}{L} \sum_{\Delta k_t} \delta[E - E'(\Delta k_t)] = D_0 \frac{\mid E \mid}{\sqrt{E^2 - (E_G/2)^2}} \Theta(\mid E \mid - E_G/2) \tag{1.19}$$

where $D_0 = \frac{8}{3\pi a_{CC} \mid t \mid}$ is the constant metallic band DOS, $\Theta(x)$ is the step function which equals 1 for $x > 0$ and 0 otherwise. Each band produces singularities at the conduction and valence band edges, as shown in Fig. 1.4.

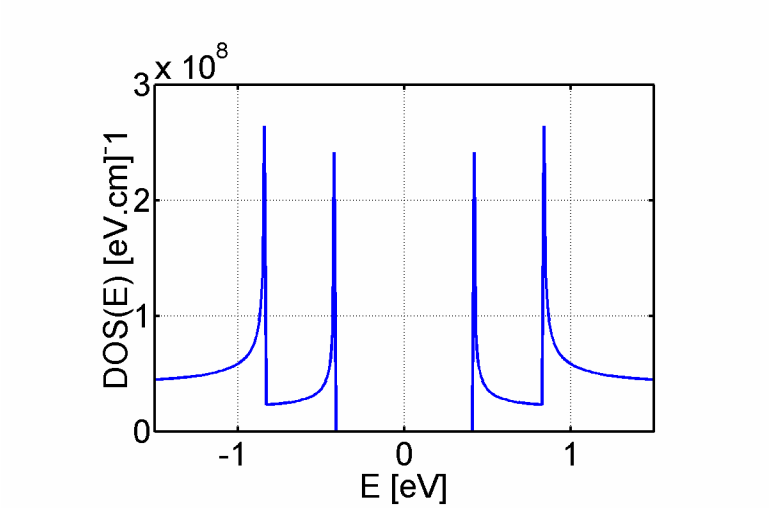


Fig. 1.4 The DOS of (13,0) CNT calculated by eqn. (1.19).

1.3 Outline of the Thesis

This thesis is organized as the following. Chapter 2 talks about the interesting electrostatic behavior of carbon nanotube devices due to its one-dimensional channel geometry. Chapter 3 describes a self-consistent quantum transport solver based on non equilibrium Green's function (NEGF) formalism for ballistic carbon nanotube transistors. Chapter 4 and 5 apply this quantum transport solver to address device related issues. Chapter 4 provides a comprehensive study of the scaling behaviors for Schottky barrier carbon nanotube transistors. Chapter 5 addresses device physics issues based on a detailed analysis a recently demonstrated short-channel, high-performance carbon nanotube transistor. Chapter 6 studies the role of phonon scattering, which is the dominating scattering mechanism for carbon nanotubes, in carbon nanotube transistors. The last chapter, chapter 7, concludes the whole thesis and also gives the directions for future research.

2. ELECTROSTATICS OF CARBON NANOTUBE DEVICES

2.1 Introduction

With the scaling limit of conventional silicon transistors in sight, there is rapidly growing interest in nanowire transistors with one-dimensional channels, such as carbon nanotube transistors [5, 6] and silicon nanowire transistors [18-21]. Due to the one-dimensional channel geometry, the electrostatics of nanowire devices can be quite different from bulk silicon devices. Previous studies of carbon nanotube p/n junctions and metal/semiconductor junctions demonstrated unique properties of nanotube junctions [22, 23]. For example, the charge transfer into the nanowire channel from the metal contacts (or heavily doped semiconductor contacts) can be significant [23, 24].

In this paper, we extend previous studies by looking at the dependence of the charge transfer on the metal/semiconductor Schottky barrier height, the insulator dielectric constant, and the metal contact geometry. We show that if an intrinsic nanowire is attached to bulk metal contacts at two ends, large charge transfer can be achieved if the Schottky barrier is low and the insulator dielectric constant is high. If, however, the intrinsic nanowire is attached to one-dimensional metal contacts, the charge density on the nanowire depends critically on the electrostatic environment rather than the properties of the metal contacts. Reducing the gate oxide thickness and the contact size decreases the distance over which the source/drain field penetrates into the nanowire channel and can, therefore, help to suppress the short channel effects and improve the transistor performance.

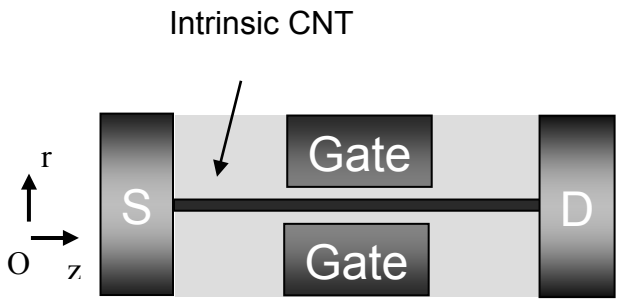


Fig. 2.1 The modeled, coaxially gated carbon nanotube transistor. The intrinsic nanotube channel has a diameter of 1.4nm and the gate work function is zero. The cylindrical coordinates for solving the Poisson equation is also shown.

2.2 Approach

We simulated the coaxially gated carbon nanotube transistor shown in Fig. 2.1. Although the calculations are for carbon nanotube transistors, the general conclusion should apply to other nanowire transistors with one-dimensional channels. The equilibrium band profile and charge density were obtained by solving the Poisson equation in cylindrical coordinates self-consistently with the equilibrium carrier statistics of the carbon nanotube. The charge density per unit length on the nanotube, $Q_L(z)$, is calculated by integrating the "universal" nanotube density-of-states (DOS) [17], $D(E)$, over all energies,

$$Q_L(z) = (-e) \cdot \int_{-\infty}^{+\infty} dE \cdot \mathrm{sgn}(E) D(E) f(\mathrm{sgn}(E)[E - \widetilde{E}_F(z)]), \tag{2.1}$$

where e is the electron charge, $\mathrm{sgn}(E)$ is the sign function, and $\widetilde{E}_F(z) = E_F - E_m(z)$ is the Fermi energy level minus the middle gap energy of the nanotube, $E_m(z)$. Since the source/drain electrodes are grounded, the Fermi level is set to zero, $E_F = 0$. The nanotube middle gap energy is computed from the electrostatic potential at the nanotube

shell, $E_m(z) = -eV(z, r = r_{cnt})$, where r_{cnt} is the nanotube radius. The electrostatic potential, V, satisfies the Poisson equation,

$$\nabla^2 V(z,r) = -\frac{\rho}{\varepsilon} \tag{2.2}$$

where ρ is the charge density, ε is the dielectric constant. The following boundary conditions were used,

$V = (Eg/2 - \phi_{bn})/e$ at the left metal contact,

$V = (E_g/2 - \phi_{bn})/e$ at the right metal contact, and

$V = V_G$ at the gate cylinder (the flat band voltage is assumed to be zero),

where E_g is the nanotube bandgap, ϕ_{bn} is the Schottky barrier height for electrons between the source/drain and the nanotube, and V_G is the gate voltage.

We numerically solved the Poisson equation by two methods, 1) the finite difference method and 2) the method of moments [25]. In order to improve the convergence when iteratively solving eqns. (2.1) and (2.2), the Netwton-Ralphson method (with details in [26]) was used. The results obtained by the finite difference method and by the method of moments agree well.

2.3 Results

We first compare the charge transfer from bulk contacts to the one-dimensional carbon nanotube to the charge transfer to a bulk silicon channel. We simulated two cases: 1) an intrinsic bulk Si channel sandwiched between two metal contacts as shown in Fig. 2.2a, and 2) an intrinsic carbon nanotube channel between metal contacts as shown in Fig. 2.2b. In both cases, the Schottky barrier heights between the metal contacts and the semiconductor channel are zero, which aligns the metal Fermi level of to the conduction band edge of the semiconductor. Electrons are transferred from metal contacts into the

intrinsic channel due to the work function difference between the metal and the semiconductor. Fig. 2.2c plots the conduction bands, and Fig. 2.2d plots the charge densities in the unit of electron per atom for the bulk Si and nanotube channel. Compared to the bulk Si channel, the barrier in the nanotube is much lower, and the charge density is much higher. Although the nanotube is $3\mu m$ long, the charge density at the center of the tube is still as high as 10^{-4}e/atom , about 5 orders of magnitude higher than that of the bulk Si in terms of electron fraction. As the result, the carbon nanotube channel is more conductive.

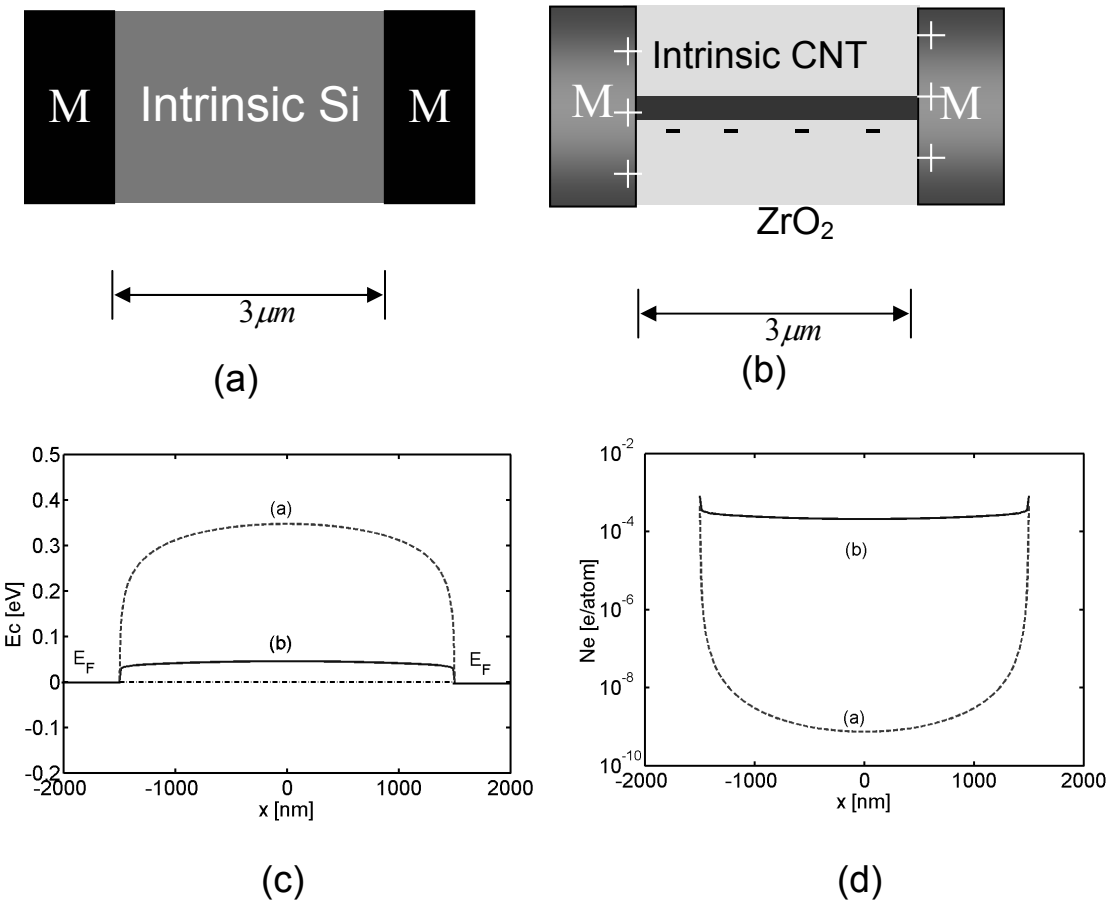


Fig.2.2 The schematic plots for (a) a bulk Si structure where the cross-sectional area is assumed to be large (b) a carbon nanotube channel between bulk metal electrodes. The Schottky barrier heights for electrons are zero. (c) The conduction band edge and (d) the electron density in the units of doping fraction. Results for the bulk Si structure are shown as dashed lines and for nanotube as solid lines.

The charge transfer to the tube is significant because the charge on tube doesn't effectively screen the potential produced by the bulk contacts. Compared to the bulk channel, the charge element on the nanotube only changes potential locally. For example, in the bulk channel, the charge element is a two-dimensional sheet charge, which produces a constant field. The charge dipole formed by charge sheet in bulk Si and metal contacts shifts the potential far away. In contrast, for the nanowire channel, the charge element is a point charge, which produces a potential decaying with distance ~1/r and has little effect far away (the potential of a point charge dipole decays even faster as ~$1/r^2$). As the result, for the one-dimensional channel, the potential produced by the bulk contacts is not screened by the charge on the nanotube near the metal/semiconductor interface. The bulk contacts tend to put the conduction band edge near the Fermi level over the whole $3\mu m$ -long tube if the metal/CNT barrier height is zero.

We next estimate the charge density in the channel. The estimation provides a simple way to understand how the charge density of the tube varies with the contact and insulator properties. For the device structure shown in Fig. 2b, if the metal contacts are grounded, and the metal/semiconductor work function difference is $U_0 = \phi_{CNT} - \phi_M$, where ϕ_{CNT} (ϕ_M) is the nanotube (metal) work function, the electron density is

$$n(z) = \overline{D}(U_0 - U(z)), \tag{2.3}$$

where *U(z)* is the electron potential energy produced by charge in the channel, and $\overline{D}$ is the average density-of-states for the energy between the nanotube middle gap energy and the Fermi level. The charge element in the one-dimensional channel only shifts the potential locally, we approximately relate the potential, *U(x)*, to the electron density at the same position, $n(z)$,

$$U(x) = e^2 n(z) / C_{ins} \tag{2.4}$$

where C_{ins} is the electrostatic capacitance per unit length between the nanotube and the bulk contacts. The electron density due to the charge transfer from the bulk contacts can be obtained from eqns. (2.3) and (2.4) as

$$n(x) = \frac{U_0 / e^2}{1/C_{ins} + 1/\overline{C}_Q}, \tag{2.5}$$

where the quantum capacitance [27] is defined as, $\overline{C}_Q = e^2 \overline{D}$, which is proportional to the average DOS of the nanotube. Equation (2.5) can be interpreted in a simple way. The bulk electrodes modulate the charge density of the nanotube through an insulator capacitor, C_{ins}, which is in series with the quantum capacitance of the nanotube.

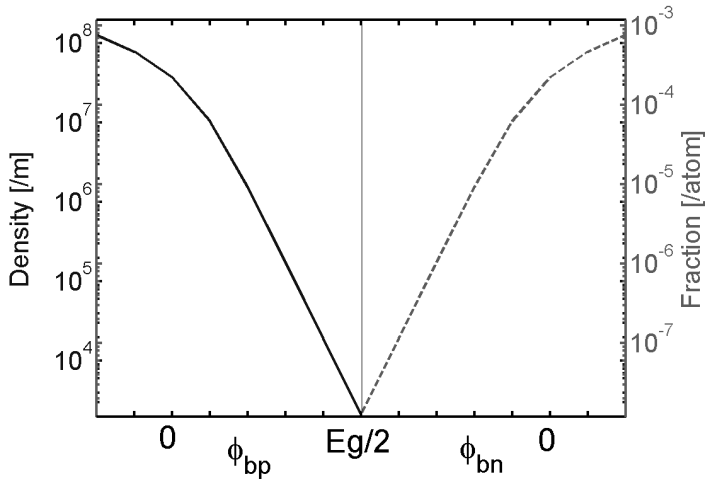


Fig.2.3 The electron density (the dashed line) and hole density (the solid line) at the center of the 3 μm-long CNT (in Fig. 2.2b) vs. the Schottky barrier height for electrons, ϕ_{bn}, and that for holes, ϕ_{bp}. The left axis shows the charge density in the unit of number of electrons (holes) per unit length and the right axis shows the same quantity in the unit of charge fraction.

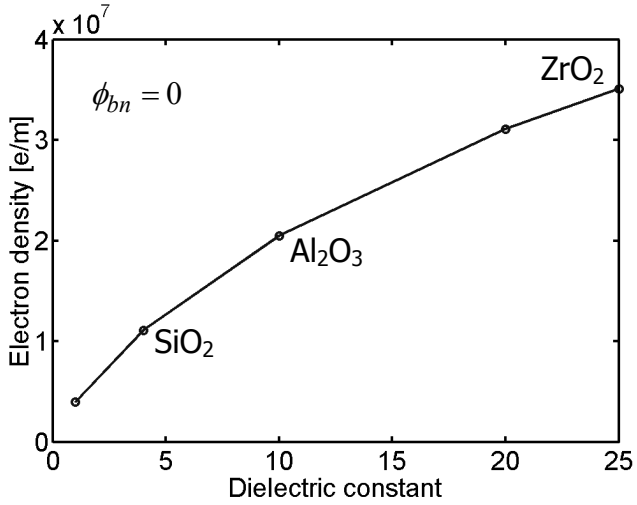


Fig. 2.4 The electron density at the center of the $3\mu m$-long tube (in Fig. 2.2b) vs. the insulator dielectric constant. The Schottky barrier height for electrons, ϕ_{bn}, is zero.

We now examine how the charge transfer varies with the Schottky barrier height and the insulator dielectric constant. Fig. 2.3, which plots the charge density at the center of the tube as shown in Fig. 2.2b vs. the barrier height, shows that when the barrier height decreases, the charge density first increases. Fig. 2.4, which plots the charge density at the center of the tube vs. the insulator dielectric constant, shows that the charge density increases as the dielectric constant increases. The dependence of the charge density on the barrier height and the dielectric constant can be easily understood based on eqn. (2.5). Lowering the barrier height increases the metal/CNT work function difference, U_0, and increasing the insulator dielectric constant increases C_{ins}, both of which increase the electron density, $n(x)$ (or hole density if the metal/semiconductor barrier height is lower for holes).

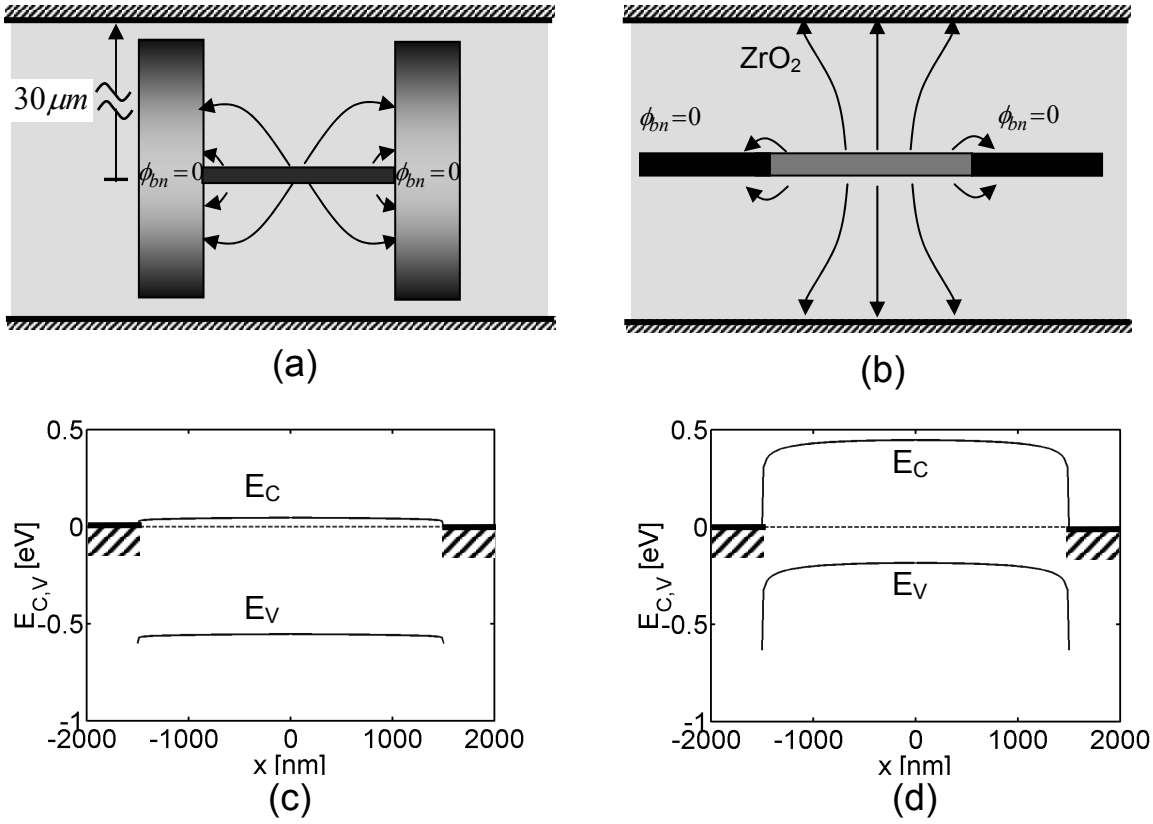


Fig.2.5. *Contact geometry.* A $3\mu m$-long CNT between (a) the bulk contacts and (b) the one-dimensional wire contacts. The tube diameter is 1.4nm. and Schottky barrier heights for electrons are zero. A coaxial gate far away with a $30\mu m$ radius is grounded. The workfunction of the gate metal equals to the semiconductor affinity plus the band gap, so that the gate tends to dope the CNT to p-type. (c) The band profile (a). (d) The band profile for (b).

The importance of charge transfer into the carbon nanotube channel by one-dimensional metal contacts has been previously discussed in [23]. We, however, reached the same conclusion that charge transfer into the one-dimensional channel is significant for a *different* contact geometry (the bulk contacts). We also explored the one-dimensional contacts. In this case, the results are quite different from bulk contacts. The charge density of the nanotube channel is critically determined by the electrostatic environment (i.e., the potential and location of nearby bulk contacts) rather than the metal-contact properties, as will be discussed in detail next.

Fig. 2.5 illustrates the important role of the contact geometry. We simulated: 1) a CNT between grounded bulk contacts as shown in Fig. 2.5a, and 2) a CNT between grounded wire contacts as shown in Fig. 2.5b. In both cases, the tube length is $3\mu m$ and a grounded, coaxial gate cylinder is far away with a radius of $30\mu m$. The S/D contacts have zero Schottky barrier heights for electrons thus tend to dope the tube n-type, while the gate has a high work function and zero barrier height for holes thus tends to modulate the tube to p-type. For the bulk contact case, the whole tube is doped to n-type by bulk contacts and the charge density on the tube is independent of the voltage on the gate cylinder. In contrast, for the wire contacts, the tube is lightly modulated to p-type and the charge density on the tube is very sensitive to the potential on the gate, although it is far away. The results shown in Fig. 2.5 can be explained as follows. For the bulk contacts, because the gate cylinder is far away, the bulk contacts at the ends collect all field lines and image all charge on the tube, as shown in Fig. 5a. For the wire contacts, however, the potential produced by the charge on the one-dimensional wire decays rapidly with distance, thus several nanometer away from the metal/semiconductor interface, the wire contacts have little effects. On the other hand, the capacitance between the gate cylinder and the tube decays slowly (logarithmically) with the tube radius, thus several nanometer away from the metal/semiconductor interface, the charge on the tube images on the gate rather than the wire contacts nearby. As a result, the charge density is determined by the potential on the gate. The charge density on the nanotube channel is essentially determined by the electrostatic environment.

One consequence of the significant charge transfer is that nanowire transistors with large gate underlap can still operate. Fig. 2.6a shows a coaxially gated CNTFET with a 500nm gate underlap and the bulk electrodes. Fig. 2.6b plots the conduction band profile at $V_G = 0$ and 0.3V. At the off state ($V_G = 0V$), a large barrier is created in the channel and the transistor is turned off. At the on-state, ($V_G = 0.3V$), the barrier under the gate is pushed down. Because the low dimensional charge on the ungated nanotube doesn't effectively screen the potential produced by the gate and S/D electrodes, the potential at the ungated region is close to the Laplace potential produced by the source and gate electrodes. The conduction band edge is approximately linear in the ungated region. If the Schottky barrier height between S/D and the channel is ~50meV, the barrier height at the ungated region at the on-state is low enough to deliver an on-current of $\sim 1\ \mu A$. This mechanism provides a possible explanation for the operation of the n-type CNTFET in a recent experiment by Javey et al. [11], in which a n-type CNTFET with large, intrinsic gate underlaps still had a good on-off ratio.

One concern about the nanowire transistors with low meta/CNT Schottky barriers is that due to the significant charge transfer, it might be difficult to turn off the transistor. To examine this concern, we simulated the coaxially gated CNTFET as shown in Fig. 7a with different gate oxide thickness. Fig. 2.7b, which plots the equilibrium band profile, shows that when the gate oxide thickness is the same as the channel length, the source/drain field penetrates into the channel the channel and the transistor cannot be turned off. When the gate oxide is thin, however, the gate still has very good control over the channel and the transistor is well turned off. By solving the Poisson equation for the CNTFET in Fig. 2.7a, the length by which the drain field penetrates into the channel (the scaling length [28]) is estimated to be the radius of the cylindrical gate, $\Lambda \sim R_G$. If the ratio between the channel length and the gate oxide thickness is large, the transistor can be well turned off.

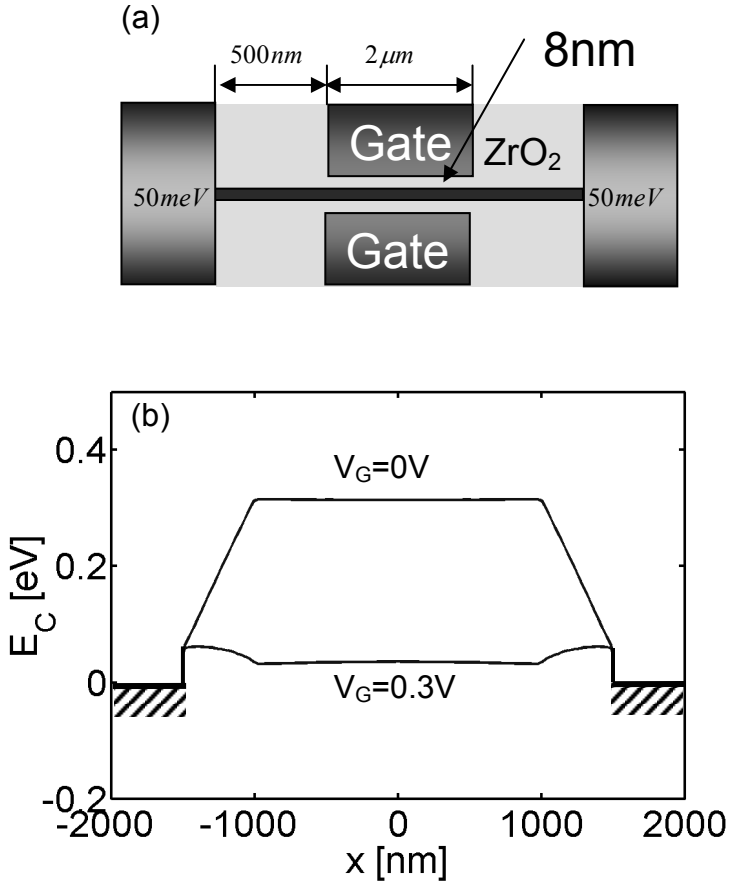


Fig. 2.6. (a) A coaxially gated CNTFET with bulk electrodes (with a radius of 500nm) and a large gate underlap. (b) The conduction band profile at V_G=0V and 0.3V. The metal/CNT barrier height for electrons is 50meV, the ZrO_2 gate oxide thickness 8nm, the tube diameter is1.4nm, the gate length is 2 μm, and the gate underlap is 500nm.

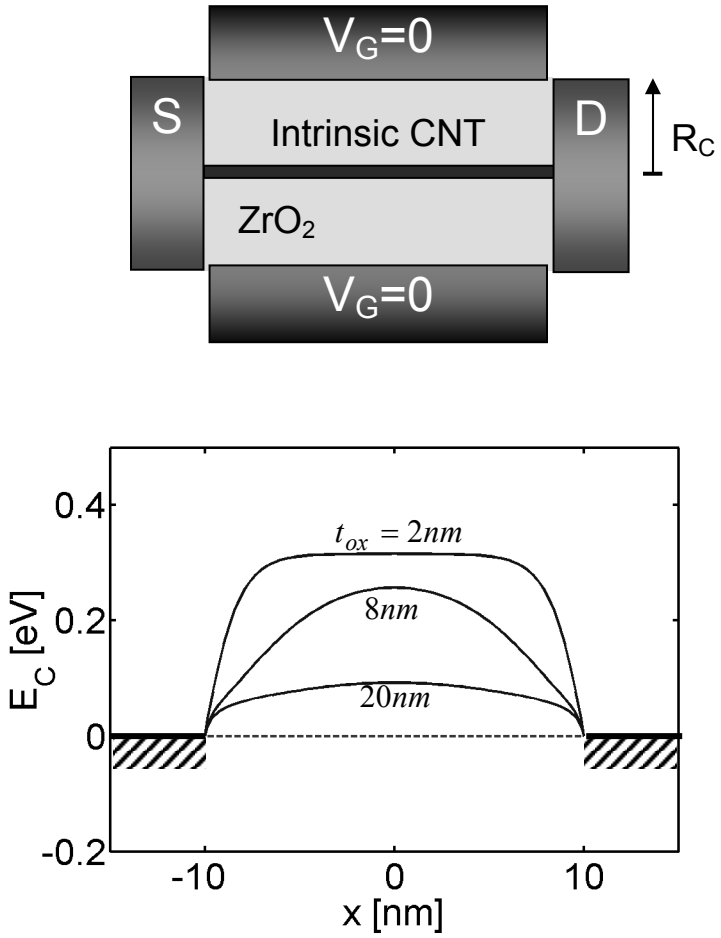


Fig. 2.7. (a) A coaxially gated CNTFET with a 20nm-long, intrinsic channel. The source/drain radius, R_C, is equal to the oxide thickness. The metal/CNT barrier height for electrons is zero, the tube diameter is 1.4nm and the dielectric constant of the gate insulator is $\varepsilon = 25$ (b) the equilibrium conduction band edge at $V_G=0$ for the gate oxide thickness *tox=2nm, 8nm and 20nm*.

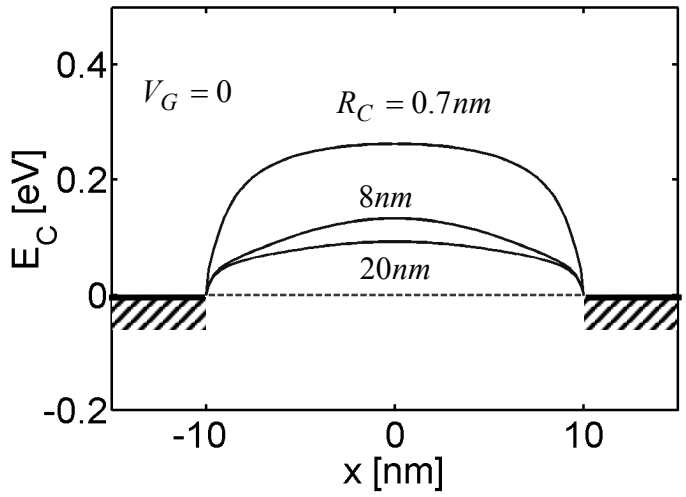


Fig. 2.8. The equilibrium conduction band edge at $V_G=0$ for the CNTFET as shown in Fig., 2.7a. The gate oxide thickness is kept constant at 20nm and the source/drain contact radius, R_C =0.7nm, 8nm, and 20nm.

Another way to reduce the penetration of the lateral field is to reduce the size of the source/drain contact. Fig. 2.8, which plots the equilibrium band profile for the CNTFET (in Fig. 2.7a) with 20nm-thick gate oxide and different contact radius, shows that the screening length for lateral fields from S/D contacts decreases when the contact radius decreases. In the limit when the source/drain electrodes are reduced to wires with the same radius as the tube, the transistor can be well turned off, although the oxide thickness is large. As discussed earlier, the reason it that the potential produced by wire contacts decays rapidly with distance. Improving transistor performance by engineering contacts has been discussed by Heinze et al, when they study the Schottky barrier CNTFETs. Smaller contacts produce thinner Schottky barriers and improve the transistor performance [8].

2.4 Conclusions

The electrostatics of nanowire transistors were explored by self-consistently solving the Poisson equation with the equilibrium carrier statistics. For an intrinsic nanowire attached to bulk contacts, charge transfer is significant if the metal/semiconductor barrier height is low and the insulator dielectric constant is high. The contact geometry also plays an important role. If the contacts are metal wires rather than bulk contacts, the charge density of the nanowire channel is essentially determined by the electrostatic environment rather than the contact properties. The penetration distance of the source/drain field can be engineered by the gate oxide thickness and the contact size, which may provide ways to suppress the electrostatic short channel effects.

3. SIMULATING QUANTUM TRANSPORT IN BALLISIT CARBON NANOTUBES

3.1 Introduction

Carbon nanotubes show promise for applications in future electronic systems, and the performance of carbon nanotube transistors, in particular, has been rapidly advancing [12, 14]. From a scientific perspective, carbon nanotube electronics offers a model system in which to explore and understand the effects of detailed microstructure of contacts, interfaces, and defects. It is also an opportunity to develop the theory and computational techniques for the atomistic simulation of small electronic devices in general. A detailed treatment of carbon nanotube electronics requires an atomistic description of the nanotube along with a quantum mechanical treatment of electron transport, both ballistic and with the effects of dissipative scattering included. As shown in Fig. 3.1, even for this simple system, multi-scale methods are essential. Metal/nanotube contacts, nanotube/dielectric interfaces, and defects require a rigorous, *ab initio* treatment, but to treat an entire device, simpler, p_z orbital descriptions must be used. Techniques connect different descriptions used for different regions of the device will need to be developed (e.g. the *ab initio* basis functions for the metal/nanotube contacts must be connected to the semi-empirical basis functions for the device itself). For extensive device optimization, continuum, effective mass level models may be necessary, and methods to relate the phenomenological parameters in those approaches to the atomistic models must be developed. For circuit simulation, even simpler, analytical models are needed, and efficient techniques for extracting circuit models from physically detailed models must be devised.

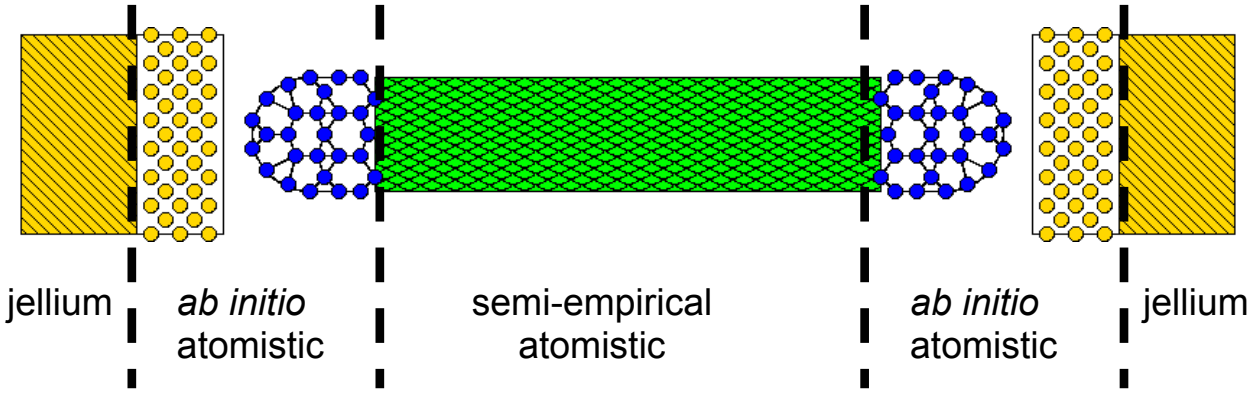


Fig.3.1 An illustration of how continuum, *ab initio*, atomistic and semi-empirical atomistic models will be combined in a multi-scale description of a carbon nanotube electronic device.

Our purpose in this paper is to describe the status of our work to develop a comprehensive, multi-scale simulation capability for electronic devices. We will focus on our initial effort that make use of a semi-empirical, p_z orbital description, and discuss briefly the challenges to be addressed in connecting this work to *ab initio* simulations, to continuum device simulations, and to circuit models. The approach has already demonstrated its usefulness in analyzing recent experimental data, suggesting experiments, and in exploring device possibilities [9].

3.2. Review of the NEGF Formalism

A carbon nanotube can be viewed as a rolled-up sheet of graphene with a diameter typically between one and two nanometers. The nanotube can be either metallic or semiconducting, depending on how it is rolled up from the graphene sheet (i.e. depending on its chirality) [15]. Semiconducting nanotubes are suitable for transistors. In order to correctly treat carbon nanotube transistors, strong quantum confinement around the tube

circumferential direction, quantum tunneling through Schottky barriers at the metal/nanotube contacts, and quantum tunneling and reflection at barriers in nanotube channel need to be considered. The non-equilibrium Green's function (NEGF) formalism, which solves Schrödinger equation under non-equilibrium conditions and can treat coupling to contacts and dissipative scattering process, provides a sound basis for quantum device simulations [29]. The NEGF simulation approach has demonstrated its usefulness for simulating nanoscale transistors from conventional Si MOSFETs [30], MOSFETs with novel channel materials [31], to CNTFETs [9, 32], and molecular transistors [33]. In this section, we give brief summary of the NEGF simulation procedure. For a more thorough description of the technique, see [34].

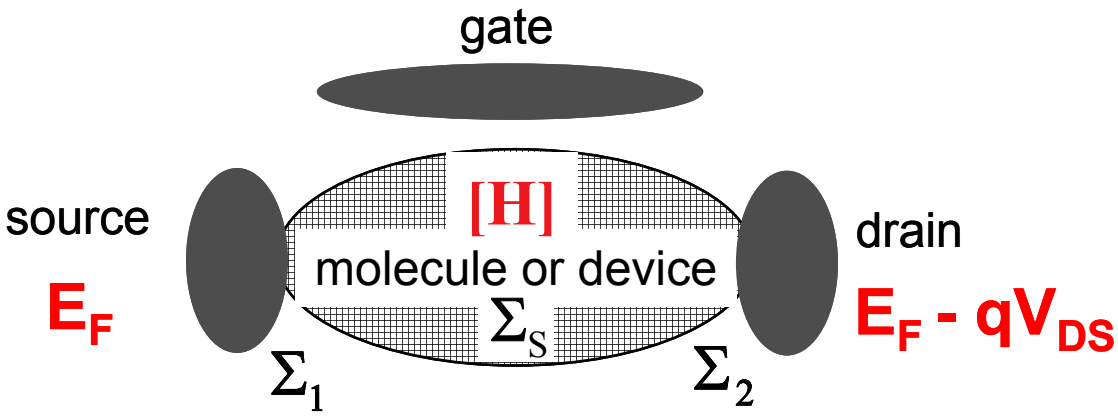


Fig.3.2 The generic transistor with a molecule or device channel connected to the source and drain contacts. The source-drain current is modulated by a third electrode, the gate. The quantities in the NEGF calculation are also shown.

Figure 3.2 shows a generic transistor and defines some terms for the NEGF simulation. The first step is to identify a suitable basis set and Hamiltonian matrix for an isolated channel. The self-consistent potential, which is a part of the Hamiltonian matrix, is included in this step. The second step is to compute the self-energy matrices, Σ_1, Σ_2 and Σ_S, which describe how the ballistic channel couples to the source/drain contacts and to the scattering process. (For simplicity, only ballistic transport is treated in this paper.) After identifying the Hamiltonian matrix and the self-energies, the third step is to compute the retarded Green's function,

$$\mathbf{G}(E) = [(E + i0^+)\mathbf{I} - \mathbf{H} - \mathbf{\Sigma}_1 - \mathbf{\Sigma}_2]^{-1}. \tag{3.1}$$

The fourth step is to determine the physical quantities of interest from the Green's function.

In the ballistic limit, states within the device can be divided into two parts: 1) states filled by carriers from the source according to the source Fermi level, and 2) states filled by the drain according to the drain Fermi level. Within the device, the source (drain) local-density-of-states (LDOS) is $\mathbf{D}_{S(D)} = \mathbf{G}\mathbf{\Gamma}_{S(D)}\mathbf{G}^+$, where $\mathbf{\Gamma}_{S(D)} = i(\mathbf{\Sigma}_{1(2)} - \mathbf{\Sigma}^+_{1(2)})$ is the energy level broadening due to the source (drain) contact. The charge density within the device is computed by integrating the LDOS, weighted by the appropriate Fermi level) over energy. The charge contributed by the source contact is

$$Q_S(z) = (-e)\int_{E_N}^{+\infty} D_S(E,z) f\left(E - E_{FS}\right) dE + e\int_{-\infty}^{E_N} D_S(E,z)\left\{1 - f\left(E - E_{FS}\right)\right\} dE$$

where e is the electronic charge, and E_N is the charge neutrality level [35, 36]. The total charge is

$$Q(z) = Q_S(z) + Q_D(z) = (-e)\int_{-\infty}^{+\infty} dE \cdot \mathrm{sgn}[E - E_N(z)]\left\{D_S(E,z) f\left(\mathrm{sgn}[E - E_N(z)](E - E_{FS})\right)\right.$$

$$+D_D(E,z)f\left(\text{sgn}[E-E_N(z)](E-E_{FD})\right)\Big\}, \qquad (3.2)$$

where sgn(E) is the sign function, and $E_{FS,D}$ is the source (drain) Fermi level. For a self-consistent solution, the NEGF transport equation is solved with iteratively the Poisson equation until self-consistency is achieved after which the source-drain current is computed from

$$I = \frac{4e}{h}\int T(E)[f_S(E) - f_D(E)]dE \qquad (3.3)$$

where $T(E) = \text{Trace}(_1\text{G}_2\text{G}^+)$ is the source/drain transmission and the extra factor of two comes from the valley degeneracy in the carbon nanotube energy band structure.

The computationally expensive part of the NEGF simulation is finding the retarded Green's function, according to eqn. (3.1), which requires the inversion of a matrix for each energy grid point. The straightforward way is to explicitly invert the matrix, whose size is the size of the basis set. This, however, is impractical for an atomistic simulation of a nanotube transistor. In the ballistic limit, the problem is simplified because only a few columns of the Greens's function are needed. Still, reducing the size of the Hamiltonian matrix and developing computationally efficient approaches are of great importance for an atomistic simulation.

3.3. Atomistic NEGF Treatment of Electron Transport in Carbon Nanotubes

3.3.1 Real space approach

In this section, we describe an NEGF simulation of ballistic CNTFETs using a real space basis. The first step is to identify a set of atomistic orbitals adequate to describe the

essential physics for carrier transport and then to write down the Hamiltonian matrix for the isolated channel in that basis. An (n, 0) zigzag nanotube as shown in Fig. 3.3 is assumed, but the method can be readily extended to armchair or chiral nanotubes. There are four orbitals in the outer electron shell of a carbon atom (s, p_x, p_y, and p_z). One p_z orbital is often sufficient because the bands involving p_z orbitals are largely uncoupled from the bands involving the other orbitals, and the bands due to the s, p_x and p_y orbitals are either well below or well above the Fermi level and, therefore, unimportant for carrier transport. With one p_z orbital per carbon atom as the basis set, the size of the Hamiltonian matrix is the number of carbon atoms in the transistor channel. For typical problems, such as the examples in 3.5, a carbon nanotube transistor will consist of several thousand carbon atoms. We use a tight-binding approximation to describe the interaction between carbon atoms, and only nearest neighbor coupling is considered. A p_z-orbital coupling parameter of $t = 3\text{eV}$ was assumed.

Figure 3.3 shows that a zigzag nanotube is composed of rings of carbon atoms in the A- and B-atom sublattices. Each ring in the A-atom sublattice is adjacent in the x-direction to a ring in the B-atom sublattice. There are *n* carbon atoms in each ring and a total of *N* atoms in the entire channel. The *N* x *N* Hamiltonian matrix for the whole nanotube channel is block tridiagonal,

$$\mathrm{H} = \begin{bmatrix} \alpha_1 & \beta_2^+ & & & & \\ \beta_2 & \alpha_2 & \beta_1 & & & \\ & \beta_1 & \alpha_3 & \beta_2 & & \\ & & \beta_2^+ & \alpha_4 & \beta_1 & \\ & & & \beta_1 & \alpha_5 & \ldots \\ & & & & \ldots & \ldots \end{bmatrix}, \tag{3.4}$$

where the *n* x *n* submatrix, $[\alpha_i]$, describes coupling within an A-type or B-type carbon ring, and the *n* x *n* $[\beta]$ matrices describe the coupling between adjacent rings.

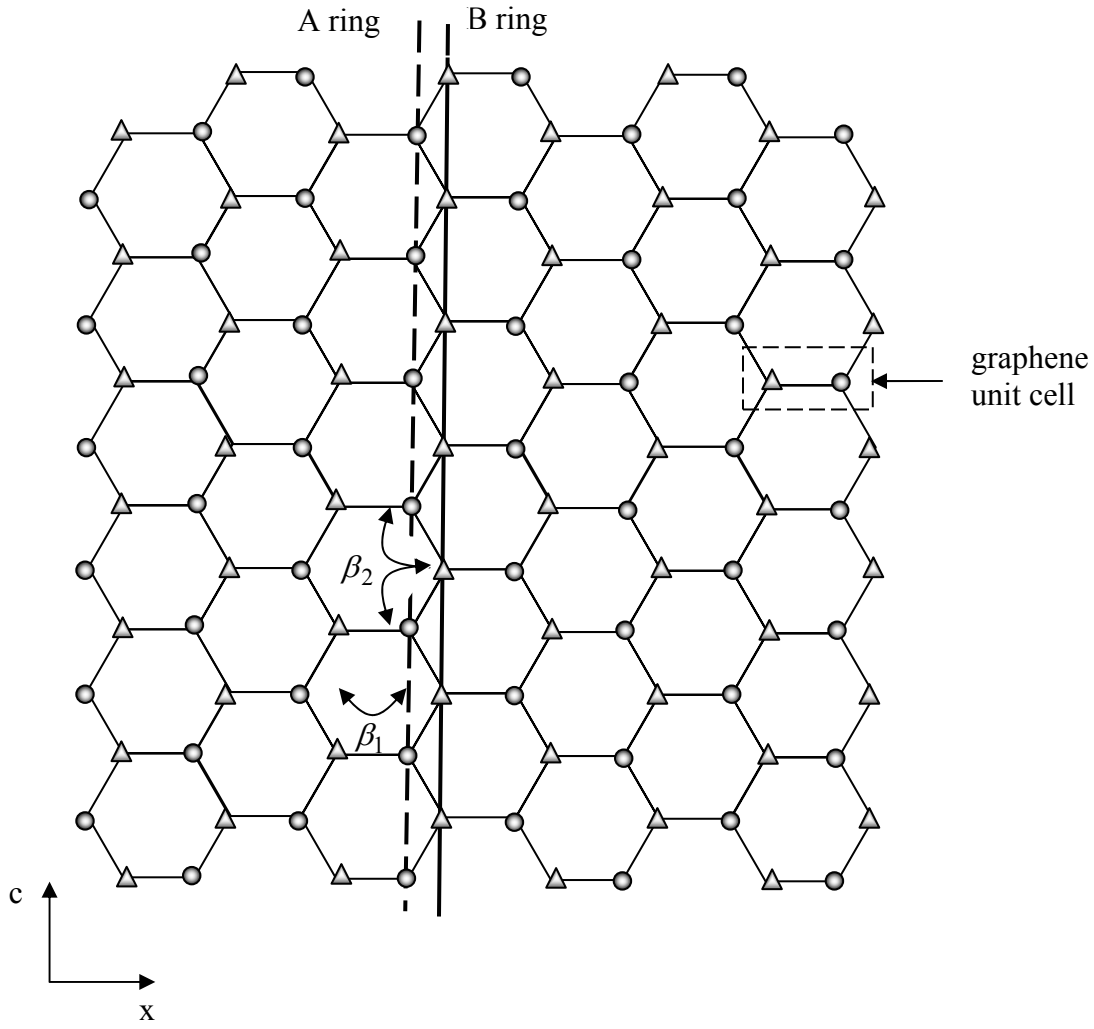


Fig.3.3 The schematic diagram of a (n, 0) zigzag nanotube (n = 6 in this case). The circles are the A-type carbon atom sublattice, and the triangles are the B-type carbon atom sublattice. The coordinate system is also shown: c is the circumferential direction, and x is the carrier transport direction.

In the nearest neighbor tight binding approximation, carbon atoms within a ring are uncoupled to each other so that $[\alpha_i]$ is a diagonal matrix. The value of a diagonal entry is the potential at that carbon atom site. If the nanotube is coaxially gated, the potential is invariant around the nanotube. The matrix, $[\alpha_i]$, therefore, is the potential at the *i*th carbon ring times the identity matrix, $[\alpha_i] = U_i[I]$.

There are two types of coupling matrices between nearest carbon rings, $[\beta_1]$ and $[\beta_2]$. As shown in Fig. 3.3, the first type, $[\beta_1]$, only couples an A(B) carbon atom to its B(A) counterpart in the neighboring ring. The coupling matrix is just the p_z orbital coupling parameter times an identity matrix,

$$[\beta_1] = t[I]\,. \tag{3.5}$$

The second type of coupling matrix, $[\beta_2]$, couples an A(B) atom to two B(A) neighbors in the adjacent ring. The coupling matrix is

$$[\beta_2] = t\begin{bmatrix} 1 & & \ldots & 1 \\ 1 & 1 & & \\ & 1 & 1 & \\ & & \ldots & \ldots \end{bmatrix}. \tag{3.6}$$

To understand eqn. (3.4), note that the odd numbered [α]'s refer to A-type rings and the even numbered one to B-type rings. Each A-type ring couples to the next B-type ring according to $[\beta_2]$ and to the previous B-type ring according to $[\beta_1]$. Each B-type ring couples to the next A-type ring according to $[\beta_1]$ and to the previous A-type ring according to $[\beta_2]$.

Having specified the Hamiltonian matrix for the channel, the next step is to compute the N x N self-energy matrices for the source and drain contacts, $[\Sigma_S]$ and $[\Sigma_D]$. The

self-energies describe the open boundary conditions for the Schrödinger equation. Only the carbon atoms on the first and last rings couple to the contacts, so $[\Sigma_S]$ and $[\Sigma_D]$ are sparse, with a structure of

$$[\Sigma_S] = \begin{bmatrix} \Sigma_{11} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, \tag{3.7}$$

where Σ_{11} is an n x n submatrix. Similarly, for $[\Sigma_D]$, the only nonzero block is the last diagonal submatrix. The derivation of these submatrices is described in detail in Appendix A.

The retarded Green's function,

$$\mathbf{G}^r = \left[(E + i0^+)\mathbf{I} - \mathbf{H} - \mathbf{\Sigma}_S - \mathbf{\Sigma}_D\right]^{-1}, \tag{3.8}$$

describes the bulk nanotube by **H** and the connection to the two contacts by the self-energy matrices. All matrices are size N x N with N being the total number of carbon atoms in the device. Solving eqn. (3.8) is equivalent to solving $\mathbf{AG}^r = \mathbf{I}$, where $\mathbf{A} = \left[(E + i0^+)\mathbf{I} - \mathbf{H} - \mathbf{\Sigma}_S - \mathbf{\Sigma}_D\right]$. The straightforward but computationally approach is to compute $\mathbf{G}^r$ is by directly inverting the **A** matrix. Significant computational savings can be achieved by exploiting the block tridiagonal structure of **A**, which allows $\mathbf{G}^r$ to be computed by a recursive algorithm without inverting a large matrix [37, 38]. If the channel consists of N_C carbon rings of a (n, 0) nanotube, the computational cost of directly inverting **A** goes as $O[(n \times N_c)^3]$ whereas with the recursive algorithm it is only $O(n^2 \times N_C)$. For the ballistic case, the solution is particularly efficient because only the first and last n columns of the Greens' function are needed.

Having computed the Green's function, the local density of states can be obtained, and the states can be filled according to the Fermi levels of the two contacts so that the charge density within the device can be found from eqn. (3.2). A method to compute the charge density from the Green's function using the recursive algorithm is also discussed in [37, 38]. By iterating between the NEGF equations to find the charge density and the Poisson equation to find the self-consistent potential, a self-consistent solution is obtained. The current is then evaluated from eqn. (3.3), where the current transmission probability, is obtained from the first diagonal block of the retarded Green's function,

$$\begin{aligned} T(E) &= \mathrm{Trace}\left({}_S\mathrm{G}^r \, {}_D\mathrm{G}^{r^+} \right) \\ &= \mathrm{Trace}\left(\Gamma_{S(1,1)} \left\{ i[G^r{}_{(1,1)} - G^r{}_{(1,1)}{}^+] - G^r{}_{(1,1)} \Gamma_{S(1,1)} G^r{}_{(1,1)}{}^+ \right\} \right) \end{aligned} \tag{3.9}$$

where $\mathbf{\Gamma}_{S,D} = i\left(\mathbf{\Sigma}_{S,D} - \mathbf{\Sigma}_{S,D}^{+}\right)/2$ is the source(drain) broadening and (1,1) denotes the first diagonal block of a matrix.

3.3.2 Mode space approach

The atomistic real space approach produces a matrix whose size is the total number of carbon atoms in the nanotube, which means that it is computationally intensive. A mode space approach significantly reduces the size of the Hamiltonian matrix. (A similar approach has been used for nanoscale MOSFETs [30]). In brief, the idea is to exploit the fact that in a carbon nanotube, periodic boundary conditions must be applied around the circumference of the nanotube, so $k_C C = 2\pi q$, where C is the circumference of the nanotube and q is an integer. Transport may be described in terms of these circumferential modes. If M modes contribute to transports, and if $M < n$, then the size of the problem is reduced from (n x N_C) unknowns to (M x N_C). If, in addition, the shape of the modes does not vary along the nanotube, then the M circumferential modes are uncoupled, and we can solve M one-dimensional problems of size, N_C, which is the

number of carbon rings along the nanotube. Mathematically, we perform a basis transformation on the (n, 0) zigzag nanotube to decouple the problem into n one-dimensional mode space lattices. The matrix is also tridiagonal, which allows the application of the efficient recursive algorithm for computing the Green's function [37].

When a zigzag nanotube is coaxially gated, the modes around the tube are simple plane waves with wave vectors satisfying the periodic boundary condition, and the mode space approach exactly reproduces the results of the real space approach. The mathematical details for obtaining the Hamiltonian matrix for a mode are provided in Appendix B. A pictorial view is shown in Fig. 3.4. After the basis transformation, the two dimensional nanotube lattice is transformed to n, uncoupled one-dimensional lattices in mode space. As shown in Appendix B, the Hamiltonian matrix for the qth mode is

$$\mathrm{H}_q = \begin{bmatrix} U_1 & b_{2q} & & \\ b_{2q} & U_2 & t & \\ & t & U_3 & b_{2q} \\ & & & \cdots \end{bmatrix}, \tag{3.10}$$

where U_i is the electrostatic potential at the ith carbon ring, t is the C-C nearest neighbor binding parameter, and $b_{2q} = 2t\cos(\pi q/n)$. Equation (3.10) should be compared with eqn. (3.4). In eqn. (3.10), each element is a number, not an n x n submatrix as in eqn, (3.4). As in eqn. (3.4), the odd-numbered diagonal entries refer to the A-type submatrix and even numbered ones to the B-type submatrices. Each A-type ring couples to the next B-type ring with the parameter, b_{2q} (analogous to β_2 in eqn. (3.4)) and to the previous B-type ring with the parameter, t (analogous to β_1 in eqn. (3.4)). Similarly, each B-type ring couple to the next A-type ring with parameter, t, and to the previous B-type ring with parameter, b_{2q}.

For an (n, 0) nanotube, there are $M = n$ circumferential modes, but the computational cost is reduced when the modes are uncoupled. The computational cost can be further reduced by noticing that typically only one or a few modes are relevant to carrier

transport. Modes with their band edges well above or below the source and drain Fermi levels are unimportant to carrier transports. The $E(k)$ relation for the qth mode as computed from eqn. (3.10) is $E(k) = \pm\sqrt{t^2 + b_{2q}^2 + 2tb_{2q}\cos\left(3ka_{CC}/2\right)}$, where $a_{cc} \approx 1.42$ Å is the C-C bonding distance. The qth mode produces a conduction band and a valence band with symmetric $E(k)$, and a band gap of $E_g = 2|t|\{1+\cos(\pi q/n)\}$. When $n \bmod 3 = 0$, the lowest subband index is $q = 2n/3$, which results in $b_{2q} = -t$ and a zero band gap. Otherwise, the nanotube is semiconducting and the lowest subband index is the integer closest to $2n/3$. By retaining only those modes whose carrier population changes with device bias or operating temperature, the size of the problem is significantly reduced.

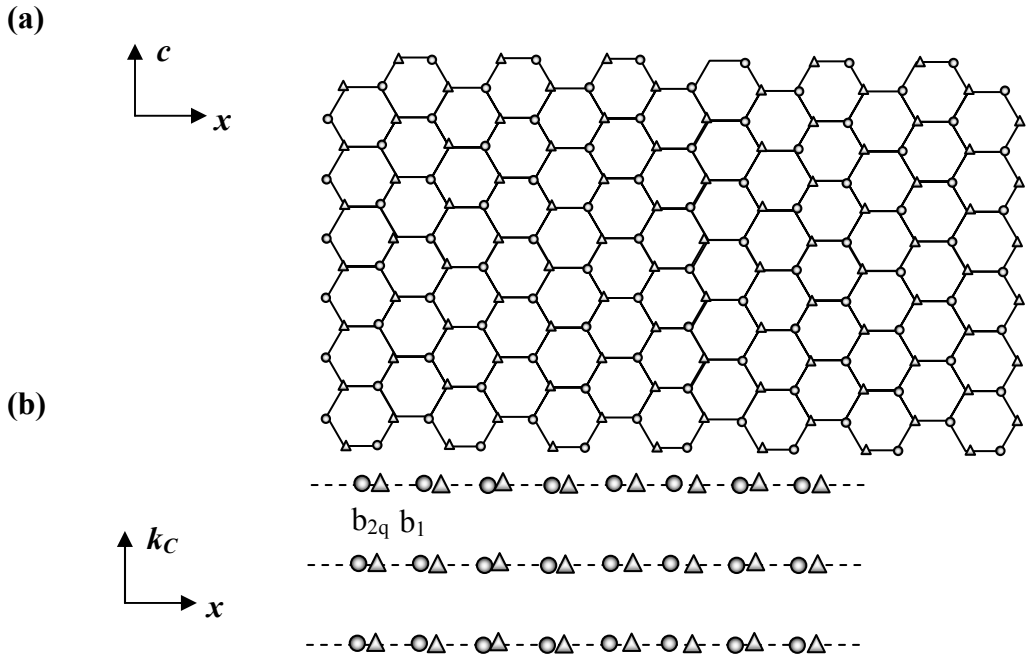


Fig.3. 4(a) The real space 2D lattice of the (n, 0) zigzag nanotube (b) The uncoupled, 1D mode space lattices. A basis transformation on the real space lattice of (a) transforms the problem to the M one-dimensional problems, where M labels a specific k_C.

The mode space source and drain self-energies can be computed using the same recursive relation for the surface Green's functions already discussed in Appendix A. The details are provided in Appendix B. The structure of the self energy matrices is the same as in eqn. (3.7) except that Σ_{11} (and Σ_{NCNC} for the drain self energy) are numbers rather than n x n submatrices. After obtaining the Hamiltonian matrix and contact self energies, the retarded Green's function is computed. Because the Hamiltonian matrix for a mode is tridiagonal and only a small part of the retarded Green's function is needed for the purpose of computing charge density and current at the ballistic limit, the recursive algorithm [37] or Gaussian elimination, rather than explicit matrix inversion, is used to compute the retarded Green's function.

3.4. Phenomenological Treatment of Metal/CNT junctions

In carbon nanotube transistors, the metal source and drain are typically attached directly to the intrinsic nanotube channel, and the gate modulates the source-drain current by changing the transmission through the Schottky barrier at the source end of the channel. To properly simulate such devices, the metal/CNT junction must be treated quantum mechanically. We currently treat this problem phenomenologically by defining an appropriate self-energy. Note that the self-energies defined in Sec. 3 do not apply here – they assume that carriers enter and leave the device without the need to tunnel through any barriers at the contact. As shown in Fig. 3.5, the phenomenological self energy must contain two parameters, one to describe the barrier height and another the density of metal-induced gap states (MIGS). Our approach mimics the effect of a real metal contact by specifying its work function and by injecting a continuous density of states near the Fermi level. This approach has proven useful in understanding transistor operations of Schottky barrier CNTFETs [9].

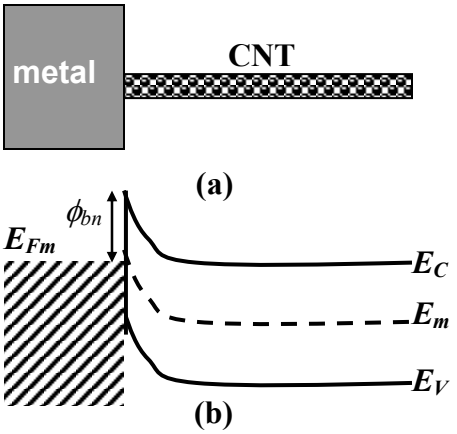


Fig.3.5 (a) The metal-carbon nanotube junction. (b) The band diagram of the junction. E_C, E_V and E_m are the conduction band edge, the valence band edge, and the middle gap energy in the nanotube, respectively. E_{Fm} is the metal Fermi level, and ϕ_{bn} is the Schottky barrier height for electrons.

The phenomenological treatment is described in Appendix C. In brief, each semiconducting mode in the semiconducting zigzag nanotube is coupled at the M/CNT interface to a mode of a metallic zigzag CNT. As shown in Appendix C, Σ_{11} in eqn. (3.7) becomes

$$\Sigma_{MS} = \alpha t^2 g_S = \alpha \frac{E - E_{m1} - \sqrt{(E - E_{m1})^2 - 4t^2}}{2}. \tag{3.11}$$

The coupling is described by two parameters. The first parameter is ϕ_{bo}, the Schottky barrier height for electrons without the presence of the interface states, which describes the band discontinuity at the interface and provides the value for E_{m1}, the mid-gap energy of the CNT at the interface. ($E_{m1} = E_{Fm} + \phi_{b0} - E_g/2$, where E_{Fm} is the metal Fermi level and E_g is the CNT band gap.) The second parameter is the tight-binding parameter, α, between the semiconducting and the metallic mode ($0 < \alpha \leq 1$),

which determines how well the metal contact is coupled to the nanotube channel, and is roughly proportional to the density of metal-induced-gap-states (MIGS). This simple model describes the interface at a level similar to those in the literature that the band discontinuity and density of interface states as input parameters [22].

3.5. The Overall Simulation Procedure

The overall simulation must be done self-consistently with Poisson's equation. Figure 3.6 shows the modeled, coaxial gate CNTFET, which provides the theoretically best gate control over the channel. The source and drain are heavily doped, semi-infinite carbon nanotubes, and the gate modulates the conductance of the channel, just like in a conventional Si MOSFET. For this device, we use the self energies described in Appendix A or Appendix B. By using a self-energy for metal/NT contacts as discussed in Appendix C, the simulation scheme can also be applied to Schottky barrier CNTFETs.

The transistor I-V characteristics strongly depend on the interplay of quantum transport and electrostatics, so we performed a self-consistent iteration between the NEGF transport equation and the Poisson equation as shown in Fig. 3.7. In brief, the procedure is as follows. For a given charge density, the Poisson equation is solved to obtain the electrostatic potential in the nanotube channel. Next, the computed potential profile is used as the input for the NEGF transport equation, and an improved estimate for the charge density is obtained. The iteration between the Poisson equation and the NEGF transport equation continues until self-consistency is achieved. Finally, the current for the self-consistent potential profile is computed.

For the coaxially gated carbon nanotube transistor, it is convenient to solve Poisson's equation in cylindrical coordinates. Since the potential and charge density are invariant around the nanotube, the Poisson equation is essentially a 2D problem along the tube (*x*-

direction) and the radial direction (*r*-direction) as shown in Fig. 3.6. Poisson's equation is written as

$$\nabla^2 E_m(r,z) = -\frac{e}{\varepsilon}\rho\,, \tag{3.12}$$

where E_m is defined as the vacuum energy level minus the work function of an intrinsic nanotube, and is exactly the middle gap energy for the grid points on the tube surface, and ρ is the charge density, which is non-zero only for grid points on the tube surface. The boundary condition applied at $r = 0$ is that the electric field along the *r*-direction is zero [39],

$$\varepsilon_r\,|_{r=0} = 0\,. \tag{3.13}$$

The potential at the gate electrode is known, so using the Fermi level of a grounded electrode as the zero energy, the electron potential at the gate electrode is,

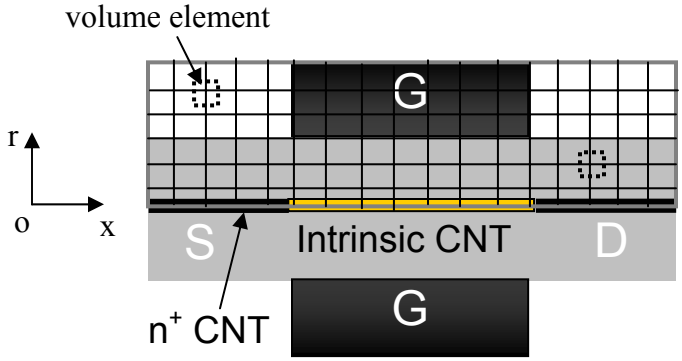


Fig.3.6 The modeled, coaxially gated carbon nanotube transistor with heavily-doped, semi-infinite nanotubes as the source/drain contacts. The channel is intrinsic and the gate length equals the channel length. Also shown are the simulated area, the simulation grid and the cylindrical coordinate system used for solving the Poisson equation. The dashed rectangular area shows the element used to discretize the Poisson equation at (x_i, r_j).

$$E_m(gate) = -eV_G + \phi_{ms}\,, \tag{3.14}$$

where V_G is the gate bias, and ϕ_{ms} is the work function difference between the gate metal and the intrinsic nanotube channel. By simulating a sufficiently large area, as shown in Fig. 3.6, Neumann boundary condition, which assumes that the electric field in the direction normal to the boundary is zero, can be applied to the remaining boundaries.

The continuous form of the Poisson equation, eqn. (3.12), is discretized for computer simulation. It is convenient to take a volume element near a grid point, as shown in Fig. 3.6, and apply the integral form of the Poisson equation to that volume element, which is a ring around the tube axis with a rectangular cross section,

$$\oint \vec{D} \cdot d\vec{S} = q_{ij}\,, \tag{3.15}$$

where q_{ij} is the charge in the total volume element, which is non-zero only on tube surface. The discretized equation for an element at the grid point (x_i, r_j) in air, is

$$\varepsilon_0 \left(\frac{r_{j-1}+r_j}{2}\Delta x \frac{E_m^{i,j-1}-E_m^{i,j}}{\Delta r} + \frac{r_{j+1}+r_j}{2}\Delta x \frac{E_m^{i,j+1}-E_m^{i,j}}{\Delta r} + r_j \Delta r \frac{E_m^{i+1,j}-E_m^{i,j}}{\Delta x} + r_j \Delta r \frac{E_m^{i-1,j}-E_m^{i,j}}{\Delta x} \right) = e r_j \Delta x (N_D - n_{net}) \,. \tag{3.16}$$

For grid points in the gate insulator, the gate insulator dielectric constant replaces ε_0 in eqn. (3.16). For the grid points at the gate insulator/air interface, the air dielectric constant is used for volume surfaces in air and the gate insulator dielectric constant is used for volume surfaces in the gate insulator.

Equation (3.16) is linear and mathematically easy to solve, but the convergence of the quantum transport and the linear Poisson equation is poor [26]. A non-linear Poisson equation, which relates the charge density to the potential through a non-linear dummy function, has been proven to be very useful in improving the convergence. The non-linear dummy function relating the charge density and the potential should be as close to the physical relation determined by carrier transport equation as possible for better convergence. Typically, semiclassical, equilibrium carrier statistics with a dummy quasi Fermi level are used as the dummy function. The non-linear Poisson equation takes the charge density computed by the transport equation as the input, and converts the charge density to a quasi Fermi level using the dummy function. Then the non-linear Poisson equation is solved for the potential by Newton-Ralphson iteration. Details of the non-linear Poisson solver can be found in [26].

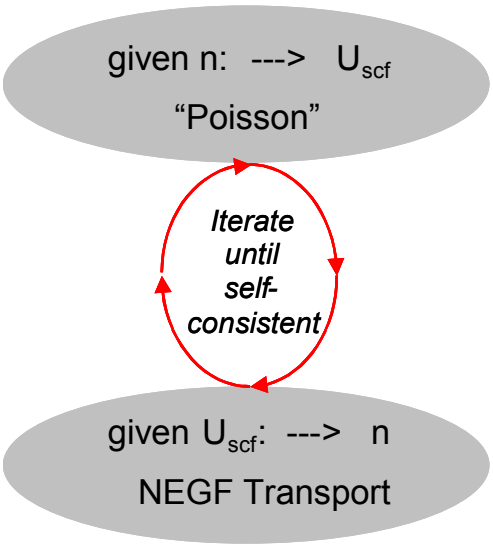


Fig.3.7 The self-consistent iteration between the NEGF transport and the electrostatic Poisson equation. These two equations are iteratively solved until self-consistency is achieved. Then the current is computed using the self-consistent potential.

3.6. Results

The simulation methods discussed in the previous sections have proven useful in several recent transistor studies [9]. The purpose of this section is to show some simple examples to demonstrate that: i) quantum effects are captured, ii) the mode space approach is valid when potential is uniform around the tube, and iii) the metal/CNT junction can be treated by our phenomenological self energy.

We first simulate a coaxially gated, MOSFET-like CNTFET as shown in Fig. 3.6. The transistor channel is a (25,0) intrinsic CNT, which results in a band gap of ~0.42eV and a diameter of ~2nm. The nanotube length is ~50nm, consisting of $\sim 1.2 \times 10^4$ carbon atoms. A self-consistent Poisson-NEGF simulation in the real space (using the recursive algorithm for computer the Green's function) is performed. Fig. 3.8a shows the energy-resolved local-density-of-states (LDOS), and the energy band profile. The band gap region with extremely low LDOS (darker in the grayscale plot) can be clearly identified. Due to the existence of the barriers, the source/drain incident wave is reflected and the quantum interference pattern between the incident and reflected waves is apparent. A quantum well is formed in the valence band of the channel, and the 1^{st} and 2^{nd} confined states with one or two LDOS maxima, respectively, can be clearly seen. The band edge of the second subband is also observed. Figure 3.8b shows the energy resolved electron density (electron density spectrum), which is obtained by filling the LDOS with the source or drain Fermi level. The bandgap, quantum interference, quantum confinement, and the second subband can still be clearly seen.

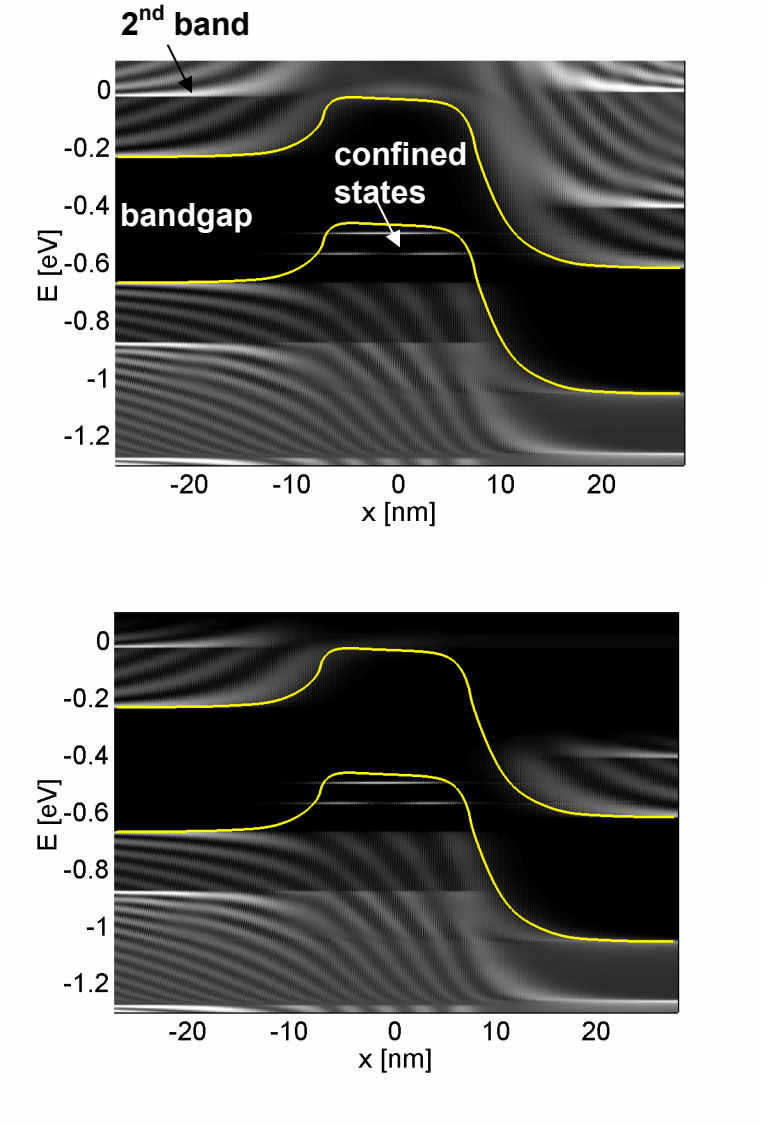


Fig.3.8(a) The local-density-of-states (LDOS) and (b) the electron density spectrum computed by the real space approach at V_G=0.25V and V_D=0.4V. The modeled transistor is shown in Fig. 7a. The nanotube is a (25,0) CNT with a diameter d~2nm and bandgap Eg~0.4eV.

Next, we explore the validity of the mode space approach by comparing the results of the real space approach to those of the mode space approach. The mode space approach theoretically should exactly reproduce the results of the real space approach when the potential is invariant around the tube, and a sufficient number of modes is included in the mode space simulation. A CNTFET as shown in Fig. 3.6 with a (13,0) nanotube channel, which results in a band gap of ~0.83eV and a diameter of ~1nm, is simulated. The carbon nanotube length is ~50nm, consisting of ~6000 carbon atoms. Because the third subband is ~1eV away from the lowest subband and the applied bias is ≤ 0.4 V, only the lowest two subbands are treated in the mode space simulation. The Hamiltonian matrix for the lowest subband is small ($\sim 500 \times 500$), and computing the Green's function for a subband using the recursive algorithm is fast even on a single CPU PC. Figure 3.9, which compares the I-V characteristics of the real space and mode space approaches, shows that the mode space simulation excellently reproduces the results of the real space approach. Figure 3.10, which plots the band profile and the charge density at on-state, again shows that the mode space approach excellently reproduces the results from the real space approach results. The good agreement between the real and mode space approach results from the equal potential around the tube direction when it's coaxially gated. The mode space is highly advantageous in reducing the computational burden, and it is valid when the potential variation around the tube is much smaller than the spacing between the subbands.

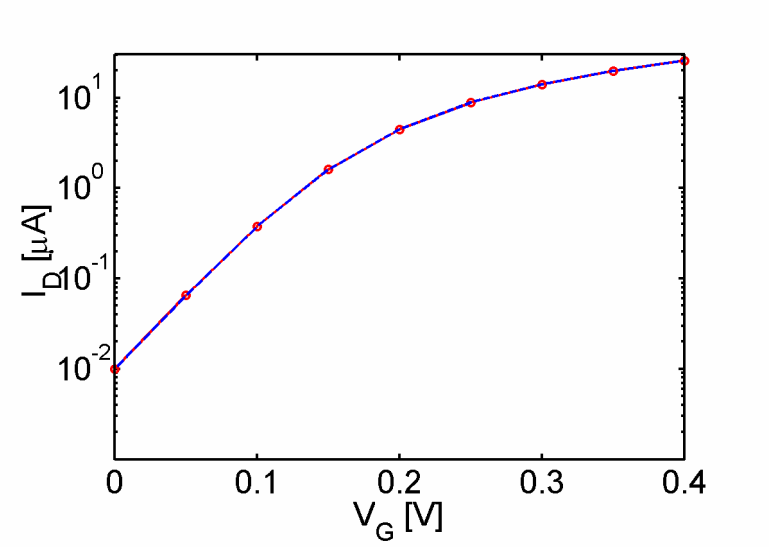


Fig.3.9 The I-V characteristics computed by the real space approach (the solid line) and the mode space approach with 2 subbands (the circles) for a CNTFET as shown in Fig. 7. The (13,0) nanotube channel length is 15nm.

Finally, we treat an SBFET-like CNTFET by self-consistent, quantum simulation. Fig. 3.11a shows the simulated transistor structure. The metal source/drain is directly attached to a (13,0) intrinsic nanotube channel, so a Schottky barriers forms between the source/drain and channel. A mid-gap Schottky barrier, with equal barrier height for electrons and holes, is simulated. Fig. 13.1b shows the local density of states at $V_D = V_G = 0.4\,\mathrm{V}$. The metal-induced gap states (MIGS) near the metal/CNT interfaces are apparent and decay rapidly with a tail of a few nanometers inside the channel. The tunneling states under the Schottky barrier in the conduction band at the source end of the channel are clear. The metal-nanotube interface is not perfectly transmitting, and the weakly confined states with the increasing number of LDOS maxima, due to the weak localization created by double metal/CNT barriers at the source and drain ends of the channel, can be seen. The atomistic-scale oscillations of the charge density spectrum along the channel direction is probably due to the charge transfer between A and B types of carbon rings in a zigzag carbon nanotube [40].

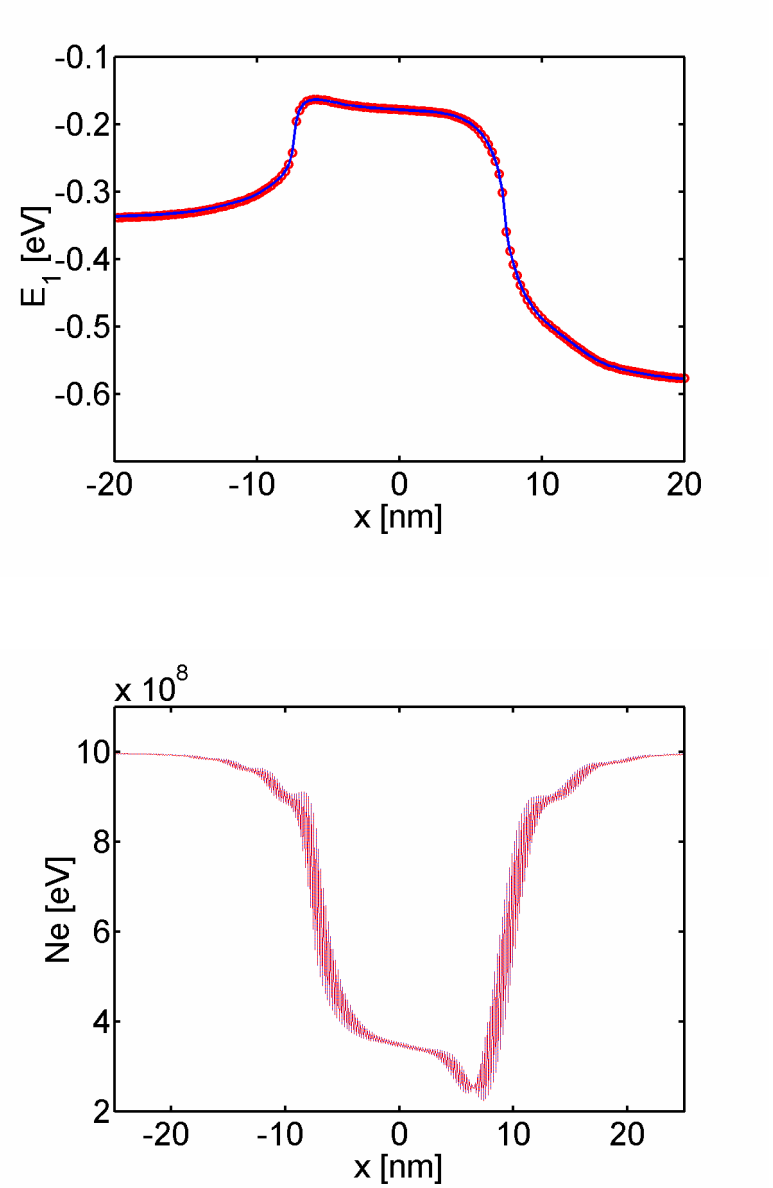


Fig.3.10(a) The conduction band profile computed by the real space approach (the solid lines) and the mode space approach (the circles) at $V_G = V_D = 0.4$ V. (b) The charge density computed by the real space approach (the solid line) and the mode space approach (the dashed line). The solid and dashed lines lie on top of each other.

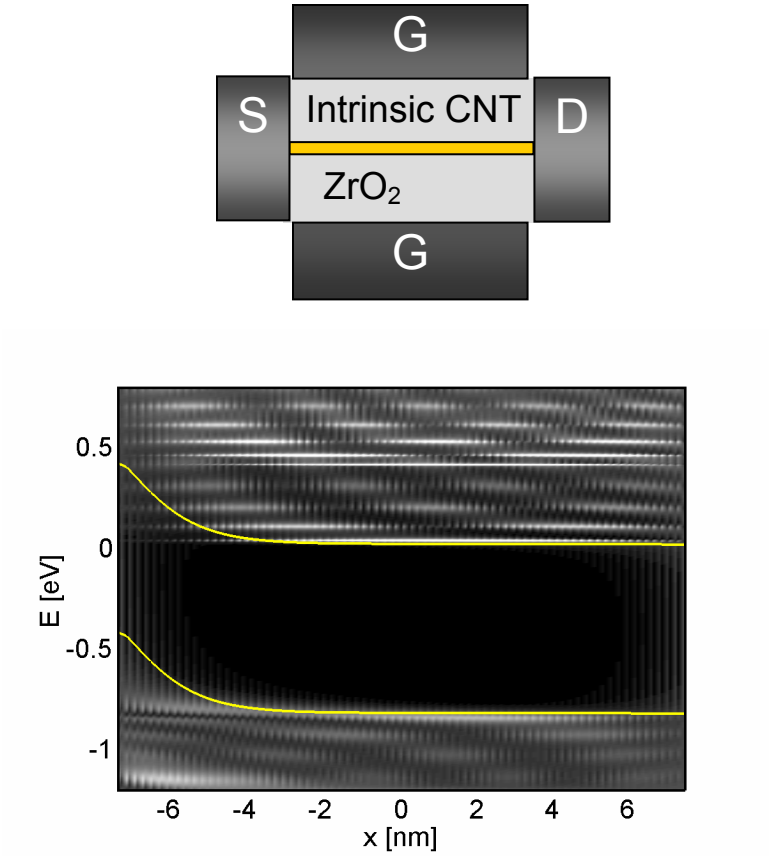


Fig.3.11(a) The coaxially gated Schottky barrier carbon nanotube transistor with an intrinsic nanotube channel directly attached to metal source and drain contacts. The nanotube channel is a (13,0) zigzag CNT with a diameter d~1nm and band gap Eg~0.83eV. The gate insulator is a 2nm-thick ZrO_2. (b) The local-density-of-states (LDOS) at $V_D = V_G = 0.4$ V, which clearly shows tunneling through the Schottky barrier at the source end of the channel, and metal induced gap states (MIGS) at the metal/CNT interfaces.

3.7. Discussion

The semi-empirical approach described in this paper is only one part of a multi-scale hierarchy shown in Fig. 3.1. More rigorous, *ab initio* methods are needed to treat the metal/CNT interface properly. Such simulations would allow first principles calculations of the barrier height and the MIGS, two parameters that we now treat as phenomenological. Such simulations may also provide useful insights into how to produce ohmic, rather than Schottky barrier, contacts when desired. The phenomenological model, however, is well-suited for device-scale simulations because of its computational efficiency. One approach to this multi-scale challenge is to relate the phenomenological parameters for the metal/nanotube contacts in the semi-empirical approach described in this paper to detailed atomistic simulations of the contact. In such an approach, the semi-empirical model would stand alone and be related to separate, *ab initio* simulations. Another possibility is the domain decomposition approach sketch in Fig. 3.1. In this approach, the key challenge is to connect the two regions, described with much different sets of basis functions, through the self-energies. This "mixed basis set approach" is already being applied to problems involving molecules on silicon contacts and is being investigate for the metal/nanotube contact as well.

The approach described in this paper assumes ballistic transport, but scattering by phonon emission is likely to be a factor in devices under realistic operating voltages [41-44]. There is a clear prescription for treating the electron-phonon interaction within the NEGF formalism [29, 34], but the computational burden increases rapidly. It is likely, therefore, that semiclassical, continuum approaches like those used to treat semiconductor devices by Monte Carlo simulation [45] will be needed when a detailed treatment of the electron-phonon coupling is necessary. NEGF simulation is the method of choice when quantum transport is the dominant factor, and phenomenological treatments of scattering can be used [46, 47]. Semiclassical approaches are the method of

choice when scattering dominates, and phenomenological quantum corrections can be made.

Finally, work at the device level needs to be coupled to circuit level models so that the system level implications of novel devices can be readily explored. Existing approaches may or may not be adequate. CNTFETs, for example, should operate near the ballistic limit, and it is not clear that traditional MOSFET models, which were developed for the scattering-dominated regime, can be extended to quasi-ballistic transistors. Recently, a new circuit model for ballistic CNTFETs has been developed [48]. The more general question of how circuit models for new exploratory devices can be rapidly developed is an important one to address.

3.8. Conclusions

Methods for the NEGF/Poisson simulation of carbon nanotube transistors were discussed and illustrated. The real space approach, which uses one p_z orbital per carbon atom as the basis, achieves atomistic resolution for quantities of interest. Significant computational saving can be achieved by using the mode space approach, which performs a basis transformation around the nanotube circumferential direction and transforms the 2D nanotube lattice to decoupled 1D mode space lattices. Each mode in the mode space approach describes one conduction subband and its corresponding valence subband, and atomistic resolution along the transport direction is retained. The simulation methods discussed in this paper have been applied to several transistor studies with the purpose of understanding experiments and exploring device physics [9]. Finally, the need to complement this semi-empirical device level model with higher level circuit models and lower level *ab initio* models was discussed.

4 A NUMERICAL STUDY OF SCALING ISSUES FOR SCHOTTKY BARRIER CARBON NANOTUBE TRANSISTORS

4.1 Introduction

Carbon nanotube field-effect transistors (CNTFETs) with promising device characteristics have recently been demonstrated [13, 14, 49, 50], so the question of the ultimate device performance capability and minimum device size that might be achievable from an optimized technology becomes important. A recent scaling study of Schottky barrier (SB) CNTFETs by Heinze et al. [51] examined the role of scaling the gate oxide thickness down and the dielectric constant up. That study found that the device performance depends in an unexpected way (as compared to a silicon MOSFET) on the gate oxide thickness and dielectric constant. In a very recent study, the issue of drain voltage scaling has been considered [10]. In this paper, we extend previous work by using a coupled Poisson-quantum transport model to treat the charge in the nanotube self-consistently. A comprehensive study of CNTFET scaling issues is performed to examine the role of gate insulator thickness and dielectric constant, nanotube diameter, Schottky barrier height, drain voltage, and channel length. In contrast to SB CNTFETs with thick gate oxides, SB CNTFETs with thin gate oxides show very strong ambipolar I-V characteristics, even if the barrier heights for electrons and holes are highly asymmetric. The ultimate scaling limit for the channel length imposed by source-drain tunneling is established. The nanotube diameter and drain voltage are shown to have a strong influence on the leakage current. In contrast to a previous study [51] that examined the subthreshold and near threshold regions, we find that increasing the gate dielectric constant improves device performance (the on-current).

In this study, we restrict our attention to Schottky barrier CNTFETs, which operate by modulating the tunneling current at the source contact. (SB CNTFETs are common in experiments at this stage [7, 8].) Note, however, that there are recent reports that

CNTFETs without Schottky barriers, which operate more like MOSFETs, can be realized [13]. We assume ballistic transport and solve the Schrödinger equation self-consistently with the Poisson equation. Because our interest is ultimate limits, we assume a coaxial geometry, rather than the planar geometry of the actual devices that have been reported. The coaxial geometry provides the best electrostatic control by the gate and, therefore, the minimum channel length for electrostatic consideration [39, 52]. A zigzag nanotube is assumed, and an atomistic description in terms of p_z orbitals is used. Our assumption of ballistic transport is a reasonable one that facilitates simulations [13, 53, 54], but it is not central to this paper. We focus here on understand the device physics of ambipolar conduction in SB CNTFETs and do not expect the conclusions to depend on whether transport is ballistic or not.

4.2 Approach

To investigate the performance of aggressively scaled CNTFETs, we simulated a coaxially gated CNTFET with a 15nm ballistic channel, as shown in Fig. 4.1 at room temperature (T=300K). The nominal device has a 2nm ZrO_2 gate oxide (a high-K gate insulator of this type has been experimentally demonstrated [55]). The diameter of the (13, 0) nanotube is $d \approx 1\ nm$, which results in a bandgap of $E_g \approx 0.83\ eV$. A power supply voltage of 0.4V is assumed, according to the value specified for the 10nm scale MOSFET in ITRS roadmap [56]. The device parameters here are the nominal ones; we explore various issues by varying these parameters.

Carbon nanotube field-effect transistors were simulated by solving the Schrödinger equation using the non-equilibrium Green's function (NEGF) formalism [29, 34, 57] self-consistently with the Poisson equation. Ballistic transport was assumed. An atomistic description of the nanotube using a tight binding Hamiltonian with an atomistic (p_z orbital) basis was used. The atomistic treatment was computationally expensive, but significant computational savings were achieved by the mode space approach [30].

Because the carbon nanotube is coaxially gated, the eigenstates around the tube circumferential direction (modes) are plane waves with wave vectors satisfying the periodic boundary conditions. The two-dimensional nanotube lattice of a (n, 0) zigzag CNT was transformed to n decoupled one-dimensional modes by doing a basis transform from the real space to the mode space in the circumferential direction (essentially Fourier transform). Under typical bias conditions, the few modes that are relevant to electronic transport are treated.

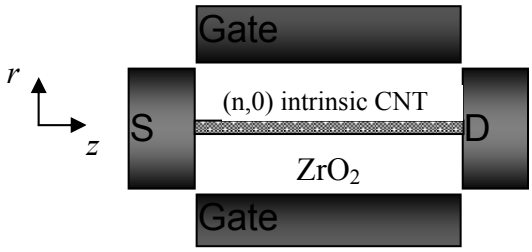


Fig.4.1 The modeled CNTFET with a coaxial gate. The gate length is the same as the source to drain spacing; the nominal value is 15nm. A 2nm-thick ZrO_2 gate insulator and a (13,0) zigzag nanotube (with the diameter *d~1nm* and the band gap E_g*~0.8eV*) are assumed.

The mode space approach reduces computation significantly yet retains atomistic resolution along the transport direction. For the *i*th mode, the charge density is computed by integrating the local density-of-states (LDOS) over energy,

$$Q_i(z) = (-e)\int_{-\infty}^{+\infty} dE \cdot \mathrm{sgn}[E - E_N(z)]\{D_{iS}(E,z) f\left(\mathrm{sgn}[E - E_N(z)](E - E_{FS})\right) + D_{iD}(E,z) f\left(\mathrm{sgn}[E - E_N(z)](E - E_{FD})\right)\}, \quad (4.1)$$

where e is the electron charge, $\mathrm{sgn}(E)$ is the sign function, $E_{FS,D}$ is the source (drain) Fermi level, and $D_{iS,D}(E,z)$ is the LDOS due to the source (drain) contact as computed by the NEGF method. Because the nanotube conduction and valence bands are symmetric, the charge neutrality level, $E_N(z)$, lies at the middle of band gap [35].

The Schottky barriers at the metal/CNT interfaces were treated phenomenologically. To mimic the continuous states injected from metal to the semiconducing nanotube modes, each semiconducting mode is coupled to the metallic mode of metallic zigzag CNTs at the M/CNT interface with the coupling described by two parameters. The first one is the band discontinuity at the interface, which is the Schottky barrier height when there are no interface states. The second parameter is the tight-binding parameter between the semiconducting and the metallic mode, which determines to the density of metal-induced-gap-states (MIGS). This simple model describes the interface at a similar level as the M/CNT models in literature with the band discontinuity and density of interface states treated as input parameters [22].

A 2D Poisson equation is solved to update the charge neutrality level in eqn (4.1), $E_N(z) = -e\phi(z, r = d/2)$, where d is the nanotube diameter and $\phi(z)$ is the electrostatic potential,

$$\nabla^2 \phi(z,r) = -\frac{\rho}{\varepsilon}. \tag{4.2}$$

The potentials at source/drain and gate electrodes are fixed as the boundary conditions, and the gate flat band voltage was assumed to be zero for simplicity. (In practice, it would depend on the gate workfunction.) In order to treat an arbitrary charge distribution on the nanotube channel, the Poisson equation (eqn. (2)) is solved by the method of moments [25]. The iteration between the atomistic quantum transport equation and the electrostatic equation continues until self-consistency is achieved, (a non-linear form of eqn (2) is used to improve the iteration convergence [26]), then the source-drain ballistic current is computed by

$$I = \frac{4e}{h} \int dE \cdot T(E)[f(E - E_{FS}) - f(E - E_{FD})], \tag{4.3}$$

where $T(E)$ is the source-drain transmission calculated by the NEGF formalism [29]. The gate leakage current is omitted in this study.

4.3 Results

We begin by simulating the nominal device and display the resulting log I_D vs. V_{GS} characteristic in Fig. 4.2a. For the CNTFET with the metal Fermi level at the middle of the bandgap, the transistor is ambipolar, showing symmetric electron and hole conduction (see the solid line in Fig. 4.2a). The minimum current occurs when the gate voltage is one-half the drain voltage at which the gate-to-source voltage equals the drain-to-gate voltage, and the conduction and valence band profiles are symmetric (see the solid line in Fig. 4.2b). Radosavljevic et al. observed similar behavior, and this bias is also the optimum bias for observing optical emission in CNTFETs [58, 59]. Several questions will be addressed in the remainder of this paper. What controls the minimum current, the on current, and the subthreshold swing? Can conduction of one type be suppressed so that SB CNTFETs can be used in conventional CMOS digital circuits? How does device performance depend on the nanotube diameter, power supply, gate insulator thickness and dielectric constant, and the channel length?

Figure 4.2 shows the effect of the metal/CNT barrier height on the I_D-V_{GS} characteristics for the nominal device with a thin (2nm) high-K (25) gate dielectric. Reducing the barrier height for electrons to zero increases the electron conduction current for $V_{GS} > V_{DS}/2$ and decreases the hole current for $V_{GS} < V_{DS}/2$. The $I_D - V_{GS}$ characteristic, however, remains approximately symmetrical; the dash-dot line in Fig. 4.2b explains why. Although the barrier height for holes is high when $\phi_{bn} = 0$ ($\phi_{bp} = E_g$) and barriers to hole conduction exist at both the source and drain electrodes, the barriers are thin. (The thickness of the Schottky barrier is approximately the thickness of the gate oxide [60]). The thin barriers are quite transparent at negative gate voltages. The observation that CNTFETs with thin gate oxide tend to be ambipolar with nearly symmetrical characteristics is consistent with recent experiments [14].

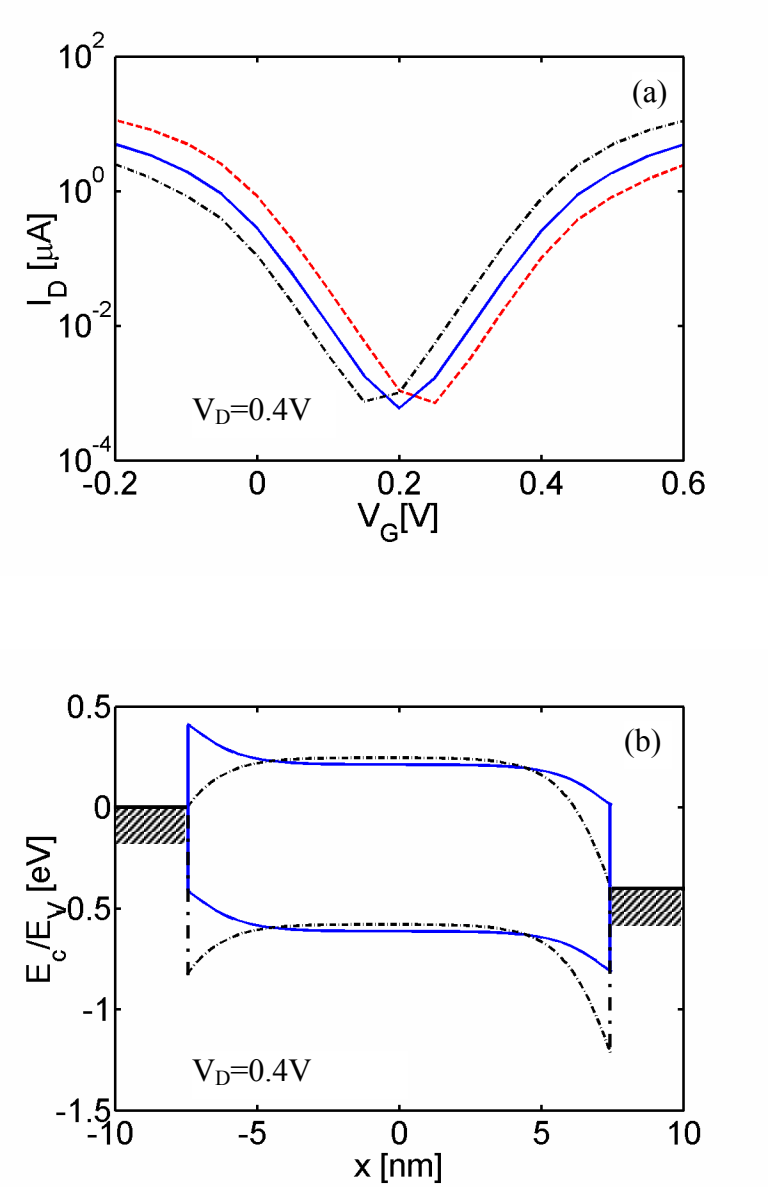


Fig.4.2 Transistor characteristics when the gate oxide is thin. (a) I_D vs. V_G characteristics for the nominal CNTFET (as shown in Fig. 4.1) with three different barrier heights. The Schottky barrier height for electrons is $\phi_{bn} = 0$ for the dash-dot line, $\phi_{bn} = E_g / 2$ for the solid line and $\phi_{bn} = E_g$ for the dashed line. (b) The conduction and valence band profile at $V_G = 0.2V$. The dash-dot line is for $\phi_{bn} = 0$ and the solid line is for $\phi_{bn} = E_g / 2$. The flat band voltage of all transistors is zero.

To further clarify the effect of Schottky barrier height on ambipolar conduction, we translated the I-V characteristics of CNTFETs in Fig. 4.2 along the x-axis so that the minimum current ($I \sim 6\times10^{-4}\,\mu A$) occurred at $V_G = 0V$. (Translating the I_D vs. V_G curve along x-axis in this way could be achieved in practice by adjusting the gate work function). The translated I-V characteristics are shown in Fig. 4.3. In the subthreshold region, the I-V characteristics of the zero barrier and the mid gap CNTFETs are nearly identical and the minimum leakage current is similar. It is interesting to note that the subthreshold swing is close to the ideal value of 60mV/dec for thermal injection over a barrier, no matter what Schottky barrier height is used. For $V_{GS} = V_{DD}$, the zero barrier height CNTFET delivers more on-current, and for $V_{GS} = -V_{DD}$, it deliver less (hole) on-current that for the mid-gap barrier. The general conclusion, however, is that the results are surprisingly symmetrical about the minimum current – no matter what the barrier height is.

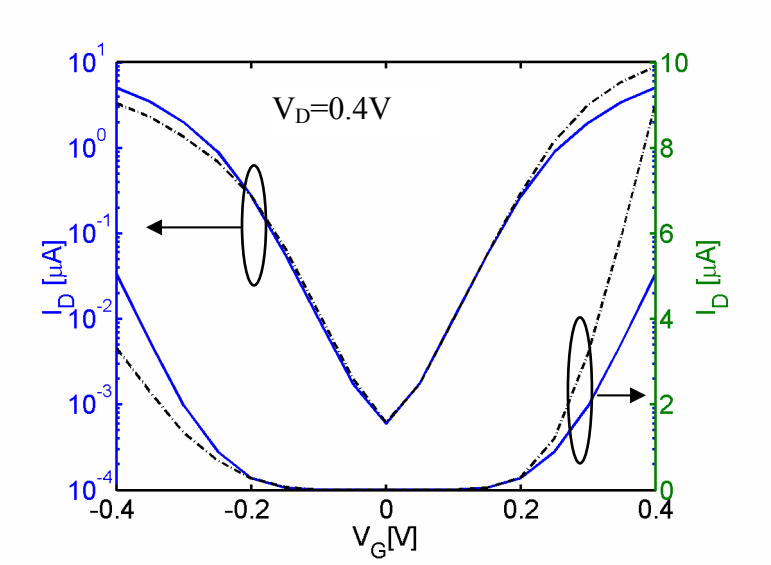


Fig.4.3 Shifted I_D vs. V_G characteristics for the nominal CNTFET (as shown in Fig. 4.1) with the barrier height for electrons $\phi_{bn} = 0$ (the solid-dash lines) and $\phi_{bn} = E_g / 2$ (the solid lines). The minimal leakage current is shifted to $V_G = 0$ by adjusting the flat band voltage for each transistor. The left axis shows the I-V on log scale and the right axis shows the same curves on linear scale.

The reason for the near-ideal subthreshold swing can be explained as follows. When the gate oxide is thin, the Schottky barrier is also thin and is essentially transparent to carriers. The current is, therefore, limited by the thermionic emission over a barrier with the height of the barrier determined by the conduction (valence) band in the interior of the channel. Tunneling through the M/S barrier varies with the barrier height and the bias, but it only plays a minor role (because the barrier is so transparent) compared to the barrier in the CNT body. Accordingly, the subthreshold swing is relatively independent of the barrier height, and the best that can be achieved is no better than what could be obtained in a MOSFET. Above the threshold, the situation is different because the barrier between the source and the CNT body is very small, so the tunneling resistance limits the on-current. In this case, the zero barrier contact delivers more on-current.

Things change when the gate insulator is thick. Figure 4.4 shows the I_D vs. V_G characteristics of a mid-gap SB CNTFET and a $\phi_{bn} = 0$ SB CNTFET with a 40nm-thick K = 25 gate oxide and 100nm channel length. In striking contrast to the thin oxide case, the I-V characteristics of these two CNTFETs are quite different. The minimum leakage current of the zero barrier CNTFET is smaller than for the thin oxide transistor, and the on-off current ratio is much better. The reason is that thicker gate insulators lead to thicker Schottky barriers so the tunneling resistance plays an important role. For the mid-gap CNTFET, the current is always limited by the Schottky barrier at the M/S contact, and the gate fringing field modulates the current by changing the tunneling barrier thickness. As a result, the subthreshold swing is ~200mV/dec – much larger than the theoretical minimum [7]. For the zero barrier height CNTFETs, however, current modulation is achieved by modulating a thermionic barrier inside the CNT body, a mechanism similar to the conventional MOSFET. As the result, the subthreshold swing is much smaller, at 90mV/dec. (This value is still larger than the ideal subthreshold swing of 60mV/dec because of short channel electrostatics and the parasitic capacitance between the source/drain contact and the channel that is large[31].) Because the SB is

thick, an asymmetric barrier height leads to quite asymmetric electron and hole conduction. For the zero barrier CNTFET, electron conduction is much better than hole conduction. Unfortunately, the thick oxide device displays a rather large subthreshold swing, and the on-current performance suffers from the tunneling barrier.

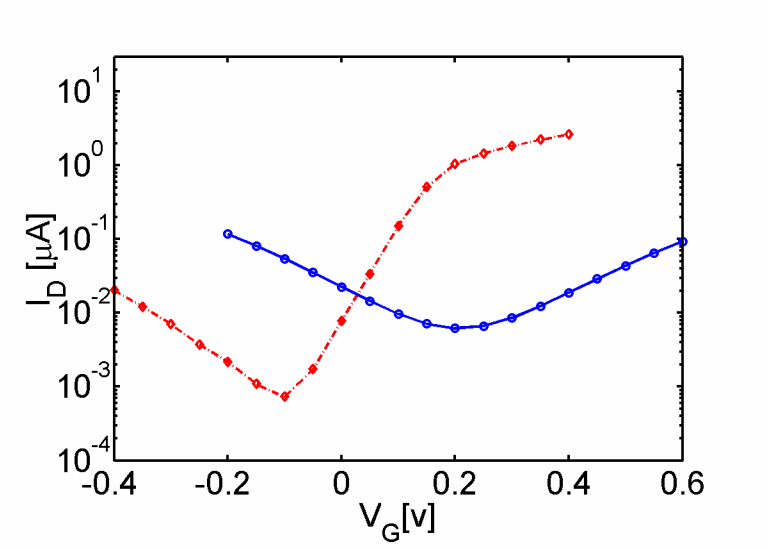


Fig.4.4 I_D vs. V_G for thick gate oxide (the oxide thickness $t_{ox} = 40nm$ and dielectric constant $\varepsilon = 25$). The channel length is 100nm. The Schottky barrier height for electrons is $\phi_{bn} = 0$ (the solid-dash lines) and $\phi_{bn} = E_g / 2$ (the solid lines). A (25,0) nanotube (with a diameter d~2.0nm and $E_g \sim 0.43eV$) is used as channel.

We turn next to the role of the nanotube diameter in determining the I-V characteristics. Figure 4.5 shows the I_D vs. V_G characteristics of the CNTFETs with three different nanotube diameters. We assume a mid-gap barrier height for all tubes, which corresponds to the same metal contact material if the work function of an intrinsic tube is independent of the tube diameter. Using a large diameter tube reduces the band gap and significantly increases the minimum leakage current at the ambipolar bias point. At the same time, the on-current is also improved, but the on-off ratio decreases significantly as the nanotube diameter increases. The small band gap of large diameter tube also leads to strong ambipolar conduction even if the gate oxide is thick and barrier heights for electrons and holes are asymmetric [61].

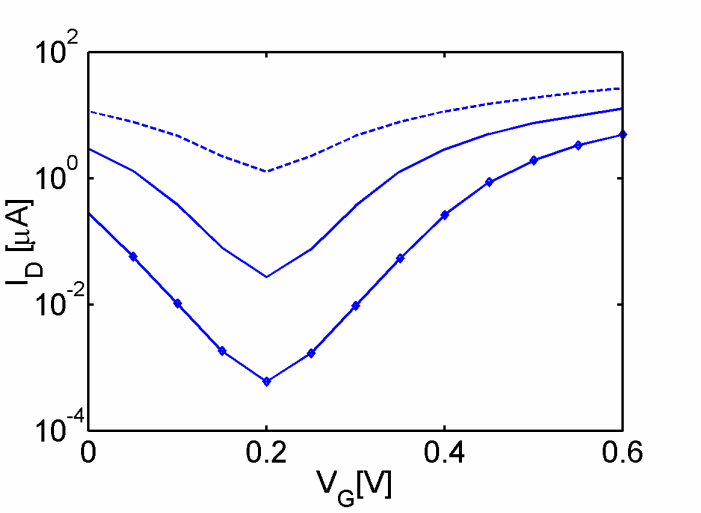


Fig.4.5 *Scaling of nanotube diameter*. I_D vs. V_G characteristics at V_D=0.4V for the nominal CNTFET with different nanotube diameter. The solid line with circles is for (13,0) CNT (with d~1nm), the sold line is for (17,0) CNT (with d~1.3nm), and the dashed line is for (25,0) CNT (with d~2nm). The flat band voltage is zero and the Fermi level lies in the middle of the band gap for all transistors.

We next examine power supply voltage scaling. Figure 4.6a shows the I_D vs. V_G characteristics of the nominal SB CNTFET with three different power supply voltages. Note that the minimum current increases exponentially with power supply voltage (as Radosavljevic et al. observed [10]). The reason is that the minimum leakage is achieved when the effective gate to source voltage is one half of the power supply voltage. Reducing the power supply voltage reduces the effective gate to source voltage at the minimum leakage point, thus exponentially reduces the minimal leakage current. Figure 4.6a also shows that the on-current increases with V_{DD}. The off-current vs. on-current for different power supply voltages is plotted in Fig. 4.6b. The trade-off for reducing the off-current by lowering the power supply voltage is the degradation of on-current. The choice of power supply voltage will depend on the type of circuit applications. For a low power design, the off-current must be small and the on-off current ratio needs to be large, which sets an upper limit for the power supply voltage. Although the low power supply voltage guarantees a small leakage current and large on-off ratio (because for small power supply voltage, the transistor operates in the exponential portion of the I_D vs V_G

characteristics), the on-current may still be too small for high-performance applications. To achieve a large on-current and a reasonable off-current simultaneously, which is required for high performance applications, the power supply voltage must be large enough.

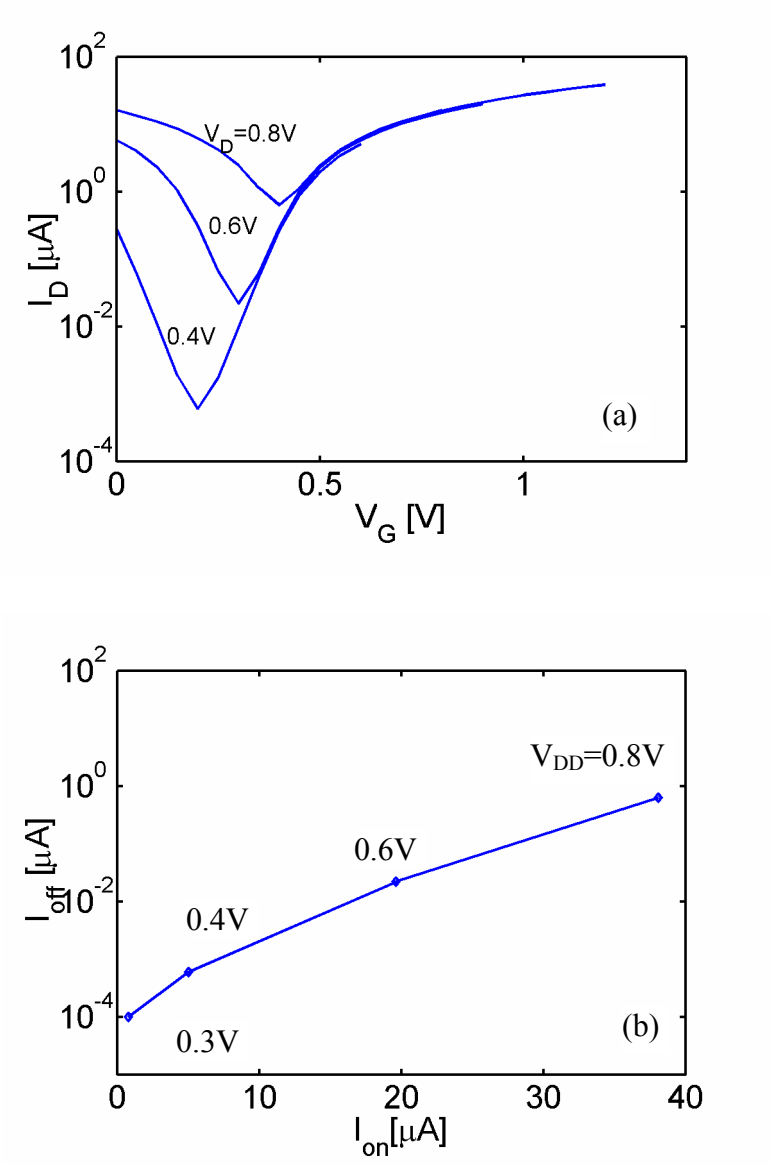


Fig.4.6 *Scaling of Power supply voltage.* (a) I_D vs V_G characteristics under different power supply voltages for the nominal CNTFET (Fig. 4.1) with mid-gap Schottky barriers. For each power supply voltage, the drain is biased at the power supply voltage, V_D=V_{DD}. The off-current is defined at the minimal leakage point ($V_{G,off}$=$V_D/2$), and the on-current is defined at $V_{G,on} = V_{G,off} + V_{DD}$. (b) The off-current vs. on-current for different power supply voltages.

Figure 4.7 explores the issue of channel length scaling. In order to establish the ultimate scaling limit imposed by source-drain tunneling, very thin gate oxide ($t_{ox} = 2nm$ for the channel length of 10nm and longer and $t_{ox} = 1nm$ for the channel length of 5nm) is used to ensure excellent gate controlled electrostatics. Although we assume a mid-gap Schottky barrier height, similar observations apply to other barrier heights. When the channel length is larger than about 15nm, the I-V characteristics are independent of the channel length because the channel is ballistic and the quantum tunneling resistance of the Schottky barrier at the source dominates the total channel resistance. Scaling the channel length down to 10nm significantly increases the off-current, but the on-off current ratio still exceeds 100, which is probably acceptable for digital logic. If the channel length is aggressively scaled down to 5nm, the on-off current ratio decreases to less than 10 due to significant source-drain quantum tunneling. Compared to Si MOSFETs with the similar channel lengths, the tunneling leakage of the CNTFET is more severe, partly due to typically smaller band gap and lighter effective mass for carriers in the CNT channel. (A parabolic E-k fit the very bottom of the conduction band of a 1nm diameter CNT gives an effective mass of ~0.08.)

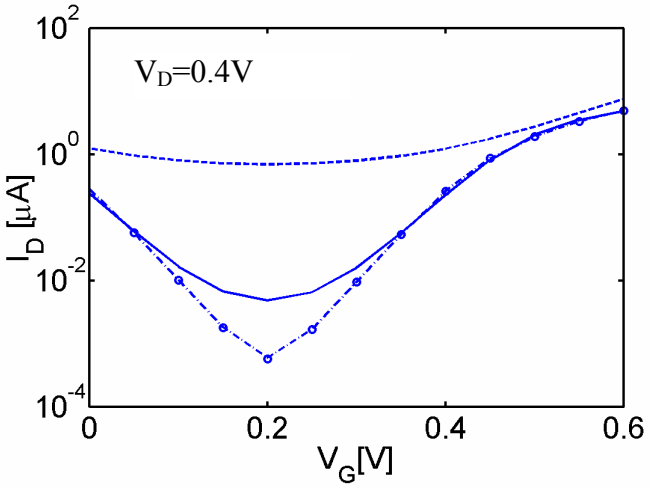


Fig.4.7 *Channel length scaling.* I_D vs. V_G characteristics of CNTFETs with different channel length. The circles are for channel length $L_{ch} = 30nm$ and gate ZrO_2 thickness $t_{ox} = 2nm$, the dash-dot line for $L_{ch} = 15nm$ and $t_{ox} = 2nm$, the solid line for $L_{ch} = 10nm$ and $t_{ox} = 2nm$, and the dashed line for $L_{ch} = 5nm$ and $t_{ox} = 1nm$. The flat band voltage is zero and mid-gap Schottky barriers are assumed for all transistors.

Figure 4.8 explores the role of the gate dielectric constant. A previous study, which omitted the charge on the nanotube, concluded that simply increasing the gate dielectric constant everywhere didn't change the band profile thus had no effect on the I-V characteristics of SB-CNTFETs [51]. Those conclusions apply below and near threshold. Figure 4.8, which shows the I_D vs. V_G characteristics for mid-gap SB-CNTFETs with a 2nm thick oxide and three different dielectric constants, shows that a high-κ gate insulator does increase the on-current. The inset, which plots the conduction band profiles for three transistors at the on-state, sheds light on these results. Because the Schottky barriers are thin (due to thin gate oxide), the charge density inside the tube is high. When the gate dielectric constant is low, this charge produces a considerable self-consistent potential. The conduction band floats up, which makes the Schottky barrier thicker and the conduction band in the interior of the channel higher. Because the tunneling current exponentially depends on the Schottky barrier thickness, the on-current of transistors with low gate dielectric constants is smaller. Calculations which omit the charge on the CNT overestimate the current when the mobile charge is important.

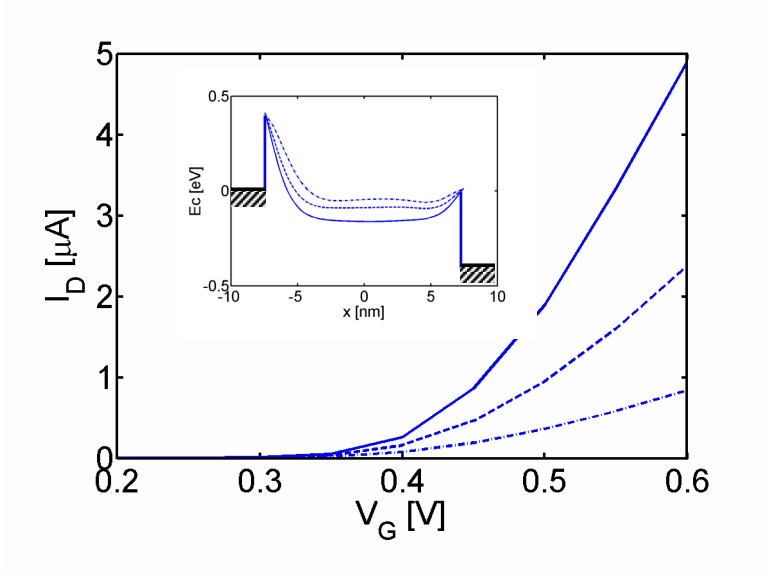


Fig.4.8 *Gate dielectric scaling*. I_D vs. V_G characteristics at V_D=0.4V for the nominal CNTFET with different gate dielectric constant. The solid line is for $\varepsilon_{ox} = 25$, the dashed line for $\varepsilon_{ox} = 4$ and the dash-dot line for $\varepsilon_{ox} = 1$. The inset shows the corresponding conduction band profile at V_G=0.6V. The flat band voltage is zero and mid-gap Schottky barriers are assumed for all transistors.

4.4 Discussions

The key point of this work, as also pointed out by Radosavljevic et al. [10], is the central importance of ambipolar conduction in SB CNTFETs when the gate oxide is thin, as it must be for high-performance transistors. To use such SB CNTFETs in conventional CMOS circuits, will require careful device design because negative gate to source voltages, and, therefore, high leakage currents, would result when transistors are stacked. The results presented in 4.3 shed light on how the leakage and on-current varies with device parameters such as gate insulator thickness, nanotube diameter, power supply voltage, etc. It is possible that gate work function engineering could be employed so that in the bias region of interest, only one branch of the ambipolar I-V is used, but this will be difficult because $V_{GS} < 0$ occurs for stacked transistors. Alternatively, one could explore MOSFET-like devices for which ambipolar conduction would not occur [55].

Finally, we should note that we also examined gate oxide thickness scaling and the effect of contact geometry. The results were similar to those of previous studies [8] (i.e. reducing the gate oxide thickness and contact size improves the subthreshold swing) were observed.

4.5 Conclusions

In summary, scaling issues for SB-CNTFETs were explored by self-consistent, atomistic scale simulations. Ambipolar conduction was found to be an important feature that must be carefully considered in transistor design. The minimum subthreshold swing is 60 mV/dec., just as it is for a MOSFET. The scaling limit for CNTFETs imposed by source-drain tunneling is between 5nm and 10nm and is determined by the small band gap and strong wave behavior of carriers in CNTs. CNTFETs, therefore, offer no scaling

advantage over a MOSFET. Larger tube diameter lowers the Schottky barrier height and delivers more on-current, but larger diameter nanotubes also result in larger leakage currents. Using a high-κ gate oxide improves the on-current, just as it improves the on-current of a MOSFET. The understanding of CNTFET device physics should prove useful in optimizing device designs.

5. ANALYSIS OF NEAR BALLISTIC CARBON NANOTUBE FIELD EFFECT TRANSISTORS

5.1 Introduction

Carbon nanotube field-effect transistors (CNTFETs) have received a lot of attention since the first demonstration [3, 4]. Significant advances have been achieved in both understanding the transistor physics and improving the device performance [5, 6]. Several important techniques for improving CNTFET performance have been successfully developed, such as high-κ gate dielectric insulator [11, 50], low Schottky barrier contacts [13], and self-aligned fabrication process for ultra-short channel transistors [14]. A high performance CNTFET, which integrates a short CNT channel, thin HfO_2 top gate insulator and low-barrier metal source/drain contacts, has recently been demonstrated [14].

In this work, we address device physics and performance issues based on an analysis of a high-performance CNTFET [14] using self-consistent quantum simulations. Techniques to characterize device parameters from electrical I-V measurements are developed and discussed. Numerical simulations are used to understand what controls the transistor performance, to suggest design optimization and identify the ultimate performance limit, and to compare the performance of CNTFETs to Si MOSFETs. The results are useful for understanding the device physics of near ballistic CNTFETs and identifying important issues for further improving CNTFET performance.

5. 2 Approach

The experimental CNTFET used in this analysis is shown in Fig. 5.1a, with the detailed fabrication process described in detail in [14]. A self-aligned process is developed to achieve an ultra-short channel, a thin high-κ top gate insulator is used to

maximize the gate modulation, and low-barrier source/drain contacts are used to optimize the metal-nanotube contact. Because of the integration of a short channel, thin high-κ insulator and low-barrier contacts, this transistor demonstrates the best performance for CNTFETs to date. Fig. 5.1b shows the measured I_D vs. V_D characteristics. A large source-drain current of ~20 μA and a near-ideal channel conductance of $0.5 \times 4e^2 / h$ are achieved at a gate overdrive $|V_G\text{-}V_T|$~1V.

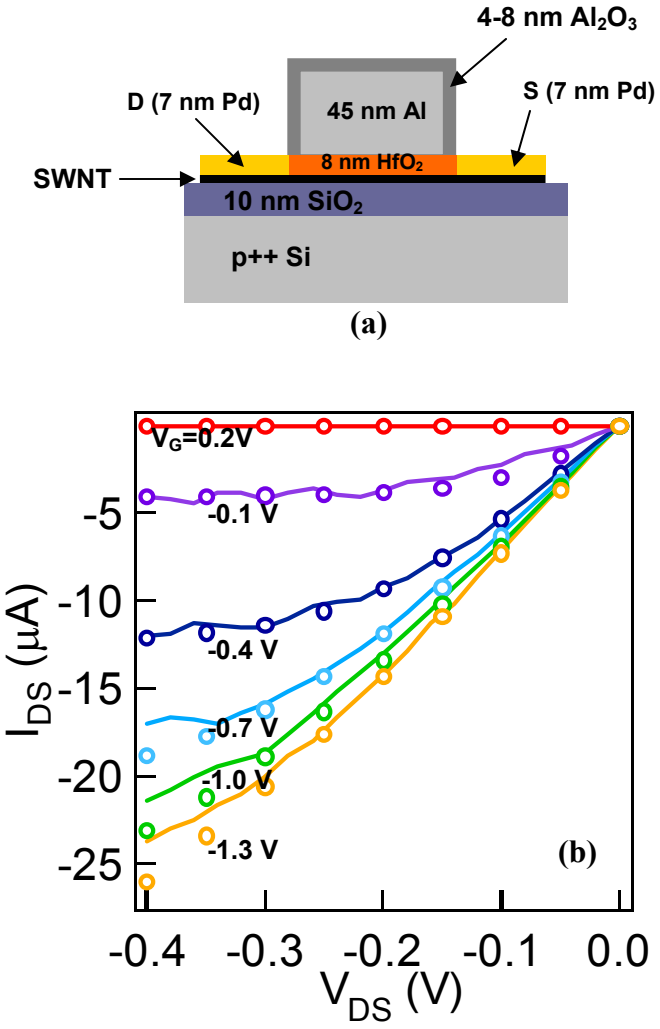


Fig. 5.1. (a) A recently reported CNTFET with Pd S/D contacts and a 50nm-long channel [14]. The HfO_2 top gate insulator is 8nm-thick with a dielectric constant $\kappa \approx 16$. The diameter of the intrinsic carbon nanotube channel is $d_{CNT} \approx 1.7$ nm. The Pd source/drain contact thickness is 7nm. (b) I_D vs. V_D of the device in (a). Solid lines are experimental data and symbols are ballistic quantum simulation.

To analyze this experimental FET, we performed self-consistent quantum simulations. The Schrödinger equation is self-consistently solved using the non-equilibrium Green's function (NEGF) formalism with a three-dimensional (3D) Poisson equation. For the NEGF formalism, an atomistic, p_z orbital description for the whole carbon nanotube channel is used. The nanotube channel is treated as a ballistic conductor. The potential variation around the tube circumferential direction is omitted, i.e., all carbon atoms in a ring around the tube feel the same electrostatic potential as the potential at the center of the ring computed by the Poisson equation. This approximation works as long as the potential drop over the circumferential direction of the nanotube is smaller than the energy spacing between the 1st and 2nd subband bottom (~0.25eV for the experimental tube [14]). A phenomenological treatment of the metal contact is used, and the metal-CNT Schottky barrier height is an input parameter to the simulation rather than computed. The detailed description for solving the quantum transport equation can be found in [32].

To determine the self-consistent potential for the quantum transport equation, a 3D Poisson equation, which treat the experimental transistor geometry, is solved using the method of moments [25]. The advantage of the method of moments is that the grid points are only needed where charge exists (the gate, source, drain and the channel). The total number of grid points, therefore, is much smaller compared to other numerical methods. The potential vector at the grid points {V} is related to the charge vector {q} by

$$\{V\} = [K]\{q\}, \tag{5.1}$$

where $[K]$ is the electrostatic kernel, whose element K_{ij} describes the potential felt by the ith grid point when there is a unit amount of charge at the jth grid element. For solving the Poisson equation, the charge on the carbon nanotube q_D (obtained from the solution of the quantum transport equation) and the potential at the source/drain/gate contacts V_C are inputs, and eqn. (5.1) becomes,

$$\begin{Bmatrix} V_D \\ V_C \end{Bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{Bmatrix} q_D \\ q_C \end{Bmatrix}, \tag{5.2}$$

where *A, B, C*, and *D* are submatrices of the electrostatic kernel matrix *K*, V_D is the potential on the nanotube channel and q_C is the charge on the source, drain, and gate electrodes. The potential on the nanotube is obtained by solving eqn. (5.2),

$$\{V_D\} = Aq_D - BD^{-1}Cq_D + BD^{-1}V_C\,. \tag{5.3}$$

The first term on the right hand side of eqn (5.3) is the potential produced directly by the charge on the tube. The second term is the potential produced by the image charge of the metal contacts, and the last term is the Laplace potential (the potential when there is no charge on the tube). The iteration between the NEGF transport equation and the 3D Poisson equation continues until self-consistency is achieved. The source-drain current is then computed,

$$I = \frac{4e}{h}\int dE \cdot trace(G\Gamma_S G^{+}\Gamma_D)\cdot[f_S(E) - f_D(E)], \tag{5.4}$$

where *G* is the retarded Green's function of the nanotube channel, $\Gamma_{S,D} = i(\Sigma_{S,D} - \Sigma_{S,D}^{+})$, and $\Sigma_{S,D}$ is the source/drain contact self-energy, and $f_{S,D}$ is the source/drain Fermi distribution function. After the I-V of the intrinsic transistor is simulated, a parasitic source/drain resistance of $R_{S,D} \sim 1.7\text{K}\Omega$ is added and the extrinsic I-V characteristics are computed [62]. A parasitic resistance is introduced to reflect the series resistance of thin Pd (~7 nm) S/D metal electrodes used in the experiment [14].

5.3 Characterization

In this section, we discuss techniques to characterize the Schottky barrier height and the CNT band gap from the transistor I-V measurement. We examin an approach for characterizing the SB height proposed by Appenzeller et al. [63]. We show that the Pd

source/drain makes a good contact to the carbon nanotube channel, and when the extracted barrier height is low, the previously proposed approach [63] may fail. Techniques to extract the tube bandgap and diameter from the electrical characteristics at the minimal leakage current point for an ambipolar CNTFET are also developed.

Fig 5.2a and 5.2b briefly summarized an approach proposed in [63] for characterizing the SB height. For a CNTFET with a positive barrier height, the source-drain current is limited by a thermal barrier in the CNT body when the magnitude of the gate voltage is low, and the thermionic source-drain current at high V_D is

$$I_D = \mathrm{T}\frac{4ek_BT}{h}\exp\left(-\frac{\Phi_B}{k_BT}\right) \tag{5.5}$$

where $0 < \mathcal{T} < 1$ is the channel transmission coefficient, T is the temperature, and Φ_B is the thermionic barrier height as shown in Fig. 2b, which can be extracted from a temperature-dependent measurement [63]. Alternatively, the thermal barrier height Φ_B can be computed from a room temperature I-V measurement,

$$\Phi_B = -k_BT\ln(I_D / I_0) \tag{5.6}$$

where $I_0 = \mathcal{T}\frac{4ek_BT}{h}$.

For a high performance CNTFET with $\mathcal{T} > 0.5$ [14], assuming $\mathcal{T} = 1$ for computing I_0 only introduces an error of $\Delta\Phi_B < 18meV$. So Φ_B can be obtained from a much simpler room-temperature measurement.

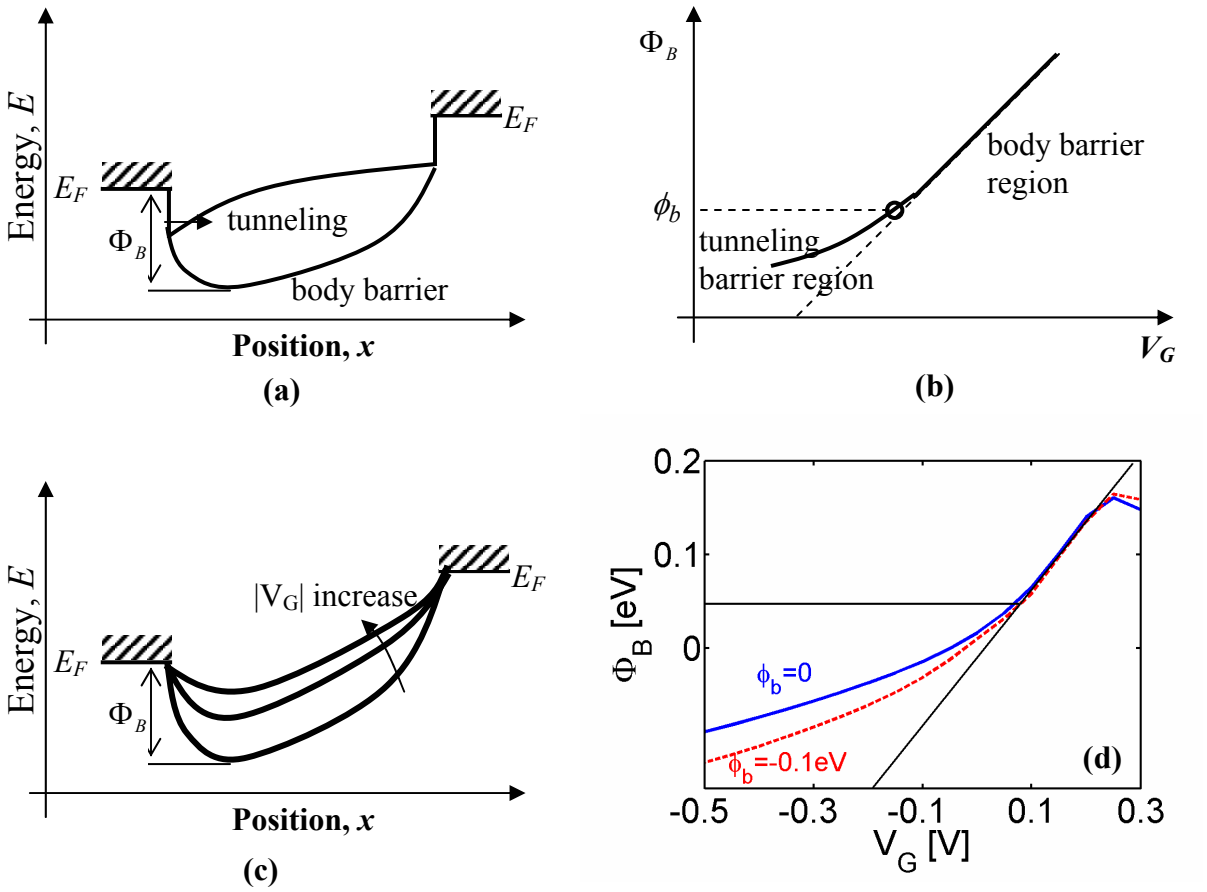


Fig. 5.2. *Extracting the SB height*. (a) The band diagram for a p-type CNTFET with a positive barrier height. When the magnitude of the gate voltage is small, a body barrier limits the source-drain current. When the magnitude of the gate voltage is large, a tunneling barrier at the source end of the channel limits the current. Φ_B is the thermal barrier height as defined in eqn. (5.6) in the text. (b) The thermal barrier height Φ_B vs. the gate voltage V_G for a CNTFET with a positive SB $\phi_b > 3k_BT$. The charge in the CNT channel is non-degenerate and negligible in the body barrier region. The Φ_B-V_G curve follows a linear relation. When the transistor begins to operate in the tunneling barrier region, the curve begins to deviate from a linear relation. The Schottky barrier height ϕ_b is identified at the transition point. (c) The sketch of E_V for a zero SB CNTFET at 3 different gate voltages. (d) The Φ_B-V_G curves for CNTFETs as shown in Fig. 1a with a zero SB $\phi_b = 0$ (the solid line) and a negative SB $\phi_b = -0.1eV$ (the dashed line). For both barrier heights, the simulated Φ_B-V_G curves begin to deviate from a linear relation at Φ_B~50meV, and the extracted value by the approach in [63] is ~ 50 meV, larger than the actual SB heights.

For each gate bias point, a thermal barrier height is extracted, and a Φ_B vs. V_G curve as shown in Fig. 5.2b is generated. As shown in Fig. 2a, when the source drain current is limited by a thermal barrier in the CNT channel, the thermal barrier is high ($\Phi_B > \phi_b$), and the small amount of charge inside the channel only produces a negligible self-consistent potential. The thermal barrier height Φ_B is directly modulated by the gate voltage V_G, and the Φ_B vs. V_G characteristic follows a linear relation as shown in the body barrier region in Fig. 2b. When the gate voltage decreases and the source-drain current begins to be limited by a tunneling barrier (the tunneling barrier region as shown in Fig. 2b), the Φ_B - V_G curve begins to deviate from the linear relation. The SB height is identified as the thermal barrier height at the point where the Φ_B - V_G curve begins to deviate from a linear relation. This technique has been shown to work when the SB is modestly high ($\phi_b > 2 - 3k_BT$) [63]. As discussed below, however, when the SB is low, the value extracted using this technique does not reflects the actual SB height.

Fig. 5.2c sketches the valence band profile at three different gate voltages for a zero SB CNTFET. In the subthreshold region when the thermal barrier is high (Φ_B>2-3k_BT) and the amount of the charge in the CNT channel is small, the Φ_B - V_G curve still follows a linear relation, but when $0 < \Phi_B < 2 - 3k_BT$, the charge in the tube channel becomes important and produces a non-negligible potential, which prevents the gate voltage from effectively lowering Φ_B. As the result, the Φ_B - V_G curve begins to deviate from the linear relation before the transistor begin to operates in the tunneling barrier region. As shown in Fig. 2d, which plots the simulated Φ_B - V_G curve for a zero SB ($\phi_b = 0$) CNTFET and a negative SB ($\phi_b = -0.1eV$) CNTFET, both curves begins to deviate from a linear relation at $\Phi_B \sim 50$ meV. The deviation is due to the transition from non-degenerate to degenerate carrier statistics in the CNT body, rather than due to the transition from the body barrier to the contact barrier region. When the extracted barrier

is low (< 2-3 $k_B T$), the extracted value may not reflect the actual barrier height, which can be even lower.

We applied the technique described in [63] to our experimental FET. As shown in Fig. 5.3, the extracted SB height is $\phi_b \approx 50\,\text{meV}$, which indicates that the Pd makes a low barrier contact to the nanotube. Because the extracted barrier height is low, the actual barrier height can be even below 50 meV, and other methods are needed to determine the exact barrier height.

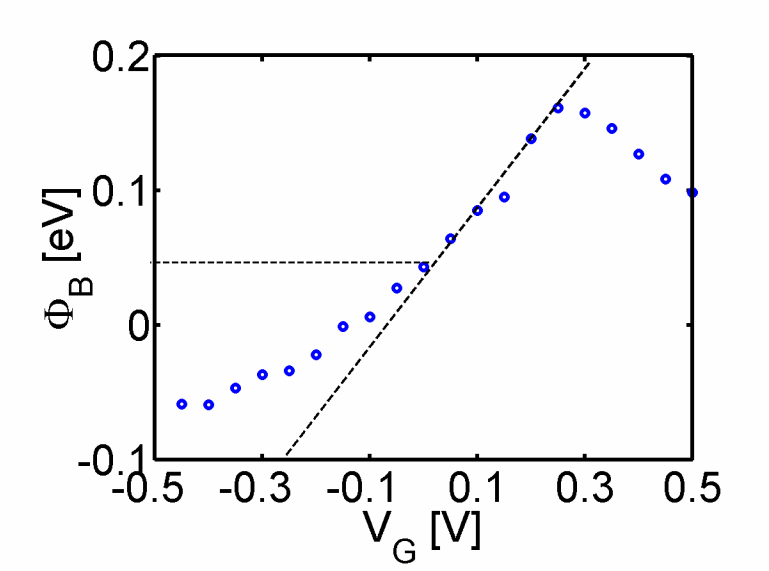


Fig. 5.3 The thermal barrier height Φ_B extracted from the measured room temperature I-V vs. the gate voltage for the transistor in Fig. 5.1.

Fig. 5.4 is an attempt to extract the SB height by fitting the whole Φ_B-V_G curve. Ballistic-channel CNTFETs with the experimental transistor geometry (Fig. 1a) and the measured parasitic resistance ($R_{S,D}$~1.7 $K\Omega$) are simulated. Fig. 4 plots the experimental Φ_B-V_G curve (the circles) and three simulated Φ_B-V_G curves for the SB heights, $\phi_b = 0.1\,\text{eV}$ (the dashed line), $\phi_b = 0$ (the solid line), and $\phi_b = -0.1\,\text{eV}$ (the dotted line).

When the SB height varies from -0.1eV to 0.1eV, the variation of the thermal barrier height Φ_B is < 50 meV, which indicates that when the gate control is good (the gate oxide is thin), the transistor characteristics become less sensitive to the SB height [Guo03b, Rad04]. On the one hand, the less important role of SB height for a thin-oxide CNTFET makes it difficult to extract an accurate SB height. On the other hand, the uncertainty of the extracted SB height for a thin-oxide CNTFET doesn't significantly affect the results. As shown in Fig. 4, the measured Φ_B-V_G curve (the circles) is between the simulated Φ_B-V_G curve for $\phi_b = 0.1\,\text{eV}$ (the dashed line) and the simulated curve for $\phi_b = 0$ (the solid line), which, again, indicates, that Pd makes a low barrier contact to the 1.7nm-diameter CNT channel.

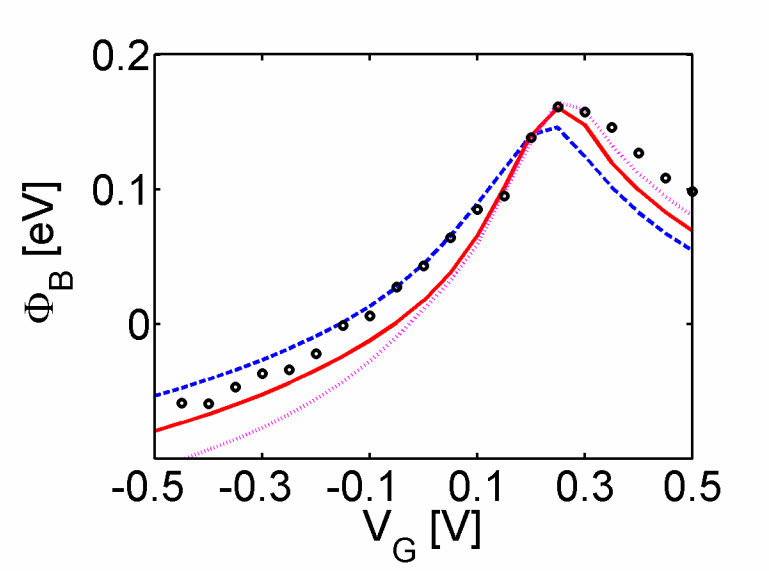


Fig. 5.4. The thermal barrier height Φ_B vs. the gate voltage V_G for extracting the SB height. The solid line is the experimental characteristic. The dashed line is for a simulated SB height $\phi_{bp} = 0.1\,\text{eV}$, the circles are for a simulated $\phi_{bp} = 0$, and the dotted line is for a simulated $\phi_{bp} = -0.1\,\text{eV}$. The tube diameter is $d_{CNT} \approx 1.7\,\text{nm}$ for all simulations.

Next we extract the bandgap and the diameter of the carbon nanotube channel. Fig. 5.5 shows the experimental Φ_B - V_G characteristic (the circles), and the simulated Φ_B - V_G curves for three different tube diameters, d_{CNT}~1.5nm (the dashed line), d_{CNT}~1.7nm (the solid line), and d_{CNT}~2.0nm (the dotted line). The maximum Φ_B in the curve (which corresponds to the minimal leakage current point in an ambipolar I_D-V_G curve) is sensitive to the tube diameter. It is therefore, possible, to extract the tube diameter by fitting the experimentally measured Φ_B - V_G curve near the largest Φ_B point (or equivalent, fitting the minimal leakage current in an ambipolar I_D-V_G curve). Fig. 5.5 indicates that the best fitting is achieved when the tube diameter d_{CNT}~1.7nm, which results in a tube band gap of E_g~0.49eV. The numerical fitting techniques only needs a room-temperature I-V and simplifies the measurement, but the analysis based on detailed numerical simulation is time-consuming. Alternatively, the tube diameter and band gap can also be extracted from temperature-dependent measurement using a simple analysis as described next.

Fig. 5.6a sketches the log (I_D) vs. V_G characteristic for a CNTFET with a thin gate oxide, which displays electron conduction at high gate voltages and hole conduction a low gate voltages [9, 10]. The band gap can be extracted from the electrical characterization of the minimal leakage point. Fig. 5.6b sketches the band diagram at the minimal leakage point at a low V_D for a CNTFET with zero barrier height for holes. Although a high SB exists for electron transport, the SB is nearly transparent for electron tunneling for two reasons. The first reason is that the SB for a high-performance transistor is thin when the gate oxide is thin. (The SB thickness is approximately the same as the gate oxide thickness [60]). The second reason is that the small carrier effective mass in CNTs greatly facilitates quantum-mechanical tunneling [63]. The barrier that limits the electron leakage current is the thermal barrier in the CNT body, Δ_n, as shown in Fig. 6b. Notice that at the minimal leakage point, the electron current is equal to the hole current. In terms of barrier height, this requires

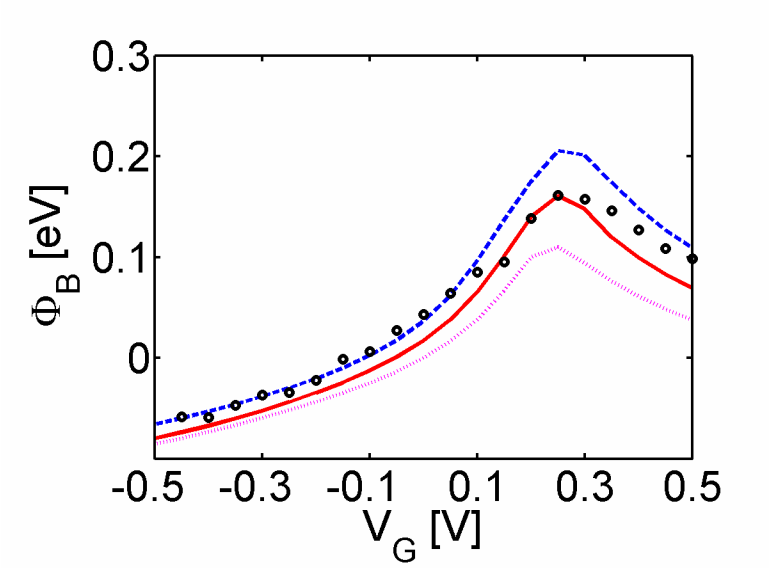


Fig. 5.5. The thermal barrier height Φ_B vs. the gate voltage V_G for extracting the tube diameter. The solid line is the experimental characteristic obtained from the measured room-temperature I_D vs. V_G at V_D=*-0.3*V. The dashed line is for a simulated tube diameter $d_{CNT} \approx 1.5$ nm with a band gap $E_g \approx 0.57$ eV, the circles are for a simulated $d_{CNT} \approx 1.7$ nm with $E_g \approx 0.49$ eV, and the dotted line is for a simulated $d_{CNT} \approx 2.0$ nm with $E_g \approx 0.43$ eV. The metal-CNT Schottky barrier height for holes of is zero in all simulations.

$$\Delta_n \approx \Delta_p \approx E_g / 2 \tag{5.7}$$

at low drain biases. The minimal leakage current can be estimated as the sum of the thermionic emission current over Δ_n for electrons and the current over Δ_p for holes,

$$I_{\min} \approx \mathrm{T}\frac{8ek_BT}{h}\exp\left(-\frac{E_g}{2k_BT}\right)\left[1-\exp\left(\frac{eV_D}{k_BT}\right)\right] \approx \mathrm{T}\frac{8e^2}{h}\exp\left(-\frac{E_g}{2k_BT}\right)V_D \tag{5.8}$$

where $0 < \mathcal{T} < 1$ is the channel transmission coefficient. The channel resistance at the minimal leakage point is

$$R_{\min} = \frac{V_D}{I_{\min}} = R_0 \exp(\frac{E_g}{2k_B T}) \tag{5.9}$$

where $R_0 = \frac{h}{8e^2}\frac{1}{T}$.

The temperature dependence of R_0 is small compared to the exponential factor in eqn. (9). The CNT band gap, therefore, can be extracted from the slope of the $\ln(R_{\min})$ vs. 1/T plot using a temperature dependent measurement,

$$\ln(R_{\min}) = \frac{E_g}{2k_B}\frac{1}{T} + \ln(R_0). \tag{5.10}$$

As shown in [14], the band-gap extracted by using this temperature-dependent measurement is $E_g \approx 0.50\,\text{eV}$, which is in good agreement with the value obtained by numerically fitting ($E_g \approx 0.49\,\text{eV}$).

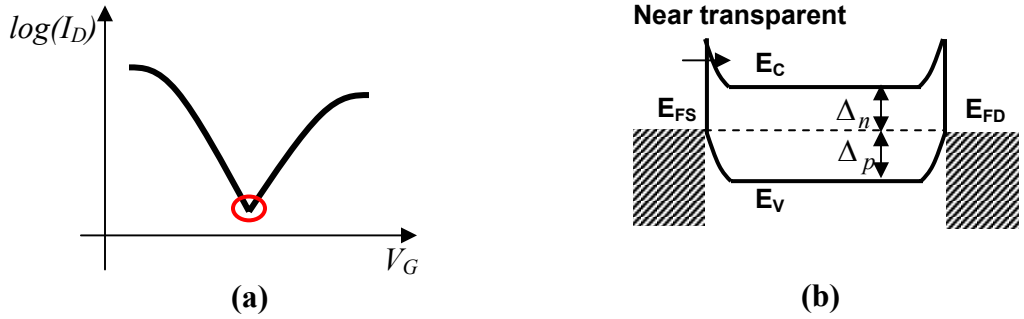


Fig. 5.6. (a) log I_D vs. V_G sketch for a thin-gate-oxide CNTFET with metal contacts. When the gate oxide is thin, the transistor shows ambipolar conduction. At the minimal leakage current (the circled point), the electron current is equal to the hole current. (b) The band diagram sketch at the minimal leakage point for a CNTFET with a thin gate oxide at a low V_D. The SB height for holes is zero ($\phi_{bp} = 0$). Δ_n (Δ_p) is the thermal barrier height for electrons (holes) in the CNT channel.

5.4 Analysis

In this section, we use quantum simulations to understand what controls the transistor performance and to suggest possible optimization to further improve transistor performance. Fig. 5.7 plots the experimental (the circles) and the theoretical (the lines) I_D vs. V_G characteristics. In the quantum simulation, the SB height $\phi_{bp} = 0$, the tube diameter d_{CNT} ~1.7nm, and the CNT channel is treated as a ballistic conductor. The simulated subthreshold swing at V_D=0.3V is S~80mV/dec, which is larger than the theoretical limit, 60mV/dec, but smaller than the measured value at V_D = 0.3V, of ~110mV/dec. The degradation of the simulated S from the ideal value is due to the thin bottom gate oxide, which is only ~10nm thick. The capacitance between the tube and the bottom gate electrode is still considerable, compared to the capacitance between the tube and the top gate. The existence of the bottom gate capacitance degrades S from the theoretical limit value to ~80mV/dec. Using a thicker bottom gate oxide can improve the subthreshold swing. Although the back gate parasitic capacitance is included in the simulation, the simulated value (S~80mV/dec) is still smaller than the measured value (S~110mV/dec), which can be due to other factors like significant interface charge states at the HfO_2 and SiO_2 interface.

The possible existence of significant interface/oxide charge is also consistent with the discrepancy between the simulated and measured threshold voltage. The threshold voltage of a transistor is determined by the gate-channel work function difference and interface/oxide charges [62]. Our simulation includes gate-CNT work function difference but no oxide/interface charges. The simulated threshold voltage is found to be ~0.75V smaller than the experimentally measured value. The mismatch of V_T between experiments and simulation after the HfO_2 top gate insulator deposition can be due to factors, such as the significant HfO_2/SiO_2 interface charges. The mismatch can also be due to the uncertainty of the top and bottom gate workfunction used as the simulation input. The oxidation of the Al top gate can change its work function from the value for pure Al and the doping density of the bottom gate can also charge the workfunction of the bottom gate.

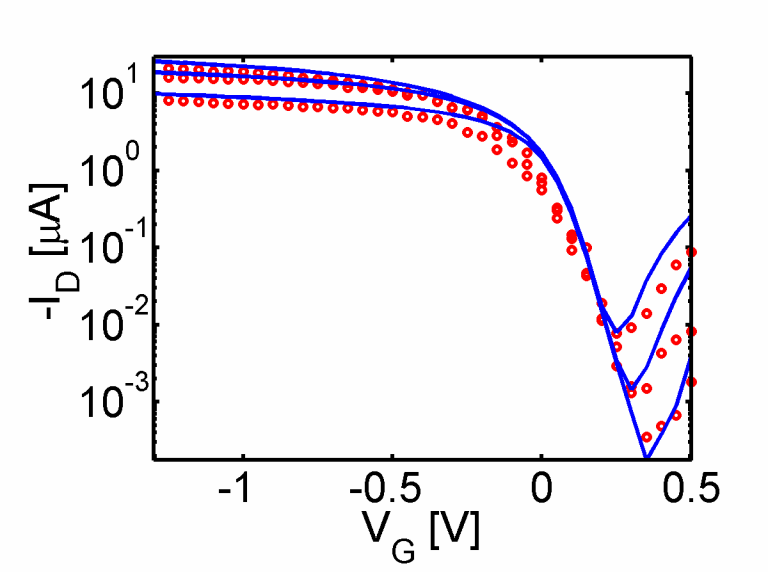


Fig. 5.7. The experimental (the dashed lines) and simulated (the solid lines) I_D vs. V_G characteristics at V_D=*-0.1, -0.2,* and *-0.3*V. For the simulated CNTFET, the Schottky barrier height for holes is $\phi_{bp} = 0$ and the tube diameter $d_{CNT} \approx 1.7$ nm. The CNT work function $\Phi_{CNT} = 4.7$ eV, the Al top gate work function $\Phi_{Al} = 4.1$ eV, and the p^+ doped bottom gate work function $\Phi_{pSi} = 5.1$ eV. No interface and oxide charges are included. The simulated curves are translated by *+0.75*V along the x-axis to match the experimental curves.

Figure 5.8 plots the experimental I_D vs. V_G characteristic at V_G=-0.4V (circles). To explore how close the experimental FET operates to its ballistic limit, we simulated the I_D vs. V_G characteristics for a ballistic CNTFET with zero SB, $\phi_{bp} = 0$, and zero parasitic source/drain resistance, $R_{S,D} = 0$ (the solid line). The experimental FET delivers >90% on-current of the ballistic current of the $\phi_{bp} = 0$ CNTFET. The performance of a zero SB transistor can be further improved if the barrier height can be made negative [64], because at on-state, a significant portion of carriers needs to tunnel through even though the SB is zero. If a negative SB could be achieved, the transistor on-current will increase and finally saturate when the SB is sufficiently negative. The on-current of a ballistic SB

transistor with a sufficiently negative barrier represents an upper-performance limit when current is limited neither by the contact nor by scattering [64]. To explore how close the experimental FET operates to the upper-performance limit, we also simulated the I_D-V_D of a ballistic CNTFET with a sufficiently negative SB, $\phi_{bp} = -0.3\,\mathrm{V}$, as shown by the dash-dot line in Fig. 8. The experimental FET delivers ~50% of the on-current of a ballistic CNTFET with a sufficiently negative barrier height. The results indicate that the experimental FET operates relatively close even to its upper-performance limit.

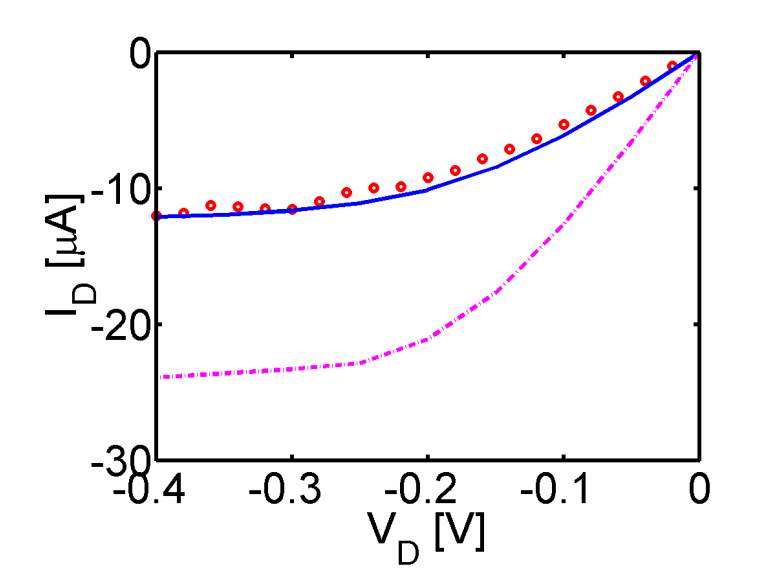


Fig. 5.8. The experimental (circles) and simulated (solid and dash-dot lines) I_D vs. V_D at $V_G = -0.4V$. The solid line is simulated for a ballistic-channel CNTFET with a zero SB for holes ($\phi_{bp} = 0$). The dash-dot line simulated for a ballistic CNTFET with a sufficiently negative SB height $\phi_{bp} = -0.3eV$. The source/drain parasitic resistance $R_{S,D} = 0$ in simulations.

Fig. 5.9 plots the channel conductance, $G_D = \partial I_D / \partial V_D \,|_{V_D=0}$, vs. the gate voltage, V_G, for the experimental FET (circles), the simulated FET with $\phi_{bp} = 0$ and $R_{S,D}=0$ (the solid line), and the simulated FET with $\phi_{bp} = -0.3V$ and $R_{S,D}=0$ (the dash-dot line). This plot further confirms that the experimental CNTFET has a source-drain conductance close to the ballistic value. The measured G_D is ~80% of the zero SB CNTFET in the measured gate voltage range, which again indicates that scattering in the CNT channel only has a small effect on the DC channel conductance. As shown in Fig. 9, the channel conductance of the $\phi_{bp} = 0$ CNTFET is still below that of the $\phi_{bp} = -0.3\,\text{eV}$ CNTFET and $G_0 = 4e^2 / h \sim 155\mu S$. The first reason is that the carrier injected near the Fermi energy level still experience quantum-mechanical reflection for a zero SB. The second reason is that at room temperature, thermal broadening of the Fermi function reduces G_D below G_0.

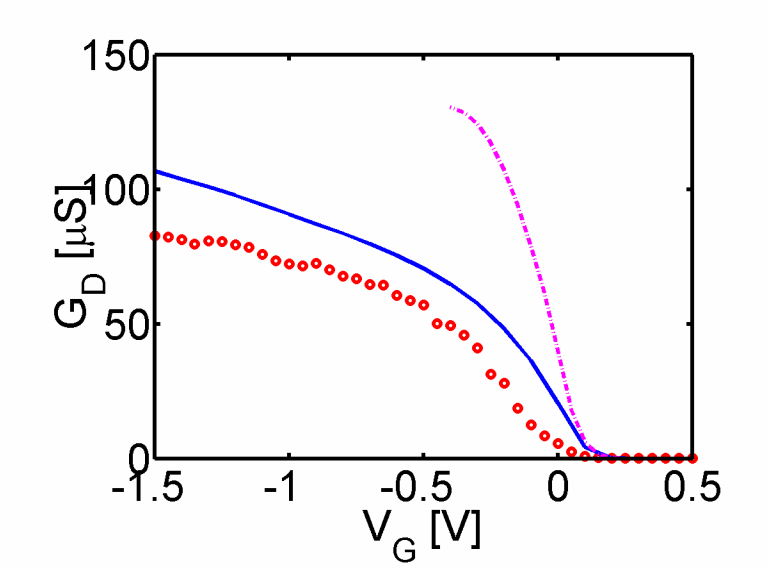


Fig. 5.9. The experimental (circles) and simulated (solid and dash-dot lines) channel conductance, $G_D = \partial I_D / \partial V_D \,|_{V_D=0}$, vs. the gate voltage, V_G. The solid line is for a simulated CNTFET with a zero SB, $\phi_{bp} = 0$, and zero parasitic resistance, $R_{S,D}=0$. The dash-dot line is for a simulated CNTFET with a negative SB, $\phi_{bp} = -0.3eV$, and zero parasitic resistance, $R_{S,D}=0$.

Figure 5.10 is an attempt to understand why scattering in the channel only has a small effect on the DC characteristics of the transistor. As pointed by previous studies [41, 43, 44], phonon scattering is the most important scattering mechanism in high-quality, single-wall carbon nanotubes. Acoustic phonon scattering has a long mean free path (~1μm) [41, 43, 44, 54, 65], therefore, it only has a small effect on the source-drain current for L_{ch}~50nm. Optical phonon (OP) scattering (with $\hbar\omega_{OP} \sim 0.16eV$) has a much shorter mean free path (~10nm) and scatters carriers even in a short channel. Due to the short mean free path, the carrier can emit an OP near the drain end of channel as shown in Fig. 5.10. After OP emission, the backscattered hole encounters a much thicker and higher SB because lose of a large OP energy $\hbar\omega_{OP} \sim 0.16eV$. The back scattered hole has little chance to tunnel through the SB and return back to the source. (OP absorption can also occur if hot phonon effects are considered, but it also has a small effect on the DC source-drain current. The treatment of hot phonon effects is beyond the scope of this paper and will be discussed in a separate study.) Although OP emission scatters carriers near the drain end of the channel, it only has a small effect on the DC characteristics of the transistor. In contrast, for a ~50nm-long metallic CNT where a potential barrier does not exist, OP emission begins to play a significant role when the applied source/drain bias is lager than the optical phonon energy $V_D > \hbar\omega_{OP}/e$ [43, 44].

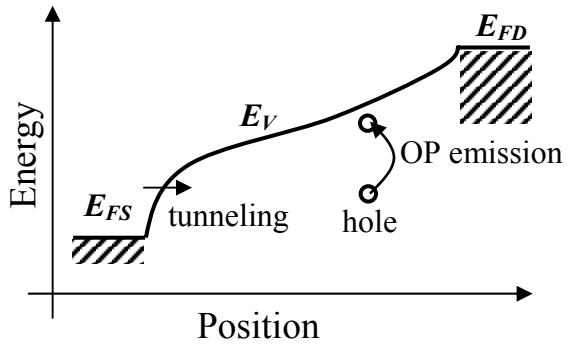


Fig. 5.10. The valence band profile at on-state for a $\phi_{bp} = 0$ CNTFET. Acoustic phonon scattering has a long mean-free-path and thus a small effect on the source-drain current for a 50nm-channel-length CNTFET [41, 43, 44, 65]. The source injected hole can emit an optical phonon with $\hbar\omega_{OP} \sim 0.16$ and get back scattered near drain, but the backscattered hole encounters a much thicker and higher SB and has little chance to return back to the source.

Next we explore the performance improvement by the use of a high-κ top gate insulator. Fig. 5.11a plots the I_D vs. V_D characteristics at V_G=-0.4V for three top gate insulators, (1) t_{top}=8nm SiO_2 as shown by the dashed line, (2) t_{top}=8nm HfO_2 used in the experiment [14] as shown by the solid line, and (3) t_{top}=4nm HfO_2 as shown by the dotted line. Using high-κ gate insulator with $\kappa \sim 16$ (HfO_2) instead of the SiO_2 gate insulator improves the transistor on-current by approximately a factor of 2. This result is in contrast with a previous study [51], which indicated that little improvements could be achieved by using high-κ gate insulator for CNTFETs. Fig. 5.11b, which plots the on-state valence band diagram, explains the reason. At the on-state, the charge in the nanotube channel can produce a considerable self-consistent potential which lowers the valence band in the CNT body and increases the SB thickness at the contact. When a high-κ gate insulator is used, the electric field produced by the charge on the tube is effectively screened by the high-κ gate insulator, and the self-consistent potential is smaller. Compared to the band diagram for SiO_2 top gate insulator as shown in Fig. 10, the Schottky barrier at the source contact is thinner, and the valence band in the CNT body is more effectively lifted up by the gate voltage by applying HfO_2 gate insulator.

On the other hand, a factor of 4 improvement of the gate dielectric by applying HfO_2 instead of SiO_2 only results in a factor of ~2 improvement of on-current for the studied CNTFET, which is smaller than what is expected for a MOSFET-like transistor. The major reason is that the mechanism of gate modulation for CNTFETs is different from that for MOSFET-like transistors. For MOSFET-like transistors, the gate modulates the channel conductance by a field in the direction normal to the channel. Increasing the gate dielectric constant results in a proportional increase of the charge at the beginning of the channel thereby the on-current. For CNTFETs, the gate modulates the tunneling conductance of the metal-nanotube contact at the source using fringing field, a modulation mechanism not as effective as that for a MOSFET-like transistor. As the result, applying a high-κ gate insulator offers smaller on-current improvement in Schottky barrier CNTFETs.

We also explore the advantages offered by vertically scaling down the gate oxide. Fig. 5.10a shows that a factor of ~2 improvement can be achieved if the HfO_2 top gate insulator thickness is scaled down to 4nm. The dotted line in Fig. 5.10b plots the valence band diagram for the CNTFET with a t_{top}=4nm HfO_2 top gate insulator. The distance by which the source field penetrates to the channel is approximately the gate insulator thickness as shown by a previous study of nanotube transistor electrostatics [60]. Using thinner gate oxide reduces the Schottky barrier thickness and improves the transistor on-current.

It is interesting to explore what gate dielectric constant is large enough to maximize the transistor performance. Fig. 5.12 plots the CNTFET on-state transconductance vs. the gate dielectric constant for four types of widely used gate insulators. The improvement of transistor performance by increasing the gate insulator dielectric constant decreases as κ increases. The transconductance tends to saturate when the gate insulator dielectric constant is large. The reason is that the high-κ gate insulator improves the CNTFET performance by reducing the self-consistent potential produced by the charge on the tube. For CNTFETs, the channel is one-dimensional and the total amount of charge in the channel is small. If the channel is already surrounded by a high-κ gate insulator and the self-consistent potential produced by the charge on the tube is already small, further improving the gate dielectric constant does not help to significantly reduce the Schottky barrier thickness and the transistor performance. When the gate insulator dielectric constant keeps increasing, the potential profile approaches the Laplace potential profile (the potential when there is no charge in the tube channel), and the transistor performance saturates at the limit set by the Laplace potential profile.

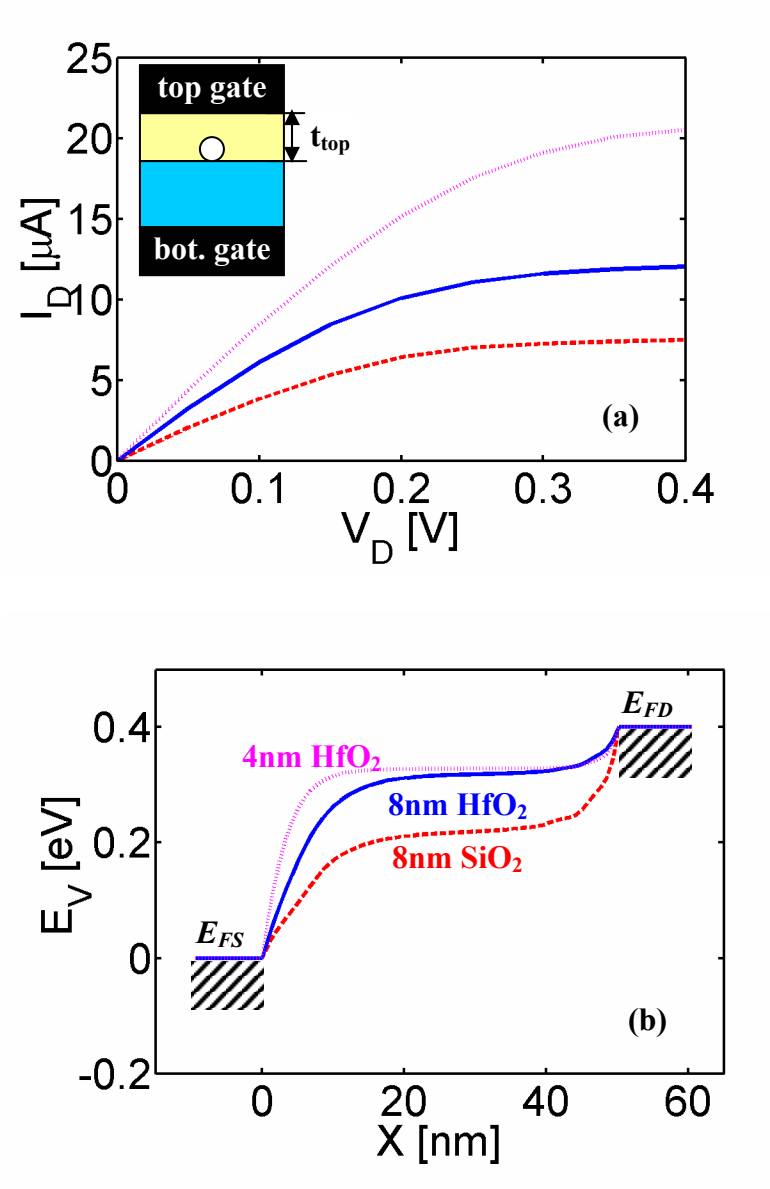


Fig. 5.11. (a) The simulated I_D vs. V_D characteristics at V_G=*-0.4*V for three different top gate insulators. The solid line is for the experiment with a top gate insulator thickness t_{top}=*8*nm and dielectric constant $\kappa = 16$ (for HfO_2). The dashed line is for SiO_2 insulator with t_{top}=*8*nm and $\kappa = 4$, and the dotted line is for a scaled, high-κ top gate oxide with t_{top}=*4*nm and $\kappa = 16$. The simulated SB height for holes is $\phi_{bp} = 0$. (b) The valence band profile at on-state (V_D=V_G=-0.4V) for three top gate insulators simulated in (a).

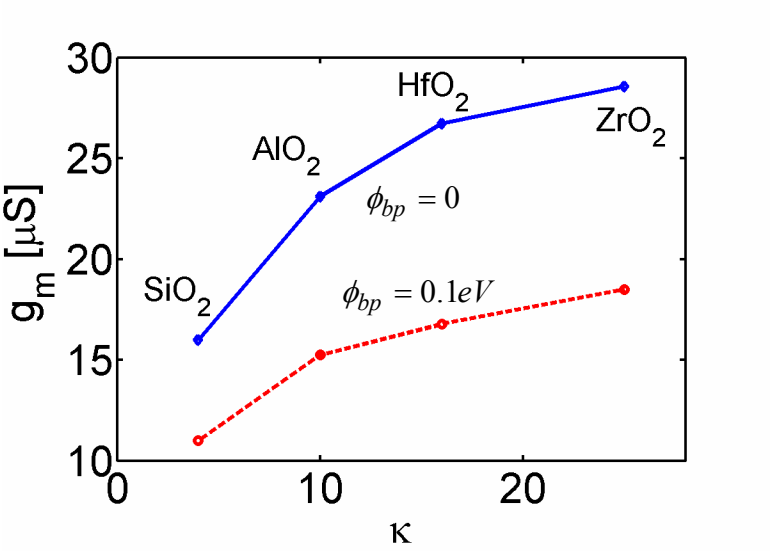


Fig. 5.12. The transconductance at $V_G = V_D = -0.4\,\mathrm{V}$ vs. the top gate insulator dielectric constant κ for $\phi_{bp} = 0$ and $\phi_{bp} = 0.1eV$. The top gate insulator thickness is t_{top}=8nm. The gate dielectric constants of four types of gate insulators are explored.

Finally, we explore the role of higher subband conduction. Fig. 5.13 plots the percentage of the 1st subband and 2nd subband current in the total current. The contribution of total current from higher subbands is small (<10%) over the whole measured gate voltage range. Fig. 5.13b, which plots the 1st and 2nd subband profile at V_G=-1.3V, explains the reason. The SB for the 1st subband at both the source and drain contacts is zero and the SB for the 2nd subband is much higher (~0.25eV) due to the large subband spacing (~0.25eV) in a small diameter tube (d_{CNT}~1.7nm). Because the tunneling probability through a SB depends exponentially on the barrier height and thickness, the current of the 2nd subband is small compared to that of the 1st subband. As a result, the 1st subband conduction dominates for the analyzed experimental FET.

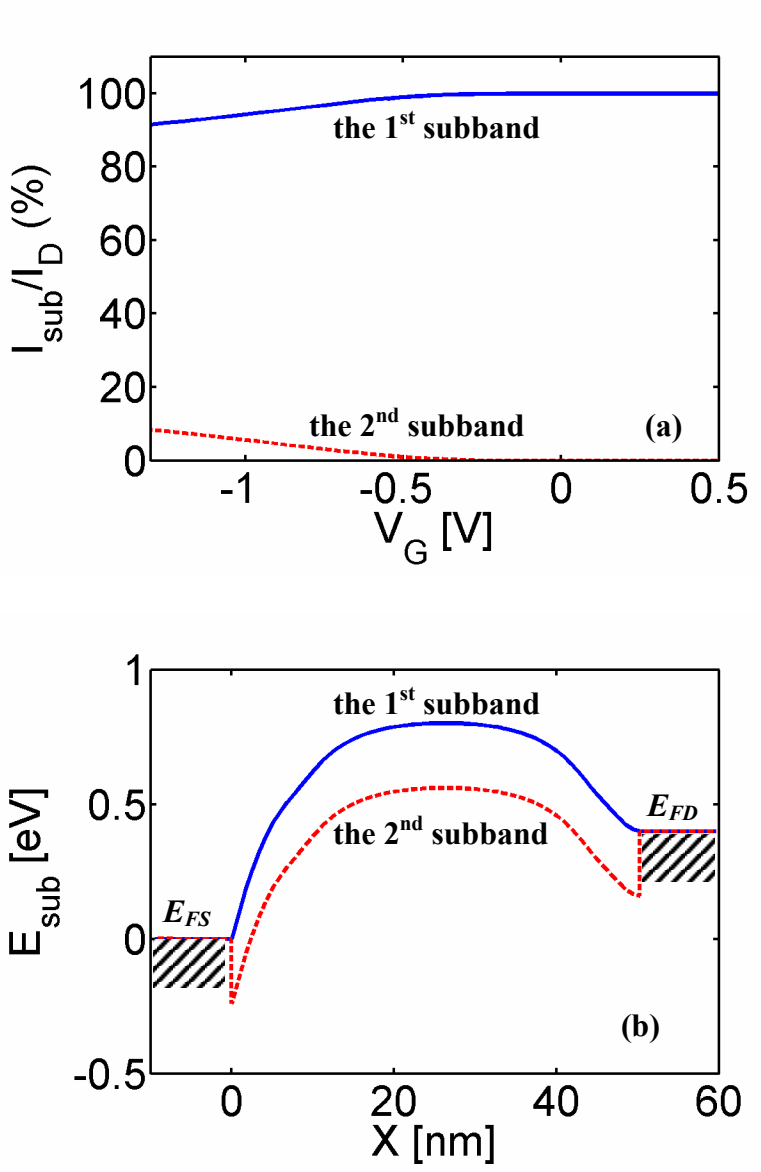


Fig. 5.13. (a) The percentages of the 1st and 2nd subband currents in the total current vs. the gate voltage. The simulated SB height for holes is $\phi_{sp} = 0$ and the drain bias $V_D = -0.4$ V. (b) The valence band profile for the 1st and the 2nd subband at a large gate overdrive, $V_G = -1.3$ V.

5.5 Discussion

As the performance of CNTFETs is progressing rapidly, it is important to compare the CNTFET performance to Si MOSFETs. A reasonable comparison requires us to: (1) assess transistor performance at a proper power supply voltage, (2) include both the on-state and off-state performance, and (3) fairly compare device metrics for different channel geometries. Fig. 14 is an attempt to compare the state-of-the-art 100nm-node-technology Si MOSFETs with the gate length L_G=50nm [66] to our experimental CNTFET also with the same L_G=50nm [14]. Fig. 14a plots the transistor intrinsic delay τ vs. I_{ON}/I_{OFF} for both transistors, which is generated using the method described in Fig. 14b. For a specified power supply voltage V_{DD}, an on-current and off-current is obtained by reading the current value at the edges of the gray window in Fig. 5.14b. The transistor intrinsic delay is computed as $\tau = C_G V_{DD} / I_{ON}$. For the Si MOSFET, the gate capacitance is obtained from the C-V measurement at the inversion operation region. For the CNTFET, C_G is extracted from the slope of the charge on the tube (Q_{CNT}) vs. the top gate voltage (V_{top}) plot above V_t for a CNT capacitor with the same gate geometry as the experimental FET [67]. A data point on the τ vs. I_{ON}/I_{OFF} plot is obtained. By sweeping the gray window along the V_G axis (which corresponds to adjusting the transistor threshold voltage), an τ vs. I_{ON}/I_{OFF} curve is generated. The advantage of using τ and I_{ON}/I_{OFF} for comparison is that they don't depend on the channel geometry, and I_{ON}/I_{OFF} considers both the off and on states. The results show that for an on-off ratio of 100, the intrinsic delay of the CNTFET at V_{DD}=0.4V is ~60% of that of the Si MOSFET at V_{DD}=1V, and a factor of 3-4 times smaller than Si FET delay at V_{DD}=0.4V because the thermal velocity in CNT is larger than that in Si, and the CNTFET operates closer to its ballistic limit.

The performance of the analyzed experimental CNTFET can be further improved by design optimization. As shown in Fig. 5.14a, the highest achievable on-off ratio for the experimental CNTFET at V_{DD} = 0.4V is only ~100, due to the ambipolar conduction of the CNTFET. The maximum on-off current ratio can be significantly improved if a

smaller diameter tube (which results in a larger band gap) is used. In addition, the 8nm high-κ gate insulator is still thick. Fig. 5.14a also plots the τ vs. I_{ON}/I_{OFF} characteristics for the analyzed FET (the dash-dot line) and the proposed optimized CNTFET (the dashed line). The optimization includes (1) using a smaller diameter (d_{CNT}~1.0nm) tube as the channel, (2) assuming that a zero SB can still be achieved even after a smaller diameter tube is used, (3) using a thinner HfO_2 top gate oxide, t_{top}=3nm, and (4) reducing the parasitic source/drain resistance. A much higher on-off ratio of $>10^4$ can be achieved at V_{DD}=0.4V after the optimization, thanks to the significant reduction of the minimal leakage current by using a smaller diameter tube. For an on-off ratio, $I_{ON}/I_{OFF}<100$, both CNTFEETs have similar intrinsic transistor delays. The reason is that after optimization, the transistor gate capacitance increases due to thinner top gate oxide, which offsets the improvement of the on-current when the transistor intrinsic delay, $\tau = C_G V_{DD} / I_{ON}$, is computed.

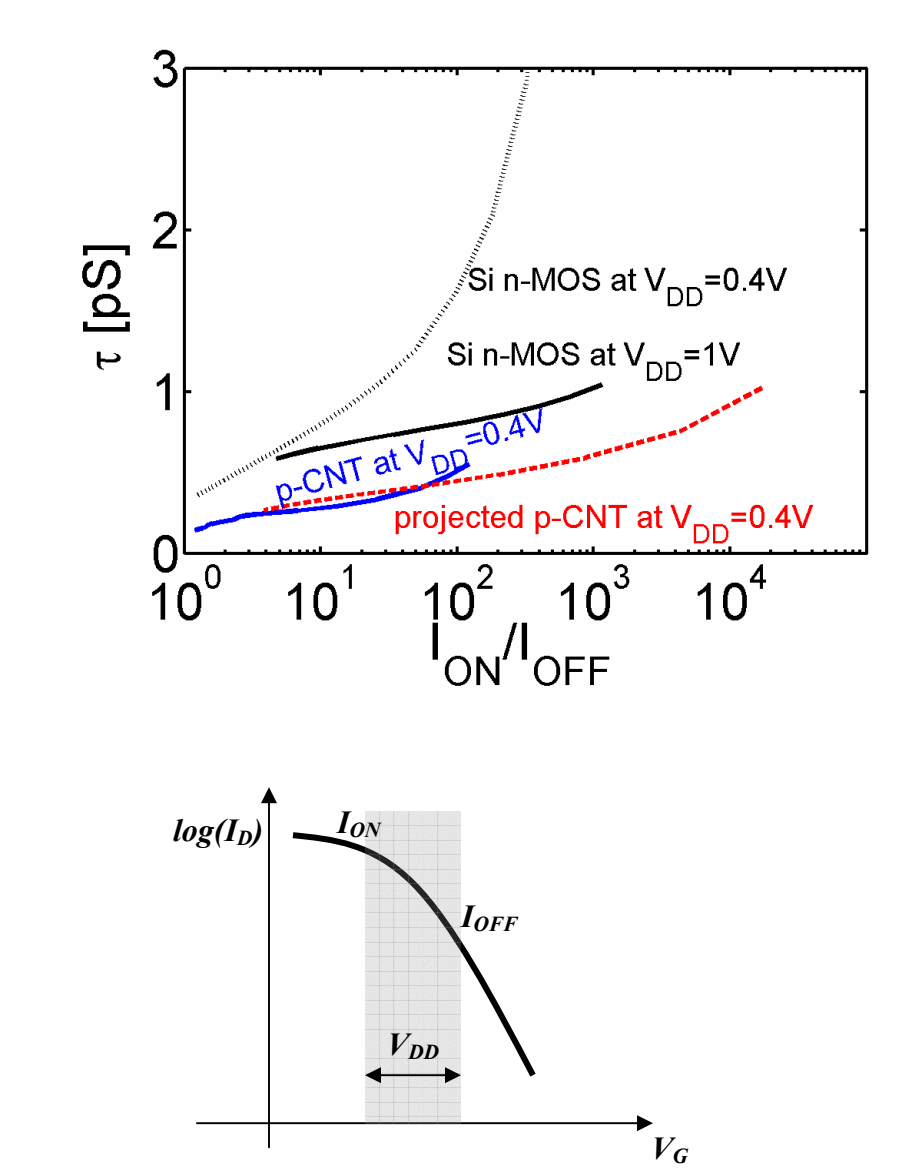


Fig. 5.14. *Comparing CNTFETs to Si MOSFETs.* (a) The intrinsic transistor delay vs. the on-off ratio for a state-of-the-art Si n-type MOSFET with the gate length L_G~50nm [Let01] at the power supply voltage V_{DD}=1V (the solid line) and V_{DD}=0.4V (the dotted line), an experimental p-type CNTFET [Jav04a] at V_{DD}=0.4V (the dash-dot line), and a theoretically projected CNTFET described in the text at V_{DD}=0.4V (the dashed line). (b) I_D vs. V_G characteristics of a p-type transistor, which shows the curves in (a) are generated (see text).

5.6 Conclusions

A recently demonstrated CNTFET is analyzed in detail using self-consistent quantum simulations. The following results are obtained in this study.

(1) Pd source/drain makes a good contact to the ~1.7nm-diameter tube channel with a Schottky barrier height $\phi_b < 80$ meV. Techniques for extracting tube band gap and diameter based on I-V measurement are proposed.

(2) Phonon scattering only has a small effect on the DC source-drain current. At V_G=-0.4V, the transistor delivers ~90% of the on-current of a ballistic CNTFET with a zero SB, and ~50% on-current of a ballistic CNTFET with a sufficiently negative SB.

(3) Applying high-κ gate insulator improves the CNTFET performance up to a certain κ value. High-κ gate insulator more effectively screening the self-consistent potential produced by the charge on the tube. Scaling down the oxide thickness results in thinner Schottky barrier and larger on-current.

(4) Conduction through higher subbands is small compared to conduction through the lowest subband.

(5) Further performance improvement can be achieved by using a smaller diameter tube as the channel, producing lower barrier contacts, applying thinner high-κ gate insulator, and reducing the parasitic source/drain resistance. Other options to significantly suppress ambipolar conduction and reduce the minimal leakage current are to produce MOSFET-like CNTFETs, which have heavily doped CNTs as source/drain extensions [55], or to use asymmetric source/drain design [68, 69].

(6) Under much lower power supply voltages, the delay metric of CNTFETs can be several times smaller than that of MOSFETs operating at the same on-off ratio.

6. ON THE ROLE OF PHONON SCATTERING IN CARBON NANOTUBE FIELD-EFFECT TRANSISTORS

6.1 Introduction

Understanding carrier transport in carbon nanotubes (CNTs) is important for both physics and applications [5, 6]. The mean-free-paths (mfps) for elastic scattering are known to be very long in carbon nanotubes (~1μm), but optical phonon (OP) scattering has been shown to be important in metallic CNTs when the applied bias is larger than the optical phonon energy (~ 160 meV) [41, 43, 44, 65]. In long, metallic CNTs, optical phonon emission causes the current to saturate at ~25 μA, and the data can be explained with an assumed mean free path (mfp) of ~10nm[41]. Recent experiments on metallic nanotubes with varying lengths showed that for short, metallic nanotubes, the current can exceed 25 μA. This data can also be explained by an OP scattering with a mfp of 10-15 nm. (Optical phonon scattering includes both intravelly scattering and intervalley scattering caused by zone boundary phonons [43, 44]). Recently, a CNTFET with channel length of L~50nm delivering nearly the ballistic current has been reported [Jav04b]. This is a surprising result, because under high drain bias the channel length is several mfps long. In this Letter, we show that even, near-ballistic DC currents can be obtained for semiconducting CNTs, even under high source-drain bias in the presence of significant inelastic scattering, which is different for metallic CNTs

Scattering in bulk CNTs [70, 71] and in metallic tubes [43] has been previously explored using semiclassical Monte Carlo simulations. In this letter, we report a simulation study of the role of phonon scattering in CNT field-effect transistors (FETs). We show that elastic scattering has a small effect on source-drain current for a short-channel CNTFET when the mfp is long, as it typically is for CNTs. If, however, a short

elastic mfp were to exist, then elastic scattering would degrade the on-current of CNTFETs much more severely than it does for Si metal-oxide-semiconductor (MOS) FETs. (This difference results from the difference in 1D transport in the CNT vs. 2D in the MOSFET.) Under high drain bias, the mfp for optical phonon emission is short, so there is significant scattering even in a short channel CNTFETs. Our results show, however, that OP scattering has little effect on the DC current in a CNTFET at modest gate voltage, (which results in a ballistic current lower than the saturation current set by the OP energy, ~25μA). This work clarifies the distinctly different roles of elastic and inelastic scattering on CNTFET on-current, as well as the different roles of elastic scattering in CNTFETs and Si MOSFETs. The insights obtained here agree with recent experiments [14], and may be applicable to other nanotube / nanowire FETs [18].

6.2 Approach

To compute the I-V characteristics of a CNTFET in presence of scattering, we performed semiclassical Monte-Carlo simulations. The validity of a semiclassical approach was first confirmed by simulations with a ballistic, quantum transport model [32]. The Monte Carlo (MC) simulation solves the Boltzmann transport equation (BTE) in a stochastic way, and has been extensively applied to study the dissipative carrier transport in Si MOSFETs [45, 72]. Monte Carlo simulation has also been previously applied to study carrier transport in bulk semiconducting tubes [70, 71] and in metallic tubes [43]. Simulations for varying metallic tube lengths agree with the experiment for tube lengths ranging from several-hundred nanometers down to ~50nm [43]. Although there are a large number of phonon bands in carbon nanotubes [15], only a few of them scatter carriers in the lowest electron subband [73]. Following the treatment of phonon scattering in metallic nanotubes, which excellently reproduced experimental measurements [41, 43], two scattering mechanisms were included: acoustic phonon scattering and optical phonon scattering (including both OP intravally valley and intervalley scattering due to zone boundary phonons). To perform these simulations, we used a simple description of CNT band structure and carrier-phonon scattering rate.

We describe the first conduction (or valence) subband of the semiconducting channel with a simple, analytical expression derived from a p_z orbital tight binding model [17],

$$E = \hbar \upsilon_F \sqrt{k^2 + k_0^2}\ , \tag{6.1a}$$

where $\hbar$ is the Planck constant, $\upsilon_F \approx 1.0 \times 10^8 cm/S$ is the Fermi velocity in metallic tubes, and $k_0 = 2/(3d)$, where d is the diameter of the tube. The corresponding density-of-states (DOS) is

$$D(E) = D_0 \cdot \frac{|E|}{\sqrt{E^2 - \Delta^2}} \mathrm{sgn}(|E| - \Delta)\ , \tag{6.1b}$$

where $D_0 = 8/(2\pi\hbar\upsilon_F)$ is the constant DOS of a metallic tube, and Δ is one half of the semiconducting tube bandgap.

In metallic tubes, the scattering rates and mfps are energy-independent due to the constant DOS near the Fermi level, but in a semiconducting tube, the mfp depends on the position-dependent kinetic energy of carriers (reflecting the 1D DOS) and is, therefore, position-dependent. Note that at high kinetic energies, the semiconducting E-k in eqn. (1) approaches that of metallic tubes and the mfps become constant. The scattering mfps in metallic tubes have been previously obtained by fitting the experimentally measured I-V for various tube lengths [41, 43, 44]. We take the values from those experimental fits and use them as nominal values for the scattering mfps of a semiconducting tube at high kinetic energies in CNTFET simulations. Acoustic phonon (AP) scattering is assumed to be elastic with a mfp at high energies of $\lambda_{elastic}^{high} = 500\,\mathrm{nm}$. For optical phonon, we assume $\hbar\omega_{OP} = 0.16\,\mathrm{eV}$ and a scattering mfp of $\lambda_{OP}^{high} = 15\,\mathrm{nm}$ [41, 43]. Because the OP energy is much larger than k_BT, only OP emission is considered. (Hot phonon effects are not treated in this study) Fig. 1 plots the electron scattering rate as a function of energy for the nominal values of the scattering mfps. To explore the role of scattering in CNTFETs,

the mfps will be varied about their nominal values. (When we quote a value for the mfp in the CNTFET simulation, we will quote the high energy mfp, which corresponds to the constant, metallic mfp.)

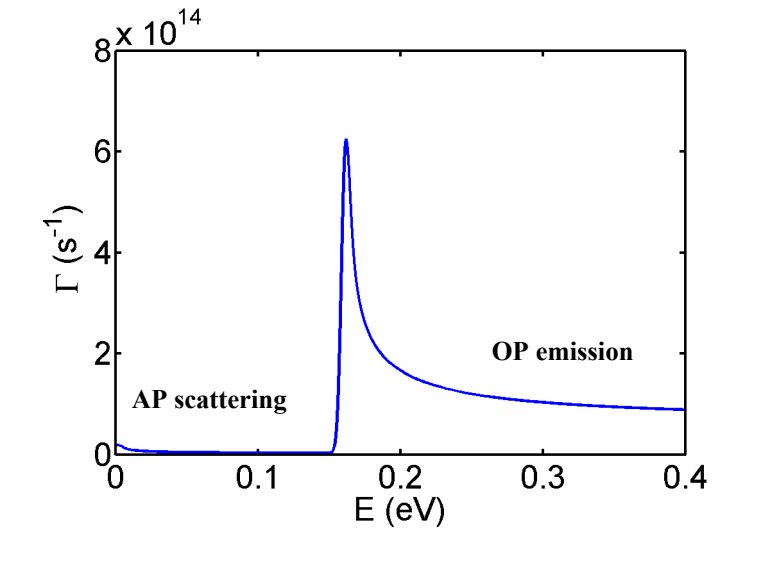


Fig. 6.1. The scattering rate vs. carrier kinetic energy in the lowest subband. The elastic scattering mfp at high energies is 500nm and the OP scattering mfp at high energies is 15nm. The OP energy is 0.16eV. Only scattering within the lowest semiconducting band is included.

Pauli blocking is an important factor that is treated using a rejection technique as described by Lugli [74]. The carrier distribution function is updated after each time step, so that when a scattering event occurs, the probability of whether a scattering final state is available can be evaluated. A random number between 0 and 1 then determines whether the scattering is permitted. To treat transistor electrostatics, Poisson's equation is solved self-consistently with the transport simulation. Two types of contacts are treated. For hypothetical MOSFET-like CNTFETs with doped tubes as source/drain [55] as shown in Fig. 2a, the contacts are assumed to be ideal (without reflection). For Schottky barrier (SB) CNTFETs [7, 8], we treat the tunneling of carriers through metal-CNT junctions as follows. For a carrier injected to the Schottky barrier at an energy, the transmission

probability through the SB at the energy is evaluated using the WKB approximation. A random number between 0 and 1 is then generated to determine whether the carrier tunnels through the SB or get reflected. Such an approach has been previously developed to treat Si SBFETs, and validated by experiments for a channel length down to ~27nm [75].

6.3 Results

We first simulated a MOSFET-like CNTFET with doped tubes as source/drain extensions, as shown in the inset of Fig. 2a. The doped S/D length is 15nm, and the intrinsic, gated tube length is 20nm, which results in a total tube length of 50nm. The solid line shows the ballistic I_D vs. V_D characteristics at high gate voltages (V_G=*0.4*V). To explore the role of elastic scattering in CNTFETs, we first included only elastic scattering in the CNT channel. Two different mfps were assumed: (i) a long elastic mfp, $\lambda_{elastic}^{high} = 500\,\text{nm}$, which is typical for a tube. (Note that $\lambda_{elastic}^{high} = 500\,\text{nm}$ is the mfp at high kinetic energies, and that the thermal average mfp near the top of the barrier is $\langle \lambda_{elastic} \rangle \sim 90\,\text{nm}$), and (ii) a short elastic mfp, $\lambda_{elastic}^{high} = 15\,\text{nm}$, which is the value expected for OP scattering in CNTs, (and corresponds to a thermal average mfp of $\langle \lambda_{elastic} \rangle \sim 3\,\text{nm}$ near the top of the barrier). The dashed line with circles in Fig. 2b shows the I-V characteristic for case (i). For a long elastic mfp with $\lambda_{elastic}^{high} = 500\,\text{nm}$ and $\langle \lambda_{elastic} \rangle \sim 90\,\text{nm}$, the transistor's on-current is ~80% of the ballistic on current, and elastic scattering has a small effect. The solid line with circles in Fig. 2b shows the I-V characteristic for case (ii) with the short elastic mfp. The transistor on-current is significantly degraded to only ~10% of the ballistic limit by short-mfp, elastic scattering.

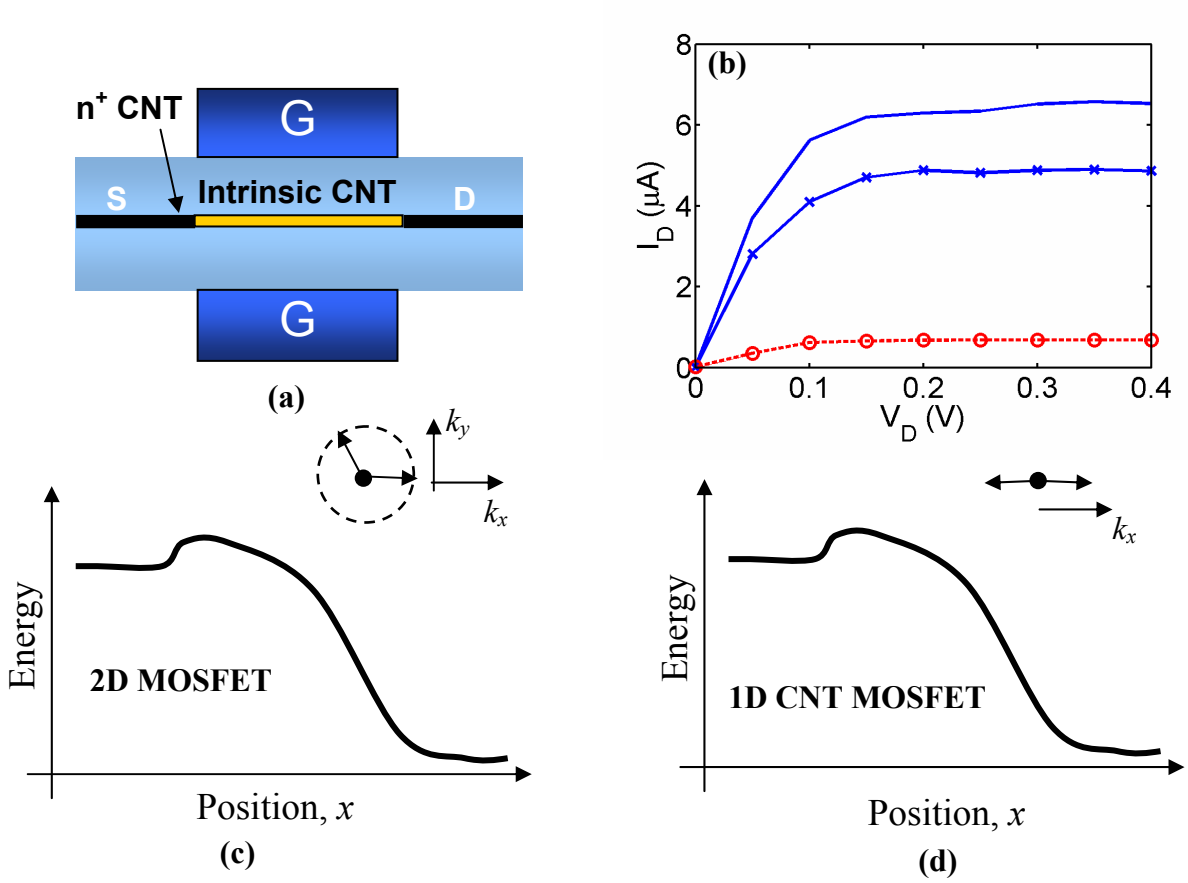


Fig. 6.2 Elastic scattering in CNTFETs. (a) A coaxially gated CNTFET with doped tubes as source/drain. The total tube length is 50nm. The intrinsic channel length is 20nm, and the doped source/drain length is 15nm. The S/D doping density of $N_D = 10^7$ /cm. The gate oxide thickness is 3nm, and the dielectric constant is 16 (for HfO_2). The tube diameter is *d~1.4*nm with a band gap E_g*~0.6*eV. (b) The simulated I_D vs. V_D characteristics at V_G*=0.4*V. The solid line is the ballistic current. The solid line with crosses is for a long elastic mfp ($\lambda_{elastic}^{high} = 500\,\text{nm}$), and the dashed line with circles is for a short elastic mfp ($\lambda_{elastic}^{high} = 15\,\text{nm}$). (c) Scattering in a 2D MOSFET channel. The final states of an elastic scattering event distribute around a circle in a 2D k-space. (d) Scattering in a 1D CNT channel. For an elastic scattering event, there is only 1 final *k*-state, which satisfies $k_{final} = -k_{initial}$

To understand why elastic scattering with a short mfp has a severe effect on the ballistic current of CNTFETs, we can compare it to the role of elastic scattering in Si MOSFETs. As shown in Fig. 2c, for scattering near the drain end of the channel of a MOSFET, the final *k*-states distribute around a circle because the channel is 2D. For most final states, carriers don not posses enough backward velocity along the channel direction to overcome the barrier and return to the source. For this reason, scattering near the drain is less important than scattering near the source end of the channel [76, 77]. (Of course, scattering near the drain causes the space charge density to build up, which has an indirect, though potentially strong, effect on the current [46]). The result is that Si MOSFETs with a channel length several times longer than the elastic scattering mfp (due to, for example, surface roughness scattering) can still operate at nearly one-half of the ballistic limit [66, 78]. For CNTFETs, however, elastic scattering anywhere in the channel has a strong effect because the channel is one-dimensional so scattering anywhere in the channel just reverses the direction of the carrier velocity and allows them to return to the source over the top of the barrier as shown in Fig. 2d. The important role of short mfp elastic scattering in CNTFETs is due to the different channel geometry. Similar effects are expected for other nanotube/nanowire FETs, so it is important to suppress elastic scattering, such as surface roughness scattering, in 1D-channel nanoelectronic FETs. Compared to Si or Ge nanowire transistors, an advantage of CNTFETs is the lack of dangling bonds and surface roughness scattering at the tube/gate oxide interface.

Because OP scattering has a much shorter mfp than elastic scattering in CNTs, it's important to explore the role of OP scattering. Fig. 3a shows I_D vs. V_D at the ballistic limit (the solid line), as well as the characteristics in presence of only elastic scattering with $\lambda_{elastic}^{high} = 500\,\text{nm}$ (the dashed line) and with both elastic scattering and OP scattering with $\lambda_{OP}^{high} = 15\,\text{nm}$ (the solid line with circles). It is interesting to note that although the OP scattering mfp is much shorter than that of elastic scattering and the tube is several times longer than the OP mfp, OP scattering has little effect on the transistor current even.

To understand this, we plotted out a snapshot of the steady-state carrier distribution in position and energy in the presence of elastic and OP scattering (Fig. 3b). For carriers injected from source, Pauli exclusion suppresses OP emission in the region before the top of the barrier. As shown in Fig. 3b, the states below the top of the barrier at the source end of the tube are filled according to the source Fermi level and are full. If OP emission were to occur, carriers would lose an energy of $\hbar\omega_{OP}$, but the final states are occupied. Such scattering, therefore, is prohibited by Pauli exclusion. After carriers travel over the top of the barrier, OP scattering can occur, but such OP emission lowers the carrier kinetic energy by $\hbar\omega_{OP}$ (~0.16eV). At modest gate bias, the top of the barrier is only modestly below the energy of source-injected carriers. After scattering, carriers do not have enough energy to overcome the barrier and return to the source. Although OP scattering occurs near the drain, it has little effect on DC source-drain current.

We also explored the role of phonon scattering under very large gate voltages. As schematically shown in Fig. 4, the top of the potential barrier, E_{top}, is pushed below the source Fermi level by a value larger than the OP energy ($\hbar\omega_{OP}$) and the ballistic source-drain current is larger than the saturation current set by the OP energy (~25μA). Under this condition, OP scattering affects the source-drain current and tends to set a saturation current of ~25μA. Because the channel length is still several OP scattering mfp long, carriers injected at high energies ($E > E_{top} + \hbar\omega_{OP}$ as shown in Fig. 4) can emit an OP and return back to the source. As the result, for the energy range of $E_{top} < E < E_{top} + \hbar\omega_{OP}$, near ballistic current is delivered, but for the energy range of $E > E_{top} + \hbar\omega_{OP}$, the current is significantly reduced by OP emission.

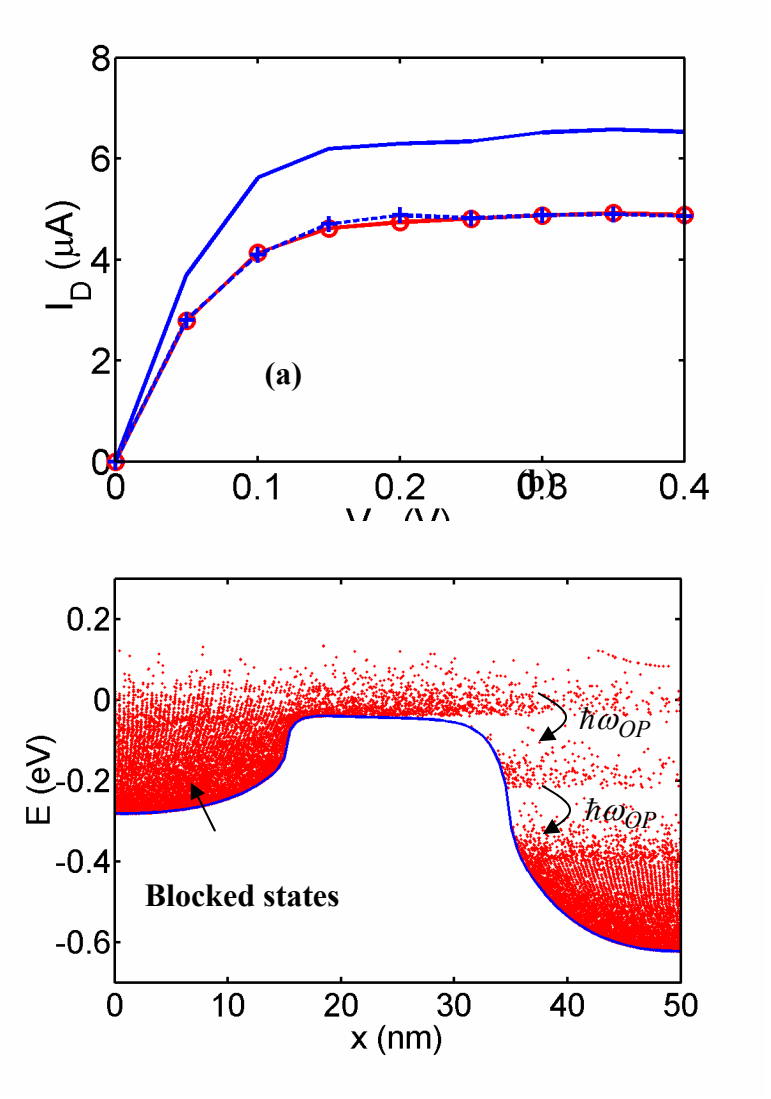


Fig. 6.3. (a) I_D vs. V_D at V_G=0.4V for the CNTFET as shown in Fig. 6.1a. The solid line shows the ballistic characteristics. The solid line with circles is the characteristic in presence of both OP scattering (with mfp $\lambda_{OP}^{high} = 15\,\text{nm}$) and elastic scattering ($\lambda_{elastic}^{high} = 500\,\text{nm}$). The dashed line with crosses shows the I-V in the presence of only elastic scattering ($\lambda_{elastic}^{high} = 500\,\text{nm}$), and without OP scattering. (b) The steady-state electron distribution at $V_D = V_G = 0.4\,\text{V}$. The OP scattering mfp is $\lambda_{OP}^{high} = 15\,\text{nm}$ and the AP scattering mfp is $\lambda_{elastic}^{high} = 500\,\text{nm}$. The solid line shows the first conduction band edge.

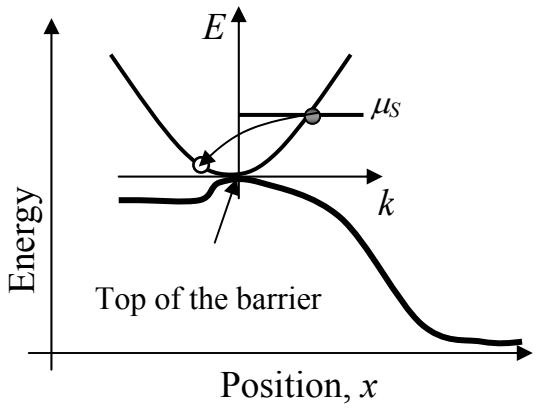


Fig. 6.4. OP scattering at high gate overdrive. The top of the barrier, E_{top} , is pushed well below the source Fermi level, μ_S. ($\mu_S - E_{top} > \hbar\omega_{OP}$). The source-injected carriers with high energies ($E > \hbar\omega_{OP} + E_{top}$) can emit an optical phonon, and still have enough energy to return back to the source. So the ballistic DC current delivered in the energy range of $E > \hbar\omega_{OP} + E_{top}$ can be significantly reduced by OP scattering.

Since many CNTFETs to date operate like SB transistors [7, 8], it is important to examine whether our understanding of MOSFET-like CNTFETs is relevant to SB CNTFETs. The simulated CNTFET are similar to a recently reported CNTFET with Pd contacts [14], except that a simplified, coaxial gate geometry is used in the simulation. As shown in the inset of Fig. 5a, the simulated FET has an intrinsic, 50nm-long tube channel and a metal/CNT SB height of $\phi_b = 0$. The main panel of Fig. 5a plots the ballistic I_D vs. V_D at V_G=*0.4*V (the solid line) and the characteristic in presence of both elastic scattering ($\lambda_{elastic}^{high} = 500\,\text{nm}$) and OP scattering ($\lambda_{OP}^{high} = 15\,\text{nm}$) (the dashed line with circles). Again, we find that the FET operates ~80% at the ballistic limit, and scattering only has a small effect on the DC source-drain current at modest gate voltages. The elastic scattering has a small effect because its mfp is long. Although the mfp for OP scattering is several times shorter than the channel length, it still has little effect on the source-drain DC current. As shown in Fig. 5b, energetic carriers can emit an OP and get back scattered, but after losing a optical phonon energy of ~160meV, (which is much larger than the optical

phonon energy in Si and GaAs,) the backscattered carrier faces a much thicker and higher SB. Since the tunneling probability exponentially decreases with the SB thickness, the chance for the scattered carrier to go back to source greatly reduces. OP scattering, therefore, has little effect on the source-drain current at modest gate biases, due to the mechanism of losing energy after scattering. Our theoretical results are consistent with a recent experiment for an ultra-short channel CNTFET (with L_{ch} ~50nm) [14]. Although the experimental FET has a channel length several times longer than the OP scattering mfp, the transistor still delivers a near ballistic DC current.

The negligible effect of OP emission in CNTFETs at modest gate biases is due to the mechanism of losing a large OP energy ($\hbar\omega_{OP} \sim 160$ meV) after scattering. In both MOSFET-like CNTFETs and SB CNTFETs, losing a large energy prevents scattered carrier to overcome (or to tunnel through) the potential barrier in the channel and to return back to the source. OP scattering begins to play an important role when the source-drain current is delivered over an energy range larger than $\hbar\omega_{OP}$ at large gate overdrives. Compared to CNTFETs, optical phonons in Si have a much smaller energy (10-60meV), and therefore, are expected to affect the source-drain current at much lower gate overdrives in Si MOSFETs or Si nanowire transistors. The large OP energy in CNTs, therefore is another advantage for delivering more current.

Only intrasubband scattering is treated in this study. At high drain voltages, carriers in the lowest subband can gain enough kinetic energy near the drain end of the channel, and experience inter-subanband scattering. After a carrier is scattered to a higher subband, the potential barrier to prevent the carrier to return back to the source increases, and it becomes more difficult for the carrier to return back to the source. Intersubband scattering, therefore, won't change the conclusion that phonon scattering only has a small effect on the DC source-drain current in the simulated CNTFETs.

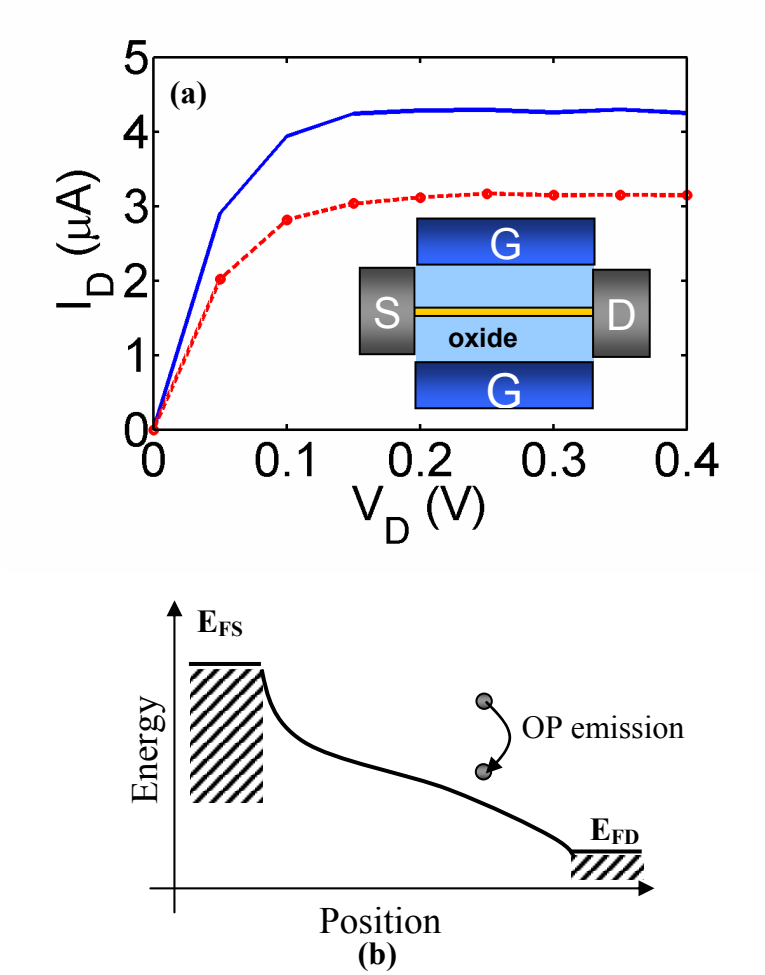


Fig. 6.5. (a) I_D vs. V_D characteristics at V_G=*0.4*V. The solid line shows the ballistic current. The circled line shows the characteristic in presence of elastic scattering (with $\lambda_{elastic}^{high} = 500\,\text{nm}$) and OP scattering (with mfp $\lambda_{OP}^{high} = 15\,\text{nm}$). The inset shows the simulated, coaxially gated CNTFET. The SB height is $\phi_b = 0$, and the intrinsic tube channel length is 50nm. The tube diameter is d~1.4nm with a band gap E_g~0.6eV. The gate oxide thickness is 8nm and the dielectric constant is 16. (b) The band profile of a SB CNTFET with a 0 Schottky barrier at on state. A source-injected carrier can emit an OP near drain, but after losing an energy of $\hbar\omega_{OP}$, the carrier faces a much thicker SB and can hardly tunnels back to the source.

6.4 Conclusions

In summary, role of phonon scattering in CNTFETs was explored using semiclassical Monte Carlo simulations. The results indicate that if elastic scattering with a short mfp exists, it has a much more severe effect on the DC current of a CNTFET than it does for a Si MOSFET. The absence of surface roughness scattering in CNTs even after depositing a high-κ gate insulator, therefore, is an important advantage for CNTFETs. Although OP scattering in CNTs has a short mfp and scatters carriers near the drain end of the channel, it has little effect on the source-drain current for both MOSFET-like CNTFETs and SB CNTFETs at modest gate biases.

7. CONCLUSIONS

Since the first demonstration of carbon nanotube transistor ~5-6 years ago[3, 4], the understanding of carbon nanotube transistor is evolving and the performance of the transistor is improving very rapidly [5, 6]. Carbon nanotube transistors with performance greatly exceeding the state-of-the-art Si MOSFETs have been recently demonstrated [14]. The purposes of this work are to develop simulation approach and tools for carbon nanotube electronics, and to use simulation to understand device physics and explore device issues, which are crucial for improving device performance. It is describe in detail as the following.

(1) Developing simulation approach for carbon nanotube transistors. Simulations have been demonstrated useful to understand the device physics, explain experiments, and explore issues, which still can not be explored experimentally. The unique carrier transport properties and electrostatic geometry makes conventional TCAD tools not applicable to carbon nanotube transistors. Developing new simulation approach and tools, therefore, becomes necessary. Guided and verified by experimental work through an on-going theory-experiment collaboration, we developed self-consistent quantum simulation approach for ballistic carbon nanotube transistors [32]. This simulator solves Schrödinger equation with an atomistic description of the nanotube channel with the non-equilibrium Green's function (NEGF) formalism, in the presence of self-consistent electrostatic field. We also investigated the role of phonon scattering in carbon nanotubes by developing a semiclassical Monte-Carlo simulator, which is coupled to quantum equations at metal/CNT contacts, for describing the quantum-mechanical tunneling through the Schottky barriers. The self-consistent quantum simulator allows us to treat the quantum-mechanical effects in the ballistic

transistor in detail, while the semiclassical Monte-Carlo simulator can be readily extended to rigorously treat various scattering mechanisms in detail.

(2) Understand device physics and explore device issues. The ultimate goal of doing device simulation is to provide physical understanding of experiments and to suggest approaches for improving device performance. In chapter 4, we did a comprehensive study of the scaling behaviors for ballistic SB CNTFETs [9]. We show that ambipolar conduction is important for SB CNTFETs. The factors that control the feature of the ambipolar I-V characteristics are explored in detail. The transistor scaling behaviors with respect to the gate oxide thickness and dielectric constant, Schottky barrier height, tube diameter and power supply voltages, are investigated. In chapter 5, we analyzed a short-channel, high-performance CNTFET [14] in detail. Numerical simulations are used to understand what controls and how to further improve the transistor performance. In chapter 6, a Monte-Carlo simulator is used to investigate the role of phonon scattering in CNTFETs, which turn out to be quite different from that in Si MOSFETs.

Numerical simulations developed in this study are also directly applied to experiments through an on-going experiment-theory collaboration for the following problems, (which are not described in this thesis but available in the cited references). i) CNTFETs with thin high-κ top gate insulator [11], ii) CNTFETs with excellent metal contacts [13], iii) phonon scattering in metallic CNTs [43], iv) MOSFET-like CNTFETs for suppressing ambipolar conduction [55], and v) self-aligned, short channel CNTFETs [14].

This work can be extended in the following ways.

(1) Device physics. A preliminary 3D Poisson solver is developed in Chapter 5 to analyze the experiments [14]. In order to efficiently and generally investigate the electrostatic control of the transistor, it is important to develop a general

3D Poisson solver, which runs efficiently and couples to the carrier transport solver nicely. Following the work in Chapter 6, more work is necessary to understand scattering mechanisms in carbon nanotubes, such as computing the scattering mean free path and calibrating it to experimental measurement, and understanding phonon transport and the role of hot phonons on electronic carrier transport.

(2) Circuit performance. After achieving understanding of a single transistor, it is important to think about how to put many CNTFETs together and assess the circuit performance. A SPICE model, which describes the behavior of experimental CNTFETs, needs to be developed and applied to investigate the performance such as delay and power for simple nanotube digital circuits. Another possibility of using CNTFETs for RF applications should also be investigated [79, 80]. Developing both simple circuit models, and detailed, time-domain device simulators are useful for assessing the RF performance potential of CNTFETs, such as cut-off frequency and Y-parameters.

(3) New application. Recent work by the IBM group demonstrates that carbon nanotubes can emit light as a direct band gap semiconductor [21, 58, 59]. It offers the chance to do both electronics and optics on one type of nanomaterial, carbon nanotubes. The possibility for optical applications should be investigated as well. Furthermore, chemical and biological sensors based on carbon nanotubes have also been demonstrated. Such sensors promise to be more sensitive, cheaper, and faster than conventional sensors used in chemistry [81, 82] and biology [83-86]. Since sensing is typically achieved by changing the electrical conduction of the nanotube, work on carbon nanotube electronics can be used as a starting point for modeling nanotube sensors.

LIST OF REFERENCES

[1] S. Iijima, "Helical Microtubules of Graphitic Carbon," *Nature*, vol. 354, pp. 56-58, 1991.

[2] R. H. Baughman, A. A. Zakhidov, and W. A. de Heer, "Carbon nanotubes - the route toward applications," *Science*, vol. 297, pp. 787-792, 2002.

[3] R. Martel, T. Schmidt, H. R. Shea, T. Hertel, and P. Avouris, "Single- and multi-wall carbon nanotube field-effect transistors," *Applied Physics Letters*, vol. 73, pp. 2447-2449, 1998.

[4] 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.

[5] P. Avouris, J. Appenzeller, R. Martel, and S. J. Wind, "Carbon nanotube electronics," *Proceedings of the Ieee*, vol. 91, pp. 1772-1784, 2003.

[6] P. L. McEuen, M. S. Fuhrer, and H. K. Park, "Single-walled carbon nanotube electronics," *Ieee Transactions on Nanotechnology*, vol. 1, pp. 78-85, 2002.

[7] J. Appenzeller, J. Knoch, V. Derycke, R. Martel, S. Wind, and P. Avouris, "Field-modulated carrier transport in carbon nanotube transistors," *Physical Review Letters*, vol. 89, pp. 126801.1-126801.4, 2002.

[8] S. Heinze, J. Tersoff, R. Martel, V. Derycke, J. Appenzeller, and P. Avouris, "Carbon nanotubes as Schottky barrier transistors," *Physical Review Letters*, vol. 89, pp. 106801.1-106801.4, 2002.

[9] J. Guo, S. Datta, and M. Lundstrom, "A numerical study of scaling issues for Schottky-Barrier carbon nanotube transistors," *Ieee Transactions on Electron Devices*, vol. 51, pp. 172-177, 2004.

[10] M. Radosavljevic, S. Heinze, J. Tersoff, and P. Avouris, "Drain voltage scaling in carbon nanotube transistors," *Applied Physics Letters*, vol. 83, pp. 2435-2437, 2003.

[11] A. Javey, H. Kim, M. Brink, Q. Wang, A. Ural, J. Guo, P. McIntyre, P. McEuen, M. Lundstrom, and H. J. Dai, "High-kappa dielectrics for advanced carbon-nanotube transistors and logic gates," *Nature Materials*, vol. 1, pp. 241-246, 2002.

[12] S. J. Wind, J. Appenzeller, R. Martel, V. Derycke, and P. Avouris, "Fabrication and electrical characterization of top gate single-wall carbon nanotube field-effect transistors," *Journal of Vacuum Science & Technology B*, vol. 20, pp. 2798-2801, 2002.

[13] A. Javey, J. Guo, Q. Wang, M. Lundstrom, and H. J. Dai, "Ballistic carbon nanotube field-effect transistors," *Nature*, vol. 424, pp. 654-657, 2003.

[14] A. Javey, J. Guo, D. B. Farmer, Q. Wang, E. Yenilmez, R. G. Gordon, M. Lundstrom, and H. Dai, "Self-aligned ballistic molecular transistors and parallel nanotube arrays," *Nano. Lett.*, in press, 2004.

[15] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, *Physical properties of carbon nanotubes*. London: Imperial College Press, 1998.

[16] L. Yang, M. P. Anantram, J. Han, and J. P. Lu, "Band-gap change of carbon nanotubes: Effect of small uniaxial and torsional strain," *Physical Review B*, vol. 60, pp. 13874-13878, 1999.
[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] Y. Cui, Z. H. Zhong, D. L. Wang, W. U. Wang, and C. M. Lieber, "High performance silicon nanowire field effect transistors," *Nano Letters*, vol. 3, pp. 149-152, 2003.
[19] Y. Cui and C. M. Lieber, "Functional nanoscale electronic devices assembled using silicon nanowire building blocks," *Science*, vol. 291, pp. 851-853, 2001.
[20] D. W. Wang, Q. Wang, A. Javey, R. Tu, H. J. Dai, H. Kim, P. C. McIntyre, T. Krishnamohan, and K. C. Saraswat, "Germanium nanowire field-effect transistors with SiO2 and high-kappa HfO2 gate dielectrics," *Applied Physics Letters*, vol. 83, pp. 2432-2434, 2003.
[21] X. F. Duan, Y. Huang, R. Agarwal, and C. M. Lieber, "Single-nanowire electrically driven lasers," *Nature*, vol. 421, pp. 241-245, 2003.
[22] F. Leonard and J. Tersoff, "Role of Fermi-level pinning in nanotube Schottky diodes," *Physical Review Letters*, vol. 84, pp. 4693-4696, 2000.
[23] F. Leonard and J. Tersoff, "Novel length scales in nanotube devices," *Physical Review Letters*, vol. 83, pp. 5174-5177, 1999.
[24] A. A. Odintsov, "Schottky barriers in carbon nanotube heterojunctions," *Physical Review Letters*, vol. 85, pp. 150-153, 2000.
[25] S. Ramo, J. R. Whinnery, and T. Van Duzer, *Fields and Waves in Communication Electronics*, third ed. New York, NY, USA: John Wiley & Sons, 1993.
[26] Z. Ren, "Ph. D. Thesis," Purdue University, West Lafayette, IN, 47907, 2001.
[27] S. Luryi, "Quantum Capacitance Devices," *Applied Physics Letters*, vol. 52, pp. 501-503, 1988.
[28] D. J. Frank, Y. Taur, and H. S. P. Wong, "Generalized scale length for two-dimensional effects in MOSFET's," *Ieee Electron Device Letters*, vol. 19, pp. 385-387, 1998.
[29] S. Datta, *Electronic transport in mesoscopic systems*. Cambridge ; New York: Cambridge University Press, 1995.
[30] R. Venugopal, Z. B. Ren, and M. S. Lundstrom, "Simulating quantum transport in nanoscale MOSFETs: Ballistic hole transport, subband engineering and boundary conditions," *Ieee Transactions on Nanotechnology*, vol. 2, pp. 135-143, 2003.
[31] A. Rahman, J. Guo, S. Datta, and M. S. Lundstrom, "Theory of ballistic nanotransistors," *Ieee Transactions on Electron Devices*, vol. 50, pp. 1853-1864, 2003.
[32] J. Guo, S. Datta, M. Lundstrom, and M. P. Anantram, "Toward multi-scale simulations of carbon nanotube transistors," *the International Journal on Multiscale Computer Engineering*, vol. in press, 2004.
[33] P. Damle, T. Rakshit, M. Paulsson, and S. Datta, "Current-voltage characteristics of molecular conductors: Two versus three terminal," *Ieee Transactions on Nanotechnology*, vol. 1, pp. 145-153, 2002.

[34] S. Datta, "Nanoscale device modeling: the Green's function method," *Superlattices and Microstructures*, vol. 28, pp. 253-278, 2000.
[35] J. Tersoff, "Schottky-Barrier Heights and the Continuum of Gap States," *Physical Review Letters*, vol. 52, pp. 465-468, 1984.
[36] J. Tersoff, "Reference Levels for Heterojunctions and Schottky Barriers," *Physical Review Letters*, vol. 56, pp. 675-675, 1986.
[37] A. Svizhenko, M. P. Anantram, T. R. Govindan, B. Biegel, and R. Venugopal, "Two-dimensional quantum mechanical modeling of nanotransistors," *Journal of Applied Physics*, vol. 91, pp. 2343-2354, 2002.
[38] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, "Single and multiband modeling of quantum electron transport through layered semiconductor devices," *Journal of Applied Physics*, vol. 81, pp. 7845-7869, 1997.
[39] C. P. Auth and J. D. Plummer, "Scaling theory for cylindrical, fully-depleted, surrounding-gate MOSFET's," *Ieee Electron Device Letters*, vol. 18, pp. 74-76, 1997.
[40] F. Leonard and J. Tersoff, "Dielectric response of semiconducting carbon nanotubes," *Applied Physics Letters*, vol. 81, pp. 4835-4837, 2002.
[41] 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.
[42] Z. Yao, C. Dekker, and P. Avouris, "Electrical transport through single-wall carbon nanotubes," *Carbon Nanotubes*, vol. 80, pp. 147-171, 2001.
[43] A. Javey, J. Guo, M. Paulsson, Q. Wang, D. Mann, M. Lundstrom, and H. J. Dai, "High-field quasiballistic transport in short carbon nanotubes," *Physical Review Letters*, vol. 92, pp.106804.1-106804.4, 2004.
[44] J. Y. Park, S. Rosenblatt, Y. Yaish, V. Sazonova, H. Ustunel, S. Braig, T. A. Arias, P. W. Brouwer, and P. L. McEuen, "Electron-phonon scattering in metallic single-walled carbon nanotubes," *Nano Letters*, vol. 4, pp. 517-520, 2004.
[45] M. V. Fischetti and S. E. Laux, "Monte-Carlo Analysis of Electron-Transport in Small Semiconductor-Devices Including Band-Structure and Space-Charge Effects," *Physical Review B*, vol. 38, pp. 9721-9745, 1988.
[46] R. Venugopal, M. Paulsson, S. Goasguen, S. Datta, and M. S. Lundstrom, "A simple quantum mechanical treatment of scattering in nanoscale transistors," *Journal of Applied Physics*, vol. 93, pp. 5613-5625, 2003.
[47] R. Venugopal, S. Goasguen, S. Datta, and M. S. Lundstrom, "Quantum mechanical analysis of channel access geometry and series resistance in nanoscale transistors," *Journal of Applied Physics*, vol. 95, pp. 292-305, 2004.
[48] A. Raychowdhury, S. Mukhopadhyay, and K. Roy, "Circuit-compatible modeling of carbon nanotube FETs in the ballistic limit of performance," presented at IEEE Conference on Nanotechnology, San Francisco, CA, 2003.
[49] S. Rosenblatt, Y. Yaish, J. Park, J. Gore, V. Sazonova, and P. L. McEuen, "High performance electrolyte gated carbon nanotube transistors," *Nano Letters*, vol. 2, pp. 869-872, 2002.
[50] S. J. Wind, J. Appenzeller, R. Martel, V. Derycke, and P. Avouris, "Vertical scaling of carbon nanotube field-effect transistors using top gate electrodes (vol 80, pg 3817, 2002)," *Applied Physics Letters*, vol. 81, pp. 1359-1359, 2002.

[51] S. Heinze, M. Radosavljevic, J. Tersoff, and P. Avouris, "Unexpected scaling of the performance of carbon nanotube Schottky-barrier transistors," *Physical Review B*, vol. 68, p. 235418, 2003.

[52] 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.

[53] S. J. Wind, J. Appenzeller, and P. Avouris, "Lateral scaling in carbon-nanotube field-effect transistors," *Physical Review Letters*, vol. 91, pp. 058301.1-058301.4, 2003.

[54] J. Kong, E. Yenilmez, T. W. Tombler, W. Kim, H. J. Dai, R. B. Laughlin, L. Liu, C. S. Jayanthi, and S. Y. Wu, "Quantum interference and ballistic transmission in nanotube electron waveguides," *Physical Review Letters*, vol. 8710, pp. art. no.-106801, 2001.

[55] A. Javey, J. Guo, D. B. Farmer, Q. Wang, D. W. Wang, R. G. Gordon, M. Lundstrom, and H. J. Dai, "Carbon nanotube field-effect transistors with integrated ohmic contacts and high-k gate dielectrics," *Nano Letters*, vol. 4, pp. 447-450, 2004.

[56] Semiconductor Industry Association, *International Technology Roadmap for Semiconductors, www.itrs.net*.

[57] S. Datta, "Electrical resistance: an atomistic view," *Nanotechnology*, vol. 15, pp. S433-S451, 2004.

[58] M. Freitag, Y. Martin, J. A. Misewich, R. Martel, and P. H. Avouris, "Photoconductivity of single carbon nanotubes," *Nano Letters*, vol. 3, pp. 1067-1071, 2003.

[59] J. A. Misewich, R. Martel, P. Avouris, J. C. Tsang, S. Heinze, and J. Tersoff, "Electrically induced optical emission from a carbon nanotube FET," *Science*, vol. 300, pp. 783-786, 2003.

[60] J. Guo, J. Wang, E. Polizzi, S. Datta, and M. Lundstrom, "Electrostatics of nanowire transistors," *Ieee Transactions on Nanotechnology*, vol. 2, pp. 329-334, 2003.

[61] A. Javey, M. Shim, and H. J. Dai, "Electrical properties and devices of large-diameter single-walled carbon nanotubes," *Applied Physics Letters*, vol. 80, pp. 1064-1066, 2002.

[62] Y. Taur and T. Ning, *Fundamentals of Modern VLSI Devices*. London, UK: Cambridge University Press, 1998.

[63] J. Appenzeller, M. Radosavljevic, J. Knoch, and P. Avouris, "Tunneling versus thermionic emission in one-dimensional semiconductors," *Physical Review Letters*, vol. 92, pp. -, 2004.

[64] J. Guo and M. S. Lundstrom, "A computational study of thin-body, double-gate, Schottky barrier MOSFETs," *Ieee Transactions on Electron Devices*, vol. 49, pp. 1897-1902, 2002.

[65] D. Mann, A. Javey, J. Kong, Q. Wang, and H. J. Dai, "Ballistic transport in metallic nanotubes with reliable Pd ohmic contacts," *Nano Letters*, vol. 3, pp. 1541-1544, 2003.

[66] A. Lochtefeld and D. A. Antoniadis, "On experimental determination of carrier velocity in deeply scaled NMOS: How close to the thermal limit?," *IEEE Electron Device Letters*, vol. 22, pp. 95-97, 2001.

[67] J. Guo, S. Goasguen, M. Lundstrom, and S. Datta, "Metal-insulator-semiconductor electrostatics of carbon nanotubes," *Applied Physics Letters*, vol. 81, pp. 1486-1488, 2002.

[68] S. Heinze, J. Tersoff, and P. Avouris, "Electrostatic engineering of nanotube transistors for improved performance," *Applied Physics Letters*, vol. 83, pp. 5038-5040, 2003.

[69] Y. M. Lin, J. Appenzeller, and P. Avouris, "Ambipolar-to-unipolar conversion of carbon nanotube transistors by gate structure engineering," *Nano Letters*, vol. 4, pp. 947-950, 2004.

[70] G. Pennington and N. Goldsman, "Semiclassical transport and phonon scattering of electrons in semiconducting carbon nanotubes," *Physical Review B*, vol. 68, p. 045426, 2003.

[71] G. Pennington and N. Goldsman, "Monte Carlo study of electron transport in a carbon nanotube," *Ieice Transactions on Electronics*, vol. E86C, pp. 372-378, 2003.

[72] K. Hess, *Monte Carlo device simulation : full band and beyond*. Boston: Kluwer Academic, 1991.

[73] G. D. Mahan, "Electron-optical phonon interaction in carbon nanotubes," *Physical Review B*, vol. 68, p. 125409, 2003.

[74] C. Jacoboni and P. Lugli, *The Monte Carlo method for semiconductor device simulation*. Wien ; New York: Springer-Verlag, 1989.

[75] B. Winstead and U. Ravaioli, "Simulation of Schottky barrier MOSFET's with a coupled quantum injection/Monte Carlo technique," *Ieee Transactions on Electron Devices*, vol. 47, pp. 1241-1246, 2000.

[76] M. Lundstrom and Z. B. Ren, "Essential physics of carrier transport in nanoscale MOSFETs," *Ieee Transactions on Electron Devices*, vol. 49, pp. 133-141, 2002.

[77] M. Lundstrom, "Elementary scattering theory of the Si MOSFET," *Ieee Electron Device Letters*, vol. 18, pp. 361-363, 1997.

[78] F. Assad, Z. B. Ren, D. Vasileska, S. Datta, and M. Lundstrom, "On the performance limits for Si MOSFET's: A theoretical study," *Ieee Transactions on Electron Devices*, vol. 47, pp. 232-240, 2000.

[79] J. Appenzeller and D. J. Frank, "Frequency dependent characterization of transport properties in carbon nanotube transistors," *Applied Physics Letters*, vol. 84, pp. 1771-1773, 2004.

[80] D. J. Frank and J. Appenzeller, "High-frequency response in carbon nanotube field-effect transistors," *Ieee Electron Device Letters*, vol. 25, pp. 34-36, 2004.

[81] Q. F. Pengfei, O. Vermesh, M. Grecu, A. Javey, O. Wang, H. J. Dai, S. Peng, and K. J. Cho, "Toward large arrays of multiplex functionalized carbon nanotube sensors for highly sensitive and selective molecular detection," *Nano Letters*, vol. 3, pp. 347-351, 2003.

[82] S. Peng, K. J. Cho, P. F. Qi, and H. J. Dai, "Ab initio study of CNT NO2 gas sensor," *Chemical Physics Letters*, vol. 387, pp. 271-276, 2004.

[83] R. J. Chen, S. Bangsaruntip, K. A. Drouvalakis, N. W. S. Kam, M. Shim, Y. M. Li, W. Kim, P. J. Utz, and H. J. Dai, "Noncovalent functionalization of carbon nanotubes for highly specific electronic biosensors," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, pp. 4984-4989, 2003.

[84] K. Besteman, J. O. Lee, F. G. M. Wiertz, H. A. Heering, and C. Dekker, "Enzyme-coated carbon nanotubes as single-molecule biosensors," *Nano Letters*, vol. 3, pp. 727-730, 2003.

[85] Y. Cui, Q. Q. Wei, H. K. Park, and C. M. Lieber, "Nanowire nanosensors for highly sensitive and selective detection of biological and chemical species," *Science*, vol. 293, pp. 1289-1292, 2001.

[86] J. Hahm and C. M. Lieber, "Direct ultrasensitive electrical detection of DNA and DNA sequence variations using nanowire nanosensors," *Nano Letters*, vol. 4, pp. 51-54, 2004.

Appendix A. The source/drain self-energies in real space

The overall size of the self-energy matrices for the source and drain contacts is the same as the Hamiltonian matrix for the channel, but the self-energy matrices are highly sparse. For example, only one carbon ring at the source end of the channel couples to the source, thus only one submatrix, the (1,1) submatrix in the basis used for eqn. (4), is non-zero for the source self energy, Σ_1. Similarly, only one submatrix is non-zero for the drain self-energy, Σ_2. The non-zero entry of the self-energies can be computed by a recursive relation for the surface Green's function, with details explained in the appendix of [30]. Here we compute the self-energy for a semi-infinite nanotube source. The self-energy approach can be readily extended to treat any type of contacts, for example, metal-nanotube contacts, as will be discussed later.

Fig. A1 shows how carbon rings are coupled for a semi-infinite nanotube source. Each circle (triangle) represents a carbon ring consisting of A(B)-type carbon atoms. The carbon ring couples to the nearest ring, with a coupling matrix of β_1 or β_2, and g_m is the surface Green's function for the *m*th ring in the source extension, ordered from the source/channel interface. The recursive relation [30] relates the surface Green's functions,

$$g_m = [(E + i0^+)I - \alpha_m - \tau g_{m+1}\tau^+]^{-1}, \tag{A1}$$

where τ is the coupling matrix between the *m*th and the *(m+1)*th carbon rings and α_m is the Hamiltonian matrix of the *m*th ring. Applying this recursive relation to the nanotube in Fig. A1, we get

$$g_1 = [(E + i0^+)I - \alpha_1 - \beta_2^+ g_2 \beta_2]^{-1} \quad g_2 = [(E + i0^+)I - \alpha_2 - \beta_1^+ g_3 \beta_1]^{-1} \tag{A2}$$

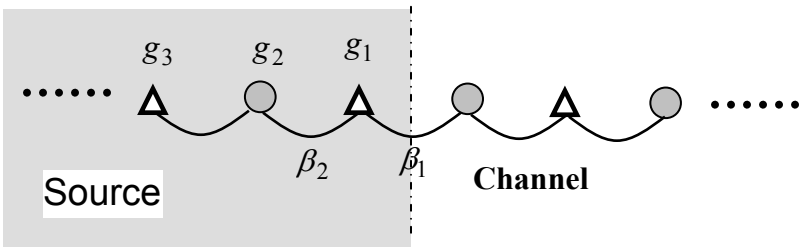


Fig. A1. Computing the source self-energy for a zigzag nanotube. The circles represent A-type carbon rings and the triangles represent B-type carbon rings. g_i is the surface Green's function for the *i*th carbon ring inside the source. β_1 (β_2) is the first (second) kind coupling matrix between neighboring rings, as described in the text.

Note that the potential is invariant inside the source, so $\alpha_1 = \alpha_2$. Furthermore, $g_1 = g_3$ due to the periodicity of the nanotube lattice. Using these relations, eqn. (A2) becomes two coupled matrix equations with two unknowns, g_1 and g_2. The surface Green's function can be numerically solved from Eqn (A2). The non-zero submatrix of the source self-energy matrix is $\Sigma_S^{1,1} = \beta_1 g_1 \beta_1^+$, where the superscript denotes that it is the (1,1) submatrix of the overall source self-energy matrix. The self-energy for the drain contact can be computed in a similar way.

Appendix B. The transistor Hamiltonian in mode space

The following basis transformation, which transforms the real space basis around the nanotube to a mode space basis, is performed to the Hamiltonian matrix in the real space,

$$H' = \begin{bmatrix} V & & & & & \\ & V & & & & \\ & & V & & & \\ & & & V & & \\ & & & & V & \\ & & & & & \ddots \end{bmatrix}^{+} \begin{bmatrix} \alpha_1 & \beta_2^{+} & & & & \\ \beta_2 & \alpha_2 & \beta_1 & & & \\ & \beta_1 & \alpha_3 & \beta_2 & & \\ & & \beta_2^{+} & \alpha_4 & \beta_1 & \\ & & & \beta_1 & \alpha_5 & \ddots \\ & & & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} V & & & & & \\ & V & & & & \\ & & V & & & \\ & & & V & & \\ & & & & V & \\ & & & & & \ddots \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_1' & \beta_2'^{+} & & & & \\ \beta_2' & \alpha_2' & \beta_1' & & & \\ & \beta_1' & \alpha_3' & \beta_2' & & \\ & & \beta_2'^{+} & \alpha_4' & \beta_1' & \\ & & & \beta_1' & \alpha_5' & \ddots \\ & & & & \ddots & \ddots \end{bmatrix} \tag{B1}$$

with

$$\alpha_i' = V^{+}\alpha_i V\,,$$

$$\beta_1' = V^{+}\beta_1 V\,, \text{ and}$$

$$\beta_2' = V^{+}\beta_2 V\,, \tag{B2}$$

where V is the transform matrix from the real space basis of a carbon atom ring to the mode space basis. Our purpose is to uncouple the modes after the basis transformation, *i.e.*, to make the Hamiltonian matrix elements between different modes equal to zero. This requires that after the transformation, α_i', β_1', and β_2', become diagonal matrices.

Notice that α_i or β_1 is a constant times an identify matrix. These matrices remain unchanged and diagonal after any basis transformation,

$$\begin{aligned} \alpha_i^{'} &= \alpha_i = U_i I \\ \beta_1^{'} &= \beta_1 = tI \end{aligned} . \tag{B3}$$

The problem now becomes to find out the eigenvectors and corresponding eigenvalues for

$$\beta_2 = t\begin{bmatrix} 1 & & \cdots & 1 \\ 1 & 1 & & \\ & 1 & 1 & \\ & & \cdots & \cdots \end{bmatrix}. \tag{B4}$$

The qth eigenvectors of β_2 is the plane wave around the nanotube

$$\psi_q(n_y) = \frac{1}{\sqrt{n}} e^{ik_q n_y}, \tag{B5}$$

where the qth wave vector, k_q , satisfies the periodic boundary condition, $k_q = 2\pi q / n$ [$0 \le q \le (n-1)$], and the qth eigenvalue is $b_{2q} = 2te^{-\pi q i / n}\cos(\pi q / n)$. After the basis transformation, β_2 becomes

$$\beta_2^{'} = V^{+}\beta_2 V = \begin{bmatrix} b_{21} & & & \\ & \cdots & & \\ & & b_{2q} & \\ & & & \cdots \end{bmatrix}. \tag{B6}$$

All submatrices, $\alpha', \beta_1^{'}$, and $\beta_2^{'}$ are diagonal, and there are no matrix elements between different modes around the nanotube after the basis transformation. If we reorder the basis according to the modes, the Hamiltonian matrix is

$$H^{'} = \begin{bmatrix} H_1^{'} & & & & \\ & H_2^{'} & & & \\ & & \ldots & & \\ & & & H_q^{'} & \\ & & & & \ldots \end{bmatrix}, \qquad \text{(B7)}$$

where $H_q^{'}$ is the Hamiltonian matrix for the qth mode,

$$H_q^{'} = \begin{bmatrix} U_1 & b_{2q}^{+} & & \\ b_{2q} & U_2 & t & \\ & t & U_3 & b_{2q} \\ & & & \ldots \end{bmatrix}, \qquad \text{(B8)}$$

and all off-diagonal submatrices of *H'* are zero because all modes are decoupled after the basis transformation. Each mode can be separately treated in the mode space, thus the Hamiltonian matrix size is greatly reduced. Furthermore, for typical terminal bias conditions, only a few modes are relevant to carrier transport, which further reduces the computational load. The phase factor of the complex number, b_{2q}, has no effect on the results such as charge density and current, thus it can be omitted and $b_{2q} = 2t\cos(\pi q/n)$ can be used instead.

The source and drain self-energies in the mode space can be computed using the same recursive relation for the surface Green's function as already shown in Appendix A. For the qth mode, the matrix β_1 in eqn. (A2) is replaced by t, and β_2 is replace by b_{2q},

$$g_{1q} = [(E + i0^{+})I - U_1 - b_{2q}^{+} g_{2q} b_2]^{-1},$$

$$g_{2q} = [(E + i0^+)I - U_1 - b_1^+ g_{1q} b_1]^{-1}, \tag{B9}$$

where g_{1q} and g_{2q} are the surface Green's functions for the first and second node inside the source as shown in Fig. A1, and U_1 is the source potential. The recursive equations in the mode space are number equations and can be analytically solved,

$$g_{1q} = \frac{(E-U_1)^2 + b_1^2 - b_{2q}^2 \pm \sqrt{\left[(E-U_1)^2 + b_1^2 - b_{2q}^2\right]^2 - 4(E-U_1)^2 b_1^2}}{2b_1^2(E-U_1)}. \tag{B10}$$

The retarded surface Green's function for the first node inside the source, g_{1q}^r, is the one with the negative imaginary part, and the source self-energy for the qth mode is $\Sigma_{Sq} = b_1^2 g_{1q}$.

Appendix C. Phenomenological treatment of metal-nanotube contacts

The metal/CNT junction is treated in the atomistic mode space. The *q*th mode of a semiconducting, zigzag CNT is

$$H_{CNT} = \begin{bmatrix} E_{m1} & b_{2q} & & & \\ b_{2q} & E_{m2} & t & & \\ & t & E_{m3} & b_{2q} & \\ & & b_{2q} & E_{m4} & \ddots \\ & & & \ddots & \ddots \end{bmatrix}, \tag{C1}$$

where E_{mi} is the middle gap potential at the *i*th carbon ring. To mimic the continuous states injected from the metal to the semiconducting nanotube, each semiconducting mode is coupled at the M/CNT interface to the metallic mode of metallic zigzag CNTs, which has a constant density of states over a large energy range. The Hamiltonian for the metallic subband is

$$H_{met} = \begin{bmatrix} \ddots & \ddots & & \\ \ddots & E_{m1} & t & \\ & t & E_{m1} & -t \\ & & -t & E_{m1} \end{bmatrix}, \tag{C2}$$

where E_{m1} is the energy of the crossing point of the metallic bands, and is the same as E_{m1} in eqn. (C1) if the mid-gap energy is assumed to be continuous at the interface for simplicity. (The simulation results are insensitive to the value of E_{m1} in eqn. (C2) due to the nearly constant density-of-states near the Fermi point of the metallic bands.) The overall Hamiltonian matrix for the metal/CNT junction is

$$H = \begin{bmatrix} H_{met} & t\sqrt{\alpha} \\ t\sqrt{\alpha} & H_{CNT} \end{bmatrix} = \left[\begin{array}{cccc|cccc} \ddots & \ddots & & & & & & \\ \ddots & E_{m1} & t & & & & & \\ & t & E_{m1} & -t & & & & \\ & & -t & E_{m1} & t\sqrt{\alpha} & & & \\ \hline & & & t\sqrt{\alpha} & E_{m1} & b_{2q} & & \\ & & & & b_{2q} & E_{m2} & t & \\ & & & & & t & E_{m3} & \ddots \\ & & & & & & \ddots & \ddots \end{array}\right]. \tag{C3}$$

The metal contact is treated by computing its self-energy to the semiconducting channel. We again use the recursive relation for the surface Green's function of the metal contact,

$$g_S = [(E + i0^+) - E_{m1} - tg_S t]^{-1} \tag{C4}$$

with the solution,

$$g_S = \frac{E - E_{m1} - \sqrt{(E - E_{m1})^2 - 4t^2}}{2t^2}. \tag{C5}$$

The solution with a negative imaginary part is the retarded surface Green's function. The self-energy for the m /CNT contact computed from the surface Green's function is

$$\Sigma_{MS} = \alpha t^2 g_S = \alpha \frac{E - E_{m1} - \sqrt{(E - E_{m1})^2 - 4t^2}}{2}. \tag{C6}$$

VITA

Jing Guo was born in Hunan, China in July, 1976. He received the B. S. degree in electronic engineering in 1998 and the M. S. degree in microelectronics and solid state electronics in 2000, both from Shanghai Jiao Tong University, China. In August, 2000, he started his Ph. D. study in school of electrical and computer engineering at Purdue University, West Lafayette, IN. His current research work centers on device physics and potential applications of carbon nanotube transistors. His previous research work includes studies of silicon nanotransistors and single-electron transistors. Since August 2004, he has been working with the department of electrical and computer engineer in University of Florida, Gainesville, FL.