Bypassing Material Limits in Graphene Transistor through Gate Engineering: Momentum filtering for low power switching

A Dissertation

Presented to

the Faculty of the School of Engineering and Applied Science

University of Virginia

In Partial Fulfillment

of the requirements for the Degree

Doctor of Philosophy (Electrical and Computer Engineering)

by

Frank Tseng

December 2013

Approval Sheet

This Dissertation is submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Electrical and Computer Engineering)

_____

Frank Tseng

This Dissertation has been read and approved by the Examining Committee:

_____

Avik Ghosh, Adviser

_____

Petra Reinke, Committee Member

_____

Robert Weikle, Committee Member

_____

Keith A. Williams, Committee Member

_____

Llyod Harriott, Committee Chair

Accepted for the School of Engineering and Applied Science:

_____

Dean, School of Engineering and Applied Science

December 2013

Abstract

Graphene is the first of now several two-dimensional materials that has garnered significant interest for its potential application as a transistor for digital and radio-frequency applications. Its natural chemical 'flat-land' has several advantages unique to its hexagonal network of carbon atoms. First, graphene has a measured mobility of 230,000 $cm^2/V-s$ compared to silicon mobility of 1400 $cm^2/V-s$, which means electrons in graphene can respond to faster changes input voltage or higher clock cycles. Second, graphene's atomically thin body allows for ease of channel conductance modulation. The third advantage its intuitive compatibility to advanced planar fabrication processes already developed for silicon complementary-metal-oxide-semiconductor (Si-CMOS) transistors. However switching a graphene field effect transistor(FET) off remains a challenge. In the literature, various methods of reducing OFF-currents and achieving output current saturation have resulted in the reduction in mobility. While it may seem graphene's future in the context of the CMOS switching paradigm is unsalvageable, how we can reduce OFF-current and extend current saturation without hurting the ON-current or mobility through momentum filtering aided by gate geometry engineering.

This work starts by investigating the limitations on electron transport in a conventional graphitic-FET with no band gap at low and high biases through a unified physics based model for graphene IV from ballistic to diffusive limits and from low to high bias. At low-bias, we show how band structure is tied to the fundamental material trade-off between opening bandgaps and mobility. We find that band gap opening increases effective mass and reduces scattering time due to increase in band-edge density of states, thus reducing mobility by a factor of $1/E^2$. This happens for all graphitic derivatives. Also at low bias, we show how the minimum conductivity behaves in the ballistic and diffusive limit in the presence of impurities. We extracted the entire phase space and showed a flip in curvature followed by a saturation with increased impurity density. At high bias, our model benchmarked with experiments and converted device model to Verliog for use in Cadence for circuit level simulations. We also show how optical phonons influence the high bias current voltage behavior leading to current saturation. Finally we address in particular the trade-off between mobility and opening a bandgap with a proof-of-concept way to bypass these material limits for

narrow band gap channel through contact engineering which is unique to the device community.

The real merit of our model is the simplicity, and the use of contact engineering alone for momentum rather than energy filtering to reduce OFF-current and extend current saturation without hurting the ON-current (Fig.5.8) that made graphene so promising in the first place. Gates are uniquely positioned and biased to cascade their local narrow bandgaps along a staircase potential profile (Fig.5.1), suppressing the transmission of intermediate conducting modes between the highest conduction band and the lowest valence band (Fig.5.4). The effective mode- filtering widen a gap in the transmission spectrum. The conventional approach of widening a real bandgap in the channel to reduce OFF-current comes at the expense of ON-current due to decrease in mobility. We show a way to bypass this fundamental material limitation. We established a proof-of-concept with a 5nm wide graphene nanoribbon and 2-D bilayer graphene, with bandgaps less than 200meV were convenient channels for simulation, but the concept of gate engineering can be generalized to other narrow bandgap materials with higher mobility.

Contents

List of Figures

Chapter 1

Introduction

Graphene is the first demonstrated incarnation of two-dimensional(2D) materials. The Graphene 'flat-land' is atomically thin making it compatible to planar fabrication processes currently developed for Silicon Complementary-Metal-Oxide-Semiconductor (Si-CMOS) devices. With electron mobility above $230{,}000cm^2/Vs$[14] for suspended samples and $10{,}000cm^2/Vs$[15] to $40{,}000cm^2/Vs$[16] for supported samples at room temperature, graphene became attractive for high-frequency and even logic applications. The large mobility at least an order of magnitude higher than Silicon allows for faster clock cycles and more operations, because the electrons are capable of responding to the fast changes in gate potential. The intrinsic two-dimensional nature allows for better modulation of channel conductance in the presence of an applied electric field. Furthermore, graphene operates in the quantum capacitance limit as it is two-dimensional by nature with massless Fermions near its Dirac point. The advantage of a device operating in the quantum capacitance limit is the reduction in load capacitance in a circuit and corresponding RC delay. However the story of graphene in context of the CMOS switching paradigm is anything but perfect.

From a device perspective, in light of the excellent electronic properties for the ON state of a field effect transistor, the main concern with graphene and its carbon derivatives is their inability to adequately suppress OFF-current in the sub-threshold biasing regime. The advantages of the high graphene ON-current are compromised by the lack of a band-gap near the intrinsic Fermi level. In theory, the lack of an energy band-gap can be remedied through edge quantization, as predicted and experimentally demonstrated in carbon-nanotubes (CNTs) and armchair-edge graphene nanoribbons (AGNRs) [17, 3, 10]. Edge quantization of graphene brings a whole set of well known practical fabrication challenges [18] to achieve device performances comparable to Si-CMOS, but what is under-appreciated are the fundamental limitations imposed by the graphene band-structure or dispersion relation that can compromise mobility as the effective mass increases

near the bottom of the band when we introduce an energy band-gap. Does the road end for graphene as a suitable switch?

While transistor scaling has satisfied the economic demand for more *functionality* in our central processing units(CPU) for nearly 5 decades, the emphasis of the device community has been solely on scaling even when evaluating new materials. Due its high mobility, graphene should be able to achieve more functionality or operations every cycle compared to silicon. The requirement for graphene to start scaling where silicon field-effect transistors have left off is not necessary. Therefore, the end goal of this dissertation is to address graphene as a field-effect transistor in a way that mitigates the obstacles (i.e., high OFF-current) while trying to preserve its strengths (i.e., excellent ON-current). In other words, bypass the material limits of imposed by the graphene band-structure through contact engineering and mode or momentum filtering.
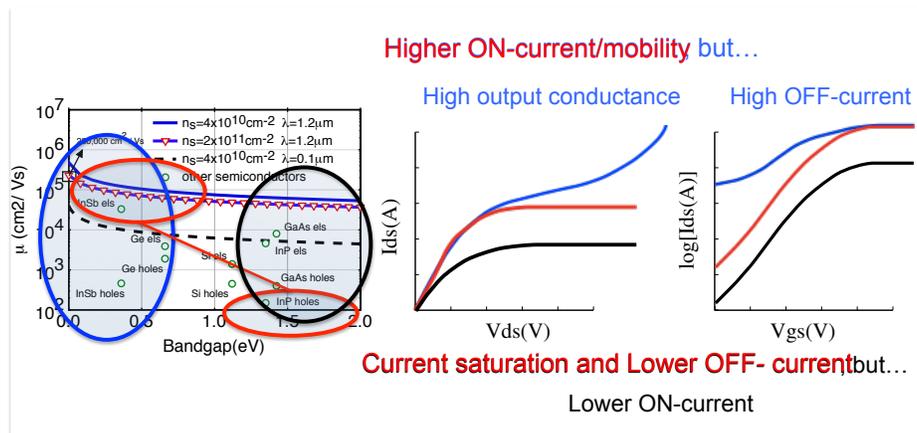
Figure 1.1: (left) Graphene and most other common group IV and III-V semiconductors follow the general trend of decreasing mobility versus bandgap. This is a fundamental trend independent of fabrication issues. In terms of devices this means materials with low bandgap and high mobility produced device output (center) and transfer (right) characteristics that follow the blue curves. Meanwhile devices with large bandgap but drastically lower mobility show device behavior resembling the red-curves. This dissertation introduces the idea of producing the black curve, where the mobility is left high but the OFF-current is lower and simultaneously current saturation as a result is more pronounced.

In addressing the material limitations of graphene as field-effect transistor, we first quantify the obstacles that stem from its chemistry and electronic band-structure. The methodology taken starts from chemistry to hierarchical device abstraction for circuit level simulations. From the bulk-bandstructure of two-dimensional graphene, we determine the scaling properties of graphene and make connections with experimental results. Particularly we look at the role of edge chemistry and roughness on the energy band-gap opening near the intrinsic Fermi level and their corresponding influence on field-effect transistor performance. Also based on what we learned from calculating the band-structure the groundworks for a hierarchical device abstraction is developed, where each element is described by *physical derivable* equations unified by the Landauer Formula.

Furthermore, by computationally tracking the evolution of the graphene band-structure from 2D to 1D, we noted a consistent departure from graphene's noted linear photon-like dispersion, the source of massless Dirac Fermions in condensed matter. It turns out that graphene's bandstructure imposes a fundamental material limited trade-off between energy band-gap and mobility, which translates to a trade-off between ON-OFF current ratio (i.e., reliability) and switching speed (delay). Projected trade-offs for different graphitic materials are presented in section(4.1). The key question and indeed the core of this dissertation is how can we bypass the fundamental limitations imposed by the graphene channel or any narrow band-gap high mobility channel.

The outline of this dissertation is as follows:

- Provide a general overview of Landauer Formula in context of building a analytical device compact model for graphene transistor, and overview the Non-Equilibrium Green's Function (NEGF) formalism used for atomistic electron transport simulations

- Quantify the electronic properties and transport properties for a conventional graphene device at low and high biases.

  – At low bias we describe the trade-off between opening a band gap and mobility (i.e. downside to energy filtering).

  – Minimum conductivity in context to impurity scattering is investigated on the phase space from ballistic to diffusive limits using a unified Landauer based model, which is also used to examine current flow at high bias.

  – NEGF with scattering within the self-consistent Born Approximation is then used to understand optical phonon drives current saturation in nanotubes.

- Getting past the trade-off in conventional graphitic-FETs, we present proof-of-concept idea based on *mode or momentum filtering* that can reduce OFF-current, extend current saturation while preserving mobility and ON-current.

Chapter 2

Synopsis of Graphene Advantages and Challenges

2.1 The Wonder Material

Regardless of imminent usefulness in conventional electronics, graphene is without doubt one of the most exciting materials in that recently emerged in condensed matter physics. The symmetry of its hexagonal lattice allows electron waves to interfere in a way that gives rise to interesting electronic properties never before seen. The orthogonal nature of the equivalent two-atom basis allows the surface p_z orbital based bonds to cross at a single point called the Dirac-point. Near the intrinsic Fermi-level at the Brillouin Zone boundaries the band crossing produces a photon-like dispersion giving rise to massless Fermions. In fact it is not straightforward to define the effective mass(m^*), which we do this in chapter 4. From the dispersion or band-structure, it has been theoretically predicted and experimentally confirmed that the low-energy Fermi velocity of graphene is 300 times smaller than the speed-of-light($10^6 m/s$). Another interesting feature of this band-crossing near the intrinsic Fermi-level is that the density-of-states vanishes at the Dirac-point in the, making graphene neither a metal or semiconductor, but rather a semi-metal.

The s and p_x and p_y orbitals hybridizes into sp^2-bonds which give graphene its inherent two-dimensional hexagonal (honeycomb) configuration, leaving a lone p_z-orbital that creates the band-crossing near the intrinsic Fermi-level. The in-plane rigidity of the hybridized orbitals ensures that out-of-phase optical phonon vibrations occur at frequencies higher than most other known crystals ($1600 cm^{-1}$)). The strengths of the in-plane hybridized bonds are seen from their energies sitting far way from the intrinsic Fermi-level. In-phase acoustic phonon vibrations are largely insignificant because the orthogonal nature of the eigenstates

representing the symmetric two-atom basis ensures that direct backscattering is disallowed. An important consequence of disallowed backscattering from acoustic phonons and far away optical phonons is that the mean-free-paths of graphene are on the order of microns. Therefore graphene has a giant electron mobility measured to be greater than $> 230,000 cm^2/V-s$ for suspended graphene exfoliated samples and $10,000$ for graphene supported on an insulating substrate. The immediate appeal is room-temperature mobility near 100 times that of silicon (Si) Long mean free paths also allow for room temperature quantum hall measurements that show signatures of resistance quantum $(h/2q^2)$.

A unique electronic property of graphene is its measured minimum conductivity of $4q/h$, where q is the electron charge, h is Planck's constant, and the pre-factor of 4 for spin and degeneracy, at the Dirac point with zero density of states is surprising at first glance. In the ballistic limit where the width of the graphene channel is much greater than the length, both experiments [19] and Landauer based theory [20] predict a minimum conductivity of $4q^2/\pi h$. In the diffusive limit, existing theory [21, 22] and experiment [23] investigate the minimum conductivity in the presence of impurities, where a Boltzmann model is used for the theory. Authors Yang et al. in their work [24] have modeled the minimum conductivity in the ballistic limit with only three data, points without identifying the origin of missing π factor. The disconnect between ballistic and diffusive limits is addressed in this dissertation using a unified Landauer based model detailed in section 4.2.

Exploring ways to open a band-gap(i.e., energy filtering) in graphene's bandstructure is a sought after way to suppress leakage currents. Confinement and breaking of sub-lattice symmetry in layered or substrated graphene can open energy band-gaps on the order of tens to a couple hundred milli-electron-Volts. For confined graphene nanoribbons, widths less than 10 nanometers are required to open a band-gap greater than 400meV. With precision of material processing to date it is still challenging to manage features less than 10 nanometers [25, 17].

A key advantage of graphene over silicon nanowires and many III-V semiconductors for high-frequency applications is that its single atom thickness allows for easy modulation of channel conductance. The approximate electric screening length for a dual gated semiconductor is,

$$\lambda = \sqrt{\frac{\epsilon_s d_s d_{ox}}{\epsilon_{ox}}} \tag{2.1}$$

,where ϵ_s and ϵ_{ox} are dielectric constants for the semiconductor and oxide, respectively. d_s and d_{ox} are semiconductor and oxide thicknesses, respectively. Eq. 2.1 provides some useful approximate insights about electrostatic response of semiconductors by giving a figure of merit for the minimum channel length below which short channel effects dominate the channel conductance [26]. The electric screening length, λ, is one

of the figure of merits that has led for the push for Silicon-on-Insulator(SOI) channels but scaling silicon
and other III-V semiconductor bodies also degrades the electron mobility. Graphene is the perfect material
because its one-atom thick body(0.22nm) ideally means that for 3nm silicon-dioxide(SiO_2) or 3nm of h-BN,
the electric screening length is on the order of several angstroms or less. This means graphene and other
emerging two-dimensional materials are highly scalable. Of all the recent emergent class of two-dimensional
materials graphene still has the highest mobility with zero-bandgap, while other 2D-like chalcogenides have
wide-bandgap with extremely low mobilities.

2.2   Graphene Device Challenges
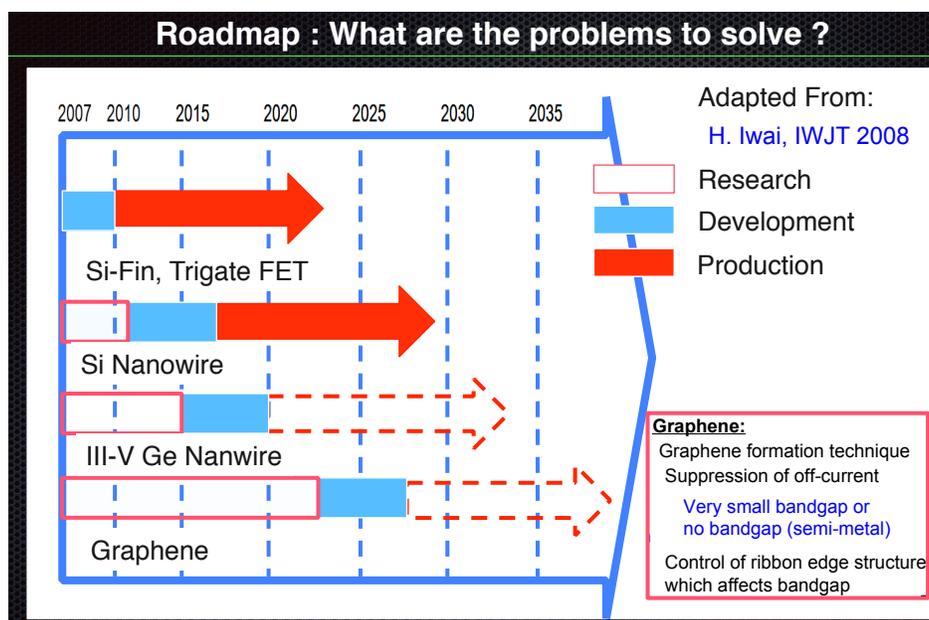
Figure 2.1: Roadmap of current and emerging technologies. Issues with graphene field-effect transistors are
highlighted red. Bandgap opening and abundance of leakage currents are key challenges that are addressed in
this dissertation. [1]

Many of today's main semiconductor device applications are in digital logic and radio-frequency applica-
tions. An economically driven semiconductor industry desires more functionality per dollar. For years the
semiconductor industry has perfected the planar fabrication processing of silicon and has managed to squeeze
more functionality through scaling. Scaling for more functionality has worked for the past 5 decades. However
the industry is down to 22nm and below, so not only are we running out of room in the bottom but the
electrostatics from three terminals are becoming increasing difficult to control as noted by Eq. 2.1. Continued
scaling also means the semiconductor industry must continue to push their fabrication equipment toward the

limit of atomic precision, as surface roughness further decreases mobility through scattering. What does this all mean for graphene?

Graphene was initially seen as prospective replacement for Silicon in the context of CMOS switching. Since there has been significant physics unraveled that causes us to rethink how computing could be done. For now these ideas are still on paper and largely far-fetched for what industry is currently ready for. As mentioned in the previous section, graphene's high mobility and inherent two-dimensional nature means that it is electrostatically more advantageous than sdilicon. On the flip slide, CMOS relies on a barrier to switch the resistance in the channel which graphene inherently lacks. However as mentioned in the section above, there are various means or opening a band-gap, however the reliability of methods is still questionable for mass-production. Assuming fabrication methods are reliably in the limit of atomic precision, there is a less appreciated fundamental trade-off associated with opening an energy band-gap in graphene that will be described in section(4.1).

For digital logic applications, traditional Silicon based N and P-type devices complement each other to increase speed and decrease static power leakage. For graphene the lack of an energy barrier means the field-effect transistor is always in ON-state. The perpetual leakage of current in a graphene field-effect transistor has a number of negative consequences discussed in more detail in Section 8.2 where a version of physics based compact-model is used to determine basic inverter circuit level performance. Due to the lack of current saturation in the output IVs and high OFF-currents from lack of a barrier a traditional graphene field-effect transistor is highly susceptible to noise. The calculated and measured voltage-transfer curves (VTC) show gain away from the $V_{DD}/2$ as the transistor never turns off. In other words, for logic it is more important to have the rail voltages saturate in a VTC. However non-saturating rail voltages also deteriorates the gain in the transition of meta-stable region in the VTC [7], see Fig. 2.2.

For radio-frequency applications, we are more interested in the gain and less worried about high OFF-current. The state-of-the art graphene field effect transistor can achieve cut-off frequencies up to 100GHz for conventional gates[27] and 300GHz with nanowire gates [28]. This is done primarily by taking advantage of the lack of short channel effects in graphene as discussed in the previous section. Less appreciated is the need to saturate the output IV, as this will increase gain (g_m/g_{ds}), where g_m is the transconductance (dI/dV_{gs}) and g_{ds} is the output conductance. So if output current saturates, g_{ds} tends to zero and gain improves. Increasing transistor cut-off frequency is frequently attributed to a decrease in parasitic resistance and capacitance, but it is clear that decreasing output conductance,g_{ds} is equally important [29].

To address the issue of leakage current in graphene, a bandgap of at least 400meV is necessary for adequate reliability in digital logic applications and gain for radio-frequency applications. Energy bandgap opening through confinement requires features less than 10 nanometer, which is possible using chemical means, but
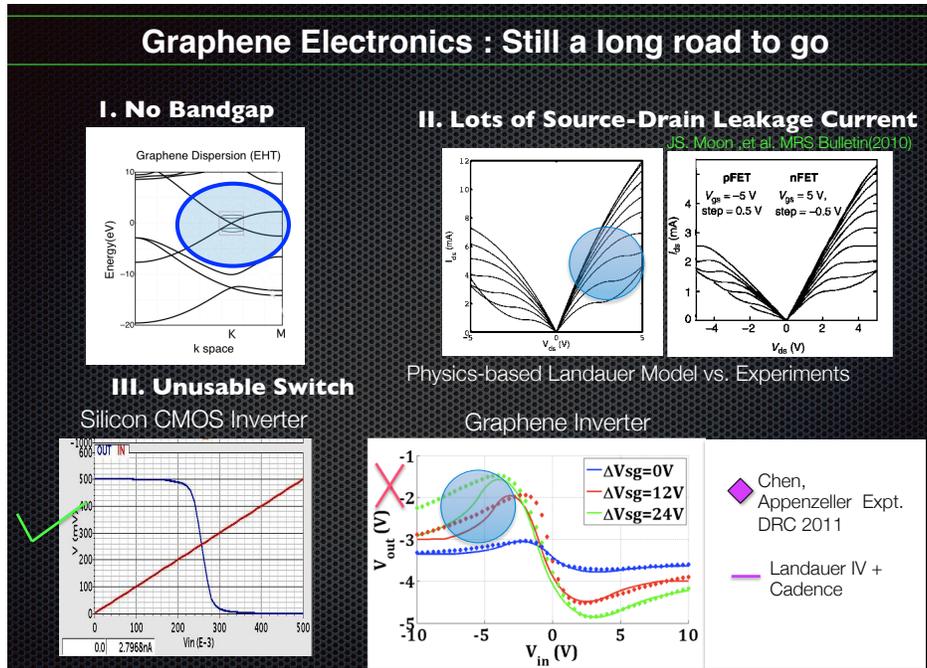
Figure 2.2: Graphene's zero band gap contributes to an abundance of source-drain leakage current. Since a graphene transistors are always ON a voltage division at the output voltage introduces gain at the input voltage extremes.

challenging for using lithographic processes designed for mass production of Silicon based electronic device technologies not to mention the difficulty of roughness control at those dimensions. Epitaxial Graphene or Graphene grown on Silicon Carbide (SiC) have demonstrated bandgaps up to 260meV[30], depending how the gaphene layers are stacked during growth. Similar interlayer symmetry breaking induced bandgaps have been demonstrated with Bernal Stacked Bilayer Graphene(BLG) with bandgaps up to 300meV, while bandgaps less than 100meV are predicted with graphene on hexagonal-Boron Nitride(h-BN). Overall confined graphene can produce the largest bandgaps; however, large scale transferring and processing of CNTs and chemically derived GNRs is not yet possible. *Until then leakage current in Graphene is still a fundamental issue, as seen in Fig. 2.1, and one which this dissertation details a possible way forward using gate engineering for momentum filtering, rather than energy filtering.*

Chapter 3

Device Modeling: Chemistry to Devices Overview

The purpose of this chapter is to present an overview of two methods used in the dissertation to describe transport and current flow in graphene devices. The first method is Landauer formalism where transport is described by lumped analytical expressions like mode density and transmission probability that can provide a more intuitive and computationally less expensive understanding of transport in graphene. The second method is the Non-Equilibrium Green's Function (NEGF) formalism which solves Schrodinger's Equation atomistically with open boundary conditions at the contacts. Since NEGF is an atomistic quantum mechanical treatment of the channel and contacts, it can be computationally expensive depending on the number of atoms in the calculation. Although the purpose of this section is to provide a brief overview and review of the NEGF and Landauer formalisms, this section presents a unique way to quantify electron transport in graphene in the presence of scattering centers.

3.1 Bottom-Up(Atomistic) Approach to Modeling Electronic Device Properties

3.1.1 Building analytical model for electron transport in Graphene

Capturing Bandstructure

Graphene's hexagonal atomic configuration has a repeatable two-atom basis and its Hamiltonian can be described by the following 2x2 simplified Hamiltonian that is valid within $\pm 1eV$ from the intrinsic , where

$v \approx 10^6 m/s$ is the constant Fermi velocity and ϵ_A and ϵ_B represent the on-site energy of the AB-sublattice atoms.

$$H_k = \begin{pmatrix} \epsilon_A & \hbar v(k_x + ik_y) \\ \hbar v(k_x - ik_y) & \epsilon_B \end{pmatrix} \tag{3.1}$$

The eigenvalues of the 2x2 Hamiltonian are $E = \pm \hbar v \sqrt{k_x^2 + k_y^2}$. An asymmetry in the AB-sublattice will open a bandgap proportional to the size of the asymmetry, $\epsilon_A - \epsilon_B$. The dispersion then becomes, $\pm \sqrt{(\epsilon_A - \epsilon_B)^2 + (\hbar v)^2 (k_x^2 + k_y^2)}$. From the bulk analytical dispersion, we can gain some meaningful estimates regarding the required scaling of graphene widths in order to yield a meaningful bandgap for logic(i.e, $E_g \leq 400 meV$). From $E = \pm \hbar v \sqrt{k_x^2 + k_y^2}$, with hard wall boundary conditions, we get:

$$E_g = \hbar v \frac{\pi}{W} \tag{3.2}$$

, where W is the confined width of a graphene nanoribbon. Eq.(3.2) describes the number of half-wavelengths that satisfy the transition across a given energy bandgap (E_g). Therefore to achieve a bandgap of $400 meV$, the width of a graphene nanoribbon has to be less than 10nm in agreement with the numerically calculated trend presented in Fig.4.7 and experiments [10].

From the analytical bulk graphene dispersion relation, the bulk density of states is:

$$D(E) = \frac{|E|}{\pi \hbar^2 v^2} \tag{3.3}$$

Eq. 3.3 represents the pristine graphene density of states with a Dirac-point; however, realistically in experiments, graphene rests on an insulating substrate, such as silicon dioxide, with embedded charged impurities that can locally dope graphene creating an in-homogenous potential distribution. Scanning tunneling microscopes (STM) have probed and characterized the potential in-homogeneity in graphene as a normal or gaussian distribution of potentials [2], as seen in Fig. (3.1). Averaging the ideal graphene density of states over a normal distribution of potentials (Eq.3.4) removes the Dirac-point and creates a non-zero raised minimum density of states (Fig.3.2).

$$\langle DOS \rangle = \sum_i \frac{|E + \mu + U_i|}{\pi \hbar^2 v^2} \frac{e^{-\frac{(E + \mu - U_i)^2}{2\sigma_o^2}}}{\sigma_o \sqrt{2\pi}} \tag{3.4}$$

When we convert Eq.(3.4) to an integral we get an exact average density of states, Eq. 3.5 with statistically relevant parameters that describe charge in-homogeneity.
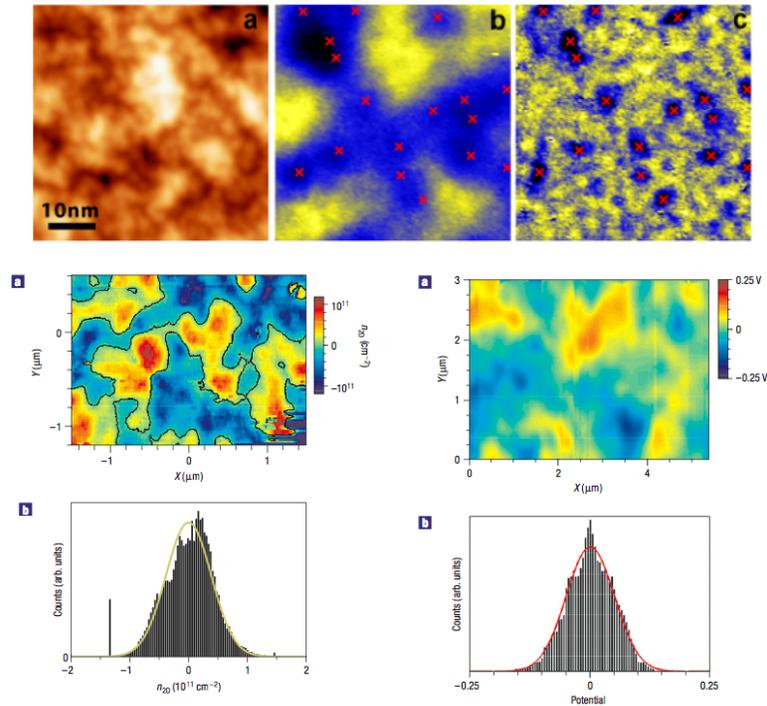
Figure 3.1: Scanning Tunneling Microscope image of electron-hole puddles that statistically follow a normal distribution.[2]
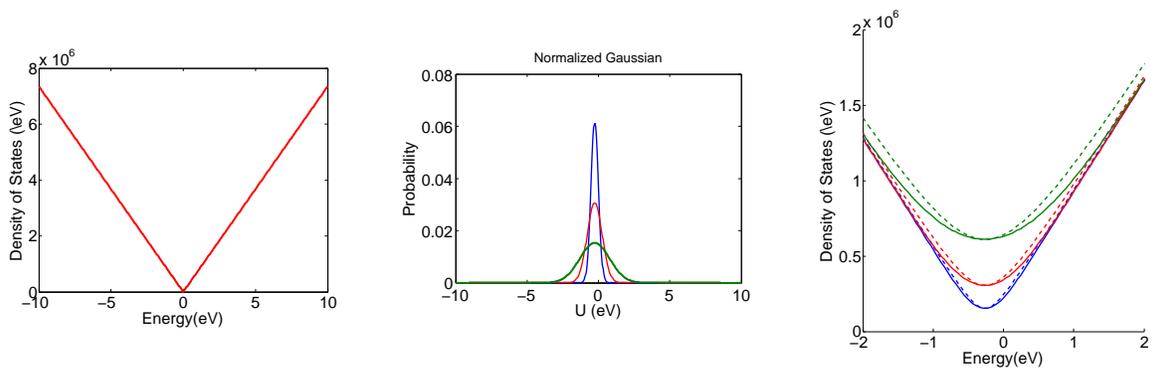
Figure 3.2: Averaging the pristine graphene density of states(left) with a normal distribution of random potentials(middle) erases the Dirac point(right).

$$\langle DOS \rangle = \frac{S}{\pi \hbar^2 v^2} \left[\sqrt{\frac{2}{\pi}} \sigma_o e^{\frac{-(E+\mu)^2}{2\sigma_o^2}} + erf\left[\frac{E+\mu}{\sqrt{2}\sigma_o}\right] |E+\mu| \right] \tag{3.5}$$

In Eq. 3.4, U_i represents a single potential in the normal distribution of potentials, σ_o represents the standard deviation and μ is the average potential of the normal distribution of charges. The average density of states can be written in a simpler form as Eq. 3.6, that incorporates the low-energy parabolic nature of the average density of states and the linear density of states at high energy. Between Eq. 3.5 and Eq. 3.6 the quantitative deviation in the non-zero minimum point in the density of states is matched with a universal factor $\alpha = 2/\pi$. Fig. 3.2 illustrates the averaging of a normal distribution of potentials and the process of matching the minimum density of states making Eq:3.6 quantitatively accurate. Both equations capture the significant role of charged impurities at low energy.

$$\langle DOS \rangle = \frac{\sqrt{(E+\mu)^2 + \alpha\sigma_o^2}}{\pi \hbar^2 v^2}. \tag{3.6}$$

The presence of charged impurities in the substrate imposes a random variation of potentials that statistically follows a normal distribution. Of particular importance is the variance (σ_o^2) of the normal distribution of potentials, which itself is a self-consistent solution dependent on the charge impurity concentration(n_{imp}) and graphene channel screening. The details of the variance(σ_o^2) are worked by Li and Sarma in [31]. Eq. 3.7, Eq. 3.8, and Eq. 3.9 is presented for completeness.

$$\sigma_o^2 = 2\pi n_{imp} q^2 \int [A_k]^2 k \, dk \tag{3.7}$$

$$A_k = \frac{2e^{-\kappa z_0} Zq \, sinh(k\,d)}{k\kappa_{ins} cosh(k\,d) + (k\kappa_v + 2\,q_{TF}\kappa)sinh(k\,d)} \tag{3.8}$$

$$q_{TF} = \frac{2\pi q^2}{\kappa} \langle DOS \rangle \tag{3.9}$$

Eq. 3.8 comes from solving Poisson's equations in cylindrical coordinates. κ_v and κ_{ins} are the respective vacuum and insulator dielectric constants, while κ is their average. Eq. 3.9 is the Thomas-Fermi screening wave-vector which depends on the average density of states (Eq. 3.6). A_k is the potential solved from Poisson's equation which accounts for the distance of the impurities (z_o) inside the oxide, thickness of the oxide (d) and the screening length $(1/q_{TF})$. Solved self-consistently we determine the variance of the normal distribution of potentials [31].

This modified density of states averaged over a random normal distribution of carriers was derived independently of Li and Sarma [31], but the conclusions are the same. However this dissertation goes further by using the averaged density of states to look at current flow in graphene all the way from low bias minimum conductivity to the high-bias current flow that shows saturation and subsequent band-to-band tunneling. Before presenting the role of impurities, we need to briefly introduce the Landauer formalism for current flow, which is the basis for all analytical models of graphene presented in this dissertation.

Intuitive Model for Electron Transport : Landauer Formula
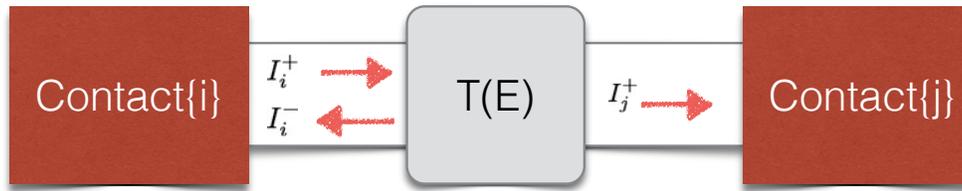
Figure 3.3: Schematic of terminal currents

The Landauer formalism is a unified way to understand charge transport and current flow from the nanoscale ballistic-limit to the large scale diffusive-limit. Landauer's formalism expresses current flow as result of a difference in agenda between source and drain contacts. Assuming reflectionless contacts and a ballistic channel transmission probability, $T(E) = 1$, the total current is the difference between the influx and out flux at a terminal, as illustrated in Fig. 3.3. To conserve current, the total current must equal the current transmitted on the opposite contact as expressed by Eq. 3.10

$$I = I_i^+ - I_i^- = I_j^+ \tag{3.10}$$

The individual currents at terminal i shown in Fig. 3.3 can be expanded to the following expression: $I_i^+ = 2q/hM[\mu_i - \mu_j]$ and $I_i^- = 2q/hM(1 - T(E))[\mu_i - \mu_j]$. μ_i and μ_j are the electrochemical potential in the contacts and $M = \hbar v/L \langle DOS \rangle$, is the number of conducting modes. Plugging terminal i currents back in to Eq. 3.10 gives the zero temperature low-bias current.

$$I = \frac{2q}{h} MT(E)[\mu_i - \mu_j] \tag{3.11}$$

Assuming non-zero temperature Eq. 3.11 becomes Eq. 3.12 ,where $f_{i,j}(E - \mu_{i,j})$ are the finite temperature Fermi functions in the contacts. The transmission probability $T(E)$ is phenomenologically expressed as

$\lambda/(\lambda + L)$, where λ is inversely proportional to the scattering rate$(1/\tau)$, determined from Fermi's Golden Rule. Appendix IV works out the scattering rate for impurities, optical and acoustic phonons.

$$I = \frac{2q}{h} \int_{-\infty}^{\infty} M\,T(E)\,[f_i(E - \mu_i) - f_j(E - \mu_j)]\,dE \tag{3.12}$$

An alternative form of Eq. 3.12 is Eq. 3.13, where we convolve the product of the number of modes and transmission probability by a thermal broadening function, $F_T(E) = df/dE = 1/4kT\,sech^2(E/2kT)$[32].

$$I = \int_{\mu_i}^{\mu_j} M(E)\,T(E) \otimes F_T(E)\,dE \tag{3.13}$$

The advantage of the Eq. 3.13 is that the limits of integration can be taken at finite, zero-temperature bounds at μ_i and μ_j versus taking infinite bounds, which serves two purposes: 1) for ease of computation and 2) easier to derive analytical expressions. The thermally broadened Landauer formula (Eq. 3.13) can be rewritten in a more intuitive form that separates the ballistic(left half of Eq. 3.14) and scattering(right half of Eq. 3.14) currents.

$$I = \int_{\mu_i}^{\mu_j} M(E) \otimes F_T(E)\,dE - \int_{\mu_i}^{\mu_j} M(E)\,\frac{L}{L + \lambda} \otimes F_T(E)\,dE \tag{3.14}$$

Given a known low energy dispersion relation or bandstructure we can construct a simplified model for current flow based on the Landauer Formula. The Landauer Formula provides a convenient and unified way to understand current flow in a nanoscale devices from the ballistic $(\lambda/L = 1)$ to diffusive limit $(\lambda/L < 1)$. In chapter 4 and in appendix II, we show both full analytical and quasi-analytical models for current voltage characteristics converted to a physics based compact model for a graphene field effect transistor that matches well with experiments and apply the models to Cadence for basic circuit level simulations.

3.1.2 Non-Equilibrium Green's Function Formalism for Atomistic Resolved Electron Transport

Chemically Sensitive Bandstructure : Tight-binding to Extended Huckel Theory

Before modeling current flow we need to construct the Hamiltonian matrix which describes the coupling between neighboring atoms. This section focuses on orthogonal tight-binding(TB) and non-orthogonal Extended Huckel Theory(EHT), the methods used to calculate the Hamiltonian matrix to determine the band structure for graphene and graphene nanoribbons.

The real-space tight-binding Hamiltonian matrix is determined using the following conditions:

$$\langle \phi_n \left| H \right| \phi_n \rangle = \epsilon$$

$$\langle \phi_n \left| H \right| \phi_m \rangle = -t \quad , m = n \pm 1$$

$$\langle \phi_n \left| H \right| \phi_m \rangle = 0 \quad , m > n \pm 1$$

ϵ is the onsite energy of atom and $t \approx 2.5eV$ is the energy coupling between each atom. For graphene, the p_z-orbital overlaps are the weakest and therefore closest in energy to the Fermi level compared to the in-plane sp^2-hybridized bonds. Therefore, in general for electron flow we can get away with a single orbital nearest neighbor TB-model. Note that the single p_z-orbital basis used for tight-binding is only applicable for an ideal in-plane hexagonal network of carbon atoms with uniform bond length of $1.42\mathring{A}$, with only the nearest neighbor interactions. When we investigate the role of edges later in this chapter we will see that this model fails to predict the lack of metallicity in armchair edge graphene nanoribbons(Fig. 3.4 far left)[3]). A third nearest neighbor tight-binding model is needed worked out by [33] to capture the accurate confinement induced band gap opening.
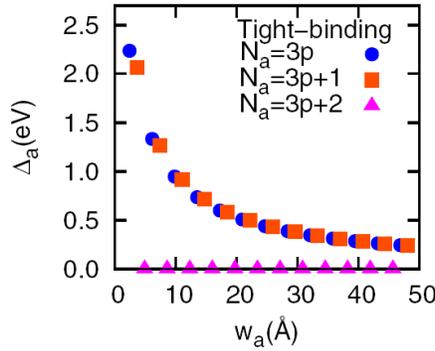
Figure 3.4: Tight-binding 1-orbital calculations show three chiral curves, one of which is metallic. Figure borrowed from [3]

To capture edge effects in graphene nano ribbons, we opted to use Extended Huckel Theory (EHT), whose more complicated, non-orthogonal and detailed valence band basis set is applicable to changes in chemical environment and bond distortions. EHT valence band orbitals are each described by Slater-type wave functions, whose general form is Eq. 3.15.

$$\Psi(r) = r^n [c_1 e^{-\zeta_1 r} + c_2 e^{-\zeta_2 r} + c_3 e^{-\zeta_3 r}] Y_{lm}(\theta, \phi) \tag{3.15}$$

n is the atomic number. $c_1, c_2, c_3, \zeta_1, \zeta_2, \zeta_3$ are the coefficients and exponents that are unique to each

orbital. In particular Carbon has 4 valence bonds: one s and three p. The coefficients and exponents for each valence orbital are determined by benchmarking the bulk band structure of graphite against density functional theory with appropriate corrections. Once the bulk-bandstructure is benchmarked, the same parameters can be used to construct a Hamiltonian for strained or bond disordered systems without the need for reparameterization. The parameters used in this work can be found in [34][35]

$$S_{\nu\mu} = \langle \phi_\nu | \phi_\mu \rangle \tag{3.16}$$

$$H_{\nu\mu} = K_{EHT} S_{\nu\mu} (H_{\nu\nu} + H_{\mu\mu})/2 \tag{3.17}$$

Operationally, Extended Huckel Theory first calculates an overlap matrix ($S_{\nu\mu}$) using Eq. 3.16, where ν and μ are atomic orbital indices. From the overlap matrix the Hamiltonian of graphene is constructed using Eq. 3.17.

$$H = \begin{bmatrix} [\alpha_{11}] & [\beta_{12}] & [\beta_{13}] & [\beta_{14}] & \cdots & [\beta_{1N}] \\ [\beta_{21}] & [\alpha_{22}] & [\beta_{23}] & [\beta_{24}] & \cdots & [\beta_{2N}] \\ [\beta_{31}] & [\beta_{32}] & [\alpha_{33}] & [\beta_{34}] & \cdots & [\beta_{3N}] \\ [\beta_{41}] & [\beta_{42}] & [\beta_{43}] & [\alpha_{44}] & \cdots & [\beta_{4N}] \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ [\beta_{N1}] & [\beta_{N2}] & [\beta_{N3}] & [\beta_{N4}] & \cdots & [\alpha_{NN}] \end{bmatrix} \tag{3.18}$$

$$H(k) = [\alpha_{nn}] + \sum_{m,m \neq n} [\beta_{nm}] e^{ikR_{mn}} \tag{3.19}$$

Extracting block elements from the general Hamiltonian in Eq. 3.18 , we can use construct a reciprocal space Hamiltonian Eq. 3.19 whose eigenvalues gives us band structure for graphene (Fig. 3.5(right)). The band crossing near E=0 arise from the p_z-orbitals while the higher energy bands are the sp^2 hybridized σ-bonds.

The bonding environment for GNRs at the edges deviates from the center as edge carbon atoms see only two carbon atoms whereas each atoms is nearest neighbor with three other carbon atoms. Along the armchair-edge, carbon atoms tend to dimerize analogous to polyacetylene. The obstruction of lateral symmetry in polyacetylene opens a known bandgap proportional to the difference between inter and intra-dimer bond coupling strengths. The bonding environment in graphene is slightly more complicated compared to polyacetylene, but the role of edge strain is nonetheless manifested in the dimerization of carbon edge
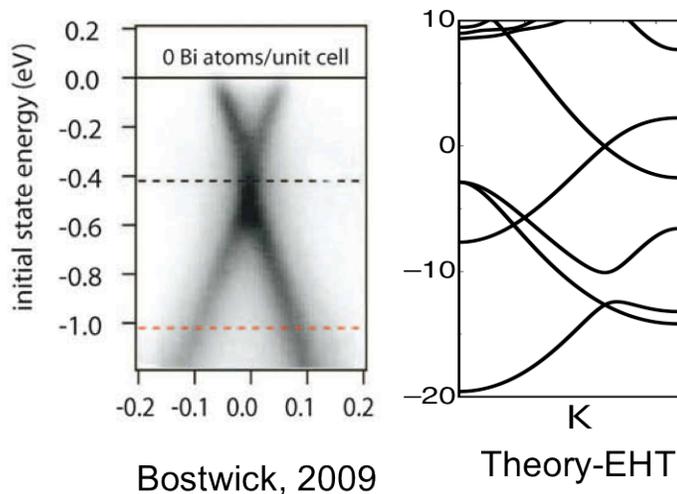
Figure 3.5: Graphene dispersion from (left) experiment from ARPS[4] and (right) from Extended Huckel Theory.
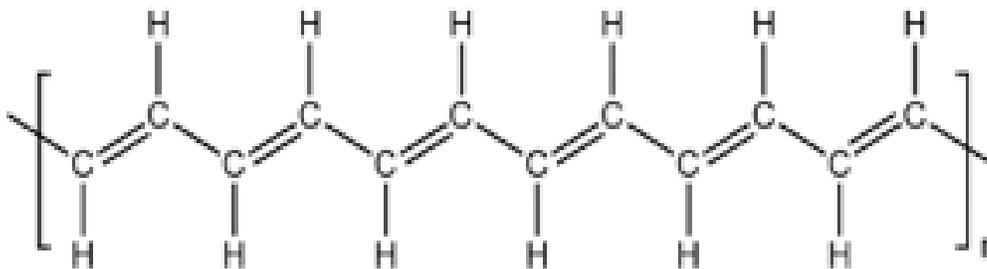
Figure 3.6: Polyacetylene molecule resembles the edge chemistry of strained armchair edge graphene nanoribbons.

atoms on bandgap opening as seen in the density of states in Fig.(3.8b).

To confirm our simple explanation on the impact of strained edge bonds on the electronic structure, we calculated the density of states (DOS) of a uniformly wide armchair GNR with edge strain using EHT. The role of edge passivation is shown in Fig. 3.8, where we can see the explicit removal of edge-induced midgap states by hydrogenation. Fig. 3.8b show the role of edge strain. In contrast to pz-orbital based nearest neighbor one orbital tight-binding theory, a small bandgap opens. While CNTs have precise periodic boundary conditions along their circumference, the edge atoms do not provide an exact hard wall boundary condition, as the electrons tend to tunnel out into the surrounding region. In the presence of edge strain, the bandgap increases because of the aforementioned dimerization, removing any semblance of metallicity from the bandgap vs. width plots Fig. 3.9. Hidden third nearest neighbor interactions are included automatically
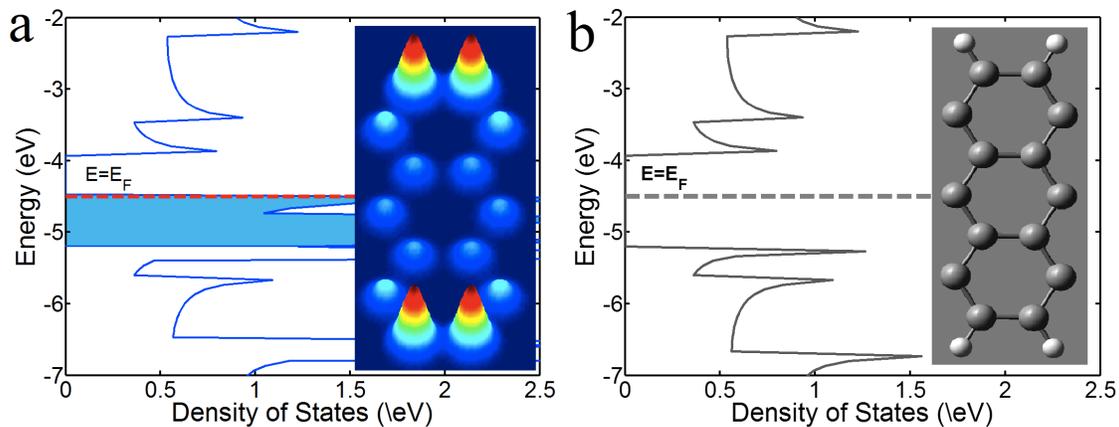
Figure 3.7: Graphene nano ribbon density of states simulated from the same Extended Huckel Theory parameters used for graphene.
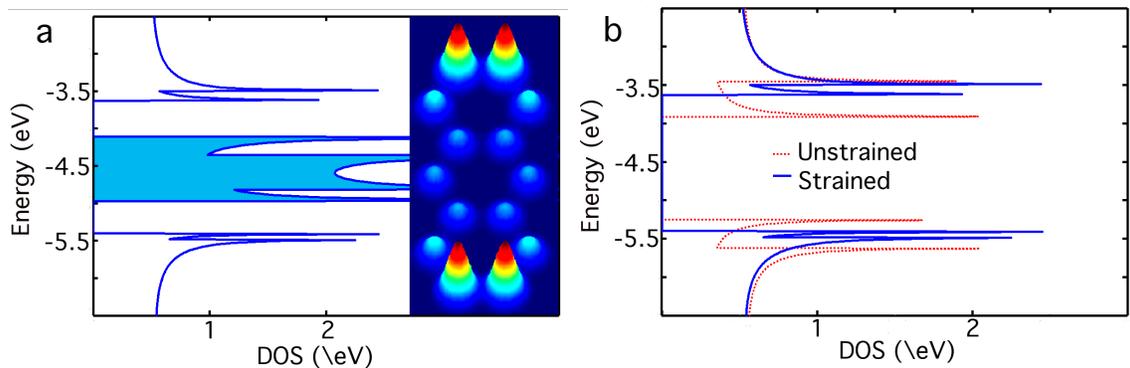
Figure 3.8: EHT captures the proper GNR chemistry, including (a) mid-gap states near the Fermi energy (-4.5 eV) arising from armchair edge dangling bonds (inset: local density of gap states). (b) H-passivation removes edge states, while soft edge boundaries prevent metallicity. A 3.5% edge strain further enhances the band-gap, E_g.

with EHT.
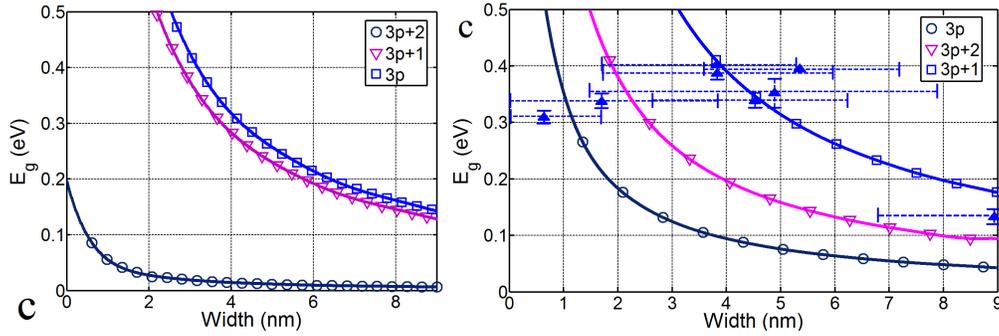
Figure 3.9: (c) Applying EHT to GNR dispersion relation across a range of sub- 10nm armchair edge widths finds an oscillating bandgap. (d) Strain of edge bonds that are hydrogen terminated widens the energy bandgap for 3p and reduces the gap for $3p+1$ GNRs. Eg vs. width results are within the range of experimental data points [5]) and also in agreement with DFT predictions

Atomistic Modeling of Current Flow

This dissertation is concerned with a current flow in the direct-current (DC) limit. Therefore, when we model nanoscale devices dominated by quantum mechanical behavior, we start with the time-independent Schrodinger's equation with open-boundary conditions for the injection and removal of carriers, Eq. 3.20.

$$E\Psi = [H + U]\Psi + \Sigma\Psi + \{s\} \tag{3.20}$$

H is the Hamiltonian matrix describing the on-site energies and hopping between neighbor atoms. Orthogonal tight-binding and non-orthogonal Extended Huckel Theory are two methods, discussed in the last section that can be used to construct, H. U is a potential matrix dependent on the electrostatics of the device, and is at the heart of any device operation. Σ is called the self-energy matrix describes the interaction between the carriers in a material system with its surrounding environment. In context of electronic devices, Σ can be broken down as the sum of contact self-energies (Σ_1, Σ_2) and scattering self-energy (Σ_s). $\Sigma\Psi$ describes the outflow of carriers, while $\{s\}$ describes some arbitrary inflow. In general, the self-energy is a complex matrix of complex numbers, where the inverse of its imaginary component(Γ) describes the lifetime of the particle.

Relating the output response of the wave function(Ψ) to an arbitrary response s is the expression, $\Psi = G\{s\}$. G is the retarded Green's function, which from Eq. 3.20 is:

$$G = [EI - H - U - \Sigma]^{-1} \tag{3.21}$$

Charge density is defined as $\langle \Psi^\dagger \Psi \rangle$, where $\langle ss^\dagger \rangle = \Sigma^{in}$ is the in-scattering . The charge density or electron correlation functions is :

$$G^n = \langle \Psi \Psi^\dagger \rangle = G \langle ss^\dagger \rangle G^\dagger \tag{3.22}$$

The expectation value of the inflow contribution, Σ^{in} can be written the inflow from all contacts including scattering, where we assume the self-energies for the contact 1 and contact 2 are equal, and $\Gamma = i(\Sigma - \Sigma^\dagger)$ is the broadening.

$$\Sigma^{in} = f_1 \Gamma_1 + f_2 \Gamma_2 + \Sigma^{in}_s \tag{3.23}$$

Similarly we can define a correlation matrix for holes, $G^p = G\Sigma_{out}G^\dagger$, which is simply difference between the spectral function, $A = G[\Gamma_1 + \Gamma_2 + \Gamma_s]G^\dagger$, and the electron correlation matrix, G^n (Eq. 3.22). Σ_{out} is the self-energy describing outflow and has the same form as Σ_{in} in Eq. 3.23, except the device contact Fermi functions, $f_{1,2}$ are replaced by $1 - f_{1,2}$.

Having defined terms that describe the inflow and outflow of carriers, we can define a terminal current analogous to Eq.3.10. From this point on in this section, we will see how get an expression for current that is similar to the Landauer formula, Eq. 3.12.

The terminal current is the net flow of charge carriers as expressed in Eq. 3.12. In terms of the terminal inflow and outflow self-energies and electron and hole correlation matrices, the terminal current is:

$$I = \frac{2q}{h} \int dE \, \Sigma^{in} G^p - \Sigma^{out} G^n \tag{3.24}$$

Using the identities, $G^p = A - G^n$ and $\Gamma = \Sigma^{out} + \Sigma^{in}$, Eq. 3.24 can be written in a form often expressed in texts, Eq.3.25.

$$I(E) = \frac{2q}{h} \int dE \, (\Sigma^{in}_i A - \Gamma_i G^n) \tag{3.25}$$

However this Eq. 3.25 does not resemble the Landauer Formula Eq. 3.12 as promised earlier, but has the advantage of explicitly including scattering. In the ballistic limit (i.e. no scattering), where $\Sigma_s = 0$, using the identities defined above($A = G[\Gamma_i + \Gamma_j]G^\dagger$ and $\Sigma^{in}_i = f_i \Gamma_i$) with the correlation matrix (Eq.3.22), the terminal current becomes:

$$I = \frac{2q}{h} \int_\infty^{-\infty} Trace\{\Gamma_i G \Gamma_j G^\dagger\} \, [f_i(E - \mu_i) - f_j(E - \mu_j)] \, dE \tag{3.26}$$

Eq 3.26 is more intuitive form of the non-Equilibrium Green's Function(NEGF) formalism as it resembles the Landauer Formula (Eq. 3.12), where $M(E)T(E) = Trace[\Gamma_i G \Gamma_j G^\dagger]$. Note that $M(E)T(E) = Trace[\Gamma_i G \Gamma_j G^\dagger]$ means that the total transmission is the sum of propagating and evanescent modes/

In problems where inelastic scattering is important, Σ_s is no longer zero and needs to calculated self-consistently with with the correlation functions (G^n and G^p). The condition for convergence is the conservation of currents with respect to all terminals including the scattering terminal defined by scattering self-energy (Σ_s). When incorporating inelastic scattering, we need to define a broadening function for scattering (Γ_s) which is proportional to the inverse lifetime of for particular scattering process. Γ_s is the sum of the in-scattering and out-scattering self-energy matrices defined in Eq. 3.27 and Eq. 3.28, respectively.

$$\Sigma_s^{in}(E) = \int d\omega \, D_o(\omega) \left[(N_w + 1)G^n(E + \hbar\omega) + N_\omega G^n(E - \hbar\omega) \right] \tag{3.27}$$

$$\Sigma_s^{out}(E) = \int d\omega \, D_o(\omega) \left[(N_w + 1)G^p(E + \hbar\omega) + N_\omega G^p(E - \hbar\omega) \right] \tag{3.28}$$

Assuming the phonons are in equilibrium, the Bose-Einstein statistic gives a temperature dependent phonon number shown in Eq. 3.29. Since later on we calculated the role of a narrow band of optical phonons which correspond to the optical density of states, $D_o(\omega)$ equals the deformational potential of $0.07eV^2$ [36] times the optical phonon density of states (Eq. 4.23)

$$N_\omega = \left[exp\left[\hbar\omega/k_B T \right] - 1 \right]^{-1} \tag{3.29}$$

The broadening function for scattering is then used to determine the scattering self-energy in Eq.3.30, which goes back in the self energy (Σ) term in the retarded Green's function (Eq. 3.21), where $\Sigma = \Sigma_i + \Sigma_j + \Sigma_s$

$$\Sigma_s = \frac{i\Gamma_s}{2} \tag{3.30}$$

Typically, Eq.3.26 is used for ballistic systems, where $\Sigma_s = 0$. For disordered systems such as graphene nano ribbons with edge roughness (Appendix I), where scattering is elastic (conserved energy and momentum), averaging over many coherent results with disorder embedded in the Hamiltonian generates phase independence and incoherence . Essentially this average is a way to capture elastic scattering by brute force numerics without having to determine Σ_s. When we are interested in phonons where Σ_s is non-zero, Eq. 3.24 or Eq. 3.25 are more convenient to use. In chapter 4, Eq. 3.24, Eq. 3.27, and Eq.3.28 will be used to determine high-bias current saturation due to optical phonon scattering.

Chapter 4

Understanding Graphene transport from Low to High Bias

4.1 Introduction

The overall purpose of this chapter is to investigate the role of band structure on low-to-high bias current using a unified model based on the Landauer Formula presented in the last chapter. The focus is on the role of charged impurities and phonons at both biasing extremes.

We will start at low-bias and break-down the fundamental band structure related trade-offs between opening a band gap to reduce OFF-current and its affect on mobility or ON-current. Quantifying this trade-off is important in motivating the purpose of the final chapter, where we present a way to bypass this mobility versus band gap trade-off. Projections and conclusions will be made on acoustic phonon limited mobility for various graphitic derivatives. Also at low bias charge impurities play a significant role on the minimum conductivity. Using a Landauer based model for conductivity we spell out for the first time the entire phase space of the minimum conductivity at both ballistic and diffusive limits in the presence of charge impurities and reveal both a high impurity saturation and an aspect ratio (W/L) depend flip in the conductivity.

Next at high bias we present a Landauer based device model that matches experiments at both the device and circuit level. Traditionally, compact models are constructed with equations fitted to experimental parameters. However the Landauer formula provides us with a solid physics based model. The first model we present is simple, assuming a pristine graphene density of states which works well for long channels. We show matches against both experimental device IVs and circuit voltage transfer curves.

4.1.1 Bandstructure limited mobility: redefining how we think of effective mass
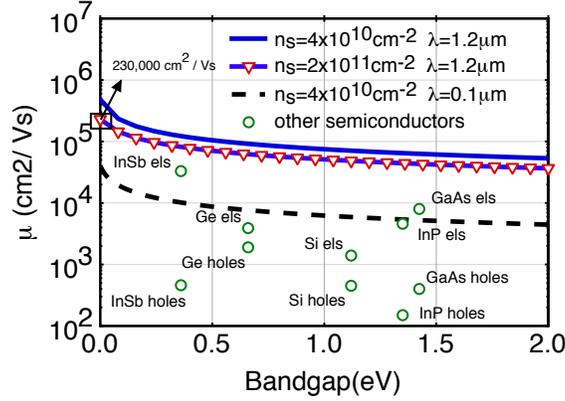
Figure 4.1: General projection of mobility versus band gap opening for graphene in comparison to the other common group IV and III-V semiconductors [6]. The projected trend is for photonics band structures with a simplified model for scattering. Experimental data points for the mobility of other semiconductors includes a large mixture of scattering mechanisms that can create variation in the $E_g - \mu$ trade-off.

Various efforts to engineer a band-gap have consistently seen gaps narrower than 400meV. However opening a band-gap in single-layer graphene(SLG), bilayer graphene(BLG), or graphene nanoribbons(GNRs) imposes a fundamental consequence on the low energy mobility due to the asymptotic constraint on the high-energy band- structure. While it is well understood that interfacial surface roughness, edge roughness due to patterning, and phonons from graphene and its interfaces can additionally degrade mobility through scattering, the scattering independent contribution from band-structure alone is often overlooked. Fig. 4.1 is a projected result as a function of bandgap.

Mobility is conventionally defined as, $\mu = q \langle \tau \rangle / m^*$, where the effective mass, m^* is usually defined by the inverse curvature at the band bottom, while scattering time, $\langle \tau \rangle$ is defined at the Fermi level. Due to band-related constraints this definition is suitable for wide-band-gap graphene or materials where the carrier kinetic energy varies slowly with change in crystal momentum. For consistency we want the effective mass at the Fermi level, m_F^*, which requires careful consideration for different graphitic derivatives 4.4.

Our general approach for defining mobility and effective mass stems from a low temperature, low bias form of Landauer formalism for transport (4.1),

$$I = \frac{2q}{h} \int_{\mu_2}^{\mu_1} dE\, T(E)\, M(E) \otimes (-\frac{df_o}{dE}) \tag{4.1}$$

where $M(E) = \frac{\pi \hbar v(E)}{L} D(E)$ is the number of modes, $T(E) = \frac{\lambda(E)}{L}$ is the transmission per mode. At low temperature and bias $\int \frac{df_o}{dE} dE = \mu_1 - \mu_2$ is just the drain bias qV_d. Normalizing with respect to drain bias gives the conductance and ultimately the conductivity, σ. To make the connection with mobility and
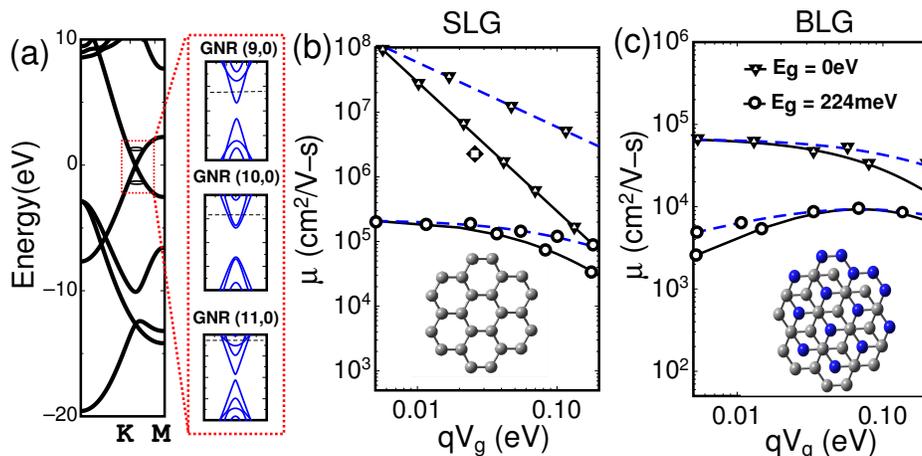
Figure 4.2: (a) Atomistic Extended Huckel Theoretical AGNR bandstructures confirm that reducing GNR width opens the bandgap while preserving the high energy bandstructure (in other words, a distortion of the bands rather than just a translation along the energy axis . (b)Carrier mobility degrades as Fermi level is biased deeper into the band (ie, increas- ing gate overdrive). (c) Low-bias mobility for bandgapped BLG shows a shows a non-monotonic dependence on gate overdrive. There is a sweet spot because while the effective mass reduces towards the band-edge, the density of states increases as well (Fig. 4.5a) and makes inter-subband scattering more effective.
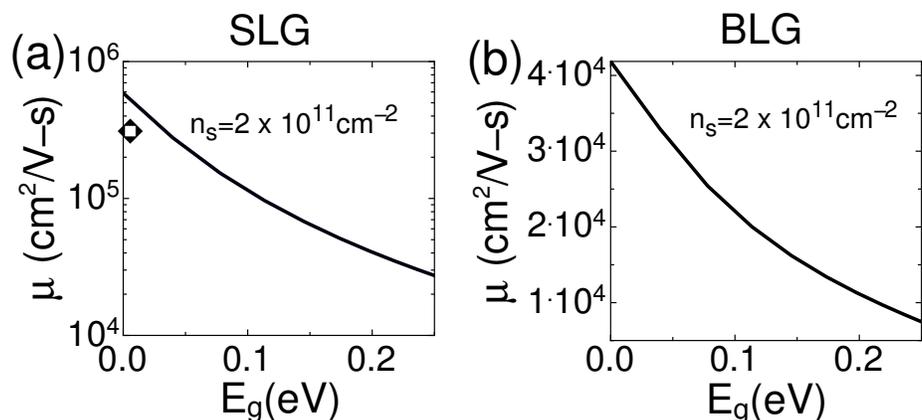
Figure 4.3: Mobility degrades with bandgap shown for (a) SLG and (b) BLG. Diamond-marker in (a) is a reference to the measured suspended SLG mobility ($\geq 230,000 cm2/Vs$)

ultimately m_F^* we relate σ from (4.1) to Einstein relation for a degenerate conductor, $\sigma = q^2 D_o D(E)$, where D_o is the diffusion constant and $D(E)$ is the density of states. To equate the degenerate conductivity in terms of mobility we must define the diffusion constant as $D_o = n_s \mu / q D(E)$, where n_s is the charge density. Sparing the details of specific graphitic derivatives for later, the general form for mobility boils down to $\mu = q\lambda/\hbar k_F$, where λ is the scattering length, $\hbar k_F$ is the crystal momentum. From which we can define effective mass as $m_F^* = \hbar k_F / v_F$, which is consistent with a carrier group velocity ($v_F = \frac{1}{\hbar}\frac{dE}{dk}$) at the Fermi level.

For a more rigorous derivation of the Fermi level dependent effective mass we start with the cyclotron effective mass defined in Eq. 4.2

$$m^* = \frac{\hbar}{2\pi} \frac{dA(E)}{dE} \tag{4.2}$$

$A(E)$ is the area in the crystal momentum space of the electron orbit in the presence of a magnetic field at a particular energy. Since graphene's Fermi surface is isotropic, we can define $A(k) = \pi k^2$. Using chain rule, Eq. 4.2 can be rewritten in the form shown in Eq. 4.3.

$$m^* = \frac{\hbar^2}{2\pi} \frac{dA_k}{dk}\frac{dk}{dE} = \frac{d(\pi k^2)}{dk}\frac{dk}{dE}\frac{\hbar^2}{2\pi} = \hbar^2 \frac{kdk}{dE} = \frac{\hbar k}{v} \tag{4.3}$$

Graphitic band-structures (Eq.(4.4)) have the general form, $E = \pm\sqrt{Eg^2/4 + (\hbar v_o k)^2}$, where $v_o \sim 10^8 cm/s$ is the carrier saturating velocity. A commonality among SLG, BLG, GNR is the high energy constraint on the band-structure, but the dynamics of bandgap (E_g) opening are particular to the different graphitic derivatives, which present similar yet unique trends in their mobility and effective mass.

$$E = \begin{cases} \sqrt{|t_1 - t_2|^2 + (\hbar v_o k)^2}, & \text{SLG} \\ \sqrt{\frac{U^2}{4} + \frac{t_\perp^2}{2} + s\frac{\sqrt{4\hbar^2 v_o^2 k^2 (U^2 + t_\perp^2) + t_\perp^4}}{2} + (\hbar v_o k)^2} & \text{BLG} \\ \sqrt{(E_n)^2 + (\hbar v_n k_x)^2}, & \text{GNR} \end{cases} \tag{4.4}$$

4.1.2 Trade-off between band gap and mobility

Having presented the general projected trends for acoustic phonon limited mobilities for bilayer and single-layer graphene,next we breakdown the Fermi level dependent effective mass and velocity and how both play a role on mobility.
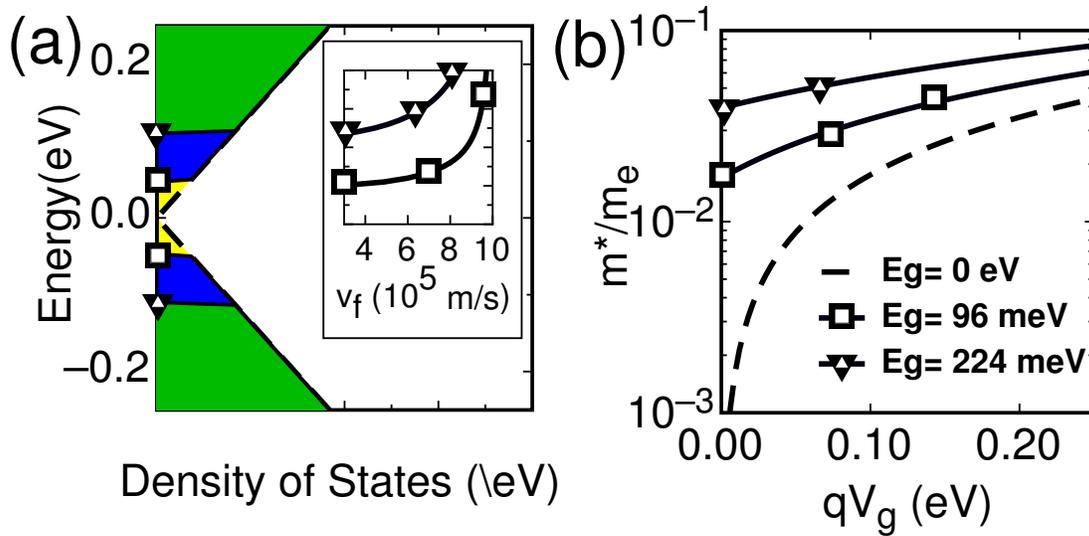
Figure 4.4: (a) SLG density of states for varying band-gaps. The increase in parabolicity with opening of a band-gap correlates with a corresponding reduction in Fermi velocity, v_F (inset), and (b) a slowly varying effective mass, m_F^*. The variability of m_F^* is higher for narrower band-gaps.

Single-Layer and Bilayer Graphene

Asymmetry induced bandgap opening is a common theme between SLG and BLG. Adsorption of Lithium atoms on SLG induces bond length distortions that create two types of couplings (t_1, t_2), inducing a bandgap equal to $2|t_1 - t_2|$ [37]. Meanwhile a interlayer coupling$(t_\perp)$ and an out-of-plane potential difference(U) in BLG extends a bandgap equal to $\frac{U t_\perp}{\sqrt{U^2 + t_\perp^2}}$ [38, 39]. As there are other methods of inducing bandgaps in two-dimensional graphene and more to be discovered, we assume that opening a bandgap in two-dimensional graphene preserves the isotropic constant energy surface, which allows us to define $k^2 = k_x^2 + k_y^2$ for both SLG and BLG. From which the density of states $(D(E))$ for SLG is $D(E) = \frac{2WLE}{\pi\hbar^2 v_o^2}\Theta(|E_f| \geq E_c)$ and for BLG is:

$$D(E) = \frac{4}{\pi\hbar^2 v_o^2} \begin{cases} \Delta_+ & \Theta\left(|E_f| \geq \frac{Ut_\perp}{2\sqrt{U^2+t_\perp^2}}\right) \\ \Delta_+ + \Delta_- & \Theta(|E_f| \geq t_\perp) \end{cases} BLG \tag{4.5}$$

,where $\Delta_\pm = \left[|E| \pm \frac{|E|(U^2+t_\perp^2)}{\sqrt{4E^2(U^2+t_\perp^2)-U^2t_\perp^2}}\right]$.

The density of states elucidates the role of distorted band-structure on effective mass and mobility. Plugging the number of modes, $M(E)$ into Landauer equation(Eq. 4.1) we get the conductivity for SLG (Eq.4.6) and BLG(Eq.4.7).

$$\sigma = \frac{4q^2}{h}\frac{\sqrt{E^2 - E_c^2}}{\pi\hbar v_o}\lambda \tag{4.6}$$
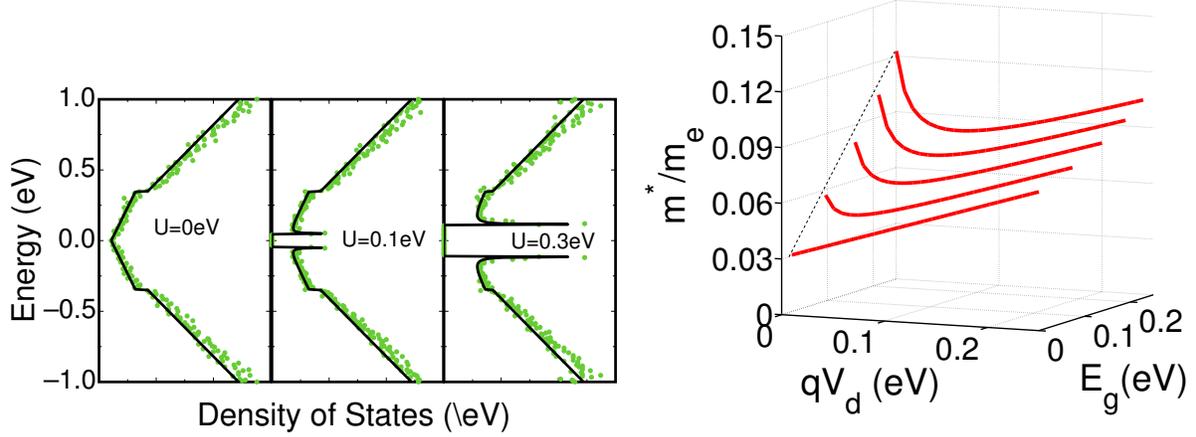
Figure 4.5: (a) Variation in BLG density of states with increasing potential difference U between the two Bernal stacked layers. The resulting Van-Hove singularities amplify phonon scattering at the band-edge. Coupled with reduction in effective mass at band-edges, the result is a maximum in mobility at intermediate values of the overdrive (Fig.4.2c)

$$\sigma = \frac{4q^2}{h} \frac{\lambda}{\hbar v_o} \sqrt{E^2 + \frac{V^2}{4} \pm \frac{\sqrt{4E^2(V^2 + t_\perp^2) - V^2 t_\perp^2}}{2}} \tag{4.7}$$

Meanwhile using Fermi's Golden Rule, $\frac{1}{\tau} = \frac{2\pi}{\hbar} \langle \psi_2 | U_s | \psi_1 \rangle^2 D(E)$ we determine the acoustic phonon scattering rate for SLG and BLG:

$$\frac{1}{\tau_F} = \begin{cases} \frac{D_A^2 k_B T}{4 \rho_m v_s^2 v_o^2 \hbar^3} E & SLG \\ \\ \frac{D_A^2 k_B T}{2 \rho_m v_s^2 v_o^2 \hbar^3} \Delta & BLG \end{cases} \tag{4.8}$$

,where the lattice strain from acoustic phonons described by the interaction potential(U_s) promotes a state transition from $\psi_1 \to \psi_2$. The deformation potential, D_A is 18±1eV for SLG and 15eV for BLG, while longitudinal acoustic phonon velocity ($v_s = 2.1 \times 10^6 cm/s$) and mass density($\rho = 7.6 \times 10^{-7} kg\, m^{-1}$) are the same for SLG and BLG [40, 41]. The energy dependent role of scattering rate on mobility is proportional to the number of available states for carriers to scatter [42] or $D(E)$(Fig. 4.4a, 4.5a), which increases with gate-overdrive and extended band-edges or band-gaps.

In addition to the $\sim 1/E$ from τ, an energy dependent, band-related $m_F^* = \hbar k_F / v_F$ further degrades mobility(Fig. 4.2b,c). The asymptotic constraint on the bandstructure enforces an energy dependent mobility trade-off. Mobility monotonically decreases with bandgap, but the non-monotonic behavior of BLG, $D(E)$ amplifies $1/\tau$ and m_F^* (Fig. 4.4b, 4.5b) near band-edge and creates a sweet-spot in the mobility that varies with band-gap:

$$max(\mu_{BLG}): \ qV_g = \frac{1}{2}\sqrt{\frac{\sqrt[3]{U^4 t_\perp^4 (U^2 + t_\perp^2)^2} + U^2 t_\perp^2}{U^2 + t_\perp^2}} \tag{4.9}$$

Eq. 4.9 is how far into the band we need to go to reach the minimum in effective mass (Fig. 4.5b) and corresponding maximum in the mobility (Fig. 4.2c). Note that the bandgap is related to the potential U between the two layers through $E_g = \frac{Ut_\perp}{\sqrt{U^2+t_\perp^2}}$. In general, $m^* \sim E_g$, $\tau \sim 1/E_g$, therefore the mobility goes as : $\mu \sim 1/Eg^2$.

Graphene Nanoribbons
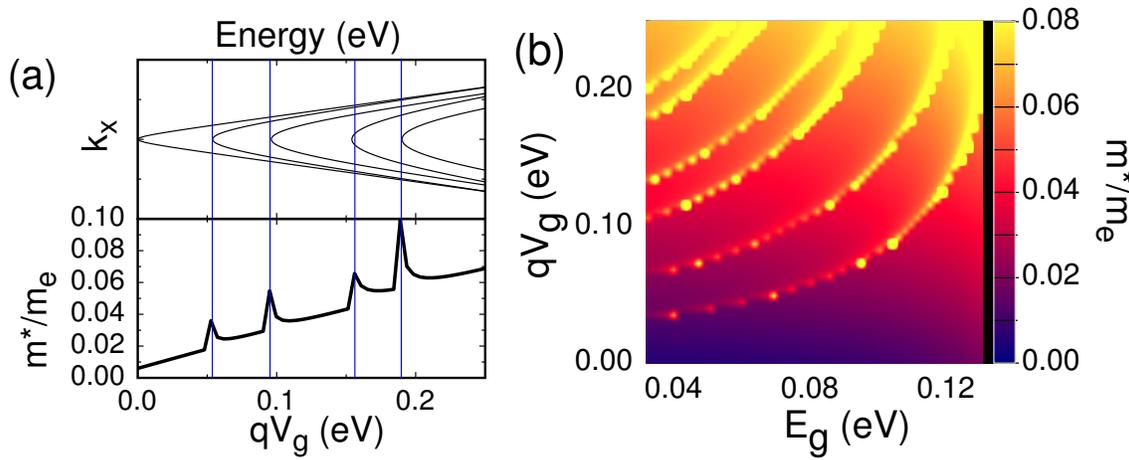
Figure 4.6: (a) Wider GNRs with a denser array of sub-bands have a greater probability of multiple sub-band crossings by the Fermi level compared to narrower GNRs. (b) Phase plot of m^* versus bandgap and Fermi level position inside the band illustrates the influence of multiple sub-band crossings on the transport effective mass.

The introduction of sub-bands in one-dimensional GNRs and CNTs raises the question of how to define m_F^* when gate-overdrive positions the Fermi level inside across multiple sub-bands? This question is particularly relevant for GNRs with widths between 20nm and 50nm as the density of sub-bands is proportional to ribbon GNR width or equivalent CNT diameter. As presented in Eq. 4.4, GNR band-structure is nearly identical to SLG, except for the n-th sub-band cut-off energy (E_n) and saturating velocity (v_n) are chiral specific quantities parameterized using third nearest neighbor tight-binding and benchmarked against Density Function Theory(DFT) [33] and Extended Huckel Theory(EHT) predicted band-structures.

In the case of crossing multiple sub-bands the mobility is defined as $\mu_F = \frac{q\lambda}{\hbar\langle k_{Fx}\rangle}$, where $\langle k_{Fx}\rangle = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{\hbar v_n}\sqrt{E_F^2 - E_n^2}$ and N is the number of sub-bands crossed. The average momentum $\hbar\langle k_{Fx}\rangle$ can be rewritten as $\langle\sum m_{Fn}^* v_{Fn}\rangle$ from which effective mass across multiple sub-bands can be extracted.

The general form for effective mass is $m_F^* = \hbar k_F / v_F$, where v_F is the average velocity across relevant sub-bands. Following the general approach presented earlier and noting that diffusion constant, $D_o = \tau \langle v_{Fx} \rangle^2 / 2$, the 1-D GNR average velocity is $\langle v_{Fx} \rangle = N \left[\sum_{n=1}^{N} \frac{|E_F|}{v_n \sqrt{E_F^2 - E_n^2}} \right]^{-1}$. From which the effective mass is just

$$m_{Fx}^* = \frac{1}{\langle v_{Fx} \rangle} \sum m_{Fn}^* v_{Fn} \tag{4.10}$$

Fig. 4.6a interestingly show jumps in m_F^* as the Fermi level reaches sub-band edges. By now m_F^* understandably increases with number of k-states, however at the edge of a sub-band the velocity contributed that particular sub-band is zero, therefore the decrease in $\langle v_{Fx} \rangle$ increases m_F^*. Fig. 4.6b is a phase plot illustrating the same trend across different overdrive and band-gaps for armchair-GNR(3p+1,0), where p are integers and $3p + 1$ denotes the GNR chirality. The brighter/ yellow curves and dots correspond to sub-band crossing and associated jumps in m_F^*. $1 - D$ graphitic mobility also follow the $1/Eg^2$ trade-off at the bottom of the band.

4.1.3 Concluding Remarks

Regardless of how scattering is or can be reduced, the band-related trade-off on mobility is unavoidable. The high energy constraint universal to the graphitic band-structure imposes increased parabolicity arises near band-edge with band-gap widening. Therefore, a parabolic to linear transition forces an increase in effective mass and decrease in mobility. Each graphitic derivative follows a similar yet unique band-gap versus mobility curve. The influence on mobility can be separated into scattering dependent (τ) and scattering independent (m_F^*) quantities. We projected the acoustic phonon limited mobility(Fig. 4.3) for suspended graphene. With the focus being on the scattering independent or band-related trade-offs we extracted effective mass for various SLG(Fig. 4.4b), Bernally-stacked BLG(Fig. 4.5b) and armchair-GNRs with edge strain(Fig. 4.6). We found that the non-monotonic nature of the density of states for BLG and GNRs had a distinct influence on mobility and effective mass compared to SLG. Thus we projected a sweet-spot in BLG's mobility versus overdrive. In general all graphitic derivatives follow a $1/E_g^2$ dependence. Understanding the role of this universal high-energy constraint on graphitic band-structure is an important design consideration for graphene-based electronic devices.

4.2   Finding the missing π: Quantifying Low-Bias Minimum Conductivity in the Presence of Impurities
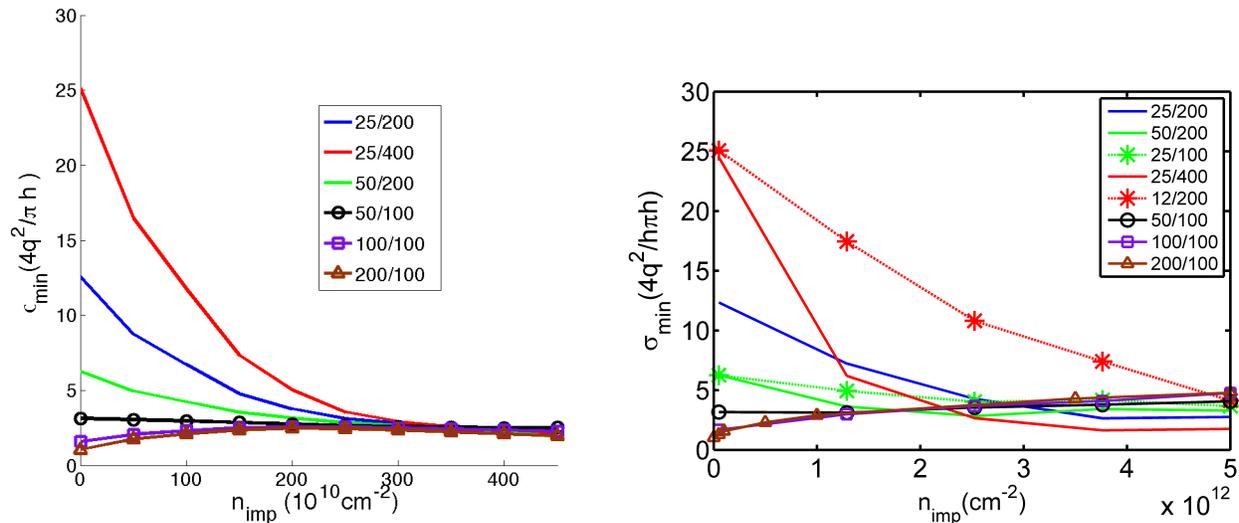
Figure 4.7: (top) Analytical results and (bottom) NEGF results courtesy of Redwan Sajjad for minimum conductivity versus n_{imp}

A unique electronic property of graphene is its minimum conductivity on the order of a single conducting mode, $4q/h$, where q is the electron charge, h is Planck's constant, and the pre-factor of 4 for spin and degeneracy, even at its zero-density point or the Dirac point. In the ballistic limit where the width of the graphene channel is much greater than the length both experiments [19] and Landauer based theory [20] predict a minimum conductivity of $4q^2/\pi h$. In the diffusive limit existing theory [21] and experiment [23] investigate the minimum conductivity in the presence of impurities, where a Boltzmann model is used for the theory. Authors Yang et al. in their work [24] have modeled the minimum conductivity in the ballistic limit with only three data points without conclusively showing the missing π.This section aims to describe away to analytically rationalize the minimum conductivity between the limits of low to high impurity concentration and small to large graphene aspect ratio.

The general Landauer conductivity is:

$$\sigma = G\frac{L}{W} = \frac{2q^2}{h}\overline{T}\frac{L}{W} \tag{4.11}$$

The total transmission is $\overline{T}$ is the number of modes times the probability per mode summed over all transverse modes, propagating and evanescent, Eq.4.12.

$$\overline{T} = \sum_{n=-\infty}^{n=\infty} T_n = M_p \frac{\lambda}{\lambda + L} + M_e T_e \qquad (4.12)$$

,where the index p refers to propagating and e refers to evanescent modes. The scattering length of the charge impurities λ is derived analytically (see Appendix iv) from Fermi's Golden Rule is only valid for propagating modes. The transmission probability T_n comes from solving a particle on a step, where the contact Fermi level is high doped compared to the graphene channel. The graphene transmission probability is similar to its free particle analogue,

$$T_n = \left| \frac{k}{k \cos(k L) + i (V_{rms}/\hbar v) \sin(k L)} \right|^2 \qquad (4.13)$$

V_{rms} represents the change in potential in the channel, $k = \sqrt{(V_{rms}/\hbar v)^2 - q_n^2}$ is the transverse wave-vector in channel that is summed over to get the total transmission. When k is real the the transverse modes are propagating, while when k is imaginary the transverse modes become evanescent. Imaginary k changes all the trigonometric functions to hyperbolic functions giving us the $1/cosh(k L)$ normally seen in the literature, assuming V_{rms} is zero.
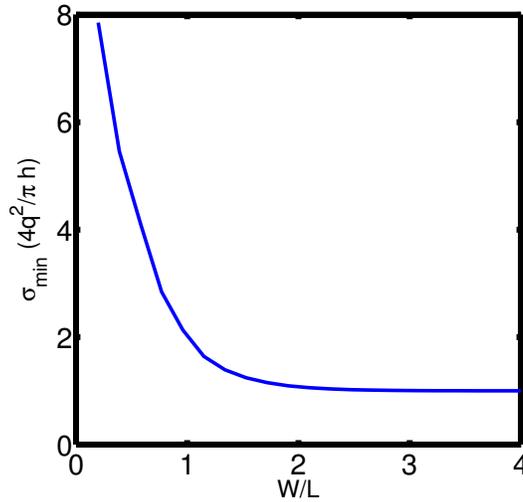
Figure 4.8: The minimum conductivity with no charged impurities at different aspect ratios.

Introduced back in chapter 3 in Fig. 3.2, the presence impurities erases the Dirac-point in the density of states. Charge impurities create a non-uniformity which means ideally we need to solve a particle in a non-uniform step potential. A particular impurity concentration will increase the minimum density of states to $\sqrt{\alpha \sigma_o^2}/(\pi \hbar^2 v^2)$, where σ_o^2 is the statistical variance of the normal distribution of potentials throughout the graphene. However, Eq. 4.13 comes from a solving the particle on a uniform step. We can approximate the

variation in the potential as a lumped element by defining the increase in the impurity affected density of states as an energy shift away from the Dirac point in the pristine graphene density of states ($eq : idealDOS$). In other words, we can treat V_{rms} in the Transmission probability(Eq. 4.13) as $\sqrt{\alpha \sigma_o^2}$.

Treating in the increase in impurities as an increase in the channel potential V_{rms} introduces more propagating modes (i.e., when $k = \sqrt{(V_{rms}/\hbar v)^2 - q_n^2}$ is real). Phase breaking incoherent scattering is accounted for in $\lambda/(\lambda + L)$. $\lambda = v\tau$ is the scattering length where v is the Fermi velocity ($\tilde{\ }10^6 m/s$) and τ is the scattering time derived in Appendix iV. from Fermi's Golden Rule and shown in Eq.4.14

$$\tau = \frac{2\pi}{\hbar} \frac{Z^2 q^4}{4\pi^2 8\epsilon_o^2 \epsilon_s^2} \left[\frac{\pi}{k(E)^2} + \frac{\pi\kappa(E)^2}{2k(E)^4} - \frac{\pi\kappa(E)^2}{2k(E)^4} \sqrt{\frac{\kappa(E)^2 + 4k(E)^2}{\kappa(E)^2}} \right] \left[\frac{\sqrt{(E+\mu)^2 + \alpha\sigma_0^2}}{\hbar^2 v_o^2} \right] n_{imp} \qquad (4.14)$$

Plugging all the elements in to Eq.4.11 while noting that only propagating modes have phase broken transmission probability per mode $\lambda/\lambda + L$(Eq.4.12), we arrive at the results shown in Fig.4.9.
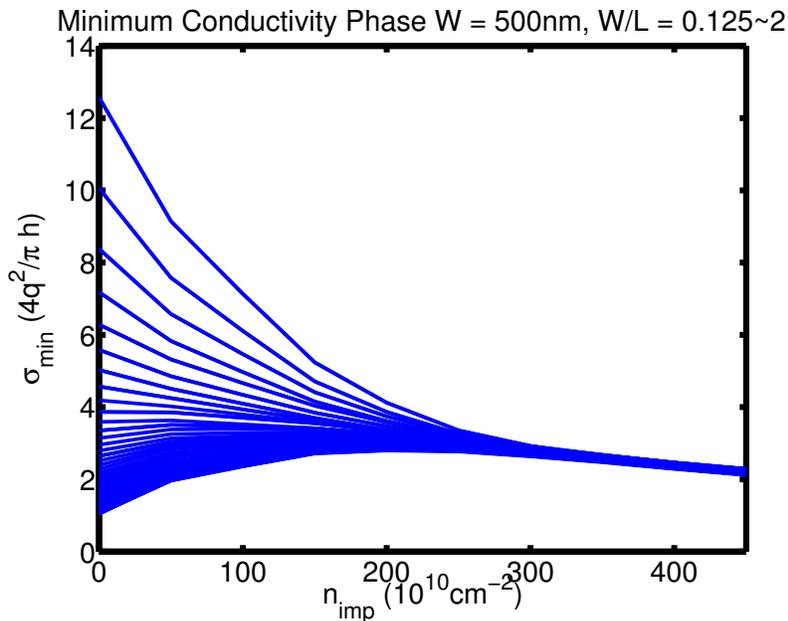
Figure 4.9: Phase plot of minimum conductivity for a fixed width of 500nm and aspect ratio varying from 0.124 to 2. Note that the smallest minimum conductivity is exactly at $\frac{4q^2}{\pi h}$. *The two important features seen for the first time in the graphene literature are (1) saturation at high impurity and (2) flip in curvature at low impurity.*

Our modified Landauer based theory matches the numerics from NEGF calculations at the same dimensions very well(Fig.4.7). Overall we see a general convergence or asymptotic behavior at high charge impurity concentration for wide width graphene. The ballistic behavior of the minimum conductivity is the rise in the minimum conductivity which agrees with the NEGF results courtesy of Redwan Sajjad, and follows the

general trend presented in [24]. The same unified Landauer based model also predicts the diffusive limit which matches qualitatively and possibly even quantitatively to [21, 23] assuming the dimension are known. *For the first time an entire phase plot of graphene's minimum conductivity is presented with a unified model.*

4.3 High-bias current : Impurity and Optical Phonon induced saturation

4.3.1 Analytical Compact Modeling of Graphene IV

From section 3.1.1 we saw that given a analytical band structure or dispersion relation, current voltage characteristics can be predicted using the Landauer formula, Eq.3.12. The band dispersion of graphitic materials, ranging from epi-G to sG, BLG, CNTs and GNRs are all described by a universal formula [43]

$$E = \pm\sqrt{E_{C,V}^2 + \hbar^2 v_0^2 k^2} \tag{4.15}$$

where the band-edges are at $E_{C,V}$ while the high energy velocity in the linear regime is $v_0 \approx 10^8$ cm/s. From the dispersion, we can readily extract the 2D density of states and band velocities

$$
\begin{aligned}
D(E) &= \left(\frac{2WL}{\pi\hbar^2 v_0^2}\right)|E|\left[\theta(E - E_C) + \theta(-E_V - E)\right]\\
v(E) &= v_0\sqrt{1 - E_{C,V}^2/E^2}
\end{aligned}
\tag{4.16}
$$

There is an additional energy dependence in the scattering length λ_{sc}. For ballistic channels, this is energy-independent, while for charge impurity and edge roughness scattering, $\lambda_{sc} \propto E$, while for acoustic phonon scattering, $\lambda_{sc} \propto 1/E$. The actual dependences are a bit more complicated, but these are reasonable approximations to adopt.

The algebra becomes particularly simple if we ignore the energy-dependence of λ_{sc}. We can then do the Landauer integral, leading to

$$I = \frac{8q}{h}\left(\frac{\lambda W}{\pi\hbar v_0 L}\right)I_0 \tag{4.17}$$

where the shape function I_0 depends on the current flow regime. Assuming we start with an n-doped graphene with a bandgap, we get

Band–to–band tunneling
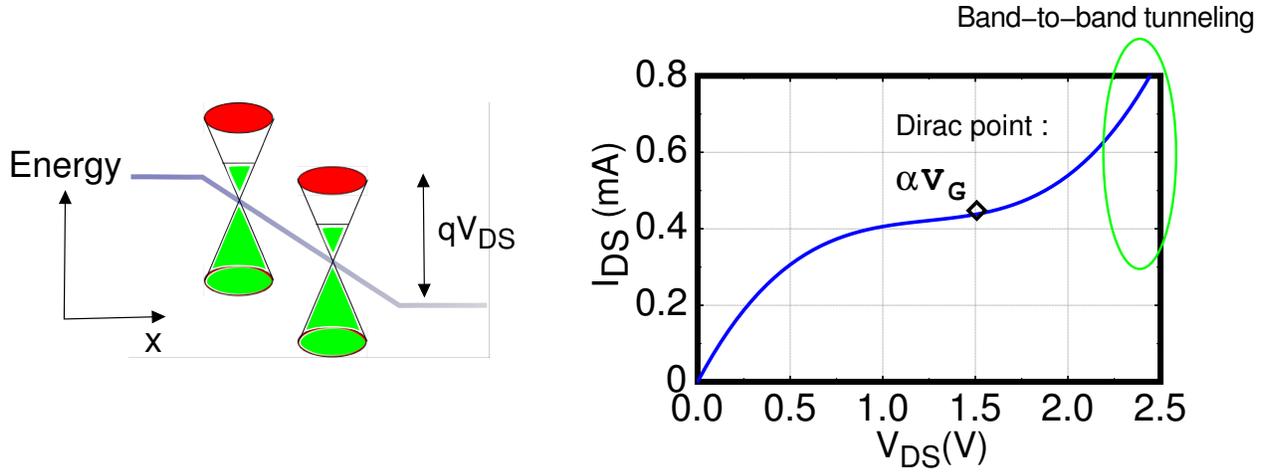
Figure 4.10: A typical I-V output shows how the I-V tends to saturate at the Dirac point even without a bandgap. The shift in the Dirac point indicates the Laplace potential drop along the channel, eventually leading to band-to-band tunneling.

$$I_0 = \begin{cases} \frac{1}{2}\left[\mu_1\sqrt{\mu_1^2 - E_C^2} - E_C^2 \cosh^{-1}\left(\frac{\mu_1}{E_C}\right) - \mu_2\sqrt{\mu_2^2 - E_C^2} + E_C^2 \cosh^{-1}\left(\frac{\mu_2}{E_C}\right)\right] \\ \qquad\qquad \text{if } qV_D < E_F - E_C, \\[2mm] \frac{1}{2}\left[\mu_1\sqrt{\mu_1^2 - E_C^2} - E_C^2 \cosh^{-1}\left(\frac{\mu_1}{E_C}\right)\right] \\ \qquad\qquad \text{if } E_F - E_C < qV_D < E_F + E_V, \\[2mm] \frac{1}{2}\left[\mu_1\sqrt{\mu_1^2 - E_C^2} - E_C^2 \cosh^{-1}\left(\frac{\mu_1}{E_C}\right)\right] - \frac{1}{2}\left[\mu_2\sqrt{\mu_2^2 - E_V^2} - E_V^2 \cosh^{-1}\left(\frac{\mu_2}{E_V}\right)\right] \\ \qquad\qquad \text{if } qV_D > E_F + E_V. \end{cases}$$

where $\mu_1 = E_F$ and $\mu_2 = E_F - qV_D$. The expressions can be further simplified. In the linear regime, the current looks like

$$I_{linear} \approx 2G_0 M \left(\frac{v_0}{v_F}\right)^2 V_D \tag{4.18}$$

where $G_0 = q^2/h$, the number of modes $M \approx 2W/(\lambda_F/2)$, and the Fermi velocity $v_F = v_0\sqrt{1 - E_C^2/E_F^2}$. The saturation current

$$I_{sat} \approx 4G_0 M \left(\frac{E_F}{2q}\right) \tag{4.19}$$

while the band-to-band tunneling current at high bias varies quadratically as

$$I_{BTB} \approx 4G_0 M \left(\frac{v_0}{v_F}\right) V_D \left(\frac{qV_D}{2E_F}\right) \tag{4.20}$$

Fig. 4.10 shows typical I-Vs based on the Eq. 4.20. These results agree with more involved, atomistic models for EHT coupled with non-equilibrium Green's function based simulations [44]. The current shows a point of inflection at the Dirac point, which is shifted by the gate bias (bandgaps would give more extended saturating regions, as we will see for our three terminal I-Vs later on). The subsequent rise in current is indicative of band-to-band tunneling. Furthermore, a prominent I-V asymmetry, consistent with experiments on SiC, can be engineered into our I-Vs (Fig. 4.11) readily by shifting the Fermi energy to simulate a charge transfer 'doping [45]' of 470meV through substrate impurities, back-bonding and/or charge puddle formation with SiC substrates. A mean-free-path(λ_{sc}) that varied inversely with gate voltage was implemented in the left figure in Fig. 4.11. For n-type conduction λ_{sc} ranged between 18nm to 40nm and 20nm to 31nm for p-type. Typically we would expect at least 100nm for low bias conductance and down to 10nm as the biasing approaches the saturation and band-to-band tunneling regions. Chosen λ_{sc} represents an average scattering length for the different regions. A more accurate model for scattering is necessary of which impurity scattering is derived in Appendix IV and acoustic phonons are presented in Eq. 4.8. In contrast, SiO_2 seems to dope the sheets minimally and the measured I-Vs show the expected symmetry between the electron and hole conducting sectors.
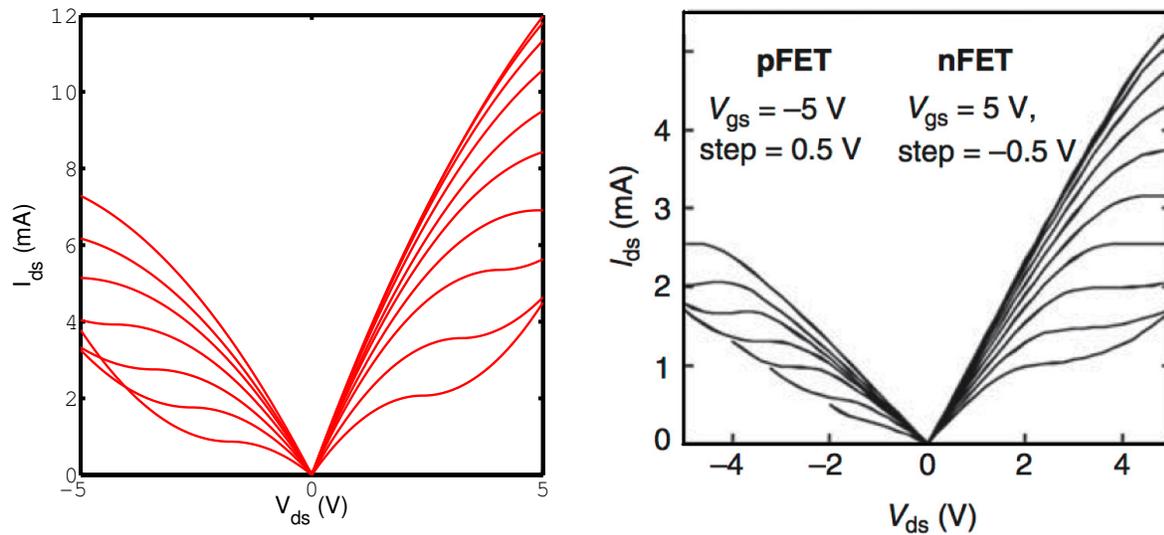
Figure 4.11: (Left) Theoretical and (Right) experimental I-Vs for graphene. The calculations on the left assume a 'doping' of the sheet by a charge density that shifts the K-point relative to neutrality. We also assume an inverse relation between the scattering length λ_{sc} and the applied voltage on the n-side, consistent with scattering by charge puddles associated with the above doping charge. The data on the right are for graphene on SiC, where charge puddles and/or back-bonding are expected to transfer a net charge density to the sheet [46].

The analytical current voltage expressions for a graphene three terminal device based on Landauer formula were translated into device model in Verilog to simulate a basic graphene CMOS-type circuits. The captured
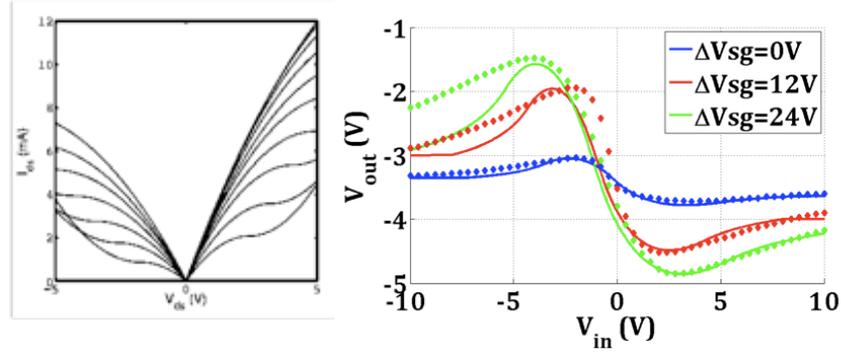
Figure 4.12: (left) Device model of graphene transistor showing an inflection in the IV due to tunneling currents. (right) Device model in Verliog used to simulate voltage transfer curve in Cadence of a graphene CMOS-type inverter. Solid line is experimental data and dotted line are from analytical compact model.The calculated and measured voltage-transfer curves (VTC) show gain away from the $V_{DD}/2$ as the transistor never turns off. In other words, for logic it is more important to have the rail voltages saturate in a VTC. However non-saturating rail voltages also deteriorates the gain in the transition of meta-stable region in the VTC [7]

upturn in the current voltage characteristic is responsible for the the poor pull-up of the output voltage at low input bias(V_{IN}) and poor pull-down of the output voltage at high input bias(V_{IN}). This is due to the non-complementary behavior of pull-up and pull-down networks as both N and P type graphene transistors are always ON creating a voltage division at the output instead of voltage railing. More details are given in Appendix II for work done in collaboration with Dincer Unluer.

4.3.2   Non-saturating IVs in the presence of impurities
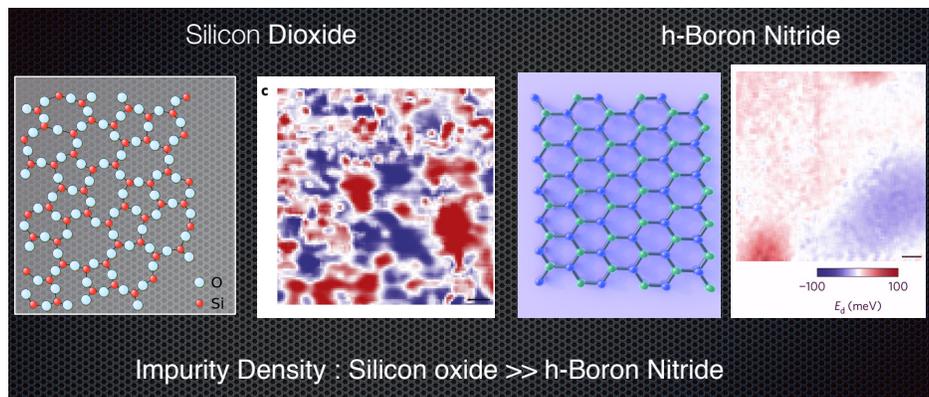
Figure 4.13: Spatially resolved potential profile in graphene on (left) silicon dioxide and (right) hexagonal-Boron Nitride [8].

Section 4.2 presented how charge impurities are responsible for increasing the minimum density of states which accounts for the missing factor of π in the minimum conductivity. At high-bias the quasi-current

saturation seen in section 4.3 turns near Ohmic for graphene field effect transistors less than 500nm on conventional oxides such as silicon dioxide and hafnium dioxide [47]. However on hexagonal-Boron Nitride, the quasi-current saturation is recovered in experiments. In a separate experiment, pulsed current-voltage(I-V) measurements which removing signatures of charge impurities trapped in the substrate recover the quasi-saturation in the IV. Scanning tunneling microscope (STM) images of the graphene potential landscape shown in Fig. 4.13 depicts a heavily doped and less uniform potential variation for graphene on conventional oxides. This section presents a high-bias current model based on Landauer formula with statistical parameters regarding charge impurity induced potential variation in the channel.

$$I = \frac{2q^2 W}{hL} \int_{\mu 1}^{\mu 2} \frac{\sqrt{(E+\mu)^2 + \alpha \sigma_o^2}}{\pi \hbar v} dE = \frac{\sqrt{\alpha}\sigma_o}{\pi \hbar v} \int \sqrt{tan^2(\Theta)+1} \sqrt{\alpha}\sigma_o sec^2(\Theta) \, d\Theta \qquad (4.21)$$

$$I = \frac{2q^2 W}{hL} \frac{\alpha \sigma_o^2}{\pi \hbar^2 v^2} \left\{ \frac{(E+\mu)\sqrt{(E+\mu)^2 + \alpha \sigma_o^2}}{\alpha \sigma_o^2} + ln \left| \frac{\sqrt{(E+\mu)^2 + \alpha \sigma_o^2} + E + \mu}{\sqrt{\alpha}\sigma_o} \right| + C \right\} \Bigg|_{E_f}^{E_f - qV_d} \qquad (4.22)$$

Using the puddled density of states(Eq.3.6) and Landauer formula (Eq.3.12) can analytical express the current voltage behavior of graphene transistor in the presence of impurities with one equation evaluated from the electrochemical potential of the source-end(E_f) to the electrochemical potential on the drain-end($E_f - qV_d$), Eq. 4.21. The complete form of the analytical I-V expression with statistical parameters is presented in Eq. 4.22 and plotted Fig. 4.14. In the short channel limit where phonon scattering is not less significant, the current-voltage characteristics in Fig. 4.14 illustrate charge impurities washing out any semblance of a Dirac point in the IV presented in section 4.3 due to the increase in the minimum density of states.
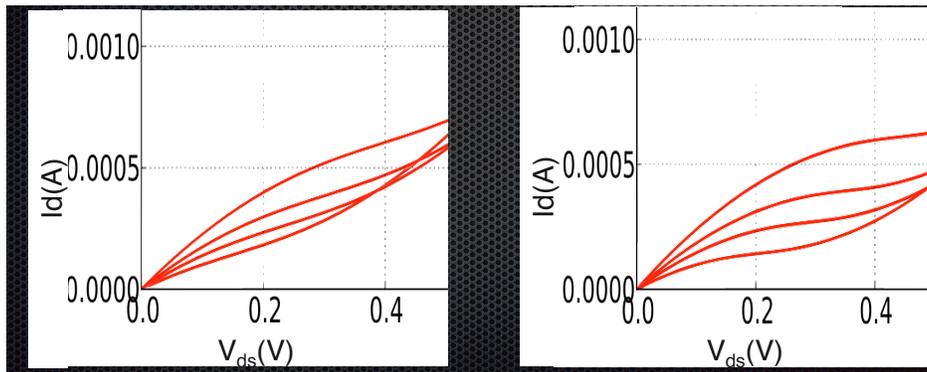
Figure 4.14: Output current voltage characteristics : (left) High impurity concentration, (right) Low impurity concentration.

This section uniquely presents a high bias current model based on Landauer formula with statistical

parameters with respect to the random potential variation that can be induced by charged impurities embedded or trapped in oxides. Conventional oxides and h-BN for the same channel length show different I-V characteristics which can be attributed to different concentration of charged impurities. *Overall the merits of a single current-voltage equation are its intuitiveness based on real statistical parameters describing a random channel potential and it being applicable as a device model abstraction for circuit level simulations as seen with a more simple example in appendix ii.*

At channel lengths on the order of microns, experimental I-Vs recover current saturation due to optical phonon scattering [48, 47]. The next section presents the role of in-elastic optical phonon scattering on current saturation in carbon nanotube using NEGF with self-consistent Born Approximation.

4.3.3   IV saturation in the presence of Phonons

At room temperature, crystal vibrations intrinsic to the channel and even from the surrounding are a non-negligible effect. Low energy vibrations are dominated by the acoustic phonon branch while the optical branch occurs much later at a high energy. Acoustic phonons are low energy phonons where the lattice vibrations are in phase. In graphene, we saw that the acoustic phonon mean free path is microns long and resulting in mobility upwards past 100,000 $cm^2/V-s$. The optical branch has a more significant effect and is known to saturate metallic carbon nanotube IVs when electron energy proportional to the applied drain voltage is large enough to emit an optical phonon corresponding to integer multiples of the optical phonon frequency (1600 cm^{-1} or 190meV) [9, 49]. The linear (Ohmic) to non-linear (saturating) transition in the metallic carbon nanotube IVs have only been reproduced with a Monte Carlo Simulations [50]. In this section, we use NEGF with self-consistent Born Approximation treatment of phonons, to reproduce this phonon induced saturation of a metallic nanotube with its length dependence, anticipating a similar effect in graphene. Experiments have shown this length dependence for long channels [48].

The general formalism to include scattering is discussed in section 3.1.2, where $\Sigma_s = D \otimes G^n$. The electron-optical phonon coupling for graphene is $0.07eV^2$ according to [36]. A simple Debye (single-mode) and a more realistic banded-mode for the optical phonons are used to simulate the role of optical phonon induced saturation graphitic (nanotube) current-voltage characteristics. Fig.4.15 illustrates difference between the Debye model and the banded model. The difference in implementation when calculating Σ_s is that Einstein model accounts for one phonon mode while the banded model requires a sum over a band of modes.

For the banded model, a Lorentzian function (Eq.4.23) similar to the thermal broadening function $F_T(E)$ is adopted to an optical phonon density of states with a finite bandwidth. The bandwidth $\Delta\hbar\omega_{op}$ is dependent
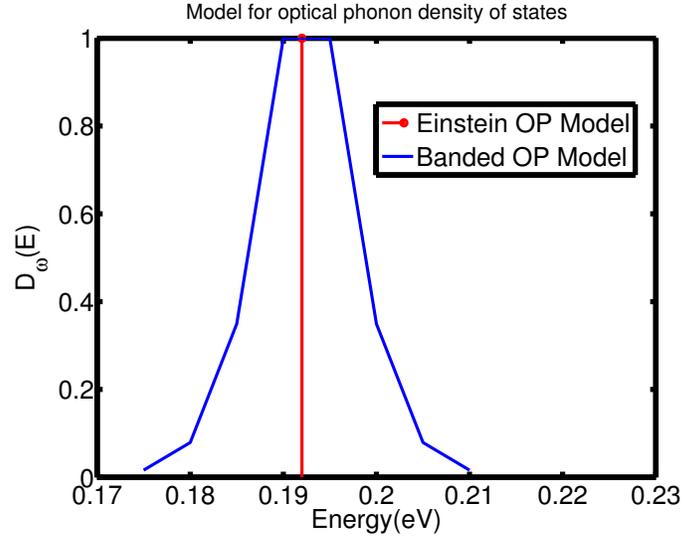
Figure 4.15: Einstein model vs Banded model for graphene optical phonon density of states.

on the phonon-phonon coupling parameter typically extracted from Density Function Theories, and outside the realm of this dissertation.

$$D_\omega(E) = sech^2(\frac{E - \hbar\omega_{op}}{2\Delta\hbar\omega_{op}}) \qquad (4.23)$$

Fig. 4.16 shows the NEGF calculation of a metallic nanotube with ohmic behavior transition to an current voltage characteristic with phonon induced saturation. The banded phonon model with the Lorenztian model for the optical phonon density of states saturates the current as the Σ_s is now a sum over all frequencies.

Fig. 4.17 shows the dependence on length seen in experiments and modeled with Monte Carlo methods. Increased channel length pronounces the current saturation as the channel length extends beyond the phonon mean-free-path increasing the chances of scattering. Carbon nanotubes are known to saturate around $29\mu A$ which is given Eq. 4.24 for an optical phonon emission at $190 meV$. Fig. 4.16 and Fig. 4.17 results do not include the double degeneracy there missing a factor of 4. Therefore our model for optical phonon induced saturation matches the expect current saturation ($29\mu A$) within a factor of 4.

$$I_{sat}^{op} = \frac{4q^2}{h}\hbar\omega_{op} \qquad (4.24)$$

This section presented showed optical phonon scattering turn-on in carbon nanotubes consistent with experiments [9]. As expected a metallic nanotube which normally would have Ohmic current-voltage characteristics saturates when the energy of the electrons proportional to an applied source-drain bias excites graphene optical phonon emission between $160 meV$ to $190 meV$. Realistically substrates with significantly
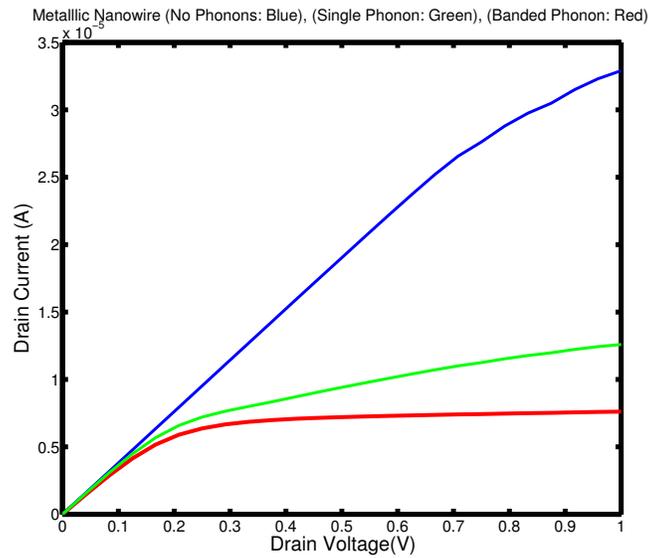
Figure 4.16: Ohmic(blue-curve) to non-linear saturating current voltage characteristics in the presence a single optical phonon(green-curve) and a Lorentzian band of optical phonons (red-curve).*Note that the IV curves on the right have not accounted for the double degeneracy in nanotubes.
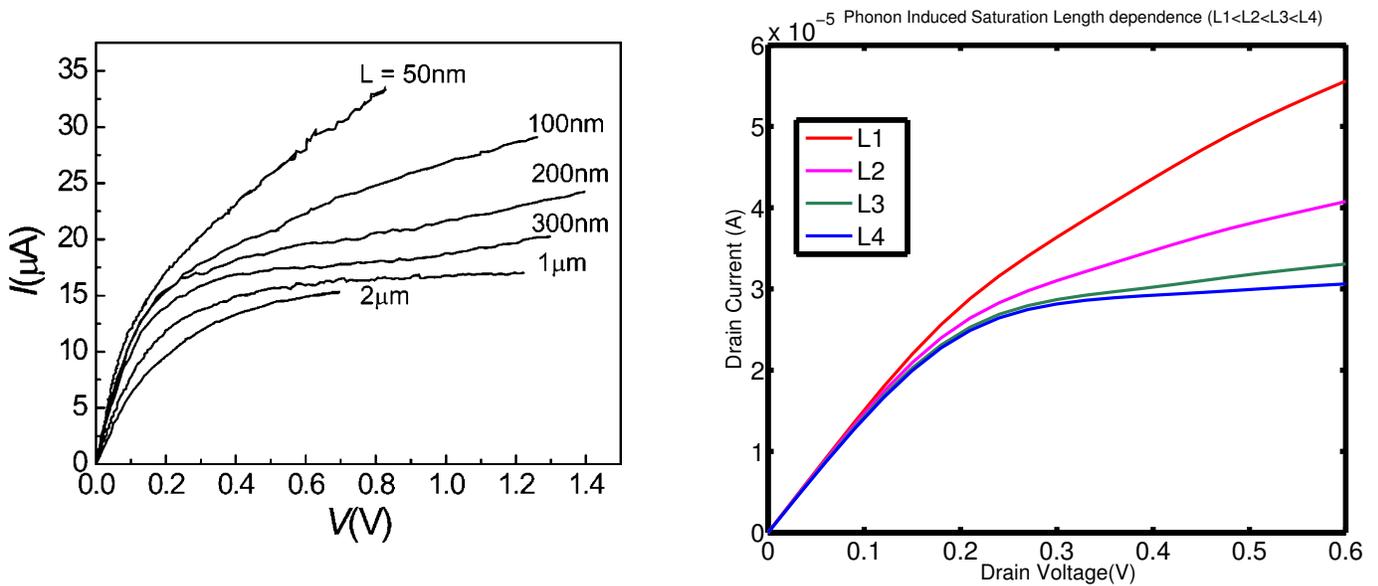
Figure 4.17: (left) Experiment of metallic nanotube current saturation due to phonons at different channel lengths. [9] (right) NEGF simulation showing current saturation with increased channel length, where $L1 < L2 < L3 < L4$. *Note that the IV curves on the right have not accounted for the double degeneracy in nanotubes.

lower optical branches (i.e., silicon dioxide: $\sim 55meV$ and h-BN: $\sim 100meV$) can remotely induced scattering and current saturation as seen with nanotubes (Fig. 4.17) at long channel lengths as demonstrated in experiments. A NEGF simulation using nano wires was presented due to the computation cost associated with simulating micron size channel length with the self-consistency within the Born Approximation.

4.4   Conclusion

This chapter builds a unique understanding of charge transport through in graphene in the presence of charged impurities across all its transport regimes, from ballistic to diffusive, low to high bias. A Landauer based device model captures the essential physics through the low energy bandstructure, where the channel is treated as lumped parameter. Charge impurities are important when conduction is through the lower energy density of states near Dirac-point. At low bias, we investigated the role of charge impurities on minimum conductivity at diffusive($W - L$) and ballistic($W >> L$) aspect ratios showing for the first time in the graphene literature the a complete phase transition with a turn-around when the aspect ratio W/L is approximately 0.5 to 1. At high-bias the presence of the Dirac-point creates an inflection in the current-voltage characteristic that disappears creating a more Ohmic curve in the short channel limit with increased impurity concentration. However at high-bias optical phonons are a competing scattering mechanism that can saturate the current. For this dissertation optical phonon induced saturation is demonstrated for a metallic nanotube. Furthermore, this chapter introduces in detail the fundamental band-related trade-off when a band gap is opened to reduce OFF-current or extended current saturation at the expense of effective mass and mobility. Quantifying the band gap-mobility trade-off is the main motivator for investigation in the next chapter where we show how to bypass material limits with gate engineering.

Chapter 5

Bypassing Material Limits: getting more from narrow bandgap graphitic channel

Recap : Graphene's mobility, structural flatness and tunable bandstructureare sought after features in the quest replace conventional silicon complementary metal oxide semiconductor(Si-CMOS) technology for both logic and RF-applications. As investigated in the previous chapters, graphene's large mobility $< 230,000 cm^2/Vs$ is a consequence of its zero-bandgap, which however makes a graphene field-effect transistor impossible to turn off by conventional means. This inability to turn off a graphene-FET was quantified at the end of chapter 3 where a quasi-analytical physics based compact model for a graphene field-effect transistor was presented. However opening a structural bandgap whether through confinement in graphene nanoribbons or broken sub-lattice symmetry with substrate doping or a perpendicular electric field in bernally stacked bilayer graphene, will invariably compromise the high-mobility due to the increased effective mass and increased scattering, as quantified in Chapter 4. The question confronted by the device community is if we can fabricate the most pristine channel material for high mobility but also have to engineer bandgap? This chapter restates it on asking a slightly differently on how to turn off a high mobility (narrow band gap) channel in a field effect transistor. So far there have been no suggestions on this to the best of our knowledge.

5.1   Introduction

The gateability of graphene is limited by its material constraint, specifically, an inverse relation that exists between its bandgap and effective mass. Indeed, its high anticipated mobility relies critically on its zero-bandgapped, linear dispersion that also makes it hard to suppress OFF-currents [51, 1]. Distorting this pristine linear dispersion with a structural bandgap, opened through quantum confinement [17, 25] or a disruption of its sublattice symmetry [52], decreases its OFF-current exponentially. At the same time, the wider bandgap also increases its transport effective mass [51, 29] and the density of states available for low energy scattering [42], hurting the overall mobility and thus the ON-current. We must thus carefully calibrate when we have a net positive in terms of ON-OFF ratio, mobility, transconductance, current drive and switching speed, as well as current saturation and its effect on the overall output conductance and device gain.

In this section, we use gate engineering (Fig.5.1 a,b) to form a staircase potential (Fig.5.1c) along a graphene sample with a small structural bandgap: a 5nm wide graphene nanoribbon (GNR) with an intrinsic bandgap of 170meV, which could in principle be single-layer graphene with a bandgap [52, 30], and a bernal stacked 2-D bilayer graphene (BLG) with a field induced bandgap of 150meV. The cascaded partially overlapping local bandgaps conspire to generate an effective wider transmission gap (Fig.5.4), filtering out transport through the intermediate conduction and valence bandgaps. Eliminating the staircase in the ON-state restores the mobility and thus preserves the ON-current even as we reduce the OFF-current and achieve a sizable current saturation in the output characteristics (Fig.5.8). Critical to the current-voltage characteristics is the biasing illustrated by the large-signal device layout in Fig.5.2.

We demonstrate proof of concept by numerically solving for the transfer $(I - V_G)$ and the output $(I - V_{DS})$ current-voltage characteristics of the cascaded gate FET from an atomistic treatment of the device, using the Non-Equilibrium Green's Function (NEGF) formalism coupled with fully self-consistent 3-D Poisson's equation, exploiting the open-source code NanoTCAD ViDES [53, 54]. Our simulation shows a reduction in OFF-current by two orders of magnitude in the $I - V_G$, and improved room temperature current saturation in $I - V_{DS}$, while fully preserving the ON-current (Fig.5.8). Since the underlying principle is quite general, *the method can be adapted to any narrow bandgap material with high mobility to improve gain in radio frequency (RF) as well as digital switching applications* [29].

5.2   Description of NanoTCAD ViDES

NanoTCAD ViDES is the primary tool used for this particular investigation. In addition to solving the same NEGF equations described in Chapter 3, it includes a 3D self-consistent Laplace-Poisson solver. Due to the computational complexity of determining the Green's function, NanTCAD ViDES uses recursive algorithm to determine the inverse of relevant diagonal blocks. The recursive algorithm reduces computational complexity from order $O(N^3)$ to order $O(N)$. NanoTCAD ViDES predominately uses nearest neighbor tight-binding model for graphene, nanoribbons, nanotubes, and bilayer graphene, but is expandable to handle non-orthogonal Hamiltonians from density functional theory(DFT) if additional chemical details are needed. Furthermore the self-consistent potential is solved using finite element with Newton-Raphson for convergence.

NanoTCAD ViDES also handles multi-threading and communication between different computers for parallelization using messaging passing interface for python(mpi4py). To conduct the work presented in this chapter NanoTCAD ViDES was integrated with University of Virginia's FIR Linux Cluster, which has 8-core computing nodes with 32GB of RAM per node. The parallelization makes the 100nm channel graphitic channel lengths for both a 5nm wide nanoribbon and 2D bernal stacked bilayer graphene computationally manageable with 3D grids and self-consistent Poisson potential.

5.3   Unconventional gate-layout.

The layouts for the GNR (Fig.5.1 a) and BLG (Fig.5.1 b) are different because the bilayer graphene requires a vertical field to induce a bandgap, and this field must stay constant across the staggered sections as we misalign the gaps. Fig.5.2 shows a bias diagram that applies to both structures. Referring to Fig.5.1 a, we simulated two vertically staggered top-gates that cover two-thirds of the graphene nanoribbon channel and are separated by gate dielectric thicknesses of d_1=4 nm and d_2=12 nm hexagonal Boron Nitride(h-BN). The back-gate covering the entire length of the channel is separated by 4nm of h-BN. The symmetry between the first top-gate dielectric thickness of 4 nm and the back-gate dielectric ensures that the cascade of bandgaps starts around $E_{fi} = 0eV$. The conductance of the remaining one-third underlapped region of the channel sitting next to the drain is modulated by the back-gate. For bilayer graphene (Fig.5.1b), three sets of top and embedded split back-gates are spaced equidistant across a 100 nm channel with 10 nm spacing between neighboring gates [55]. As long as the top to bottom differential gate bias, $|V_{TG} - V_{BG}|$, is held constant among the three set of gates while varying V_{TG} and V_{BG}, the field induced bandgap between the gates will stay the same even as we electrostatically dope the different segments of the graphitic channel away from $E_{Fi} = 0\ eV$ by the average potential between top and bottom layers, $(V_{TG} + V_{BG})/2$ [38].
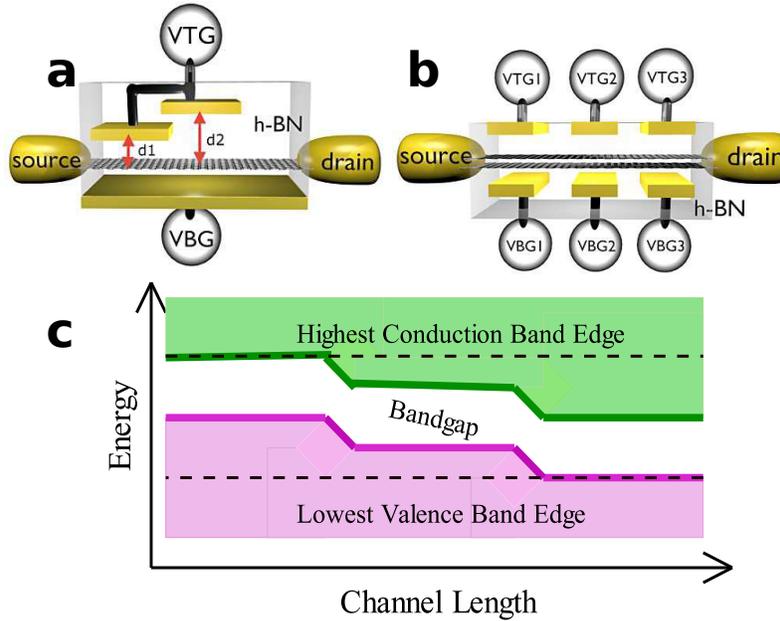
Figure 5.1: (a) Layout for $5nm$ nanoribbon channel with bandgap of $170meV$. $d1 = 4nm$, $d2 = 12nm$, top-gate lengths $= 35nm$, $|V_{TG}| = |V_{BG}| = 0.310V$ (b) Layout for $2D$ bilayer graphene. Gate lengths are $25nm$ and spaced apart by $10nm$. $|V_{TG} - V_{BG}| = 4.00V$ maintains field induced bandgap of $150meV$: $V_{TG1} = 2.00V$, $V_{BG1} = -2.00V$, $V_{TG2} = 2.20V$, $V_{BG2} = -1.80V$, $V_{TG3} = 2.40V$, $V_{BG3} = -1.60V$. (c) Staircase potential profile synonymous with the layouts from (a,b) promotes conduction through the highest conduction-band the lowest valence-band to prolong current saturation and reduce OFF-currents (Fig.5.8). The contact doping is $0.20eV$ (Fig.5.3)
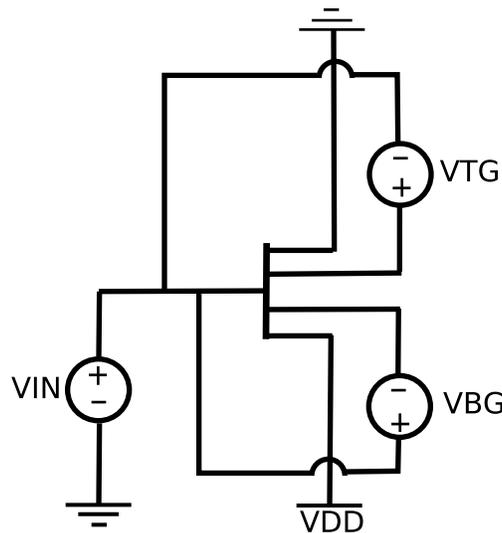
Figure 5.2: Large signal diagram for devices in Fig.5.1a and Fig.5.1b where DC voltages (V_{TG}, V_{BG}) are tied to input voltage V_{IN}. The supply voltage (V_{DD}) is the source-drain bias (V_{ds}) for the devices.
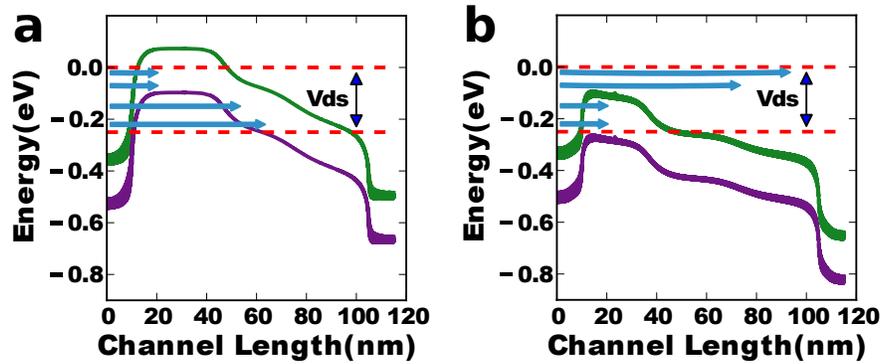
Figure 5.3: Computed self-consistent band diagram of the staircase potential profile at (a) $V_{IN} = 0V$ (OFF-state) and (b) $V_{IN} = 0.24\ V$ (ON-state) both constant drain-source bias of $Vds = 0.30\ V$. The intrinsic Fermi energy in each case is assumed at $0\ V$. Band-profile begins to flatten due to potential build-up from charges in the channel as the cascaded bandgap transistor is biased toward ON-state. The influence of staggered gates on channel potential depends on each gates transfer factor which decreases with distance from the channel. Shift in channel potential will be different per V_{IN}.

5.4   Transmission Gap: Filter conducting modes with a narrow band gap.

The deliberate misalignment of the local bandgapped regions allows transmission of electrons only above the highest available conduction band and below the lowest available valence band (Fig.5.1c). Between these two limits, the intermediate modes can transmit through inelastic processes, and to a lesser extent, by direct, ballistic quantum tunneling. The ballistic Landauer resistance is given by $R = (h/2q)M\bar{T}$, where $\bar{T}$ is the mode-averaged transmission and M is the number of modes at the Fermi energy, proportional to the band velocity and density of states. The series combination of Landauer resistances along the staircase potential is set by an equivalent parallel combination of the conducting modes available in each gated segment, $M \approx \left(M_1^{-1} + M_2^{-1} + \ldots\right)^{-1}$. The idea behind the cascaded gate device is to have all modes available in the ON-state (hence its high conductivity), but misalign the modes so that the overall mode-averaged transmission $\bar{T}$ goes down in the OFF-state.

Fig.5.3 illustrates the NEGF-computed energy band profile biased for the ON-state at $V_{IN} = 0.456\ V$ (Fig.5.3b) and for the OFF state at $V_{IN} = 0V$ (Fig.5.3a), with a fixed source-drain voltage $V_{ds} = 0.3\ V$. We apply the concept of parallel modes to both 1-D GNR and 2-D BLG and see a consistent widening of the transmission gap (Fig.5.4a and Fig.5.4c), even as the the mode averaged transmission above the highest conduction band and below the lowest valence band remains unaffected. Inside the transmission gap, the exact quantum mechanical tunneling probabilities are shown in Fig.5.4b and Fig.5.4d extracted from the retarded Green's function, and the broadening matrix obtained from a recursive treatment of the contacts,

assumed to be gate free extensions of the same core device structure to infinity.

The transmission coefficient computed through NEGF formalism is a numerically exact quantum mechanical solution that can be approximated with the Wentzel-Kramer-Brillouin(WKB) Approximation. Within the full energy range of a bandgap the elliptical nature of the energy dependent wavevector cannot be captured by real wavevectors describing propagating modes [56, 57]. Instead complex wavevectors are needed to describe the evanescent modes inside the bandgap, which are simply determined by solving the $k \bullet p$ 2x2 graphene Hamiltonian with a bandgap (E_g) and the relation $k = i\kappa$.

$$
\begin{bmatrix}
E - U(x) & \hbar v \left(k_x - i k_y \right) \\
\hbar v \left(k_x + i k_y \right) & E - U(x) - Eg
\end{bmatrix}
\tag{5.1}
$$

Hence, complex energy dependent wavevector inside the bandgap with a spatially varying potential along the channel, $U(x)$ can be expressed as,

$$
\kappa(E) = 2\sqrt{[E - U(x)][E_g + U(x) - E]}/\hbar v
\tag{5.2}
$$

Its reciprocal ($1/\kappa$) describes the decay length which increases with decreasing bandgap. The total tunneling probability created by the staircase $U(x)$ is just $T_{WKB} \simeq exp[-2 \int_0^L \kappa(E)dx]$. The approximate OFF-current which depends on how well the modes inside the transmission gap are suppressed by the staircase $U(x)$ for a given narrow bandgap is $I_{tunneling} = \int T_{WKB} \otimes F_T \, dE$ [32, 58], where the F_T is the thermal broadening function and the limits are defined within the transmission gap with bandwidth proportional to the drain-source bias($\sim qV_{ds}$). Fig.5.6 is comparison of T_{WKB} against numerically calculated transmission from NEGF. Due to the split gate geometry for BLG and the self consistent Poisson potential, the NEGF potential variation between each gate is slightly more complicated than the simple cascaded potentials (Fig. 5.5) used in our quasi-analytical T_{WKB} results, which is the source of a small quantitative discrepancy in Fig.5.6b.

The qualitative difference between bandgap and transmission gap is clearly seen in Fig.5.4b and Fig.5.4d. In contrast to an increasing structural bandgap where the OFF-current reduces exponentially, the increasing transmission gap for the cascaded gate device merely expands the low current region to the wings of the transmission gap that extend beyond the actual bandgap (here, < 200 meV). The net effect is still a hundred fold decrease in OFF-current and corresponding increase in saturation proportional to the width of the transmission gap. The *minimum* transmission, however, does not reduce upon further stacking of the cascaded gates, although it still benefits the saturation and thus the output impedance (Fig.5.8b , Fig.5.8d).

For N-type conduction a downward staircase potential profile cascades the bandgaps toward the original valence band energies. Fig.5.3 is an example for the nanoribbon where the downward staircase potential
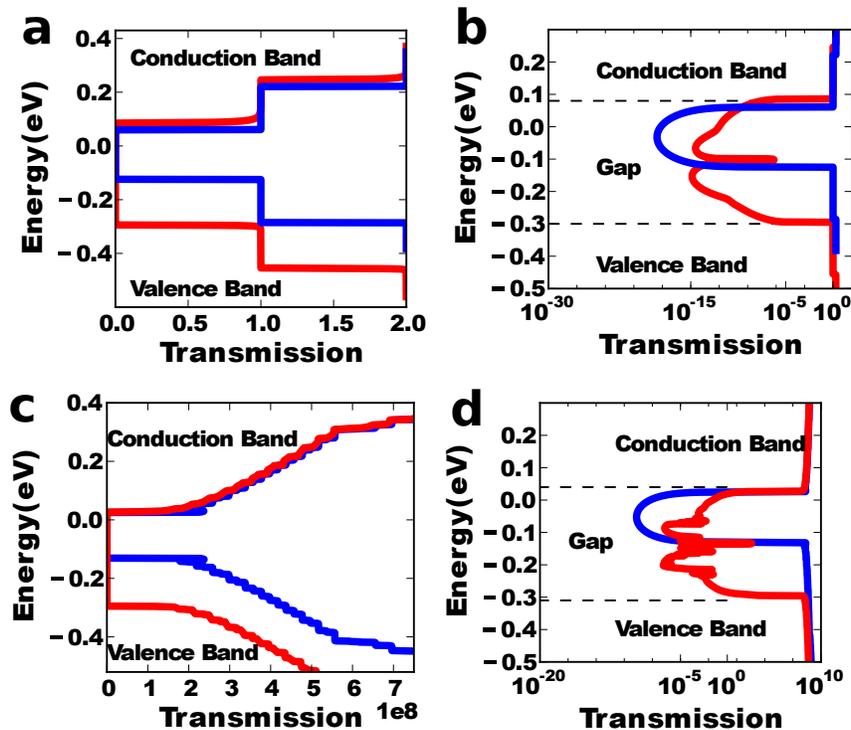
Figure 5.4: (a) Widening transmission gap for 5nm GNR with (red) staggered top-gate layout compared to (blue) conventional uniform gating. The transmission of the conducting modes stays pristine while (b) intermediate modes between the highest conduction band and the lowest valence band, revealed on a semilog-plot, are suppressed over a wider energy window as tunneling modes. (c) Widening transmission gap for 2-D BLG with an initial field induced gap of 150meV. The staggered gates widen the transmission gap to $\sim$ 400meV, even as the transmission of the conducting modes outside gap (and thus the ON-state mobility) are preserved. (d) Plotted on a log scale, we see the effect on the tunneling intermediate modes.
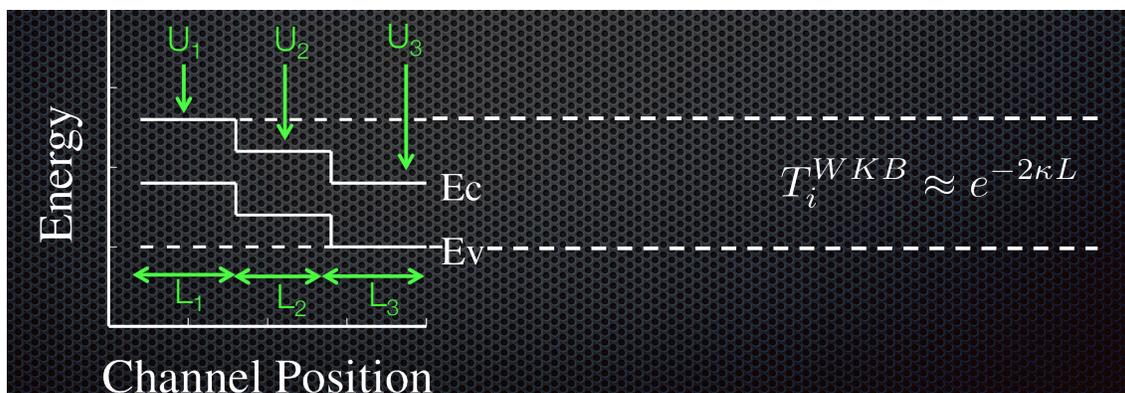
Figure 5.5: Cascade step potential. The total transmission is the product of the tunneling transmission probability from each section. Results are shown in Fig. refTranslog

profile is closest to E_{fi}. V_{TG} and V_{BG} have positive and negative bias, respectively. For P-type conduction an upward staircase potential profile does the opposite and cascades the bandgaps toward the conduction band, ensuring the lowest valence band is closest to E_{fi}. Likewise the V_{TG} and V_{BG} polarities can be reversed for
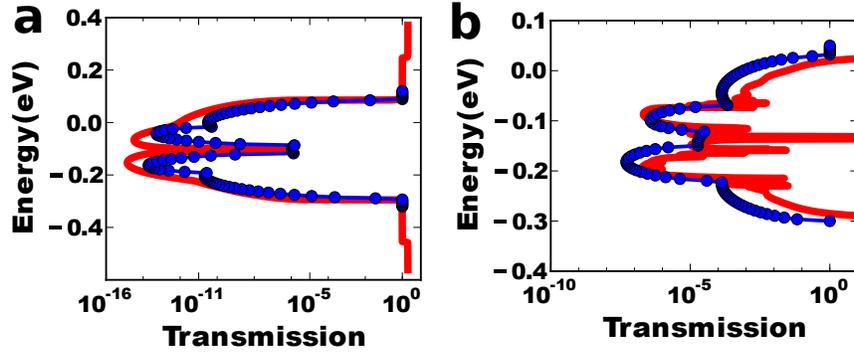
Figure 5.6: Comparison between (red) NEGF transmission and (blue circle-marker) T_{WKB} using $\kappa(E)$ from Eq. 1 for (a) graphene nanoribbon and (b) bilayer graphene with a staircase potential profile. T_{WKB} results are approximated with a simple series of three cascaded steps.

P-type conduction. See. Fig. 5.7.
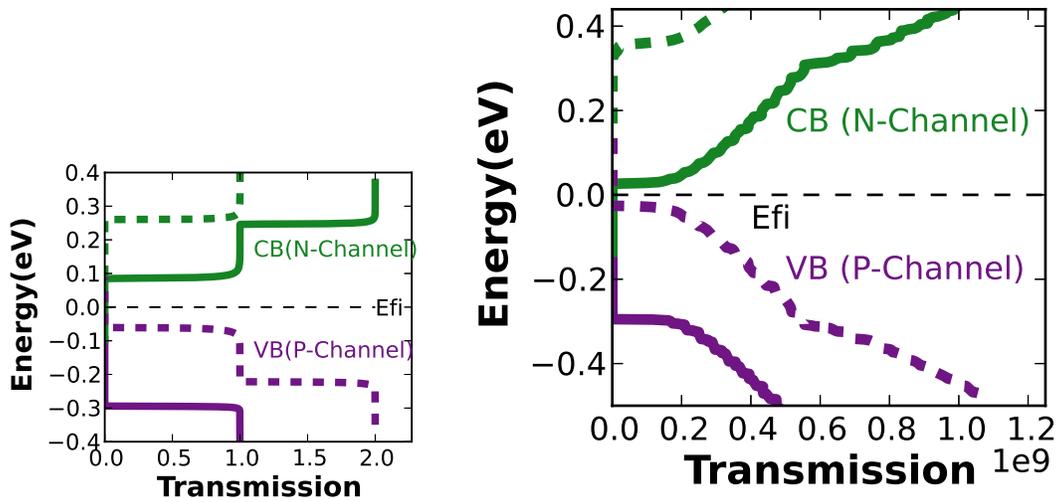
Figure 5.7: Transmission spectra that are preferential for N (solid) and P (dotted) conduction for (a) GNR and (b) BLG. A downward staircase potential(Fig.5.3) makes N-type conduction easier while an upward staircase potential makes P-type conduction easier.

5.5   Realizing the best of both worlds: Reduced OFF-current , Preserving ON-current / high mobility

The usefulness of the cascaded gate geometry is captured by Fig.5.8. In the gate transfer characteristics,

the effect is a decrease in OFF-current (and thus standby power dissipation) by two orders of magnitude

a

10^4
10^3
10^2
10^1
10^0

Ids (A/m)

Vds = 0.30V

−0.4 −0.2 0.0 0.2 0.4
VIN (V)

b

1.2
1.0
0.8
0.6
0.4
0.2
0.0

Ids(1e4 A/m)

0.0 0.2 0.4 0.6 0.8
Vds (V)

c

10^5
10^4
10^3
10^2
10^1
10^0

Ids (A/m)

(dotted) Vds= 0.20V
(solid) Vds= 0.35V

−0.4 −0.2 0.0 0.2 0.4
VIN (V)

d

14.0
12.0
10.0
8.0
6.0
4.0
2.0
0.0

Ids (1e3 A/m)

0.0 0.2 0.4 0.6 0.8 1.0
Vds (V)
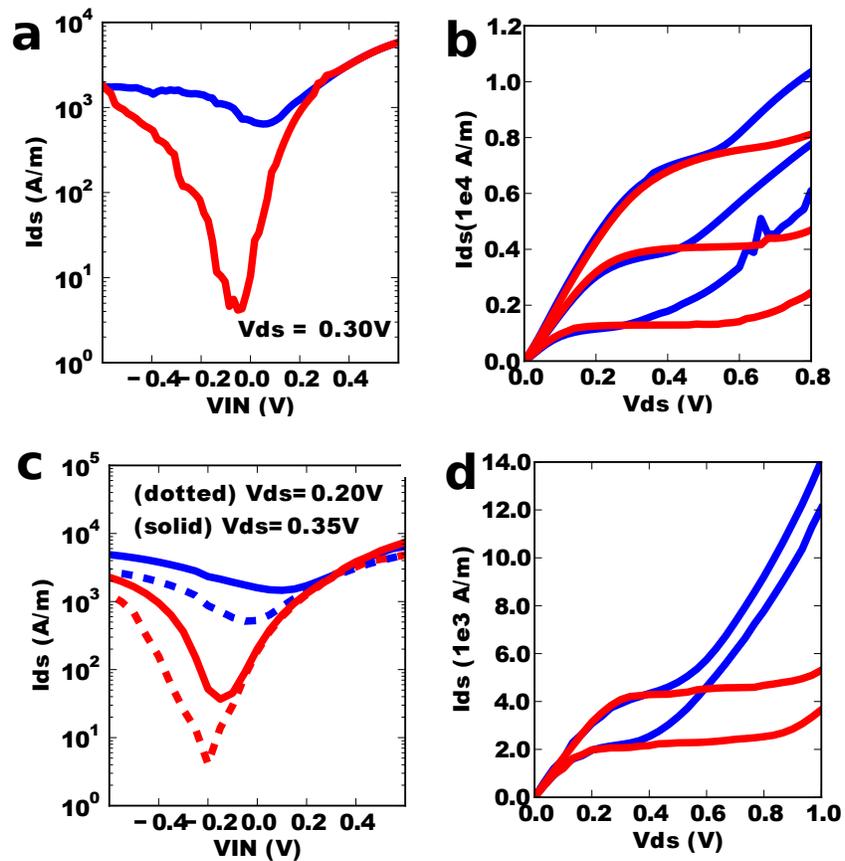
Figure 5.8: (a) Transfer and (b) output characteristics comparing a uniform potential profile(blue) and staircase potential profile(red) for 5nm GNRFET (Fig.5.1 a). (c) Transfer and (d) output characteristics comparing a uniform potential profile(blue) and staircase potential profile(red) for 2D-BLGFET (Fig.5.1 b).

(red curve) even as the ON-current stays pinned to the original single gate device (blue curve). This also corresponds to a significant improvement in saturation in the $I - V_{ds}$ characteristics across a 400 meV source-drain bias (red), compared to original (blue). Since all the modes are aligned and preserved in the ON state but merely misaligned in the OFF-state, the ON-current is demonstrably unaffected. This is also clear from the quantitative entries in Table I. Another significant improvement is a greater than 80% reduction in subthreshold slope, still bound by thermionic emission at $60 meV/decade$. Along with OFF-current, the direct source-to-drain tunneling current also decreases by at least a factor of two for a source-drain bias up to 800 mV.

One possible concern is that electrons can also cascade down the staggered potential through optical phonon emission when the source-drain bias reaches the optical phonon energy of either the graphene (190meV) or the BN-substrate(100meV). The ON-current is then expected to decrease while reduced OFF-current due

	Conventional GNR	Staircase Potential GNR
SS (mV/decade)	553	77
I_{OFF}(A/m)	625	4.01
g_{ds} (A/V)	1.44×10^{-5}	1.10×10^{-6}
g_m (A/V)	6.64×10^{-5}-5	6.64×10^{-5}
	Conventional BLG	Staircase Potential BLG
SS (mV/decade)	1000	180
I_{OFF}(A/m)	1990	56
g_{ds}	3.48×10^{-3}	1.10×10^{-3}
g_m	13.21×10^{-3}	17.98×10^{-3}

SS: subthreshold slope, I_{OFF}: OFF-current, g_{ds}: output conductance, g_m: transconductance

Table 5.1:

to the cascaded geometry can increase. However the phonon absorption process is more difficult and an the upward cascade for P-type conduction would see a reduced effect of phonon induced currents. [59]. In general, we believe the affect of phonons should be small; however at the consequence of a longer channel, a combination of downward followed by an upward cascade potential could filter phonon induced electron currents from emission, which is a topic of future work.

The split gate geometries shown here have not been optimized for electrostatics, in particular, parasitic capacitances between split gates. This is especially true for the BLG layout, where there are three sets of voltages for V_{TG} and V_{BG}. The parasitic problem could be mitigated for a GNR by the progressive stacking of top-gates tied to a common supply voltage V_{TG}. A possible way to remedy the parasitics is to use side-gates [60] that can also be engineered to dope a staircase profile, for the GNR, BLG, and other 2D materials. The advantage of side-gates is that the field lines are less dense compared to top-gates, and this should help with the parasitics [51]. The optimization of the electrostatics is an important concern that needs further design. We defer that to future studies, and focus here on establishing proof of concept - i.e., *a way to preserve the ON-current and reduce the OFF-current - thereby exploiting the strength of graphene while bypassing its obvious material weakness.*

5.6 Conclusion

The real merit of our model is the simplicity, and the use of contact engineering alone for momentum rather than energy filtering to reduce OFF-current and extend current saturation without hurting the ON-current (Fig.5.8) that made graphene so promising in the first place. Gates are uniquely positioned and biased to cascade their local narrow bandgaps along a staircase potential profile (Fig.5.1), suppressing the transmission of intermediate conducting modes between the highest conduction band and the lowest valence band (Fig.5.4).

The effective mode- filtering widen a gap in the transmission spectrum. The conventional approach of widening a real bandgap in the channel to reduce OFF-current comes at the expense of ON-current due to decrease in mobility. We show a way to bypass this fundamental material limitation. We established a proof-of-concept with a 5nm wide graphene nanoribbon and 2-D bilayer graphene, with bandgaps less than 200meV were convenient channels for simulation, but the concept of gate engineering can be generalized to other narrow bandgap materials with higher mobility.

Chapter 6

Conclusion

In conclusion, this dissertation has made the following contributions:

- Built a unified physics based model for graphene I-V from ballistic to diffusive limits and from low to high bias.

- Benchmarked with experiments and converted device model to Verliog and Cadence.

- Show how the minimum conductivity behaves in the ballistic and diffusive limit in the presence of impurities. We extracted the entire phase space and showed a flip in curvature followed by a saturation with increased impurity density.

- Showed how band structure is tied to the fundamental material trade-off between opening bandgaps and mobility. We find that band gap opening increases effective mass and reduces scattering time due to increase in band-edge density of states, thus reducing mobility by a factor of $1/E^2$. This happens for all graphitic derivatives.

- Show how optical phonons influence the high bias current voltage behavior leading to current saturation.

- Finally show how we can reduce OFF-current and extend current saturation without hurting the ON-current or mobility through momentum filtering aided by gate geometry engineering.

The work presented in this dissertation also contributes a series of unique physics based models derived from Landauer formula were developed which would useful for building device models for circuit level simulations. A circuit simulation is demonstrated with a simpler analytical physics based device model.

The real merit in addressing the first question is the simplicity and affiliation with traditional gate structures and conventional electron transport physics to reduce OFF-current and extend current saturation

without hurting the ON-current (Fig.5.8). Gates are uniquely positioned and biased to cascade local narrow bandgaps forming a staircase potential profile (Fig.5.1) that suppresses the transmission of intermediate conducting modes between the highest conduction band and the lowest valence band (Fig.5.4). The effective mode- filtering widens a gap in the transmission spectrum. The conventional approach of widening a real bandgap in the channel to reduce OFF-current comes at the expense of ON-current due to decrease in mobility. We it is possible to bypass this fundamental mobility and band gap limitation. A 5nm wide graphene nanoribbon and 2-D bilayer graphene with bandgaps less than 200meV were convenient channels for simulation, but the concept of gate engineering can be generalized to other narrow bandgap materials with higher mobility.

The work done in addressing the final two questions contributed to the understanding of conduction in a conventional graphene device from low to high bias in the presence of phonons and charge impurities, which the later is unique to the community. The beauty of using Landauer based models to look at mobility, minimum conductivity, and high bias current is its unified framework for conduction. This demonstration across a wide range of conduction is unique in the community. Future work could apply the more complicated models with charge impurities to investigate influences on circuit level performance.

.

Chapter 7

Appendix i

Graphene nanoribbons (GNRs) are widely being explored as potential channel materials for nanoelectronic devices [61]. GNRs share many advantages with carbon nanotubes (CNTs), including high mobility [61, 62] and minimal top surface dangling bonds. In addition, they raise the possibility of wide-area fabrication using top-down lithographic methods compatible with well established planar silicon CMOS technologies, as well as bottom-up chemical control through edge doping and intercalation. While the structural robustness of CNTs enables high-quality devices, it creates considerable challenges with chemical tunability and placement for circuit-level integration [63, 64]. Conversely, the structural amorphousness of GNRs opens them up to tunability and control, but increases their sensitivity to structural non-idealities that could degrade their current voltage (I-V) characteristics.
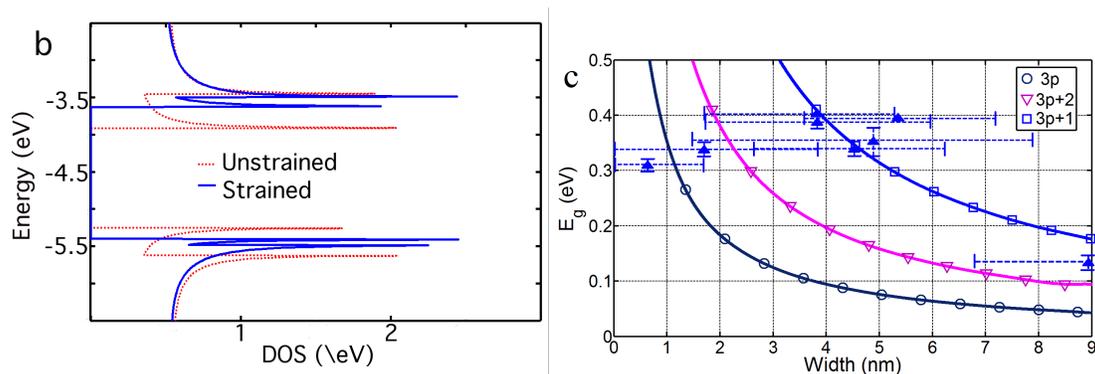
Figure 7.1: EHT captures the proper GNR chemistry, including (a) mid-gap states near the Fermi energy (-4.5 eV) arising from armchair edge dangling bonds (inset: local density of gap states). (b) H-passivation removes edge states, while soft edge boundaries prevent metallicity. A 3.5% edge strain further enhances the band-gap E_g. (c) The experimental lack of chirality and a strong clustering instead around AGNR (3p+1,0) [10] are attributed to chiral 'mixing' by edge roughness that favors the largest transmission bandgap among different segment widths (Fig. 7.2).

Crucial to GNRFET analysis is the employment of a well-benchmarked electronic structure theory that circumvents the computational complexity (and bandgap inaccuracies) of LDA-DFT [34], while capturing the complex chemistry and severe strain effects through transferable parameters that simple tight-binding cannot capture. Hence in this work we employ a bandstructure model based on EHT to properly capture the distortion chemistry in graphene [34], coupled with a NEGF formalism to explore the role of structural non-idealities on GNRFET performance. Starting with a reconciliation of the anomalous bandgapped structure and absence of prominent chirality dependence in experimental GNRFETs [5, 10], we catalogue some of the most influential non-idealities. We also identify critical parameters where these scattering events are most effective, as well as features in the I-Vs that are most robust to these variants. Since our analyses focus on the atomic and electronic structures, we considerably simplify the device electrostatics with a capacitive network, and discuss the complexities of their electrostatic potentials elsewhere [57].

Edge roughness and absence of chirality dependence. Experiments show that chemically derived, ultra-smooth GNRs narrower than 10nm have 'anomalous' bandgaps [10], in contrast to straightforward single-orbital tight-binding predictions. Similarly our EHT predictions in Fig. 7.1 shows this suppression of metallicity by the porosity of the GNR edges to transverse waves, amplified for narrow (< 10nm) ribbons by a 3.5% strain [3] at the edges introduced using geometry-optimized using classical molecular dynamics. Ab-initio [3, 17] and EHT theories both predict an oscillatory dependence of bandgap on ribbon width, superposed on an inverse power law (Fig. 7.1(c)). Such chirality dependences are absent in experiments of "ultrasmooth" GNRs [10] (dashed segments), showing instead a cluster around the 3p+1 nanoribbon results. Our simulations suggest (Fig. 7.2) that this arises from chiral mixing by edge roughness, that promotes the largest transmission band-gap among different width segments.
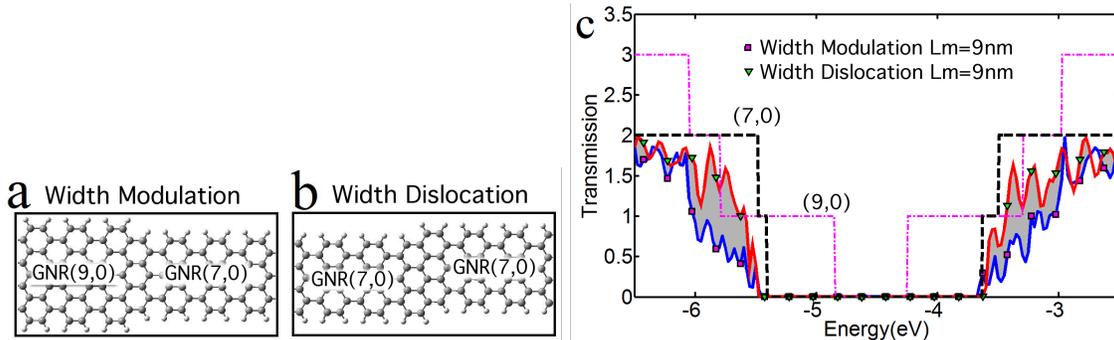
Figure 7.2: Edge roughness can be categorized as (a) mixed width (Δ_m=.246nm) vs (b) width dislocations (Δ_m=.123nm). (c) The transmissions with both kinds of roughness ($\Delta_m = .246$ and $.123$ nm, $L_m = 9$ nm, ribbon length L = 18nm) increase at energies corresponding to GNR (3p+1,0) band-edges showing the dominance of the largest band-gapped segment. However this bandgap dominance comes at the expense of suppressed transmission from mode-misalignment and electron backscattering near band-edges .
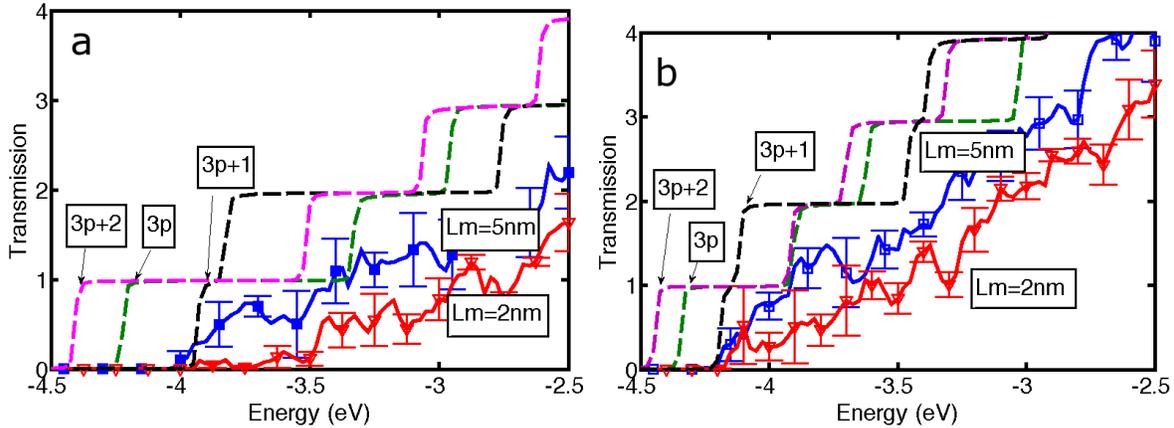
Figure 7.3: Using the same roughness patterns (Δ_m=.123nm and L_m =2 nm, Lm= 5nm) with multiple seeds on a narrow and wide AGNR we found that(a) narrower ($\sim$ 1.2 nm) nanoribbons (GNR(10,0)) with more roughness (L_m 2nm) suppresses transmission near band-edges to create psuedo-bandgap in the transmission. (b) While on a wider($\sim$ 2 nm) nanoribbon (GNR(19,0) with L_m= 2nm raises the transmission at the dominant 3p+1 GNR band-edge.

Atomistic fluctuations at GNR edges can be classified as width modulation, width dislocation or a combination of both (Fig. 7.2 (a,b)). We model edge roughness with a stochastic distribution following an exponential autovariance function [65] parametrized by a roughness amplitude (Δ_m) and a correlation length (L_m). The two classes of edge roughness produce significantly differing results, especially within the first 0.5 eV of the band-edge. Indeed, the modulating width creates significant backscattering at low energies where the mode count is sparser [62]. For longer correlations ($L_m = 5$ nm), Figs. 7.3 demonstrate that *edge roughness is dominated by the largest energy bandgap* corresponding to the (3p+1) segment. The absence of interband transitions in elastic edge scattering favors the higher HOMO (highest occupied molecular orbital) or lower LUMO (lowest unoccupied molecular orbital) level depending on n or p-type conduction. Faster edge oscillations with a small $L_m = 2$ nm, suppresses higher energy modes, opening an additional 0.50 eV pseudo-transmission bandgap (Fig. 7.3 (a)). *The pseudo-bandgap explains how the rough-edge band-gaps could exceed the (3p+1) prediction*, although the effect vanishes for wider segments (Fig. 7.3b). The net effect is an effective wash-out of chirality dependences, thus classifying AGNRs into *(a) ultra-wide, semi-metallic, and (b) ultra-narrow, semiconducting segments, in agreement with experiments* [5, 10].

Device Performance and Design. The salient features in the transmission from even small amounts of edge roughness are evident in the current-voltage characteristics of FETs (Fig. 7.4). I-Vs are calculated using NEGF by integrating the transmission energetically over the operating bias window imposed at the contacts. Since OFF currents depend exponentially on threshold voltage while ON currents vary linearly for a ballistic device, the former is affected more than the latter. To make a meaningful comparison, however, we adjusted the I-Vs to achieve a match of the OFF current and focused instead on the ON-OFF ratio at V_ds=500mV.
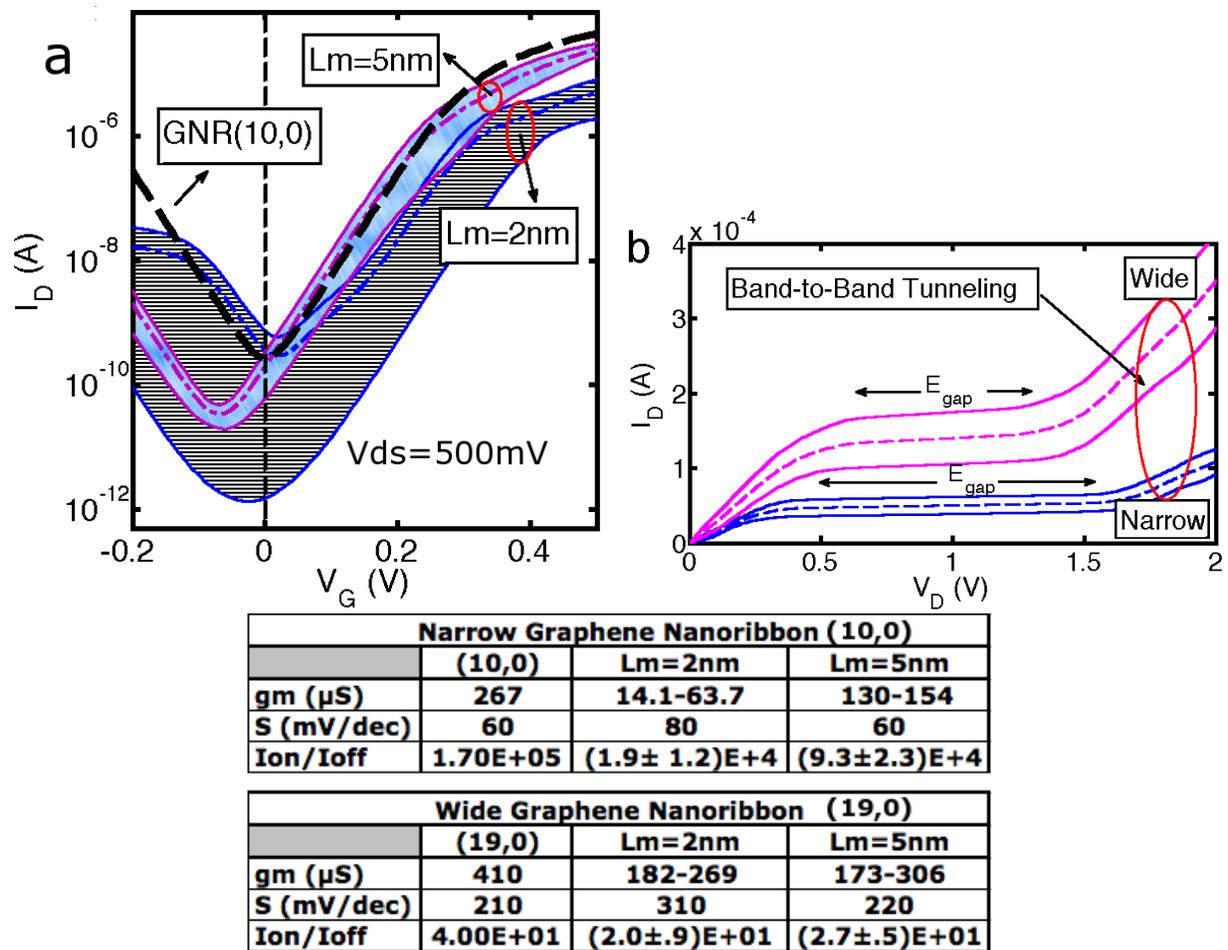
Figure 7.4: (a)Variance in drain current versus gate voltage characteristics for a narrower GNR(10,0) under multiple seeded roughness parameters (Δ_m=.123nm and L_m =2 nm (shaded) , L_m= 5nm (hatched)) with average OFF currents pinned at V_G=0. Better quality GNR edges (L_m=5nm) have less ON current and transconductance variance . (b) Threshold voltage is constrained by any variance in the transmission bandgap before additional applied V_{ds} causes Zener tunneling. Bottom table shows certain device performance metrics for wide and narrow GNR channel FETs.

As expected, shorter L_m and larger Δ_ms degrade the device mobility, sub-threshold slope, on-current and increases variance in device performance. The transconductance(gm) which scales with mobility can be remedied with wider widths as conduction mode count is denser. Narrower nanoribbons with $L_m > 5$ nm can achieve a faster turn off of drain current with a sub-threshold slopes (S) at 60 mV/dec comparable to that of a corresponding smooth-edge GNR. While shorter edge roughness correlation lengths degraded sub-threshold slope by at least 20 mV/dec. Similarly OFF current and threshold voltage is influenced as there is a large variance in the transmission around the transmission band-edges that can potentially open pseudo-transmission bandgaps. The variance in threshold voltage of at least 0.5V is an important design parameter considered for CMOS circuits such as SRAM (static random access memory).

As drain voltage has less influence on the GNR channel, threshold voltage is limited by size of the transmission bandgap and any variance due to edge roughness. *An applied drain bias beyond the transmission bandgap creates band-to-band (Zener) tunneling (Fig. 7.4 (b))* [66]. We note that our channel length (L=10nm) [67]is subject to tunneling, which slightly raises the OFF current and lowering ON-OFF ratio. Furthermore performance metrics in the table show that the denser mode count in wider GNRs (2.3nm) are less susceptible to effects of edge roughness as variance in ON-OFF ratio is less compared to the narrower GNR (1.2nm). We also found that stiff C-C bonds in graphene lattice prevent out-of-plane step heights even up to 0.5nm from influencing electronic properties and subsequently device performance GNRFETs [68].

In summary, chiral signatures enjoyed by CNTs are washed away for GNRs by edge roughness, especially for narrow ribbons. Hence ribbon width becomes the dominant parameter determining metallicity. This implies that absolute control of GNR widths is unnecessary, and the concept of an all-graphene wide-narrow-wide GNRFET is plausible even with structural nonidealities. The resulting large band-gaps make them gateable, but open them up to larger device-to-device fluctuations. Nonetheless, the devices are characterized by excellent sub-threshold slope, Ion/Ioff, and on-current, not to mention the considerable benefits that 2-D electrostatics, Ohmic contacts and high-k dielectrics in a wide-narrow-wide geometry can bring in addition [57].

Chapter 8

Appendix ii

8.1 Compact Model Equations

To simulate the performance of such a circuit, let us first outline a compact model. This will require us to outline (a) an equation for the bandstructure that includes effects due to edge strain and roughness, (b) an equation for the scattering length that depends on the phonon spectrum and edge roughness, (c) equations for the 2D electrostatics due to the source and drain contacts, and (d) the resulting I-Vs obtained by integrating the transmission over the relevant energy window.

To recap, the bandstructure of GNRs, including edge strain, can be written in a tight-binding form as $E = \pm\sqrt{E_{C,V}^2 + \hbar^2 v_0^2 k^2}$. Specific expressions for $E_{C,V}$ and v_0 for variously strained graphitic materials exist in Ref. [69].

The next term is the scattering λ_{sc}, which is related to the scattering time through an angle averaged geometrical factor and the overall Fermi velocity. The scattering time is extracted from Fermi's Golden Rule. For short range scattering by edge roughness and phonons, $\tau_{sc} \propto 1/|E|$, while for long ranged unscreened Coulomb scattering, $\tau_{sc} \propto |E|$. Explicit expressions exist in the literature [70] [71].

The tricky part that does not exist in the literature are the electrostatic capacitances associated with the 2D electrostatics from the planar source and drain contacts, competing with the top and bottom gates through their individual dielectrics. We are in the process of extracting formulae based on knowledge of planar micro-strip line electrostatics, with geometrical factors calibrated with our numerical MOM solutions for a variety of geometries [72].

Once we have the electrostatic, band and scattering parameters, we can then use Eq. 3.12 to extract the I-Vs. For energy-independent λ_{sc}, this was already shown earlier. We will generalize it to various scattering

configurations in our future work.

We thus have a comprehensive compact model that captures the chemistry and bandstructure, scattering, electrostatic and transport parameters needed for our circuit simulations. We will report one example here, and report further results in our subsequent publications.

8.2 Digital circuits:

Static complementary CMOS gates utilize pull-up (PUN) and pull-down (PDN) networks to achieve low power dissipation and large noise margin in logic circuits such as the inverter, NAND, and NOR gates. CMOS logic circuits are composed of some series and parallel combinations of n and p-type FETs. An inverter illustrated in Fig. 8.1 is the simplest logic element and the focus of this section of the review.
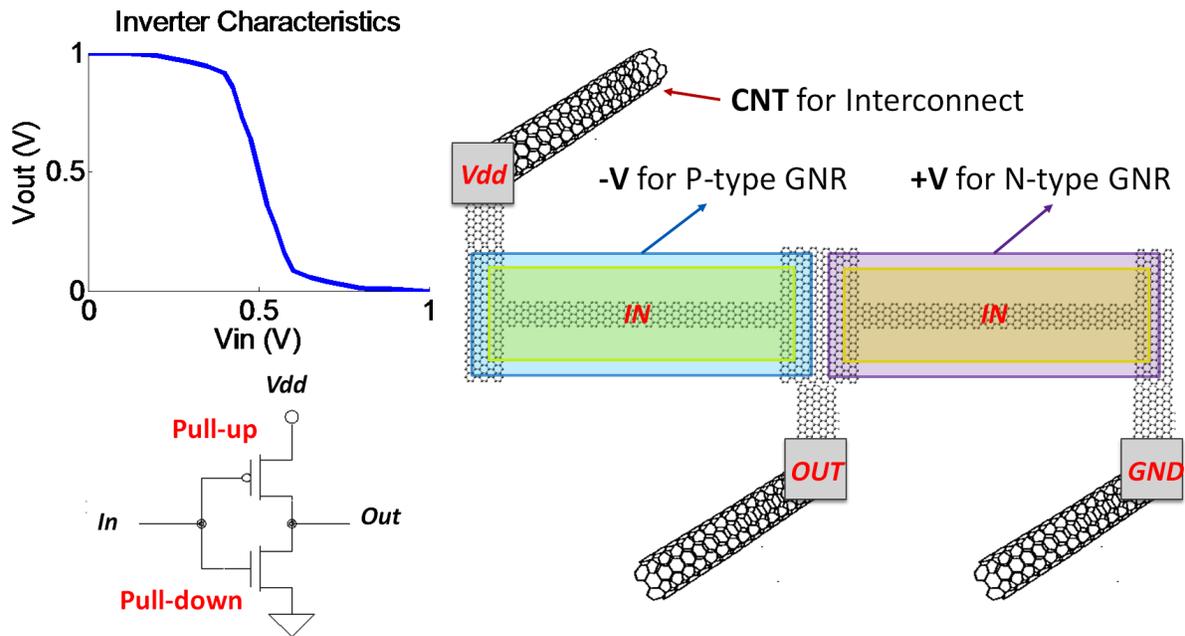
Figure 8.1: GNR inverter geometry and voltage transfer curve. This inverter design uses the WNW (metal-semiconductor-metal) all graphene structure for pull-up and pull-down networks. In this design CNT interconnects make direct contact with device level graphene. CNT/graphene interface has been experimentally demonstrated by Fujitsu Laboratories Ltd [11, 12].

When the input into the common gate is Vin=0 the p-type FET (PUN) is active while the n-type FET (PDN) is cut-off, hence the circuit will pull the output voltage up toward the supply voltage (Vdd) or high, Vout=1. Likewise when Vin=1, n-type FET is active and p-type FET is cut-off pulling the circuit down toward ground, Vout=0. Usually it is impossible to pull-up or pull-down to exact values of 1 or 0, so threshold voltage and tolerance are designed for each circuit to help distinguishing between these two logic levels.

Circuit designers allow some tolerance in the voltage levels used to avoid conditions that generate intermediate levels that are undefined. For example, 0 to 0.2V on the output can represent logic (0) and 0.3 to 0.5V can show (1), making the 0.2 to 0.3V range invalid, not metastable, since the circuits cannot instantly change voltage levels.

The voltage-transfer curve (VTC) of an inverter circuit captures the DC or steady-state of specific input versus output voltages and provides a figure of merit for the static behavior of the inverter. VTCs for logic circuits provide information on operating logic-levels at the output, noise margins, and gain. Ideally we want the VTC to appear as an inverted step-function, indicating precise switching between the on and off stages, but in real devices there is a continuous transition between on and off states. From the VTC we can extract noise margin (Fig. 8.2), which provides a measure of circuit reliability and predictability. Biasing outside the noise margin puts the logic circuit in an unpredictable state. Circuit designers want to maximize the noise margins.
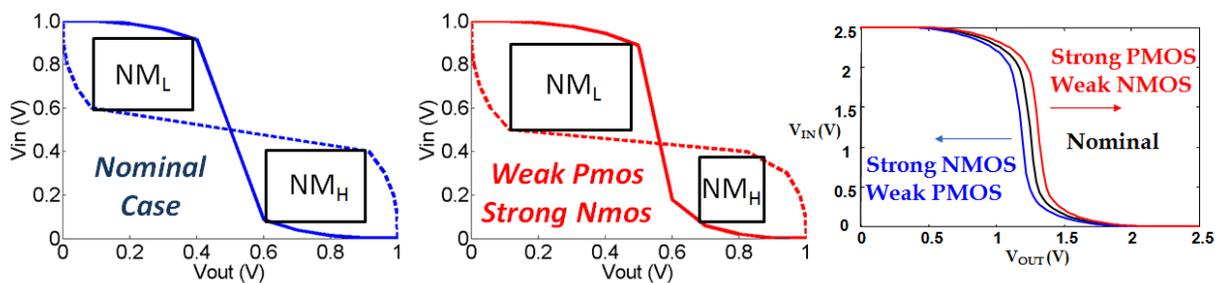
Figure 8.2: showing the importance of balancing CMOS transistor sizes to achieve equal high and low noise margins(NM). The noise margin is graphically represented by the largest square that fits inside the enclosed space outlined by normal and rotated VTCs.

A significant advantage of graphene is its intrinsic electron-hole effective mass symmetry. In the absence of extrinsic doping a graphene based FET the I-V characteristics for n and p-type conduction would be the symmetric. However, asymmetry can be introduced into the system through charge-transfer doping [45] (e.g. Fig. 4.11) or by contact induced doping [73]. Significant screening of charge impurities in the substrate should bring Fermi level closer to its intrinsic value at the Dirac or $K-$point, therefore recovering symmetric n and p-type I-V characteristics. *On a circuit level this symmetry means the response of PUN and PDN would be equal and opposite, which is important for circuit reliability, and not to mention ease of circuit design.* In conventional Si-CMOS logic circuits, the asymmetry in the electron-hole effective mass is compensated by scaling the physical width of the p-type FETs in the PUN so the I-Vs are equal and opposite with the PDN. Graphene's natural electron-hole symmetry would allow circuit designers to bypass this design issue.

A major impediment to GNR based logic circuits is its narrow bandgap ($\leqslant 200 meV$), as the device

elements in the PUN and PDN are prone to sub-threshold leakage from band-to-band tunneling. The two-fold effect on an GNRFET-based inverter where the channel has a narrow bandgap is demonstrated in Fig. 8.3. The first effect is a large voltage swing of approximately 0.4V. The second effect is a significantly diminished noise margin. Band-to-band tunneling in narrow bandgap GNRFETs prevents either the PUN or the PDN from completely cutting off when its complement network is active.
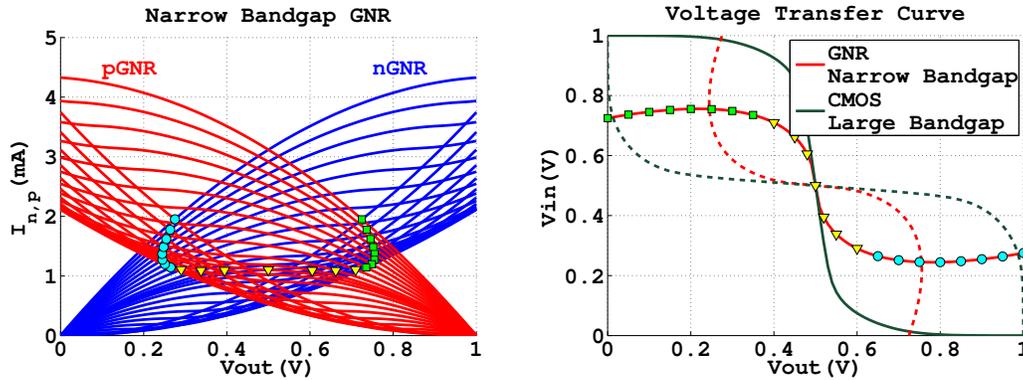
Figure 8.3: Comparison of VTC curves for narrow bandgap GNR and 45nm CMOS Technology. Narrow bandgap GNRFETs will be more susceptible to noise than CMOS due to smaller noise margins.

Fig. 8.1 shows the physical layout of a functional graphene inverter composed of WNW P-type and N-type GNR device arrays and the voltage transfer curve. The inverter voltage-transfer curve and gain can be calculated readily from the current-voltage characteristics. As expected the gain of the device determined by the electrostatics, geometrical parameters, and mobilities which ultimately determine the P and N-type GNR transconductances. The VTC above with gain of 4 is derived from the I-V shown in Fig. 8.1 for the 8.66nm device by using the methodology described in detail in [74]. These I-Vs generated in SPICE can be used to simulate other complex layouts such as NAND or NOR gates (The results of these logic gates will be reported in future publications).

Propagation delay can be measured by pulsing the input voltage between 0 and 1 and observing the output transient response. The transit time for a GNRFET is approximately L/v, where L is the length of the channel and v is an energy dependent velocity defined in Eq. 4.16. Intrinsic and extrinsic device level scattering mechanics could also influence transit time. However, a cascade of inverters or some other logic elements in series, the load capacitance between each logic stage typically dominated by $C_{interconnect}$ would be responsible for the majority of the delay.

Beyond individual logical elements (ie., inverter, NAND, NOR), an important CMOS circuit design

parameter is fan-out, which estimates the number of logic stages or CMOS gates that can be consecutively driven before signal attenuation is no longer tolerable. Past the maximum fan-out a repeater or amplifier is necessary to drive subsequent logic stages in a circuit. The maximum fan-out scales proportionally with propagation delay; therefore circuits designed for low frequency applications have a larger maximum fan-out compared to circuits designed for higher frequency applications. If graphene is to indeed follow the MOSFET and CMOS paradigm fan-out would be important circuit design trade-off to consider and a topic we will discuss in an upcoming work.

Chapter 9

Appendix iii
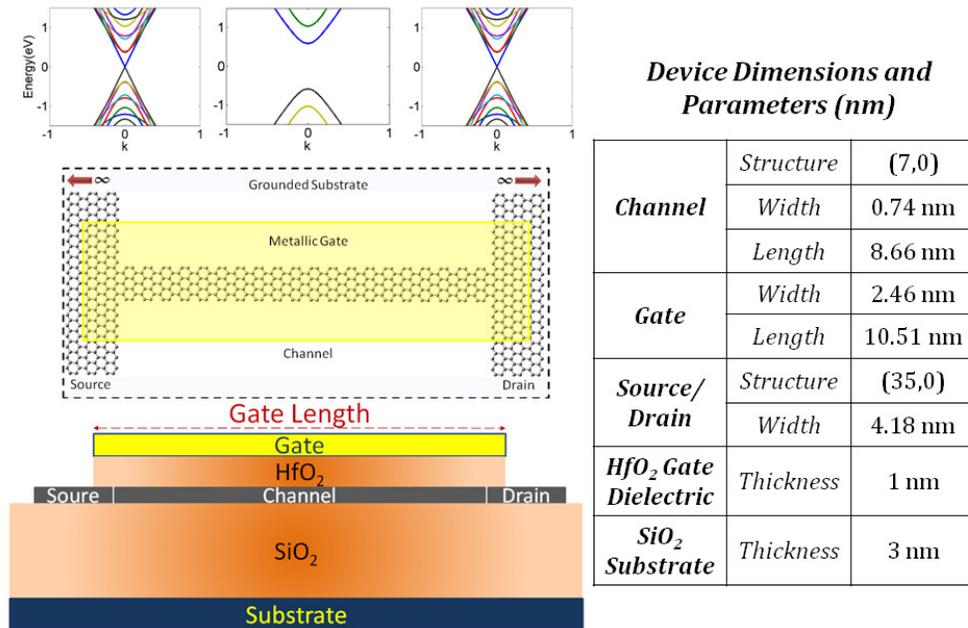
Figure 9.1: WNW dual gated all graphene device, showing local $E - k$s (top), top view (center) and side view (bottom) with the device parameters listed.

The structure of an imagined WNW graphene nanoribbon field-effect transistor (GNRFET) is shown in Fig. 9.1. The wide regions are metallic and the narrow ones semiconducting. There are planar gates both at the top and the bottom, the top ones for gating and the bottom ones for electrostatic 'doping' (see figure later for inverters). Let us first discuss how we simulate the I-V of one of these WNW devices.

65

Solving quantum transport and electrostatic equations

The calculations we will show couples a suitable bandstructure/density of states for the graphene channel with full 3D Poisson's equation for the electrostatics and the Non-Equilibrium Green's Function (NEGF) formulation for quantum transport [43]. The wider contact regions are captured recursively by computing their surface Green's functions $g_{1,2}(E)$. The corresponding energy-dependent self-energy matrices $\Sigma_{1,2}(E) = \tau_{1,2}\, g_{1,2}\, \tau_{1,2}^{\dagger}$ project the contact states onto the channel subspace, where the τ matrices capture the bonding between the contact and channel regions. In order to capture the interfacial chemistry properly, we extend the device a couple of layers into the wider regions and calculated its Hamiltonian H matrix. The Coulomb matrix U is computed using the method-of-moments, described below [75].

From the above matrices, the retarded Green's function $G = (ES - H - U - \Sigma_1 - \Sigma_2)^{-1}$ is computed, and thence the charge density matrix $\rho = \int dE\, G \Sigma^{in} G^{\dagger}/2\pi$, whose trace gives us the total charge. $\Sigma^{in} = (\Gamma_1 f_1 + \Gamma_2 f_2)$ in the simple limit where the only scattering arises at the contact channel interface. In the previous equation, $\Gamma_{1,2} = i(\Sigma_{1,2} - \Sigma_{1,2}^{\dagger})$ give the contact broadenings (the matrix analogues of the injection rates $\gamma_{1,2}$ introduced in section 18.2.1), while $f_{1,2}(E) = 1/[1 + e^{(E-\mu_{1,2})/k_B T}]$ represent the contact Fermi-Dirac distributions, with $\mu_{1,2}$ being the bias-separated electrochemical potentials or quasi-Fermi energies in the contacts [44]. The charge density matrix is then used to recompute the Coulomb matrix U self-consistently through Poisson's equation. Finally, the converged Green's function is used to compute the current $I = (2q/h) \int dE\, T(E)[f_1(E) - f_2(E)]$, where the transmission $T(E) = trace(\Gamma_1 G \Gamma_2 G^{\dagger})$ [43].

Let us now get into a few details on the 3D Poisson equation we solve, using the method of moments (MOM) numerically. MOM captures the channel potential by setting up grid points on the individual device atoms with a specific charge density δn_D, and on the contact atoms with a specific applied voltage ϕ_C [75]. Using the notations 'C' for Contact and 'D' for Device, we get

$$\phi_d = \underbrace{(U_{dC} U_{CC}^{-1})\phi_C}_{Laplace} + \underbrace{(U_{dd} - U_{dC} U_{CC}^{-1} U_{Cd})}_{Single\ Electron\ Charging\ Energy} \Delta n_d \tag{9.1}$$

where we imply vector notations for the potentials ϕ and matrix notations for the Coulomb kernels U. Δn_d is calculated relative to its neutrality value N_0 by tracing over ρ above, while N_0 is calculated analogously, while grounding all the contact potentials (this would depend on the workfunction of the contacts, as in MOS electrostatics). The matrix elements in U need to be computed with the correct dielectric constants. Let us describe it in the simpler case with a dielectric constant κ for the top gate and a dielectric constant unity for

the bottom (trivially generalized to multiple dielectrics). Using the method of images,

$$
\begin{aligned}
U(\vec{r}_1, \vec{r}_2) &= \frac{q}{4\pi\epsilon_0\epsilon_1}\left[\frac{1}{|\vec{r}_1-\vec{r}_2|} - \left(\frac{\epsilon_2-\epsilon_1}{\epsilon_2+\epsilon_1}\right)\frac{1}{|\vec{r}_1-\vec{r}_2'|}\right] \quad \text{(in the same medium)} \\
&= \frac{q}{2\pi\epsilon_0(\epsilon_1+\epsilon_2)|\vec{r}_1-\vec{r}_2|} \quad \text{(in different media)}
\end{aligned}
\tag{9.2}
$$

where $\vec{r}_2'$ is the image of the charge at $\vec{r}_2$ [76] [77]. A tricky point is to avoid the infinities at the onsite locations, for instance, when $x_1 = x_2$ and $y_1 = y_2$. We can avoid these using the Mataga-Nishimoto approximation, where we replace terms like $1/|\vec{rs}_1 - \vec{r}_2|$ with an atomistic correction $1/\sqrt{|\vec{r}_1-\vec{r}_2|^2+a^2}$, with the cut off parameter a adjusted to represent the correct onsite Coulomb (Hubbard) charging energy given by the difference between the atomic ionization energy and the electron affinity [78].

Let us now discuss the observed electrostatic characteristics in the WNW device, which explains the geometric advantages of this particular structure.

Improved electrostatics in 2-D
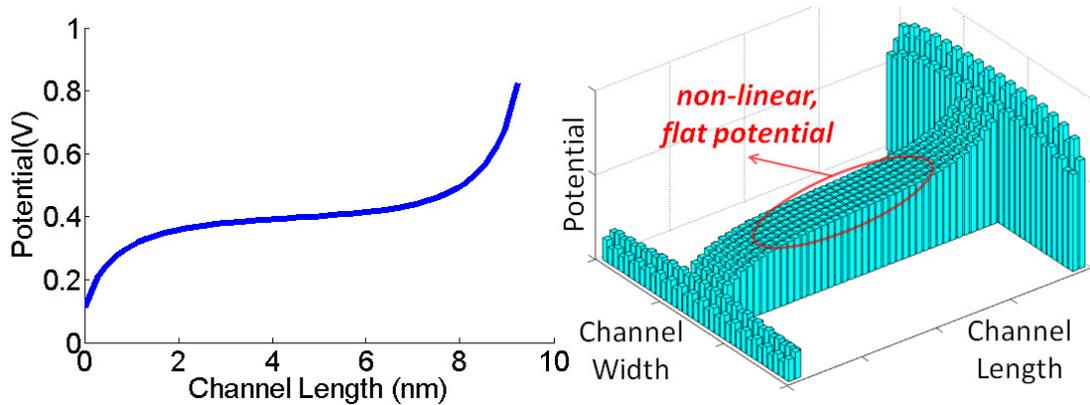
Figure 9.2: (Left) The two-terminal potential shows the vanishing fields near the channel, implying the superior gate control and the improved short-channel effects with the 2D contacts. (Right) The 3D potential shows the non-linear flat potential in the middle of the channel.

We simulate a device patterned monolithically from a two-dimensional sheet of graphene with a wide dilution of widths from the source and drain contacts to the active channel region. Simulated WNW (35-7-35) GNRFETs compose of (7,0) armchair graphene nanoribbon (GNR) narrow regions for the channel and (35,0) armchair GNR regions for the contact and interconnect regions. A metallic gate is placed on top of the channel region, while a wide grounded substrate is placed at the bottom of the channel. For calibration with the conventional CMOS technologies, the unique two-dimensional (2D) contacts of the GNRFET are replaced

with 3D bulk metal contacts (whose surfaces act as parallel plate capacitors) for the same device, gate, and dielectric geometry.

A particular advantage of the WNW structure is the low capacitance of the 2D source drain contacts [72]. In a conventional MOSFET, the gate electrode needs to compete electrostatically with the source and drain for control of the channel charge. Indeed, a majority of developments in transistor technology over the last few decades have concentrated on making the field lines gate controlled rather than source/drain controlled. This is becoming harder with aggressive size scaling. The 2D source and drain contacts with a top gate makes the S/D capacitances lower, as they can only influence the channel through their fringing fields. Note that a 2D side gate geometry, as advocated in many device designs, would eliminate that electrostatic advantage, as the gate needs to compete with the S/D electrodes.
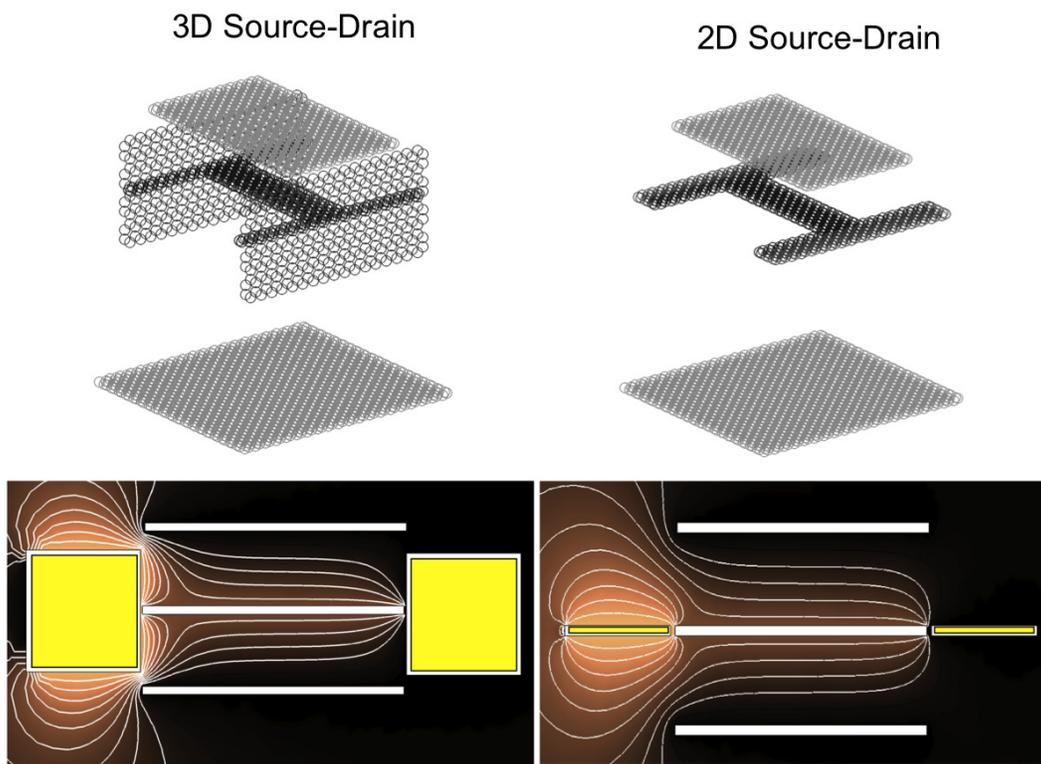
Figure 9.3: Comparison of planar source drain vs 3D source-drain. Denser field lines on the channel from the 3D contacts correlate to stronger source coupling and DIBL. For the given material and geometrical parameters listed in Fig. 9.1, the C_G/C_D ratios are 4.95 and 5.80 respectively. Top and bottom gates were grounded while the source was simulated with a potential of 0.3V and conducting channel had a potential of 0.1V

As the channel length gets shorter with the aggressively scaled technologies, the 3D contacts start to influence the channel potential as their surfaces act as parallel capacitor plates flanked by the insulator at the

top and bottom. In the case of the monolithically patterned 2D GNR contacts, the charges on the contact surfaces are line charges so that the applied source-drain field decays into the channel, creating a *non-linear channel potential even in the absence of a gate* (Fig. 9.2, left). Moving on to a three-terminal, dually gated structure, Fig. 9.2 (right), shows that the gate contact holds the channel potential flat against the action of the drain, thereby reducing short-channel effects.
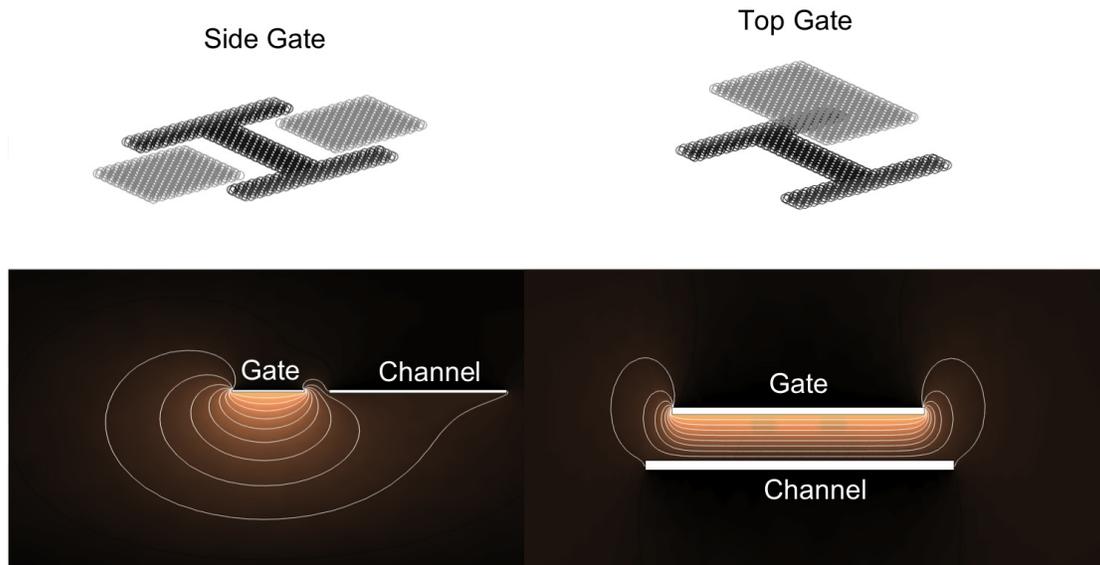
Figure 9.4: Comparison of top vs side gate. Denser field lines from the top gate ensure better gate control which is reflected by larger gate capacitance. For the given material and geometrical parameters listed in Fig. 9.1, the C_G/C_D ratios are 5.80 and 3.82 respectively. Gates were biased at 0.4V , while the conducting channel had a potential of 0.1V.

Figs. 9.3, 9.4 show that for the same channel geometry, the top gate with 2D side contacts has the largest capacitance, followed by the top gate with 3D side contacts and finally the lowest gate to drain capacitance ratio is obtained when all electrodes are co-planar. The corresponding field line diagrams are also shown in these figures. Note also that in addition to the source, drain and dual gate electrodes, one needs to worry about the quantum capacitances, which are automatically included from our density matrix calculations that enter Poisson's equation.

The capacitance ratio can be extracted by plotting the channel transmission (T) for two scenarios: maintaining a constant drain voltage (V_d) while sweeping gate voltage (V_g) , and analogously, maintaining a constant V_g while sweeping V_d. Sweeping the V_g creates a larger energy shift in the transmission of the GNRFET channel than the sweeping of the V_d. From the shifting rates of these transmissions and the charge density calculations from the MOM, we can extract the capacitance values of the contacts. With shifts in transmission plots, we once again find that 2D contacts indeed help the gate exercise superior control over

the channel [72].

We will now explore the effect of the improved short channel effect on the computed I-V characteristics.

9.0.1   Three terminal I-Vs

The computed three terminal I-Vs (Fig. 9.5) show excellent short channel effects, at least over a small voltage range given by the bandgap. Plotted vs gate voltage, the current shows excellent saturation characteristics with a large output impedance. Plotted vs gate voltage, the current shows little drain bias dependence (so-called drain induced barrier lowering or DIBL). Taken together, the curves signify that the device electrostatics in the geometry is nearly ideal, making the outputs relatively robust with process variations. It is interesting to note that instead of enhancing the gate capacitance as in regular CMOS devices, the trick in WNW devices has been to reduce the source and drain capacitances in comparison.

The simulations results of the model in the Fig. 9.5 demonstrate a Subthreshold Swing (SS) of 84.3 mV/dec and a Drain-Induced Barrier Lowering (DIBL) of 24 mV/V. We note that unless otherwise specified all simulations refer to material and geometrical parameters shown in Fig. 9.1 The value of the DIBL and the SS can be further improved by increasing the length of the channel (currently 1:8.6 ratio of HfO$_2$ thickness to channel length). These values calculated are better (smaller) than the estimated values of DIBL = 122 mV/V and the SS = 90 mV/dec for the double gate, 10 nm scaled Si MOSFETs [79]. Also in addition to showing improved short-channel effects, the GNRFET structure with the 2D contacts also shows controlled switching behavior. The on-current (Ion) of the system equals to 2670.62 $A/\mu m$ with the off-current (Ioff) set at 4.07 $A/\mu m$; thus giving a Ion/Ioff ratio of 656. The ON-OFF ratio, however, ends up being modest, and is a critical challenge in GNRFETs, especially in the light of its seemingly inverse relation with the charge mobility (section 18.2.2).

With the scaling of the channel length, the short-channel effects started to have a huge influence on the device parameters such as the DIBL and SS. As the channel lengths get shorter, the DIBL and SS of the device increases due to the C_g/C_d ratio of decreasing with length. The line charges with the 2D contacts endow the gate with more control over the channel and interface states compared to the 3D contacts by lowering the drain capacitance.

9.0.2   Pinning vs. Quasi-Ohmic contacts

In today's semiconductors, Ohmic contacts are a desired to help achieve linear and asymmetric I-V characteristics. The potential profile inside the channel can be influenced by increasing the drain-source voltage (V_{ds}) or the gate voltage (V_{gs}). For carbon nanotubes, this has been a particular challenge, as the metal
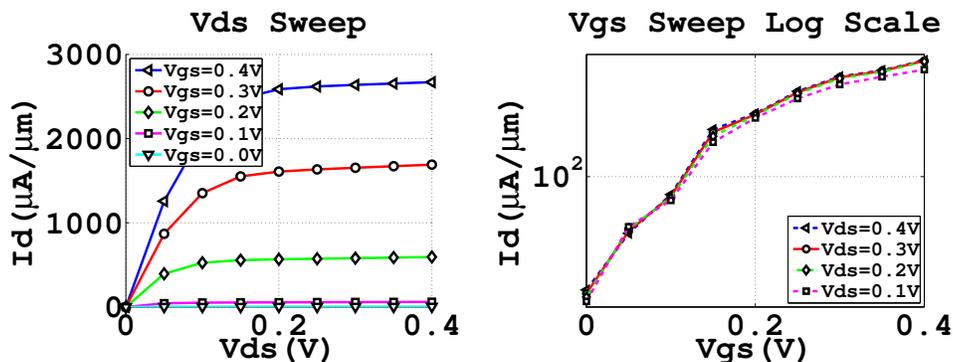
Figure 9.5: I-V curves for a n-type GNRFET confined to create a large bandgap (in this case, a (7,0) armchair GNR with a bandgap nearly 1 eV). Such an extreme geometry postpones the onset of band-to-band tunneling. More importantly, the point of the I-V is to show the effect of better electrostatics which is independent of bandgap issues – resulting in a high current saturation, low DIBL and SS.

carbon bonds at the ends have predominantly created Schottky barriers [80]. In our WNW geometries, since the bulk metal contacts are relegated to the ends of the device array, the bonding configuration near the wide-narrow interfaces are controlled by C-C covalent bonding. As our simulations show, this seems to promote a quasi-Ohmic behavior. The better bonding increases the decay lengths of the corresponding metal-induced gap states (MIGS) entering from the wide regions. The partial delocalization reduces the single-electron charging energy (that enters through our MOM treatment), thus making it harder for the contact regions to pin the Fermi energy and reducing the effectiveness of the Schottky barrier.

Schottky barrier FETs behave qualitatively different from MOSFETs In the latter, an applied gate bias reduces the channel potential and controls the thermionic emission over the voltage-dependent interfacial barrier. In the former, the gate reduces the thickness of the Schottky barrier and controls the tunneling of electrons through a voltage-independent, pinned barrier height. The question is what the potential profile looks like in the channel, and whether the contact MIGS are effective in pinning this potential adequately.

As seen in the Fig. 9.6, the lowering of the potential throughout the entire graphene channel region with applied gate bias is a characteristic of the regular ohmic contact FETs rather than the Schottky barrier FETs, whose potentials would otherwise be pinned to the midgap by the charging of the interfacial states [81].

The MIGS due to the tail ends of the metal states in the contacts, leak in the semiconductor. Even with this 2D contact geometry, the MIGS will be present because of the contact-channel interfaces [72]. Our WNW all-graphene structure can filter these quickly decaying states, resulting in no significant contribution to the electron transmission. In the case of our device with the channel length of 8.66 nm, the MIGS do not travel all the way from source to the drain, but only extend approximately 0.7 nm into the semiconducting channel (Note a typo in one of our earlier papers, where we wrongly quoted this as 0.07 nm) [72]. The decay
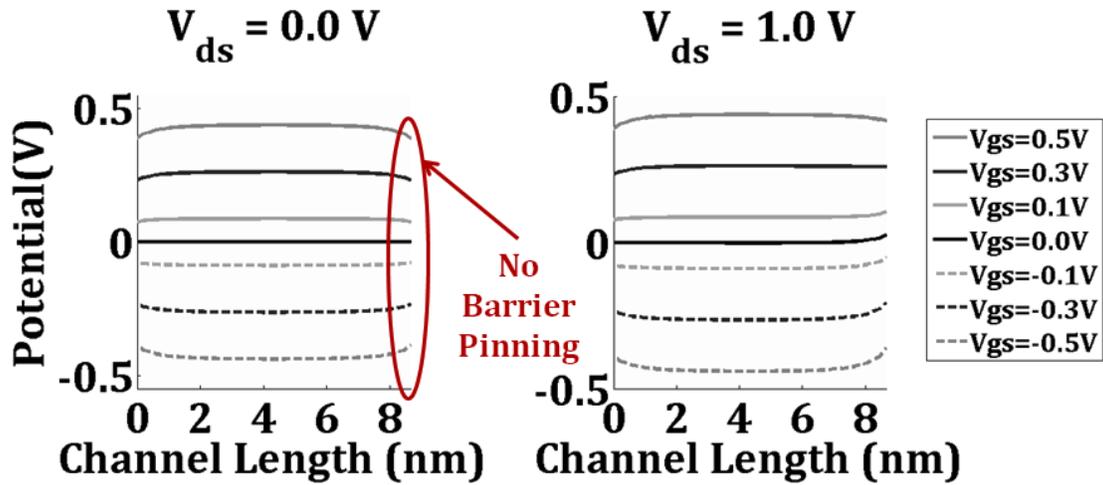
Figure 9.6: At $V_{ds} = 0.0$V and $V_{ds} = 1.0$V, variation of channel potential with different gate voltages shows no barrier pinning at the contacts, implying Ohmic contacts.

length of these MIGS can be calculated by plotting the wavefunction of the channel electrons at specific energies, as well as by evaluating the complex E-k diagram. The intensity of these MIGS at a given distance x can be calculated by using the equation $I_0 * e^{(-x/2\lambda)}$, where I_0 is the intensity of MIGS at the interface and λ is the decay length.

Note that issues similar to those discussed here have been discussed in the context of pentacene molecules with CNT contacts. While CNTs would offer even better 1D electrostatic gains, a trade-off arises with the increasing series resistance in CNTs due to a paucity of modes. For GNR source/drain analogously, we will need to imagine wide blocks simultaneously contacting many GNR devices, so that the contact resistance is minimized by extending its width.

Chapter 10

Appendix iv.

10.1 Charge Impurity Scattering fo Graphene

$$T_{i \to f} = \frac{2\pi}{\hbar} \sum \left| \left\langle \vec{k} \, |qV(\vec{r})| \, \vec{k'} \right\rangle \right|^2 \delta(E - E_k) \, n_i$$

Impurity Potential

$$V(\vec{r}) = \frac{Zq^2 e^{-\kappa r}}{4\pi\epsilon_o \epsilon_s r}$$

Working out Matrix Element : $\left| \left\langle \vec{k} \, |qV(\vec{r})| \, \vec{k'} \right\rangle \right|$

$$= \quad \frac{Zq^2}{4\pi\epsilon_o\epsilon_s} \frac{1}{2} \int \begin{pmatrix} 1 \\ 1 \end{pmatrix}^{\dagger} e^{-i\vec{k'}\vec{r}} \frac{e^{-\kappa r}}{r} \begin{pmatrix} 1 \\ e^{i\phi} \end{pmatrix} e^{i\vec{k}\vec{r}} d^2\vec{r}$$

$$= \quad \frac{Zq^2}{4\pi\epsilon_o\epsilon_s} \frac{\left(1 + e^{i\phi}\right)}{2} \int_0^{2\pi} \int_0^{\infty} \frac{e^{-\kappa r}}{r} e^{i(\vec{\Delta k} \, cos(\theta))\vec{r}} r \, dr \, d\theta$$

Using (G+R): 3.915(2) $n = 0$:

$$\int_0^{\infty} dr \, e^{-\kappa r} \int_0^{2\pi} e^{i\Delta kr \, cos(\theta)} d\theta = \int_0^{\infty} dr \, e^{-\kappa r} \left\{ 2\pi J_0(\Delta kr) \right\}$$

Using (G+R): 6.611 (1) $\nu = 0$

$$\frac{Zq^2}{4\pi\epsilon_o\epsilon_s}\int_0^\infty dr\, e^{-\kappa r}\left\{2\pi J_0(\Delta kr)\right\} = \frac{Zq^2}{2\epsilon_o\epsilon_s\sqrt{\kappa^2 + \Delta k^2}}$$

Note that :

$$\left|\frac{1 + e^{i\phi}}{2}\right|^2 = \frac{1}{4}(1 + e^{-i\phi})(1 + e^{i\phi}) = \frac{1 + cos(\phi)}{2}$$

Therefore the matrix element is:

$$\left|\left\langle \vec{k}\,|qV(\vec{r})|\,\vec{k'}\right\rangle\right|^2 = \frac{Z^2q^4(1 + cos(\phi))}{8\epsilon_o^2\epsilon_s^2\left[\kappa^2 + \Delta k^2\right]}$$

Working out the Scattering Rate

And the Scattering rate is:

$$T_{i\to f} = \frac{2\pi}{\hbar}\sum \frac{Z^2q^4(1 + cos(\phi))(1 - cos(\phi))}{8\epsilon_o^2\epsilon_s^2\left[\kappa^2 + \Delta k^2\right]}\delta(E - E_k)\,n_i$$

Note $\Delta k = |\vec{k'} - \vec{k}|$ is the scattering wavevector :

$$|\vec{k'} - \vec{k}|^2 = 2k^2(1 - cos(\phi)) = 4k^2 sin^2(\phi/2)$$

$$T_{i\to f} = \frac{2\pi S}{\hbar}\int\int_0^{2\pi}\frac{kdkd\phi}{(2\pi)^2}\frac{Z^2q^4(1 + cos(\phi))(1 - cos(\phi))}{8\epsilon_o^2\epsilon_s^2\left[\kappa^2 + 4k^2sin^2(\phi/2)\right]}\delta(E - E_k)$$

Take the integral with respect to ϕ first we get:

$$T_{i\to f} = \frac{2\pi S}{\hbar}\frac{Z^2q^4}{4\pi^2 8\epsilon_o^2\epsilon_s^2}\int kdk\delta(E-E_k)\left[\frac{\phi(\kappa^2 + 2k^2)}{4k^2} + \frac{sin(\phi)}{2k^2} + \frac{\sqrt{\kappa^4 + 4k^2\kappa^2}}{2\kappa^2}tan^{-1}\left\{\frac{\sqrt{\kappa^2 + 4k^2}}{\kappa}tan(\frac{\phi}{2})\right\}\right]_0^{2\pi}$$

$$T_{i\to f} = \frac{2\pi S}{\hbar}\frac{Z^2q^4}{4\pi^2 8\epsilon_o^2\epsilon_s^2}\int kdk\delta(E - E_k)\left[\frac{\pi}{k^2} + \frac{\pi\kappa^2}{2k^4} - \frac{\pi\kappa^2}{2k^4}\sqrt{\frac{\kappa^2 + 4k^2}{\kappa^2}}\right]$$

Defining k and κ

If we assume no charge puddling then k and κ are defined as follows:

$$k(E) = \frac{|E|}{\hbar v_o}$$

For the screening term κ we use the Thomas Fermi wavevector [Ashcroft+Mermin pg.342]:

$$\kappa(E) = \frac{4\pi q^2}{\epsilon_o \epsilon_s} \frac{dn}{d\mu}$$

where $dn/d\mu$ is really just the density of states per unit area:

$$\kappa(E) = \frac{4\pi q^2}{\epsilon_o \epsilon_s} \frac{|E|}{\pi \hbar^2 v_o^2}$$

Assuming we have charge puddles:

The density of states per unit area in the presence of charge puddles is :

$$DOS = \frac{\sqrt{(E + \mu)^2 + \alpha \sigma_o^2}}{\pi \hbar^2 v_o^2}$$

μ is the mean potential and $\sigma = n_i \pi \hbar^2 v_o^2$ is the variance the potential distribution defined in terms of the number of impurities per unit area.

We can work backwards from :

$$k = \sqrt{4\pi n}$$

,where $n = \int DOS \, dE$. Using trigonometric substitution DOS can be expressed in terms of common trigonometric functions. Trigonometric identities help simplify the integral.

$$tan(\Theta) = \frac{E + \mu}{\sqrt{\alpha \sigma_o}}, \quad sec^2(\Theta) d\Theta \, \frac{dE}{\sqrt{\alpha \sigma_o}}$$

$$\int \frac{\sqrt{(E + \mu)^2 + \alpha \sigma_o^2}}{\pi \hbar^2 v^2} dE = \frac{\sqrt{\alpha \sigma_o}}{\pi \hbar^2 v^2} \int \sqrt{tan^2(\Theta) + 1} \sqrt{\alpha \sigma_o} sec^2(\Theta) \, d\Theta$$

$$\frac{\alpha \sigma_o^2}{\pi \hbar^2 v^2} \int sec^3(\Theta) \, d\Theta = \frac{\alpha \sigma_o^2}{\pi \hbar^2 v^2} \{sec(\Theta) tan(\Theta) + ln\,|sec(\Theta) + tan(\Theta)| + C\}$$

$$sec(\Theta) = \frac{\sqrt{(E+\mu)^2 + \alpha\sigma_o^2}}{\sqrt{\alpha}\sigma_o}, \ tan(\Theta) = \frac{E+\mu}{\sqrt{\alpha}\sigma_o}$$

$$\frac{\alpha\sigma_o^2}{\pi\hbar^2 v^2} \{sec(\Theta)tan(\Theta) + ln\,|sec(\Theta) + tan(\Theta)| + C\}$$

$$\frac{\alpha\sigma_o^2}{\pi\hbar^2 v^2} \left\{ \frac{(E+\mu)\sqrt{(E+\mu)^2 + \alpha\sigma_o^2}}{\alpha\sigma_o^2} + ln\,\left|\frac{\sqrt{(E+\mu)^2 + \alpha\sigma_o^2} + E + \mu}{\sqrt{\alpha}\sigma_o}\right| + C \right\}$$

Therefore:

$$k(E) = \sqrt{\frac{4\pi\alpha\sigma_o^2}{\pi\hbar^2 v^2} \left\{ \frac{(E+\mu)\sqrt{(E+\mu)^2 + \alpha\sigma_o^2}}{\alpha\sigma_o^2} + ln\,\left|\frac{\sqrt{(E+\mu)^2 + \alpha\sigma_o^2} + E + \mu}{\sqrt{\alpha}\sigma_o}\right| + C \right\}}$$

and

$$\kappa(E) = \frac{4\pi q^2}{\epsilon_o\epsilon_s} \frac{\sqrt{(E+\mu)^2 + \alpha\sigma_o^2}}{\pi\hbar^2 v_o^2}$$

Note that the integral with respect to k with the delta function gives us the density of states which we already know since we actually worked backwards from a known density of states :

$$\int kdk \left[\frac{\pi}{k^2} + \frac{\pi\kappa^2}{2k^4} - \frac{\pi\kappa^2}{2k^4}\sqrt{\frac{\kappa^2 + 4k^2}{\kappa^2}}\right]\delta(E - E_k) = |M_{k,k'}(E)|^2 \ DOS(E)$$

Putting everything together and expressing the scattering rate in terms of energy I get:

$$T_{i\to f} = \frac{2\pi}{\hbar}\frac{Z^2 q^2}{4\pi^2 8\epsilon_o^2\epsilon_s^2}\int kdk\delta(E - E_k)\left[\frac{\pi}{k^2} + \frac{\pi\kappa^2}{2k^4} - \frac{\pi\kappa^2}{2k^4}\sqrt{\frac{\kappa^2 + 4k^2}{\kappa^2}}\right]n_i$$

$$T_{i\to f} = \frac{2\pi}{\hbar}\frac{Z^2 q^4}{4\pi^2 8\epsilon_o^2\epsilon_s^2}\left[\frac{\pi}{k(E)^2} + \frac{\pi\kappa(E)^2}{2k(E)^4} - \frac{\pi\kappa(E)^2}{2k(E)^4}\sqrt{\frac{\kappa(E)^2 + 4k(E)^2}{\kappa(E)^2}}\right]\left[\frac{\sqrt{(E+\mu)^2 + \alpha\sigma_o^2}}{\hbar^2 v_o^2}\right]n_i$$

Typically n_i is between $3 - 15 \times 10^{11}cm^{-2}$ [82]and standard deviation σ_o of potential distributions is around 50meV[13].

Useful Integrals : Gradshteyn and Ryzhik (G+R)

3.915 (2) (5)

$$\int_0^\pi e^{i\beta cos(x)} cos(nx) dx = i^n \pi J_n(\beta)$$

6.611 (1) pg.729

$$\int_0^\infty e^{-\alpha x} J_\nu(\beta x) dx = \frac{\beta^{-\nu} \left[\sqrt{\alpha^2 + \beta^2} - \alpha \right]^\nu}{\sqrt{\alpha^2 + \beta^2}}, \ Re(\nu) > -1$$
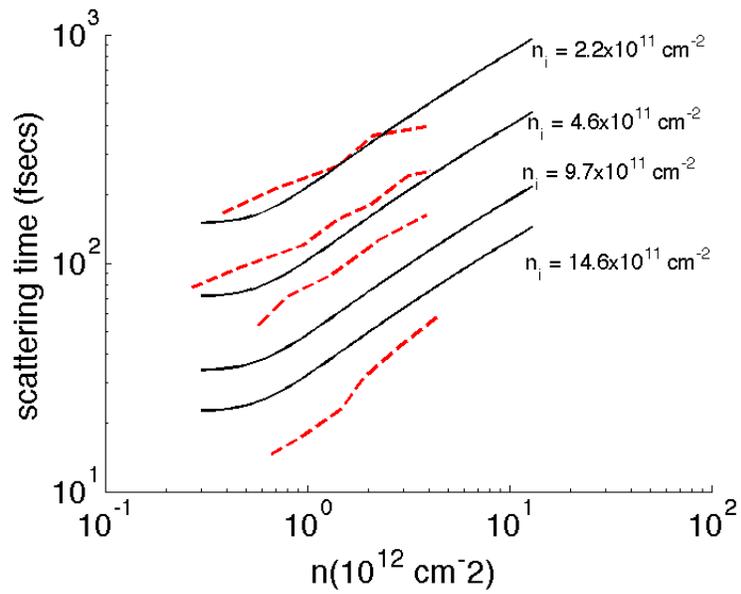
Figure 10.1: Black solid lines come from the analytical expression for scattering time derived in this write-up. The red dashed line is from Ref.([13]) where scattering rate is measured. Impurity density was set between $2.2 \times 10^{11} cm^{-2}$ and $14.6 \times 10^{11} cm^{-2}$.

Chapter 11

Appendix iv.

11.1 Contact Resistance

$$V(x + \Delta x) - V(x) = I(x)R_2 = \frac{I(x)R_s}{W}\Delta x$$

$$I(x + \Delta x) - I(x) = \frac{V(x)}{R_1} = \frac{V(x)W}{\rho_c}\Delta x$$

$$\frac{dV(x)}{dx} = \frac{I(x)R_S}{W} \quad , \quad \frac{dI(x)}{dx} = \frac{V(x)W}{\rho_c}$$

$$\frac{d^2V(x)}{dx^2} = \frac{dI(x)}{dx}\frac{R_s}{W} = \frac{V(x)R_s}{\rho_c}$$

$$\frac{d^2V(x)}{dx^2} - V(x)\frac{R_s}{\rho_c} = 0 \quad \rightarrow \quad V(x) = c_1 e^{\sqrt{R_s/\rho_c}x} + c_2 e^{-\sqrt{R_s/\rho_c}x}$$

conditions:

$$V(0) = V_o = c_1 + c_2$$

$$V(x) = \frac{V_o}{2}cosh(\sqrt{R_s/\rho_c}x) \quad , \quad I(x) = \frac{W V_o}{2R_s}\sqrt{\frac{R_s}{\rho_c}}sinh(\sqrt{R_s/\rho_c}x)$$

therefore the resistance of one contact is:

$$R_c(x) = \frac{\sqrt{R_s\rho_c}}{W} cotanh(\sqrt{R_s/\rho_c}x)$$

Bibliography

[1] H. Iwai. Roadmap for 22nm and beyond (invited paper). *Microelectronic Engineering*, 86:1520 – 1528, 2009.

[2] Yuanbo Zhang, Victor W. Brar, Caglar Girit, Alex Zetti, and Michael F. Crommie. Origin of spatial charge inhomogeneity in graphene. *Nature Physics*, 5:722–726, 2009.

[3] Young-Woo Son, Marvin L. Cohen, and Steven G. Louie. Energy gaps in graphene nanoribbons. *PHYS.REV.LETT.*, 97:216803, 2006.

[4] Aaron Bostwick, Taisuke Ohta, Thomas Seyller, Karsten Horn, Eli, and Rotenberg. Quasiparticle dynamics in graphene. *Nature Physics*, 3:36–40, 2007.

[5] Melinda Y. Han, Barbaros Özyilmaz, Yuanbo Zhang, and Philip Kim. Energy band-gap engineering of graphene nanoribbons. *Phys. Rev. Lett.*, 98:206805, May 2007.

[6] S.M. Sze and Kwok K. Ng. *Physics of Semiconductor Devices*. Wiley, 2007.

[7] D. Unluer, F. Tseng, and A.W. Ghosh. Physics-based gnrfet compact model for digital circuit design. In *Semiconductor Device Research Symposium (ISDRS), 2011 International*, pages 1–2, 2011.

[8] Jiamin Xue, Javier Sanchez-Yamagishi, Danny Bulmash, Philippe Jacquod, Aparna Deshpande, K Watanabe, T Taniguchi, Pablo Jarillo-Herrero, and Brian J LeRoy. Scanning tunnelling microscopy and spectroscopy of ultra-flat graphene on hexagonal boron nitride. *Nature materials*, 10(4):282–285, 2011.

[9] Zhen Yao, Charles L. Kane, and Cees Dekker. High-field electrical transport in single-wall carbon nanotubes. *Phys. Rev. Lett.*, 84:2941–2944, Mar 2000.

[10] Xiaolin Li, Xinran Wang, Li Zhang, Sangwon Lee, and Hongjie Dai. Chemically derived, ultrasmooth graphene nanoribbon semiconductors. *Science*, 319(5867):1229–1232, 2008.

[11] M. Katagiri, Y. Yamazaki, N. Sakuma, M. Suzuki, T. Sakai, M. Wada, N. Nakamura, N. Matsunaga, S. Sato, M. Nihei, and Y. Awano. Fabrication of 70-nm-diameter carbon nanotube via interconnects by remote plasma-enhanced chemical vapor deposition and their electrical properties. In *Interconnect Technology Conference, 2009. IITC 2009. IEEE International*, pages 44 –46, 2009.

[12] Y. Awano. Carbon nanotube lsi via interconnects. The Fullerenes and Nanotubes Research Society, Mar. 2008.

[13] Y.-W. Tan, Y. Zhang, K. Bolotin, Y. Zhao, S. Adam, E. H. Hwang, S. Das Sarma, H. L. Stormer, and P. Kim. Measurement of scattering rate and minimum conductivity in graphene. *Phys. Rev. Lett.*, 99:246803, Dec 2007.

[14] K.I. Bolotin, K.J. Sikes, Z. Jiang, M. Klima, G. Fudenberg, J. Hone, P. Kim, and H.L. Stormer. Ultrahigh electron mobility in suspended graphene. *Solid State Communications*, 146:351 – 355, 2008.

[15] J.S. Moon, D. Curtis, S. Bui, M. Hu, D.K. Gaskill, J.L. Tedesco, P. Asbeck, G.G. Jernigan, B.L. Van-Mil, R.L. Myers-Ward, C.R. Eddy, P.M. Campbell, and X. Weng. Top-gated epitaxial graphene fets on si-face sic wafers with a peak transconductance of 600 ms/mm. *Electron Device Letters, IEEE*, 31(4):260 –262, april 2010.

[16] C.R. Dean, A.F. Young, I. Meric, C. Lee, L. Wang, S. Sorgenfrei, K. Watanabe, T. Taniguchi, P. Kim, K.L. Shepard, and J. Hone. Boron nitride substrates for high-quality graphene electronics. *Nature Nanotechnology*, 5:722–236, 2010.

[17] Vernica Barone, Oded Hod, and Gustavo E. Scuseria. Electronic structure and stability of semiconducting graphene nanoribbons. *Nano Letters*, 6(12):2748–2754, 2006. PMID: 17163699.

[18] Levente Tapaszto, Gergely Dobrik, Philippe Lambin, and Laszlo P Biro. Tailoring the atomic structure of graphene nanoribbons by scanning tunnelling microscope lithography. *Nature Nanotechnology*, 3(7):397–401, 2008.

[19] F. Miao, S. Wijeratne, Y. Zhang, U. C. Coskun, W. Bao, and C. N. Lau. Phase-coherent transport in graphene quantum billiards. *Science*, 317(5844):1530–1533, 2007.

[20] J. Tworzydło, B. Trauzettel, M. Titov, A. Rycerz, and C. W. J. Beenakker. Sub-poissonian shot noise in graphene. *Phys. Rev. Lett.*, 96:246802, Jun 2006.

[21] Shaffique Adam, E H Hwang, Victor Galitski, and S Das Sarma. A self-consistent theory for graphene transport. *Proceedings of the National Academy of Sciences of the United States of America*, 104(47):18392–7, 2007.

[22] S. Das Sarma, Shaffique Adam, E. H. Hwang, and Enrico Rossi. Electronic transport in two-dimensional graphene. *Rev. Mod. Phys.*, 83:407–470, May 2011.

[23] J H Chen, C Jang, M S Fuhrer, E D Williams, and M Ishigami. Charged impurity scattering in graphene. *Nature Physics*, 4(5):19, 2007.

[24] Yang Sui, Tony Low, Mark Lundstrom, and Joerg Appenzeller. Signatures of disorder in the minimum conductivity of graphene. *Nano Letters*, 11(3):1319–1322, 2011.

[25] F. Tseng, D. Unluer, K. Holcomb, M. R. Stan, and A. W. Ghosh. Diluted chirality dependence in edge rough graphene nanoribbon field-effect transistors. *Applied Physics Letters*, 94(22):223112, 2009.

[26] Sheneve Z. Butler, Shawna M. Hollen, Linyou Cao, Yi Cui, Jay A. Gupta, Humberto R. Gutirrez, Tony F. Heinz, Seung Sae Hong, Jiaxing Huang, Ariel F. Ismach, Ezekiel Johnston-Halperin, Masaru Kuno, Vladimir V. Plashnitsa, Richard D. Robinson, Rodney S. Ruoff, Sayeef Salahuddin, Jie Shan, Li Shi, Michael G. Spencer, Mauricio Terrones, Wolfgang Windl, and Joshua E. Goldberger. Progress, challenges, and opportunities in two-dimensional materials beyond graphene. *ACS Nano*, 7(4):2898–2926, 2013.

[27] Y.-M. Lin, C. Dimitrakopoulos, K. A. Jenkins, D. B. Farmer, H.-Y. Chiu, A. Grill, and Ph. Avouris. 100-ghz transistors from wafer-scale epitaxial graphene. *Science*, 327(5966):662, 2010.

[28] Lei Liao, Yung-Chen Lin, Mingqiang Bao, Rui Cheng, Jingwei Bai, Yuan Liu, Yongquan Qu, Kang L Wang, Yu Huang, and Xiangfeng Duan. High-speed graphene transistors with a self-aligned nanowire gate. *Nature*, 467(7313):305–308, 2010.

[29] Frank Schwierz. Graphene transistors. *Nature Nanotechnology*, 5:487–496, 2010.

[30] S. Y. Zhou, G.-H. Gweon, A. V. Fedorov, P. N. First, W. A. de Heer, D.-H. Lee, F. Guinea, A. H. Castro Neto, and A. Lanzara. Substrate-induced bandgap opening in epitaxial graphene. *Nat Mater*, 6:770–775, 2007.

[31] Qiuzi Li, E. H. Hwang, and S. Das Sarma. Disorder-induced temperature-dependent transport in graphene: Puddles, impurities, activation, and diffusion. *Phys. Rev. B*, 84:115442, Sep 2011.

[32] Philip F. Bagwell and Terry P. Orlando. Landauer's conductance formula and its generalization to finite voltages. *Phys. Rev. B*, 40:1456–1464, Jul 1989.

[33] D. Gunlycke and C. T. White. Tight-binding energy dispersions of armchair-edge graphene nanostrips. *Phys. Rev. B*, 77:115116, Mar 2008.

[34] D. Kienle, J. I. Cerda, and A. W. Ghosh. Extended h[u-umlaut]ckel theory for band structure, chemistry, and transport. i. carbon nanotubes. *Journal of Applied Physics*, 100(4):043714, 2006.

[35] J. Cerdá and F. Soria. Accurate and transferable extended hückel-type tight-binding parameters. *Phys. Rev. B*, 61:7965–7971, Mar 2000.

[36] Youngki Yoon, Dmitri E. Nikonov, and Saye. Role of phonon scattering in graphene nanoribbon transistors: Nonequilibrium greens function method with real space approach. *Appl. Phys. Lett.*, 98:203503, 2011.

[37] M. Farjam and H. Rafii-Tabar. Energy gap opening in submonolayer lithium on graphene: Local density functional and tight-binding calculations. *Phys. Rev. B*, 79:045417, Jan 2009.

[38] M. Cheli, G. Fiori, and G. Iannaccone. A semianalytical model of bilayer-graphene field-effect transistor. *Electron Devices, IEEE Transactions on*, 56(12):2979 –2986, dec. 2009.

[39] Edward McCann. Asymmetry gap in the electronic band structure of bilayer graphene. *Phys. Rev. B*, 74:161403, Oct 2006.

[40] E. H. Hwang and S. Das Sarma. Acoustic phonon scattering limited carrier mobility in two-dimensional extrinsic graphene. *Phys. Rev. B*, 77:115449, Mar 2008.

[41] K. M. Borysenko, J. T. Mullen, X. Li, Y. G. Semenov, J. M. Zavada, M. Buongiorno Nardelli, and K. W. Kim. Electron-phonon interactions in bilayer graphene. *Phys. Rev. B*, 83:161402, Apr 2011.

[42] Yang Zhao, Albert Liao, and E. Pop. Multiband mobility in semiconducting carbon nanotubes. *Electron Device Letters, IEEE*, 30(10):1078–1080, 2009.

[43] Supriyo Datta. *Quantum Transport: Atom to Transistor*. Cambridge University Press, 2005.

[44] S. Datta. The non-equilibrium green's function (negf) formalism: An elementary introduction. In *Electron Devices Meeting, 2002. IEDM '02. Digest. International*, pages 703 – 706, 2002.

[45] Charge transfer between epitaxial graphene and silicon carbide. *Applied Physics Letters*, 97(11):112109, 2010.

[46] P.N. First, W. A. deHeer, T. Seyller, C. Berger, J. A. Stroscio, and J.S. Moon. Epitaxial graphenes on silicon carbide. *MRS Bulletin*, 35, 2010.

[47] Inanc Meric, Cory Dean, Andrea Young, Jim Hone, Philip Kim, and Kenneth L Shepard. Graphene field-effect transistors based on boron nitride gate dielectrics. In *Electron Devices Meeting (IEDM), 2010 IEEE International*, pages 23–2. IEEE, 2010.

[48] Inanc Meric, Melinda Y. Han, Andrea F. Young, Barbaros Özyilmaz, Philip Kim, and Kenneth L. Shepard. Current saturation in zero-bandgap, top-gated graphene field effect transi. *Nature N*, 3:654–658, 2008.

[49] Ji-Yong Park, Sami Rosenblatt, Yuval Yaish, Vera Sazonova, Hande stnel, Stephan Braig, T. A. Arias, Piet W. Brouwer, and Paul L. McEuen. Electronphonon scattering in metallic single-walled carbon nanotubes. *Nano Letters*, 4(3):517–520, 2004.

[50] Ali Javey, Jing Guo, Magnus Paulsson, Qian Wang, David Mann, Mark Lundstrom, and Hongjie Dai. High-field quasiballistic transport in short carbon nanotubes. *Phys. Rev. Lett.*, 92:106804, Mar 2004.

[51] F. Tseng, D. Unluer, M.R. Stan, and A.W. Ghosh. *Graphene Nanoribbons: From Chemistry to Circuits*, chapter Graphene Nanoelectronics: Metrology, Synthesis, Properties and Applications, pages 555–585. Springer, 2012.

[52] Gianluca Giovannetti, Petr A. Khomyakov, Geert Brocks, Paul J. Kelly, and Jeroen van den Brink. Substrate-induced band gap in graphene on hexagonal boron nitride: *Ab initio* density functional calculations. *Phys. Rev. B*, 76:073103, Aug 2007.

[53] G. Fiori and G. Iannaccone. Simulation of graphene nanoribbon field-effect transistors. *Electron Device Letters, IEEE*, 28(8):760–762, 2007.

[54] G. Fiori and G. Iannaccone. On the possibility of tunable-gap bilayer graphene fet. *Electron Device Letters, IEEE*, 30(3):261–264, 2009.

[55] Stephen Y. Chou and Paul B. Fischer. Double 15-nm-wide metal gates 10 nm apart and 70 nm thick on gaas. volume 8, pages 1919–1922. AVS, 1990.

[56] X. Guan, D. Kim, K.C. Saraswat, and H.S.P. Wong. Complex band structures: from parabolic to elliptic approximation. *Electron Device Letters, IEEE*, 32(9):1296–1298, 2011.

[57] Dincer Unluer, Frank Tseng, Avik W Ghosh, and Mircea R Stan. Monolithically patterned wide-narrow-wide all-graphene devices. *IEEE Transactions on Nanotechnology*, 10(5):9, 2011.

[58] Avik W. Ghosh, Titash Rakshit, and Supriyo Datta. Gating of a molecular transistor: Electrostatic and conformational. *Nano Letters*, 4(4):565–568, 2004.

[59] A.W. Ghosh. *Comprehensive Semiconductor Science and Technology*, chapter Electronics with Molecules, pages 383–497. Elsevier, 2011.

[60] Hong-Yan Chen and J. Appenzeller. Complementary-type graphene inverters operating at room-temperature. In *Device Research Conference (DRC), 2011 69th Annual*, pages 33 –34, june 2011.

[61] Phaedon Avouris, Zhihong Chen, and Vasili Perebeinos. Carbon-based electronics. *Nature Nanotechnology*, 2:605–615, 2007.

[62] Walt A De Heer, Claire Berger, Xiaosong Wu, Phillip N First, Edward Conrad, Xuebin Li, Tianbo Li, Michael Sprinkle, Joanna Hass, Marcin L. Sadowski, Marek Potemski, and Gerard Martinez. Epitaxial graphene. *Solid State Communications*, 143:92–100, 2007.

[63] Alan M. Cassell, Nathan R. Franklin, Thomas W. Tombler, Emory M. Chan, Jie Han, and Hongjie Dai. Directed growth of free-standingsingle-walled carbon nanotubes. *J. Am. Chem. Soc.*, 121:7975–7976, 1999.

[64] Yuegang Zhang, Aileen Chang, Jien Cao, Qian Wang, Woong Kim, Yiming Li, Nathan Morris, Erhan Yenilmez, Jing Kong, and Hongjie Dai. Electric-field-directed growth of aligned single-walled carbon nanotubes. *App. Phys. Lett.*, 79, 19, 2001.

[65] Jing Wang, Eric Polizzi, Avik W Ghosh, Supriyo Datta, and Mark Lundstrom. Theoretical investigation of surface roughness scattering in silicon nanowire transistors. *Appl. Phys. Lett.*, 87:043101, 2005.

[66] Gengchiau Liang, Neophytos Neophytou, Dmitri E. Nikonov, and Mark S. Lundstrom. Performance projections for ballistic graphene nanroribbon field-effect transistor. *IEEE Trans. Electron Devices*, 54, 4:677–682, 2007.

[67] Y. Ouyang, Y. Yoon, and J. Guo. Scaling behaviors of graphene nanoribbon fets: A three-dimensional quantum simulation study. *IEEE Trans. Electron Devices*, 54,9:2223–2231, 2007.

[68] N. P. Guisinger, G.M. Rutter, J.N Crain, C. Heiliger, P.N. FIrst, and J.A. Stroscio. Atomic-scale investigation of graphene formation on 6h-sic(0001). *J. Vac. Sci. Technol. A*, 26, 932, 2008.

[69] D. Gunlycke and C. T. White. Tight-binding energy dispersions of armchair-edge graphene nanostrips. *Phys. Rev. B*, 77(11):115116, Mar 2008.

[70] Shah R. and T. M. Mohiuddin. Charge carrier mobility degradation in graphene sheetunder induced strain. arXiv:1008.4425v3.

[71] T. Fang, A. Konar, H. Xing, and D. Jena. Mobility in semiconducting graphene nanoribbons: Phonon, impurity, and edge roughness scattering. *Phys. Rev. B*, 78(20):205403, Nov 2008.

[72] D. Unluer, F. Tseng, A. Ghosh, and M. Stan. Monolithically patterned wide-narrow-wide all-graphene devices. *Nanotechnology, IEEE Transactions on*, PP(99):1 –1, 2010.

[73] D. B. Farmer, R. Golizadeh-Mojarad, V. Perebeinos, Y.-M. Lin, G. S. Tulevski, J. C. Tsang, and P. Avouris. Chemical doping and electron-hole conduction asymmetry in graphene devices. *Nano Letters*, 9(1):388–392, 2009.

[74] J. M. Rabaey, A. Chandrakasan, and B. Nikolic. *Digital Integrated Circuits - A Design Perspective (2nd Ed)*. Prentice Hall, 2003.

[75] S. Ramo, J. R. Whinnery, and T. Van Duzer. *Field and Wave in Communication Electronics*. Wiley, 1993.

[76] J.D. Jackson. *Classical Electrodynamics*. Wiley, 1975.

[77] N. Neophytou, J. Guo, and M. Lundstorm. Electrostatics of 3d carbon nanotube field-effect transistors. In *Computational Electronics, 2004. IWCE-10 2004. Abstracts. 10th International Workshop on*, pages 175 – 176, 2004.

[78] J. N. Murrell and A. J. Harger. *Semi-Empirical SCF MO Theory of Molecules*. Wiley, 1972.

[79] S. Hasan, J. Wang, and M. Lundstrom. Device design and manufacturing issues for 10 nm-scale mosfets: a computational study. *Solid-State Electronics*, 48(6):867–875, 2004. Silicon On Insulator Technology and Devices.

[80] J. Guo, A. Javey, H. Dai, and M. Lundstrom. Performance analysis and design optimization of near ballistic carbon nanotube field-effect transistors. In *Electron Devices Meeting, 2004. IEDM Technical Digest. IEEE International*, pages 703 – 706, 2004.

[81] F. Leonard and J. Tersoff. Role of fermi-level pinning in nanotube schottky diodes. *Phys. Rev. Lett.*, 84(20):4693–4696, May 2000.

[82] J. Martin, N. Akerman, G. Ulbricht, T. Lohmann, J. H. Smet, K. von Klitzing, and A. Yacoby. Observation of electron-hole puddles in graphene using a scanning single-electron transistor. *Nat Phys*, 4(2):144–148, February 2008.