Chapter 4

Electronic Devices

4.1 Introduction

Given the semiconducting character of two thirds of carbon nanotubes, their high aspect ratio and structural robustness, it is natural to ask if semiconducting carbon nanotubes can be used as active elements in nanoscale electronic devices. Indeed, there has now been many demonstrations of such devices, ranging from two-terminal rectifiers to field-effect transistors. These demonstrations have spurred tremenduous interest in the field of carbon nanotube electronics for several reasons. First, for a single nanotube device, the channel width is on the order of one nanometer, much smaller than state-of-the-art silicon transistors, promising higher device densities. Second, carbon nanotubes have low defect density, and electronic transport is expected to be less affected by defect scattering. Third, the carrier distribution is not as sensitive to temperature variations due to the van Hove singularities at the band edges. Finally, quantum confinement should be easier to achieve because of the small channel width, and thus single-electron devices should be more readily available. Balancing these advantages is the fact that the knowledge base, fundamental understanding, fabrication infracture, and device design are much less developed for carbon nanotubes compared to silicon. For example, while some of the fabricated carbon nanotube devices bear resemblance with traditional silicon devices, this chapter emphasizes the much different physics that governs the operation of the carbon nanotube devices. We have already seen in the previous chapter how one part of a carbon nanotube electronic device—the contacts—behaves much differently from that in conventional materials. The current chapter goes beyond these concepts to address the issue of the whole device, building from a discussion of simple two-terminal devices like *p-n* junctions to more complex devices such as transistors.

4.2 Rectifiers

Rectifiers are simple two-terminal devices that essentially allow current to flow for only one polarity of the applied voltage, the simplest examples being *p-n* junction diodes and Schottky diodes. While these are simple devices, they are used extensively by themselves and as part of more complex circuits; just as important however, is the fact that these systems serve as excellent testbeds to study and highlight the differences between nanotube-based devices and conventional ones. This section presents experimental and theoretical work aimed at realizing carbon nanotube rectifiers, and at understanding their basic operating principles. A central result of this section is that the electrostatics of carbon nanotube devices is much different than that of bulk devices because of the reduced dimensionality, and more specifically because of the weak electrostatic screening. We will see in Chapter 7 that this weak electrostatic screening also has an important impact on the optical properties of carbon nanotubes.

4.2.1 Experimental Realizations of Carbon Nanotube *p-n* Junctions

There are many possible strategies to achieve *p-n* junctions with carbon nanotubes. Examples include substitution of boron and nitrogen in the carbon lattice, doping by charge transfer from electrodes, atoms or molecules, or electrostatic control of the band bending. Figure 4.1 shows one of the strategies that has been implemented to fabricate such a device [Zho00]. The method hinges on the fact that the synthesized carbon nanotubes assembled into devices in air are predominantly *p*-type (this is discussed at length in Chapter 8). Thus, it is only necessary to reverse the doping on one side of the nanotube to obtain a *p-n* junction. This can be accomplished by protecting half of the nanotube with the photoresist PMMA, and exposing the uncovered half to potassium, which is an electron donor. The associated current-voltage curve for such a device (Figure 4.1) shows similarities with an Esaki diode, i.e. it shows negative differential resistance.

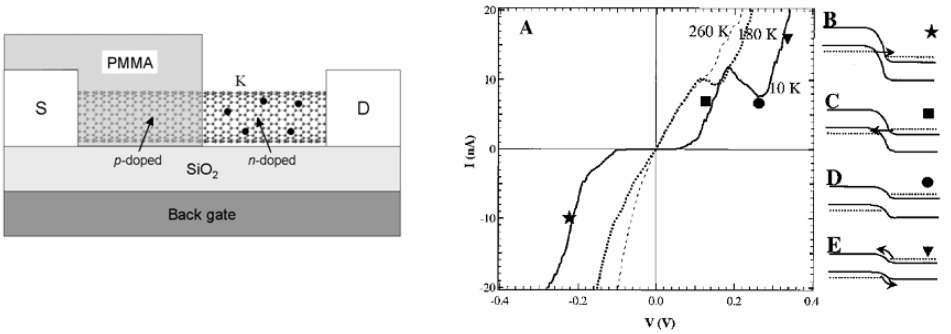

Figure 4.1: Schematic of a chemically doped carbon nanotube *p-n* junction and the associated I-V curve. The sketches labeled B-E show the band-bending along different points of the I-V curve, and the different transport regimes. *Figures from, and after Ref. [Zho00].*

The role of dopants in *p-n* junctions is to create an electrostatic potential step at the junction. In traditional planar devices, doping is essentially the only way to generate such a potential step. In nanotubes however, one can take advantage of the quasi-one-dimensional geometry and use an external electrostatic potential to form the *p-n* junction. An example of this strategy [Lee04] using a buried split-gate structure is shown in Figure 4.2. The advantages of this technique are that no chemical doping of the nanotubes is required, and that the device can be operated in several different modes in a controlled manner. Figure 4.2 shows the I-V curve for this device for three regimes of operation, allowing the transformation of the device from a *p-n* diode, to a *n-p* diode, and to a transistor-like device.

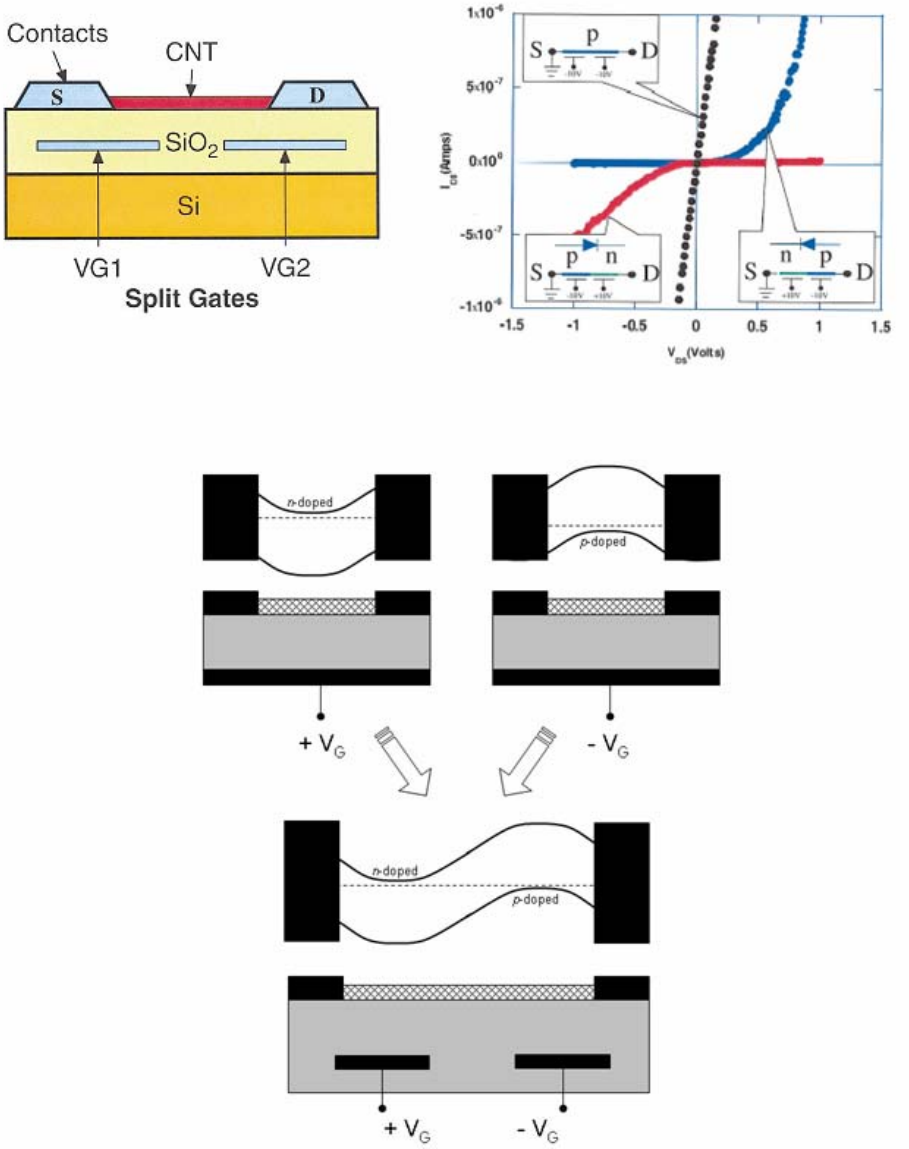

Figure 4.2: Top left: Split-gate structure to create a carbon nanotube *p-n* junction without the need to chemically dope the nanotube. Top right: Experimental current-voltage characteristics of such a device. *Figures from Ref. [Lee04].* Bottom: Schematic illustrating the concept of the split gate architecture, including a sketch of the band-bending.

To obtain an equation describing the current-voltage characteristics of carbon nanotube *p-n* junctions we consider the equation for the current through a ballistic conductor

$$I = \frac{4e^2}{h}\int T(E)\left[f_R\left(E - E_F^L\right) - f_L\left(E - E_F^R\right)\right]dE \qquad (4.1)$$

where the subscripts R and L refer to the right and left leads, respectively. For non-degenerate doping of the carbon nanotube, the transport across the *p-n* junction is through

thermionic emission over the potential step; to model this situation we assume that the transmission probability $T(E)=1$ for electrons (holes) at energies above (below) the conduction (valence) band. The current is then

$$I=\frac{8e}{h}\int_{E_c^\infty}^{\infty}\left[f_R\left(E-E_F^L\right)-f_L\left(E-E_F^R\right)\right]dE \tag{4.2}$$

where E_c^∞ is the energy of the conduction band edge on the *n*-type side far from the junction. When the doping is not too large, $E_c^\infty - E_F^R \gg kT$, the Fermi functions can be approximated as

$$f=\frac{1}{1+\exp\left(\dfrac{E-E_F}{kT}\right)}\approx\exp\left(\frac{E_F-E}{kT}\right)\quad\text{when } E-E_F\gg kT\,. \tag{4.3}$$

Substituting this approximation in the expression for the current gives

$$\begin{aligned} I&=\frac{8e}{h}\int_{E_c^\infty}^{\infty}\exp\left(\frac{E_F^R-E}{kT}\right)dE-\frac{8e}{h}\int_{E_c^\infty}^{\infty}\exp\left(\frac{E_F^L-E}{kT}\right)dE \\ &=\frac{8ekT}{h}\left[\exp\left(\frac{E_F^R-E_c^\infty}{kT}\right)-\exp\left(\frac{E_F^L-E_c^\infty}{kT}\right)\right] \end{aligned}\,. \tag{4.4}$$

Assuming that the band edge simply tracks the Fermi levels in the leads (i.e. far away from the junction), the difference $E_c^\infty - E_F^R$ is independent of applied voltage; furthermore, since we have $E_F^R - E_F^L = eV$ we obtain

$$I=I_0\left[\exp\left(\frac{eV}{kT}\right)-1\right] \tag{4.5}$$

with

$$I_0=\frac{8ekT}{h}\exp\left(\frac{E_F^R-E_c^\infty}{kT}\right). \tag{4.6}$$

Equation (4.5) is the celebrated ideal diode equation describing rectifying behavior, except that here it was derived under the assumption of ballistic transport.

Further development of the buried-gate approach has shown that such carbon nanotube diodes can achieve this ideal rectifying behavior. Figure 4.3 shows the measured I-V curves for a *p*-*i*-*n* diode, with the solid lines in the figure representing the modified diode equation

$$I=I_0\left[\exp\left(\frac{eV}{nkT}\right)-1\right] \tag{4.7}$$

where n is the ideality factor. The right panel in Figure 4.3 shows that at low bias, the experimental data satisfies this expression with an ideality factor n=1.2, close to the ideal diode value n=1. The inverse of the slope of the $d\ln(I)/dV$ curve

$$\alpha = \frac{dV}{d\ln(I)} = nkT/e \tag{4.8}$$

is plotted in the bottom inset in Figure 4.3 as a function of temperature clearly indicating the linear dependence of α on temperature. The inset in the left panel indicates that at high bias, the measured values of α does not approach zero at small temperatures as Equation (4.8) would suggest; the origin of this effect is unclear. Further analysis of the I-V curves shows that the transmission probability is about 0.2, indicating that there is some amount of scattering in the nanotube. The device behavior can thus be improved by reducing scattering; one source of scattering is due to the interaction of the carbon nanotube with the substrate, which can cause geometrical deformations of the nanotube and also distortions of the nanotube electronic structure. This source of scattering can be removed by suspending the nanotube over a trench in the channel region, as shown in Figure 4.4. This device is fabricated by first growing a 400 nm thermal oxide on a heavily-doped Si substrate, on top of which two Mo split gates are fabricated using standard lithography, with a gate spacing between 0.5 micron and 1 micron. These split gates are then used as an etch mask to etch 250 nm of oxide. 150 nm of oxide is then deposited to form the gate dielectric for the split gates, and lithography is used to define source and drain electrodes. Carbon nanotubes were grown by chemical vapor deposition from catalyst particles on the electrodes. The current-voltage characteristics of this device follow the ideal diode equation with $n = 1$.

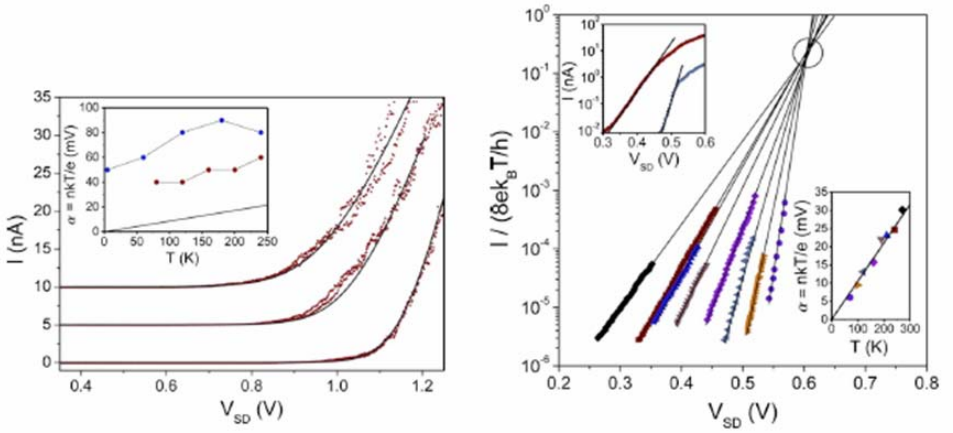

Figure 4.3: Measured current-voltage characteristics of carbon nanotube *p-i-n* junctions. The left panel shows fits of the ideal diode equation with a series resistor (1–6 MΩ) in the high forward bias regime at 80, 160, and 240 K (bottom to top, offset for clarity). Inset: α for two different devices. The solid line indicates the expected temperature dependence for an ideal diode. The right panel shows *I-V* curves in the low forward bias regime at various temperatures. Top inset: Full *I-V* curves in forward bias at *T*=240 K (filed squares) and 120 K (filled triangles). Bottom inset: Temperature dependence of the inverse slope $\alpha = dV/d(\ln I)$. *Figures from Ref. [Bos06].*

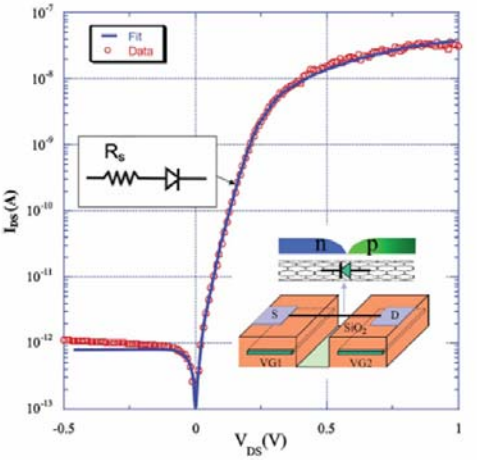

Figure 4.4: The inset shows a sketch of a buried-gate carbon nanotube *p-n* junction where the nanotube is suspended over a trench in the channel region. The main figure shows the measured current-voltage diode behavior (open symbols) as well as a fit to the ideal diode equation (solid line) including a contact resistance of 18MΩ. *Figure from Ref. [Lee05].*

4.2.2 Theory of Carbon Nanotube *p-n* Junctions

While the experimens discussed in the previous section indicate that carbon nanotube *p-n* junctions can be fabricated and can behave as ideal diodes, such devices need to be examined more carefully to establish commonalities and differences with traditional devices. The behavior of nanotube *p-n* junctions can be understood by performing self-consistent calculations of the charge and electrostatic potential along the nanotube. Assuming azimuthal symmetry, the simplest model for the charge on the nanotube is

$$\sigma(z) = \frac{e}{\varepsilon} f - \frac{e}{\varepsilon} \int D(E,z) F(E) dE \tag{4.9}$$

where ε is the dielectric constant of the medium in which the nanotube is embedded, f is the doping fraction, $D(E,z)$ is the nanotube density of states at position z along the tube, and $F(E)$ is the Fermi function. The density of states can be expressed as

$$D(E,z) = \frac{a\sqrt{3}}{\pi^2 R V_0} \frac{\left|E + eV(z)\right|}{\sqrt{\left(E + eV(z)\right)^2 - \left(E_g/2\right)^2}} \tag{4.10}$$

where $V(z)$ is the electrostatic potential along the nanotube. This expression for the spatial variation of the density of states is simply a rigid shift with the local electrostatic potential; while there are more sophisticated methods to calculate the actual local density of states and the occupation of the states that enters in the calculation of the charge,

Equations (4.9) and (4.10) are sufficient to illustrate the general properties of nanotube *p-n* junctions.

The other equation necessary for the computations is the electrostatic potential $V(z)$ generated by the charge density $\sigma(z)$:

$$V(z) = \int K(z-z')\sigma(z')dz' \tag{4.11}$$

where σ is the charge per unit area on the nanotube and $K(z-z')$ is the free space electrostatic kernel for a hollow cylinder:

$$K(z-z') = \frac{R}{4\pi\varepsilon_0}\int_0^{2\pi} \frac{d\theta}{\sqrt{(z-z')^2 + 2R^2 - 2R^2\cos\theta}}. \tag{4.12}$$

The procedure is therefore to solve self-consistently Equations (4.9) and (4.11) for a given doping on the nanotube. Figure 4.5 shows results of such calculations for two doping fractions. Clearly, the band-bending in the nanotube is similar to what is observed in planar devices: a potential step at the junction and essentially flat bands away from the junction. The behavior is quite different, however, if one looks at the charge distribution. In a planar device, there is a region of constant charge near the junction, and no charge outside of that so-called depletion region. In contrast, for nanotubes, there is significant charging away from the junction. In fact, the charge decays only logarithmically away from the junction. This difference between planar and nanotube devices is again due to the different electrostatics of dipole sheets and dipole rings (see Chapter 3). In the planar device, having a dipole sheet at the junction is sufficient to ensure that the potential stays constant far away from the junction. For the dipole ring however, the potential decays away from the junction. Since the potential must be constant far away from the junction, charge must continuously be added along the nanotube to keep the potential from falling. To illustrate this important effect in more detail, we consider mathematically the behavior of bulk and nanotube *p-n* junctions.

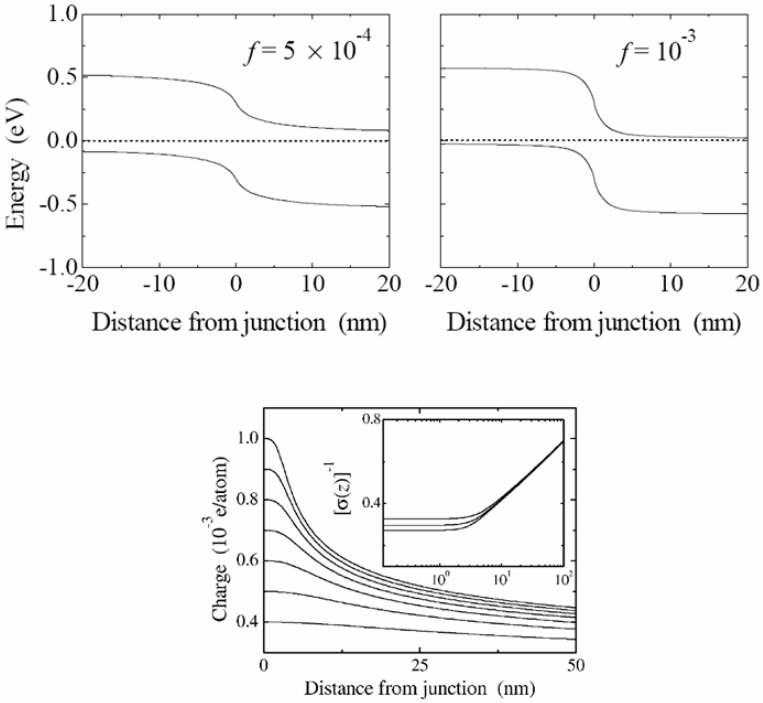

Figure 4.5: Calculated self-consistent band-bending and charge along a carbon nanotube *p-n* junction. The top two panels show the band-bending for two different doping fractions. The dotted line is the Fermi level. The bottom panel shows the charge distribution on the *n*-type side of the junction for different doping fractions, which can be read-off from the value of the vertical intercept. The inset in the bottom figure shows a scaling of all of the curves for the different doping fractions. *Figure from Ref. [Leo99].*

Bulk *p-n* junction

All derivations of the electrostatics of bulk *p-n* junctions essentially start with the assumption of a depletion region near the junction where the dopants are completely depleted and where the charge is a constant. It is usually assumed that this depletion region ends abruptly at the depletion width W as illustrated in Figure 4.6. For equal and opposite doping on the two sides of the *p-n* junction, Poisson's equation for this charge distribution is

$$\frac{d^2V(z)}{dz^2} = \begin{cases} -\dfrac{\rho}{\varepsilon} & 0 < z < W \\ \dfrac{\rho}{\varepsilon} & -W < z < 0 \\ 0 & |z| > W \end{cases} \tag{4.13}$$

where ρ is the charge density. This second order differential equation is augmented by the boundary conditions that the potential reach a value at the depletion edge that gives charge neutrality in the semiconductor:

$$V\left(\pm W\right)=V_{\pm}. \tag{4.14}$$

The difference $V_{+}-V_{-}$ gives the so-called built-in potential V_{bi}. Integrating Equation (4.13) with this boundary condition leads to the solution

$$V(z)=\begin{cases}\frac{V_{bi}}{2}-\frac{\rho}{2\varepsilon}\left(z^{2}+W^{2}\right)+A_{+}\left(z-W\right) & 0<z<W\\ -\frac{V_{bi}}{2}+\frac{\rho}{2\varepsilon}\left(z^{2}+W^{2}\right)+A_{-}\left(z+W\right) & -W<z<0\\ \frac{V_{bi}}{2} & z\geq W\\ -\frac{V_{bi}}{2} & z\leq -W\end{cases}. \tag{4.15}$$

The constants $A_{\pm}$ are determined from the continuity of the derivative of the potential at $z=\pm W$. Applying this condition gives the solution for the potential

$$V(z)=\begin{cases}\frac{V_{bi}}{2}-\frac{\rho}{2\varepsilon}\left(z-W\right)^{2} & 0<z<W\\ -\frac{V_{bi}}{2}+\frac{\rho}{2\varepsilon}\left(z+W\right)^{2} & -W<z<0\\ \frac{V_{bi}}{2} & z\geq W\\ -\frac{V_{bi}}{2} & z\leq -W\end{cases}. \tag{4.16}$$

Continuity of the potential at $z=0$ gives the expression for the depletion width

$$W_{bulk}=\sqrt{\frac{\varepsilon V_{bi}}{\rho}}. \tag{4.17}$$

The spatial variation of the potential is sketched in Figure 4.6. The important point is that the assumption of a depletion region with constant charge naturally leads to a potential that is constant outside of the depletion region, and thus the semiconductor is charge-neutral even at long distances from the junction. This result turns out to be entirely fortuitous and is a special property of bulk, planar interfaces. Nanotubes, with their reduced dimensionality have a much different behavior, as we now discuss.

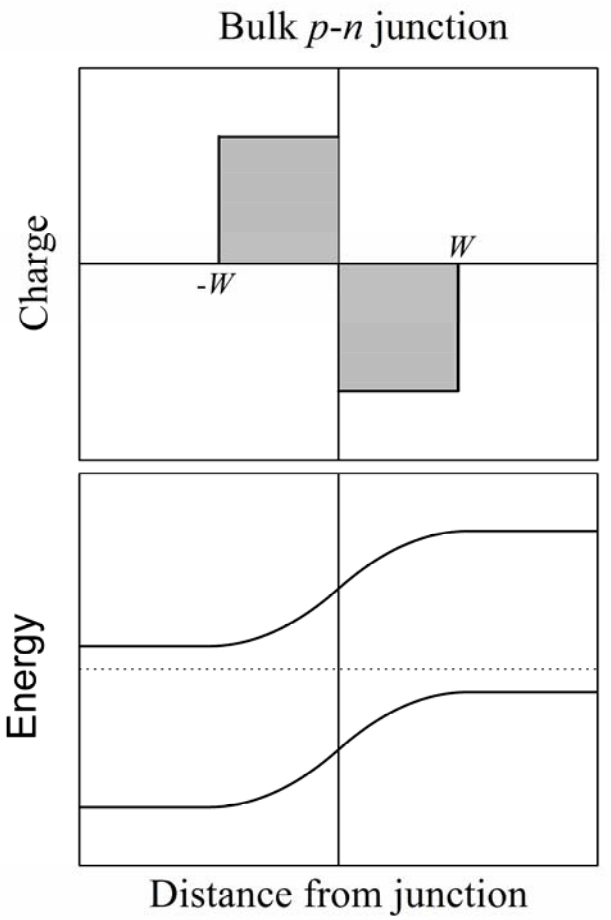

Figure 4.6: Charge distribution and band-bending for a symmetric bulk *p-n* junction. Dotted line is the Fermi level.

Carbon nanotube *p-n* junction

We first illustrate the problem that arises when the assumption of a constant depletion region is made for a carbon nanotube *p-n* junction. The critical difference with the bulk junction is that we no longer have an infinite dipole sheet, but rather have a nanometer-size dipole ring. The potential due to a dipole ring can be calculated as

$$V(z) = \frac{R}{4\pi\varepsilon_0}\int dz'\sigma(z')\int_0^{2\pi}\frac{d\theta}{\sqrt{(z-z')^2+2R^2-2R^2\cos\theta}} \tag{4.18}$$

with

$$\sigma(z') = \begin{cases} \rho & 0 < z < W \\ -\rho & -W < z < 0 \\ 0 & |z| > W \end{cases}. \tag{4.19}$$

The potential due to the *n*-type side is obtained by direct integration

$$\begin{aligned} V_n(z) &= \frac{R\rho}{4\pi\varepsilon_0}\int_0^W dz' \int_0^{2\pi} \frac{d\theta}{\sqrt{(z-z')^2 + 2R^2 - 2R^2\cos\theta}} \\ &= \frac{R\rho}{4\pi\varepsilon_0}\int_0^{2\pi} d\theta \ln\left(\frac{z + \sqrt{2R^2(1-\cos\theta)+z^2}}{z - W + \sqrt{2R^2(1-\cos\theta)+(z-W)^2}}\right). \end{aligned} \tag{4.20}$$

Similarly, the potential due to the *p*-type region is

$$V_p(z) = -\frac{R\rho}{4\pi\varepsilon_0}\int_0^{2\pi} d\theta \ln\left(\frac{-z + \sqrt{2R^2(1-\cos\theta)+z^2}}{-z - W + \sqrt{2R^2(1-\cos\theta)+(z+W)^2}}\right), \tag{4.21}$$

so that the total potential is given by

$$\begin{aligned} V(z) &= V_n(z) + V_p(z) \\ &= \frac{R\rho}{4\pi\varepsilon_0}\int_0^{2\pi} d\theta \ln\left(\frac{z + \sqrt{2R^2(1-\cos\theta)+z^2}}{z - W + \sqrt{2R^2(1-\cos\theta)+(z-W)^2}} \frac{z + W - \sqrt{2R^2(1-\cos\theta)+(z+W)^2}}{z - \sqrt{2R^2(1-\cos\theta)+z^2}}\right) \end{aligned} \tag{4.22}$$

For large z this expression becomes

$$V(z) \approx \frac{R\rho}{\varepsilon_0}\frac{W^2}{z^2} \quad |z| \gg R, W. \tag{4.23}$$

Thus, under the assumption of a finite depletion region with constant charge, the potential decays far away from the junction. This is in contrast to the bulk planar junction where the potential is constant outside of the depletion region. The decay of the potential far from the junction implies that the bands would fall below the charge neutrality position and that the nanotube would be charged far from the junction. To prevent this unphysical situation from happening, charge must be added to the nanotube outside of the depletion region to prevent the potential from decaying. This is the basis of the long-range charging in nanotube junctions. To obtain an expression for the charge outside of the depletion region, we consider the charge distribution

$$\sigma(z)=\begin{cases}\rho & 0<z<W\\ \alpha V(z) & z>W\\ -\rho & -W<z<0\\ \alpha V(z) & z<-W\end{cases}, \tag{4.24}$$

which is essentially a linear response model for the charge outside of the depletion region. The electrostatic potential is then

$$V(z)\approx V_0(z)+\alpha\int_{-\infty}^{-W}V(z')K(z-z')dz'+\alpha\int_{W}^{\infty}V(z')K(z-z')dz'. \tag{4.25}$$

where $V_0(z)$ is the potential due to the finite width depletion regions. Since the potential is small in the depletion region, we extend the limits of integration in the integrals to get

$$V(z)\approx V_0(z)+\alpha\int_{-\infty}^{-\infty}V(z')K(z-z')dz'. \tag{4.26}$$

To obtain the asymptotic dependence of this integral equation, we transform the potential and the electrostatic kernel to Fourier space. The Fourier representation of the kernel is

$$K(q)=I_0(qR)K_0(qR) \tag{4.27}$$

where I_0 and K_0 are modified Bessel functions. The equation for the potential in Fourier space is

$$V(q)=V_0(q)+\alpha V(q)K(q) \tag{4.28}$$

giving the solution

$$V(q)=\frac{V_0(q)}{1-\alpha K(q)}. \tag{4.29}$$

For small q (long distances) the Bessel functions have asymptotic behavior $I_0(qR)\sim$ constant and $K_0(qR)\sim-\ln(qR)$ giving the potential

$$V(q)\sim\frac{V_0(q)}{\ln(qR)}. \tag{4.30}$$

The real space potential is given by the inverse transform

$$V(z)\sim\int_0^{\infty}V(q)\sin(qz)dq=\int_0^{\infty}\frac{V_0(q)\sin(qz)}{\ln(qR)}dq\sim\int_0^{\infty}\frac{\sin(qz)}{q\ln(qR)}dq. \tag{4.31}$$

For large z, and because the integrand is dominated by small q, the integral can be approximated as

$$V(z) \sim z\int_0^{q_c} \frac{1}{\ln(qR)} dq \tag{4.32}$$

where the cut-off wavenumber $q_c \sim 1/z$. The integral gives

$$V(z) \sim zEi\left[\ln\left(q_c R\right)\right] \sim \frac{ze^{\ln\left(q_c R\right)}}{\ln\left(q_c R\right)} \sim \frac{1}{\ln(R/z)}. \tag{4.33}$$

Thus, because the charge is linearly related to the potential through Equation (4.24) the charge outside of the depletion region follows the asymptotic behavior

$$\sigma(z) \sim \frac{1}{\ln(z/R)} \quad \text{large } z; \tag{4.34}$$

this function varies extremely slowly with distance, leading to long range charging of the carbon nanotube. This expression is in excellent agreement with the numerical results presented in Figure 4.5.

While there is long-range charging of the carbon nanotube, there is still a depletion-like region near the junction where the carriers are fully depleted and the charge is constant. It is interesting to obtain expressions for the width of this depletion region as a function of doping. To do so, we first obtain an equation for the built-in potential by assuming that the potential step at the junction is due solely to the charge in the depletion regions:

$$\begin{aligned} V_{bi} &= V_n(W) - V_p(-W) = 2V_n\left(W\right) \\ &= \frac{R\rho}{2\pi\varepsilon_0}\int_0^{2\pi} d\theta \ln\left(\frac{W + \sqrt{2R^2(1-\cos\theta) + W^2}}{\sqrt{2R^2(1-\cos\theta)}}\right) \\ &= \frac{R\rho}{2\pi\varepsilon_0}\int_0^{2\pi} d\theta \ln\left(\frac{\frac{W}{R} + \sqrt{2(1-\cos\theta) + \left(\frac{W}{R}\right)^2}}{\sqrt{2(1-\cos\theta)}}\right). \end{aligned} \tag{4.35}$$

For not too high doping, $W \gg R$ and Equation (4.35) can be approximated as

$$V_{bi} \approx \frac{R\rho}{\varepsilon_0}\ln\left(\frac{2W}{R}\right), \tag{4.36}$$

giving the expression for the depletion width

$$W_{NT} \sim R\exp\left(\frac{\varepsilon_0 V_{bi}}{\rho R}\right). \tag{4.37}$$

As indicated by Equation (1.47), the built-in potential depends logarithmically on the doping and is therefore a slowly varying function of doping, and the main dependence of

the depletion width on the doping [Equation (4.37)] comes from the $1/\rho$ factor in the exponential. As a consequence the depletion width is an extremely sensitive function of the doping, in constrast to the mild dependence on doping of bulk junctions, Equation (4.17).

While we discussed the long distance charging in the context of *p-n* junctions, Schottky junctions between nanotubes and planar metals are also expected to show the same behavior, and this has been demonstrated experimentally. Figure 4.7 shows a scanning electron micrograph image of a nanotube connecting two Au electrodes, and the associated charge distribution away from the contact. The long distance charging is observed, as predicted theoretically.

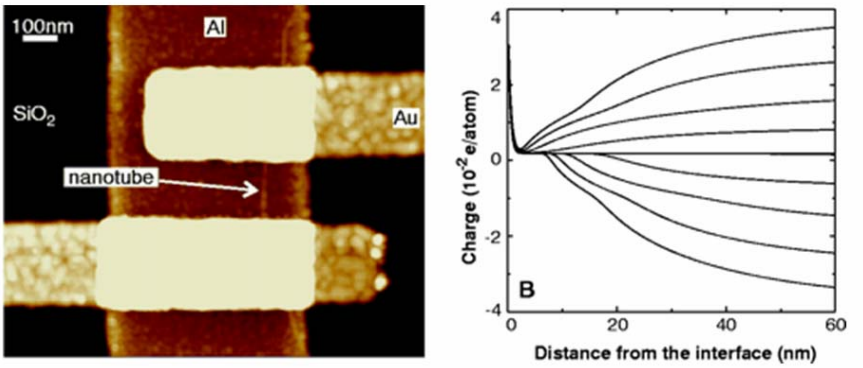

Figure 4.7: The left panel shows a scanning electron micrograph image of a nanotube between two electrodes. The right panel shows the charge along the nanotube, indicating long-distance charge transfer from the electrodes. *Figure from Ref. [Bac01].*

The much different charge distribution and electrostatics in nanotube junctions has dramatic impact on device design. For example, in traditional devices, the height of the potential step can be tailored by changing the doping. The depletion width in such devices depends weakly on the doping, thus allowing for precise control of the device properties. For nanotubes, however, the situation is quite different. Figure 4.8 shows the calculated depletion width for the nanotube *p-n* junction as a function of doping.

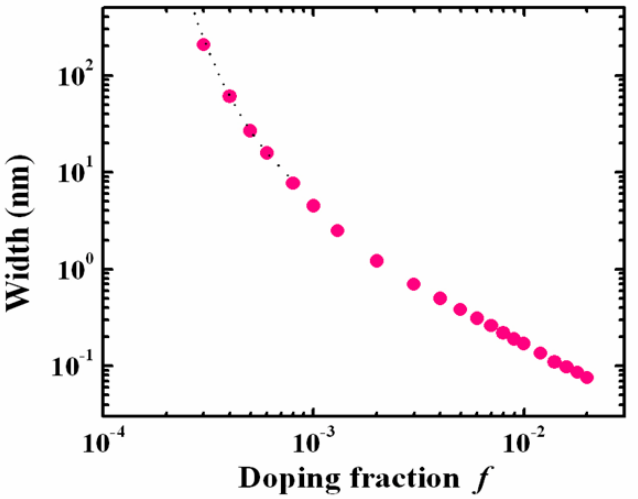

Figure 4.8: Calculated depletion width for a nanotube *p-n* junction as a function of doping. *Figure from Ref. [Leo99].*

Clearly the depletion width is extremely sensitive to the doping, and thus fluctuations in dopant levels from device to device can significantly affect the device characteristics. Furthermore, at high doping, the depletion width is so small that tunneling across the potential step prevents the device from rectifying. This tunneling phenomenon is the basic operating principle behind negative differential resistance devices, and is observed in the experimental device of Figure 4.1. It is thus interesting to model the properties of such devices. To do so requires computing the I/V curve, and this is done using the expression for the current in Equation (4.1) where the transmission probability $T(E)$ for tunneling is calculated using the WKB approximation. Figure 4.9 shows the results of such calculations, which indicate negative differential resistance, with a large ratio of maximum to minimum current.

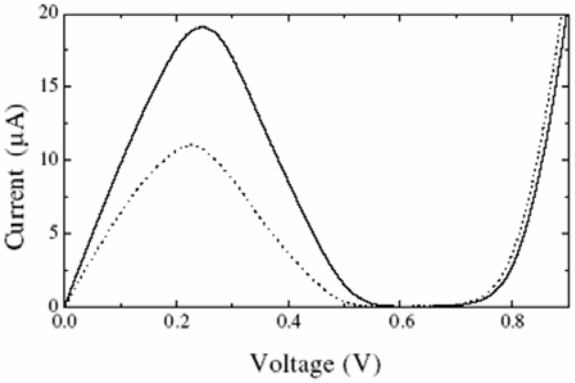

Figure 4.9: Calculated current-voltage curve for a nanotube *p-n* junction with high doping. The current-voltage curve shows negative differential resistance between 0.25 and 0.6 Volts. *Figure from Ref. [Leo00].*

4.2.3 Metal-Nanotube Junctions

Sections 4.2.1 and 4.2.2 described *intratube p-n* junctions, where rectification comes from modulation of the doping within a single nanotube. In addition, we discussed how contacts between nanotubes and metals can also act as Schottky diodes. In this section, we are concerned with metal-semiconductor rectifiers where both the metal *and* the semiconductor are carbon nanotubes. Such devices can be fabricated by combining two different nanotubes: Figure 4.10 shows an experimental realization of one such device, consisting of two crossing nanotubes. Measurement of the individual conductance is used to determine the semiconducting or metallic character of each of the two nanotubes. Figure 4.11 indicates that the current between the metallic and semiconducting nanotubes (curve labeled MS) shows rectification. This rectification behavior can be understood from the fact that the band gap in a semiconducting nanotube arises from the opening of a symmetric gap around the Fermi points of a graphene sheet. Thus, the Fermi level is in the middle of the nanotube bandgap, and is at the same energy as the Fermi level in a metallic tube. This leads to the presence of a Schottky barrier at the crossing point between the two nanotubes equal to half the badgap, as illustrated in Figure 4.11.

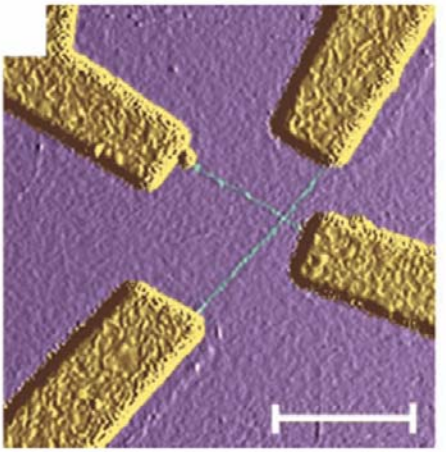

Figure 4.10: A nanotube device made of two crossing nanotubes. *Figure from Ref. [Fuh00].*

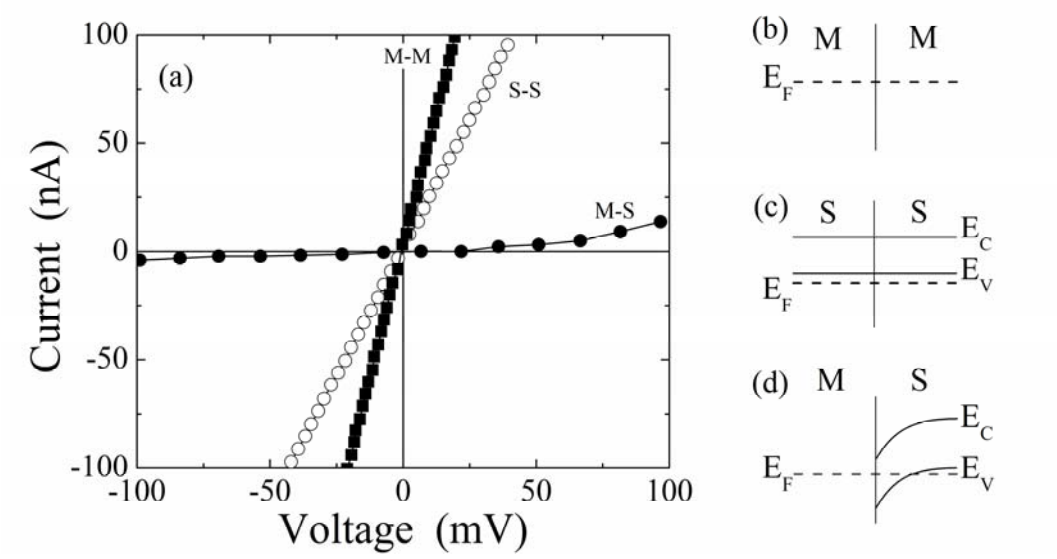

Figure 4.11: Measured I-V curves for devices like the one in Figure 4.10. The metal-semiconductor (MS) junction shows rectifying behavior, due to a Schottky barrier at the junction, as illustrated in panel (d). *Figure from Ref. [Fuh00].*

The same Schottky barrier concept can be used to create *intra*-tube metal-semiconductor junctions. Figure 4.12 shows an image of a nanotube in a four probe measurement configuration, with a kink between the middle electrodes. Two-probe measurements show that one end of the nanotube is semiconducting, while the other end is metallic. Thus, the two segments of the nanotube correspond to different chiralities, and the angle at which they meet is determined by the presence of topological defects which allow a seamless junction. This type of *p*-*n* junction shows strong rectification, and its working principle has been described theoretically in reference [Odi00].

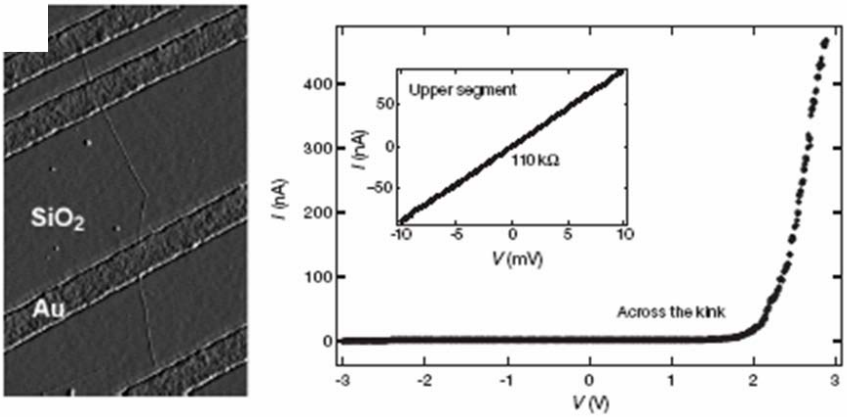

Figure 4.12: Intra-tube metal-semiconductor junction, and associated rectifying behavior. *Figure from Ref. [Yao99].*

4.3 Field Effect Transistors

Ever since its invention, the transistor has been the workhorse of the electronics industry, with field-effect transistors dominating the mass consumer market. Simple field-effect transistors are easier to realize, as compared to other types of transistors, such as bipolar junction transistors, in that no intricate doping control is necessary to demonstrate the field-effect. It is no surprise then that some of the initial devices made with nanotubes have been field-effect transistors [Tan98, Mar98]. Figure 4.13 shows an atomic force micrograph image of one of the early nanotube transistors, which consists of a semiconducting nanotube bridging two Pt electrodes, and sitting on SiO_2 between the electrodes. A heavily doped Si substrate serves as a back gate, which controls the switching action of the transistor.

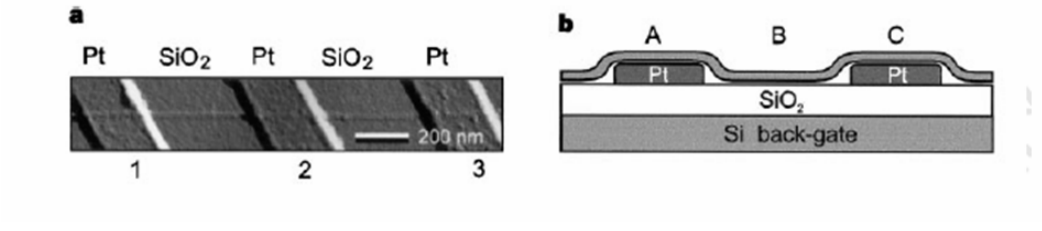

Figure 4.13: Atomic force micrograph image and sketch of original nanotube transistor. The transistor action is controlled by changing the voltage on the Si back gate. *Figure from Ref. [Tan98].*

The drain current versus drain voltage characteristics of this transistor are shown in Figure 4.14. In going from a gate voltage of -3 volts to +6 volts, the device changes from high to low conductance, which correspond to the ON and OFF states, and thus providing the switching action of the transistor.

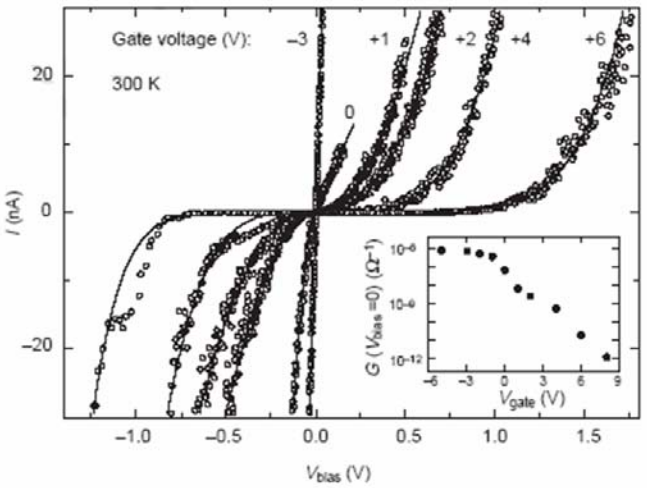

Figure 4.14: Current/voltage characteristics of an early nanotube transistor. *Figure from Ref. [Tan98].*

Since this original device, there has been much experimental and theoretical progress in the understanding of the physics that governs the transistor action, and in improvement of the device performance. An important outcome of this work is the fact that the type of contact (ohmic or Schottky) has a profound influence on the device behavior.

4.3.1 Ohmic Contacts

As discussed in Chapter 3, ohmic contacts to nanotubes have been reported in the literature. Because of the ohmic contacts, the physics governing the transistor action is bending of the bands in the channel by the applied gate voltage. Theoretical work [Leo02a] to explain this behavior has been presented in the literature. Results of such work, based on quantum transport calculations, are presented in Figure 4.15. As illustrated in this figure, the device consists of a single-wall, semiconducting carbon nanotube, embedded in metal contacts on either side, defining the source and drain. Between the source and drain electrodes, an insulating dielectric surrounds the nanotube up to a radius of 10 nm. A cylindrical gate of radius 10 nm wraps the dielectric and serves to control the device behavior. In the calculations, the nanotube and the metals are separated by a van der Waals distance of 0.3 nm. (The insulator has dielectric constant of 3.9, as for SiO2, and it is also separated from the tube by 0.3 nm.) A zigzag nanotube of index (17, 0) is considered, which has a radius of 0.66 nm and band gap 0.55 eV. The metal Fermi level is chosen to be 1 eV below the nanotube midgap before self-consistency. (For the nanotube midgap 4.5 eV below the vacuum level, this corresponds to a metal work function of 5.5 eV.) Panel (a) in the figure shows the calculated zero bias conductance as a function of the gate voltage. The device shows three regimes: in regime I the conductance is high, corresponding to the ON state of the transistor. In this regime, the bands are essentially flat [Figure 4.15b] so there is little scattering of electrons at the Fermi level. (Even though the transport is “ballistic”, a spatially-varying potential can

cause scattering of electrons.) Since the conduction band has a degeneracy of two, the conductance in this regime saturates to a value close to two quanta of conductance. As the gate voltage is increased, the conductance decreases sharply, and the transistor enters the OFF regime. This regime is characterized by a large barrier in the middle of the nanotube that blocks the electrons (there is a small leakage current due to source-drain tunneling). As the gate voltage is further increased, the channel is driven into inversion. While for micron-sized channels this inversion would lead to a permanent turn-on of the conductance, for nanometer-sized channels the situation is quite different. In this case, the band bending creates an electrostatic quantum dot in the middle of the nanotube, leading to the appearance of localized energy levels. Thus, the inversion regime in nanotube transistors with nanoscale channels consists of resonant tunneling through these discrete levels, leading to a peak in the conductance in regime III. This regime is expected to have intriguing behavior such as high frequency response and has yet to be explored experimentally.

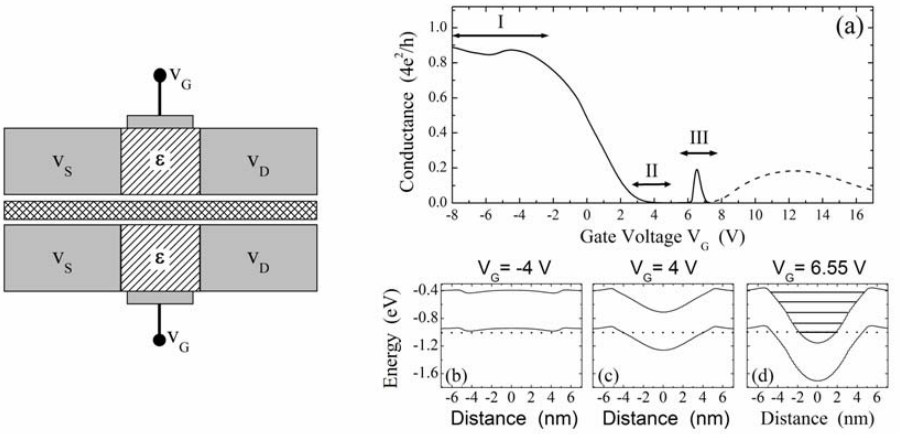

Figure 4.15: Left panel: Schematic cross section of a carbon nanotube field-effect transistor with a cylindrical gate. The gray areas are the gate and the metallic source and drain contacts to the nanotube. The hatched areas represent the dielectric that surrounds the nanotube, and the cross-hatched area is the nanotube. The source-drain separation is 10 nm; the cylindrical gate has a radius of 10 nm. (a) The calculated zero-bias conductance of the nanotube field-effect transistor. (b-d) Band-bending associated with regimes I, II and III of panel (a). Dotted line is the Fermi level. *Figure from Ref. [Leo02a].*

As we discussed in Chapter 3, ohmic contacts to carbon nanotubes have been realized experimentally using palladium as the contact metal. Nanotube field-effect transistors with palladium contacts and channel lengths of a few hundred of nanometers have achieved ON state conductances close to two quanta of conductance [Jav03]. However, increasing the channel length to a few microns severely decreases the ON state conductance, consistent with the electron-phonon scattering mean-free path of 1.5 microns. Thus for ohmically-contacted nanotubes, the ON state conductance is limited by scattering in the channel, and channel lengths less than the electron-phonon scattering mean-free path are essentially ballistic.

The temperature dependence of the ON state conductance also provides further evidence for the presence of ohmic contacts. Indeed, using Equation (2.11) and assuming perfect transmission through the contacts and the nanotube, we can obtain the temperature-dependent ON state conductance as [Leo06b]

$$G = \frac{4e^2}{h} \frac{e^{\Delta / kT}}{1 + e^{\Delta / kT}} \tag{4.38}$$

where $\Delta = E_v - E_F$ represents the position of the Fermi level in the valence band. The monotonically decreases with increasing temperature in agreement with the data presented in Figure 3.8. Thus, even in carbon nanotube field-effect transistors without electron-phonon scattering it is expected that the conductance will decrease with increasing temperature, and can be reduced by as much as a factor of two at room temperature compared to its low-temperature value.

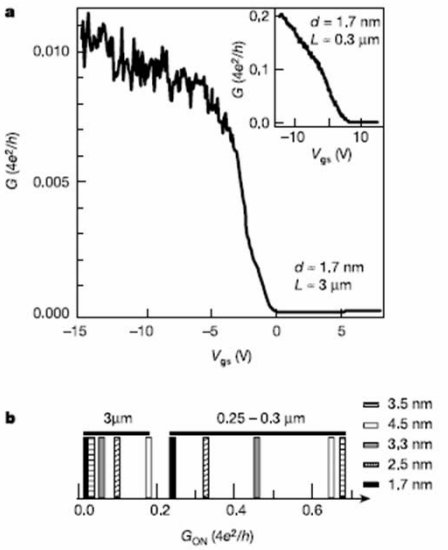

Figure 4.16: (a) Transfer characteristics of carbon nanotube field-effect transistors with ohmic palladium contacts. The main panel shows that a device with a channel of 3 microns reaches only 1/100th of the maximum possible ON conductance. However, when the channel length is reduced to 0.3 microns, the ON state conductance is within a factor of 5 of the maximum possible value. (b) Distribution of ON state conductance for a number of devices illustrates the role of the channel length. *Figure from Ref. [Jav03].*

The main conclusion of this section is that the behavior of ohmic nanotube transistors is determined by changes in the band-bending and scattering in the channel region. As we will see in the next section, nanotube transistors with Schottky contacts behave much differently.

4.3.2 Schottky Contacts

As we have discussed in Chapter 3, electrical contacts to carbon nanotubes are often characterized by the presence of Schottky barriers. Normally, the current across such contacts is dominated by thermionic emission, where electrons must be thermally excited over the Schottky barrier. However, if the band bending near the contact is very sharp, electrons can tunnel through the barrier, leading to a much increased current. This is precisely the effect that governs the operation of Schottky barrier nanotube transistors, as illustrated in Figure 4.17.

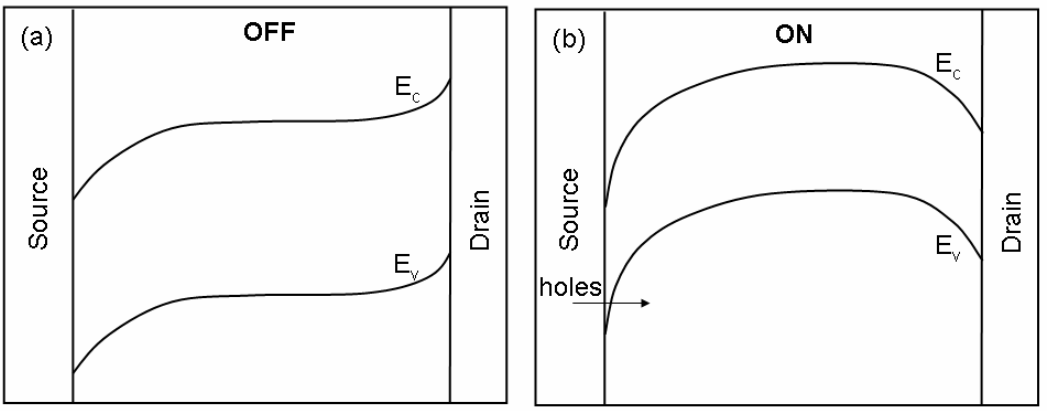

Figure 4.17: Band bending in a Schottky barrier nanotube transistor, for two values of the gate voltage. *Figure after Ref. [App02].*

In part (a) of this figure, the band-bending is sketched for the OFF state of the transistor. At this gate voltage, the band-bending near the contact is small, the tunneling length is long, and consequently the tunneling current is small. Increasing the gate voltage as in panel (b) raises the bands in the middle of the nanotube, leading to a much sharper band-bending at the contacts. This reduces the tunneling distance at the contacts and leads to larger current. The device operation is thus controlled by modulation of tunneling at the contacts, a mechanism that is entirely different form conventional transistors and ohmic contact nanotube field-effect transistors. This effect of the gate voltage on the contact behavior has been demonstrated using computer simulations [Hei02]. Figure 4.18 shows a cross-section of the nanotube transistor considered for the calculations. The calculations proceed as described earlier, with the charge and the potential on the nanotube calculated self-consistently; the transmission due to tunneling through the band-bending at the contact is obtained from the WKB approximation, and the conductance is calculated from the transmission using the Landauer formula, Equation (4.1). For a calculated potential variation $V(z)$ along the length of the carbon nanotube, the transmission probability within the WKB approximation is given by

$$T(E) = \exp\left(-2\int_{z_1}^{z_2} k(E,z)dz \right) \tag{4.39}$$

where $k(E,z)$ is the imaginary part of the wavevector as a function of position along the carbon nanotube and z_1 and z_2 are the classical turning points. Tunneling through the

contact band-bending is illustrated in Figure 4.19, indicating that the imaginary part of the wavevector depends on distance along the nanotube because of the band-bending. To obtain an expression for the imaginary part of the wavevector as a function of energy and distance, we begin by deriving a simple expression for its dependence on energy. As discussed in Chapter 1, near the Fermi points the bandstructure is given by

$$E = \frac{3a\gamma}{2}\sqrt{\left(\frac{E_g}{3a\gamma}\right)^2 + k^2} \tag{4.40}$$

and expressing k in terms of E from this equation gives

$$k(E) = \frac{2}{3a\gamma}\sqrt{E^2 - \left(E_g/2\right)^2}. \tag{4.41}$$

For values of $|E| < E_g/2$ the wavevector is imaginary; this imaginary value is zero at the

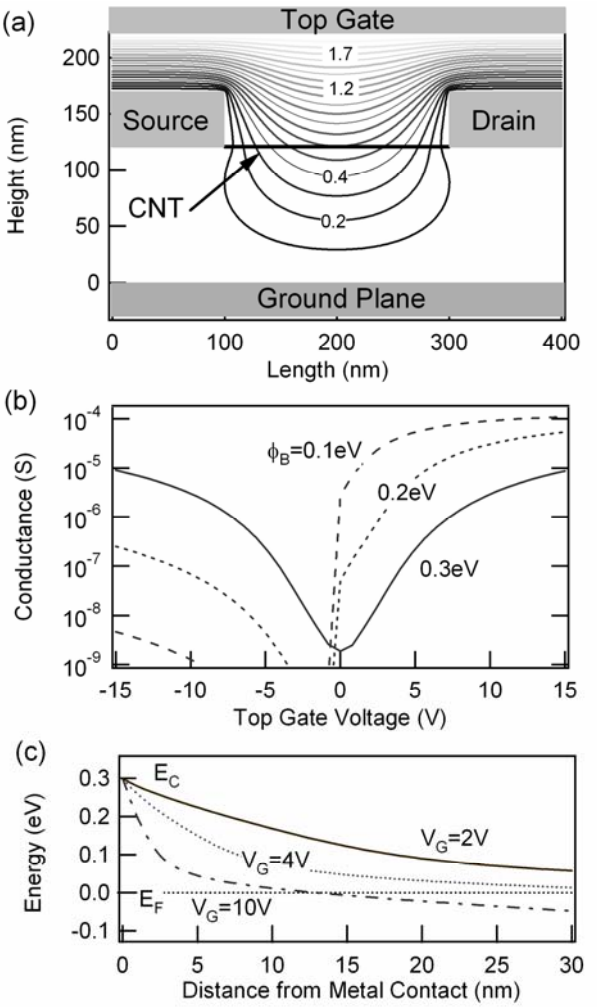

Figure 4.18: (a) Sketch of a carbon nanotube transistor with Schottky barriers at source and drain. Solid lines are lines of constant electrostatic potential for a gate voltage of 2 V. (b) Room temperature conductance as a function of gate voltage for different values of the electron Schottky barrier height. (c) Band-bending near the contact for two values of the gate voltage. *Figure from S. Heinze.*

band edges and reaches a maximum at midgap. The dependence of Im(*k)* on energy is shown in Figure 4.19 for a nanotube with 0.6 eV bandgap. For this nanotube, the maximum value of the imaginary part of k is about 0.5 nm^{-1} corresponding to a decay length of about 2 nm; one might therefore expect that tunneling will become important when the tunneling length is less than about 10 nm.

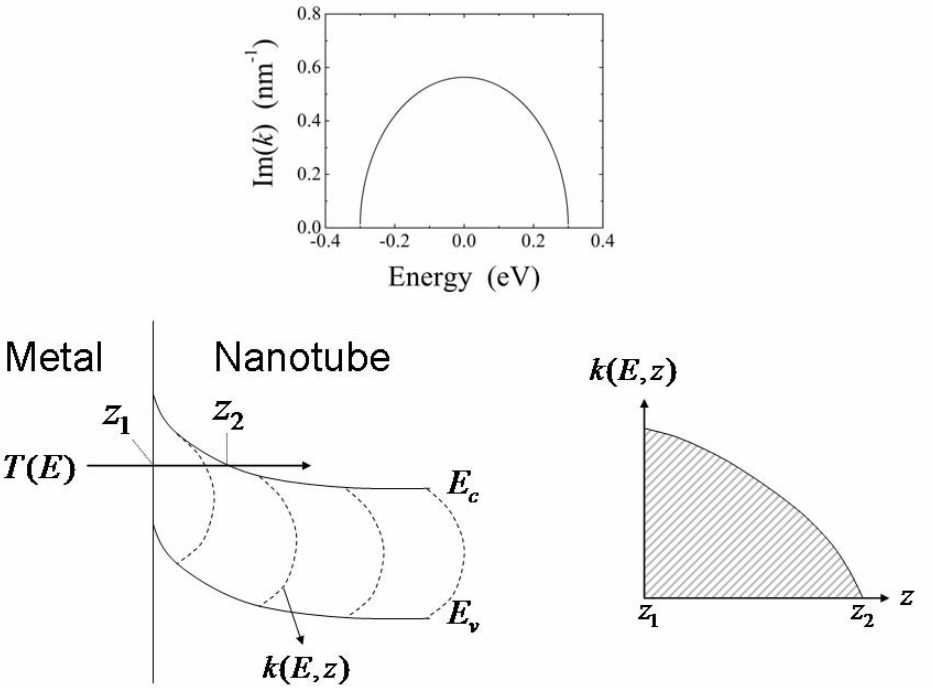

Figure 4.19: Top panel: Imaginary part of the electron wavevector for a carbon nanotube with a bandgap of 0.6 eV. The imaginary part of the wavevector is zero at the band edges and reaches a maximum at midgap. Bottom left panel: Illustration of tunneling through band-bending at a nanotube/metal contact, with the transmission probability $T(E)$, the imaginary part of the wavevector $k(E,z)$ and the two classical turning points z_1 and z_2. Bottom right panel: Sketch of the imaginary part of the wavevector as a function of distance along the nanotube. The area under the curve enters the calculation of the transmission probability.

To calculate the tunneling probability from Equation (4.39) the spatial dependence of k is also needed. We assume that the electrostatic potential $V(z)$ is known and simply shift the nanotube bandstructure locally by this electrostatic potential. Therefore, the energy and spatial dependence of the imaginary part of the wavevector is

$$k(E,z)=\frac{2}{3a\gamma}\sqrt{\left(E_g/2\right)^2-\left[E+eV(z)\right]^2} \tag{4.42}$$

and the transmission probability is

$$T(E)=\exp\left\{-\frac{4}{3a\gamma}\int_{z_1}^{z_2}\left(\frac{E_g^2}{4}-\left[E+eV(z)\right]^2\right)^{1/2}dz\right\}. \tag{4.43}$$

Once the transmission probability is known for the relevant range of energies, the conductance is calculated from

$$G=\frac{4e^2}{h}\int_{-\infty}^{\infty}T(E)\left(-\frac{\partial f}{\partial E}\right)dE. \tag{4.44}$$

Figure 4.18b shows the calculated conductance as a function of gate voltage for different values of the electron Schottky barrier (the barrier for holes is simply the nanotube bandgap minus the electron barrier). The main curve in this figure corresponds to a situation where the metal Fermi level is exactly in the middle of the nanotube bandgap. The barriers for electrons and holes are equal and the device turns on at both positive (electron conduction) and negative (hole conduction) values of the gate voltage. For unequal electron and hole barriers, the conductance versus gate voltage curve is asymmetric, with much larger current for the side with the smallest barrier (in the case of Figure 4.18b the smallest barrier is for electrons). The origin of the conductance modulation by the gate voltage is illustrated in Figure 4.18c. There it is shown that an increase of the gate voltage leads to a sharper band-bending in the vicinity of the contact, substantially reducing the tunneling length at the Fermi level (from 40 nm to 10 nm in this particular case). Since tunneling probabilities depend exponentially on tunneling length, the device conductance increases very rapidly with increase in gate voltage.

4.3.3 Subthreshold Swing

A key performance parameter of field-effect transistors is the subthreshold swing S defined as

$$S=\left|\frac{d\log_{10}G}{dV_g}\right|^{-1} \tag{4.45}$$

and is a measure of the effectiveness of the gate in changing the channel conductance. In a traditional channel-controlled transistor, the smallest subthreshold swing is achieved when a change in the gate voltage leads to the same shift of the bands, i.e. the gate fields are unscreened. This corresponds to a situation where

$$\frac{dE_{c,v}}{d\left(eV_g\right)}=-1. \tag{4.46}$$

The conductance can be calculated from the expression in Equation (4.44), assuming that the transmission probability is equal to 1 for energies below the minimum of the valence band. Under this assumption, the conductance reads

$$G = 2G_0 \exp\left(\frac{E_v - E_F}{kT}\right) \tag{4.47}$$

and the subthreshold swing is

$$S = \left|\frac{d\log_{10} G}{dV_g}\right|^{-1} = \left(\frac{d\log_{10} G}{d\left(E_v / e\right)}\right)^{-1} = \frac{kT}{e}\ln 10 \tag{4.48}$$

which is equal to 60 meV/decade at room temperature.

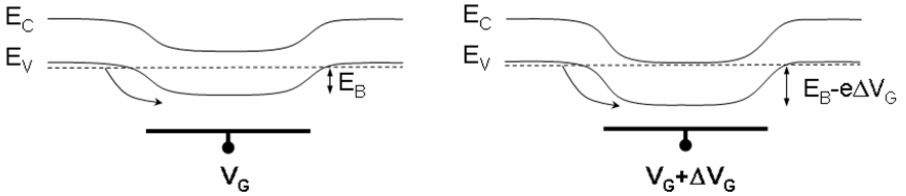

Figure 4.20: Illustration of the ideal subthreshold swing in a *p*-type field-effect transistor. At finite temperature, the small bias conductance is governed by thermal excitation of holes over a potential barrier E_B in the middle of the channel. The ideal device behavior occurs when a change ΔV_G in the gate voltage is unscreened and leads to a change $-e\Delta V_G$ in the position of the band edge in the middle of the channel.

This traditional derivation of the subthreshold swing assumes that the device characteristics are controlled by band-bending in the channel and that conduction is determined by thermal excitation of carriers; it applies to carbon nanotube field-effect transistors with ohmic contacts. In practice, the gate oxide screens the gate fields, and the coupling between the gate and the channel is not perfect, which causes the subthreshold swing to be larger than the ideal value given by Equation (4.48). The typical strategy to reduce the subthreshold swing is to decrease the oxide thickness to improve the coupling, and this will be discussed in the next section. However, because the switching mechanism is fundamentally different in Schottky barrier nanotube transistors, the geometrical requirements for improving device performance are also quite different. Indeed, in the case of the Schottky barrier nanotube transistor the band-bending at the contact is the key issue that affects device performance, and enhancing the effect of the gate on this band-bending is the key challenge in device design. For example, it is clear from the equipotential lines in Figure 4.18a that the contacts significantly screen the gate-induced electric fields. Two ways have been proposed to improve the impact of the gate voltage on the band-bending at the contact [Hei02]. The first approach follows that of conventional transistors and consists of decreasing the gate oxide thickness. Figure 4.21 shows the calculated conductance versus gate voltage curve for the device of Figure 4.18a, for a fixed contact thickness of 50 nm, but with oxide thicknesses ranging from 60 to 120 nm. As expected, this improves the device performance, with a larger ON conductance and lower threshold voltage. More importantly, the dash-dotted line in Figure 4.21 shows that reducing the thickness of the contact from 50 nm to 5 nm leads to

further performance improvement, a mechanism that does not occur in traditional channel-controlled devices. From this insight, it is clear that the ultimate behavior is obtained for a needle-like contact and a cylindrical gate, as the open circles in Figure 4.21 indicate.

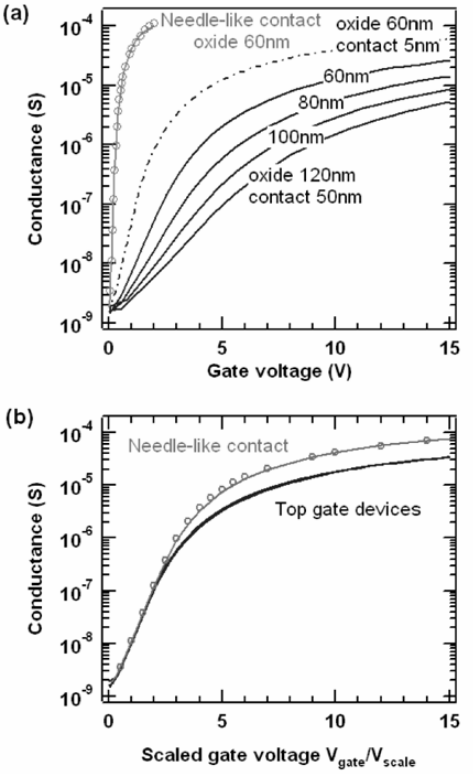

Figure 4.21: (Top): Influence of transistor geometry on the device characteristics of Schottky barrier nanotube transistors. The four right-most curves correspond to different thicknesses of the oxide above the nanotube, as labeled, with all other parameters as in Figure 4.18. (The curve for 100 nm is the same as that in Figure 4.18b for a Schottky barrier height of 0.3 eV.) The dot-dashed curve shows the conductance when the contact thickness is reduced to 5 nm. The curve at the left corresponds to a needlelike metal electrode and cylindrical gate. Open circles are calculated as for the other curves, solid curve uses the exact electrostatic kernel. All calculations are at room temperature. (Bottom): The graph shows the same curves, but for each curve the gate voltage is rescaled by the voltage at which the conductance is 10^{-8} S. (The five rightmost curves cannot be distinguished on this scale, forming a single line). *Figure from S. Heinze.*

The much different physics behind the operation of Schottky barrier nanotube transistors has important implications on the scaling of various performance parameters with device dimensions. As discussed above, it was predicted that reducing the thickness of the gate insulator improves the subthreshold swing because it allows the gate to more effectively modulate the band-bending at the contact. Such a behavior has been verified experimentally by fabricating nanotube transistors with gate oxide thickness between 2 and 20 nm. The current-voltage characteristics of these transistors shown in Figure 4.22 clearly indicate that the devices with the smaller gate oxides have smaller subthreshold swings. This behavior is quantitatively plotted in the right panel of Figure 4.22 which shows that the subthreshold swing decreases rapidly as the gate oxide thickness is

reduced, but that the improvement is smaller for reduction of the gate oxide thickness below 5 nm. This can be understood by considering the unscreened band-bending at the contact for a double-gate geometry with a vanishingly thin contact[Hei03], as illustrated in the top inset in the right panel of Figure 4.22. (By unscreened, we mean that the charge on the carbon nanotube is not considered in the calculation. This is a good approximation as long as the Fermi level is more than kT away from the band edge.) For this device geometry, the potential near the contact is [Hei03]

$$V(z) = \frac{2V_g}{\sqrt{\pi}} \left(\frac{z}{t_{ox}} \right)^{1/2} \tag{4.49}$$

where t_{ox} is the gate oxide thickness. Substituting this expression in Equation (4.43) gives

$$\begin{aligned} T(E) &= \exp\left\{ -\frac{4}{3a\gamma} \int_{z_1}^{z_2} \left(\frac{E_g^2}{4} - \left[E + \frac{2eV_g}{\sqrt{\pi}} \left(\frac{z}{t_{ox}} \right)^{1/2} \right]^2 \right)^{1/2} dz \right\} \\ &= \exp\left\{ -\left(\frac{\tilde{V}}{V_g} \right)^2 h(s) \right\} \end{aligned} \tag{4.50}$$

where

$$h(s) = \int_{\min(0,-1-s)}^{1-s} t\left(1 - \left[s+t\right]^2\right)^{1/2} dt \tag{4.51}$$

and

$$\tilde{V} = \left(\frac{\pi E_g^3}{12a\gamma} \right)^{1/2} t_{ox}^{1/2}. \tag{4.52}$$

Since the transmission depends on the ratio $\tilde{V}/V_g$, $\tilde{V}$ plays the role of a rescaling voltage. The conductance is calculated from

$$G = \frac{4e^2}{h} \int_{-\infty}^{\infty} T(E) \left(-\frac{\partial f}{\partial E} \right) dE = \frac{2e^2 E_g}{hkT} H(x,y) \tag{4.53}$$

with

$$H(x,y) = \int_{-\infty}^{\infty} e^{-h(s)/x^2} \left(-\frac{e^{sy}}{\left(1+e^{sy}\right)^2} \right) ds \tag{4.54}$$

and

$$x = \frac{V_g}{\tilde{V}}, \quad y = \frac{E_g}{2kT}. \tag{4.55}$$

These expressions can be simplified by assuming that $y = E_g / 2kT \gg 1$ to obtain

$$G = \frac{4e^2}{h} \exp\left[-\frac{1}{3} \left(\frac{\tilde{V}}{V_g} \right)^2 \right]. \tag{4.56}$$

The conductance increases exponentially with gate voltage and scales with $t_{ox}^{1/2}$ since $\tilde{V}$ is proportional to $t_{ox}^{1/2}$. The threshold voltage can be defined as the gate voltage at which the conductance decreases by a factor of 1000 from its saturation value:

$$V_{th} = \frac{1}{\sqrt{9 \ln 10}} \tilde{V} \propto t_{ox}^{1/2}. \tag{4.57}$$

Thus the threshold voltage is also zproportional to $t_{ox}^{1/2}$. The subthreshold swing at the threshold voltage is then

$$S = \left(\frac{d \log G}{dV_g} \right)^{-1} \Bigg|_{V_{th}} \propto \tilde{V} \propto t_{ox}^{1/2} \tag{4.58}$$

and is also proportional to $t_{ox}^{1/2}$. The important point is that this scaling behavior of the subthreshold swing with gate oxide thickness is different from that in conventional MOSFETS because the switching is controlled by Schottky barriers at the contacts.

As the thickness of the oxide becomes quite small, the above analysis breaks down since the fields at the contact no longer follow the simple expression used for the calculations. But in the limit of very thin oxides, the subthreshold swing must approach the theoretical limit of $(kT/e)\ln 10$; a simple empirical expression for *S* is

$$S = \left[\alpha t_{ox} + \left(\frac{kT}{e} \ln 10 \right)^2 \right]^{1/2}. \tag{4.59}$$

The solid lines in the right panel of Figure 4.22 show that such an expression describes well the experimental and computational results. It is important to note that the scaling with $t_{ox}^{1/2}$ is a consequence of the particular form of the electrostatic potential at the contact. For a different device geometry this functional form may be different giving another scaling form. In particular, for an infinitely thick contact [Hei03], the electrostatic potential varies as $z^{2/3}$ leading to a scaling $t_{ox}^{2/3}$.

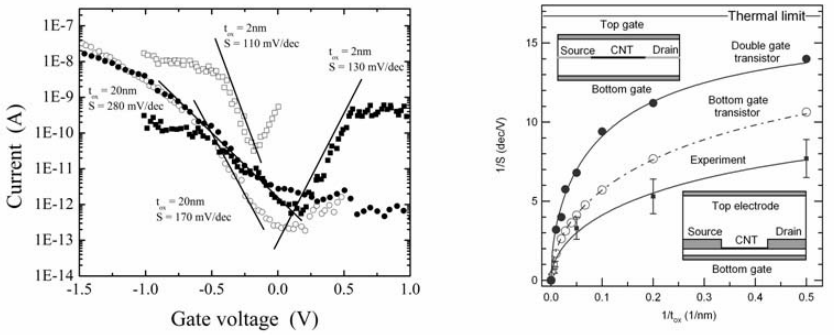

Figure 4.22: The left panel shows the current vs gate voltage characteristics of Schottky barrier carbon nanotube transistors for different gate oxide thicknesses. The right panel shows the measured and calculated (see text for detail) subthreshold swing as a function of the inverse gate oxide thickness. *Figures from Ref. [Hei03] and from S. Heinze.*

4.3.3 High-κ Dielectrics

High-κ dielectrics for gate insulators have been the subject of intense research in the semiconductor industry. The driving force behind this activity is that reducing the channel length in transistors also requires a reduction of the gate oxide thickness to maintain the gate efficiency. For SiO_2, current technologies utilize a thickness of only a few atomic layers, and further reduction of this thickness leads to channel-gate leakage current because of tunneling. For this reason, alternative materials to SiO_2 are being explored, with the requirement that they have the same or larger gate capacitance and low leakage current. Because capacitance is essentially the ratio of the dielectric constant to the gate insulator thickness, high-κ dielectrics allow the gate oxide thickness to be increased (to limit leakage) while maintaining or even increasing the capacitance. At the time of writing of this manuscript, major semiconductor manufacturers are announcing that chips using HfO_2 ($\kappa = 25$) as the gate insulator will soon be available.

Integrating new high-κ dielectrics in existing silicon technology is a complex process which is very much dependent on the ability to grow high-quality thin oxide layers and the ability to form good interfaces with the channel material. The same issues are faced in carbon nanotube transistors. While it has been demonstrated that SiO_2 is a relatively good substrate for carbon nanotubes, it is by no means obvious that high-κ dielectrics will form good interfaces with the nanotubes. An example of this is shown in Figure 4.23 [Lu06]. There the goal is to coat a carbon nanotube with a thin layer of HfO_2 as a gate insulator. Unfortunately, the HfO_2 has poor wetting properties on bare carbon nanotubes, and it is impossible to form a continuous, uniform layer directly on the nanotubes. However, Figure 4.23 also shows that coating the nanotubes with DNA allows the formation of a good oxide layer. Transistors fabricated with such DNA-coated nanotubes and a 3nm HfO_2 gate insulator can achieve low subthreshold swings because of the large gate capacitance.

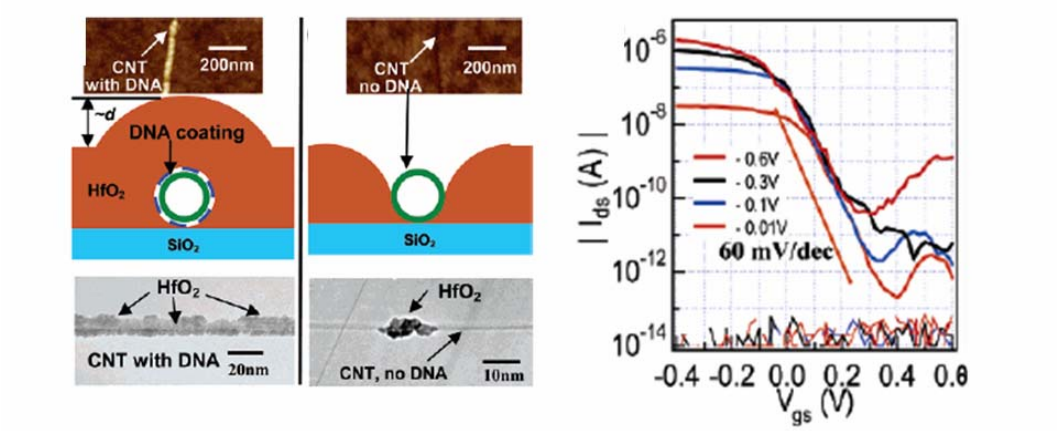

Figure 4.23: Left image: Growth of HfO_2 by atomic layer deposition on carbon nanotubes with (a) and without (b) a DNA coating. Top images in both panels are atomic force microscope images of carbon nanotubes with about 5 nm of HfO_2. Middle panels are sketches of the expected wetting behavior. The bottom panels show transmission electron microscopy images of a 5 nm layer of HfO_2 on the carbon nanotubes. Good wetting is seen in the case of DNA-functionalized nanotubes. Right image: Transfer characteristics of a carbon nanotube transistor with 3 nm of HfO_2 as the gate dielectric. The subthreshold swing approaches the theoretical room-temperature limit of 60 mV/decade. *Figures from Ref. [Lu06].*

One high-κ material that has been found to form good interfaces with nanotubes when grown using atomic layer deposition is ZrO_2 [Jav02a]. Figure 4.24 shows that the deposition of ZrO_2 over a carbon nanotube sitting on SiO_2 gives a highly conforming interface between the ZrO_2 and the nanotube. This property allows the fabrication of top-gated field-effect transistors with thin layers of high-κ dielectric, which show low subthreshold swing and low leakage current (Figure 4.25).

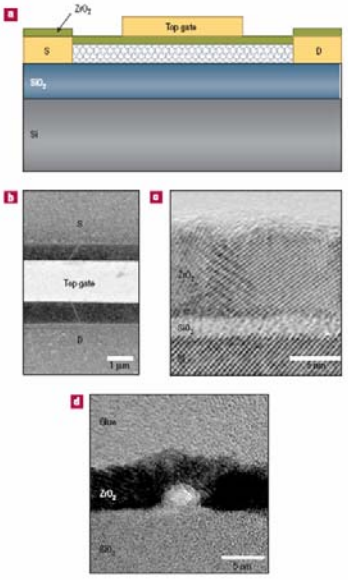

Figure 4.24: ZrO_2 as a high-κ dielectric for carbon nanotube field-effect transistors. Panel (a) shows a sketch of the device with the top gate geometry. (b) Scanning electron microscopy image of a carbon nanotube transistor viewed from the top. (c) Tranmission electrom microscopy (TEM) image of the interface between ZrO_2 and SiO_2. (d) TEM image of the ZrO_2/SiO_2 interface including a carbon nanotube (bright circular spot in the image). *Figure from Ref. [Jav02a].*

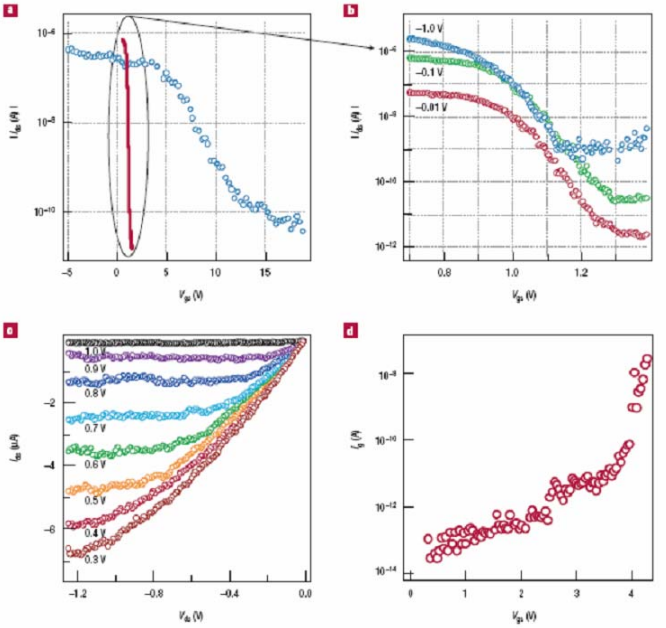

Figure 4.25: Characteristics of a *p*-type nanotube field-effect transistor with high-*κ* gate insulator. a) I_{ds}–V_{gs} curves recorded under V_{ds}=–100 mV with bottom Si/SiO_2 gate (circles) and top-gate/ZrO_2 (solid curve) respectively.The bottom-gate was grounded during top-gate operation. b)I_{ds}–V_{gs} curves recorded with top-gate/ZrO_2 at bias voltages of V_{ds}=–10 mV,–0.1 V and –1 V,respectively. c) I_{ds}–V_{ds} curves of the transistor recorded for various top-gate voltages at 0.1 V steps. d) Gate leakage current versus top-gate voltage. The leakage current is negligible (at the pA level) until V_{gs} > 3 V. *Figure from Ref. [Jav02a].*

4.3.4 Logic Circuits

The discovery of transistor behavior in carbon nanotubes was critical in demonstrating their potential for future electronic devices. The next step in this progress is to demonstrate that combining several devices can lead to logic function. While this may appear relatively simple at first glance, it requires the ability to control and combine the properties of *p*-type and *n*-type devices, and carbon nanotube devices are typically *p*-type. One strategy to form *n*-type nanotubes is to utilize the strong effects of gases on the nanotube electronic properties (this will be discussed in more detail in Chapter 8). For example, Figure 4.26 shows that vacuum annealing or exposure to potassium of initially *p*-type carbon nanotube transistors converts them to *n*-type devices. For the vacuum annealed nanotubes, this conversion to *n*-type behavior can be reversed by exposing the nanotube to oxygen, as indicated in Figure 4.27.

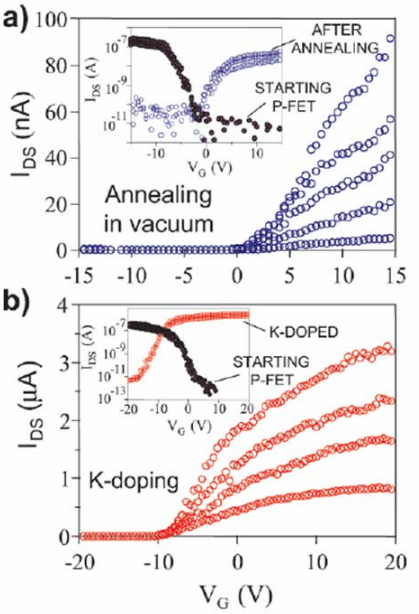

Figure 4.26: Conversion of *p*-type carbon nanotube transistors to *n*-type by (a) vacuum annealing at 700K for 10 min, (b) exposure to potassium. *Figure from Ref. [Der01].*

Therefore, one strategy to make both *p*-type and *n*-type devices for logic circuits is to first convert the nanotubes to *n*-type by vacuum annealing. The nanotubes that are to stay *n*-type are protected by PMMA, while the others are exposed to oxygen and converted to *p*-type. The two types of transistors are then electrically connected to form an inverter circuit (illustration Figure 4.27c). An example of the measured V_{in} versus V_{out} characteristics of such devices is shown in panel (d) of Figure 4.27, indicating that a positive V_{in} gives a negative V_{out}, and vice-versa; the device thus inverts the input voltage.

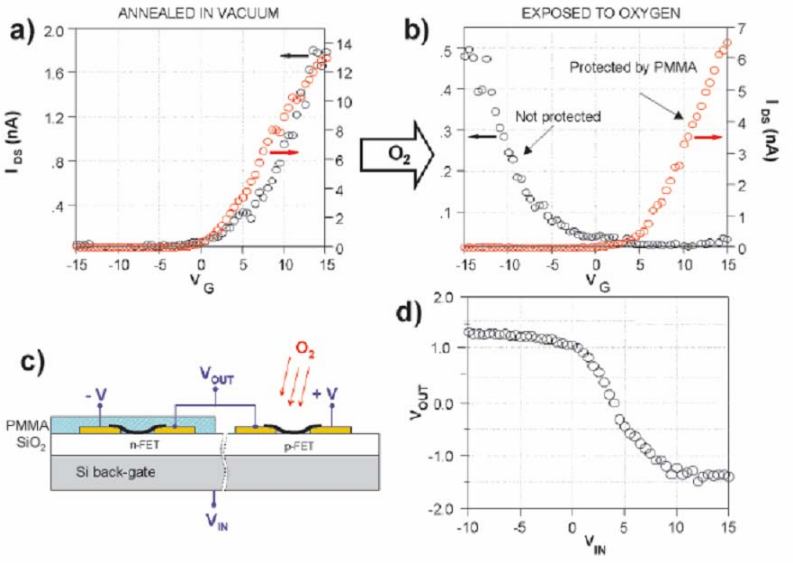

Figure 4.27: Fabrication of a voltage inverter ("NOT" logic gate) using two nanotube field-effect transistors. Initially the two transistors are *p*-type. One of them is protected by PMMA, the other is not. (a) After vacuum annealing both transistors are converted to *n*-type. (b) The two devices are exposed to oxygen (10^{-3} Torr of oxygen for 3 min). The unprotected transistor (black curve) converts back to the original *p*-type, while the protected nanotube (red curve) remains *n*-type. (c) The two complementary transistors are wired as shown in the schematic. (d) Characteristics of the resulting intermolecular inverter are shown. *Figure from Ref. [Der01].*

More complicated logic circuits using a larger number of individual nanotube transistors have also been demonstrated. Figure 4.28 shows the operation of NOR, OR, NAND, and AND gates, as well as an oscillator, utilizing up to six individual nanotube transistors. While the frequency of the ring oscillator is rather low, this is a consequence of the simple connection of the transistors through coaxial cables, and the performance can in principle be significantly improved by direct interconnects on the chip.

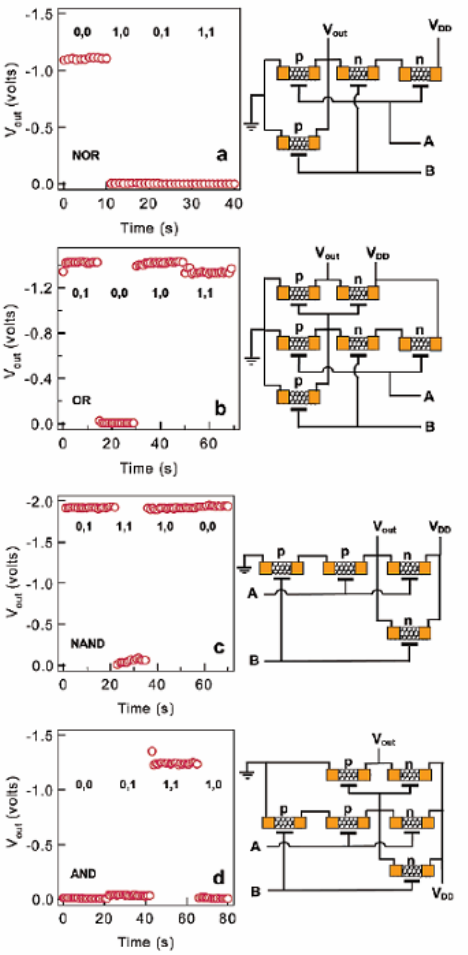

Figure 4.28: Several types of logic circuits using several carbon nanotube transistors. (a) NOR; (b) OR; (c) NAND; (d) AND gates. The bottom figure shows a ring oscillator. *Figure from Ref. [Jav02b].*

Recent progress in carbon nanotube logic circuits is the fabrication of these circuits on a *single* carbon nanotube. Figure 4.29 shows a single bundle of carbon nanotubes making contact to three electrodes, producing two *p*-type transistors in series. This nanotube is then covered with PMMA, and a window is opened between two of the electrodes. Exposure of this channel to potassium causes a transformation from *p*-type to *n*-type, and a common back gate to the two channels is used to operate the device as an inverter.

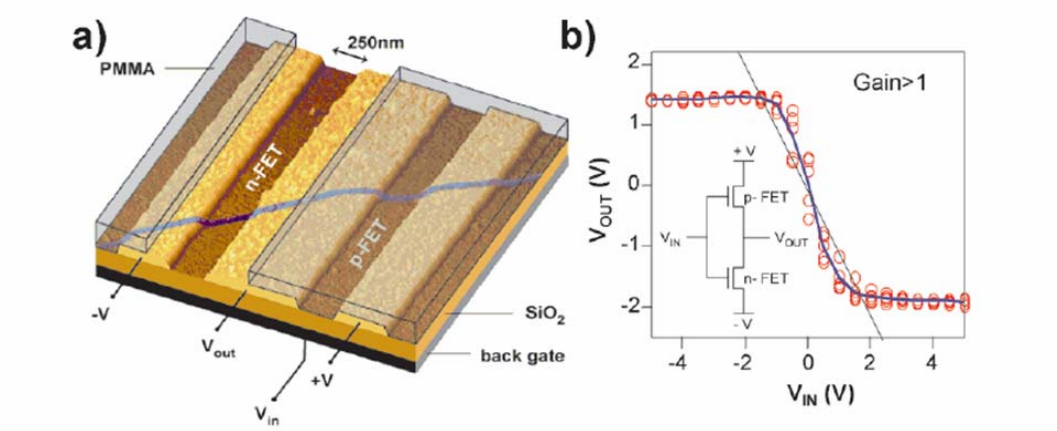

Figure 4.29: a) Atomic force microscope image showing the geometry of a carbon nanotube inverter. A single nanotube bundle is positioned over the gold electrodes to produce two *p*-type nanotube field-effect transistors in series. The device is covered by PMMA and a window is opened by e-beam lithography to expose part of the nanotube. Potassium is then evaporated through this window to produce a *n*-type field-effect transistor, while the other nanotube remains *p*-type. (b) Characteristics of the resulting intramolecular voltage inverter. Open red circles are raw data for five different measurements on the same device (V = ± 2 V). The blue line is the average of these five measurements. The thin straight line corresponds to an output/input gain of one. *Figure from Ref. [Der01].*

Extension of this idea of making a circuit entirely on a single carbon nanotube has recently shown that multistage ring oscillators can be fabricated using long carbon nanotubes. Figure 4.30 shows a scanning electron microscope image of an 18 micron carbon nanotube with multiple source, drain and gates. The arrangement actually consists of 10 nanotube transistors (5 *p*-type, 5 *n*-type) arranged in goups of two (1 *p*-type, 1 *n*-type) to form 5 inverters in series. Interestingly, the approach to make the *p*-type and *n*-type transistors is different from what we have already discussed. Here, two different metals (Pd and Al) are used as the gate material to make the two types of transistors. The idea is that the large workfunction difference between Pd and Al (about 1 eV) effectively acts like an extra gate voltage, shifting the I-V curve on the V_g axis. Thus, one takes advantage of the ambipolar nature of the transport, and brings the positive V_g turn-on in the window of the negative V_g turn-off: as one device turns-off the other turns on giving the inverter characteristics. This is illustrated in Figure 4.30b which shows about a 1 V shift of the Pd I-V curve to the left when Pd is replaced with Al.

The frequency dependence of the signal is shown in Figure 4.30d for different values of V_{dd} showing resonance frequencies in the range of 13 to 52 MHz. The magnitude of the resonant frequency and dependence on V_{dd} can be understood as follows: the expected frequency for a N-stage ring oscillator is [Che06]

$$f = \frac{1}{2\tau N} \tag{4.60}$$

where τ is the stage delay time given by the RC constant

$$\tau = RC = \frac{VC}{I}. \tag{4.61}$$

Therefore, the resonance frequency can be expressed as

$$f = \frac{I}{2CVN}. \tag{4.62}$$

The capacitance per stage in the oscillator is estimated to be 1.8 fF, while for $V_{dd} = 0.92$ V the DC current in a single transistor is about 1 μA. For a 5 stage oscillator this gives a frequency $f = 59$ MHz in good agreement with the experiment. The point is that this measured resonance frequency is due to the parasitics and not to the intrinsic properties of the nanotube. The increase of the resonance frequency with increasing V_{dd} is a consequence of the increase in the current through a single transistor as V_{dd} is increased (i.e. the current is not saturated with V_{gs}). Another consequence of this effect seen in Figure 4.30d is the increase of the signal power with increase in V_{dd}. The signal attenuation is due to the impedance mistmatch between the output of the ring oscillator and the spectrum analyzer.

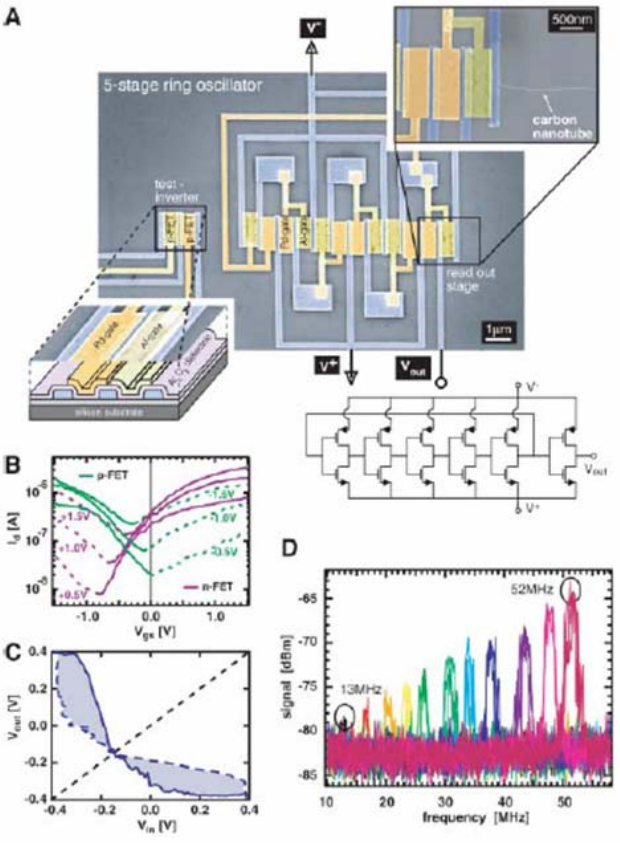

Figure 4.30: Multistage ring oscillator fabricated on a single carbon nanotube. (a) Scanning electron microscope image of the five-stage ring oscillator, as well as test inverter. (b) Current-

voltage characteristics of the *p*-type transistors with Pd gate, and the *n*-type transistors with Al gate. (c) Characteristics of the test inverter. (d) Frequency response of the ring oscillator. The different colored curves correspond to different values of V_{dd} equal to 0.5 V and 0.56 V to 0.92 V (in 0.04 V imcrements) from left to right. *Figure from Ref. [Che06].*

4.3.5 Mobility

While carbon nanotubes are lauded for their promise as ballistic conductors, for long enough channel lengths, scattering with defects, phonons, etc. eventually leads to diffusive behavior, and the introduction of a mobility μ.

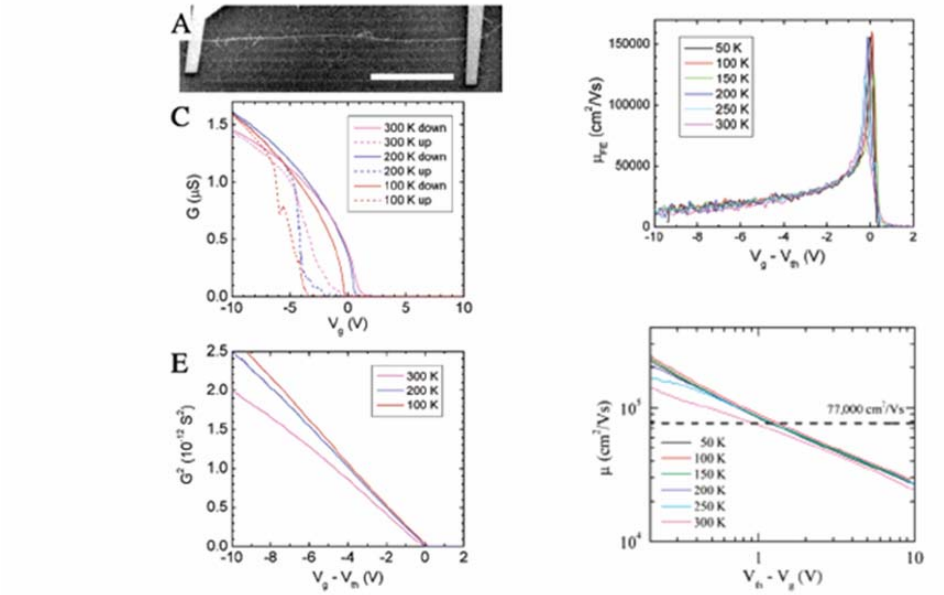

Figure 4.31: Behavior of long-channel carbon nanotube transistors, indicating extremely large field-effect and intrinsic mobilities. The scale bar in A is 100 μm long. *Figure from Ref. [Dur04].*

The mobility of the charge carriers is defined through the relation

$$\mu = \sigma / n \tag{4.63}$$

where σ is the conductivity and n is the carrier concentration. For diffusive transport, the conductivity is related to the conductance G as

$$\sigma = GL \tag{4.64}$$

where L is the length of the nanotube. For a *p*-type channel near threshold, the charge density is proportional to deviations of the gate voltage from the threshold voltage:

$$n = C\left(V_{th} - V_g\right) \tag{4.65}$$

and the mobility is therefore

$$\mu = \frac{L}{C}\frac{G}{V_{th} - V_g}. \tag{4.66}$$

Another quantity of interest when evaluating device performance is the field-effect mobility defined as

$$\mu_{FE} = \frac{L}{C}\frac{\partial G}{\partial V_g}. \tag{4.67}$$

Figure 4.31 shows the measured field-effect mobility, which attains a value of 79 000 cm^2/Vs at room temperature, almost two orders of magnitude larger than that typically achieved in Si MOSFETs. The intrinsic mobility exceeds the highest reported mobility in bulk materials, 77 000 cm^2/Vs, the Hall-mobility of InSb. (It should be noted that subsequent experiments [Zho05] have obtained field effect mobilities in the range 1500-20 000 cm^2/Vs and have suggested that the very large mobilities could origininate from a non-uniform response of the nanotube to the gate field.)

Without a detailed atomistic study of the particular nanotube used in the experiments, it is difficult to ascertain the role of defects in causing the diffusive behavior. However, phonons undoubtedly cause electron scattering, leading to diffusive behavior. Theoretical work has addressed this issue by considering electron-phonon scattering in semiconducting carbon nanotubes [Per05]. The approach is similar to what was presented in Chapter 2 for electron-phonon scattering in metallic carbon nanotubes: the scattering time and scattering length are calculated using Fermi's Golden rule for the electron-phonon interaction and the Boltzman equation is solved in the presence of an electric field to obtain the carrier velocity and mobility as a function of the electric field from the non-equilibrium carrier distribution. The essential difference here is the much different band structure of semiconducting and metallic nanotubes.

The results of such calculations are presented in Figure 4.32 for a (25,0) semiconducting carbon nanotube. We first note that the binding energy, representing the distortions of the bandstructure due to the electron-phonon interaction, tends to be small and relatively independent of energy. This effect was discussed in Chapter 1 in the context of the bandgap renormalization by phonons, and will not be discussed further here. Instead, we focus on the scattering time as a function of energy in the conduction band, as shown in the panels of Figure 4.32. It is clear from this picture that the scattering time is sensitive to both energy and temperature. The scattering time can be as small as a few hundreths of picoseconds, but can also reach picoseconds for energies not too far from the conduction band edge. Thus, for devices where injection of carriers is a fraction of an eV in the valence or conduction bands, ballistic transport should be possible in defect-free nanotubes over long distances. It turns out that almost all of the scattering is from phonons near the Γ and K points of the graphene Brillouin zone. The lowest energy phonon band gives negligible scattering. The second lowest energy phonon mode is an acoustic mode that gives large scattering near the band edge, and correspondingly small scattering time. For the (25,0) nanotube, the third lowest energy phonon mode is the radial breathing mode with energy of 15 meV, leading to a dip in the scattering time at about 0.02 eV above the conduction band edge. The next phonon mode with important

electron-phonon scattering is a longitudinal acoustic mode which gives a scattering length of about 5 microns (the scattering length is simply the scattering tim e times the Fermi velocity). Longitudinal optical phonons at the Γ and *K* points give the strongest scattering, significantly reducing the scattering length to the 20-40 nm range.

For device applications and comparison with experiment, what matters is the total effect of all of the phonon modes on the conductivity, as captured by the mobility. Figure 4.33 shows the calculated drift velocity and mobility as a function of the electric field. The maximum drift velocity for this particular nanotube is 5×10^7 cm/s which is about a factor of five larger than silicon. The mobility obtained from the saturation velocity is plotted in Figure 4.33b, which attains values of 5×10^3 cm^2/Vs at fields of 1V/μm. More importantly, the graph indicates that a simple relation exists between the mobility and the electric field:

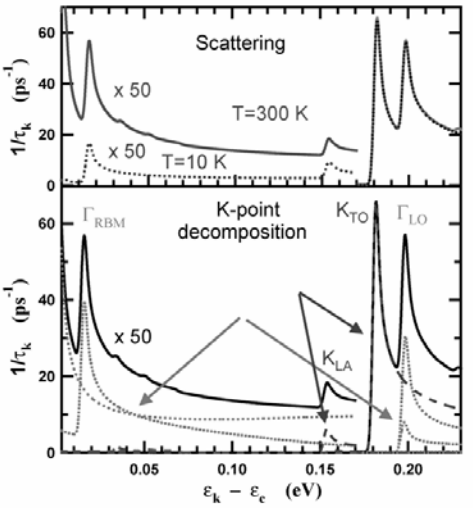

Figure 4.32: Electron-phonon scattering time vs electron energy relative to conduction band edge, for a (25,0) tube. Solid curves are for T = 300 K, dotted curves are for T = 10 K. *Figure from V. Perebeinos.*

$$\mu^{-1} = \mu_0^{-1} + v_s^{-1}E \tag{4.68}$$

where v_s is the saturation velocity and E is the electric field. The zero-field mobility μ_0 takes a wide range of values depending on the diameter of the nanotube: from 15 000 cm^2/Vs for a (13,0) nanotube to 65 000 cm^2/Vs for a (25,0) nanotube. Indeed this dependence on diameter is plotted in Figure 4.33 for different temperatures. The behavior can be captured by the phenomenological relation

$$\mu_0(t,d) = \mu_1 \frac{300K}{T}\left(\frac{d}{1\text{ nm}}\right)^{\alpha} \tag{4.69}$$

with $\mu_1 = 12000$ cm^2/Vs and $\alpha = 2.26$.

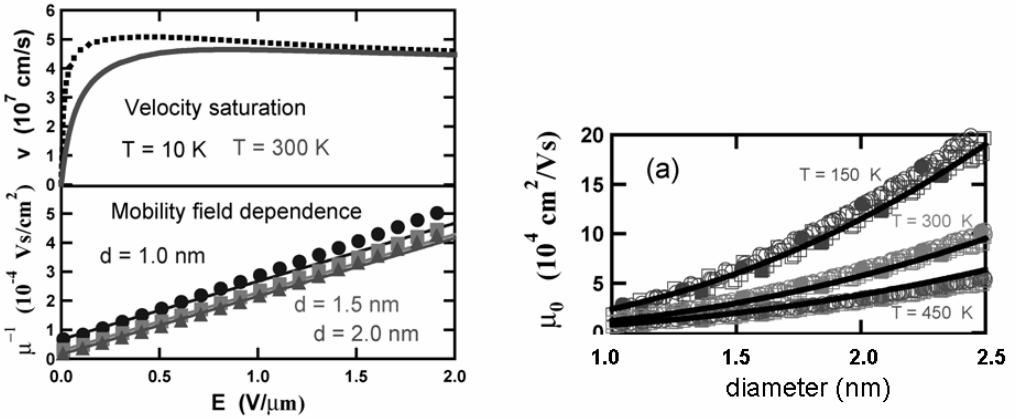

Figure 4.33: Left frame: (Top) Drift velocity vs electric field at T = 300 K (solid curve) and T = 10 K (dotted curve) in a (25,0) nanotube. (bottom) Inverse mobility at T = 300 K vs electric field for a (25,0) nanotube (triangles); (19,0) nanotube (squares); and (13,0) nanotube (circles). Right frame: Zero-field mobility vs tube diameter, for tubes of many different chiralities. Solid lines are from analytical model, Equation (4.69). *Figures from V. Perebeinos.*

Since the mobility due to electron-phonon scattering is sensitive to the presence or absence of available states, carrier concentration can have a strong impact on the mobility because states in the conduction or valence bands become occupied by the carriers. These carriers may originate from doping, or from capacitive coupling to an external gate (see for example Equation 4.65 and the experimental results of Figure 4.31). To explore this effect, theoretical calculations have used a Boltzman transport approach with electron-phonon interactions to calculate the mobility as a function of the carrier density [Per06]. Results of such calculations at low fields are presented in Figure 4.34. The mobility initially increases with an increase in the carrier density, reaches a maximum and decreases at high carrier concentrations. To understand the origin of this behavior, we consider the classical relationship between the mobility, the scattering time and the average effective mass

$$\mu = \frac{e\tau}{m_{av}}. \tag{4.70}$$

The introduction of an average effective mass is because states in the bands are occupied over an energy range that reflects the carrier density, so this leads in principle to an energy-dependent effective mass. Introducing an energy-averaged effective mass allows for a simple analysis using Equation (4.70). This average effective mass is obtained from the equation

$$\frac{1}{m_{av}} = \frac{\int_{E_g/2}^{E_F} m^{-1}(E)dE}{\int_{E_g/2}^{E_F} dE} = \frac{1}{E_F - E_g/2}\int_{E_g/2}^{E_F} \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k^2}dE \tag{4.71}$$

where E_F is the position of the Fermi level for a given doping level. At zero temperature, it is obtained from the expression

$$f = \int_{E_g/2}^{E_F} D(E)dE = D_0 \int_{E_g/2}^{E_F} \frac{E}{\sqrt{E^2 - \left(E_g/2\right)^2}}dE = D_0\sqrt{E_F^2 - \left(E_g/2\right)^2} \tag{4.72}$$

or

$$E_F = \sqrt{\left(f/D_0\right)^2 + \left(E_g/2\right)^2}\ . \tag{4.73}$$

In these equations, f is the doping fraction in electrons/atom and $D_0 = a\sqrt{3}/\pi^2 R\gamma$. To calculate the average effective mass we use the expression for the subband energy

$$E(k) = \frac{E_g}{2}\sqrt{1 + 9R^2k^2} \tag{4.74}$$

to get

$$\frac{\partial^2 E}{\partial k^2} = \frac{9E_g^4 R^2}{8E^3} \tag{4.75}$$

and the average effective mass is

$$m_{av} = \frac{4\hbar^2}{18D_0E_g^2R^2}\frac{4f^2 + \left(D_0E_g\right)^2}{\sqrt{4f^2 + \left(D_0E_g\right)^2} + D_0E_g}\ . \tag{4.76}$$

For typical nanotubes, the product D_0E_g is small, and the above equation simplifies to

$$m_{av} = m_0\sqrt{1 + \frac{4f^2}{D_o^2E_g^2}} \tag{4.77}$$

where m_0 is the effective mass at the band edge. The important point here is that the average effective mass monotonically increases with the carrier concentration (see Figure 4.34b), and is only weakly dependent on temperature, so that the non-monotonic behavior of the mobility with carrier density and its temperature dependence arise from the scattering time. Indeed, from the mobility calculated using the Boltzman transport approach and average effective mass from Equation (4.77), the scattering time can be obtained from Equation (4.70). Figure 4.34c shows the scattering time as a function of

the carrier density. The scattering time initially increases with increasing carrier density since the number of available states for scattering decreases as the Fermi level is higher in the conduction band. At a certain carrier concentration *f*, the Fermi level reaches the second subband, opening up additional channels for scattering; this leads to a decrease of the scattering time. Additional calculations [Per06] show that the scattering time is essentially proportional to the diameter of the carbon nanotube.

The high field behavior is studied by calculating the drift velocity v_d from the occupation function in the Boltzman equation [Per06]. Figure 4.34d shows the field-dependence of v_d for different values of the carrier concentration. In the absence of carriers, the drift velocity reaches a maximum at a field of 1 V/μm and then decreases at larger fields. In contrast, at high carrier concentrations the drift velocity increases continuously. It is tempting to invoke the occupation of the second subband as the reason for the non-monotonic field dependence of the drift velocity. However, comparisons of the results with those of a single-band model (Figure 4.34e) indicate that when the single band is described by a hyperbolic band profile, the Boltzman results are well reproduced. A parabolic model for the single band does not match the full numerical calculations, and fails to give a maximum in v_d. Therefore, the non-parabolicity of the first subband is responsible for the non-monotonic behavior.

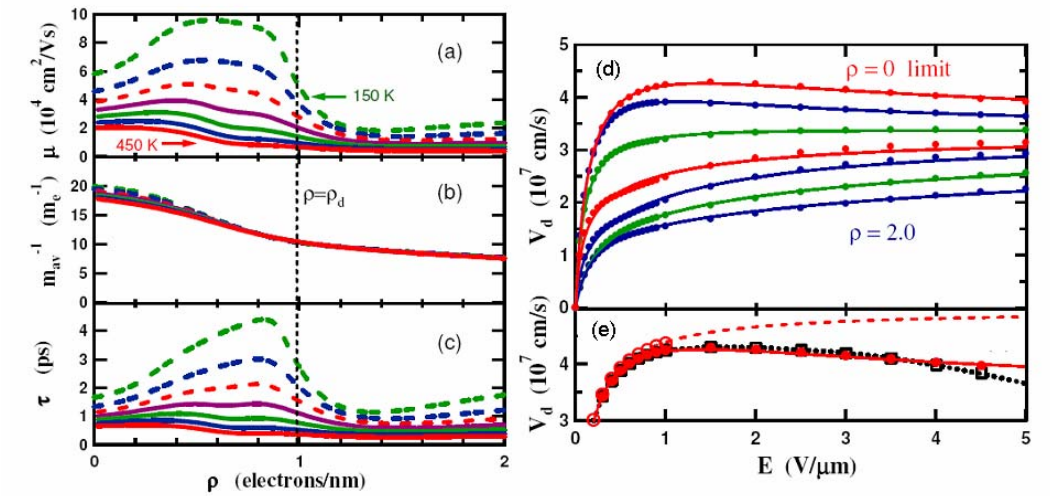

Figure 4.34: Left frame: (a) Electron mobility at low fields for a (19,0) carbon nanotube as a function of the carrier density. (b) Contribution to the mobility of the inverse effective mass. (c) Contribution to the mobility of the scattering time. Temperatures from top to bottom in each panel are 150, 200, 250, 300, 350, 400 and 450 K. The vertical dashed line indicates the doping density at which the second subband becomes occupied. (d) Drift velocity at room temperature as a function of electric field for different carrier densities for a (19,0) nanotube. Curves from top to bottom correspond to carrier densities of 0, 0.25, 0.5, 0.75, 1.0, 1.5 and 2.0 electrons/nm. (e) Results for $\rho = 0$ are compared with single band models: single hyperbolic band (dotted line) and single parabolic band (dashed line). *Figures from Ref. [Per06].*

4.3.6 Short-Channel Effects

Short channel effects in transistors can lead to serious degradation of the device characteristics as the different length scales (channel length, gate insulator thickness, etc.) are reduced to smaller and smaller dimensions. In conventional transistors made with silicon or other bulk semiconductor materials, short channel effects have been extensively studied, and the basic physics is well understood. In contrast, short channel effects in carbon nanotube transistors has received relatively little attention, despite the obvious importance for applications. The key issue here is that, because of the reduced dimensionality of carbon nanotubes, the Coulomb interaction has different implications than for bulk devices. In this section we discuss modeling work aimed at studying short channel effects in carbon nanotube field effect transistors with ohmic contacts [Leo06b].

Figure 4.35 shows a sketch of the device under consideration: it consists of a semiconducting carbon nanotube embedded in source and drain electrodes made of palladium. In the channel region, the nanotube is surrounded by SiO_2, and the SiO_2 is itself wrapped by a cylindrical gate of radius R_G. The simulations use a (17,0) zigzag nanotube, which has a bandgap of 0.55 eV in the tight binding model used for the calculations. The calculations proceed by obtaining self-consistently the charge and the potential on the carbon nanotube. The potential is calculated by solving Poisson's equation in three-dimensions, with the charge on the nanotube and the boundary conditions at the source and drain contacts, the gate, and at the dielectric/air interface. The charge and ultimately the current are calculated using the non-equilibrium green's function technique, allowing for a quantum description of the system in the ballistic transport regime. The metal Fermi level in the contacts is taken to be 1 eV below the nanotube midgap before self-consistency. Results of these calculations for the current as a function of gate voltage are shown in Figure 4.36, for a gate radius of 3 nm. For a channel length of 10 nm, the ON/OFF ratio is about 1000, and the subthreshold swing is 160 mV/decade. An increase of the source-drain voltage from -0.1 V to -0.3 V causes a shift of the threshold voltage by 310 mV, a signature of drain-induced barrier lowering (DIBL), as will be discussed below. Increasing the channel length to 20 nm significantly improves both the subtreshold swing and the DIBL, which become 69 mV/decade and 11.5 mV/decade, respectively. It is worth mentioning that at $V_g = 0$, the nanotube is significantly hole-doped due to the long-range charge transfer from the contacts (left inset in Figure 4.36).

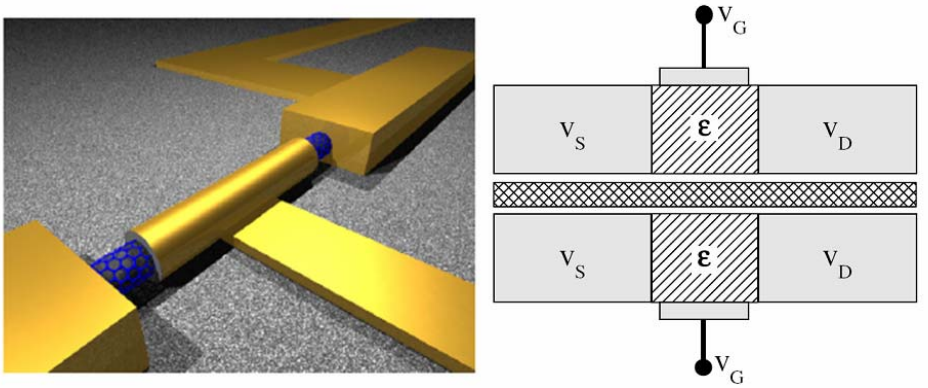

Figure 4.35: Sketch of the nanotube transistor used for device simulations. The nanotube is embedded in metals at its two ends, and in a dielectric in the channel region. The dielectric is wrapped by a cylindrical gate. In the left panel, the separation between the contacts and the central dielectric region is to illustrate the structure in the channel; in the calculations, the contacts touch the dielectric as shown in the right panel. *Figures from Ref. [Leo06b] and [Leo02a].*

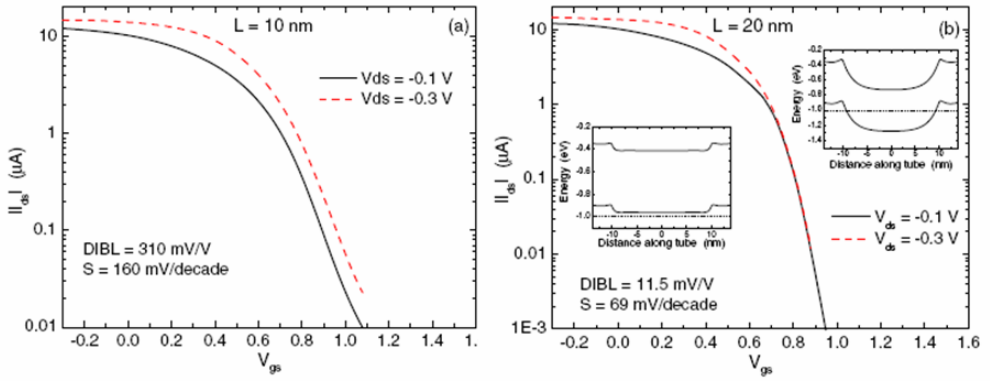

Figure 4.36. Current as a function of gate–source voltage for channel lengths of (a) 10 nm and (b) 20 nm. The gate radius is 3 nm. The insets in (b) show the band bending for V_{gs} = 0 (left) and V_{gs} = 1 V (right). *Figure from Ref. [Leo06b].*

The short channel effects are also seen in the current versus source-drain voltage, as indicated in Figure 4.37. The 20 nm channel device shows saturation of the current for large values of the source-drain voltage, but the 10 nm device shows no saturation at all. The origin of this effect is DIBL, and plots of the band-bending along the carbon nanotube confirm this behavior (Figure 4.38). At $V_g = 0.8\,\mathrm{V}$ and $V_{ds} = 0$ there is a barrier of 93 meV that significantly impairs hole conduction; as the source-drain voltage is increased, the barrier is reduced, causing an increase in the current and a lack of current saturation at large source-drain voltages. This voltage-dependent barrier is also at the origin of the DIBL in Figure 4.36.

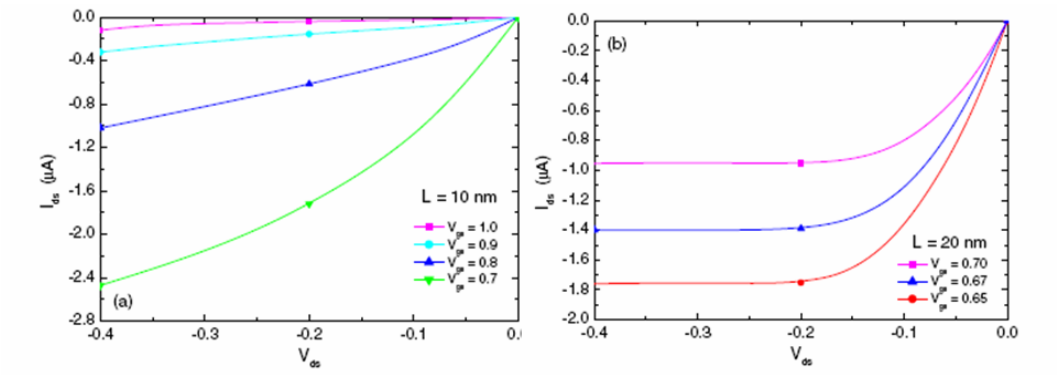

Figure 4.37: Current as a function of drain–source voltage for the carbon nanotube transistor of Figure 4.35 for (a) 10 nm channel length and (b) 20 nm channel length. In both panels the gate radius is 3 nm. *Figure from Ref. [Leo06b].*

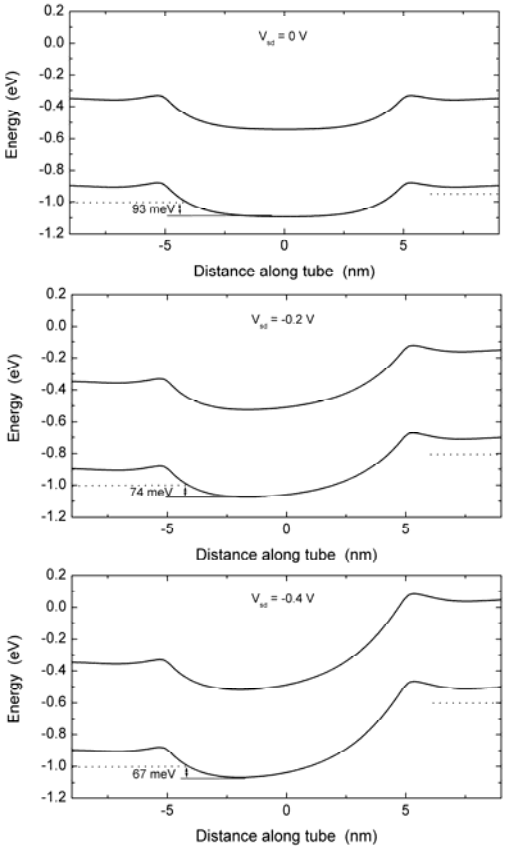

Figure 4.38: Calculated self-consistent band bending for the 10 nm channel device with a 3 nm gate radius, for a gate–source voltage of 0.8 V. The solid lines are the valence and conduction band edges; horizontal dotted lines are the metal Fermi levels in the contacts. *Figure from Ref. [Leo06b].*

One can expect that for longer channel lengths the source and drain contacts will have less impact on the barrier in the middle of the channel, and that the gate will be more effective at modulating the height of that barrier. This is indeed the case, as Figure 4.39 indicates. There it is shown that both the DIBL and the subthreshold swing decrease rapidly with an increase of the channel length. Typical device requirements are that DIBL be less than 100 mV/V and for the subthrehold swing to be less than 80 mV/decade. For the cylindrical nanotube device with a 3 nm radius SiO_2 gate insulator and a (17,0) nanotube, this requires the channel to be larger than 15 nm. This constraint depends on the thickness of the gate insulator. The inset in Figure 4.39 shows that scaling of the DIBL is obtained if the channel length is scaled by the factor

$$\alpha = \sqrt{R_g + l}. \tag{4.78}$$

For the particular device geometry and nanotube discussed here, it is found that $l = 1$ nm. It is important to point out that electrostatic analyses for cylindrical gate transistors [Oh00, Joh03] predict a scaling quantity proportional to the oxide thickness. This type of scaling relation does not lead to collapse of the data onto a single curve. It is thought that the unusual dielectric response of the nanotube [Leo02a], strong charge transfer from the contacts and the actual device geometry render the conventional analyses inapplicable; more work is needed to address these issues.

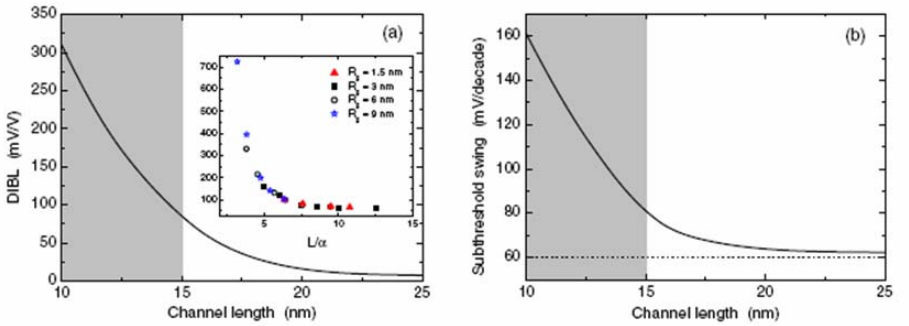

Figure 4.39: Panels (a) and (b) show the variation of DIBL and subthreshold swing on channel length, respectively. Shaded areas are regions where the short channel effects are larger than typical device requirements. The inset in panel (a) shows collapse of the data for several devices of different channel lengths and gate radii upon scaling of the channel length. The horizontal dotted line in panel (b) is the theoretical limit for the subthreshold swing. *Figure from Ref. [Leo06b].*

4.3.7 Crosstalk

Ever increasing computing power not only requires devices with short channels, but also a large density of these devices on a single chip. At high density, interactions between neighbor devices (crosstalk) can become significant, even more so for carbon nanotubes since all of the carriers are constrained to a surface shell and can be impacted by electrical fields. Thus, it is important to assess the role of interactions between nearby

carbon nanotube devices, a topic that is also important for multi-nanotube devices that are becoming more common (an example is shown in Figure 4.40). Computer simulations [Leo06c] are addressing this issue by considering the impact of nanotube density on the characteristics of carbon nanotube transistors, including both semiconducting-semiconducting and semiconducting-metallic interactions. Figure 4.40 shows sketches of planar carbon nanotube transistors with multiple parallel nanotubes in the channel. The key question is how the spacing *d* between nanotubes influences the ON and OFF states of the transistor, as well as the properties of the contacts.

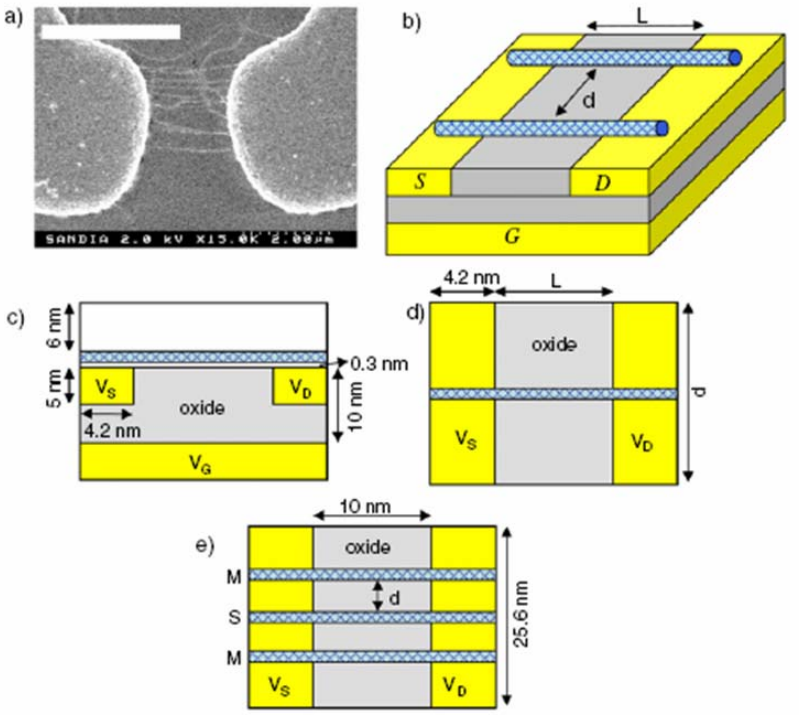

Figure 4.40: Panel (a) shows a scanning electron micrograph image of a carbon nanotube device containing multiple nanotubes. Panel (b) shows a sketch of the device used in the calculations, with the distance *L* indicating the channel length and *d* the nanotube separation. Panel (c) shows a side view of the device, while panels (d) and (e) show top views of the unit cells used to study semiconducting–semiconducting and semiconducting–metallic interactions, respectively. *Figure from Ref. [Leo06c].*

Figure 4.41 shows the results of self-consistent quantum transport calculations for the conductance of a single semiconducting nanotube in the array. The results indicate that the overall conductance decreases as the separation between nanotubes decreases. This behavior has two origins: the first is that in the ON state, there is significant charge on the nanotubes, which interacts with the nearby charge due to other nanotubes. This is energetically costly, and to lower its energy, the system lowers the charge on the nanotubes, effectively decreasing the conductance. The second cause is that because of the charge repulsion, the band alignment in the contact is modified such that the charge

on the nanotubes is decreased; this leads to the metal Fermi level being in the bandgap, and the resulting Schottky barrier decreases the current in the ON state.

To quantify the impact of the nanotube density on the conductance, Figure 4.42 shows the conductance at $V_{gs} = 0$ as a function of the nanotube spacing for different channel lengths. Irrespective of the channel length, the conductance follows a dependence $A\left[1 - \exp\left(-d/\beta\right)\right]$ indicating that it is exponentially sensitive to the nanotube separation. From the figure, it is clear that the parameter β depends on the length of the channel, with a larger value of β for larger channel lengths. Indeed, if one defines a 10% reduction in the conductance as the onset of crosstalk, then plotting the separation d at which this reduction happens as a function of the channel length gives the diagram of Figure 4.42. Clearly, very short channel devices can have very high packing densities, while long channel devices are limited to tube separations of 15 nm. A notable aspect of the results of Figure 4.42 is that the value of d is independent of the channel length for large L. At first glance, one would expect that a larger channel length leads to larger total charge on the nanotubes and thus larger interaction energy. However, screening of the Coulomb interaction by the planar gate leads to a different behavior, as we now discuss.

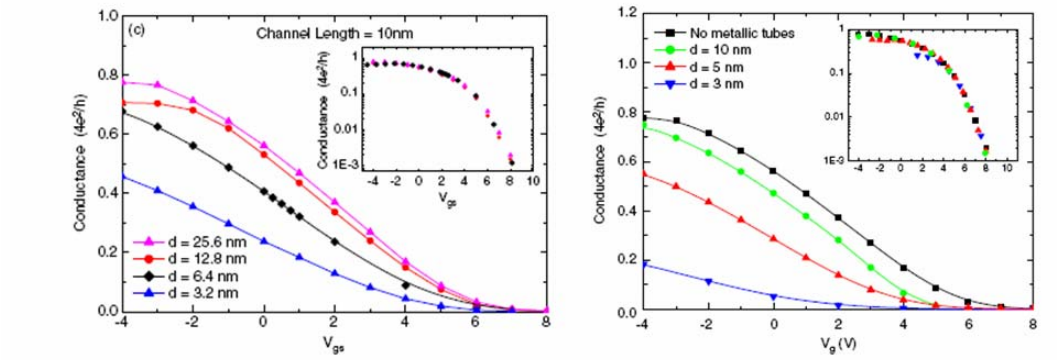

Figure 4.41: Conductance of a (17, 0) nanotube transistor as a function of gate–source voltage, for different separations of the nanotube from other semiconducting nanotubes (left) or metallic nanotubes (right). In each figure, the inset shows the collapse of the data after rescaling and shifting of the gate voltage. *Figure from Ref. [Leo06c].*

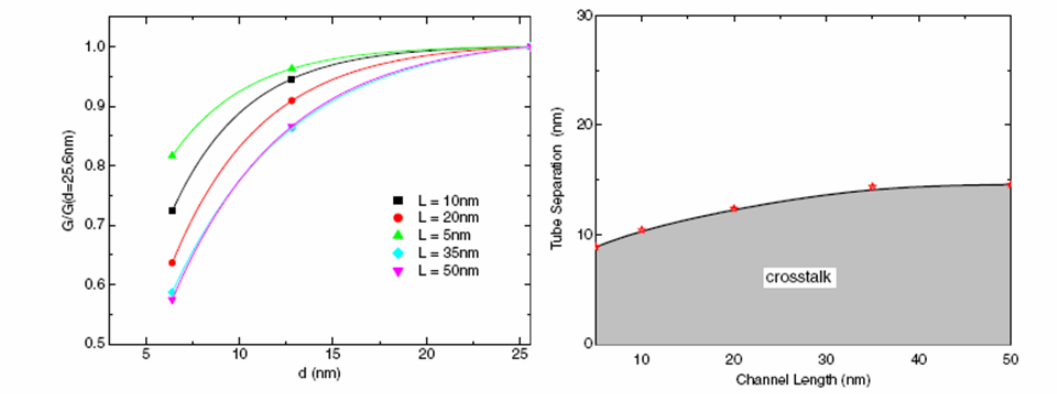

Figure 4.42: The left figure shows the conductance at V_{gs} = 0 versus nanotube separation, for different channel lengths. Solid lines are fits of the form A[1 − exp(−d/β)]. The figure on the right shows the nanotube separation below which crosstalk becomes important. *Figures from Ref. [Leo06c].*

To understand the role of interactions between semiconducting carbon nanotubes, we consider the simplified geometry of Figure 4.43. There, two uniformly charged semiconducting nanotubes of length L are separated from each other by a distance d, and both are at a distance L_G from an infinite metallic plane held at potential V_G.

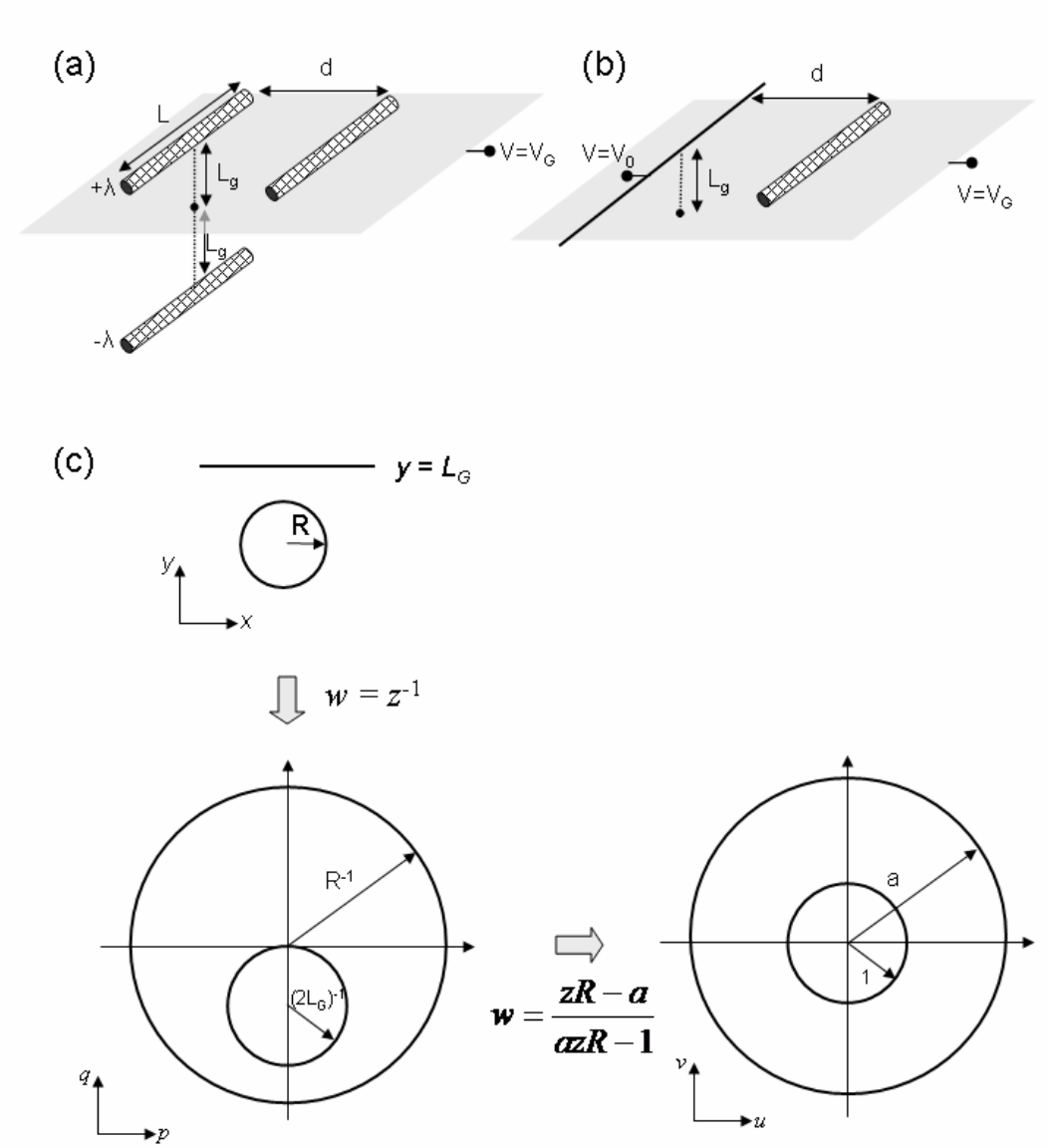

Figure 4.43: Sketch of situations considered for analytical calculation of interactions between (a) semiconducting nanotubes, (b) semiconducting and metallic nanotubes. *Figures from Ref. [Leo06c].* Panel (c) illustrates the conformal mapping procedure to solve the electrostatic problem associated with (b), see text for details.

The electrostatic potential on one of the nanotubes due to the other can be calculated using an image potential construction (Figure 4.43a) to be

$$V = V_g + \frac{\lambda}{4\pi\varepsilon} \ln \left(\frac{\left(L + \sqrt{L^2 + 4d^2}\right)\left(-L + \sqrt{L^2 + 4d^2 + 16L_g^2}\right)}{\left(-L + \sqrt{L^2 + 4d^2}\right)\left(L + \sqrt{L^2 + 4d^2 + 16L_g^2}\right)} \right) \tag{4.79}$$

where λ is the charge per unit length on the nanotubes. In the long channel limit $L \gg d$ the last equation becomes

$$V = V_g + \frac{\lambda}{4\pi\varepsilon} \ln \frac{\sqrt{d^2 + 4L_g^2}}{d}. \tag{4.80}$$

Thus in this limit the potential is independent of the channel length, and the length scale that competes with the nanotube separation is the gate insulator thickness, due to screening of the fields by the gate. Taking a potential of ΔV as a criterion for the importance of intertube effects gives the nanotube separation below which crosstalk becomes important:

$$d^* = \frac{2L_g}{\sqrt{\exp\left(\frac{8\pi\kappa\varepsilon_0\Delta V}{\lambda}\right) - 1}}. \tag{4.81}$$

Thus the gate insulator thickness L_G sets a length scale for d^*, but is exponentially reduced due to the screening by the dielectric insulator. Since d^* is only 15 nm for uncovered nanotubes on SiO_2, very high device densities should be possible by embedding the nanotubes in the insulator and using high-κ dielectrics.

The general behaviour of interactions between semiconducting nanotubes can be understood from further analysis of Equation (4.80). Near threshold, the charge on the nanotube can be written as $\lambda = \lambda_0 \left(V_g - V_{th}\right)$, where V_{th} is the threshold voltage. Substitution of this expression in Equation (4.80) gives

$$V = \left(1 + \frac{\lambda_0}{2\pi\varepsilon} \ln \frac{\sqrt{d^2 + 4L_g^2}}{d}\right) V_g - \frac{\lambda_0 V_{th}}{2\pi\varepsilon} \ln \frac{\sqrt{d^2 + 4L_g^2}}{d} \tag{4.82}$$

Hence, the gate voltage is rescaled and shifted by the interaction between nanotubes. This behavior is indicated in the inset of Figure 4.41, showing good collapse of the data after appropriate rescaling and shifting of V_g.

As mentioned above, interactions between semiconducting and metallic nanotubes are also important for multi-nanotube devices, and may also be important for all-nanotube devices where metallic nanotubes are used as interconnects. Figure 4.41 plots the conductance of a semiconducting carbon nanotube in close proximity to metallic nanotubes. The same qualitative behavior observed for semiconducting-semiconducting interactions is seen here as well. This behavior can be understood by considering the situation of Figure 4.43b: a semiconducting nanotube carrying charge density λ is at a distance *d* from a metallic nanotube. Both nanotubes are at distance L_G from an infinite metallic plane held at potential V_G. The metallic nanotube plays a different role than the semiconducting nanotube in the case of semiconducting-semiconducting interactions in

that the metallic nanotube has plenty of charge to screen the electric fields; thus, the metallic nanotube can be approximated has a line of constant electrostatic potential, equal to V_0. To calculate the electrostatic potential on the semiconducting nanotube due to the metallic nanotube we solve Laplace's equation

$$\nabla^2 V = 0 \tag{4.83}$$

with the boundary conditions

$$V = \begin{cases} V_0 \text{ for } x^2 + y^2 = R^2 \\ V_G \text{ for y} = L_g \end{cases}. \tag{4.84}$$

The solution to this problem can be obtained using conformal mapping. The general idea behind conformal mapping is to take advantage of the properties of complex functions to transform a complicated differential equation problem to a simpler problem in a new coordinate system. In the particular case of Equations (4.83) and (4.89), we will show that this problem can be transformed to one consisting of coaxial cylinders held at constant potential, for which the Laplace equation has a simple solution. Once that solution is obtained, it can be converted back to the original coordinate system using coordinate transformations.

We consider the complex function $w_1 = p(x,y) + iq(x,y)$ which satisfies the Laplace equation by virtue of the Cauchy-Rieman relations between the real and imaginary part. We first perform a conformal mapping

$$w_1 = f_1(z) = \frac{1}{z} \tag{4.85}$$

where $z = x + iy.$ This transforms the straight line $y = L_g$ into a circle of diameter L_G^{-1} touching the origin in the (p,q) plane, and transforms the circle $x^2 + y^2 = R^2$ into the circle $p^2 + q^2 = R^{-2}$ as illustrated in Figure 4.43. From Equation (4.85) we can obtain

$$\begin{aligned} p &= -\frac{y}{x^2 + y^2} \\ q &= \frac{x}{x^2 + y^2} \end{aligned}. \tag{4.86}$$

Furthermore, a second mapping $w_2 = u(p,q) + iv(p,q)$

$$w_2 = f_2(z) = \frac{zR - a}{aRz - 1} \tag{4.87}$$

with $z = p + iq$ and

$$a = \frac{L_G}{R}\left(1+\sqrt{1-\left(R/L_G\right)^2}\right) \approx \frac{L_G}{R} \tag{4.88}$$

shifts the center of the circle obtained from the straight line $y = L_g$ to the origin, and rescales its radius to 1. This mapping leaves the circle centered at the origin undisplaced, but rescales its radius to the value a. The relationships between (u,v) and (p,q) are

$$\begin{aligned} u &= \frac{aR^2\left(p^2+q^2\right)-\left(a^2+1\right)Rp+a}{a^2R^2\left(p^2+q^2\right)-2aRp+1} \\ v &= \frac{\left(a^2-1\right)Rq}{a^2R^2\left(p^2+q^2\right)-2aRp+1} \end{aligned} . \tag{4.89}$$

Note that the sequence of mappings has created a boundary value problem of two coaxial circles, and Laplace's equation can be easily solved for this problem. The solution is

$$V(u,v) = V_0 + \frac{V_G - V_0}{\ln\left(a^2\right)}\ln\left(u^2+v^2\right). \tag{4.90}$$

The potential in the original (x,y) space is obtained with the use of Equations (4.86) and Equation (4.89). Since we are interested in the potential caused on a nanotube parallel to the nanotube held at potential V_0 and at the same distance from the gate, we focus on the potential $V(x,0)$. Then we have $(p,q) = \left(0,x^{-1}\right)$ and

$$\begin{aligned} u &= \frac{aR^2+ax^2}{a^2R^2+x^2} \approx \frac{L_G}{R}\frac{x^2}{L_G^2+x^2} \\ v &= \frac{\left(a^2-1\right)Rx}{a^2R^2+x^2} \approx \frac{L_G^2}{R}\frac{x}{L_G^2+x^2} \end{aligned} . \tag{4.91}$$

This gives the solution at a distance *d*

$$V(d,0) = \frac{V_G}{2\ln\left(L_G/R\right)}\ln\left(\frac{d^2L_G^2}{R^2\left(L_G^2+d^2\right)}\right) + V_0\left[1-\frac{1}{2\ln\left(L_G/R\right)}\ln\left(\frac{d^2L_G^2}{R^2\left(L_G^2+d^2\right)}\right)\right]. \tag{4.92}$$

Much like the situation for semiconducting-semiconducting interactions, the interaction of semiconducting nanotubes with metallic nanotubes leads to a rescaling and shift of the gate voltage as shown in the inset of Figure 4.41.

4.3.7 Noise

Noise in electronic devices can have many sources, including $1/f$ noise, thermal noise, shot noise, random-telegraph-signal (RTS) noise, etc. Of these, $1/f$ noise has raised the most scientific interest because it is ubiquitous in many systems but yet its origin is unclear. The name "$1/f$ noise" originates from the dependence of the current power spectrum $P(f)$ on the frequency f

$$P(f) = \int \left(\langle I(t)I(0) \rangle - \langle I \rangle^2 \right) e^{ift} dt = A \frac{I^2}{f}. \tag{4.93}$$

In traditional semiconductor electronic devices, two origins for this behavior have been proposed: (1) fluctuations in carrier density due to fluctuations in charge trapping at surface states [McW57]; and (2) mobility fluctuations [Hoo69]. In the case of mobility fluctuations, the power spectrum takes the form

$$P(f) = \frac{\alpha_H I^2}{N_c f} \tag{4.94}$$

where α_H is known as Hooge's constant and N_c is the total number of carriers in the system. This expression is also utilized as an empirical equation describing $1/f$ noise in general. Central to this equation is the assumption that fluctuations in the number of carriers are responsible for the current fluctuations; as we now discuss, this hypothesis is valid for long carbon nanotubes where the transport is diffusive but breaks down for ballistic carbon nanotube transistors.

Figure 4.44 shows the inverse of the noise power in carbon nanotube transistors with long channel lengths such that the transport is diffusive. It is clear from the power spectrum that the main component of the noise is a $1/f$ behavior, with some additional noise component observed in some devices (Figure 4.44b). This additional noise component is well described by adding a small contribution from RTS noise to the total noise:

$$P(f) = A \frac{I^2}{f} + B \frac{I^2}{1 + \left(f / f_0 \right)^2}. \tag{4.95}$$

More importantly however is the dependence of the noise power on the number of carriers in the nanotube. To extract this dependence, note that the number of carriers in a field-effect transistor is equal to

$$N_c = C_g L \left| V_g - V_{th} \right| / e \tag{4.96}$$

and the noise power is given by

$$P(f) = A \frac{I^2}{f} = \frac{\alpha_H e I^2}{C_g L \left| V_g - V_{th} \right| f}. \tag{4.97}$$

Figure 4.45 shows the measured value of $1/A$ as a function of $|V_g - V_{th}|$ indicating a linear relationship; this behavior is consistent with an inverse dependence on the number of carriers. Equation (4.97) is further supported by measurements on devices with different channel lengths, which show a linear dependence of the noise power spectrum on the channel length (Figure 4.45).

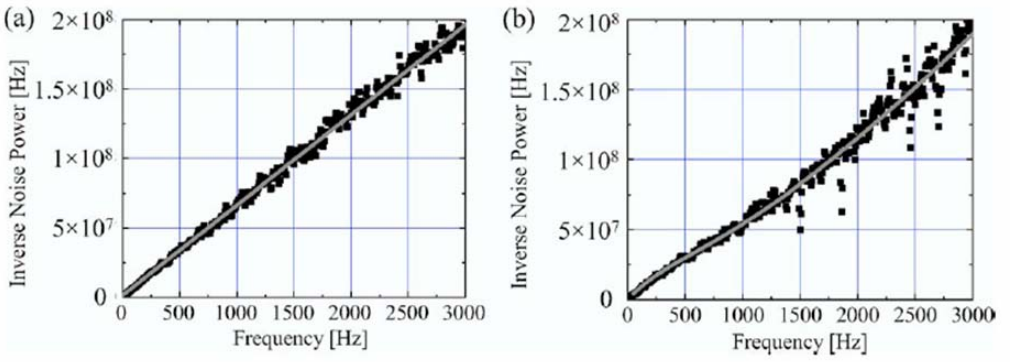

Figure 4.44: Noise behavior in single nanotube transistors. Panel (a) shows a $1/f$ noise behavior. Panel (b) shows that in addition to the $1/f$ behavior, an additional noise component is seen in some devices. *Figure from Ref. [Ish06].*

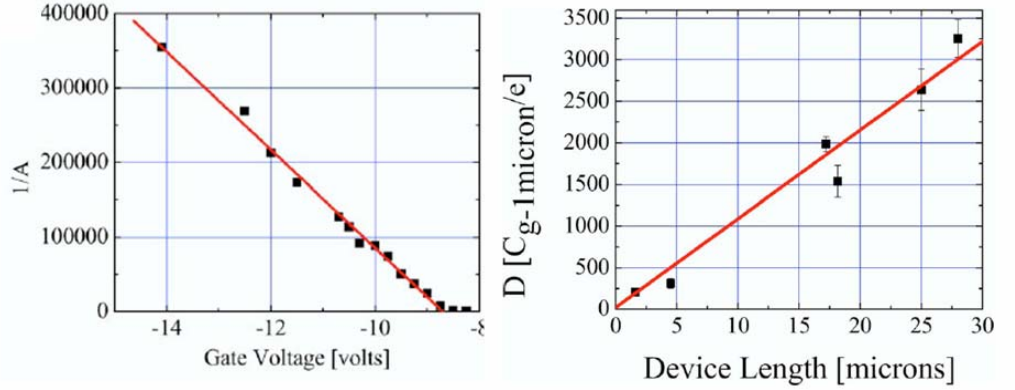

Figure 4.45: The left panel shows the dependence of the coefficient of the noise spectrum on the gate voltage. The right panel is the dependence on the length of the carbon nanotube channel. *Figure from Ref. [Ish06].*

The above discussion focused on long carbon nanotubes where the transport is diffusive. For such diffusive transport, a body of work has established the fundamental aspects of the $1/f$ noise in terms of mobility fluctuations. For carbon nanotubes shorter than the scattering mean-free path, the transport is ballistic, and not much is known about noise in ballistic systems. To address this gap, recent experiments and theory have begun to look

at this situation for both Schottky barrier and channel-controlled carbon nanotube transistors [Ter07].

Figure 4.46 shows the measured current-gate voltage characteristics of a Schottky barrier carbon nanotube transistor with a 600 nm channel as well as the amplitude of the power spectrum A from Equation (4.93). Because the channel length is less than the electron-phonon scattering mean-free path, this device is believed to operate in the ballistic regime. Clearly, the amplitude of the power spectrum is much larger near the threshold voltage. To explain the origin of this behavior we consider the impact of fluctuating charge traps in the gate oxide on the current. It is well known that such charge fluctuations lead to $1/f$ noise because the charge state is thermally activated [Dut81].

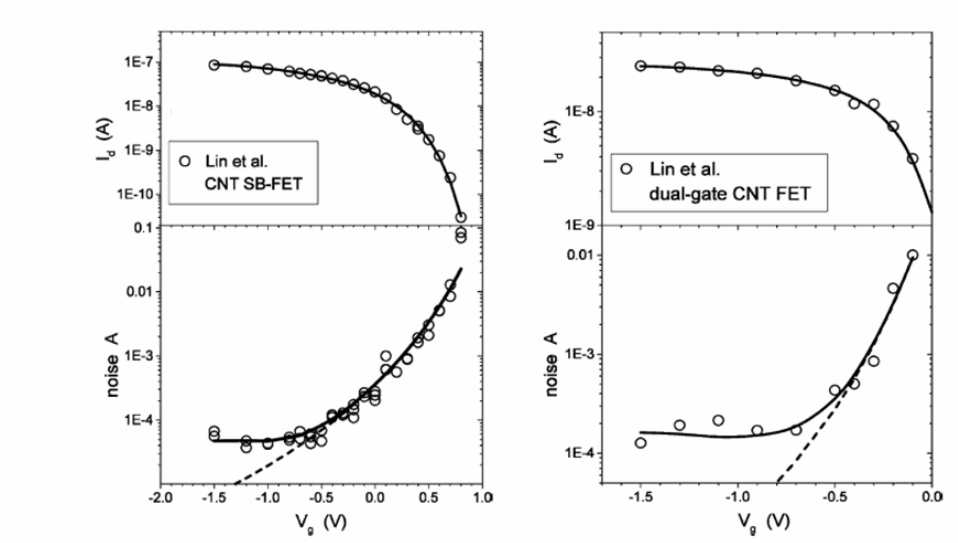

Figure 4.46: Measured and calculated noise in a Schottky barrier (left) and a channel-controlled (right) carbon nanotube transistor. Solid and dashed lines are theoretical fits, see text for details. *Figure from Ref. [Ter07].*

For a Schottky barrier device with ballistic transport, the current is a function of the electric field at the contact which in turn depends on the gate voltage and the device geometry. The geometrical factors can be captured in a parameter S_g and the electric field at the contact is then equal to V_g / S_g. The current is then a function of V_g / S_g:

$$I = I\left(V_g / S_g\right). \tag{4.98}$$

Fluctuating charge traps lead to fluctuating electric fields in the vicinity of the carbon nanotube, including near the contact. These fluctuating electric fields can be modeled as $\gamma F(t)$ where γ is the strength of the fluctuations and $F(t)$ is a dimensionless function of time with $1/f$ power spectrum. The total electric field at the contact is $V_g / S_g + \gamma F(t)$ and the current will be

$$I = I\left(V_g / S_g + \gamma F(t)\right) \approx I_0\left(V_g\right) + \gamma S_g \frac{dI}{dV_g} F(t) \tag{4.99}$$

where the last expansion assumes a small electric field due to the noise. From this expression, we can calculate the current power spectrum as

$$\begin{aligned} P(f) &= \int \left(\left\langle I(t)I(0) \right\rangle - \left\langle I \right\rangle^2 \right) e^{ift} dt \\ &= \gamma^2 S_g^2 \left(\frac{dI}{dV_g} \right)^2 \int \left\langle F(t)F(0) \right\rangle e^{ift} dt \\ &= \gamma^2 S_g^2 \left(\frac{dI}{dV_g} \right)^2 \frac{1}{f}. \end{aligned} \tag{4.100}$$

Therefore, the amplitude of the power spectrum is

$$A = \gamma^2 S_g^2 \left(\frac{d \ln I}{dV_g} \right)^2 . \tag{4.101}$$

This immediately shows that the amplitude of the noise is related to the term $d \ln I / dV_g$ which is largest in the subthreshold regime, explaining the general behavior of the experimental measurements of Figure 4.46. By fitting the experimentally measured current with a smooth function, the noise can be calculated from Equation (4.100); the dashed line in Figure 4.46 shows very good quantitative agreement with experiment near the threshold voltage.

Deviations from the theoretical model arise when the current is large, corresponding to a larger transmission probability through the contact Schottky barrier. In that case, the fluctuations in the channel become important and can no longer be ignored. To model this situation, we consider the transmission probability through two scattering regions in series, corresponding to the Schottky barrier (transmission T_{SB}) and the channel (transmission T_{ch}). The procedure for calculating the total transmission probability for scatterers in series is illustrated in Figure 4.47.

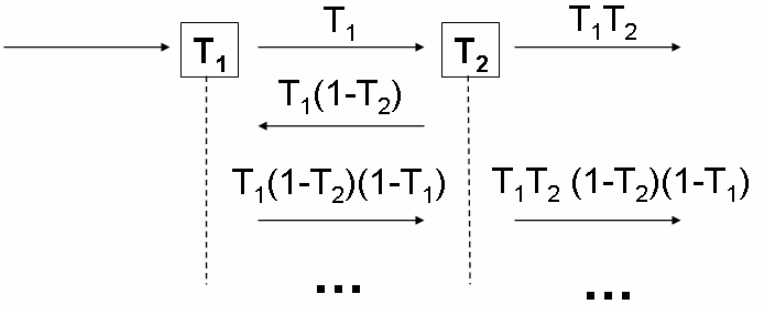

Figure 4.47: Illustration of the procedure to calculate the total transmission probability across two scatterers in series.

The total transmission is given by summing each of the scattering paths

$$\begin{aligned} T &= T_1T_2 + T_1T_2\left(1-T_1\right)\left(1-T_2\right) + T_1T_2\left(1-T_1\right)^2\left(1-T_2\right)^2 + \ldots \\ &= \frac{T_1T_2}{1-\left(1-T_1\right)\left(1-T_2\right)}. \end{aligned} \tag{4.102}$$

In terms of the transmission probabilities for the Schottky barrier and the channel, we have

$$T = \frac{T_{SB}T_{ch}}{T_{SB} + T_{ch} - T_{SB}T_{ch}}. \tag{4.103}$$

Assuming that both the transmission probabilities are small this equation simplifies to

$$T = \frac{T_{SB}T_{ch}}{T_{SB} + T_{ch}} \tag{4.104}$$

and from the Landauer formula the total transmission translates into the conductance

$$G = G_0\frac{T_{SB}T_{ch}}{T_{SB} + T_{ch}} \tag{4.105}$$

and the resistance

$$R_{tot} = \frac{1}{G_0}\frac{T_{SB} + T_{ch}}{T_{SB}T_{ch}} = \frac{1}{G_0T_{SB}} + \frac{1}{G_0T_{ch}} = R_{SB} + R_{ch}. \tag{4.106}$$

Thus the transmissivities in series lead to a total resistance that consists of two series resistances due to the Schottky barrier and the channel. For series resistors with fluctuations δR_{ch} and δR_{SB}, the total noise amplitude is given by

$$A = \frac{\left(\delta R_{SB}\right)^2 + \left(\delta R_{ch}\right)^2}{\left(R_{SB} + R_{ch}\right)^2}. \tag{4.107}$$

Re-arranging this equation leads to a noise amplitude that is the sum of the Schottky barrier and channel fluctuations

$$A = A_{SB} + A_{ch}\left(\frac{R_{ch}}{R_{tot}}\right)^2 = \gamma^2 S_g^2\left(\frac{d\ln I}{dV_g}\right)^2 + \alpha_{ch}I^2 \tag{4.108}$$

where $\alpha_{ch} = A_{ch}\left(R_{ch}/V_{sd}\right)^2$ and $A_{ch} = \left(\delta R_{ch}\right)^2/R_c^2$ is the noise amplitude of the channel resistance. The important point is that the additional series resistance due to the channel increases the noise, with an amplitude that is proportional to I^2. In regions where the current changes little with the gate voltage (as in the ON state) this contribution to the

noise dominates. Indeed, a numerical fit of Equation (4.108) to the experimental data shows an improved agreement in the large current regime (solid line in Figure 4.46).

For carbon nanotube transistors with ohmic contacts, the current is controlled by the height of the barrier in the middle of the channel. Still in that case the current can be written as $I = I\left(V_g / S_g\right)$ where the parameter S_g now describes the effectiveness of the gate at changing the barrier height in the middle of the channel. Charge trap fluctuations cause fluctuations in the height of the barrier which can again be written in the form $\gamma F(t)$. Therefore the analysis presented above for the Schottky barrier transistor carries directly over to the case of channel-controlled devices. Such channel-controlled devices can be achieved by fabricating ohmic contacts or by using a double-gate approach: the back gate is used to modulate the contacts in a Schottky barrier transistor to increase their transparency while a top gate is used to modulate the channel conductance. Figure 4.46 shows the measured current as a function of the top gate voltage for such a device. The noise is well described by the ballistic noise theory including both the near and far threshold behavior.