Electronic Transport in Graphene Nanoribbons

Melinda Young Han

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Fu Foundation School of Engineering and Applied Science

COLUMBIA UNIVERSITY
2010

© 2010
Melinda Young Han
All Rights Reserved

ABSTRACT

Electronic Transport in Graphene Nanoribbons

Melinda Young Han

This dissertation examines the electronic properties of lithographically fabricated graphene "nanoribbons" (GNRs) with widths in the tens of nanometers. Low temperature and temperature-dependent measurements reveal a length- and orientation-independent transport gap whose size is inversely proportional to GNR width. Electronic states at energies in the gap are localized, and charge transport exhibits a transition between variable range hopping at lower temperatures and simple thermally activated transport at higher temperatures. A comparison between devices with different geometric capacitance shows that charging effects constitute a significant portion of the activation energy. Graphene devices operated at high source-drain bias show a saturating I–V characteristic. This decrease in conductivity at high applied electric field is described by carrier velocity saturation due to optical phonon emission. Finally, graphene nanoribbons that have been treated with a diaminopropane solution exhibit a width-dependent charge doping indicative of edge-selective covalent chemical functionalization.

Contents

Acknowledgments **v**

1 Introduction **1**
1.1 Graphene's short history . . . 1
1.2 Potential for electronics . . . 4
1.3 Physical and electronic structure . . . 6
1.4 Band gap engineering . . . 10
1.5 Overview of this work . . . 12

2 Band gap engineering **14**
2.1 Background: Theoretical predictions for band gaps in GNRs . . . 15
2.2 Device Fabrication . . . 18
2.2.1 The "Scotch tape" method for graphene deposition . . . 20
2.2.2 Graphene nanoribbon fabrication . . . 23
2.3 Experiment . . . 26
2.4 Results and Discussion . . . 29
2.4.1 Conductance near the charge neutrality point is suppressed . . . 29
2.4.2 GNRs have a well-defined sheet conductivity in the on-state . . . 30
2.4.3 The energy gap size is measured with 2-D conductance plot . . . 32
2.4.4 The energy gap follows a robust scaling rule . . . 36
2.5 Conclusions . . . 41

3 Disordered graphene nanoribbons **43**
3.1 Background . . . 44
3.1.1 Models for transport in disordered systems . . . 44
3.1.2 Previous work on disordered graphene nanoribbons . . . 49
3.2 Experiment . . . 53
3.3 Results and discussion . . . 55
3.3.1 Signatures of hopping transport . . . 55
3.3.2 The gap in bias voltage corresponds to a critical electric field . . . 63
3.3.3 Coulomb charging contributes to E_a . . . 65
3.4 Conclusions . . . 68

4 Graphene devices at high bias **70**
4.1 Background: Current saturation in carbon nanotubes 70
4.2 Experiment: Measurement at high source-drain bias 71
4.3 Results and discussion 72
4.3.1 Saturating behavior fits a velocity saturation model 72
4.3.2 Enhanced current saturation in top-gated devices 79
4.3.3 Heating effects can overcome transport gap at high bias 80
4.4 Conclusions 88

5 Chemical edge functionalization **90**
5.1 Motivation and Background 90
5.2 Experiment 92
5.3 Modeling 94
5.4 Results and Discussion 99
5.5 Conclusions and avenues for future work 103

List of Figures

1.1 Bravais lattice and reciprocal lattice of graphene 7
1.2 Graphene dispersion relation 9

2.1 Naming conventions for CNT and GNR directions 16
2.2 Theoretical predictions for GNR gap sizes 19
2.3 Deposition of graphene by mechanical exfoliation 21
2.4 Optical microscope images of graphene on silicon dioxde 23
2.5 Process flow for graphene nanoribbon fabrication 24
2.6 Atomic force microscope image of GNR devices 26
2.7 Scanning electron microscope images of GNR devices 27
2.8 Contact resistance measurement of GNRs 29
2.9 Conductance vs. gate voltage in GNRs of different widths 30
2.10 On-state conductance behavior of GNRs 31
2.11 Schematic of band gap measurement 33
2.12 2-D plots of device response and measurement of E_{gap} 34
2.13 Scaling of GNR energy gaps as a function of width 37
2.14 Comparison with data of different scaling predictions 40

3.1 Back-gated vs. dual-gated device geometries 55
3.2 Measurement of ΔV_g 56
3.3 Temperature dependence of the GNR conductance minimum 58
3.4 Characteristic GNR transport energy scales 59
3.5 Temperature dependence of wide GNRs 62
3.6 Measurement of ΔV_b 64
3.7 Length dependence of ΔV_b 65
3.8 Temperature dependence of back-gated vs. dual-gated devices 67
3.9 Electric field simulations for back-gated and dual-gated GNRs 67

4.1 High bias transport in a back-gated GNR 72
4.2 Saturation behavior in a back-gated GNR 73
4.3 Fit of velocity saturation based current model 75
4.4 Saturation velocity values from fits to the model 76
4.5 Current-voltage characteristics for a dual-gated GNR 81
4.6 The “kink” behavior in top-gated devices at high bias 82
4.7 High bias I–V for ribbons vs. wide devices 84
4.8 Current density for a ribbon and a wide device at high bias 85

4.9 GNR gap sizes compared to room temperature thermal energy 86
4.10 Thermal modeling of back-gated and dual-gated GNRs 88

5.1 Finite element modeling of GNR electrostatics 96
5.2 Model of surface charge density and gate capacitance of GNRs 97
5.3 Expected shift in V_{CNP} due to edge or plane doping of GNRs 99
5.4 Measured shift in V_{CNP} of GNRs after various treatments 101
5.5 Fits of doping models to measured ΔV_{CNP} 102

Acknowledgments

This dissertation represents an effort made possible only by the solid support team that has surrounded me throughout this process.

First, my sincere thanks to my advisor, Philip Kim; you have repeatedly impressed me in the consistency of your support, guidance, and patience. Throughout my graduate education, it has always been clear that you are behind me in the pursuit of my goals, whatever they may be. Beyond that, you taught me what it means to be a good experimentalist, both directly and by example, and you taught me to hold my work to a high standard.

The students and postdocs I've worked with have made this process not only endurable, but enjoyable. Thank you to Mikey Purewal, my big brother in the lab; your encouragement, apparent total confidence in my abilities, and friendship was especially essential on the rough days, and was key to my acclimation to life as a grad student. Thank you Erik Henriksen, with whom I shared many lunches and pots of coffee; you are a true friend, and through your approach to research – always building, tinkering, and asking questions, you've also taught me a few lessons in how to be a scientist. Thanks to Yuri Zuev; I feel like we grew up together as students and as researchers, and I know you'll do great on that home stretch without me. Thank you Yue Zhao, who can always make me smile, and Dmitri Efetov, who can always relate, thanks Fereshte Ghahari Kermani, for your reliable sweetness, and thank you Andrea Young, for lots of good pasta and good conversation. Thanks also to my predecessors in the lab, Yuanbo Zhang, Josh Small, and Alex Henderson, who showed me the ropes. And thank you Kirill Bolotin and Barbaros Özyilmaz, two very creative, fun, and supportive former postdocs. You are actually nothing alike, but I thank you both for bringing perspective to my view of academic life.

I'm also indebted to several faculty members who have served has my mentors and teachers. Thank you Jim Yardley, who I learned early on I could rely on for encouragement and a fresh point of view. Thanks to Ioannis Kymissis, who's energy and expertise have been invaluable in hunting out my next career step. And thank you Irving Herman, to whom I've already awarded an informal teaching award, I meant it!

I also want to thank my family, for getting me this far, and for cheering me on from a distance. My mom, the best in the world, this accomplishment is in no small part a result of years of your love and support. You are also more of a scientist than you will admit – I credit you for sparking my interest in science and holding my interest in math. Thanks Mommy. And my dad, who's hard work has brought me

many opportunities; I think its no accident that I followed closer to your footsteps than I ever planned. You've been a source of both confidence and inspiration to me, thank you.

Thank you to my big brother Steven, who, for as long as I can remember has offered his services to me as an expert consultant on life decisions, big and small. After you graduated from college and entered the "real world", you advised me to get good grades and stay in school as long as possible. Good call, thank you! And since I've been in grad school, you've worked to steer me to use my skills for good (solar cells, etc.) rather than evil (mind-controlling nanobots, etc.). Again, thanks!

Thank you to my good friend Heidi Perry, who suffered through the first draft of this thesis. More than that, you've been there through the toughest parts of this process, and you've put up with me when research made me crazy or late for dinner. When I am at my most erratic, you are somehow the steady and insightful support I need. Thanks for your help.

Finally, I acknowledge the scientific contributors to the work in this thesis. I thank Yuanbo Zhang for getting me started on the nanoribbon project, and for teaching me how to make graphene. Thanks to Barbaros Özyilmaz for helping to develop fabrication procedures that are used throughout this thesis. Thank you Juliana Brant for very careful work on the chemical functionalization project and for help with fabrication and measurements on the transport gap temperature dependence. Thanks to Noah Tremblay and Colin Nuckolls for developing and carrying out the chemistry for the edge functionalization project. I also thank Inanc Meric, Andrea Young, and Kenneth Shepard for their contributions to the high bias project. And thanks to the National Science Foundation for three years of my paychecks.

Thank you everybody, its been fun!

Chapter 1

Introduction

1.1 Graphene's short history: the youngest member of a distinguished family

Graphene is a single layer of carbon atoms bound in a honeycomb arrangement. With its first experimental discovery in 2004, graphene takes its place as the two-dimensional member in the family of graphitic carbon allotropes. The three-dimensional, and most familiar member of this family, graphite, is made up of stacks of graphene which are weakly bound together; this is what gives graphite the slipperiness that makes it useful as a solid lubricant or as pencil "lead". A carbon nanotube is conceptually[1] made of graphene rolled into a tube with a diameter of typically < 2 nm [1], and can be thought of as a one-dimensional form of graphitic carbon. Fullerenes, such as C_{60}, are ball-shaped cages of carbon atoms, with graphene's hexagons recognizable in the soccer ball arrangement of the atoms. These materials, also known as

[1] and only conceptually, contrary to a relatively prevalent popular understanding of carbon nanotube fabrication

buckyballs, constitute the zero-dimensional member of the family. Its interesting to note that although graphene is in some sense the basic building block for all of these materials, it was the last to be isolated and identified.[2]

Graphene has long existed on paper (no pun intended) as a theoretical building block of solid graphite. As early as 1947, the electronic band structure of graphite was calculated with the tight-binding model by first considering the structure of individual sheets of graphene[4]. At this time, however, graphene was mainly considered a theoretical or toy material, and if fact it had been predicted by some that 2-D crystals could not exist at room temperature [5].

Over the decades, various attempts have been made to separate graphite into its layers. Early attempts involved intercalating bulk graphene with other molecules, which produces new materials interesting in their own right [6], which, in some cases have been used to separate graphite into scrolled graphene sheets in solution [7]. A more successful method for producing thin graphite on substrate was the so called "nanopencil". This technique involves attaching a carefully patterned piece of bulk graphite to an atomic force microscopy (AFM) tip and rubbing it across a clean surface. Using this method, few-layer graphite samples as thin as 30 atoms thick can be achieved [8].

It was not until 2004 that graphene saw its first experimental realization, by Andre Geim's group at the University of Manchester [9]. The method they used is almost incomprehensibly simple. They separated single atomic layers of graphene from bulk

[2] 0-D buckyballs were discovered in 1985 [2], 1-D carbon nanotubes in 1991 [3], and 3-D graphite centuries before.

graphite using scotch tape. This discovery spurred the explosion of a new field of science where “the scotch tape method” for graphene production is the industry standard for achieving high quality graphene samples.

In the years following the first experimental isolation of graphene, the amount of research on graphene grew at literally an exponential rate. After its initial discovery, the first experimental works on graphene focused the novel quantum hall effect that results from graphene’s unique band structure [10, 11]. From there, theoretical results poured in with a wide range of ideas and perspectives, and experimentalists with experience in two-dimensional electron gases (2DEGs) or carbon nanotubes used their expertise in these related materials to study graphene’s electronic structure. Along with this came a push in research related to graphene’s potential for electronics applications, and then its chemical and mechanical properties. As of 2010, at Columbia University alone, at least six different departments[3] are pursuing some graphene-related research effort.

This dissertation focuses on the electronic properties of graphene, specifically the electronic properties of graphene which has been patterned into “nanoribbons” with widths in the tens of nanometers. I will address questions regarding the size and nature of the transport gap that is induced by etching graphene into nanoribbons, the potential for graphene nanoribbons in electronic devices, and possible routes for improving the quality of graphene nanoribbons.

[3] Physics, Applied Physics and Applied Math, Chemistry, Electrical Engineering, Mechanical Engineering, and Chemical Engineering

1.2 Potential for electronics: the appealing properties of carbon nanotubes, made controllable by lithography

After its first experimental realization, graphene was quickly recognized as a material with great promise for use as a semiconducting material in commercial electronics. Graphene has many appealing electronic properties.

Most prominent is its high room temperature carrier mobility. Electron and hole mobilities of up to $200,000$ $\mathrm{cm}^2/\mathrm{V}\cdot\mathrm{s}$ have been measured in suspended graphene samples, where the underlying silicon dioxide substrate has been etched away [12], but the properties of substrate-supported graphene are more relevant for electronics applications. In graphene on substrate, mobilities of up to $20,000$ $\mathrm{cm}^2/\mathrm{V}\cdot\mathrm{s}$ have been measured [5]. For comparison, electron and hole moblities in clean silicon are 1500 and 450 $\mathrm{cm}^2/\mathrm{V}\cdot\mathrm{s}$, respectively [13].[4]

Perhaps even more relevant for electronics than carrier mobility is saturation velocity. Graphene has been shown to have a saturation velocity of up to $\sim 5\times 10^7$ cm/s on a silicon dioxide substrate [14]. This is a five-fold increase over the saturation velocity in silicon, 1×10^7 cm/s. Other appealing properties include high thermal conductivity, reported to be 5×10^3 W/mK [15], and high maximum current densities of ~ 2 mA/μm in width, or 6×10^9 $\mathrm{A/cm}^2$ if we take the thickness of a graphene layer to be 3.41 Å, the interlayer spacing in graphite.

From a broad view, the electronic properties and resulting attention surrounding

[4]Some III-V semiconductors have mobilities higher than the state-of-the-art in substrate-supported graphene. For instance, InSb has mobilities of $80,000$ $\mathrm{cm}^2/\mathrm{V}\cdot\mathrm{s}$ for electrons and 1250 $\mathrm{cm}^2/\mathrm{V}\cdot\mathrm{s}$ for holes. However, since these materials are rare and expensive, they have little potential for widespread commercial use.

graphene is reminiscent of the spotlight on carbon nanotubes in the late 90's and early 2000's. Carbon nanotubes have promising electronic properties analogous to those described above, a similarity which is unsurprising given their closely related physical structure. For a time, carbon nanotubes were heralded as the future of electronics, and it seemed that they would be the semiconductor industry's much sought after post-silicon material. All that was needed was a method to controllably place the desired types of nanotubes in the desired locations on a wafer. As of this writing, such a method has not been discovered, and predictions for commercially successful carbon nanotube electronics are increasingly rare. Graphene, however, has one major advantage that may allow it to succeed where carbon nanotubes have not. Graphene is a planar material, which means in principle[5] it can be seamlessly integrated into well-developed top-down fabrication processes, effectively eliminating the controllability problem.

Graphene's two-dimensional structure has other advantageous differences from both bulk 3-D materials and conventional 2DEGs. Since all of its atoms are exposed on the surface, it can be easily doped either chemically or electrostatically with local precision. In combination with the ambipolar band structure and the linear density of states at low energies, this allows for the fabrication of graphene heterostructures with n- and p-type graphene separated by sharp electric field gradients [23], and also makes graphene a useful material for chemical sensing [24], or for applications like solar cells, where control of the workfunction is important [25].

[5] Assuming the development of a suitable wafer-scale growth process, which is to be expected [16–22]

1.3 Physical and electronic structure of graphene: a honeycomb of atoms leads to “Dirac cones” in k-space

Graphene is made up of one single layer of carbon atoms, arranged in a “honeycomb” pattern of repeating hexagons (Fig 1.1 (a)). The unit cell contains two carbon atoms, labeled A and B in Fig 1.1(a). The length of the carbon-carbon bond is known to be 1.42 Å. The unit vectors are

$$\vec{a_1} = \left(\frac{\sqrt{3}}{2}a, \frac{a}{2}\right), \ \vec{a_2} = \left(\frac{\sqrt{3}}{2}a, -\frac{a}{2}\right) \tag{1.1}$$

where the lattice constant $a = 1.42$ Å $\times \sqrt{3} = 2.46$ Å, so that $|\vec{a_1}| = |\vec{a_2}| = a$. This gives us a reciprocal lattice that is also hexagonal, rotated 90° from the Bravais lattice. The reciprocal lattice has with a lattice constant $4\pi/\sqrt{3}a$ and unit vectors

$$\vec{b_1} = \left(\frac{2\pi}{\sqrt{3}a}, \frac{2\pi}{a}\right), \ \vec{b_2} = \left(\frac{2\pi}{\sqrt{3}a}, -\frac{2\pi}{a}\right) \tag{1.2}$$

shown in Fig 1.1(b).

Carbon atoms have four valence electrons. Away from the edges, each carbon atom in a sheet of graphene shares a bond with three neighbors. If the graphene lies in the x–y plane, then these three σ bonds correspond to sp^2 hybridized orbitals formed from the $2s$, $2p_x$ and $2p_y$ orbitals. This leaves the $2p_z$ orbital, perpendicular to the plane, to form π covalent bonds. The resulting π energy bands are the main contribution to the electronic structure of graphene. These bands result from the

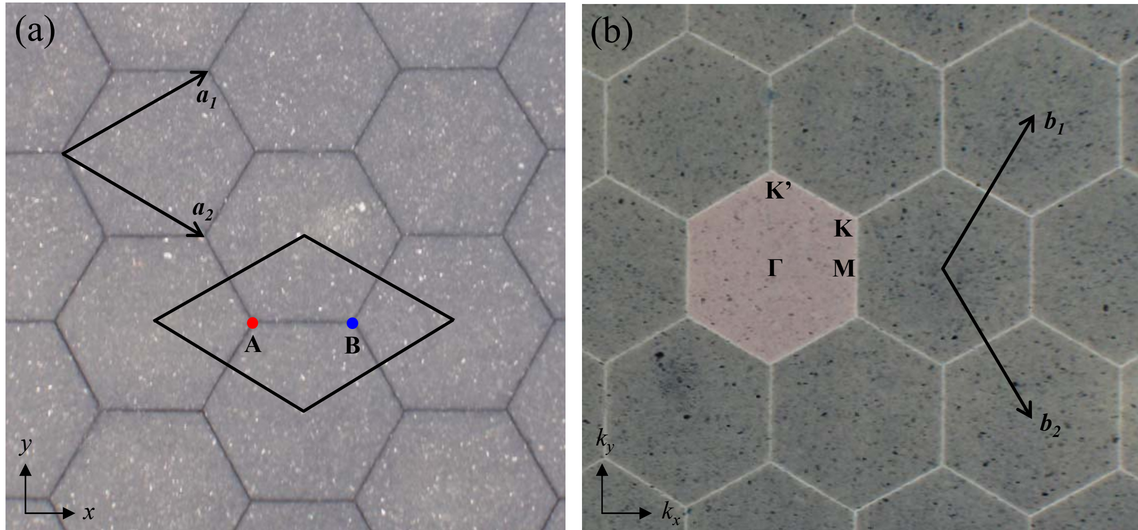

Figure 1.1: (a) Model of the graphene lattice structure, with unit cell and primitive lattice vectors, taken from College Walk on the Columbia University campus. Red and blue dots correspond to A and B carbon atoms, respectively. (b) Graphene reciprocal lattice, with the first Brillouin zone shaded and high symmetry points and unit vectors labeled.

one electron per atom which is not involved in nearest-neighbor σ bonding, so in calculations of the band structure, we treat only this one electron per carbon atom.

Note here the similarities between the electronic orbitals in graphene and those in conjugated molecules with sp^2 hybridization, such as polyacetylene. In this way, graphene can be thought of as a giant aromatic molecule. In fact, the analogy to conjugated molecules is sometimes used in the literature in studying the electronic structure of small pieces of graphene or narrow graphene ribbons. [1, 26–28]

We carry out a tight binding model calculation on the lattice described above in order to find the electronic band structure. A well known approach is that in Reference [1], which considers only nearest-neighbor interactions and does not neglect the overlap of the p_z wavefunctions centered on different atoms. This approach results

in the dispersion given by:

$$E(\vec{k}) = \frac{\epsilon_{2p} \pm tw(\vec{k})}{1 \pm sw(\vec{k})} \tag{1.3}$$

where the atomic energy of the $2p_z$ orbital, ϵ_{2p}, is usually set to zero, t is the nearest-neighbor transfer integral, s is the overlap integral between nearest A and B atoms, and $w(\vec{k})$ is given by:

$$w(\vec{k}) = \sqrt{1 + 4\text{cos}\frac{\sqrt{3}k_x a}{2}\text{cos}\frac{k_y a}{2} + 4\text{cos}^2\frac{k_y a}{2}} \tag{1.4}$$

Figure 1.2 shows a visualization of the dispersion in Equation 1.3 with $\epsilon_{2p} = 0$, $t = -3.033$ eV, and $s = 0.129$. Neglecting the higher order contributions, we obtain a simple approximation for the graphene dispersion relation which is symmetric around $E = \epsilon_{2p} = 0$ with $s \approx 0$:

$$E(k_x, k_y) \approx \pm t\sqrt{1 + 4\text{cos}\frac{\sqrt{3}k_x a}{2}\text{cos}\frac{k_y a}{2} + 4\text{cos}^2\frac{k_y a}{2}} \tag{1.5}$$

For the purposes of this dissertation, there are two important features to notice about this band structure.

First, graphene is a zero-gap semiconductor. The tight binding calculation outlined above gives two π energy bands (one bonding and one anti-bonding). These two bands touch at two non-equivalent K points, which we call K and K', found at the corners of the Brillouin zone. There are two atoms per unit cell, so the bottom band is fully occupied. The shape of the bands at high energies depends to some extent

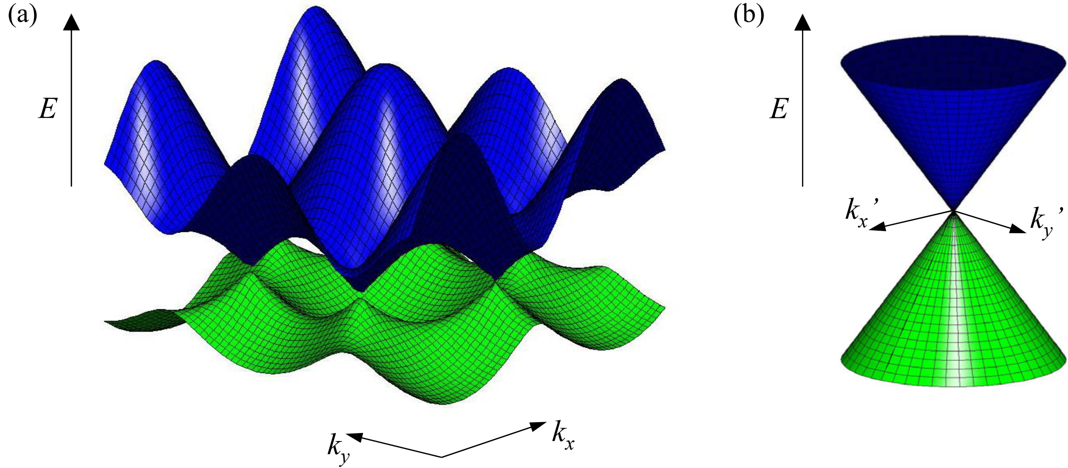

Figure 1.2: (a) Dispersion relation for graphene π energy bands. Plotted here is the tight binding result in Eqn. 1.3 with $t = -3.033\ eV$ and $s = 0.129$, following Ref [1]. (b) Detail of a low energy "Dirac cone" found near the K points, with $\vec{k'} = \vec{k} - \vec{K}$.

on the approximations used in the calculation, but the degeneracy at the K points is unaffected by choice of approximation. Instead it comes from the symmetry of the lattice structure, and the fact that the two atoms in the unit cell are the same [1].

Second, at low energies, the tight-binding result reduces to a linear dispersion relation:

$$E \approx \hbar v_F |\vec{k'}| \tag{1.6}$$

where $\vec{k'} = \vec{k} - \vec{K}$. That is, near the K points, the electronic dispersion is linear in wavevector k, in stark contrast to the parabolic dispersion usually seen for electrons. The equi-energy contours are circular around the K points, producing so-called Dirac cones in the dispersion.

This linear dispersion has several interesting implications. First, the band velocity of charge carriers is constant. Electrons move with a constant "Fermi velocity" v_F in

the lattice which is independent of the wavevector k, in contrast to the parabolic dispersion seen in most other materials. Using the tight binding parameter $t = 3.1$ eV, the Fermi velocity is estimated to be $v_F \approx 10^6$ m/s. Second, since the linear bands have no curvature, the effective mass approximation breaks down. Taken together, the consequence is a breakdown of the usual Schrödinger approach to quantum mechanics with a massive carrier. Instead we employ the analogous relativistic quantum mechanics, using the Dirac equation with the speed of light replaced by v_F:

$$\hat{H} = \hbar v_F \begin{pmatrix} 0 & k_x - ik_y \\ k_x + ik_y & 0 \end{pmatrix} = \hbar v_F \vec{\sigma} \cdot \vec{k} \tag{1.7}$$

1.4 Band gap engineering: How carving graphene into a ribbon induces a band gap

The above description of the graphene band structure shows that this material is a zero-gap semiconductor. However, many electronics applications rely on the use of a semiconductor with a finite band gap, where this energy barrier to charge transport can be used to control electron flow. This brings us to the central theme of this dissertation. In this thesis, I will show that it is possible to induce a transport gap in the graphene spectrum by carving graphene into a narrow ribbon, or a "graphene nanoribbon". The result is something like an unrolled carbon nanotube [29, 30], that could potentially have the promising electronic properties of carbon nanotubes with the large-scale controllability of a planar material.

From a heuristic approach, we can understand the induced transport gap in

graphene nanoribbons in terms of an energy gap arising from simple quantum confinement. In carbon nanotubes, the band gap depends on the diameter and chirality (orientation with resect to the graphene plane) of the tube. Likewise, we expect that the band gap of a graphene nanoribbon will depend on its width and crystallographic orientation.

As a first estimate of the size and the width dependence of the band gap in a graphene nanoribbon, we carry out a "back-of-the-envelope" calculation for the allowed energies expected in a graphene nanoribbon. The two ingredients to this calculation are (i) quantum confinement in one dimension, and (ii) graphene's linear dispersion at low energies. Quantum confinement in the x-direction results in allowed wavefunctions:

$$\psi(x) \propto \sin(k_x x) \text{ with } k_x = n\pi/W \tag{1.8}$$

where n is an integer, and W is the ribbon width. We then introduce the dispersion relation

$$E \approx \hbar v_F \sqrt{k_x^2 + k_y^2} \tag{1.9}$$

to obtain allowed energy levels

$$E_n = n\frac{\hbar v_F \pi}{W} \tag{1.10}$$

so that the gap between levels is

$$E_{gap} = \frac{\hbar v_F \pi}{W} \tag{1.11}$$

This simple calculation gives us an idea of the basic form of $E_{gap}(W)$. For curiosity's sake, the coefficient $\hbar v_F \pi$ in the W dependence above comes to 2.06 eV $\cdot$ nm. More detailed approaches to calculations of the graphene nanoribbon band structure and gap size will be presented in Chapter Two.

1.5 Overview of this work

This dissertation will examine electronic transport in graphene nanoribbons and address the potential and remaining technological hurdles in using graphene nanoribbons as a material for electronic devices.

Chapter Two introduces graphene nanoribbons as a semiconducting material, including their fabrication and measurement. In this chapter, I experimentally determine a general scaling rule for the size of the "band gap" as a function of ribbon width. Many theoretical and experimental results have been reported between the time the work in this chapter was originally published and the writing of this thesis, so I will present this work in the context of the published results at the time.

Chapter Three contains a more detailed approach to the electronic structure of graphene nanoribbons, particularly the nature of the gap, which I now refer to more generally as a "transport gap". Experiments in this chapter reveal conduction through localized states inside the transport gap. Here I will present an up-to-date understanding of transport in graphene nanoribbons as reported in the literature as well as in my own work.

Chapter Four explores the possibilities for use of graphene and graphene nanoribbons as real electronic devices, specifically the properties of graphene nanoribbon devices when operated at high source-drain biases. This includes measurements of velocity saturation behavior in graphene devices operated at high bias, as well as simulations of the relevant heat dissipation.

Chapter Five addresses the possibilities for engineering of the etched ribbon edges through chemical functionalization. I present one potential method for achieving such functionalization, with computational predictions and preliminary experimental results.

Chapter 2

Band gap engineering in graphene nanoribbons

In this chapter, we introduce graphene nanoribbon fabrication and measurement, and find a general scaling rule for the size of the "band gap" as a function of ribbon width. The results presented in this chapter were published in Ref. [31] as the first experimental report on graphene nanoribbons, though this publication was followed soon after by a related study by the IBM group [32] and preceded by a report on nanoribbons made from few-layer epitaxial graphene by the Georgia Tech group [16].

2.1 Background: Theoretical predictions for band gaps in graphene nanoribbons

Here we review the theory of the electronic structure of graphene nanoribbons (GNRs), focusing on predictions for band gaps in graphene nanoribbons. In this section we address only the background to the experiment in this chapter, that is, the theory that was done by early 2007, before the experiment presented here was originally published. Specifically, this means tight-binding method and density functional theory studies on ribbons with "perfect" edges. "Perfect" edges are the atomically smooth edges obtained by cutting a straight line in a graphene sheet, either in the high symmetry armchair and zigzag directions, as described below, or in low-symmetry directions. Later studies on disordered graphene nanoribbons, with random vacancies and other structural disorder at the edges, will be discussed in Chapter Three.

In the honeycomb structure of a sheet of graphene, there are two high symmetry directions, known as "armchair" and "zigzag".The armchair direction can be found by drawing a straight line through any two nearest-neighbor carbon atoms. Tracing a path in this direction though nearest-neighbor bonds gives a shape reminiscent of an overstuffed armchair, shown by the red lines in Fig. 2.1(a) and (c). The zigzag direction is 30° away from the armchair direction in the graphene plane. Tracing along nearest-neighbor bonds in this direction gives a path that "zigzags" neatly back and forth(Fig. 2.1(b) and (d)). Note that there is a 60° periodicity to crystallographic angles in the lattice.

The naming conventions in the literature for CNT and GNR crystallographic

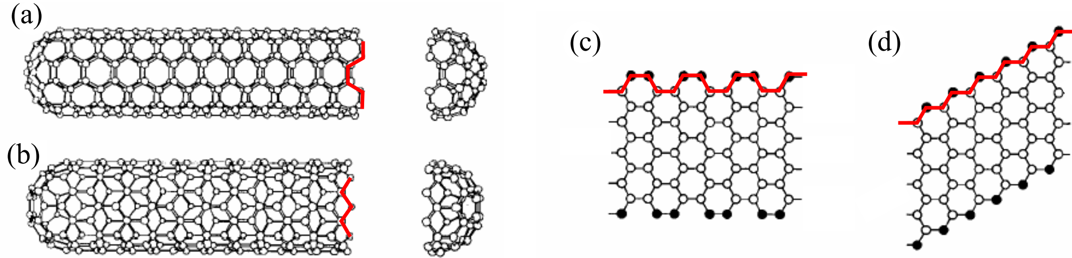

Figure 2.1: Naming conventions for (a) armchair CNTs, (b) zigzag CNTs, (c) armchair GNRs, and (d) zigzag GNRs. Diagrams reproduced with permission from [33, 34] Red lines added to highlight the armchair and zigzag edges.

orientation require careful clarification. A carbon nanotube is conceptually a rolled up piece of graphene, with the lattice seamlessly joined together along the longitudinal axis. A graphene nanoribbon can then be imagined as an unrolled carbon nanotube. If a graphene nanoribbon runs in the armchair direction, that is, if the long edge of a ribbon shows an armchair pattern, then it is referred to as as an armchair nanoribbon. However, an armchair carbon nanotube refers to a tube with an armchair pattern in its lateral direction, the edge that would be exposed of you cut off the end of a nanotube. Likewise, a zigzag GNR has a zigzag edge in its longitudinal direction, and a zigzag CNT has a zigzag pattern in its lateral direction. Since the armchair and zigzag directions run perpendicular to each other in the graphene lattice, this has the confusing consequence that an “unrolled” zigzag nanotube is an armchair nanoribbon, and an “unrolled” armchair nanotube is a zigzag ribbon (See Fig. 2.1). Not surprisingly, we will see that the basic band structures of zigzag (armchair) nanoribbons have some similarities with that of armchair (zigzag) nanotubes.

The first theoretical works on graphene nanoribbons and their band gaps appeared

several years before graphene was discovered. In 1987, Tanaka *et. al.* [26] carried out tight binding calculations for polymer chains they called "One-dimensional graphite", and noted a decrease in the band gap for molecules with a wider graphene structure added to the polymer backbone. Stein *et. al.* [27] studied hexagonally symmetric aromatic molecules, essentially small pieces of graphene, and found a decrease in the HOMO-LUMO gap with increasing molecule size. The mid-90s saw some renewed interest in GNRs, sometimes called "nanographites" or "carbon nanoribbons". Though the bulk of this research focused on the anomalous edge state predicted for zigzag edges [35], several works also predict a band gap with a size inversely proportional to ribbon width [34, 36].

Early calculations of GNR electronic structure [28, 34] take the same basic approach that worked well for carbon nanotubes, that is, first carry out a tight-binding calculation for graphene, then project it onto the appropriate axis using the zone-folding technique. The difference in boundary conditions is tacked on as a correction. As might be expected, this gives roughly the same result for GNRs as for CNTs. Armchair ribbons are expected to come in three "families", one of which is metallic, and zigzag ribbons are, in the simplest picture, metallic. In zigzag ribbons, however, a new prediction arises in performing this calculation on GNRs instead of CNTs: the degenerate state near the Fermi level comes in the form of a pair of flat bands corresponding to localized edge states. These zigzag edge states have been the subject of much theoretical [37], and some experimental [38] work, though they will not be discussed any further in this dissertation. A more sophisticated approach by Brey

and Fertig [39], using the Dirac equation with a tight-binding Hamiltonian and appropriate boundary conditions, produced qualitatively similar results for the band gap dependence on ribbon width for different GNR varieties.

Later work by Son *et. al.* [40] takes an *ab initio* approach that captures some features of the electronic structure missed by previous approaches. Most prominently, this work also finds that armchair ribbons come in three families, but in their local density approximation (LDA) calculations they find that all three of these families have a non-zero band gap that varies inversely with ribbon width, so that all varieties of GNRs have band gaps.

Barone *et. al.* [41] also use density functional theory, and they make specific predictions of scaling laws for band gaps of ribbons with different edges and terminations. They predict that GNR band gaps scale as $E_g(W) = aW^{-b}$ where E_g is band gap and W is ribbon width, and a and b are constant parameters. Depending on the crystallographic orientation, chemical edge termination, and the families discussed above, they find values of a from 0.30 to 1.6 eV and values of b from 0.76 to 1.1.

It is with this theoretical background that we conducted the following experiment to determine the dependence of the band gap on GNR width and crystallographic orientation.

2.2 Device Fabrication

Samples discussed throughout this dissertation were fabricated following the same basic procedure, which will be described in this section. First we describe the method of

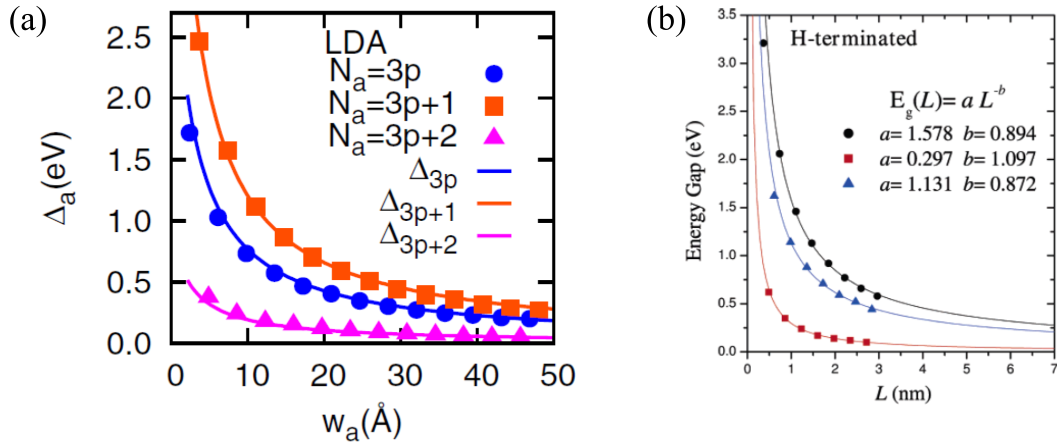

Figure 2.2: Theoretical predictions for GNR gap sizes versus ribbon width. (a) LDA calculations for energy gap size Δ_a as a function of ribbon width w_a for armchair GNRs. Gap behaviors fall into three families, depending on the number of dimer lines N_a contributing to the ribbon width; all three families are semiconducting. Reproduced with permission from Reference [40]. (b) Calculations for energy gap vs. ribbon width L for hydrogen terminated armchair GNRs. Reproduced with permission from Reference [41].

single layer graphene deposition and identification first discovered by the Manchester group in 2004 [9], then we describe our method for fabricating graphene nanoribbons from graphene sheets.

2.2.1 The "Scotch tape" method for graphene deposition

As of the time of this writing, the best quality graphene samples are made by a process called mechanical exfoliation, or the "Scotch tape" method. The basic idea of this method is to start with bulk graphite and cleave apart the graphene planes using adhesive tape. The result is single and few layer graphene pieces on a silicon dioxide substrate, which can then be identified using an optical microscope.

Graphene production by mechanical exfoliation is performed as follows. A small piece of graphite (a few millimeters across, and one hundred microns or less in thickness, as shown in Fig 2.3(a)) is placed on the sticky side of a piece of adhesive tape. The tape is folded onto itself, and then gently peeled apart, so that the graphite is cleaved apart into two thin pieces, mirror images of the original piece. This fold-and-peel step is repeated several (3-10) times, until a region of a few square centimeters of the tape is covered with thin graphite, as in Fig 2.3(b). Next, the tape is pressed, sticky side down, onto a clean chip from a silicon wafer with ~ 300 nm SiO_2, shown in Fig. 2.3(c). The chips used here were cleaned with a 10 minute bath in "piranha", a 3:1 sulfuric acid to hydrogen peroxide solution. Once in contact with the chip, the tape is gently rubbed with pair of teflon tweezers for several minutes to ensure good contact across the entire chip, and then the tape is carefully peeled off. In some places, the graphite comes off cleanly with the tape, leaving no graphite on the chip.

Figure 2.3: (a) Bulk kish graphite used for graphene exfoliation. (b) A piece of tape prepared with graphite and ready for pressing onto a chip. (c) Clean Si/SiO_2 substrates for graphene exfoliation.

In other places, Van der Waals forces keep the graphite stuck to the chip, and a few of the top layers of graphite are pulled away with the tape, so that slabs of graphite will remain after the tape is lifted. But in some places, the graphite sticks to the chip, and it cleaves between planes so that all but one layer pulls away with the tape, leaving behind the single layers of graphene. This method relies on a lucky balance of forces; the interplane binding between the graphite layers is weak, and the Van der Waals forces binding the graphene to the silicon dioxide surface are comparatively strong.

Through trial and error, we found a few parameters that anecdotally correlate with a higher yield of large, single layer pieces of graphene. The fold-and-peel step described above should be done as few times as possible. When the tape is ready to be placed on the chip, the graphite on the tape should be thin, with much of the graphite appearing matte in finish rather than shiny, and the graphite should be flat, that is, it should not sparkle when viewed at different angles of incident light. The rubbing step to increase contact between the graphite and the chip should be done

with light pressure; just the weight of the tweezers is enough. And the peeling step should be done slowly, with minimal force, and with a small angle between the tape and the chip.

The simplicity of this method is a key factor in the quick growth of experimental interest in graphene. Another major factor is the ease and reliability with which graphene can be found and identified using an optical microscope. Particularly on ≈ 300 nm SiO_2 substrates [42], single layers of graphene are easy to see in an optical microscope, and layers of different thicknesses are reliably distinguishable by their colors. Single and bilayer graphene can be unambiguously identified by their unique signatures in either Raman spectroscopy [43] or quantum hall transport measurements [10, 11], so that we have a known mapping between colors as seen in the optical microscope and number of layers (single layer, bilayer, or more). As a result, we can identify single and bilayer graphene by color alone, allowing us to quickly screen large numbers of samples without the need for more sophisticated or time-consuming procedures. Once identified based on the optical image, the layer thickness can be verified with Raman spectroscopy.

Figure 2.4 shows optical microscope images of exfoliated graphite at 2x, 10x, and 100x magnifications. The 100x image shows a piece of few-layer graphite that has cleaved with several different thicknesses. The lightest color, highlighted with a red arrow, corresponds to single-layer graphene. Also visible in this image are areas where two pieces of single layer graphene overlap, forming a region of bilayer graphene, as highlighted by the blue arrow. This image shows the clear visible difference between

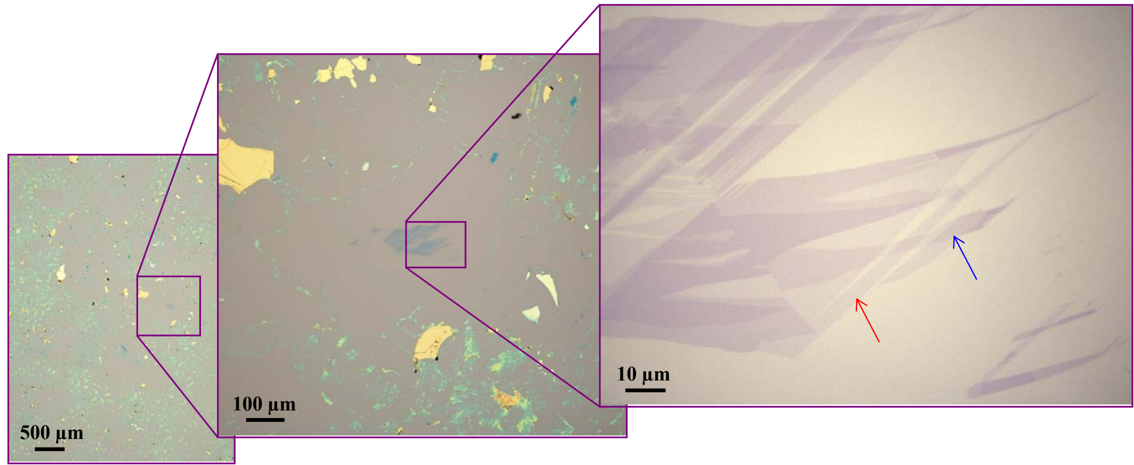

Figure 2.4: Optical microscope images of Si/SiO_2 substrate after graphene deposition using the piece of tape in Fig. 2.3(b). In the highest magnification image (right), single-layer graphene is visible as the faintest purple colored layer (red arrow). The distinctly darker color of bilayer graphene can be seen where two pieces of single-layer graphene overlap (blue arrow). More than two layers give an even darker color. In the lower magnification images (left and center), thick graphite pieces and adhesive are visible in yellow and blue, respectively.

single and bilayer graphene on this substrate.

2.2.2 Graphene nanoribbon fabrication

Here we describe the process for fabricating back-gated graphene nanoribbons that is used in experiments throughout this dissertation. The process flow is outlined in Fig. 2.5. Briefly, we begin with exfoliated graphene, fabricate metal electrodes using standard electron beam (e-beam) lithography procedures, pattern an etch mask using an negative e-beam resist, and etch away unprotected graphene using an oxygen plasma etch. An atomic force microscope (AFM) image of a finished device is shown in Figure 2.6.

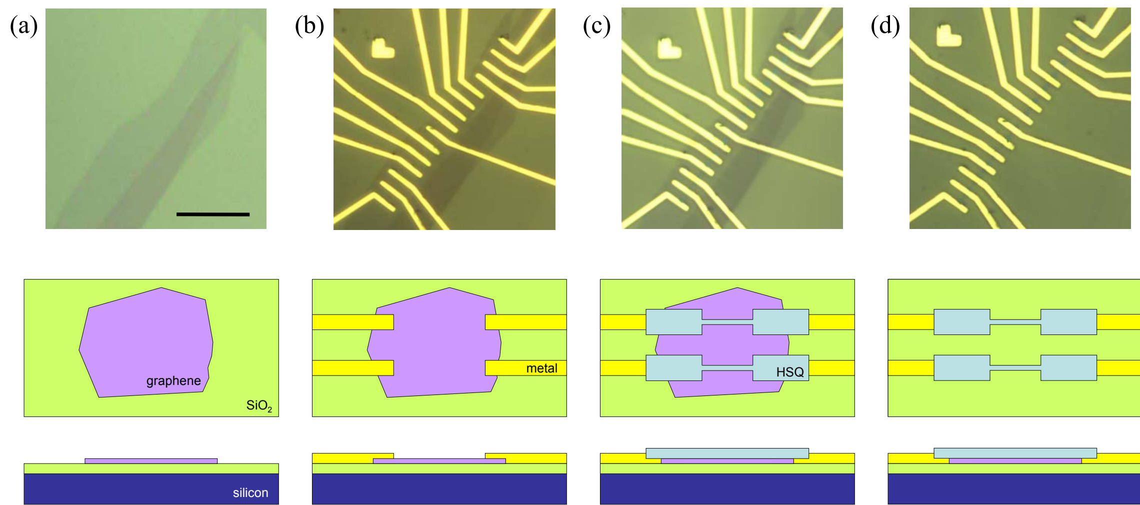

Figure 2.5: Process flow for graphene nanoribbon fabrication. Optical microscope image (top row), cartoon top view (middle row), and cartoon side view (bottom row), for each of four major processing steps. (a) Graphene deposition on Si/SiO_2 substrate. (b) E-beam lithography fabrication of metal electrodes. (c) Patterning of negative e-beam resist etch mask. (d) Removal of unprotected graphene by oxygen plasma etching. Scale bar in optical image is 20 μm, all four optical images have the same scale.

Once a suitable piece of graphene has been deposited and identified using the procedure described above, the next step is to electrically contact the graphene with metal electrodes using e-beam lithography. We begin by spinning on a layer of poly(methyl methacrylate) (PMMA) e-beam resist[1] and baking on a hotplate at 180°C for 2 minutes. Then we use e-beam lithography to write a 2 mm by 2 mm grid of alignment marks at roughly the location of the graphene, and develop in a solution of methyl isobutyl ketone:isopropal alcohol (MIBK:IPA) 1:3 for 5–10 seconds. This quick development leaves alignment mark "holes" in the PMMA, which we use for alignment in the following e-beam lithography step, eliminating the need for metal alignment mark deposition or another PMMA spin step. Electrodes are patterned in this PMMA layer with e-beam lithography, using an optical image of the sample with the alignment mark holes for design and alignment. Thermal evaporation is then used to deposit 1–2 nm of chrome and 25–50 nm of gold, and the chip is placed in acetone overnight at room temperature for lift-off (Fig 2.5(b)).

Once the graphene has been successfully contacted with Cr/Au electrodes, we create an etch mask to define the nanoribbons. A negative tone e-beam resist, hydrogen silsesquioxane (HSQ) (2% in MIBK) is spun on to the chip (at 4000 rpm, for a typical film thickness of 14 nm). We use HSQ as the resist for this step because a negative resist is ideal for creating a small etch mask, and because with HSQ we can obtain small feature sizes. The etch mask is written at a relatively high e-beam dose (1300 $\mu C/cm^2$ for the ribbons in our 30 keV system, with lower doses for larger features)

[1] typically PMMA A6 950K at 4000 rpm for a film thickness of $\sim$ 300 nm, though we have also used other thicknesses and materials, including copolymer (MMA(8.5)MAA)

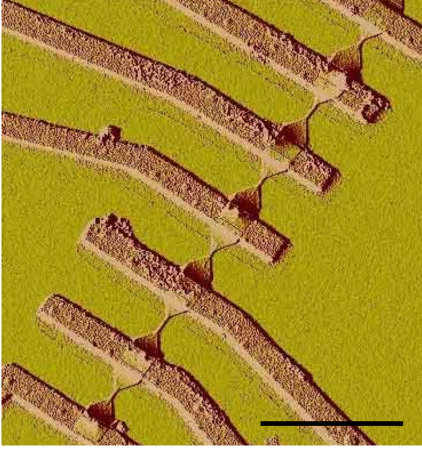

Figure 2.6: Atomic force microscope (AFM) image of the device in Figure 2.5. The ribbons in this image correspond to the lower six ribbons in the optical image in Figure 2.5(d). Scale bar is 4 μm.

and developed in a solution of 0.26N tetramethylammonium hydroxide (TMAH) in water for 1 minute (Fig. 2.5(c)).

After defining the etch mask, the graphene is ready to be etched. The device is exposed to oxygen plasma in a Technics reactive ion etcher (RIE) with 200 mTorr O_2 at 50 W for 5–10 seconds. These conditions etch graphene at a rate of about one layer per second, so that unprotected single layer and few-layer graphene are etched away cleanly (Fig. 2.5(d)). The finished device (Fig. 2.6) is then ready to be wirebonded and measured.

2.3 Experiment

In the experiment described in this chapter, we study two different types of device sets: sets P1–P4 each contain many ribbons of varying width running parallel (Fig 2.7(a)),

and device sets D1 and D2 contain ribbons of (nearly) uniform width and varying relative orientation (Fig 2.7(b)). All devices from a given set are etched from the same sheet of graphene, so that although the absolute crystallographic orientation of the graphene sheet has not been measured, the relatives orientations of the GNRs within a given set are known.

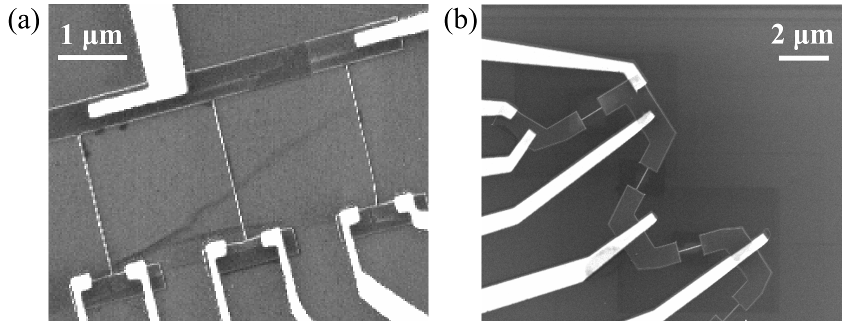

Figure 2.7: Scanning electron microscope (SEM) images of finished GNR devices. Ribbons connect to larger blocks of graphene which are contacted by the metal electrodes (white) (a) GNRs from set P1 of parallel devices. (b) GNRs from set D2 of devices with varying relative orientation.

Ribbons in this experiment have widths ranging from 10 to 100 nm and lengths of 1 to 2 μm. The dimensions of each GNR were measured using a scanning electron microscope (SEM) after the transport measurements were performed. Since the HSQ etch mask was not removed from the GNRs for this imaging, this measurement provides an upper bound to the true width of the GNR.

Each GNR is a three terminal device with a source and drain metal electrode contacting the graphene, and a back gate electrode connected to the doped silicon substrate which is used to modulate the carrier density in the graphene. In general, conductance of GNR devices is measured using a standard lock-in technique with a

small bias AC voltage (100 μV@13.5 Hz) applied to the channel. A DC voltage source is used for the gate voltage, and can also be used to apply a larger bias voltage along the channel.

We remark that each GNR connects two blocks of wider ($\sim 1\ \mu m$) graphene, which are in turn contacted by metal electrodes (Fig 2.7). Thus, unlike carbon nanotubes, Schottky barrier formation at the metal electrodes is absent in our GNR devices.

Furthermore, by contacting each of the wide "graphene leads" with two metal electrodes, it is possible to carry out a four-terminal measurement of the graphene device resistance without the effect of contact resistance from the graphene-metal interface. This geometry does not eliminate the contact resistance at the nanoribbon-graphene lead interface, which is assumed to be small. The measurement is shown schematically in Figure 2.8(a). Using this geometry, we measured the contact resistance $R_C = (V_{source} - V_{device})/I_{device}$ of a few characteristic devices and found that in general R_C is small (~ 10 kΩ) relative to GNR resistance, and varies little with temperature, as shown in Figure 2.8(b) for one device from set D1. Based on this result, we use a two-terminal geometry for rest of the measurements throughout this dissertation. Note that measurement of R_C diverges at gate voltages near the gap. This is an expected artifact of the measurement; when the GNR resistance becomes comparable to the ~ 10 MΩ input impedance of the voltmeter, V_{device} no longer provides a non-invasive probe of the voltage drop across the device.

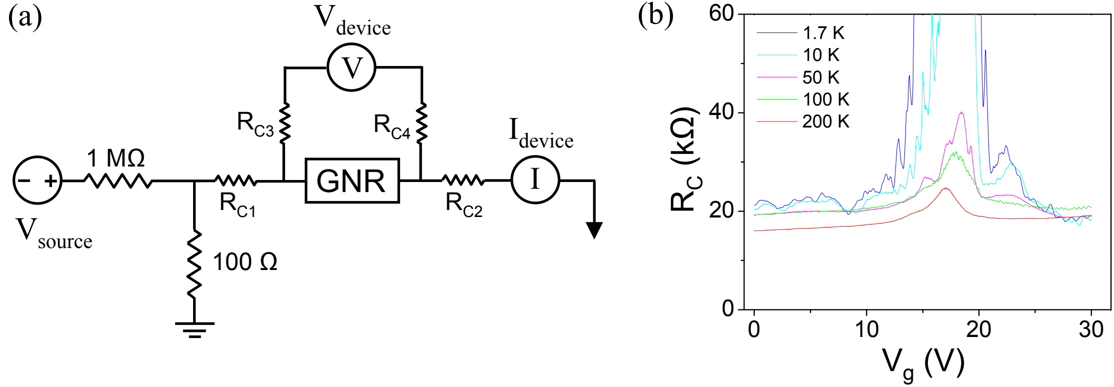

Figure 2.8: Contact resistance measurement of GNRs. (a) Schematic circuit diagram of contact resistance measurement. Total contact resistance $R_C = R_{C1} + R_{C2}$. (b) Contact resistance R_C of one GNR device vs. gate voltage at varying temperatures. This measurement is valid when the GNR resistance is much less than the input impedance of the voltmeter.

2.4 Results and Discussion

2.4.1 Conductance near the charge neutrality point is suppressed

The conductance G of each GNR was measured as a function of gate voltage V_g at different temperatures, as shown in Figure 2.9 for three representative GNR devices of varying width ($W = 24 \pm 4, 49 \pm 5$, and 71 ± 6 nm) and uniform length ($L =$ 2 μm). All curves exhibit a region of suppressed G for a region of V_g. In "bulk" (*i.e.*, unpatterned) graphene, this dip in G is well understood and corresponds to the minimum conductivity $\sim 4e^2/h$ at the charge neutrality point, $V_g = V_{CNP}$, where e and h are the electric charge and Plank constant, respectively.

Unlike the bulk case, GNRs with width $W < 100$ nm show a decrease in G_{min} of more than an order of magnitude at low temperatures. The narrowest GNRs show the greatest suppression of G_{min}. For example, the GNR with $W = 24$ nm

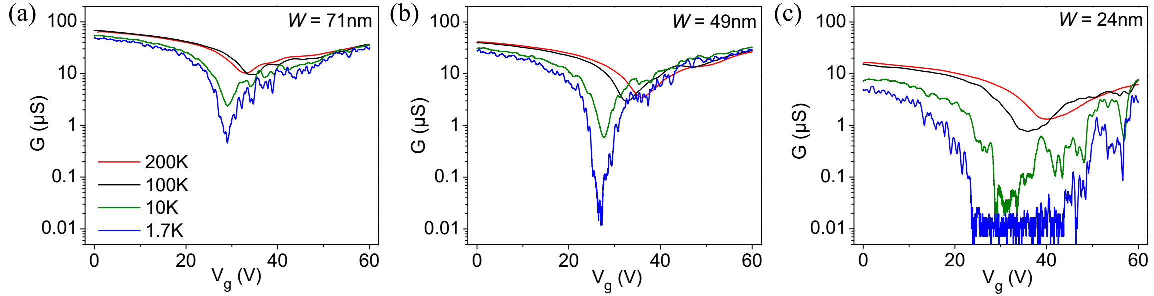

Figure 2.9: Conductance as a function of gate voltage at varying temperatures for three GNRs of different widths. All three devices have $L = 2\mu$m and were etched in parallel from the same sheet of graphene.

(Fig. 2.9(c)), a large "gap" region appears for $\sim 25 < V_g < \sim 45$ V, where G_{min} is below our detection limits ($< 10^{-8}\Omega^{-1}$). This strong temperature dependence of $G(V_g)$ in GNRs is in sharp contrast to that of bulk graphene samples, where G_{min} at the charge neutrality point changes less than 10% in the temperature range 30 mK to 300 K [44]. The suppression of G near the charge neutrality point suggests the opening of an energy gap. We observe (Fig. 2.9) stronger temperature dependence of G for a broader range of V_g values in narrower GNRs, suggesting the presence of larger energy gaps in narrower GNRs.

2.4.2 GNRs have a well-defined sheet conductivity in the on-state

We now turn to the "on-state" conductance, that is, the behavior outside of the gap region, at densities away from the charge neutrality point. In this regime, conductance scales with GNR width.

Figure 2.10(a) shows the conductance G of a set P4 of parallel GNRs, with widths ranging from 14 to 63 nm, at varying temperatures. Here the gate voltage is fixed

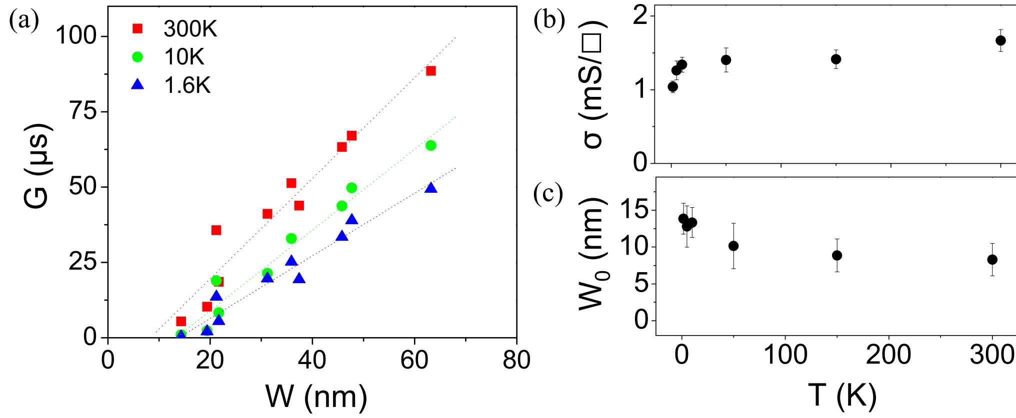

Figure 2.10: Conductance for a set of parallel GNRs in the on-state, away from the gap. (a) Conductance vs. ribbon width for three different temperatures. Dotted lines are linear fits to the data, $G = \sigma L/(W - W_0)$ (b) and (c) show σ and W_0, respectively, from the linear fits at different temperatures.

at $V_g = V_{CNP} - 50\ V$, which corresponds to a hole carrier density of $n \approx 3.6 \times 10^{12}$ cm^{-2}. The conductance at each temperature is well described by the linear fit $G = \sigma(W - W_0)/L$. Here we can interpret σ as the GNR sheet conductivity in the active GNR width participating in transport, $(W - W_0)$. From the fit, the room temperature sheet conductivity for this set of ribbons is $\sigma \sim 1.7$ mS/□ and decreases with decreasing temperature, reaching $\sim$ 75% of its room temperature value at $T = 1.6$ K (Fig. 2.10(b)).

The inactive GNR width W_0 increases from 10 nm at room temperature to 14 nm at 1.6 K. A reduced active channel width in graphene electronic devices was first reported in GNRs fabricated from epitaxially grown multi-layer graphene films [16], where much larger inactive edges ($W_0 \sim 50$ nm) were estimated compared to our GNR samples.

We suggest two possible contributions to W_0. First, W_0 may represent an overestimation of the actual graphene ribbon width, and second, W_0 may represent an region of the ribbon, presumably at the edges, where graphene is present but is not conducting, *i.e.*, localized edge states due to structural disorder from the etching process. To investigate the first contribution, we removed the HSQ etch mask from several GNRs on one device set with a hydrofluoric acid vapor etch, and found that the actual width of the graphene ribbon is often $\sim$ 10 nm narrower than the HSQ protective mask. Difficulties in removing the HSQ etch mask and leaving the device intact prevented us from doing this measurement on more than a few devices in one set. This suggests that the inactive region due to localized edge states is small ($<$ 2 nm) at room temperature and spreads to as much as $\sim$ 5 nm at low temperatures.

2.4.3 The energy gap size is measured with 2-D conductance plot

We now address the "off-state" in the ribbon transport, specifically, the quantitative scaling of the transport gap size.

Figure 2.11(a) shows a schematic energy band diagram for a semiconductor connected to source and drain electrodes, with a charge density controlled by a gate electrode. Here the difference in source and drain potential creates a bias window equal to the applied source-drain bias V_b. The gate voltage adjusts the position of the Fermi level, effectively moving the source-drain levels relative to the gap. When the bias window lies entirely inside the band gap, as shown in the diagram, there are no conducting states are inside the bias window, and charge cannot flow. There is no conduction, so the device is "off". As the bias window increases, the source and

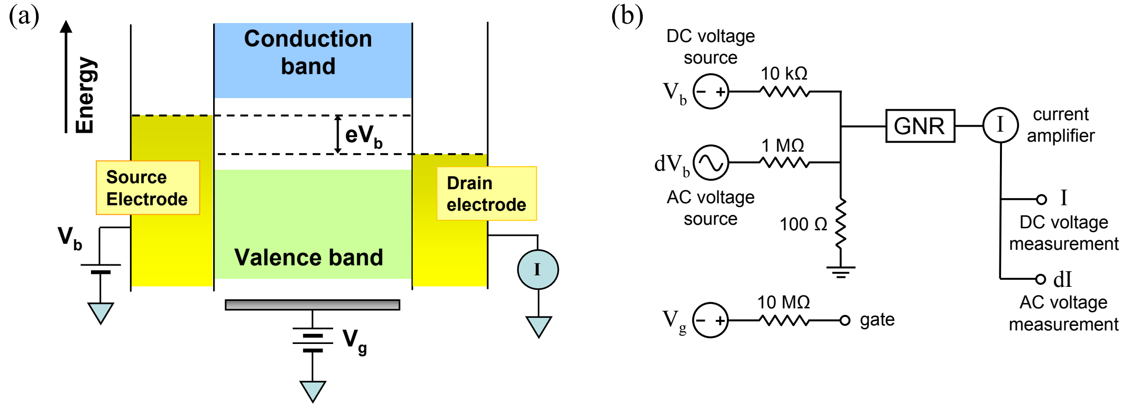

Figure 2.11: (a) Cartoon of a semiconducting material connected to source and drain electrodes and a gate electrode. (b) Schematic circuit diagram of the 2-D conductance measurement.

drain potentials approach the valence and conduction band edges. When the bias window contains part of the conduction or valence band, we measure electron or hole transport through the device, and the device is "on".

For a complete view of the low temperature transport behavior of the GNRs, we measure differential conductance as a function of both gate voltage and source-drain bias voltage with a setup shown schematically in Figure 2.11(b). With this technique we simultaneously measure DC current I and the small signal AC current dI for the full range of V_g–V_b space. This results in a 2-D plot of the differential conductance dI/dV_b which describes the low-bias device response, as well as DC I–V_b curves for each value of V_g, from which we obtain a direct measure of the size of the energy gap.

Figures 2.12(a)–(c) shows 2-D plots of differential conductance for three representative GNR devices, with widths $W = 22$, 36, and 48 nm. The color indicates dI/dV_b on a logarithmic scale. The blue region in each plot corresponds to the "off" state,

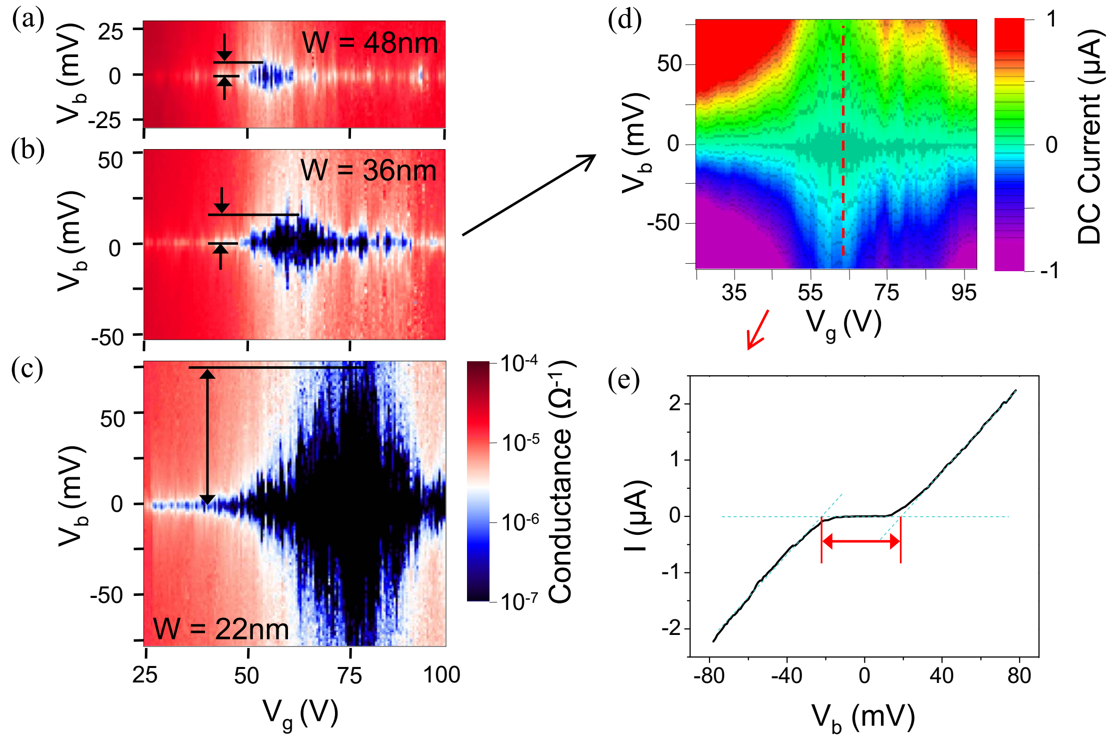

Figure 2.12: (a)–(c) 2-D plots of differential conductance dI/dV_b for three GNRs of different widths, with ΔV_b highlighted by the black arrows. The color scale is the same for all three plots. (d) DC current vs V_b and V_g for the same device as in (b). Dashed red line at $V_g = 64\ V$ shows the constant V_g cut plotted in the following panel. (e) Procedure for extracting $E_{gap} = e\Delta V_b$ from an I–V_b plot taken with the gate voltage fixed at the center of the gap. Black curve is the data, dashed blue lines show $y = 0$ and linear fits to the curve away from the gap, red lines show $2\Delta V_b$.

when both band edges are outside of the bias window. The diamond shape of this region indicates that both V_b and V_g adjust the position of the band edges relative to the source and drain energy levels, analogous to nonlinear transport in quantum dots [45]. The fluctuations and deviations from a clean diamond outline are indicative of a more complex band structure and will be discussed in more detail in Chapter Three.

The value of $V_b = \Delta V_b$ at the vertex of the diamond gives a direct measure of the size of the energy gap. The upper vertex of the diamond $V_b = \Delta V_b$ corresponds to a bias voltage configuration with the source at the edge of the valence band and the drain at the edge of the conduction band, and the lower vertex $V_b = -\Delta V_b$ corresponds to the opposite scenario, with the source at the conduction band and the drain at the valence band, so that the total height of the diamond $2\Delta V_b$ corresponds to twice the gap size. This is a somewhat simplified picture of the GNR band structure, and a more subtle interpretation will be given in Chapter Three.

Figures 2.12(d)–(e) show the method used in this chapter for extracting gap size. Fig. 2.12(d) shows a plot of DC current I vs V_b and V_g. We take a constant gate voltage cut at the largest part of the diamond-shaped "off" region, giving the non-linear I–V_b curve shown in Figure 2.12(e). We then take a linear fit to this curve far from the gap, and extrapolate to the $y = 0$ axis. Following Figure 2.11(a), we take this voltage difference ΔV_b to be potential difference when the source and drain levels are at the edges of the gap. Multiplying this potential by electron charge e gives $E_{gap} = e\Delta V_b$ in units of energy. It should be noted that this "gap" value corresponds

to the energy band gap only when the entire GNR behaves like a clean semiconductor. Other methods of evaluating the "gap" in cases with disorder induced intragap states or in systems consisting of quantum dots will will be discussed in Chapter Three.

2.4.4 The energy gap follows a robust scaling rule

Motivated by the expected gap scaling $E_{gap} \sim 1/W$ [35, 40, 41], we plot E_{gap}^{-1} against W in Figure 2.13(b) for the 13 parallel GNRs in set P4. A linear fit (dashed line) results in the scaling law

$$E_{gap} = \alpha/(W - W^*) \tag{2.1}$$

where we obtained $\alpha = 0.2$ eV $\cdot$ nm and $W^* = 16$ nm from the fit. The density functional theory study in Ref [41] calculates values of α ranging from 0.3 to 1.5 eV $\cdot$ nm, depending on crystallographic orientation and edge termination. Reasons why the measured value may be smaller than theoretical predictions will be discussed in Chapter Three, for now let us say that the form of our scaling law matches with theory, with agreement of the coefficient to within an order of magnitude. We also note that $W^* \approx W_0$, in good agreement with the independent estimation of GNR edge effects above.

A similar scaling behavior holds even across GNR device sets made from different pieces of graphene and running in different crystallographic directions. Figure 2.13(a) shows the overall scaling of E_{gap} as a function of W for six different device sets. Four device sets (P1–P4) have parallel GNRs with W ranging from 15–90 nm, and two device sets (D1, D2) have GNRs with similar $W \approx 40$ nm and 25 nm, respectively,

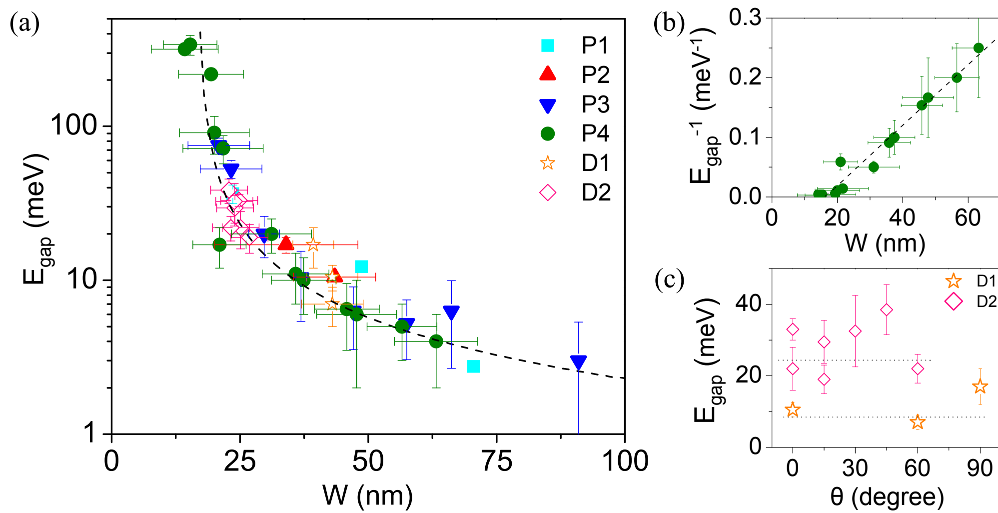

Figure 2.13: Scaling of GNR energy gaps as a function of width and crystallographic orientation (a) Energy gaps E_{gap} for 36 ribbons from 6 different device sets as a function of ribbon width. Dotted line is a fit to $E_{gap} = \alpha/(W - W^{*})$. (b) Inverse energy gap vs ribbon width for the parallel GNRs in device set P4. Dashed line is a linear fit. (c) Energy gaps as a function of relative crystallographic orientation for two device sets. Dashed lines show E_{gap} for each set as predicted by the average ribbon width of that set.

but different crystallographic directions. The energy gap behavior of all devices is well described by the scaling $E_{gap} = \alpha/(W - W^*)$ as discussed above, indicated by the dashed line. Remarkably, energy gaps as high as $\sim$ 200 meV are achieved by engineering GNRs with widths $W \sim 15$ nm. Based on the empirical scaling determined here, a narrower GNR may show an even larger band gap, making the use of GNRs for semiconducting device components in ambient conditions a possibility.

Following other theoretical predictions, we also check the fit to other predicted scaling laws. Figure 2.14 shows four different possible scaling rules. We now fit to the data from all six device sets, as opposed to the the fit above, which was done only on the data from set P4. In each panel, the blue curve corresponds to a fit that was performed on E_{gap}^{-1} vs W first, then inverted before plotting. The red curve is the result from fitting to the data directly. In general, fitting to the inverse produces more accurate results, mostly because it gives more weight to the long "tail" for higher W values, where E_{gap} values vary more slowly. The parameters listed in the figure correspond to the blue curve.

The first fit, shown in Fig. 2.14(a), is the same fit as in Equation 2.1, only now the fit is done on the full data set, producing slightly different values. Panel (b) of the same figure shows the prediction from Ref. [41] mentioned in Section 2.1, that is, $E_g(W) = aW^{-b}$. Here we have modified the expression to accommodate W^*, resulting the relation

$$E_g(W) = a(W - W^*)^{-b} \tag{2.2}$$

Note that this is only a more general expression of Equation 2.1, so we expect the

quality of the fit to be at least as accurate. In fact, we find that the exponent b, which is a free parameter, turns out to be very close to -1, producing a result similar to that in panel (a). In panel (c), we fix the exponent to $b = -2$, and obtain a negative value for W^* that does not make physical sense in the context of our physical picture for W^*.

In Fig. 2.14(d), we address Ref. [46], a theoretical study that was done in response to our publication of the results in this chapter. In this work, the authors propose the scaling rule

$$E_g \approx \frac{e^2}{W} e^{-W/W_0} \tag{2.3}$$

This corresponds to a physical picture where the energy gap is caused by Coulomb blockade in constrictions throughout the ribbon. In their publication, they show a fit to our data, using only the data for $W > 20$ nm. Here we fit the expression to the complete data set, and, as for Eqn. 2.2, we add another free parameter W^*. Fig. 2.14 shows the resulting fit, which describes the gap size well for wider ribbons, but does not capture the scaling behavior for the full range of widths.[2] It should be noted that although this particular model for energy gap size resulting from Coulomb blockade does not fit the data well, this does not rule out Coulomb blockade as an explanation for the origin of the energy gap.

Finally, we remark on the crystallographic directional dependence of E_{gap}. Figure 2.13(c) shows E_{gap} versus the relative orientation angle θ for two sets of GNRs.

[2]Note that with three free parameters, this fit did not converge. If we set $W^* = 0$, we get a converging fit with $\alpha = 631.5$ meV $\cdot$ nm and $W_0 = 90.49$ nm, although here we have plotted the non-converging three parameter fit, which is a better match to the data.

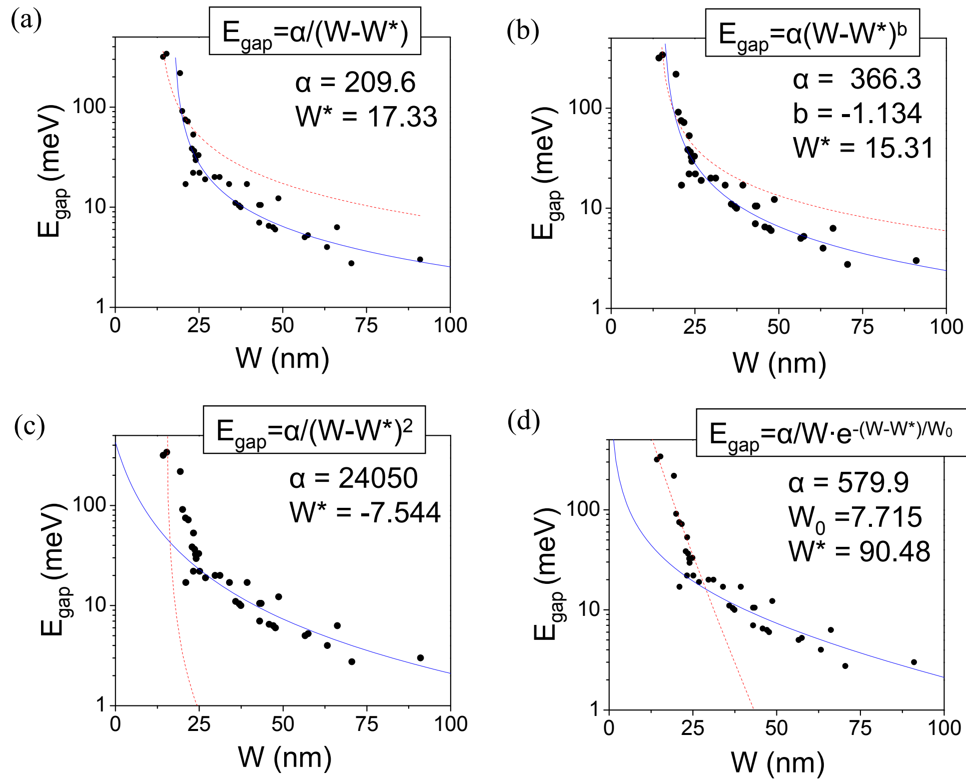

Figure 2.14: Comparison of different scaling rules, fit to data from all six device sets. Blue curve corresponds to the result from fitting to E_{gap}^{-1}, red curve is the result from fitting to the data directly. Parameters listed in each panel are for the blue curve. In all panels, α has units of meV $\cdot$ nm, W^* and W_0 have units of nm, and b is a dimensionless exponent.

Each of these ribbon sets contains GNRs of nearly the same width cut in different crystallographic directions from the same sheet of graphene. In principle, we expect $E_{gap}(\theta)$ for each set to be periodic in θ, provided all GNRs have similar edge structures. However, experimental observation shows randomly scattered values around the average E_{gap} corresponding to the ribbon width W and no sign of crystallographic direction dependence. In fact in Figure 2.13(a) we see that the the variations in E_{gap} between the ribbons in the ribbons in these two sets (D1 and D2) can be more readily attributed to small variations in W. In this sense, the scaling rule $E_{gap} = \alpha/(W - W^*)$ is more robust than expected, since it holds for ribbons with all crystallographic orientations.

This lack of crystallographic orientation dependence suggests that detailed edge structure plays a more important role than the overall crystallographic direction in determining the properties of the GNRs. Indeed, theory for ideal GNRs predicts that E_{gap} depends sensitively on the boundary conditions at the ribbon edges [28, 35, 36, 39–41]. The lack of directional dependence indicates that our device fabrication process does not give us the atomically precise control of the GNR edges necessary to reveal this effect.

2.5 Conclusions

In this chapter we have demonstrated the fabrication and measurement of graphene nanoribbons. We observe that GNRs have an energy gap that scales inversely with ribbon width and has no dependence on crystallographic orientation. Transport away

from the gap region shows a well-defined, temperature dependent, sheet conductivity. A temperature dependent non-conducting region of the width exists in addition to errors in width measurement, and is detectable both in on-state measurements (as W_0) and energy gap measurements (as W^*).

Chapter 3

Electronic transport in disordered graphene nanoribbons

In this chapter, we delve deeper into the behavior of transport in graphene nanoribbons and explore the nature of the energy gap with more detailed experiments. We will see that the energy gap we found in the previous chapter is not, in fact, a “band gap”, in that it is not a range of energies for which there are no electronic states. For this reason we will henceforth refer to the energy gap more generally as a “transport gap”.

From the work in the previous chapter, we have already seen that etched graphene nanoribbons behave quite differently from the theoretical predictions for perfect ribbons. The orientation-independence of our empirical scaling law was an early indication of a more subtle system at play; another clue was the existence of a nonconducting ribbon width $W_0 \sim W^*$. Other indications that the transport gap is not a band gap (to be discussed in more detail in this chapter) include the existence of reproducible

conductance peaks at energies inside the gap, and a thermally activated behavior which deviates from the simple thermal activation of a clean band gap. These inconsistencies between the data and our simple band gap model motivate the experiments described in this chapter.

In particular, we would like to answer the following questions: what is the nature and origin of the conducting states inside the transport gap, and what is the dominant transport mechanism at energies inside the gap?

Here we explore these question by carrying out both low temperature and temperature dependent transport measurements on GNRs of various dimensions. From the scaling of several characteristic energies with GNR width (W) and length (L), we find evidence of a transport mechanism in disordered GNRs based on hopping through localized states whose size is close to the GNR width. The work in this chapter was published as Ref. [47].

3.1 Background

3.1.1 Models for transport in disordered systems

Before discussing graphene devices specifically, we review some models for electronic transport in disordered systems and systems with reduced dimensions.

Anderson localization and variable range hopping in disordered systems

Anderson localization describes conduction in materials with random disorder. In this model, if the disorder strength surpasses a certain threshold, diffusive electron transport can no longer occur. Electron wavefunctions become localized, and transport occurs by means of "quantum jumps between localized sites." [48] At low enough carrier densities, where no diffusion can take place, the system is described as an "Anderson insulator". The critical energy separating localized states, through which transport occurs only by hopping, and extended states, through which electrons can travel by diffusion, is known as the "mobility edge" [49].

A consequence of Anderson localization is transport by means of variable range hopping (VRH), the mechanism for conduction by hopping between localized states. Conductance G is proportional to the jump frequency between states:

$$G \propto e^{-2R/\xi} e^{-\Delta E/k_B T} \tag{3.1}$$

where ξ is the localization length of a state, R is the distance between two states, and ΔE is the energy difference between two states [50]. Conduction by hopping though localized states is a balance between the probability of hops with a large energy cost ΔE and hops with a large distance R. As temperature increases, hops with a larger energy difference become more favorable and hopping distance decreases, hence the term "variable range" hopping.

We seek an expression for $G(T)$ as a function of T only, one that does not explicitly invoke R, ΔE, or ξ. $G(T)$ must also depend on the dimensionality of the system

through the number of available states with a given energy difference. Assuming a constant density of states $g(\epsilon) = N_0$, we have, in 3 dimensions:

$$\frac{4}{3}\pi R^3 = \frac{1}{N_0 \Delta E} \tag{3.2}$$

More generally, in d dimensions:

$$\Delta E \propto \frac{R^{-d}}{N_0} \tag{3.3}$$

So that equation 3.1 can be written:

$$G \propto e^{-2R/\xi - R^{-d}/N_0 k_B T} \tag{3.4}$$

Conductance will occur at the hopping distance R which maximizes the exponent, and therefore maximizes the hopping probability:

$$R = \left(\frac{2N_0 k_B T}{\xi d}\right)^{-1/(d+1)} \tag{3.5}$$

So that the conductance as a function of temperature is given by:

$$G \propto e^{-(T_0/T)^{1/(d+1)}} \tag{3.6}$$

This expression for the VRH temperature dependence, derived assuming a constant density of states, is known as Mott variable range hopping [50]. We can write a more

general expression

$$G \propto \exp(-(T_0/T)^\gamma) \tag{3.7}$$

where γ depends on the dimensionality of the system and the form of the density of states; for Mott-VRH, $\gamma = 1/(d+1)$. Other VRH models have been derived with different forms of $g(\epsilon)$ and hence γ. The Efros-Shklovskii VRH [51] picture incorporates long-range Coulomb interactions in the form of a soft Coulomb gap in the density of states. The form $g(\epsilon)$ of the soft Coulomb gap depends on d, and has a zero only at $\epsilon = 0$. The "2/5" law [52] describes systems with very low impurity concentrations, which results in a different form of $g(\epsilon)$, again depending on d. Fogler *et. al.* [52] summarize the VRH exponents seen for various systems as $\gamma = (\mu+1)/(\mu+d+1)$, with $g(\epsilon) = \epsilon^\mu$.

To summarize, variable range hopping is a single particle picture of electron transport through disorder-induced localized states, and has a temperature dependence with an exponent that depends on dimensionality and other properties of the system.

Coulomb charging effects in transport through small systems

In certain systems, the electrostatic effects of adding or removing one electron has measurable consequences. If charging effects prevent the tunneling of electrons to an island of charge, "Coulomb blockade" is said to prohibit transport through the system.

Single-electron charging effects become relevant if two conditions are fulfilled. First, we consider the charging energy, the change in electrostatic potential associated

with adding or removing one electron, $E_C = e^2/C$, where C is the total capacitance between the charge island and the rest of the system. For charging effects to be relevant, this charging energy must be larger than the thermal energy $k_B T$:

$$e^2/C >> k_B T \tag{3.8}$$

Second, the charge island must be well isolated from the rest of the system, so that the amount of charge on the island is a well-defined quantity. To meet this condition, the resistance of the leads R_l must be large enough such that quantum fluctuations of the number of charges on the island are less than one. Equivalently, the fluctuations in the energy of the island must be small compared to $E_C = e^2/C$. From the Heisenberg uncertainty relation $\Delta E \Delta t > h$ with the time to discharge the island $\Delta t = R_l C$, this requires that

$$R_l >> h/e^2 \tag{3.9}$$

In systems where the level spacing in the island can be neglected, if conditions 3.8 and 3.9 are met, the charging energy E_C is the addition energy for adding or removing one electron to the island. At constant gate voltage, this addition energy prevents an electron from tunneling to the island, resulting in the barrier to transport known as Coulomb blockade [53]. Transport in such a system can only occur when an applied bias voltage is large enough to overcome the charging energy, $V_b > E_C/|e|$, or when a change in gate voltage changes the electrostatic potential of the island by an amount corresponding to one electron charge, $\Delta V_g = eE_C = e/C_g$. In small, semiconducting

systems where the single particle level spacing ΔE within one island is relevant, the addition energy becomes $\Delta E + E_C = \Delta E + e^2/C$.

3.1.2 Previous work on disordered graphene nanoribbons

Disordered GNR theory

Since the publication of the experiment presented in Chapter Two [31], many theoretical works on GNRs with imperfect edges have been published. The bulk of the theoretical literature on transport in disordered graphene nanoribbons describes a picture based on edge disorder induced localization of the electronic wavefunctions. [54–60]. Most of those works explicitly describe this effect as a phenomenon similar to Anderson localization, though it is noted that a transport gap due solely to Anderson localization would have a size dependent on the strength of disorder only [55], in contrast to the width dependence seen experimentally.

Querlioz *et. al.* [54] calculate the local density of states (LDOS) for perfect GNRs and for GNRs with disordered edges. In perfect GNRs, they predict a band gap with a size dependent on the crystallographic orientation of the ribbon, consistent with earlier predictions such as those in Refs. [39–41]. In contrast, for disordered GNRs, they find no gap in the density of states, but instead predict a threshold energy below which electron wavefunctions are localized. Above the threshold energy, localization lengths increase rapidly, and most states are extended along the ribbon. The authors interpret this threshold energy as a "quasi-mobility edge".[1] They calculate values

[1] The mobility edge is referred to as "quasi" because in this case the transition is due in part to the finite size of the system, whereas in a traditional Anderson insulator the transition between

of the quasi-mobility edge for ribbons of different crystallographic orientations, and find that the orientation dependence is almost completely suppressed in GNRs with as little as two rows of edge disorder. They find a scaling law for the quasi-mobility edge in disordered GNRs which matches the experimental rule we found in Chapter Two, $\alpha/(W - W^*)$, with a value of $\alpha = 4$ eV $\cdot$ nm.

Martin and Blanter [55] reconcile the the Anderson localization picture with the width dependence of the gap by demonstrating an edge disorder induced segmentation of the wavefunctions into blocks with lengths on the same order as the ribbon width. This leads to a conductivity at energies inside the gap which is governed by variable range hopping at low temperatures and crosses over to nearest-neighbor hopping at higher temperatures. Similarly, Evaldsson *et. al.* [56] calculate the LDOS in edge disordered ribbons, and find "hills and canyons" in the density of states which can extend across the width of the ribbon, blocking conductive paths.

Yoon and Guo [57] do not explicitly use the language of Anderson localization, but they find that edge variation "induces localized states in the band-gap energy range" which leads to an increase of current in the "off" state of the device.

A different description of the origin of the transport gap is offered by Sols *et. al.* [46] in a paper written in response to our work Ref. [31]. In this work, the authors suggest that edge disorder leads to the formation of "necks" in the ribbon, so that the GNR behaves as a series of quantum dots. The transport gap is then a result of Coulomb blockade. Their prediction for a scaling law was seen in Chapter Two to match poorly with experiment (see Eqn. 2.3 and Fig. 2.14(d)), but this does not

localized and extended states is purely a function of disorder strength.

preclude the possibility that Coulomb blockade plays a significant role in the formation of the transport gap.

One theoretical work, Ref. [61], focuses on density inhomogeneity [62] as the origin of the transport gap, and describes the gap behavior in terms of a percolation metal-insulator transition. They apply this model to our data from Ref. [31] and their own data, and find that ribbons with widths as narrow as ($W = 24$ nm) are in the 2D percolation universality class. They predict that further decreasing ribbon width would reveal a transition to 1D percolation.

Experimental studies of the transport gap in etched GNRs

Several experimental works on graphene nanoribbons address questions about the nature of the transport gap and conduction at energies inside the gap, in particular from the ETH group [63, 64], the Stanford group [65, 66], and the Delft group [67]. The physical models proposed in these works are heavily influenced by the existence of Coulomb diamond-like features in the data (see also Refs. [68–73]) which necessitate a many-body picture, inconsistent with the Anderson localization model favored by theorists.

Stampfer *et. al.* [64] propose a picture where the gap is formed by quantum dots along the ribbon. The suggest that these quantum dots are the result of a quantum confinement-induced energy gap in the presence of strong potential fluctuations from edge and bulk disorder, so that there are two energy scales to the problem. One energy scale, the characteristic charging energy, is related to measurements of the gap using the source-drain bias, and is of the order of 10 meV. The other, the disorder potential

strength, is seen in measurements using the gate voltage, and is of the order of 100 meV.

Collaborators Molitor *et. al.* [63] suggest that the width dependence of the gap measured by gate voltage is due to an increase in the probability of a percolating conductive path through the constriction. They also find that the gaps measured both with gate voltage and with source-drain bias are dependent on ribbon length, and claim that these two measures of the gap have a proportionality independent of constriction geometry.

Gallagher *et. al.* [66] carry out a careful study of transport in GNRs with constant width and varying length, and support a similar physical picture with two energy scales corresponding to the energy gap measured with gate voltage or with source-drain bias. The authors study the effect of annealing the ribbons, and find that annealing reliably decreases the gate voltage gap, but changes the source-drain gap unpredictably and non-monotonically. They also find a correlation between the size of the gate voltage gap and its distance from zero volts. Taken together, they interpret this as evidence that potential inhomogeneity in the device is due to charged impurities of a dominant sign. Finally, in one particular GNR device, for a range of carrier densities, they see the periodic conductance oscillations of Coulomb blockade. They interpret this as evidence that conductance peaks in GNRs are due to Coulomb blockade even in cases where the peaks do not have regular spacing, and they assert that, because of their periodic spacing, the peaks cannot be caused by Anderson localization.

Liu *et. al.* [67] measure a different device geometry, a graphene nanoribbon with a local top gate, so that the density under the top-gated center region of the ribbon can be controlled independently from the density of the ribbon on either side of the top gate. When the top gate region and the outside regions are oppositely doped, so that the devices are in either the *pnp* or *npn* configuration and *p*-*n* junctions are formed at the edges of the top gate, the devices exhibit Coulomb blockade with an addition energy dependent on the area of the gated region, as expected. When the center and the outside regions are doped with different densities of the same polarity (*pp'p* or *nn'n*), Coulomb blockade is seen with an addition energy that is largely independent of the device dimensions. The authors attribute this Coulomb blockade in the absence of *p*-*n* junctions to Anderson localization of carriers due to edge and substrate induced disorder. The addition energy in the case of unipolar doping is lower than that in devices with *p*-*n* junctions, implying that the charge island size in the unipolar case is larger than the size of the top-gated region.

3.2 Experiment

The experiments in this chapter differ from those discussed above in that: (1) We measure the temperature dependent behavior, which is essential to understanding transport mechanisms in disordered systems. (2) As a result of (1), we find two additional characteristic energy scales to the problem, both related to temperature dependent transport. (3) We draw from a large sample size of devices with a range

of widths and lengths, allowing us to identify trends in the data, and (4) We directly compare transport in devices engineered to have difference capacitive coupling, allowing us to quantify the contribution of electrostatic charging effects.

GNRs were fabricated following the procedures described in Chapter Two. Most experiments in this chapter were performed on back-gated GNRs on a substrate of highly doped silicon with a 285 nm thick SiO_2 gate dielectric, like the devices in Chapter Two. We measured electron transport in a total of 41 of these back-gated GNRs with $20 < W < 120$ nm and $0.5 < L < 2$ μm. Additionally, we fabricated dual-gated GNRs for a comparative study, as shown in Figure 3.1. Dual-gated GNRs have both a top gate and a back gate. These were fabricated by first making a back-gated GNR, then depositing a 10 nm gate dielectric of HfO_2 using atomic layer deposition (ALD), then patterning metal top gates. The final gate dielectric material consists of the ≈ 15 nm HSQ etch mask, which is not removed from the ribbons, in addition to the 10 nm HfO_2. The top gates are designed to cover both the entire ribbon and part of the wide "graphene lead", in order to avoid the formation of p-n junctions in the channel. For the comparative study in this chapter, dual-gated and back-gated devices were fabricated together from the same sheet of graphene, so that back-gated devices are also covered with the same HfO_2 dielectric, controlling for any surface phonon or screening effects. Since the entire graphene nanoribbon is covered with HSQ, HfO_2, and metal in the completed device, is it difficult accurately measure the width of the underlying ribbon. In this experiment we control for width by making several ribbons on one graphene flake with identical design and electron beam dose,

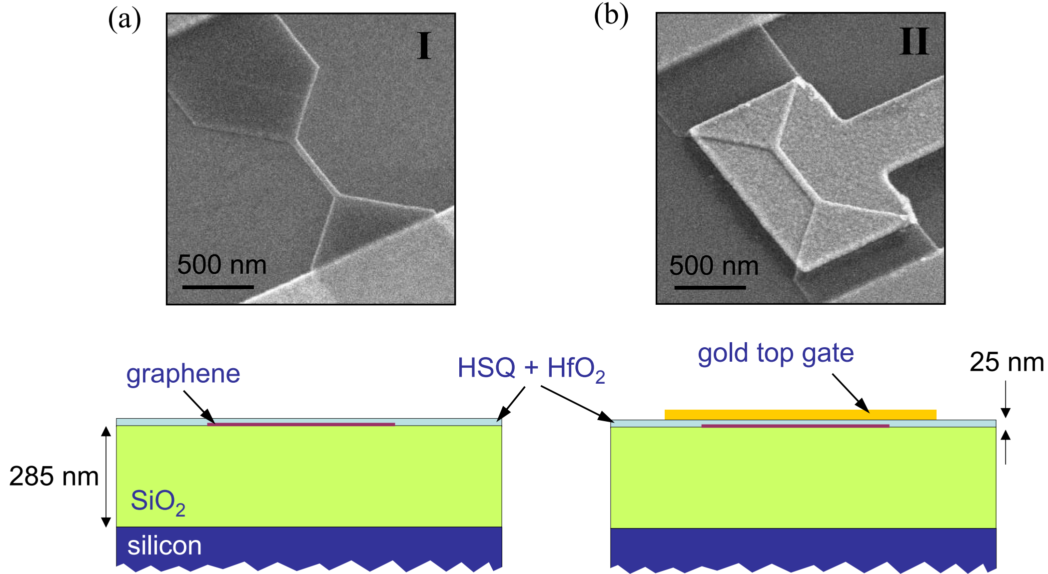

Figure 3.1: Scanning electron microscope (SEM) images (top) and side view cartoons (bottom) of (a) a back-gated device and (b) a dual-gated device used in the comparative study.

then covering half of the ribbons with top gates.

3.3 Results and discussion

3.3.1 Signatures of hopping transport

Characteristic energies emerge from V_g and T dependent measurements

First we examine GNR transport as a function of back gate voltage V_g and extract a measure of the transport gap in terms of V_g. As in Chapter Two, we measure differential conductance $G = dI/dV$ at low source-drain bias using a standard lock-in technique. Again, GNR conductance is strongly suppressed for a region of back gate voltages V_g near the graphene charge neutrality point, suggesting the formation of a

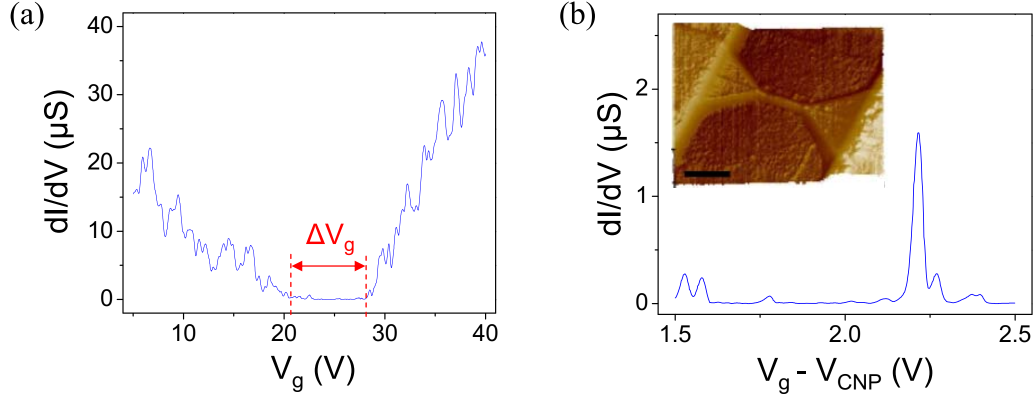

Figure 3.2: (a) dI/dV of a GNR with $W = 36$ nm and $L = 500$ nm, plotted as a function of V_g. Dashed lines highlight measurement of ΔV_g. (b) A close-up of dI/dV within the gap regime plotted as a function of $V_g - V_{CNP}$, where $V_{CNP} = 21$ V is the gate voltage for the charge neutrality point. Inset shows an atomic force microscope image of the device, scale bar is 500 nm.

transport gap. Fig. 3.2(a) shows dI/dV as a function of V_g for a typical GNR. The transport gap region as measured in back gate voltage, ΔV_g, can be identified in this curve by extrapolating the smoothed dG/dV_g to zero [63]. We note that reproducible conductance peaks appear in the gap region, as shown in Fig. 3.2(b) [63, 64, 66]. These peaks are indicative of resonant conduction paths through localized states inside the transport gap. In general, resonance peaks in the gap are less than 10 % of the G values outside of the gap region.

The observed transport gap, ΔV_g corresponds to an energy in the single particle energy spectrum: $\Delta_m = \hbar v_F \sqrt{2\pi C_g \Delta V_g/|e|}$, where $v_F = 10^6$ m/sec is the Fermi velocity of graphene [74] and C_g is the capacitive coupling of the GNR to the back gate. This geometric capacitance is strongly dependent on ribbon dimensions, and we calculate it using a finite element model, obtaining, for example, $C_g = 69.0$ nF/cm^2

and $\Delta_m = 200$ meV for the particular device in Fig 3.2.

At energies inside the transport gap but away from the resonant conduction peaks, the conductance is strongly suppressed, and the charge transport can be described by thermally excited hopping. We study the thermal activation of the off-resonant conduction in this regime by measuring G_{min}, the minimum conductance for a given sweep of gate voltage V_g, at different temperatures as shown in Figure 3.3(a). Figure 3.3(b) shows an Arrhenius plot for $G_{min}(T)$. Evidently, thermally excited transport exhibits two distinct regimes of behavior, separated by a characteristic temperature T^*. At high temperatures ($T > T^*$), the transport is simply activated: $G_{min} \sim \exp(-E_a/2k_BT)$, where $E_a = 285$ K is obtained from a linear fit of the Arrhenius plot for the device in Fig. 3.3(b) (dashed red line). At lower temperatures ($T < T^*$), however, G_{min} deviates from the simple activation behavior and decreases more slowly with decreasing temperature than the activated transport would imply. In this low temperature regime, the overall behavior is consistent with variable range hopping (VRH), where $G \sim \exp(-(T_0/T)^\gamma)$, with $\gamma = 1/2$ and a constant T_0, determined by the characteristics of the localized states. This fit is shown as the dotted blue line in Figure 3.3(b), though we note that reliable determination of γ and T_0 requires decades of data on both axes, which is not available here due to the limited temperature range exhibiting this behavior and our experimental limitations for very low currents and temperatures.

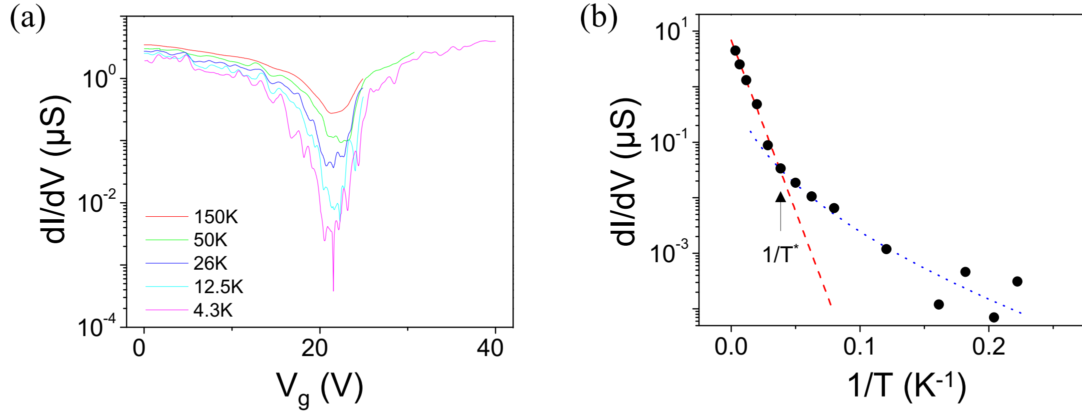

Figure 3.3: (a) $dI/dV(V_g)$ at several temperatures for the same device as in Fig. 3.2 (b) T dependence of the minimum conductance at each temperature. The red dashed and blue dotted lines are a fit to simple activated behavior and variable range hopping with $\gamma = 1/2$, respectively. An arrow highlights the position of T^*.

Scaling of GNR parameters suggests hopping transport through localized states

The GNR transport gap and temperature dependent characteristics described above are typical of all GNRs with $W \lesssim 80$ nm, so that Δ_m, E_a, and $k_B T^*$ can be determined for each of these narrow GNRs. These three representative energy scales are plotted as a function of W in Figure 3.4.

In this graph, we note that (i) there is a clear separation between these energy scales, setting a general relation: $\Delta_m > E_a > k_B T^*$ for given W; (ii) Δ_m, E_a, and T^* depend sensitively on W but not L (iii) the energy scales are reasonably well described by inverse proportion to the lateral confinement of the GNR.

The length independence of each of these parameters can be observed by comparing values from GNRs with similar W but different lengths L. In Figure 3.4,

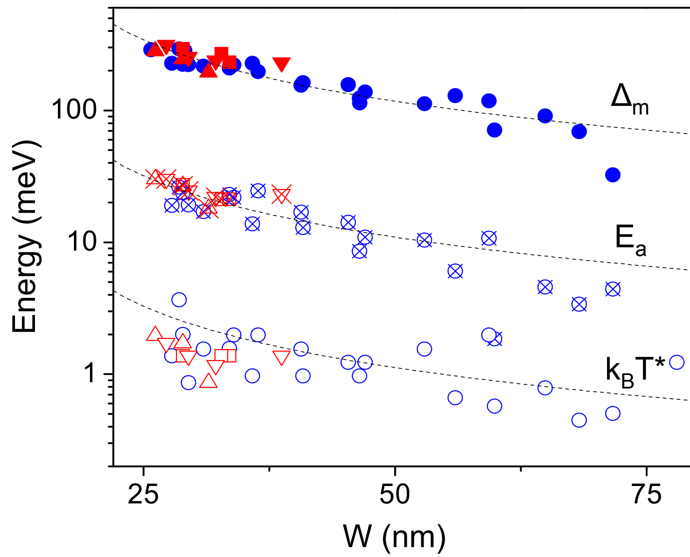

Figure 3.4: GNR transport energy scales : Δ_m (solid), E_a (crosshatched), and k_BT^* (open) plotted as a function of GNR width. Blue circles correspond to ribbons with $L = 0.5$ μm. Red triangles, squares, and stars correspond to ribbons ribbons with L =1, 1.5, and 2 μm, respectively. The dashed lines are the fits described in the text.

the blue circles represent ribbons with $L = 500$ nm, and the red symbols represent longer ribbons, up to $L = 2$ μm, as detailed in the figure caption. The longer GNRs fit the same W dependent trend as the short ribbons, indicating these parameters have no strong dependence on ribbon length. This suggests that these three energy scales are 1D intensive properties of GNRs. To show this, we define the normalized width $w = (W - W_0)/a_0$, where $a_0 = 0.142$ nm is the carbon-carbon bond length and W_0 is an offset introduced phenomenologically, in agreement with Chapter Two and Ref. [31]. Then, we find that all energy scales can be reasonably fit (dotted lines): $\Delta_m = \Delta_m^0/w$; $E_a = E_a^0/w$; $T^* = T_0^*/w$ with the proportionality parameters $\Delta_m^0 = 36.3$ eV, $E_a^0 = 3.39$ eV, and $k_B T_0^* = 0.347$ eV, respectively, with $W_0 = 12$ nm held fixed for all three fits.

Edge disorder in the GNRs tends to induce wavefunction localization, with a localization length that decreases rapidly with decreasing energy, resulting in a transport gap with strongly localized states at energies between the mobility edges [58]. The size of this mobility gap is larger than the clean band gap of an ideal ribbon; Querlioz *et. al.* calculate the scaling prefactor $\Delta_m^0 \approx 32.2$ eV, averaged over many configurations of edge disorder [54]. The close match of our data to theoretical prediction supports the view that atomic defects at the graphene edges create localized states. We point out, however, that the observed energy scales lie within the range of disorder potential fluctuation created by the charged impurities in the SiO_2 substrate [61], making it difficult to exclude the contribution of a substrate disorder induced transport gap, as discussed in a recent experiment on transport in thermally annealed GNRs [66].

On the other hand, $E_a^0/\Delta_m^0 \approx 0.1$; that is, the activation energy at higher temperatures is an order of magnitude smaller than Δ_m. This observation excludes the scenario that extended states carry current via thermal activation across the transport gap. Instead, we interpret the simply activated behavior as a signature of 1D nearest neighbor hopping (NNH) through localized states within the transport gap [55]. In this picture, disorder at the edges tends to produce a rapid variation in the local density of states over the whole width of the ribbon, blocking the conductive paths and leading to a quasi-1D arrangement of localized states [56]. Martin and Blanter predict [55] that the energy spacing between nearest neighbor states is determined by $\sim t'/w$, where $t' \approx 0.2t$ is the hopping matrix element between second nearest neighbor carbon atoms in graphene, so that $E_a^0 \sim 2t' = 1.2$ eV. Our measured value for this scaling prefactor, 3.39 eV, is somewhat larger than this prediction, which may be explained by the contribution of a charging energy to the hopping energy E_a, discussed in more detail in Section 3.3.3.

The change of the transport behavior across the temperature T^* allows a further comparison of our data to theory. In a recent theoretical work, the NNH and VRH crossover is calculated to occur at $T^* = E_a/k_B\alpha$, where $\alpha \approx 8$ was estimated numerically [75]. In our experiment, we obtain $E_a^0/k_BT^* = 9.8$, reasonably consistent with this theoretical prediction, lending further support to a model of charge transport via thermally activated hopping between localized states.

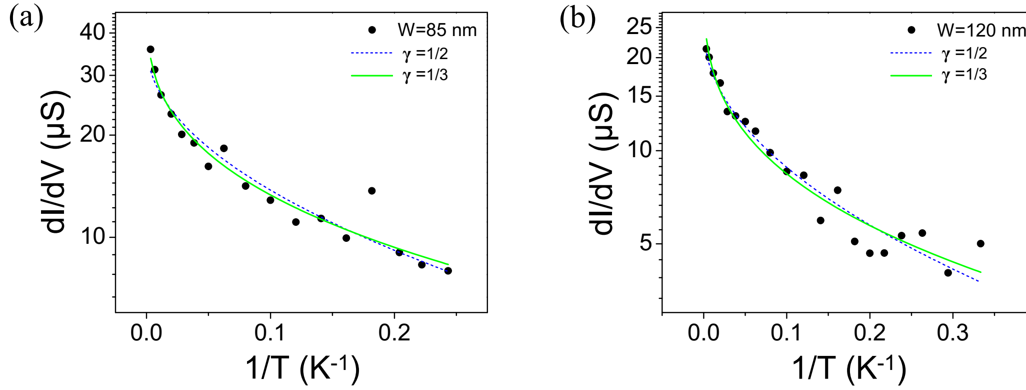

Figure 3.5: Two representative examples of temperature dependent behavior in wide ribbons with $W > 80$ nm. Temperature dependence of the minimum conductance for (a) a 85 nm wide GNR and (b) a 120 nm wide GNR. The solid green curve shows a fit to variable range hopping behavior $G \sim \exp(-(T_0/T)^\gamma)$ with $\gamma = 1/3$. The dashed blue curve shows $\gamma = 1/2$.

Wider ribbons show a qualitatively different temperature dependence

The temperature dependent characteristics described above are typical of all measured GNRs with $W \lesssim 80$ nm. For ribbons with $W > 80$ nm, we observe a different temperature dependence, shown in Figure 3.5. In these wider ribbons, high temperature behavior is not simply activated, and we do not see two distinct temperature regimes at high and low temperatures as we do for narrower ribbons (as in Fig. 3.3(b)). Rather, the entire range of data can be fit with a variable range hopping model. The curve fits well with $\gamma = 1/3$, although, as above, the exponent can not be reliably extracted with this range of data, and $\gamma = 1/2$ may also fit the data.

Variable range hopping with $\gamma = 1/3$ corresponds to 2D Mott variable range hopping. A value of $\gamma = 1/2$ can indicate Mott-1D VRH, and it can also occur in a 2D system with a soft Coulomb gap in the Efros-Shklovskii VRH model. The

notable feature here is not the precise exponent, but the fact that wider ribbons show a distinctly different temperature dependent behavior than narrower ribbons. The absence of a crossover to NNH behavior for $W > 80$ nm lends additional insight into the mechanism for that behavior. In particular, the 1D behavior breaks down above a certain width, and that this width is significantly greater than the ∼10–20 nm length scale of charge inhomogeneity observed in bulk graphene [76, 77]. This supports a model where the observed simply activated behavior corresponds to nearest neighbor hopping through a 1D arrangement of states induced by ribbon-wide local minima in the density of states.

We note that, based on the scaling rules above, an 80 nm GNR is expected to have an E_a corresponding to $\sim 80\ K$. The absence of simply activated behavior cannot be attributed to measurement in a range of temperatures too high to resolve E_a.

3.3.2 The gap in bias voltage corresponds to a critical electric field

An alternative approach to probing the GNR transport gap is measurement of the non-linear transport characteristics [31]. Figure 3.6(a) shows differential conductance, dI/dV_b as a function of both V_g and source-drain bias voltage V_b. Transport through the GNR at finite V_b shows a strong non-linear $I - V_b$ characteristic when E_F is in the transport gap regime, which is most extreme when V_g is near the charge neutrality point of the GNR (Figure 3.6(b), black curve). The non-linear gap ΔV_b can be defined where a steep increase of current appears in logarithmic scale (Figure 3.6, green curve). Note that this definition of ΔV_b is slightly different from the one used in the previous chapter. As others have observed [63, 66], the precise definition has

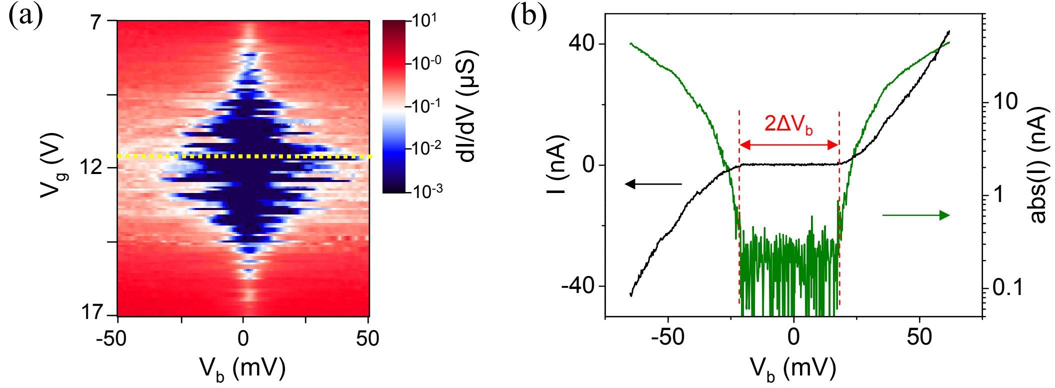

Figure 3.6: (a) Differential conductance as a function of both V_g and V_b measured in a GNR with $L = 1$ μm and $W = 31$ nm. (b) Current as a function of V_b with V_g fixed in the off-resonant condition marked by the dotted yellow line in (a). ΔV_b is highlighted by the dashed red lines.

little affect on the overall scaling behavior.

In our previous study [31], the energy corresponding to $e\Delta V_b$ was interpreted as the band gap of the GNR. However, this naive interpretation should be carefully reconsidered for edge disordered GNRs, where the charge transport is dominated by hopping through localized states. Indeed, from the plot of ΔV_b vs. W (Figure 3.7(a)), we notice that ΔV_b depends strongly on L, and is not well determined by W alone, in contrast to the three characteristic energy scales discussed earlier (Δ_m, E_a, and $k_B T^*$). Since the charge transport in disordered GNRs is diffusive, it is likely that electric field is driving transport in the transport gap. Indeed, if we convert ΔV_b into the corresponding critical electric field $\mathcal{E}_{cr} = \Delta V_b / L$, we restore a reasonable scaling behavior, where $\mathcal{E}_{cr}$ depends only on W and not on L (Figure 3.7(b)).

In disordered systems in which transport is dominated by hopping through localized states, applied electric field $\mathcal{E}$ plays a similar role to temperature. Thus we

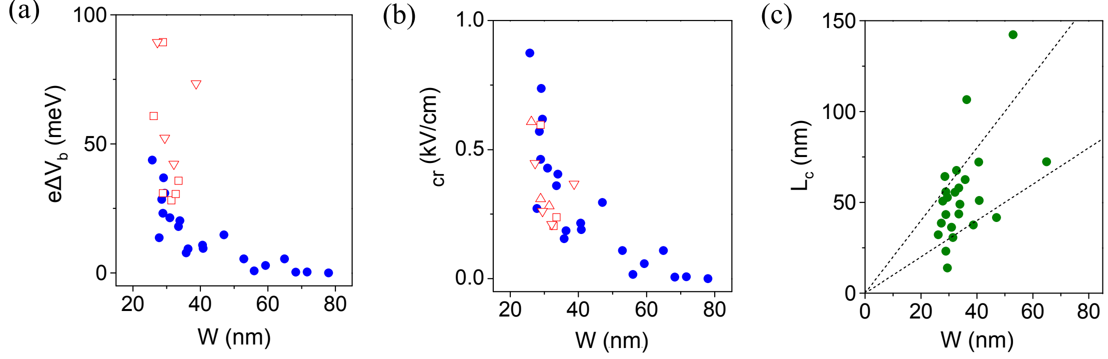

Figure 3.7: (a) ΔV_b as a function of W. Symbols follows the convention set in Fig. 2. (b) The critical electric field $\mathcal{E}_{cr}$ versus W converted from the data set in (a). (c) L_c as a function of W. Dashed lines at $L_c = W$ and $L_c = 2W$ are a guide to the eye.

can treat the electric field as an effective temperature: $k_B T_{eff} = e\mathcal{E}L_c$, where L_c is the average hopping length between localized states [78]. Noting that the transition from NNH dominated transport to VRH transport occurs at T^*, we relate T^* to the transition occurring at $\mathcal{E}_{cr}$ and estimate $L_c \approx k_B T^*/e\mathcal{E}_{cr}$. For most GNRs in this experiment we find that $W \lesssim L_c < 2W$, as seen in Fig. 3.7(c). The fact that $L_c \gtrsim W$ supports our claim that hopping transport through the ribbons is effectively 1D. We note that this L_c is distinct from the wavefuction localization length, which should be smaller than L_c and is expected to be comparable to W [60, 79].

3.3.3 Coulomb charging contributes to E_a

Finally, we discuss the effect of Coulomb charging in GNRs. Several previous works have discussed the role of Coulomb blockade and charging effects on the transport gap in GNRs and graphene constrictions [46, 64, 66]. In principle, in a GNR with hopping between localized states, we expect Coulomb interactions to open a soft Coulomb gap

near the Fermi surface, which can be incorporated into the total nearest-neighbor hopping energy E_a in addition to the single particle energy level spacing t'/w, so that $E_a \approx t'/w + E_C$, where E_C is the Coulomb charging energy. [51, 52, 55].

In order to quantify the contribution of charging energy E_C to the hopping energy E_a, we perform a comparative transport measurement on GNRs with different gate coupling (Figure 3.1). Figure 3.8 shows the temperature dependent minimum conductance $G_{min}(T)$ for a back-gated GNR (device I) and dual-gated GNR (device II) with similar W and L. While device I has the usual capacitive coupling to the back gate, (i.e., $C^I \approx C_g$), device II is much closer to the top gate, leading to a larger capacitance: $C^{II}/C^I \approx 4$. This ratio was calculated using a finite element method in COMSOL Multiphysics and is lower than the analogous ratio for bulk graphene devices. This is because back-gated only ribbons have an enhanced gate capacitance due to fringing electric fields, which are screened in the dual-gated device, as shown in Figure 3.9. From the thermally activated Arrhenius behavior in the high temperature regime (dashed lines), we obtain the activation energies of the two devices, $E_a^I = 15$ meV and $E_a^{II} = 8.4$ meV averaged over two devices of type I and four of type II. Considering the reduced charging energy contribution in the dual gated device, smaller values of the activation energy are indeed expected, if Coulomb effects are appreciable in the GNR.

Employing the ratio $E_a^{II}/E_a^I \approx 0.55$, we now can estimate the charging energy contribution quantitatively. Assuming that the single particle energy level spacing t'/w is similar for both GNRs due to their similar dimensions, we obtain $E_a^{II} - E_C^{II} =$

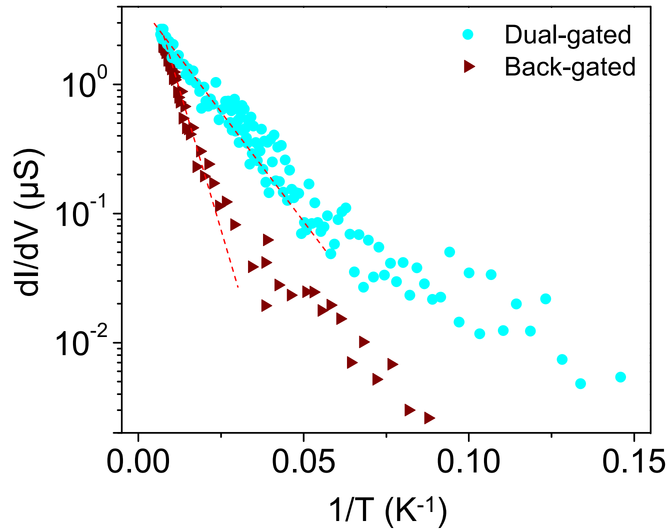

Figure 3.8: Temperature dependence of the conductance minimum for dual-gated (blue circles) and back-gated (brown triangles) GNRs with similar W and L. The dashed red lines are Arrhenius fits in the high temperature regime.

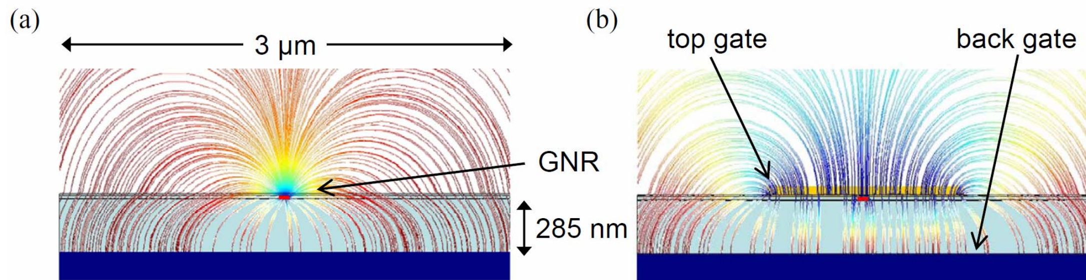

Figure 3.9: A simulated cross-sectional side view of (a) a back-gated device and (b) a dual-gated device. Electric field lines from the back gate were calculated using finite element methods in COMSOL Multiphysics.

$E_a^I - E_C^I = t'/w$, where the charging energy ratio of device I and II are given by $E_C^I/E_C^{II} = C^{II}/C^I \approx 4$. The resulting estimate for the charging energy contribution, $E_C^I/E_a^I \approx 0.6$, indicates that the Coulomb charging effect provides a substantial portion of the activation energy.

3.4 Conclusions

In this chapter we have studied the transport gap in etched graphene nanoribbons in detail, and found evidence of hopping transport through localized states inside the transport gap, with features of both Anderson localization and Coulomb charging effects. The main results of the chapter follow below.

The gap as measured in gate voltage, Δ_m gives the mobility edge, the threshold energy below which electron wavefunctions are localized, and above which states are extended along the length of the ribbon. Our measured trend for $\Delta_m(W)$ matches well with predictions for mobility gap size by Querlioz *et. al.*

Temperature dependent behavior in the transport gap shows variable range hopping behavior at low temperatures and crosses over to simply activated behaviors at higher temperature in narrow ribbons. The corresponding activation energy is an order of magnitude lower than the mobility gap, indicating transport by hopping through nearest neighbor localized states inside the transport gap. The crossover to NNH at high temperatures, as well as the absence of NNH in ribbons with $W > 80$ nm, suggests a segmentation of the local density of states by ribbon-wide local minima in the LDOS.

Attempts to measure the gap based on the source-drain bias at the onset of conduction in the non-linear regime produce a length-dependent value which is more meaningful when converted to a critical electric field. This critical electric field for the onset of conduction in the gap can be related to the transition temperature T^* at the onset of NNH to derive a nearest-neighbor hopping length scale L_c. We find that L_c is usually between one and two times the ribbon width, consistent with the model for transport by quasi-1D hopping.

Charging energy constitutes a significant portion of the activation energy, but importantly, it does not account for activation energy in its entirety. Understanding the transport for energies inside the gap requires consideration of both Anderson-type localization and single electron charging in localized states. This combination of a disordered single-particle picture with strong electron-electron interaction is a non-trivial problem for which no suitable theoretical model is yet available. We note that GNRs are a well suited platform to study this combination of effects, since they provide access to both disorder strength (through, for example, annealing, as in Ref. [66]) and charging energy (through capacitance) as tunable parameters.

Chapter 4

Graphene devices at high bias

This chapter examines the high bias transport characteristics of graphene nanoribbons. So far in this thesis, we have focused mainly on low-bias transport, with the source-drain bias V_{sd} in the microvolt or millivolt range. Here we investigate the transport behavior of devices biased up to a few volts, a regime more relevant for electronics applications. We first address characteristics of graphene at high bias which are not specific to graphene nanoribbons, then we address GNRs at high bias specifically. Much of the work present in this chapter was published in Ref [14], though here we rely on a different dataset and use slightly different analysis.

4.1 Background: Current saturation in carbon nanotubes

Since graphene and graphene nanoribbons have a similar band structure to carbon nanotubes, we first review related nanotube experiments.

In a well known experiment, Yao *et. al.* [80] found that current in metallic single

wall carbon nanotubes saturates at high electric field. Their result is explained in terms of zone-boundary optical phonon emission from high energy electrons. At high electric fields, a steady-state population is developed between right and left moving charge carriers with a maximum energy difference corresponding to the phonon energy $\hbar\Omega = 160$ meV, leading to a saturated current of $(4e/h)/(\hbar\Omega) \approx 25$ μA.

A slightly different behavior was reported in semiconducting single wall carbon nanotubes by Chen and Fuhrer [81]. In these devices, current does not saturate completely, and the transport is described by an electric field dependent carrier velocity. The authors fit their data with a model based on a carrier velocity that saturates to a constant value at high electric field and a carrier density dependent on the local potential along the device. They find a saturation velocity of 2×10^7 cm/s in their device.

4.2 Experiment: Measurement at high source-drain bias

The devices measured in this experiment are back-gated and dual-gated etched graphene devices made following similar fabrication procedures to those described in Chapters Two and Three. Graphene devices often fail or change drastically and irreversibly when the current density per unit width exceeds a threshold of ~ 2 mA/μm. For measurements described in this chapter, we operate the device at currents below this threshold. Current-voltage characteristics at varying gate voltages were measured for 17 ribbon devices with a range of widths and lengths, and three "wide" devices with $W = 200$ nm, in order to compare to the behavior of "non-ribbon" devices.

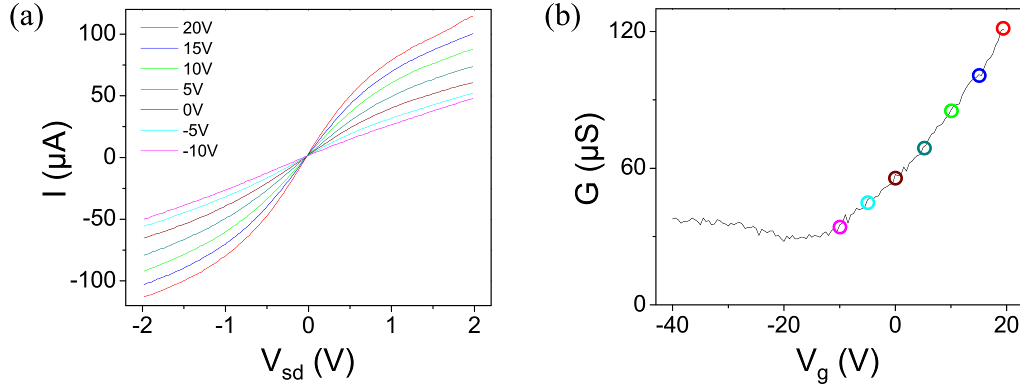

Figure 4.1: High bias transport data from a ribbon with W=70 nm and L=500 nm measured at room temperature in vacuum. (a) Current vs. source drain bias at varying gate voltages V_g, as shown in the legend. (b) Conductance vs. gate voltage for the same device at a source drain bias of V_{sd} =200 meV.

4.3 Results and discussion

4.3.1 Saturating behavior fits a velocity saturation model

Figure 4.1 shows a plot of current vs source-drain bias for varying gate voltages in a back-gated device. We focus here on the curves taken at densities far from the charge neutrality point, such as the curve singled out in Figure 4.2. Here we see that at low bias the slope of the curve is constant, and at high bias the curve turns down, approaching a linear behavior with a reduced slope.

To describe this saturating decrease in conductivity, we propose a model based on an electric field-dependent carrier velocity $v_d(E)$ of the form:

$$v_d(E) = \left(\frac{1}{\mu_0 E} + \frac{1}{v_{sat}} \right)^{-1} \tag{4.1}$$

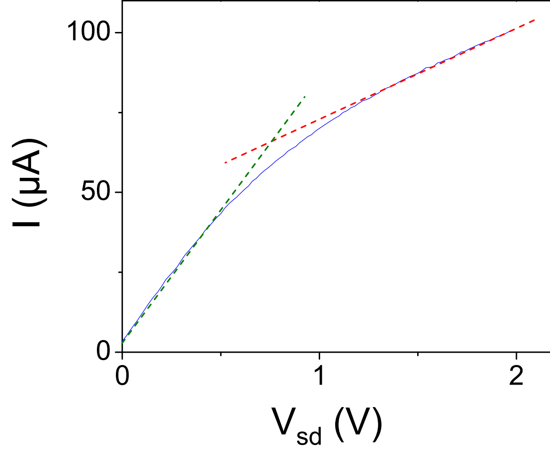

Figure 4.2: Current-voltage characteristic for the curve with $V_g - V_{CNP}$ = -30V shown in Figure 4.1(a). The distinct slopes at low and high V_{sd} resulting from the saturation behavior are highlighted with green and red dashed lines, respectively.

where μ_0 is the low field mobility and v_{sat} is a phenomenologically introduced saturation velocity. The total current through the device is given by

$$I = jW = -nev_dW \tag{4.2}$$

We assume that the capacitance to the back gate dominates in determining the charge density in the channel, so that

$$ne = C_g\left(V(x) - (V_g - V_{CNP})\right) = C_g(V - V_0) \tag{4.3}$$

where $V = V(x)$ is the potential at position x along the channel, and we have defined

$V_0 \equiv V_g - V_{CNP}$. Using the relation $E = dV/dx$, we have

$$I^{-1} = -\frac{1}{WC_g(V - V_0)} \left(\frac{1}{\mu_0 dV/dx} + \frac{1}{v_{sat}} \right) \tag{4.4}$$

Rearranging terms and integrating gives the current

$$\int_0^{V_b} \left(\frac{WC_g\mu_0}{I}(V_0 - V) - \frac{\mu_0}{v_{sat}} \right) dV = \int_0^L dx \tag{4.5}$$

$$\frac{WC_g\mu_0}{I} \left(V_0 V_b - \frac{1}{2} V_b^2 \right) - \frac{\mu_0 V_b}{v_{sat}} = L \tag{4.6}$$

$$I = WC_g(V_0 - V_b/2) \frac{\mu_0 V_b/L}{1 + \mu_0 V_b/v_{sat}L} \tag{4.7}$$

In its limiting forms, Eqn. 4.7 for the current qualitatively gives the behavior seen in Figure 4.2. At low V_b, current is linear in V_b with a conductivity $WC_gV_0\mu_0/L$, determined by the low field mobility, as expected. At high V_b, current is again linear in V_b, but now with a conductivity of $WC_gv_{sat}/2$ and an offset determined by the gate voltage. At low fields, the variation in carrier density is small and the linear I–V results from the linear form of $v_d(E) \approx \mu_0 E$ in this regime. At high fields, v_d approaches a constant value v_{sat}, and the linear dependence of the carrier concentration on V_b is responsible for an I–V characteristic approaching linear behavior. Note this is in contrast to the case of carbon nanotubes, where there are a set number of conducting channels, so that the current saturates with the drift velocity.

The expression in Eqn. 4.7 for $I = I(V_b)$ was fit to the I–V characteristics in Figure 4.1; the result is shown in Figure 4.3. For ribbon devices, the geometry is not well approximated by a parallel plate capacitor, so the gate capacitance was calculated

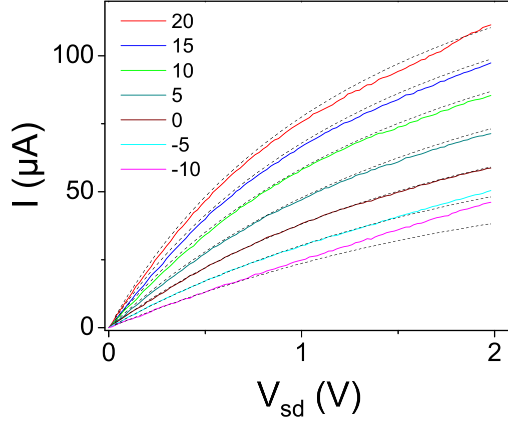

Figure 4.3: Fit of the current model in Equation 4.7 to the data in Figure 4.1. The legend gives the gate voltage for each sweep, dashed lines are fits to each sweep.

numerically.[1] The model fits well for curves taken at densities far at high carrier densities, and begins to break down for curves measured near the charge neutrality point, as seen in Figure 4.3 for $V_g = -10$ V. This fit has two free parameters, v_{sat} and μ_0. For this dataset, this model gives μ_0 values between 400 and 600 cm^2/Vs, compared to the value of 700 cm^2/Vs from low bias sweeps of G–V_g.

The values of v_{sat} obtained from this fit are plotted against V_g in Figure 4.4(a). In Figure 4.4(b), we plot v_{sat} against the inverse of the Fermi energy

$$E_F = \hbar v_F \sqrt{\pi C_g (V_g - V_{CNP})} \tag{4.8}$$

Converting V_g to E_F involves the value of V_{CNP}, which commonly drifts throughout measurement due to changes in adsorbed molecules and positions of trapped charges.

[1]For the device in Figure 4.3, the capacitance was calculated to be 47.5 nF/cm^2; the method for calculating the capacitance will be discussed in more detail in Chapter Five.

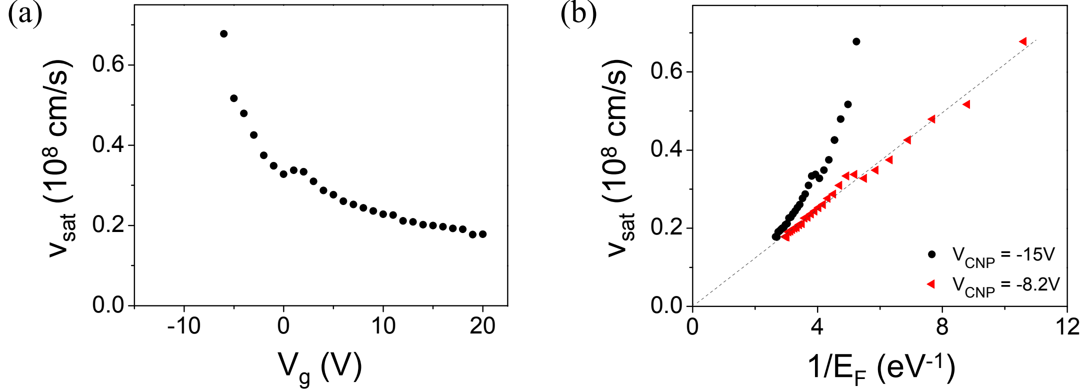

Figure 4.4: Saturation velocity values extracted from the fits in Figure 4.3. (a) Saturation velocity v_{sat} as a function of the gate voltage V_g. (b) v_{sat} vs the inverse of the Fermi energy E_F. Black circles and red triangles correspond to a conversion of V_g to E_F using $V_{CNP} = -15$ V and $V_{CNP} = -8$ V, respectively. Dashed line is a linear fit to this data.

Here the black circles correspond to conversion of V_g to E_F using $V_{CNP} = -15$ V, the same value used in Equation 4.3 for the original fit. Red triangles represent a conversion to E_F using $V_{CNP} = -8$ V so that the a linear fit of v_{sat} vs. E_F^{-1} intersects the origin.

In order to understand the inverse relationship between v_{sat} and E_F, we seek a physical understanding of the electric field dependent carrier velocity, or drift velocity, in Equation 4.1. This expression corresponds to scattering by optical phonons, which would produce an electric field dependent mean free path. By Matthiessen's rule, mean free paths add as

$$\frac{1}{l} = \frac{1}{l_{sc}} + \frac{1}{l_{op}} \tag{4.9}$$

where l is the total mean free path and l_{sc} is the mean free path for elastic impurity scattering and quasi-elastic acoustic scattering, and l_{op} is the mean free path for

optical phonon emission. If electrons are immediately scattered upon reaching the optical phonon energy, so that

$$l_{op} = \frac{\hbar\Omega}{eE} \tag{4.10}$$

where E is the electric field and Ω is the relevant optical phonon frequency, then the mobility μ is given by

$$\frac{1}{\mu} = \frac{1}{\mu_0} + \frac{E}{v_{sat}} \tag{4.11}$$

This form of the mobility results in the expression for the drift velocity $v_d = \mu E$ given in Equation 4.1. For electrons and holes in graphene, which have a constant carrier velocity of v_F, drift velocity can be understood as the time averaged velocity of carriers when scattering is taken into account.

From the above calculation we see that our phenomenological velocity saturation model can be understood in terms of a picture where electrons scatter by optical phonon emission upon reaching the phonon energy $\hbar\Omega$ under the influence of the applied electric field. With this in mind, we derive an expression for current density using a different approach, in order to gain insight into our measured values for the saturation velocity. Current density is given by

$$\vec{j} = -e \int d\vec{k} D_k \vec{v}(\vec{k}) g(\vec{k}) \tag{4.12}$$

where $D_k = 2/(2\pi)^2$ is the density of electronic states in k-space, $\vec{v}(\vec{k}) = v_F$ is the electron velocity, and $g(\vec{k})$ is the distribution function. In the relaxation time

approximation, we have

$$g(\vec{k}) = g^0(\vec{k}) - e\vec{E}\cdot\vec{v}(\vec{k})\tau(\epsilon(\vec{k}))\left(-\frac{\partial f}{\partial \epsilon}\right) \tag{4.13}$$

where $g^0(\vec{k})$ is the equilibrium distribution function, τ is the relaxation time, and f is the Fermi-Dirac distribution function. For a device with its length in the x direction, we seek $\vec{j} = j\hat{x}$, so we consider only $\vec{E} = E\hat{x}$, and

$$\vec{E}\cdot\vec{v}(\vec{k}) = Ev_F\cos\theta \tag{4.14}$$

where θ is the angle between $d\vec{k}$ and $\vec{E}$. We assume that electrons are immediately scattered upon reaching the energy threshold for phonon emission, giving

$$\tau = \frac{\hbar\Omega}{eEv_F} \tag{4.15}$$

So that for Eqn. 4.12 we have

$$j = e\int \frac{dk}{\pi^2} v_F\cos^2\theta\hbar\Omega\left(-\frac{\partial f}{\partial \epsilon}\right)\Big|_{\epsilon=\hbar v_F k} \tag{4.16}$$

In polar coordinates

$$\begin{aligned} j &= e\int_0^\infty \frac{dk}{\pi^2}\int_0^{2\pi} d\theta v_F\cos^2\theta\hbar\Omega\delta(\hbar v_F k - E_F) \\ &= \frac{e}{\pi}\Omega\frac{E_F}{\hbar v_F} \end{aligned} \tag{4.17}$$

At high fields, we assume $j = nev_{sat}$ and use $E_F = \hbar v_F \sqrt{\pi n}$ to obtain

$$\frac{v_{sat}}{v_F} = \frac{\hbar\Omega}{E_F} \tag{4.18}$$

Using this expression with $v_F = 10^8$ cm/s [10, 11], we obtain a value of $\hbar\Omega = 62.0$ meV from the linear fit (dashed line) in Figure 4.4(b). This is well below the value of the longitudinal zone-boundary phonon for graphene, which has $\hbar\Omega = 200$ meV [43]. We suggest that our measured phonon energy corresponds to the SiO_2 surface phonon energy $\hbar\Omega = 55$ meV [82–84], although we note that values measured in other ribbon devices of different geometries vary widely (from ≈ 22 meV to ≈ 120 meV), possibly due to discrepancies in determining the relevant device geometry, the corresponding capacitance, and the position of the charge neutrality point.

4.3.2 Top-gated graphene devices show an enhanced current saturation effect

In dual-gated devices, we observe a velocity saturation behavior similar to that the back-gated device behavior described above. However, we also see an enhanced current saturation at certain gate voltage combinations, as first reported in Reference [14]. Figure 4.5 shows current-voltage characteristics and corresponding conductance-gate voltage sweeps for a dual-gated device with $W = 35$ nm and $L = 2$ μm. At combinations of V_{bg} and V_{tg} near the charge neutrality point, we see a "kink" in the I–V curve, where the current first begins to flatten out, then turns upwards again. Figure 4.6 highlights this behavior in one I–V curve from the same device. This effect

is specific to top-gated devices, where the strong capacitive coupling allows the bias voltage to dominate the carrier density in the channel.

The "kink" effect in graphene is similar to pinch-off in traditional MOSFETs, where a strong bias voltage pulls the quasi-Fermi level at one end of the channel into the charge-depleted bandgap. In graphene, where there is no bandgap, this results in a transition within the channel from one carrier type (electrons or holes) to the other. In a device that is n-type, as in Figure 4.6(c)(I), a positive source-drain voltage (applied to the source) depletes the electron density in the channel near the source (II). At sufficiently strong positive bias voltage, the bias voltage begins to pull holes into the channel, so a region of the channel is at charge neutrality and contributes a large resistance (III). As bias is further increased, hole density at the source also increases, so conductivity increases again (IV). In Reference [14] we showed that wide plateaus in current could be achieved when this "kink" effect is made to coincide with velocity saturation.

4.3.3 Heating effects can overcome transport gap at high bias

The results discussed so far in this chapter come from graphene nanoribbons measured at high bias, but the key features of the data, saturation velocity at strong electric fields and the "kink" effect in the current for top-gated devices, are also seen in wide graphene devices [14]. This leads to the question, how are nanoribbons different from wide, non-ribbon devices when operated at high bias? Here we present the preliminary results of a comparison between dual-gated ribbons and wide devices and so far find no major differences in their performance. This result is only preliminary

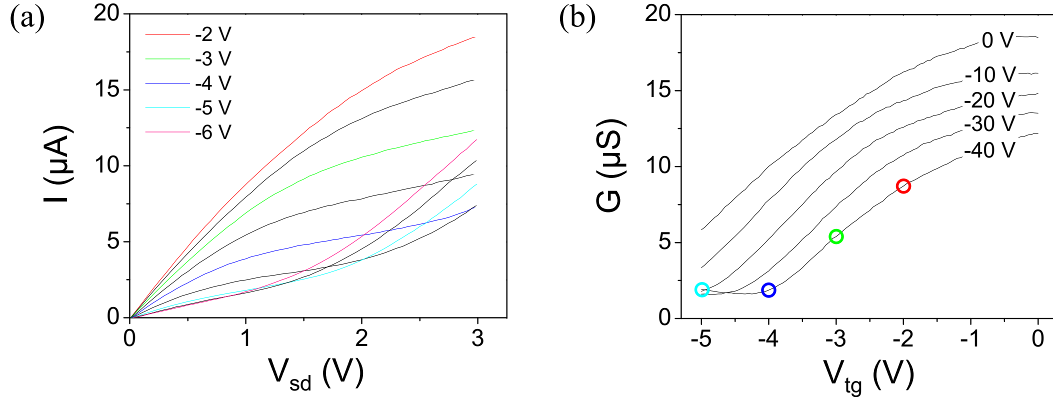

Figure 4.5: Current-voltage characteristics for a dual-gated device with W=35 nm and L=2 μm measured at room temperature in vacuum. (a) I–V_{sd} at constant V_{bg} = -40 V and varying V_{tg}. Select curves are shown in color and have V_{tg} values as noted in the legend, the black curves fall between these curves at 0.5 V increments. (b) G-V_{tg} at varying V_{bg} values, as noted on each curve, measured with $V_{sd} = 1$ meV. Colored dots correspond to the gate voltage positions where the colored curves in Figure (a) were measured.

because the widths of the GNRs in this experiment are not well specified within the range of $W \approx$ 20–60 nm. Recall from Chapter Three that in dual-gated devices, the widths of nanoribbons lying underneath the dielectric and metal layers cannot be accurately measured. Estimates of the width can be made based on the expected width dependence of the low-temperature transport characteristics Δm and ΔV_b from the analysis in Chapter Three. From these comparisons, it is estimated that the ribbons used in this experiment have $W \approx 50$ nm. Ribbons of this width are narrow enough to behave distinctly from "wide" ($W \gtrsim 100$ nm) devices at low temperatures and low bias, but as we shall see below, they may not be narrow enough show a difference in transport characteristics at high bias. Ribbons as narrow as $W \approx$ 15–20 nm are achievable by our fabrication methods, so measurements of narrower devices

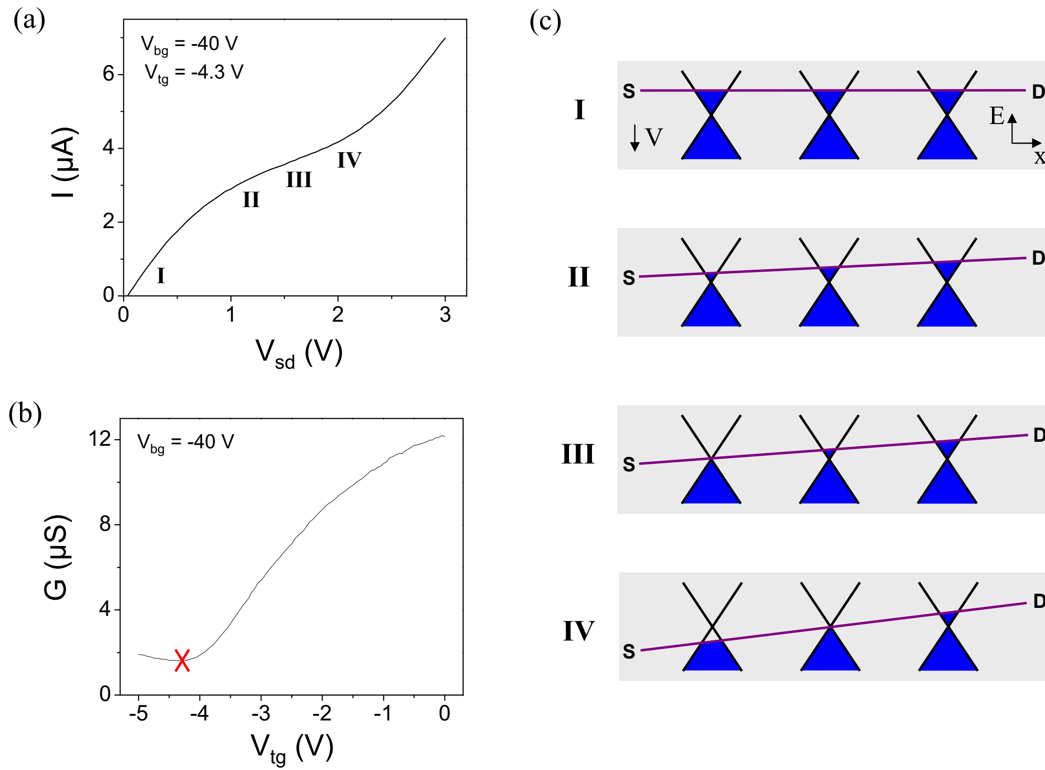

Figure 4.6: (a) An I–V_{sd} curve for the device in Figure 4.5, highlighting the current saturation "kink" behavior. (b) G-V_{tg} for the same device; the red "x" highlights the low bias (V_{sd} = 1 meV) conditions corresponding to the curve in (a). (c) Cartoon schematic of the Fermi level in the channel for each condition (I–IV) marked in (a). Dirac cones with different Fermi levels along the length of the channel result from carrier density variation along the channel under the influence of a strong V_{sd}.

with larger transport gaps may still reveal distinct device behavior.

In comparing gapped graphene nanoribbons to wide graphene with no gap, there are several differences we may expect to see. First, since graphene nanoribbons have a strongly suppressed current at energies inside the gap, we may see an increased transconductance. Also, we could see larger and more fully saturated current in the "kink" region, as the presence of a gap causes the "kink" to more closely resemble pinch-off in a traditional MOSFET. We may also see the effects of edge roughness. In narrow ribbons where edge roughness constitutes a significant portion of the total ribbon width, this could lead to a decrease in maximum current carrying capabilities, or cause the devices to degrade more quickly.

Figure 4.7 shows I–V characteristics for graphene devices taken at two different temperatures, 77 K and 300 K. The I–V curves do not change significantly between the two temperatures. Since we expect to see thermal effects in the conductivity even away from the charge neutrality point, this suggests that the effective temperature in the device is similar at both 77 K and 300 K, in other words, other heating in the system dominates over the ambient temperature up to 300 K.

As a straightforward method to directly compare ribbon devices with wide devices, we compare the scaled current density per width $j = I/W$ for two devices, a ribbon device with width $W \approx 50$ nm and a wide device with $W = 200$ nm, both with length $L = 500$ nm, shown in Figure 4.8. Here we can see that in the ribbon device, there is no difference in the size or shape of the kink behavior (Figure 4.8(b)), and only a minor difference in transconductance. From this data we see that a ~ 50 nm wide

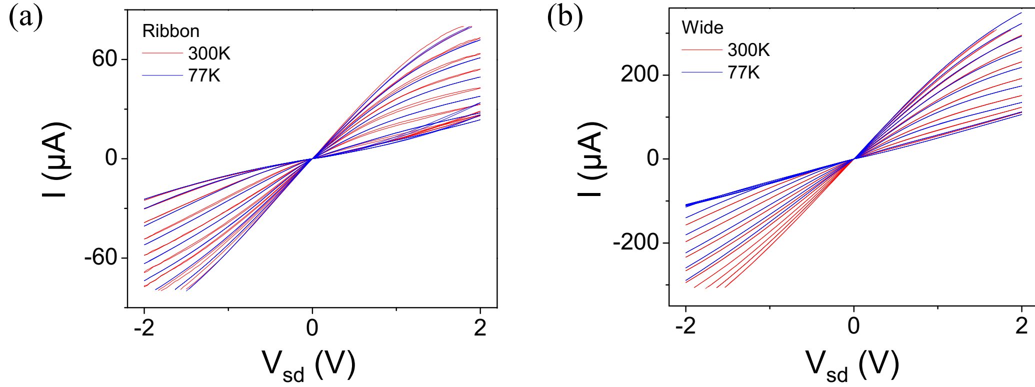

Figure 4.7: I–V_{sd} characteristics at constant $V_{bg} = 0$ with V_{tg} varying from 0 to -8 V, measured at T=77 K and T=300 K for (a) a ribbon device ($W \approx 50$ nm) and (b) a wide graphene device ($W = 200$ nm). Both devices have $L = 500$ nm

graphene nanoribbon shows no major differences in behavior from a wide device when operated at high bias.

To understand the similarity in behavior between 50 nm and 200 nm wide devices, we compare the gap size of the ribbon device with the relevant thermal effects in the system; if the available thermal energy in the system is larger than the gap, the effect of the gap will be washed out by thermally activated charge carriers. In Chapter Three we found that there are three different ways to measure the size of the gap: Δ_m, from the gate voltage, ΔV_b, from the bias voltage, and E_a, from the activation energy for nearest neighbor hopping. Here we are concerned with current flow at high bias, so ΔV_b is the most relevant of these scales for distinguishing the on and off states of the device, though E_a will determine the leakage current in the off state. For the 500 nm long devices studied here, these values are similar. Since ΔV_b has a strong length dependence, if we wish to increase ΔV_b we can increase the device length L,

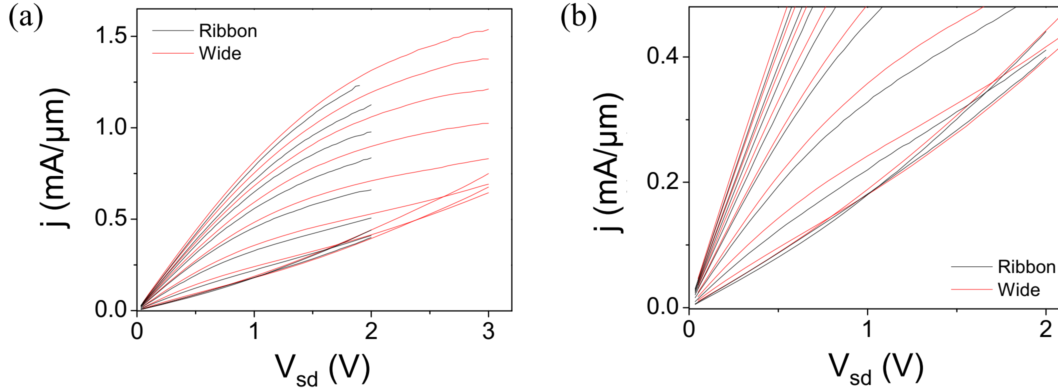

Figure 4.8: (a) Current density $j = I/W$ at high V_{sd} at constant V_{bg} and varying V_{tg} for the two devices in Figure 4.7. (b) Same data as in (a), enlarged to show behavior near in the region of the "kink".

with the trade-off of an increased the resistance and therefore a decreased current.

We consider two heating effects in this experiment. First, we compare the size of the gap with the thermal energy at room temperature, $k_B T \approx 26$ meV. Figure 4.9 shows how the relevant gap sizes (from the data in Chapter Three) compare with the room temperature thermal energy. We see that ΔV_b can be easily made greater than 26 meV by decreasing the ribbon width to below 30 nm or increasing ribbon length. However, only the narrowest ribbons shown here have a large enough E_a; narrower ribbons would be needed to ensure a low thermally activated leakage current.

If heating effects raise the device temperature above room temperature, then heating effects will be more relevant than ambient temperature effects. Several recent works [85–87] address the topic of heating in graphene at high bias. From the ribbon device data in Figure 4.7(a), we can expect to see dissipated electrical power

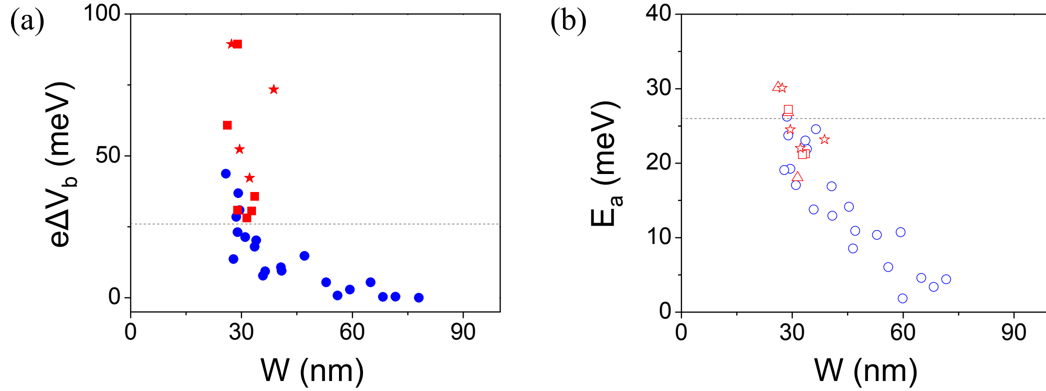

Figure 4.9: Measures of gap size vs width from the data in Chapter Three, with the room temperature thermal energy 26 meV highlighted with the dashed line.

$P = IV$ of up to ≈ 350 kW/cm^2, though power dissipation may be lower in the optimal operating regime for device applications. From the results in Reference [87] for temperature vs. power per area, this power dissipation corresponds to a temperature of 1350 K, or a thermal energy of 116 meV. From this it is clear that the thermal energy from heating greatly exceeds that from the ambient temperature, but this result was from a back-gated device. In a dual-gated device geometry, the top-gate dielectric and electrode may act as a heat sink and decrease the effect of heating.

In Figure 4.10, we model the heat sinking effects of a gate dielectric and top gate on a hot ribbon. This was done in the COMSOL Multiphysics finite element modeling package by assigning a heat flux to the ribbon such that maximum temperature in a back-gated device is ~ 1100 K, shown in Figure 4.10(a). The graphene ribbon and graphene leads were assigned a thermal conductivity of 5000 K [15] and a thickness of 3.4 Å; heat dissipation was also allowed through the 285 nm SiO_2 layer to the

Si substrate below. In Figure 4.10(b), a 30 nm SiO_2 gate dielectric and a 30 nm gold top gate are added to the same model, again allowing heat dissipation into the gate dielectric and the top gate. Here, SiO_2 was used in place of HSQ because they are expected to have similar material properties. In this model, the maximum temperature is decreased to 825 K. If the top gate thickness is increased to 100 nm to allow for more heat sinking, the temperature decreases only slightly more, to 812 K.

The gate dielectric actually used in the experiment consists of HSQ/HfO_2 with thicknesses of 15/15 nm. Hafnium dioxide has a much higher thermal conductivity than silicon dioxide (23 $W/m \cdot K$ for HfO_2 versus 1.4 $W/m \cdot K$ for SiO_2). When the model is changed to include the proper layer thicknesses of each dielectric, the maximum nanoribbon temperature decreases to 680 K, which corresponds to an energy of 59 meV. This is the behavior we can expect to see in the actual device measured in Figures 4.7 and 4.8. From Figure 4.9 it is clear that E_a and ΔV_b are both far below this energy, so thermally activated carriers easily wash away any gap-related effects we might have seen in the transport at high bias.

Heat sinking could be greatly improved by removing the HSQ, such as by an hydrofluoric acid etch, and depositing $\approx$ 15 nm of hafnia only as the dielectric. In this geometry, the dielectric would be thinner and more thermally conductive, allowing for more efficient heat dissipation to the metal top gate. For the same heating conditions, this device construction would result in a maximum nanoribbon temperature of 460 K. The corresponding energy, 40 meV, is a gap size easily achievable by our nanoribbon fabrication methods. We note that from the results in Chapter Three, the addition

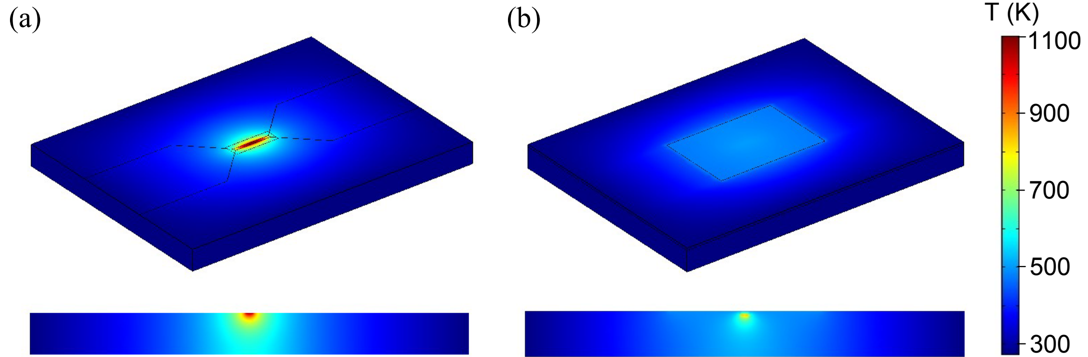

Figure 4.10: Thermal modeling of (a) back-gated and (b) dual-gated graphene nanoribbon devices. Top row is a full 3D view of each device (on a 3 μm by 4 μm rectangle), bottom row is a cross-sectional slice taken midway across the ribbon (showing the oxide thickness of 285 nm), colorbar is the same for all plots.

of a top gate tends to decrease E_a, but as the top-gated geometry provides very good heat sinking, and top gates will ultimately be needed for optimized device design, we see this as the best route for development of a graphene nanoribbon device that retains its gapped behavior at high bias.

4.4 Conclusions

In this chapter we have observed a saturating I–V characteristic in graphene devices operated at high source-drain bias, and described the behavior using a model where surface phonon emission results in a carrier velocity that saturates to a Fermi energy dependent value at high applied electric field. We showed that for top-gated graphene devices have an enhanced current saturation effect at certain gate voltage combinations. This effect results from the introduction of the charge neutrality point into the

channel, and is similar to pinch-off in MOSFET devices. We see that heating effects in graphene at high bias are significant, and very narrow ribbons with a strongly heat sinking device design are required to produce a device where confinement-induced gap effects dominate over the effects of heating.

Chapter 5

Chemical edge functionalization of graphene nanoribbons

This chapter addresses a major hurdle in the use of etched graphene nanoribbons in semiconductor applications, and begins to explore one possible solution. Specifically, this chapter will discuss the use of chemical functionalization of the etched graphene edges as first step in controlling edge disorder in GNRs with the goal of improving GNR electronic properties.

5.1 Motivation and Background: Control of etched graphene edges is essential for applications

The on/off ratio in etched GNRs is so far too low to allow for their use in digital transistor applications. The on/off ratio, I_{on}/I_{off}, describes the ratio between the current in the "on" state of the device (I_{on}), at energies far from the gap, to the current

in the “off” state of the device (I_{off}), at energies inside the gap. In etched graphene nanoribbons, this quantity is typically on the order of 10^2 at low temperatures and 10^1 at room temperature, whereas for digital transistor applications, a larger on/off ratio is needed.

The low on/off ratio in GNRs is due to the high off current I_{off} through localized states at energies inside the transport gap, and will persist as long as those localized states are present in GNRs. For an idea of how to decrease the formation of localized states in the gap, we turn to the results in Chapter Three or Ref. [47]. These studies showed that disorder at the etched graphene edges is likely to play a major role in the localization of states at low energies. Control over the edges could ultimately allow us to produce GNRs with fewer localized states and, consequently, a lower off-current and higher on/off ratio. As a first step towards controlling the ribbon edges, we seek a selective, covalent, chemical functionalization of the etched graphene edges.

In addition to a process that is covalent and selective to the edges, we would like our edge functionalization process to be compatible with large-scale fabrication. Since graphene’s compatibility with lithographic processes was a major draw to graphene-based electronics, we would like an edge-processing step that maintains this advantage.

A few approaches so far in the literature attempt to gain control over graphene edges. Li *et. al.* [88] have developed a method for creating ribbons in solution through chemical exfoliation. After they are deposited onto a substrate, the ribbons have good electronic properties including a very high on/off ratio of 10^7 at room temperature.

The very low off current suggests that the ribbon edges may be structurally very clean compared to etched ribbons, but no conclusive evidence (*e.g.*, transmission electron microscope images) has so far been published. Jia *et. al.* [89] have demonstrated that Joule heating can lead to a rearrangement and vaporization of carbon atoms resulting in sharp zigzag or armchair edges. An approach by Campos *et. al.* [90], following an early demonstration by Datta *et. al.* [91] uses thermally activated metal nanoparticles to etch graphene preferentially along the zigzag direction, though isolated nanoribbon devices have not yet been produced. Other locally controlled etching methods by scanning tunneling microscope lithography [92] and local anodic oxidation lithography [93] have also been reported.

All of these methods take the approach of controlling ribbon edge structure directly during fabrication. It is possible to imagine a fabrication process with a Joule heating step that could be somewhat compatible with planar fabrication; combining other fabrication processes with large scale fabrication is likely to be more challenging. We envision a chemical modification process that could be applied *en masse* to already etched ribbons on a wafer scale. To date, no such process has been demonstrated.

5.2 Experiment: Chemical functionalization of graphene edges

The goal of this experiment is to make covalent, chemical modifications that are selective to etched graphene edges. Additionally, we would like to show that after this initial chemical modification, we can perform additional chemistry on the newly

added end groups.

We quantify the effectiveness of the chemical functionalization by measuring changes in the position of the charge neutrality point (CNP), which reflects the charge doping level in the graphene. In order to determine the edge-selectivity of the process, we will perform the functionalization on ribbons of varying width, and look for a width dependent response. If the process is selective to the edges, we expect to observe an enhanced effect on narrower ribbons, which have a larger number of edge sites relative to basal plane sites.

The basic structure of the experiment is as follows. First we fabricate a set of back-gated GNRs with constant length and varying width using the methods described in Chapter Two. Then we measure conductance as a function of gate voltage, and find the position of the charge neutrality point in each ribbon. Once the initial CNP has been determined, we subject the devices to a series of chemical treatments. After each treatment, we repeat the CNP position measurement to measure the change in doping from each treatment.

Our main data point at each stage is the position of the charge neutrality point, which reflects doping due to chemical functionalization, but is also sensitive to physical adsorption of molecules. In order to focus on the effects from covalent chemistry, it is necessary to minimize this incidental charge doping. For this reason, the measurements are done in a vacuum probe station, and a thermal anneal step (1 hour at 110C in vacuum) is used at each stage to drive off physiosorped charge dopants. In

some cases, the CNP was measured both before and after the anneal step for comparison. In general, the CNP was moved to a lower voltage after annealing, suggesting a desorption of p-type dopants (such as oxygen) during the anneal.

We study chemical treatment with diaminopropane, which is expected to donate electrons and thus act as a n-type dopant, followed by methylation with methyl iodide, which is expected to partially reverse the n-doping. The diaminopropane treatment was done with a solution of 1,3-diaminopropane (1.99 g) in a solvent of pyridine (17.3 g) with 1-ethyl-3-(3-dimethylaminopropyl carbodiimide(EDCI)) (0.325 g). Graphene ribbon devices (complete with electrodes) were soaked in this solution for 24–48 hours, followed by a 24 hour rinse in pyridine or acetone. Methylation was done by soaking the devices in a 3:1 solution of acetone:methyl iodide for 12 hours, followed by a rinse in acetone and a 24 hour soak in isopropyl alcohol or deionized water. Control experiments were done with pyridine only and with pyridine and EDCI only.

5.3 Modeling: Accurate capacitance estimates are needed to find the expected CNP shift from doping

The gate voltage at the minimum conductance point corresponds to the charge density required to modulate the Fermi level to the charge neutrality point through

$$V_{CNP} = \frac{Q}{C_g} = \frac{ne}{C_g} \tag{5.1}$$

where V_{CNP} is the gate voltage at the charge neutrality point, Q is the total induced charge, n is the number of charge carriers (electrons or holes) per area, e is the electron charge, and C_g is the geometrical gate capacitance per area associated with adding charge to the device. For a given device, the position of the charge neutrality point moves in proportion to the charge density of the device.

We are interested in the charge density contributed to the device though dopants acting only on the edges. We use E to represent the number of charges donated from edge sites per length of the ribbon (including both edges), so that density of charges per area is E/W, or, if we take E and W to have units of charge/nm and nm, respectively, then

$$n = \frac{E \times 10^{14}}{W} \tag{5.2}$$

where n has units of cm^{-2}. The analogous relation for doping in the basal plane has a more straightforward form, $n = P$, where P is the number of charges donated per unit area and has units of cm^{-2}. Thus charge density from edge doping has a width dependence that distinguishes it from plane doping.

The relation between width and CNP position is complicated by a width dependent geometric capacitance. The graphene nanoribbons studied here have widths as small as 20 nm, and are separated from the silicon backgate by 285 nm of silicon dioxide, so the geometry is very different from a parallel plate capacitor. Fringing electric field lines from the back gate to the ribbon device result in a more strongly enhanced capacitance for narrower ribbons, as shown in Figure 5.1. The width dependence of the capacitance has the effect of reducing the width dependence of the CNP shift

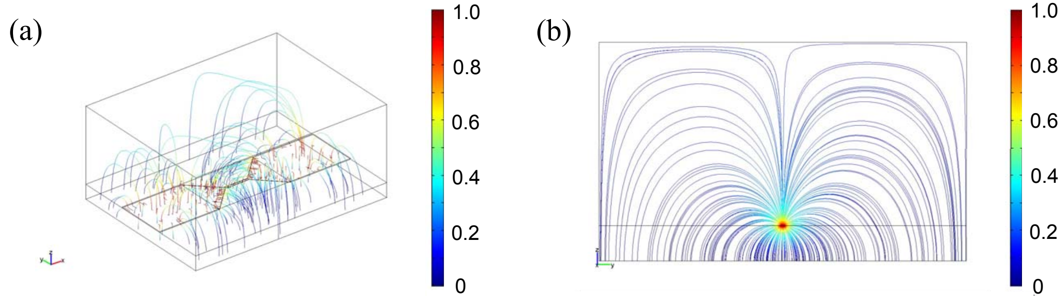

Figure 5.1: Electric field line simulations, calculated by finite element modeling with the COMSOL Multiphysics electrostatics package. Color bar represents electric potential in volts (a) Complete view of electric field lines in the device model to show the geometry of the graphene device, with a 40 nm wide ribbon connected to 1 μm wide graphene leads on 285 nm of SiO_2. Extra lines in device geometry have been added to define increased grid density near the graphene. (b) Cross-sectional side view showing fringing electric fields converging on the narrow ribbon.

from edge doping. While edge doping has a stronger effect on the CNP for narrow ribbons, the increased capacitance in narrow ribbons means that more charge density is needed to produce a certain shift in voltage. This will make the width dependent CNP shift due to edge doping more difficult to detect, but as we will see below, it aids us in separating the effects of plane doping from those of edge doping.

To determine the geometric capacitance of GNR devices with varying widths, we used the COMSOL Multiphysics software to carry out finite element method (FEM) calculations. Capacitance for a given device geometry was determined by setting the device voltage to 1 V, the backgate voltage to 0 V, and calculating the resulting charge density accumulated on the ribbon, shown as a color plot in Figure 5.2(a). This gives a correction factor to the capacitance $S(W) = C_r/C_p$, where C_r is the ribbon capacitance found through the FEM calculation and C_p is the parallel plate

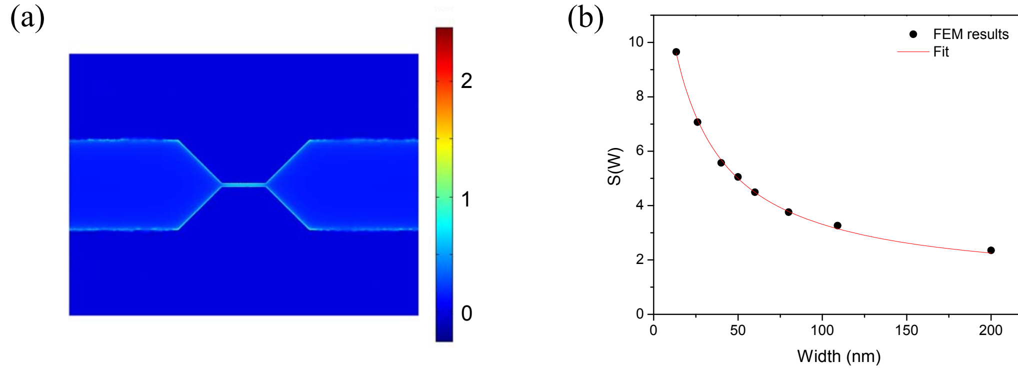

Figure 5.2: (a) Surface charge density on graphene device and SiO_2 surface, from FEM model. Note the increased charge density on the ribbon (W=40 nm, L=500 nm as compared to the 1 μm wide "graphene leads". Color bar represents surface change in mC/m^2(b) Correction factor $S(W) = C_{\text{ribbon}}/C_{\text{parallel plate}}$. Black circles are FEM model results, red curve is a fit to the model results.

capacitance for the same dielectric. This process for determining $S(W)$ was carried out for ribbons of several different widths, and then extrapolated to an empirical rule for ribbons of any width. The scaling factor,

$$S(W) = 1 + \frac{1}{.067 + .0037W} \tag{5.3}$$

was designed to go to 1 for large W by fitting the FEM results to $S(W) = 1 + 1/(a + bW)$. This rule holds only for ribbons with $L = 0.5$ μm. It can be modified by another multiplicative scaling factor to account for the (small) change in capacitance resulting from variations in ribbon length.

The width dependent capacitance and the carrier concentrations due to doping can be folded into Equation 5.1 to obtain expressions for the shift in CNP from either

edge or plane doping in ribbons of varying width.

$$\Delta V_{CNP}^{edge} = \frac{(E \times 10^{14}/W)e}{C_p(1 + \frac{1}{.067+.0037W})} \tag{5.4}$$

$$\Delta V_{CNP}^{plane} = \frac{Pe}{C_p(1 + \frac{1}{.067+.0037W})} \tag{5.5}$$

where we use $C_p = 11.5$ nF/cm^2 for the parallel plate capacitance on our substrates. If E or P are negative, then a positive charge must be induced to bring the Fermi level to the charge neutrality point, and ΔV_{CNP} is negative. The expected CNP shifts for varying levels of edge and plane doping are plotted versus ribbon width in Figure 5.3. We see that edge doping still produces a width dependent effect, with larger CNP shifts for narrower ribbons, and that the CNP shift from plane doping is also dependent on width, with smaller CNP shifts for narrower ribbons.

Ultimately, the width dependent GNR capacitance makes it easier to detect edge-specific doping in GNRs using the width dependence of the CNP shift. The width dependent C_g works to muffle the CNP shift width dependence for edge-specific doping, but it also introduces an opposite width dependence to the CNP shift from plane specific doping. This serves to distinguish edge and plane site doping effects and aids in our goal of detecting charge donation from edge-specific dopants.

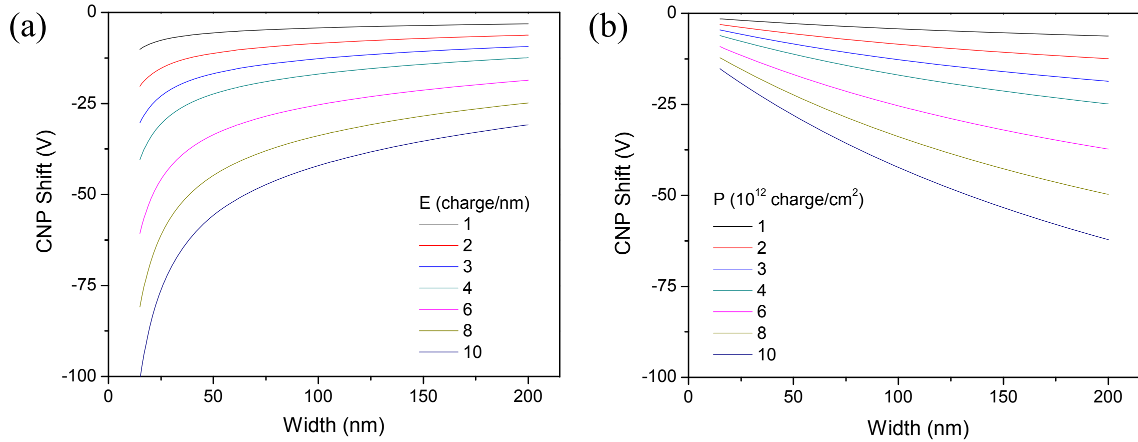

Figure 5.3: Expected charge neutrality point shift as a function of ribbon width for different values of edge and plane doping. (a) Edge doping E is measured in number of charges per nanometer. (b) Plane doping P in number of charges per square centimeter.

5.4 Results and Discussion: CNP shift implies diaminopropane donates charge at graphene edge sites

The results from one set of devices is shown in Figure 5.4. Figure 5.4(a) shows conductance vs. gate voltage sweeps for one ribbon in the set at each stage of the experiment, including before and after each anneal step. Note that although in this analysis we are only extracting the position of the CNP, more information might be obtained by considering other aspects of this same data, such as the field-effect mobility, or the degree of electron-hole symmetry.

In Figure 5.4(b), the position of the CNP is plotted at each experimental stage for each device in the set as a function of ribbon width. In general, the devices are p-type as fabricated, the CNP is brought close to zero during the first annealing step as p-type dopants are driven off, and after subsequent chemical treatments the devices

are, n-type. In Figure 5.4(c), we focus in on two key steps of the experiment: the annealed diaminopropane treated devices, and the annealed methylated devices. In this panel, we plot the shift in CNP relative to the CNP of the same device at the pristine annealed stage, which for most devices was close to zero.

In Figure 5.4(c), the CNP shift for the diaminopropane treatment (black) has a suggestive width dependence. In Figures 5.5(a–c), we focus on fitting this dataset to edge or plane doping. Figure 5.5(a) shows a fit to plane doping only, using equation 5.5 with P as a fitting parameter. Clearly doping in the plane only does not account for the CNP shifts in the data. Figure 5.5(b) shows the data fit to edge doping only, using Equation 5.4 with E as the free parameter. This describes the data quite well. For an even better fit, we fit the data a sum of both expressions, with two free parameters P and E. This results in a curve that fits the data very well, with a large amount of edge doping and a smaller amount of additional plane doping. In Figure 5.5(d), we fit the data from the methylation step to the same two-parameter fitting curve, and find that this data can be roughly described by a model where the edge doping is greatly reduced from its value after the diaminopropane step, and the plane doping is somewhat increased.

The quality of the fits to edge doping for the CNP shifts after diaminopropane treatment suggest that charge donation in the device is occurring from edge selective chemical functionalization. The improved quality of the fit when plane doping is included could mean that either diaminopropane or another molecule is attaching covalently on the plane, perhaps at defect sites, or adsorping with a binding energy

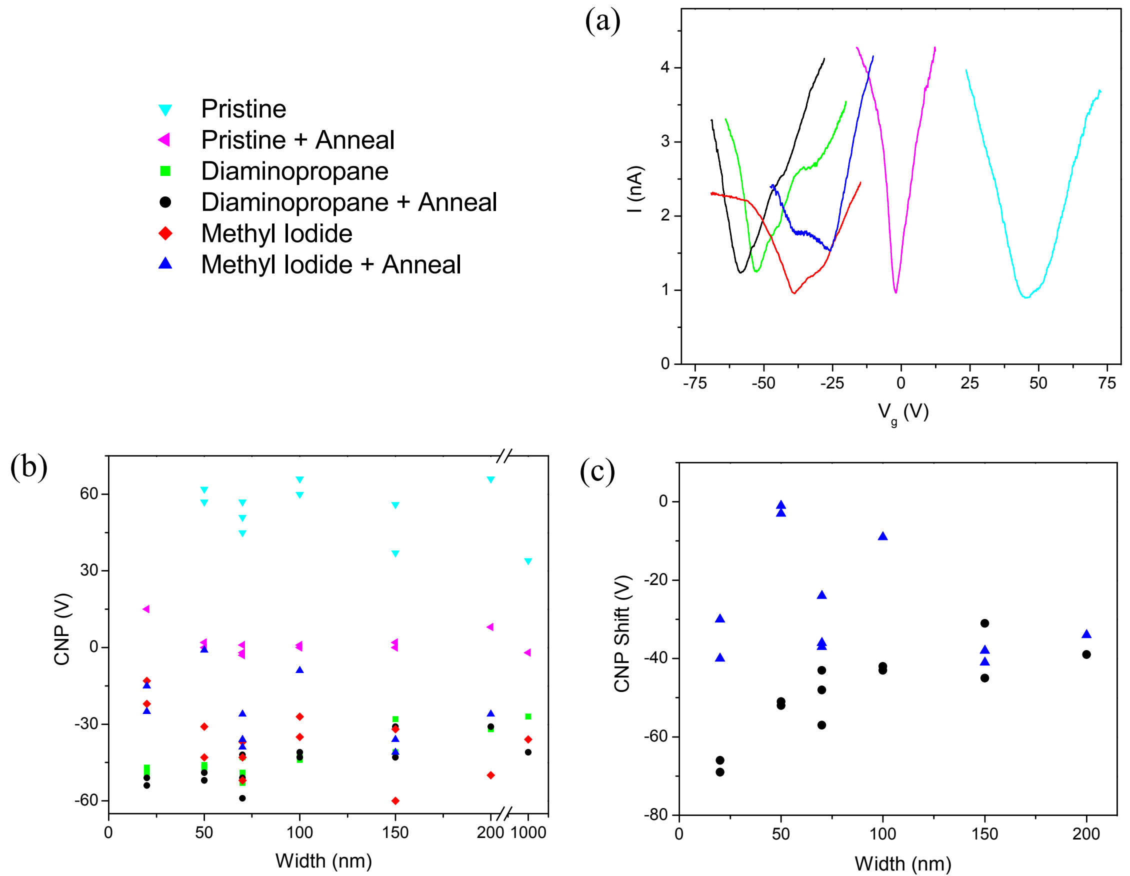

Figure 5.4: (a) Typical charge neutrality point shift corresponding to the annealing and chemical treatment steps. Data shown is for a 70nm wide ribbon. (b) Charge neutrality point for ribbons of different widths after each treatment. (c) Charge neutrality point shift relative to annealed sample before chemical treatment.

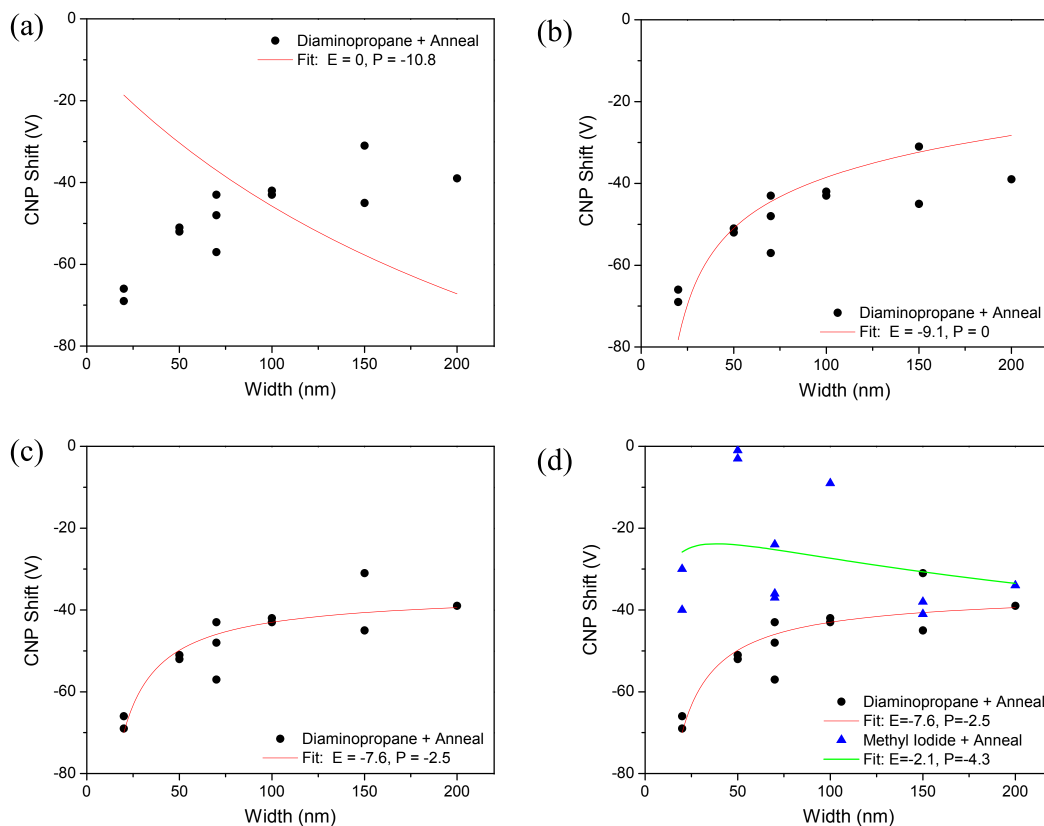

Figure 5.5: Different fits of edge and plane doping levels to the same data set. In all cases, E is edge doping in charge/nm and P is plane doping in 10^{12} charge/nm^2.

too high to be removed by our annealing step. The results for the methyl iodide step are less conclusive, partially due to the large amount of scatter in the data, but the generally positive shift in the charge neutrality point would be consistent with methylation of the diaminopropane.

It should be noted that although edge doping fits well to the data, the amount of edge doping found from the fit is unexpectedly high. The fit in Figure 5.5(c) predicts $E = 7.6/\text{nm}$. For context, a zigzag-edge GNR with one charge donated for every single edge site corresponds to $E = 16.2/\text{nm}$. Even if a molecule bonds to

every exposed edge site, we do not necessarily expect each molecule to donate one full charge; it may only donate a fraction of a charge. The value for E predicted by the fit corresponds to a large fraction of edge sites functionalized with molecules donating a large amount of charge. Alternatively, the ribbon edges could be so so rough that there are many more edge sites per length of ribbon than in a ribbon with clean edges.

5.5 Conclusions and avenues for future work

In this chapter we have presented evidence that treatment of etched graphene nanoribbons with diaminopropane results in charge doping from covalent chemical functionalization that is selective to the graphene edges. We also showed that further treatment with methyl iodide may partially reverse the edge doping effect, suggesting methylation of the attached diaminopropane.

A different experimental design could be used to more simply and clearly prove that a treatment is covalent and selective to the edges. One possibility is to abandon the ribbon geometry of the graphene device, and instead etch the graphene into quantum dots. This would allow direct measurement of the device capacitance, and also result in devices with more edge sites per area. However, detecting chemical functionalization would still rely mainly on the position of the CNP, and thus it would still be heavily dependent on many other environmental factors. A more robust method of probing for edge functionalization would be one that made use of on other techniques, such as imaging, rather than transport. One idea, currently being pursued

by members of the Kim group and the Brus group at Columbia, is to attach molecules that can be detected with florescence microscopy or atomic force microscopy. Using imaging techniques rather than transport would give direct feedback as to the quantity and location of functionalized sites.

Bibliography

[1] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, Physical Properties of Carbon Nanotubes (Imperial College Press, London, 1989).

[2] H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, and R. E. Smalley, Nature **318**, 162 (1985).

[3] S. Iijima, Nature **354**, 56 (1991).

[4] P. R. Wallace, Physical Review **71**, 622 (1947).

[5] A. K. Geim and K. S. Novoselov, Nature Materials **6**, 183 (2007).

[6] M. S. Dresselhaus and G. Dresselhaus, Advances in Physics **30**, 139 (1981).

[7] H. Shioyama, Journal of Materials Science Letters **20**, 499 (2001).

[8] Y. Zhang, J. P. Small, W. V. Pontius, and P. Kim, Applied Physics Letters **86**, 073104 (2005).

[9] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, Science **306**, 666 (2004).

[10] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov, Nature **438**, 197 (2005).

[11] Y. Zhang, Y.-W. Tan, H. L. Stormer, and P. Kim, Nature **438**, 201 (2005).

[12] K. Bolotin, K. Sikes, Z. Jiang, M. Klima, G. Fudenberg, J. Hone, P. Kim, and H. Stormer, Solid State Communications **146**, 351 (2008).

[13] S. M. Sze, Physics of Semiconductor Devices; 2nd ed. (Wiley, New York, 1981).

[14] I. Meric, M. Y. Han, A. F. Young, B. Özyilmaz, P. Kim, and K. L. Shepard, Nature Nanotechnology **3**, 654 (2008).

[15] A. A. Balandin, S. Ghosh, W. Bao, I. Calizo, D. Teweldebrhan, F. Miao, and C. N. Lau, Nano Letters **8**, 902 (2008).

[16] C. Berger, Z. Song, X. Li, X. Wu, N. Brown, C. Naud, D. Mayou, T. Li, J. Hass, A. N. Marchenkov, E. H. Conrad, P. N. First, and W. A. de Heer, Science **312**, 1191 (2006).

[17] Q. Yu, J. Lian, S. Siripongleit, H. Li, Y. P. Chen, and S.-S. Pei, Applied Physics Letters **93**, 113103 (2008).

[18] P. W. Sutter, J.-I. Flege, and E. A. Sutter, Nature Materials **7**, 406 (2008).

[19] X. Li, W. Cai, J. An, S. Kim, J. Nah, D. Yang, R. Piner, A. Velamakanni, I. Jung, E. Tutuc, S. K. Banerjee, L. Colombo, and R. S. Ruoff, Science **324**, 1312 (2009).

[20] S. Bae, H. K. Kim, Y. Lee, X. Xu, J.-S. Park, Y. Zheng, J. Balakrishnan, D. Im, T. L. Y. I. Song, Y. J. Kim, K. S. Kim, B. Özyilmaz, J.-H. Ahn, B. H. Hong, and S. Iijima, arXiv:0912.5485 (2009).

[21] K. S. Kim, Y. Zhao, H. Jang, S. Y. Lee, J. M. Kim, K. S. Kim, J.-H. Ahn, P. Kim, J.-Y. Choi, and B. H. Hong, Nature **457**, 706 (2009).

[22] K. V. Emtsev, A. Bostwick, K. Horn, J. Jobst, G. L. Kellogg, L. Ley, J. L. McChesney, T. Ohta, S. A. Reshanov, J. Roehrl, E. Rotenberg, A. K. Schmid, D. Waldmann, H. B. Weber, and T. Seyller, Nature Materials **8**, 203 (2009).

[23] A. F. Young and P. Kim, Nature Physics **5**, 222 (2009).

[24] T. O. Wehling, K. S. Novoselov, S. V. Morozov, E. E. Vdovin, M. I. Katsnelson, A. K. Geim, and A. I. Lichtenstein, Nano Letters **8**, 173 (2007).

[25] Y.-J. Yu, Y. Zhao, S. Ryu, L. E. Brus, K. S. Kim, and P. Kim, Nano Letters **9**, 3430 (2009).

[26] K. Tanaka, S. Yamashita, H. Yamabe, and T. Yamabe, Synthetic Metals **17**, 143 (1987).

[27] S. E. Stein and R. L. Brown, Journal of the American Chemical Society **109**, 3721 (1987).

[28] M. Ezawa, Physical Review B **73**, 045432 (2006).

[29] D. V. Kosynkin, A. L. Higginbotham, A. Sinitskii, J. R. Lomeda, A. Dimiev, B. K. Price, and J. M. Tour, Nature **458**, 872 (2009).

[30] L. Jiao, L. Zhang, X. Wang, G. Diankov, and H. Dai, Nature **458**, 877 (2009).

[31] M. Y. Han, B. Özyilmaz, Y. Zhang, and P. Kim, Physical Review Letters **98**, 206805 (2007).

[32] Z. Chen, Y.-M. Lin, M. J. Rooks, and P. Avouris, Physica E **40**, 228 (2007).

[33] M. S. Dresselhaus, G. Dresselhaus, and R. Saito, Carbon **33**, 883 (1995).

[34] K. Nakada, M. Fujita, G. Dresselhaus, and M. S. Dresselhaus, Physical Review B **54**, 17954 (1996).

[35] Y. Miyamoto, K. Nakada, and M. Fujita, Physical Review B **59**, 9858 (1999).

[36] K. Wakabayashi, M. Fujita, H. Ajiki, and M. Sigrist, Physical Review B **59**, 8271 (1999).

[37] Y.-W. Son, M. L. Cohen, and S. G. Louie, Nature **444**, 347 (2006).

[38] K. A. Ritter and J. W. Lyding, Nature Materials **8**, 235 (2009).

[39] L. Brey and H. A. Fertig, Physical Review B **73**, (2006).

[40] Y. Son, M. L. Cohen, and S. G. Louie, Physical Review Letters **97**, (2006).

[41] V. Barone, O. Hod, and G. E. Scuseria, Nano Letters **6**, 2748 (2006).

[42] P. Blake, E. W. Hill, A. H. C. Neto, K. S. Novoselov, D. Jiang, R. Yang, T. J. Booth, and A. K. Geim, Applied Physics Letters **91**, 063124 (2007).

[43] A. C. Ferrari, J. C. Meyer, V. Scardaci, C. Casiraghi, M. Lazzeri, F. Mauri, S. Piscanec, D. Jiang, K. S. Novoselov, S. Roth, and A. K. Geim, Physical Review Letters **97**, 187401 (2006).

[44] Y.-W. Tan, Y. Zhang, H. L. Stormer, and P. Kim, The European Physical Journal Special Topics **148**, 15 (2007).

[45] L. Kouwenhoven, C. Marcus, P. McEuen, S. Tarucha, R. Westervelt, and N. Wingreen, in Proceedings of the Advanced Study Institute on Mesoscopic Electron Transport, edited by L. Sohn, L. Kouwenhoven, and G. Schön (Kluwer, Dordrecht, 1997).

[46] F. Sols, F. Guinea, and A. H. C. Neto, Physical Review Letters **99**, 166803 (2007).

[47] M. Y. Han, J. C. Brant, and P. Kim, Physical Review Letters **104**, 056801 (2010).

[48] P. W. Anderson, Physical Review **109**, 1492 (1958).

[49] N. F. Mott, Advances in Physics **16**, 49 (1967).

[50] N. F. Mott, Journal of Non-Crystalline Solids **1**, 1 (1968).

[51] B. I. Shklovskii and A. L. Efros, Electronic Properties of Doped Semiconductors (Springer, New York, 1984).

[52] M. M. Fogler, S. Teber, and B. I. Shklovskii, Physical Review B **69**, 035413 (2004).

[53] D. V. Averin and K. K. Likharev, Journal of Low Temperature Physics **62**, 345 (1986).

[54] D. Querlioz, Y. Apertet, A. Valentin, K. Huet, A. Bournel, S. G. Retailleau, and P. Dollfus, Applied Physics Letters **92**, 042108+ (2008).

[55] I. Martin and Y. M. Blanter, Physical Review B **79**, 235132 (2009).

[56] M. Evaldsson, I. V. Zozoulenko, H. Xu, and T. Heinzel, Physical Review B **78**, 161407 (2008).

[57] Y. Yoon and J. Guo, Applied Physics Letters **91**, 073103 (2007).

[58] D. Gunlycke, D. A. Areshkin, and C. T. White, Applied Physics Letters **90**, 142104 (2007).

[59] E. R. Mucciolo, A. H. Castro Neto, and C. H. Lewenkopf, Physical Review B **79**, 075407 (2009).

[60] G. Schubert, J. Schleede, and H. Fehske, Physical Review B **79**, 235116 (2009).

[61] S. Adam, S. Cho, M. S. Fuhrer, and S. D. Sarma, Physical Review Letters **101**, 046404 (2008).

[62] J. Martin, N. Akerman, G. Ulbricht, T. Lohmann, J. H. Smet, K. von Klitzing, and A. Yacoby, Nature Physics **4**, 144 (2008).

[63] F. Molitor, A. Jacobsen, C. Stampfer, J. Güttinger, T. Ihn, and K. Ensslin, Physical Review B **79**, 075462 (2009).

[64] C. Stampfer, J. Güttinger, S. Hellmüller, F. Molitor, K. Ensslin, and T. Ihn, Physical Review Letters **102**, 056403 (2009).

[65] K. Todd, H. Chou, S. Amasha, and D. Goldhaber-Gordon, Nano Letters **9**, 416 (2009).

[66] P. Gallagher, K. Todd, and D. Goldhaber-Gordon, arXiv:0909.3886v1 (2009).

[67] X. Liu, J. B. Oostinga, A. F. Morpurgo, and L. M. K. Vandersypen, Physical Review B **80**, 121407 (2009).

[68] X. Liu, J. B. Oostinga, A. F. Morpurgo, and L. M. Vandersypen, arXiv:0812.4038v1 .

[69] J. Güttinger, C. Stampfer, F. Libisch, T. Frey, J. Burgdörfer, T. Ihn, and K. Ensslin, Physical Review Letters **103**, 046810 (2009).

[70] C. Stampfer, E. Schurtenberger, F. Molitor, J. Güttinger, T. Ihn, and K. Ensslin, Nano Letters **8**, 2378 (2008).

[71] C. Stampfer, J. Güttinger, F. Molitor, D. Graf, T. Ihn, and K. Ensslin, Applied Physics Letters **92**, 012102 (2008).

[72] J. Güttinger, C. Stampfer, S. Hellmüller, F. Molitor, T. Ihn, and K. Ensslin, Applied Physics Letters **93**, 212102 (2008).

[73] S. Schnez, F. Molitor, C. Stampfer, J. Güttinger, I. Shorubalko, T. Ihn, and K. Ensslin, Applied Physics Letters **94**, 012107 (2009).

[74] A. H. C. Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Reviews of Modern Physics **81**, 109 (2009).

[75] A. S. Rodin and M. M. Fogler, Physical Review B **80**, 155435 (2009).

[76] Y. Zhang, V. W. Brar, C. Girit, A. Zettl, and M. F. Crommie, Nature Physics **5**, 722 (2009).

[77] E. Rossi and S. Das Sarma, Physical Review Letters **101**, 166803 (2008).

[78] S. Marianer and B. I. Shklovskii, Physical Review B **46**, 13100 (1992).

[79] H. Xu, T. Heinzel, and I. V. Zozoulenko, Physical Review B **80**, 045308 (2009).

[80] Z. Yao, C. L. Kane, and C. Dekker, Physical Review Letters **84**, 2941 (2000).

[81] Y.-F. Chen and M. S. Fuhrer, Physical Review Letters **95**, 236803 (2005).

[82] S. Fratini and F. Guinea, Physical Review B **77**, 195415 (2008).

[83] M. V. Fischetti, D. A. Neumayer, and E. A. Cartier, Journal of Applied Physics **90**, 4587 (2001).

[84] J.-H. Chen, C. Jang, S. Xiao, M. Ishigami, and M. S. Fuhrer, Nature Nanotechnology **3**, 206 (2008).

[85] M. Freitag, M. Steiner, Y. Martin, V. Perebeinos, Z. Chen, J. C. Tsang, and P. Avouris, Nano Letters **5**, 1883 (2009).

[86] D.-H. Chae, B. Krauss, K. von Klitzing, and J. H. Smet, Nano Letters **2**, 466 (2010).

[87] S. Berciaud, M. Y. Han, L. E. Brus, P. Kim, and T. F. Heinz, Under review by Physical Review Letters .

[88] X. Li, X. Wang, L. Zhang, S. Lee, and H. Dai, Science **319**, 1229 (2008).

[89] X. Jia, M. Hofmann, V. Meunier, B. G. Sumpter, J. Campos-Delgado, J. M. Romo-Herrera, H. Son, Y. Hsieh, A. Reina, J. Kong, M. Terrones, and M. S. Dresselhaus, Science **323**, 1701 (2009).

[90] L. C. Campos, V. R. Manfrinato, J. D. Sanchez-Yamagishi, J. Kong, and P. Jarillo-Herrero, Nano Letters **9**, 2600 (2009).

[91] S. S. Datta, D. R. Strachan, S. M. Khamis, and A. T. C. Johnson, Nano Letters **8**, 1912 (2008).

[92] L. Tapasztó, G. Dobrik, P. Lambin, and L. P. Biró, Nature Nanotechnology **3**, 397 (2008).

[93] S. Masubuchi, M. Ono, K. Yoshida, K. Hirakawa, and T. Machida, Applied Physics Letters **94**, 082107 (2009).