Electron transport in one and two dimensional materials

Samuel William Lagasse

University at Albany, State University of New York

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Condensed Matter Physics Commons, Nanoscience and Nanotechnology Commons, and the Quantum Physics Commons

Recommended Citation

https://scholarsarchive.library.albany.edu/legacy-etd/2317

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive. Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
Electron transport in one and two dimensional materials

by

Samuel William LaGasse

A Dissertation
Submitted to the University at Albany, State University of New York
in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy

College of Nanoscale Science and Engineering
2019
Abstract

This dissertation presents theoretical and experimental studies in carbon nanotubes (CNTs), graphene, and van der Waals heterostructures. The first half of the dissertation focuses on cutting edge tight-binding-based quantum transport models which are used to study proton irradiation-induced single-event effects in carbon nanotubes [1], total ionizing dose effects in graphene [2], quantum hall effect in graded graphene $p$-$n$ junctions [3], and ballistic electron focusing in graphene $p$-$n$ junctions [4]. In each study, tight-binding models are developed, with heavy emphasis on tying to experimental data. Once benchmarked against experiment, properties of each system which are difficult to access in the laboratory, such as local density of states, local current density, and quantum transmission probability, are extracted to build our physical intuition. The second half of the dissertation covers experimental work on transport in van der Waals heterostructures. High-quality samples, evidenced by measurements of quasi-ballistic graphene $p$-$n$ junctions, are enabled by encapsulation in hexagonal boron nitride, assembled using a modified dry transfer technique. The Schottky-Mott limit, previously only a textbook example, is probed in gated graphene-WSe$_2$ heterojunctions [5]. Schottky barrier measurements as a function of gate voltage reveal perfect barrier tuning, following the Schottky-Mott rule. Enabled by the lack of Fermi-level pinning at the graphene-WSe$_2$ interface, a method for dynamically tuning the Schottky diode ideality factor is demonstrated. Finally, an analytical model describing tuning of the junction is developed.
To my parents.
Acknowledgments

Throughout the completion of my doctorate, I have become indebted to many people. While each Ph.D. is awarded to a single person, each one is the culmination of the influence of many separate people on an individual. In this section, I will do my best to acknowledge those people who have strongly influenced me in the years leading up to my defense.

Throughout my Ph.D., I have been closely mentored by my advisor, Professor Ji Ung Lee. Dr. Lee has taught me so much that it is hard to put to words- how to solve problems as a physicist, how to be critical of ones work, how to distill a problem to its most important and relevant components. He has taught me how to carry and present myself as a professional scientist. The level of patience he has and the care with which he teaches, even for seemingly basic topics, has been extremely important in my development.

I am so grateful for the numerous extensive conversations Dr. Lee and I have had in his office, often longer than an hour, and usually without prior scheduling. One conversation in particular happened around the time I began to focus primarily on experimental physics. At this point, Dr. Lee had already expressed that I could write my thesis and defend based off my theoretical work. Despite this, Dr. Lee was willing to give me a chance to move into this new and exciting area. As I move on in my scientific career, I will always be thankful to Dr. Lee for giving me this chance.

I would like to acknowledge all of the Lee-lab members, past and present, for their contributions throughout my Ph.D. Everett Comfort and Surajit Sutar, former students whom I unfortunately did not overlap with in the lab, were instrumental in the creation of the substrates used throughout the later chapters of this dissertation. Paul David (another former student from the lab) and I have known each other as far back as my undergraduate days, and we have become very good friend throughout the course of graduate school. Paul trained me on a great deal of the different tools in our lab and was always extremely helpful and willing to lend a hand.

I am also very appreciative to Luke Blanco and Sharadh Jois, with whom I have spent many hours with in the lab. Additionally, I have had the pleasure of working with an extremely bright student, Collin Sanborn, who has assisted me a great deal, beginning with
theory work and moving into experiment. I look forward to seeing where he goes over the next several years. I would be remiss not to thank Prathamesh Dhakras, whom of everyone, I have likely worked the most closely with. Prathamesh has been my primary partner in developing the pick-and-place technique for our lab and we spent many hours performing measurements together in the development of the Schottky barrier work presented later in this dissertation. To everyone in Dr. Lee’s lab, I am very appreciative for all the help and all of the discussions, and I wish them all the best.

Outside of Dr. Lee’s lab, there are a number of graduate students which deserve thanks. My good friends Horace Bull, Ryan Porrazzo, Leigh and Lauren Lydecker, and Michael T. Murphy have all been such a good support group throughout my studies, it have not been nearly as much fun without them. Horace was the best man at my wedding and discussions with him have always been a cathartic experience. Mike has been a great friend not just personally, but in the lab as well. He was instrumental in the formulation of the solutions used in pick-and-place, and a constant source of good advice for any polymer-related science, and lab work in general.

Outside of students at CNSE, I am indebted to the academic support group. They provide so much day-to-day assistance in performing research- without them, I don’t know how I would succeeded in experimental work. They tirelessly support graduate education at CNSE, their contribution is so important, I know I cannot do it justice. Steve Warfield has trained me on nearly every fabrication tool I use, most importantly the electron beam evaporator. Any issue I have had, he has always responded in less than an hour- his assistance has been invaluable. Stephen Stewart has been such a help in almost all areas of graduate school. Steve has been a fantastic sounding board for various research problems, and is always willing to lend a hand (or tools) when solving problems with research apparatuses. The HP4156B unit which I performed most of my electrical measurements arrived in our lab thanks to Steve saving it from the trash! Kyle Unser helped me a great deal in the design and machining of my 2D material transfer stage, going through a number of iterations without any complaints- I expect his contribution will help a number of graduate students over the next several years. Finally, Brian Taylor deserves huge thanks, he helps out in so many ways, but one vivid story I have in my mind is when our $5000 order of liquid helium arrived and
did not have the proper fitting for our transfer line. We called Brian and told him we had an emergency—within a few hours he came by with a custom machined part that worked perfectly—enough said, I think.

Cory D. Cress has been a collaborator and mentor of mine from Naval Research Lab, from which led to two of the papers which will be discussed in this dissertation. Cory has provided a great deal of assistance in that work, especially in the area of writing and presenting physics research. His advice has been invaluable and I have learned quite a bit from his help. I greatly look forward to seeing what will come from our future collaboration.

Without the love of my parents, Anne and Bernie, I know I would not have gotten to this point at all. Their constant love and support over the years goes beyond anything I could even imagine. Growing up, they vigorously supported any activity or hobby (of which there were many) which I showed interest in, which in no small part led me to studying physics. Their contribution is so immense, there is no possible way I can do it justice here. I simply hope that over the years I continue to be worthy of their support and make them proud.

Finally, I need to acknowledge my strongest supporter, my wife, Kate. I met Kate shortly after joining Dr. Lee’s lab, and knew I wanted to marry her almost immediately. She is my best friend and has been my greatest supporter these last several years. I am constantly in awe of her spirit and kindness. I am so thankful that she found me, and I am glad for every moment which led up to us meeting. Kate has inspired me to persevere through the hard times of research, without her I likely would have quit. If any of the results in this dissertation can be considered of high-quality, it is because of her.
Contents

1 Foreword 1
  1.1 Motivation ................................................. 1
  1.2 Outline for this thesis .................................... 4

I Quantum transport theory in carbon nanotubes and graphene 7

2 Introduction 8
  2.1 Graphene- basic electronic properties ......................... 8
  2.1.1 Tight-binding model for simulations ........................ 14
  2.2 Quantum transport basics .................................. 15
  2.3 Graphene p-n junctions ..................................... 19
    2.3.1 Angle-dependent transport and Klein tunneling .......... 19
    2.3.2 Angle-dependent transport and Klein tunneling .......... 24
    2.3.3 Conductance of a graphene p-n junction .................. 24
  2.4 Graphene in magnetic fields ................................ 26
  2.5 Tight-binding transport calculations ........................ 32

3 Quantum transport modeling of carbon nanotubes under proton irradia-
tion 37
  3.1 Introduction ............................................... 37
  3.2 Transport Model ........................................... 38
    3.2.1 Non-equilibrium Green’s Function method ................. 38
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.2</td>
<td>CNT Hamiltonian</td>
<td>40</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Computational details</td>
<td>41</td>
</tr>
<tr>
<td>3.3</td>
<td>Results and Discussion</td>
<td>42</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Local Density of States</td>
<td>44</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Transmission</td>
<td>44</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Drain Current</td>
<td>47</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Ion adsorbates</td>
<td>47</td>
</tr>
<tr>
<td>3.3.5</td>
<td>Encapsulated CNTs</td>
<td>51</td>
</tr>
<tr>
<td>3.4</td>
<td>Chapter conclusions</td>
<td>51</td>
</tr>
<tr>
<td>4</td>
<td>Quantum transport modeling of graphene devices with radiation induced Coulomb scatterers</td>
<td>53</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>53</td>
</tr>
<tr>
<td>4.2</td>
<td>Transport model</td>
<td>56</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Contact model</td>
<td>56</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Coulomb scatterer model</td>
<td>57</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Gate voltage considerations</td>
<td>58</td>
</tr>
<tr>
<td>4.3</td>
<td>Results and Discussion</td>
<td>59</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Transfer Characteristics</td>
<td>59</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Transport visualization</td>
<td>63</td>
</tr>
<tr>
<td>4.4</td>
<td>Chapter conclusions</td>
<td>68</td>
</tr>
<tr>
<td>5</td>
<td>Landau level mixing in heavily graded graphene p-n junctions</td>
<td>71</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction- Quantum Hall effect in graphene p-n junctions</td>
<td>71</td>
</tr>
<tr>
<td>5.2</td>
<td>Quantum transport model</td>
<td>75</td>
</tr>
<tr>
<td>5.3</td>
<td>Results and Discussion</td>
<td>77</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Benchmarking against experiment</td>
<td>79</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Mapping Landau level transport</td>
<td>82</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Factors influencing Landau level mixing</td>
<td>84</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Landau level shape</td>
<td>86</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Chapter conclusions</td>
<td>89</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>7.4.2</td>
<td>Electron beam lithography</td>
<td></td>
</tr>
<tr>
<td>7.4.3</td>
<td>Heterostructure etching</td>
<td></td>
</tr>
<tr>
<td>7.4.4</td>
<td>Metallization</td>
<td></td>
</tr>
<tr>
<td>7.5</td>
<td>Electrical measurements</td>
<td></td>
</tr>
<tr>
<td>7.6</td>
<td>Transport case studies</td>
<td></td>
</tr>
<tr>
<td>7.6.1</td>
<td>Graphene p-n junction device</td>
<td></td>
</tr>
<tr>
<td>7.6.2</td>
<td>Graphene-WSe$_2$ hybrid devices</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Gate-tunable graphene-WSe$_2$ ideal Schottky diodes at the Schottky-Mott limit</td>
</tr>
<tr>
<td>8.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>8.2</td>
<td>Results and discussion</td>
</tr>
<tr>
<td>8.2.1</td>
<td>Electrical characterization</td>
</tr>
<tr>
<td>8.2.2</td>
<td>Schottky barrier height extraction</td>
</tr>
<tr>
<td>8.2.3</td>
<td>Capacitance model</td>
</tr>
<tr>
<td>8.2.4</td>
<td>Dynamic ideality factor tuning</td>
</tr>
<tr>
<td>8.3</td>
<td>Remarks on measurement stability, yield</td>
</tr>
<tr>
<td>8.4</td>
<td>Chapter conclusions</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>Graphene crystal structure</td>
<td>9</td>
</tr>
<tr>
<td>2-2</td>
<td>Tight-binding band structure of graphene</td>
<td>12</td>
</tr>
<tr>
<td>2-3</td>
<td>Comparison of the low energy band structure of graphene and semiconductors</td>
<td>13</td>
</tr>
<tr>
<td>2-4</td>
<td>Electronic transport regimes</td>
<td>17</td>
</tr>
<tr>
<td>2-5</td>
<td>Depiction of the Fermi window</td>
<td>20</td>
</tr>
<tr>
<td>2-6</td>
<td>Graphene $p$-$n$ junction schematic</td>
<td>21</td>
</tr>
<tr>
<td>2-7</td>
<td>Angle-dependent transmission function of a symmetric graphene $p$-$n$ junction</td>
<td>23</td>
</tr>
<tr>
<td>2-8</td>
<td>Transverse modes in graphene</td>
<td>25</td>
</tr>
<tr>
<td>2-9</td>
<td>Ballistic conductance asymmetry in graphene $p$-$n$ junctions</td>
<td>27</td>
</tr>
<tr>
<td>2-10</td>
<td>Comparison of zigzag graphene nanoribbon tight-binding band structure in zero magnetic field and a large magnetic field</td>
<td>29</td>
</tr>
<tr>
<td>2-11</td>
<td>Half-integer quantum Hall effect in graphene</td>
<td>30</td>
</tr>
<tr>
<td>2-12</td>
<td>Edge vs. bulk transport for graphene in the quantum Hall regime</td>
<td>32</td>
</tr>
<tr>
<td>2-13</td>
<td>Generalized scattering problem for a sample with several leads</td>
<td>33</td>
</tr>
<tr>
<td>3-1</td>
<td>3D model summarizing the simulations performed in this chapter</td>
<td>43</td>
</tr>
<tr>
<td>3-2</td>
<td>CNT local density of states under the influence of an ion at different distances</td>
<td>45</td>
</tr>
<tr>
<td>3-3</td>
<td>Transmission function calculations for a CNT under the effect of ions at various distances</td>
<td>46</td>
</tr>
<tr>
<td>3-4</td>
<td>Comparison of the simulated drain current for different ion distances with the results of Bushmaker <em>et al</em></td>
<td>48</td>
</tr>
<tr>
<td>3-5</td>
<td>LDOS and transmission calculation assuming an ion adsorbate with charge transfer to the CNT</td>
<td>49</td>
</tr>
</tbody>
</table>
3-6 Resonant tunneling effects resulting from ion adsorbates .......................... 50
3-7 Comparison of the effect of different dielectric environments .................... 52

4-1 Visual summary of the graphene TID problem ............................................. 54
4-2 Electric potential due to trapped charges at different densities .................. 58
4-3 Benchmarking tight-binding simulation against the experimental data of Cress

 et al .................................................. 64
4-4 Simulated transmission function corresponding to the current calculation in

 Fig. 4-3a .................................................. 65
4-5 Energy-resolved LDOS in the presence of a trapped charge ....................... 67
4-6 Tight-binding calculation of local current density in the presence of trapped

 charges .................................................. 69

5-1 Experimental studies on QHE in graphene p-n junctions ................................. 74
5-2 Schematic of quantum Hall transport in a graphene p-n junction .................. 78
5-3 Quantum Hall simulation benchmarking ..................................................... 81
5-4 Maps of quantum Hall transport in a disordered graphene p-n junction at

 small and large junction widths .................................................. 84
5-5 Mapping Landau level transport at different Fermi energies ....................... 85
5-6 Junction width dependence on the quantum Hall conductance of a graphene

 p-n junction .................................................. 87
5-7 Comparison of individual Landau levels separated at an extremely graded p-n

 junction with analytical expression derived by Lukose et al .......................... 90

6-1 Early transverse electron focusing experiment in GaAs 2DEGs ..................... 93
6-2 Device for ballistic electron focusing by Chen et al ................................. 95
6-3 Schematics for transverse electron focusing modeling ................................ 98
6-4 Comparison of particle current density with and without dephasing edge contacts100
6-5 Off-resonance characteristics of the symmetric graphene p-n junction .......... 103
6-6 Benchmarking tight-binding transport model against results of Chen et al .... 106
6-7 Slices through the simulated non-local resistance map ................................ 108
6-8 Spatially resolved maps of current density in the first and second TEF resonance configurations ........................................... 110
6-9 Mode resolved current density .................................................. 111
7-1 Crystal structure of MX$_2$ transition metal dichalcogenides ................. 118
7-2 van der Waals heterostructures .................................................. 119
7-3 Demonstration of the mechanical exfoliation technique ........................ 128
7-4 Example of an early encapsulated graphene sample .......................... 131
7-5 Hot transfer stage for performing the van der Waals assembly process .... 132
7-6 Example process flow for the creation of a stamp used in the transfer process 134
7-7 Interfacial cleaning in the graphene pick-up step ............................. 140
7-8 Example of a pristine encapsulated graphene sample resulting from the interfacial cleaning technique .......................................... 142
7-9 Example of interfacial cleaning of a completed heterostructure during the final set-down .......................................................... 143
7-10 Example of an etched h-BN encapsulated graphene sample .................. 149
7-11 Cross-sectional schematic of a edge contact to graphene ................... 150
7-12 Basic schematic depicting the metal lift-off process used in this dissertation 151
7-13 Heterostructure and layout of a high quality graphene p-n junction device 153
7-14 Transport measurements on a high quality graphene p-n junction .......... 155
7-15 Drude mobility and mean free path extracted from the slices in Fig. 7-14b 156
7-16 Two methods of assembling graphene-contacted WSe$_2$ devices encapsulated in h-BN .......................................................... 158
7-17 Two completed graphene-WSe$_2$ hybrid devices using the two different process flows in Fig. 7-16 ........................................... 159
7-18 Electrical characterization of a reconfigurable graphene-WSe$_2$ device .... 161
8-1 Textbook-style example of the formation of an n-type Schottky diode which follows the Schottky-Mott rule ........................................ 165
8-2 Demonstration of the effect of Fermi-level pinning on the Schottky barrier height from Myburg et al ............................................. 167
8-3 Images and schematics of the Schottky barrier device studied in this chapter ............... 169
8-4 AFM image of the device studied in this chapter ............................................... 170
8-5 Electrical characterization of the graphene-WSe$_2$ device .................................... 172
8-6 Demonstration of diode fitting and leakage current extrapolation below the measureable current limit ......................................................... 174
8-7 Extraction of Schottky barrier heights as a function of gate voltage ..................... 176
8-8 Schematic depiction of the capacitor model used to describe our perfect tuning of the Schottky barrier height .............................................. 178
8-9 How to achieve dynamic tuning of the diode ideality factor ................................. 180
8-10 Demonstration of dynamic control over the diode ideality factor ....................... 181
8-11 More examples of dynamically tuned diode measurements .............................. 183
8-12 Demonstration of decoupling of the diode leakage current and $n_{\text{Eff}}$ ........ 184
8-13 Demonstration of measurement repeatability .................................................... 186
8-14 Measurements on an alternate Schottky diode sample .................................... 188
Chapter 1

Foreword

1.1 Motivation

Semiconductor-based electrical devices have historically acted as a fantastic platform for studying solid state physics. A flurry of experimental discoveries culminating the point contact transistor by Bardeen, Brattain, and Shockley in 1947 (for which they received the Nobel Prize in physics for in 1956) cemented semiconductors as a fixture of modern physics [6]. Additionally, the creation of the metal oxide semiconductor field effect transistor (MOSFET) by John Atalla and Dawon Kahng in 1959 represented a paradigm shift for modern electronics. Indeed, since then, integrated circuits formed from billions of aggressively scaled MOSFETs have transformed modern society as well.

The creation of the MOSFET drew massive attention in the world of physics and acted as the basis for many modern experiments in electronics and condensed matter physics. Physicists throughout the world sought to understand how electrons behaved in these devices, resulting in some earth-shaking discoveries. One outstanding question regarded resistance in mesoscopic devices. Physicists asked; “if an electron does not scatter as it transports through a device, will the resistance be zero?”. History has shown that the resistance is, in fact, not zero, but the story of how we got there sets the stage for the modern field of quantum transport.

In 1957, Rolf Landauer proposed a new viewpoint on electron transport in solids which was in contrast with the standard Drude picture [7]. Landauer derived a simple equation
which re-envisioned solids as electron waveguides. In the Landauer picture, the conductance of a wire was not simply from current flowing due to an external electric field, but from the transmission probability of electrons populating quantum modes. While it would take some years for Landauer’s picture to be verified, he did provide an important viewpoint on transport which motivated the field.

A series of experiments stemmed from the creation of the first MOSFET, each of which have driven the efforts of countless physicists over the years. One of the most important discoveries was the integer quantum Hall effect by Klaus von Klitzing. von Klitzing studied a seven terminal degenerately n-doped silicon MOSFET under a high magnetic field (up to 15 Tesla) and observed something particularly peculiar [8].

Classically, the Hall conductivity continuously and linearly was varied by the carrier density of the sample. von Klitzing, however, observed that the Hall conductance came in discrete, integer values, of $e^2/h$, $e$ the charge of the electron and $h$ Planck’s constant. Indeed, as von Klitzing swept the magnetic field applied to his device, he observed that the Hall conductance did not continuously change, but was quantized. For this discovery, von Klitzing received the Nobel prize in physics in 1985. Since then, quantum Hall effect measurements have been ubiquitous in the area of mesoscopic physics and have emerged as an extremely important tool for understanding electron transport.

Whilst von Klitzing’s discovery established the existence of transport quantization in mesoscopic systems, it did not fully put to bed the questions posed regarding ballistic electron transport. It was not until 1988, when B.J. van Wees et al discovered the quantum point contact in a high mobility two-dimensional electron gas (2DEG) inside of a gallium arsenide quantum well [9]. By using a pair of split-gates, van Wees et al electrostatically defined a constriction in the 2DEG. When the constriction was sufficiently small, comparable to the Fermi wavelength, the conductance measured through the device became quantized!

By sweeping the split-gate voltages, making the constriction smaller, steps of $2e^2/h$ emerged in the measurement of what we now refer to as a quantum point contact (QPC). This experiment represented a major leap forward, providing strong evidence for Landauer’s transport picture from the 1950s. Thanks to the QPC experiment, we began to understand that electron transport in the absence of scattering- ballistic transport- does not occur with
an absence of resistance.

Shortly after the original QPC experiment, the same group published results from a transverse electron focusing (TEF) experiment exploiting the QPC geometry [10]. By using small magnetic fields to bend the trajectory of electrons emitted by one QPC and collecting them in another QPC, van Houten et al observed interesting peaks in the non-local resistance when measured in a Hall bar geometry. These peaks emerged when the cyclotron radius of the electrons matched integer fractions of the spacing between the QPC. These peaks in the resistance represented ballistic skipping orbits which the electrons took followed in the device. More interestingly, however, were the aberrations in the focusing spectra which could only be understood using quantum mechanical interference effects.

The QPC experiment and subsequent TEF experiments made it evident that the classical Drude model of transport would not work well for describing modern transport experiments. One convenient way to understand these modern transport problems is through the Landauer-Büttiker transport model [11]. However, Landauer’s convenient “transmission is conductance” model would necessitate an upgrade to our modeling toolbox. For that, we look to quantum mechanics, giving rise to the field of quantum transport.

The QPC experiment spawned a series of quantum transport measurements on 2DEG heterostructures. As the sample quality was improved, researchers routinely observed ballistic transport. Many studies have been performed since then, creating exotic structures like quantum dots, zero dimensional artificial atoms based off the original QPC design. While these experiments allowed researchers to aggressively probe quantum transport, the work typically would be done in dilution refrigerators at milli-Kelvin temperatures. However, starting in 2004, a series of experiments were reported on a new material made from a single layer of carbon atoms, graphene, which gave us the prospect of ballistic transport at room temperature [12][13]. The synthesis of graphene using the “scotch tape” method and further electrical measurements by Novoselov and Geim [12][13] represented a paradigm shift in the world of condensed matter physics. Their discoveries lead to a new age of physics focusing on the study of two-dimensional materials, which persists today.
1.2 Outline for this thesis

This thesis will describe my work in Professor Ji Ung Lee’s lab over the last several years. It will be organized into two parts: Part I describes the theoretical and modeling work which I focused on at the beginning of my work in Professor Lee’s lab. The work focuses on quantum transport modeling of experiments performed on carbon nanotubes and graphene and have been summarized in four publications [1][3][2][4]. While the research areas vary, the goal throughout the work was to consistently connect theoretical models with experimental results. The logical conclusion of this work is Part II of my thesis, or as Professor Lee refers to it, “my second thesis”, which relates to experimental nanoelectronics in 2D materials.

Chapters 2-6, section 6 of Chapter 7, and chapter 8 will all be presented in a pedagogical manner where I will be switching to the third-person voice as I guide the reader through the different topics. However, for sections 1-5 of Chapter 7 I will switch back to the first-person voice, where I will describe my personal experiences in learning nanoscale device fabrication methods. These changes in voice are intentionally chosen to best convey the topics for each section.

In Chapter 2 I will introduce the fundamentals of graphene as well as the quantum transport theory which I use throughout this dissertation. I will discuss the origin of Dirac electrons in graphene and highlight the interesting implications from that, such as the half-integer quantum hall effect and angle-dependent transport in graphene p-n junctions. From there, I will look how tight-binding transport models can be used to understand transport in these systems.

Chapters 3 and 4 cover the details of my radiation effects modeling work. During my graduate study I wrote two publications on radiation effects in carbon nanoelectronics [1][2] in collaboration with Hap Hughes and Cory D. Cress of the US Naval Research Laboratory. The first paper modeled a proton irradiation study of a suspended carbon nanotube transistor. I used the non-equilibrium Green’s function (NEGF) technique to model electron transport in a carbon nanotube subjected to the electric potential of a nearby ion. In the second paper, I applied a similar tight-binding transport model to study total ionizing dose effects due to gamma radiation in graphene devices.
In Chapter 5 I will describe the first magnetotransport study I performed, modeling quantum Hall measurements made on heavily graded graphene p-n junctions (GPNJs) [3]. In this work I modeled experiments performed at NIST, which showed interesting Landau level mixing effects in their devices. I found that by modeling a smooth p to n transition region in the GPNJ, as well as by adding disorder and roughness to the junction, that I could match the experimental results with tight-binding transport. With this benchmark in hand, I was able to calculate maps of the transport within the device, building out our understanding of the QHE in GPNJs.

While Chapter 5 discusses a high magnetic field transport study in graphene, Chapter 6 describes a low magnetic field study [4]. Shortly after the publication of my QHE work, interesting results were being shown in the literature on magnetic focusing in graphene. This inspired me to see if the same model which described QHE in GPNJs could capture the vibrant results being demonstrated in the highest quality graphene samples available. Indeed, it could, however to succeed in modeling these experiments I found I had to include dephasing into the model. With proper application of the Landauer-Büttiker equation, the match between the model and experiment was striking. From there, spatially resolved maps of the current density in the model unveiled a fascinating interplay between classically predicted cyclotron orbits, quantum interference, and quantum Hall states.

During the process of research for my first four publications, the major focus was modeling experimental results. Indeed, the models were only as good as their agreement with experiment, and to do this I found I had to think like an experimental physicist. The natural course of events, seeking more experimental results to model, led to me trying out experimental device work for myself. In Chapter 7 I will discuss the experimental techniques which I used in our lab for the second part of my thesis. In particular, I will describe my efforts in developing the "pick and place" technique for our lab, whereby I used hexagonal Boron Nitride (h-BN) supplied by our collaborators, Takashi Taniguchi and Kenji Watanabe at the NIMS institute in Japan, to encapsulate graphene and two-dimensional semiconductor transition metal dichalcogenides (TMDs) such as tungsten diselenide (WSe$_2$), tungsten disulfide (WS$_2$), and molybdenum disulfide (MoS$_2$).

My first major success in experimental device physics calls back to some of the earliest
work in semiconductors, the metal-semiconductor junction. In Chapter 8, I will discuss measurements of an $h$-BN encapsulated graphene-WSe$_2$ Schottky junction which exhibited a complete absence of Fermi-level pinning. The very clean interface between the metal (graphene) and semiconductor allowed us to rigidly control the height of the Schottky barrier in the device over a large range of gate voltages. Furthermore, the unique Schottky barrier tunability the device displayed enabled us to also digitally tune the ideality factor of the diode, beating the previously defined limit of $n = 1$. 


Part I

Quantum transport theory in carbon nanotubes and graphene
Chapter 2

Introduction

Chapters 2-6 will cover four different studies of theoretical quantum transport simulations on carbon nanotubes (CNTs) [1] and graphene [2][3][4]. This chapter will introduce the theoretical concepts which underpin each of these chapters. We will begin by introducing the basic physics of graphene, looking at its structure and electronic properties. From there we will dig deeper and develop tight-binding models of the band structure for each system, which will be used heavily throughout this thesis. At this point, it will become necessary to introduce the fundamental concepts of quantum transport within the Landauer picture. From there, we will present a variety of relevant topics, such as transport in graphene $p$-$n$ junctions and the physics of graphene in a magnetic field. Finally, we will discuss the tight-binding transport method and practical aspects of tight-binding calculations.

2.1 Graphene- basic electronic properties

The discovery of graphene in 2004 changed the face of condensed matter physics and continues to shape the field today [12]. Graphene is a single layer of carbon atoms arranged in a honeycomb structure which result in a rich playground for studying nanoscale physics [14]. These exotic properties stem from the fact that electrons in graphene behave as massless Dirac fermions [13].

The crystal structure of graphene (Fig. 2-1) is described by a 2D hexagonal lattice with
Figure 2-1: Graphene crystal structure (A) Real-space crystal structure of graphene. The carbon atoms are arranged in a honeycomb structure with $sp^2$ bonding. The lattice vectors, eq. (2.1), and nearest neighbor vectors, eq. (2.5), are denoted. (B) Reciprocal space structure of graphene. The first Brillouin zone is outlined with a hexagon and the high-symmetry points are marked as well as the reciprocal lattice vectors (eq. (2.2)). This figure is reproduced from Neto et al [14] (copyright American Physical Society).

a two atom basis set. The lattice vectors for the triangular lattice are

\[
\mathbf{a}_1 = \frac{3}{2} a\mathbf{x} + \frac{\sqrt{3}}{2} a\mathbf{y} \\
\mathbf{a}_2 = \frac{3}{2} a\mathbf{x} - \frac{\sqrt{3}}{2} a\mathbf{y}
\]  \hspace{1cm} (2.1)

where $a \approx 1.44$ Å is the the carbon-carbon bond distance. In reciprocal space, the graphene lattice vectors are

\[
\mathbf{b}_1 = \frac{2\pi}{3a} \mathbf{x} + \frac{2\sqrt{3}\pi}{2a} \mathbf{y} \\
\mathbf{b}_2 = \frac{2\pi}{3a} \mathbf{x} - \frac{2\sqrt{3}\pi}{2a} \mathbf{y}
\]  \hspace{1cm} (2.2)

In this dissertation, we are primarily concerned with the electronic properties of graphene, therefore we will begin by studying its electronic structure. We begin with a nearest neighbor tight-binding model

\[
\hat{H} = -t \sum_{i,j}^N \hat{a}_i^\dagger \hat{b}_j - t \sum_{i,j}^N \hat{b}_i^\dagger \hat{a}_j, 
\]  \hspace{1cm} (2.3)

where $a_i^\dagger/a_i$ and $b_i^\dagger/b_i$ are creation/annihilation operators for electrons on lattice site $i$ for the
A and B sublattices, respectively. The quantity $t$ is the hopping parameter, which describes the overlap of the $\pi$-bonds of the graphene lattice and is set to -2.71 eV. In this case, the summations only occur over first nearest neighbor atoms.

If we consider that each atom in the graphene lattice has three nearest neighbors, computing the summations in eq. (2.3) yields

$$\hat{H} = -t \sum_i a_{r_i}^\dagger b_{r_i+\delta_1} - t \sum_i a_{r_i}^\dagger b_{r_i+\delta_2} - t \sum_i a_{r_i}^\dagger b_{r_i+\delta_3} + \text{h.c.},$$

(2.4)

where the first three terms represent the connection of the A sublattice to the neighboring B sublattice sites, and the Hermitian conjugate term represents the connection of the B sublattice site to the neighboring A sublattices. $r_i$ is the position of a lattice site $i$ and $\delta_{1,2,3}$ are the nearest neighbor vectors (shown in Fig. 2-1) defined as

$$\delta_1 = a/2(1, \sqrt{3}), \delta_2 = a/2(1, -\sqrt{3}), \delta_3 = -a(1, 0).$$

(2.5)

To begin calculating the energy eigenvalues of this Hamiltonian, eq. (2.3), we need to move from real space to reciprocal space \[15\]. To do this, we define the inverse Fourier transform of the creation/annihilation operators:

$$a_i = \frac{1}{\sqrt{N}} \sum_k a_k e^{-ik \cdot r}, b_i = \frac{1}{\sqrt{N}} \sum_k b_k e^{-ik \cdot r}$$

(2.6)

where $k = k_x + k_y$ is the crystal momentum and $N$ is the number of real-space lattice sites. By completing this transformation for all the terms in eq. (2.4) and making some simplifications, we arrive at

$$H = -t \sum_k a_k^\dagger b_k \left[ e^{-ik \cdot \delta_1} + e^{-ik \cdot \delta_2} + e^{-ik \cdot \delta_3} \right]$$

$$- t \sum_k b_k^\dagger a_k \left[ e^{ik \cdot \delta_1} + e^{ik \cdot \delta_2} + e^{ik \cdot \delta_3} \right]$$

(2.7)
which may be written in a more compact matrix form as

\[ H = \sum_k \begin{pmatrix} a_k^\dagger & b_k^\dagger \end{pmatrix} \begin{pmatrix} 0 & h_0(k) \\ h_0(k)^* & 0 \end{pmatrix} \begin{pmatrix} a_k \\ b_k \end{pmatrix} \] (2.8)

where \( h_0 = -t \left[ e^{-ik\delta_1} + e^{-ik\delta_2} + e^{-ik\delta_3} \right] \).

By solving for the eigenenergies of eq. (2.8), we arrive at the energy dispersion relationship of the first nearest neighbor tight-binding model for graphene:

\[ \epsilon_k = \pm t \sqrt{3 + 2 \cos \left( \frac{k_y a\sqrt{3}}{2} \right) + 4 \cos \left( \frac{3}{2} k_x a \right) \cos \left( \frac{\sqrt{3}}{2} k_y a \right)} \] (2.9)

Inspection of eq. (2.9), plotted in Fig. 2-2, shows that the valence and conduction bands in graphene meet at the \( K \) and \( K' \) points of the Brillouin zone, forming a what we refer to as Dirac points. Interestingly, in the vicinity of the Dirac points, graphene’s bands are linear in momentum. Indeed, the electrons in graphene near the \( K \) and \( K' \) points are described by a massless Dirac Hamiltonian \[ H_{K'} = \hbar \nu_F \sigma \cdot k \] (2.10)

(or \( H_{K'} = \hbar \nu_F \sigma^* \cdot k \) for the \( K' \) point) where \( \sigma = (\sigma_x, \sigma_y) \), \( k = (k_x, k_y) \), and \( \nu_F \approx 10^6 \text{m/s} \) is the Fermi velocity.

The eigenenergies of eq. (2.10) are given by \( E = \pm \hbar \nu_F k \) where \( k \) is the momentum relative to the \( K \) (or \( K' \)) point. Interestingly, the electrons in graphene behave similarly to photons, except with a reduced velocity compared to the speed of light. This is in stark contrast to electrons in conventional semiconductors which are described by a massive Hamiltonian, \( H = -\frac{\hbar^2}{2m^*} \nabla^2 \), with \( E = \frac{\hbar^2 k^2}{2m^*} \) where \( m^* \) is the effective mass of the electron.

The massless Dirac behavior of electrons in graphene has significant ramifications for its electronic properties. For example, we can calculate the density of states for electrons in graphene near the Dirac point. Starting from the definition for two-dimensional density of
Figure 2-2: Tight-binding band structure of graphene

Energy dispersion relationship of graphene calculated using a first nearest neighbor tight-binding model, plotted from eq. (2.9). The left panel shows the band structure around the first Brillouin zone. The right panel zooms into the band structure around the K point where the band structure is linear.

states,

\[ D_{2D}(E) = \frac{g_\nu g_s}{A} \sum_K \delta(E - E(k)), \]  

(2.11)

noting that for \( E > 0 \), \( \sum_K \delta(E - E(k)) = \frac{A}{(2\pi)^2} \int_0^\infty \delta(E - E(k))2\pi kdk \) and inserting \( E = \hbar \nu_F k \) we arrive at \( D(E) = \frac{g_\nu E}{\pi \hbar^2 \nu_F^2} \). Here \( g_\nu = 2 \) is the valley degeneracy and \( g_s = 2 \) is the spin degeneracy, and by noting the symmetry of the dispersion relationship about the Dirac point, we can re-write the graphene density of states as

\[ D_{\text{graphene}}(E) = 2 \frac{|E|}{\pi \hbar^2 \nu_F^2} \]  

(2.12)

In contrast to conventional three-dimensional semiconductors, which have a square root dependence for the density of states and an energy band gap, \( E_G \), graphene’s density of states is linear with respect to momentum and gap-less (the low energy band structure of graphene is compared with semiconductors in Fig. 2-3). Indeed, graphene is typically referred to as a Dirac semimetal.
Graphene is described by a massless Dirac Hamiltonian, which leads to an energy dispersion relationship which is linear near the $K$ and $K'$ points (left panel). This is in contrast to semiconducting materials which are generally described by massive Hamiltonians, leading to parabolic-like bands characterized by an energy band gap, $E_G$ (a direct band gap semiconductor is shown here).
2.1.1 Tight-binding model for simulations

While the tight-binding Hamiltonian in eq. (2.3) is useful for analytical calculations in graphene, such as determining the energy dispersion relationship, there is a more practical version which we will use in the simulations throughout this dissertation,

\[ \hat{H} = \sum_{i}^{N} \epsilon_{i} \hat{c}_{i}^{\dagger} \hat{c}_{i} + \sum_{i,j}^{N} t_{i,j} \hat{c}_{i}^{\dagger} \hat{c}_{j}, \]  

(2.13)

where \( \hat{c}_{i}^{\dagger} / \hat{c}_{j} \) are Fermionic creation/annihilation operators, \( \epsilon_{i} \) is the on-site energy at site \( i \), and \( t_{i,j} \) is the hopping energy between sites \( i \) and \( j \). The creation/annihilation operators are defined such that \( \hat{c}_{i}^{\dagger} \hat{c}_{i} = 1 \), \( \hat{c}_{j}^{\dagger} \hat{c}_{i} = 1 \) if sites \( i \) and \( j \) are first nearest neighbors, and \( \hat{c}_{i}^{\dagger} \hat{c}_{j} = 0 \) otherwise. In the absence of any magnetic field, we write the hopping energy as \( t_{i,j} = t_{0} \approx 2.7 \text{ eV} \), to fit to the energy bands given by the \( \pi \)-bond overlap of the \( sp^2 \)-bonded carbon atoms. This Hamiltonian may also be used to describe CNTs by including a periodic boundary condition around the circumference of the nanotube.

There are a few caveats which are important for proper use of this Hamiltonian. First, we restrict the Hamiltonian to just the first nearest neighbors around each carbon atom, neglecting any higher order terms. This is done mainly for reducing the size of the computation we perform. Additional terms only slightly effect the electronic properties, as well. Second nearest neighbor hoppings, for example, add a slight asymmetry to the band structure around the Dirac point, however this effect is minimal, especially for energies close to the Dirac point.

Our tight-binding Hamiltonian also neglects the spin degree of freedom. While it is mathematically trivial to include a spin term to the Hamiltonian, it doubles the size of the matrix which we will need to solve. Also, as we will see, none of the experiments which we work to model displayed spin-dependent phenomena, thus we decide to neglect it in this discussion. In the same vein, our Hamiltonian also neglects any electron-electron interaction or exchange-correlation effects, which would dramatically increase the computational complexity of the problems we consider.

Some version of this tight-binding Hamiltonian will be used in each theory-based chapter
of this dissertation. The on-site energy term will be used throughout each chapter to define
the energy landscape of each simulation, such as in including p-n junctions and contact-
induced doping effects. Furthermore, the hopping energy will be modified with Peierl’s
phase (to be described in more detail, later on) to include the effect of externally applied
magnetic fields. We will see how this tight-binding Hamiltonian is a general platform from
which to build more complex simulations which show good agreement with experimental
data.

The linear dispersion relationship of graphene results in many special electrical properties
which we will study throughout this dissertation. In the following sections of this chapter
we will introduce some of these special properties. First, it is now necessary to lay the
groundwork of the Landauer transport formalism which will be used heavily throughout
the the first part of this dissertation. We will then investigate transport across potential
barriers in graphene, which exhibit the phenomena of Klein tunneling and angle-dependent
transport. From there we will discuss quantum Hall effect in graphene.

2.2 Quantum transport basics

Before continuing our discussion of graphene, now is a good time to lay the groundwork
for how we will interpret transport phenomena throughout this thesis. We consider a two-
terminal system of two-dimensional electrons, consisting of a channel connected to two reser-
voirs of electrons. Traditionally, the conductance of such a channel is governed by Ohm’s
law [17]

\[ G = \frac{W}{L}\sigma, \]  

(2.14)

where \( W \) is the width of the channel, \( L \) is the length of the channel, and \( \sigma = e^2\rho(E_F)D \)
is the Drude conductivity of the channel with \( e \) the fundamental electron charge, \( \rho(E_F) \) the
carrier density at the Fermi energy, and \( D \) the electron diffusivity.

Very generally, the diffusivity of electrons is described by some mean scattering length,
\( l \). This scattering could be due to many mechanisms, such as inelastic scattering, elastic
scattering, or even dephasing. This diffusive picture of transport assumes that the dominant
scattering length is much smaller than the dimensions of the channel, however in modern
nanoelectronics this is often not the case.

Indeed, in high quality samples, the mean scattering length may be much larger than the sample dimensions, such that there is little to no scattering for an electron transporting between the two reservoirs of carriers. In this case, we call the transport “ballistic”. There is also an intermediate “quasi-ballistic” case where there are just a few scattering events inside the channel. These cases were summarized in an excellent figure in a review by Beenakker et al, which we reproduce in Fig. 2-4. For some time it was debated what would happen if the mean scattering length was shorter than the channel length. Would the resistance become zero?

Rolf Landauer proposed a picture of ballistic transport which related the conductance to transmission. In Landauer’s picture, the conductivity is no longer a useful quantity for describing the conductance of the device channel. In this case, we abandon the diffusive transport formalism and instead adopt a picture of transport which relates the conductance of the channel to quantum mechanical transmission, $T$:

$$G = \frac{e^2}{h} T(E_F) \tag{2.15}$$

In this Landauer picture, conductance is purely determined by the transmission of electrons at the Fermi energy. In this case, the channel is analogous to a waveguide which consists of a number of propagating modes, $M(E)$, each of which transmit with some probability, $t(E)$, and we define $T(E) = M(E)t(E)$. In the Landauer picture, a transmission probability of unity (ballistic transport) tells us that the maximum conductance of a ballistic conductor is $G = \frac{e^2}{h} M(E)$.

Indeed, in a ballistic conductor the resistance does not drop to zero. van Wees et al provided strong experimental evidence for the Landauer picture of transport in their measurement of the first QPC, where by electrostatically constricting the dimensions of their ballistic device channel they modulated $M(E)$, showing discrete steps in the conductance. This picture of conductance as discrete modes of transmitting electrons is why we refer to the study of ballistic electron transport as “quantum transport”.

Not only does the Landauer picture of transport give us a way to interpret transport
Figure 2-4: Electronic transport regimes. This figure depicts the three major regimes of electronic transport. When the mean scattering length is much smaller than the channel dimensions, the transport type is (A) diffusive and may be described within a Drude model. As the mean scattering length is increased (or the sample dimensions decreased) the amount of scattering within the channel begins to decrease, eventually reaching the (B) quasi-ballistic regime. When there is little to no scattering within the channel the sample is in the (C) ballistic transport regime and can be described well by the Landauer picture of transport. This figure was reproduced from a review by Beenakker et al. [17], copyright Academic Press Inc., 1991.
in ballistic electronics, it also informs how one would model them. Indeed, Landauer’s picture reduces the problem to simply calculating the transmission of a system at the Fermi energy. Throughout most of this dissertation our primary task will revolve around calculating transmission in different systems.

Eq. (2.15) is derived under the assumption that the electrical device is at zero temperature, however in most cases this not realistic. As such, it becomes necessary to modify Landauer’s equation to include a finite temperature, which is written as

\[
G = \frac{2e^2}{h} \int_{-\infty}^{+\infty} T(E) \left( -\frac{d f_0(E)}{d E} \right) dE, \tag{2.16}
\]

where \( f_0(E) = \frac{1}{e^{E-E_F/k_B T}+1} \) is the Fermi-Dirac distribution function which describes the occupation of states near the Fermi energy and \( k_B \) is Boltzmann’s constant. Additionally, a factor of 2 has been included representing the spin degeneracy, however one must be careful to know when it is required.

The Fermi-Dirac distribution, or Fermi function, is particularly important in quantum transport. The derivative of the Fermi function with respect to energy is referred to as the thermal broadening function \[18\], \( F_T(E) = -\frac{d f_0(E)}{d E} = \frac{1}{4k_B T} \sec h \left( \frac{E}{2k_B T} \right) \). The conductance in eq. (2.16) is given by a weighted average of the transmission of different states around the Fermi energy within an envelope described by \( F_T(E) \), which we call the “Fermi window”. In the limit of low temperature, which we define as essentially liquid helium temperature and below \((T < 4.2 \text{ K})\) eq. (2.16) reduces to eq. (2.15), where the conductance is due to only the electrons exactly at the Fermi energy.

It is also important to understand current in a ballistic conductor. Generally in the lab one does not measure conductance directly, instead one applies a voltage bias and measures a current. In the Landauer picture, for finite temperature, the current is given by \[18\]

\[
I = \frac{2e}{h} \int_{-\infty}^{+\infty} T(E) \left( f_1(E, \mu_1) - f_2(E, \mu_2) \right) dE \tag{2.17}
\]

where here \( f_{1,2}(E, \mu_{1,2}) \) are the Fermi-Dirac distributions of contacts labeled 1 and 2, each with chemical potentials (quasi-Fermi energies) \( \mu_{1,2} \). The \((f_1(E, \mu_1) - f_2(E, \mu_2))\) term in eq. (2.17) is analogous to the thermal broadening function in eq. (2.16), except this time there
is an additional broadening which is due to the bias applied, expanding the Fermi window by approximately $\mu_1 - \mu_2$. An example of this effect is shown in Fig. 2-5 for $(\mu_1 - \mu_2) = 100$ meV and $T = 300$ K. We will use eq. (2.17) heavily throughout chapters 3 and 4 when calculating current flow through CNTs and graphene at finite temperatures.

2.3 Graphene $p$-$n$ junctions

Some of the richest physics of graphene become evident in the study of graphene-based $p$-$n$ junctions. Whilst graphene does not have a band gap, it is still possible to form a $p$-$n$ junction in graphene. Importantly, a graphene $p$-$n$ junction is not a rectifying diode. The band structure of graphene leads to a new form of $p$-$n$ junction with unique properties compared to semiconducting $p$-$n$ junctions. Two chapters of this dissertation will be dedicated to studying quantum transport in graphene $p$-$n$ junctions, one on high magnetic field transport in the quantum Hall regime (chapter 5) and the other on low magnetic field transverse electron focusing (chapter 6). Additionally, at the end of chapter 7, we will see experimental measurements on a high quality graphene $p$-$n$ junction sample which exhibits quasi-ballistic transport at room temperature.

Typically, a metal gate is capacitively coupled to graphene and by biasing the gate, one can adjust the graphene Fermi energy. When the Fermi energy is below the Dirac point, the graphene is $p$-type, and when it is above the Dirac point, the graphene is $n$-type. It is possible to locally gate two sides of the a graphene device through a number of schemes, such as buried split-gates [19], a top gate/bottom gate scheme [20], or local bottom gates [21].

2.3.1 Angle-dependent transport and Klein tunneling

Many of the interesting properties of graphene result from interesting angle-dependent transport phenomena, which we will describe in this section.

Early in the study of graphene, there were a flurry of papers predicting angle-dependent transport phenomena. Katsnelson et al predicted that electrons incident normal to a potential barrier in graphene (a $p$-$n$-$p$ or $n$-$p$-$n$ junction) would perfectly transmit to the other
Figure 2-5: Depiction of the Fermi window for a two-terminal system with a small bias at 300 K. The two solid lines are the Fermi-Dirac distributions for two reservoirs, one at $\mu_1 = 0$ eV and one at $\mu_2 = -0.1$ eV. The difference in the chemical potentials puts the contacts out of equilibrium, causing current to flow. The yellow highlighted region depicts the difference between the two Fermi functions, as appears in eq. (2.17). In eq. (2.17), the Fermi window is convoluted with the transmission function, where the thermal broadening and bias effects "smear" the observed characteristics, providing a weighted average of the transmission functions at different energies.
Figure 2-6: Graphene p-n junction schematic This figure shows the real space energetic landscape (top) and top-down schematic (bottom) for an abrupt graphene p-n junction (left) and a graded graphene p-n junction (right). Dirac cones are superimposed on the band diagrams, to show the Fermi energy relative to the Dirac point for the different configurations. The lower left panel defines the incident, reflected, and transmitted angles referred to in the main text. For the graded junction, a finite transition region between the n- and p-type regions is shown, with a width labeled $D_W$. The top-down schematics each represent a semi-infinite graphene channel oriented parallel to the x-axis, with a finite width $W$ along the y-axis.
side, exhibiting a phenomena known as Klein tunneling [22]. Cheianov and Fal’ko separately derived that a p-n junction would act as an angle-dependent low pass filter for electrons, where normally incident electrons would Klein tunnel whilst high-angled electrons are reflected [23]. Cheianov et al further went on to predict that the angle-dependent transport at a graphene p-n junction could be exploited to create a Veselago lens for electrons, a flat lens which necessitates a negative index of refraction material [24].

We will begin by considering a graphene device, which is shown in Fig. 2-6, which is of some width $W$ and is oriented along the $x$-direction (the transport direction). For $x < 0$ the Fermi energy is moved into the conduction band and for $x > 0$ the Fermi energy is in the valence band, forming an abrupt $n - p$ junction (we will consider the smooth junction later). For each side the Fermi energy is close to the Dirac point such that the electrons (or holes) are described by eq. (2.10). We will assume that an electron is incident on the junction from the left side with an angle $\theta_i$ (measured from perpendicular to the junction interface). It will either transmit at an angle $\theta_t$ or be reflected at an angle $-\theta_i$.

The incident electron has a Fermi wave vector, $k_{F_i}$ with $k_x = k_{F_i} \cos \theta_i$ and $k_y = k_{F_i} \sin \theta_i$. The transverse momentum ($k_y$) is conserved across the junction interface at $x = 0$, which constrains the momentum of transmitted electrons such that

$$k_{F_i} \sin \theta_i = k_{F_t} \sin \theta_t$$  \hspace{1cm} (2.18)

with $k_{F_t}$, the Fermi wave vector of the transmitted hole. Eq. (2.18) is analogous to Snell’s law from geometrical optics, with the indices of refraction of two mediums replaced by the Fermi wave vectors on each side of the junction. In order for the transmitted holes to continue propagating to the right, the sign of the $x$-component of the wave vector must flip, resulting in $\theta_t = -\theta_i$. Therefore, the transporting carriers are refracted similar to how you would expect light to be refracted in a material with negative refractive index [24].
Figure 2-7: Angle-dependent transmission function of a symmetric graphene $p-n$ junction. This figure depicts the transmission functions, as a function of incident angle, for an electron transporting across a graphene $p-n$ junction. The blue line shows the transmission due to an abrupt junction, described by eq. (2.19) in the text. The green line is for a graded junction (described by eq. (2.20)) with $k_F D_W = 15$, which corresponds to a Fermi energy of approximately $\pm 0.1$ eV with a junction width of 100 nm, similar to values for experimental samples in chapter 7.
2.3.2 Angle-dependent transport and Klein tunneling

If we consider an abrupt symmetrical (each side has equal, but opposite doping) junction, Cheianov and Fal’ko derived that the transmission probability is given by

\[ T = \cos^2 \theta_i. \]  

(2.19)

For electrons which are normally incident to the junction, Eq. (2.19) predicts perfect transmission, an example of Klein tunneling in graphene. Transmission is reduced for higher angled electrons. For a graded junction with a smooth transition between the n and p regions of length \( D_W \), by invoking the Wentzel-Kramers-Brillouin (WKB) approximation and assuming \( k_F D_W >> 1 \), the angle dependent transmission function becomes

\[ T = e^{-\pi k_F D_W \sin^2 \theta_i}. \]  

(2.20)

Each of these solutions are shown in Fig. 2-7, comparing the abrupt transmission probability to a graded junction. In this case, the graded junction is defined as \( k_F D_W = 15 \), which corresponds to \( E_F \approx 0.1 \) eV and \( D_W = 100 \) nm, both of which are routinely achievable using buried gates. In the case of a graded junction, there is significantly enhanced filtering of higher angled modes.

2.3.3 Conductance of a graphene p-n junction

By restricting the transverse (in this case, the y-direction) direction of the graphene sheet to be of width \( W \), the transverse momentum, \( k_y \), becomes quantized such that there are \( M = 2Wk_F/\pi \) transverse modes (the 2 is for spin) with a spacing between each mode, \( \Delta k_y = \pi/W \). This is the minimum number of modes with spacing \( \Delta k_y \) which can fit inside the Fermi surface assuming a radius of \( k_F \). This effect is depicted in Fig. 2-8.

Quantization of the transverse momentum also tells us that the angle of incidence of electrons transporting towards a p-n junction have discrete values determined by the Fermi energy. In order to calculate the conductance of a p-n junction, we can use Landauer’s equation and sum the transmissions of each mode, \( G = G_0 \sum_M T(\theta) \), with \( G_0 = 4e^2/h \).
Figure 2-8: Transverse modes in graphene The left panel shows a top-down schematic of a uniform n-type sheet of graphene with a width $W$, similar to previous figures. The top right panel depicts the low-energy dispersion relationship of the graphene sheet and the lower right panel shows a top-down view of the Fermi surface at Fermi energy $E_F$. The finite width of the graphene channel confines the electrons into different transverse modes, depicted by the slices through the Fermi surface. As discussed within the main text, these modes are characterized by a Fermi wave vector with magnitude $k_F$ and an angle $\theta$ measured relative to the $k_x$-axis. The modes are spaced in $k_y$ by $\Delta k_y = \pi/W$ and have an angular spread which is given by $\Delta \theta = \frac{\Delta k_y}{k_F \cos \theta}$. For large values of $W$, the distribution of modes becomes continuous, allowing for the conversion between summation and integration used during the derivation of the junction conductance in the main text.
If we assume there are a large number of modes, this summation may be re-written as an integral

$$G = G_0 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} T(\theta) \Delta \theta d\theta$$ \hspace{1cm} (2.21)

where, $\Delta \theta = \Delta k_y \frac{k_F}{\cos \theta}$ is the angle between each transverse mode. By performing the integration in eq. (2.21), assuming an abrupt symmetric $p-n$ junction, we arrive at

$$G = G_0 \frac{2}{3} M(E).$$ \hspace{1cm} (2.22)

By noting that the conductance for uniformly doped ballistic graphene conductor is $G = G_0 M(E)$, eq. (2.22) tells us that the conductance of a symmetric junction will be $2/3$ that of a uniformly doped graphene device, which is purely a consequence of the angle-dependent transport. Similarly, if we repeat the integration in eq. (2.21) for the graded junction with transmission from eq. (2.20),

$$G = G_0 M \text{erf} \left( \frac{\sqrt{\pi} k_F D_W/2}{2 \sqrt{k_F D_W/2}} \right).$$ \hspace{1cm} (2.23)

These results are compared as a function of $k_F$ in Fig. 2-9 where we normalize each conductance to the uniform (no $p-n$ junction) ballistic channel $G = G_0 M$. In the result in Fig. 2-9 we assume a junction width of 100 nm, which (due to enhanced filtering of higher angled modes) produces significantly more conductance asymmetry than the abrupt junction.

### 2.4 Graphene in magnetic fields

In both chapters 5 and 6 we will discuss magnetotransport in graphene $p-n$ junctions. Transport measurements in magnetic fields have been a powerful tool for characterizing materials for an extremely long time and in graphene it is not any different. Indeed, quantum Hall measurements in graphene provided the strongest early evidence that graphene supported massless Dirac fermions- that graphene truly was a two-dimensional material \[13\] \[25\].

To account for applied magnetic fields, the vector potential $\mathbf{A}$, a gauge potential, is included in our continuum Hamiltonian by modifying the momentum operator such that
Figure 2-9: Ballistic conductance asymmetry in graphene p-n junctions: Comparison of the conductance of a uniformly doped ballistic graphene with the conductance predicted for an abrupt symmetric junction (eq. (2.22)) and a graded junction (eq. 2.23) of $D_W = 100$ nm.
The magnetic field is defined as $\mathbf{B} = \nabla \times \mathbf{A}$. Typically, in experimental measurements, the magnetic field is oriented normal to the surface of the sample. For a graphene sheet oriented parallel to the $xy$ plane, a magnetic field of magnitude $B$ in the $z$-direction may be included by adopting the Landau gauge for the vector potential, $\mathbf{A} = (0, -Bx, 0)$.

Practically, the vector potential is included in our tight-binding Hamiltonian by modifying the hopping parameter with Peierl’s phase, $t_{i,j} = t \exp \left[ i \frac{e}{\hbar} \int_{\mathbf{r}_i}^{\mathbf{r}_j} \mathbf{A} \cdot d\mathbf{r} \right]$, where $\mathbf{r}_{i,j} = (x_{i,j}, y_{i,j})$ are the positions of the two neighboring carbon atoms. We can compute the line integral above analytically, which yields,

$$t_{i,j} = t \exp \left[ -i \frac{e}{\hbar} B (x_i - x_j) \frac{y_i + y_j}{2} \right] \quad (2.24)$$

which is the quantity used in our numerical calculations. We note that in real simulations, the Landau gauge as we have defined it is used for leads which are periodic along the $y$-direction. For leads oriented along the $x$-direction, the Landau gauge should be changed to $\mathbf{A} = (-By, 0, 0)$. The idea here is that the Peierl’s phase should only change along the transverse axis of the lead, preserving the periodicity of the semi-infinite leads.

Classically, the inclusion of a magnetic field will induce cyclotron motion in graphene, characterized by a cyclotron radius $r_c = \frac{\hbar \sqrt{\pi n}}{e|B|}$, where $n$ is the charge carrier density. In a strong enough magnetic field (hundreds of mT and up) the electrons will become confined into Landau levels from which one may observe the half-integer quantum Hall effect \[13\] \[25\]. In this dissertation we will look at each of these regimes in detail, but in this section we will primarily focus on high magnetic field effects.

Under the effect of a magnetic field, neglecting Zeeman splitting, the energy levels of graphene become quantized with energies

$$E_{LLn} = \text{sgn}(n) \sqrt{|n| 2eB \hbar v_F^2}, \quad (2.25)$$

where here $n$ is now the Landau level index taking on values $0, \pm 1, \pm 2, \ldots$ where the plus signs indicate electron-like Landau levels and minus signs are hole-like. Each Landau level is spin and valley degenerate, except for the 0th Landau level which is a special case relating
Figure 2-10: Comparison of zigzag graphene nanoribbon tight-binding band structure in zero magnetic field and a large magnetic field: The left panel shows the tight-binding band structure for a wide zigzag-edge graphene nanoribbon oriented along the \(x\)-direction. The \(\mathbf{K}\) and \(\mathbf{K}'\) points are shown, depicting Dirac-like bands described in the main text. The right panel shows the same nanoribbon in a large magnetic field, which deforms the bands, putting the sample into the quantum Hall regime. A more detailed depiction of graphene’s quantum Hall transport is included in Fig. 2-12.

To graphene’s gapless band structure. The lowest Landau level is shared by an electron-like and hole-like Landau level at the Dirac point energy. This phenomenon gives rise to the half-integer quantum Hall effect in graphene [13][25].

The half-integer quantum Hall effect was first observed simultaneously in graphene by two separate groups; Novoselov et al [13] and Zhang et al [25]. We show the result by Novoselov et al, reproduced in Fig. 2-11 [13] (copyright Springer Nature, 2005), although the results of Zhang et al are effectively identical. It is important to note that while we refer to this effect as the “half-integer” quantum Hall effect, it should not be confused with the fractional quantum Hall effect, which deals with Landau levels of fractional charge. Instead, the shared nature of the lowest Landau level between electrons and holes simply offsets the conductance of standard quantum Hall Landau levels.

**Band structure and transport in strong magnetic fields**

Fig. 2-10 shows tight-binding band structure calculations for a zigzag-edged graphene ribbon in zero magnetic field (left panel) and a strong magnetic field (right panel). In each figure
Figure 2-11: Half-integer quantum Hall effect in graphene: Experimental measurement of the half-integer quantum Hall effect in graphene by Novoselov et al [13]. The inset shows measurements on bilayer graphene, which shows standard quantum Hall plateaus for massive Fermions. This figure is reproduced from ref. [13], copyright Nature, 2005.
we see sub-bands in the band structure which correspond to the different transverse modes which arise from truncating the ribbon along the $y$-direction. In the zero magnetic field case, the sub-band spacing is uniform. In a strong magnetic field, however, the sub-band spacing switches to follow eq. (2.25) and takes on a square root dependence in energy.

In addition to the sub-band spacing changing, the shape of each sub-band is deformed in response to the magnetic field. Each sub-band, each of which effectively “is” the Landau level, is deformed from V-shaped cone to a U-shape with a flat bottom. We refer to the flat part of each sub-band as a bulk state and the sloped portions of the bands are edge states (labeled in Fig. 2-12a). This naming becomes intuitive if we look at the local current density within the graphene ribbon for different Fermi energies.

In Fig. 2-12b and c we show maps of the local current density (whose calculation will be explained later) for the Fermi energy tuned to a bulk state and an edge state. When the Fermi energy is tuned to the bulk state energy, the entire graphene sheet conducts, with higher levels of current density denoted by the dark blue color shading. When the Fermi energy is on an edge state, conduction occurs along the edges of the sample and the bulk becomes insulating. In each of these cases, we only show carriers which are injected from the left side of the simulation, which especially highlights the complete filling of the ribbon in the bulk state configuration.

These results can be used to intuitively understand the longitudinal resistivity, $\rho_{xx}$ shown in Fig. 2-11. In their measurement, Novoselov et al are able to sample the band structure of graphene by tuning the Fermi energy of the sample with a global gate. When they place the Fermi energy between two bulk levels, transport at the Fermi energy only occurs along the edges of the sample. As a result, neighboring Hall probes will float to the same voltages (understanding multiterminal measurements will be explained in detail in chapter 6) and $\rho_{xx}$ drops to zero.

However, when the Fermi energy is tuned to a bulk level, carriers transport everywhere in the sample and result in peaks in $\rho_{xx}$. While the total description is quite complex, one of the primary contributors of the width of these peaks in $\rho_{xx}$ is the level of disorder in the sample, which serves to broaden the energy of the bulk level. In early graphene samples, this effect was quite strong. In our simulations, the bulk level is essentially a delta function in
energy. Thus, to be able to see the effect of bulk conduction we added a very small amount of disorder ($\pm 0.001$ eV, randomly distributed) to the bulk of our simulation cell.

In chapter 5, we will see how the inclusion of a $p-n$ junction in the quantum Hall system lets us localize these bulk states at the junction interface. Furthermore, we will see how they are actually spatially delocalized in a graded $p-n$ junction.

### 2.5 Tight-binding transport calculations

Except for the simulations in chapter 3 (based on non-equilibrium Green’s functions), all of the quantum transport calculations in this dissertation are the result of the tight-binding transport formalism, implemented in the python package Kwant [27]. In this section we will describe the tight-binding transport formalism, essentially repeating the details laid out in
Figure 2-13: Generalized scattering problem for a sample with several leads: This figure depicts the typical scattering problem which we will be solving using Kwant throughout this dissertation. The simulation consists of a central, finite, scattering region which is contacted by (in this case, three) semi-infinite leads.

One of the greatest contributions from the Kwant package and its developers is that in practice, most of these details may be ignored in favor of focusing on the physics of the individual problems of interest. However, understanding the construction of the tight-binding transport equations can be quite useful; for example, in chapter 6 we will decompose the different lead-modes to assist in understanding the transverse electron focusing problem.

In tight-binding transport problems, generally, every problem may be broken down into a system consisting of a central scattering region which is connected to an arbitrary number of semi-infinite leads (an example is shown in Fig. 2-13). The system will be described by some tight-binding Hamiltonian matrix (not necessarily the graphene Hamiltonian we have discussed), which will have a form

$$H = \begin{bmatrix} H_{\text{leads}} & V \\ V^\dagger & H_{\text{channel}} \end{bmatrix},$$

(2.26)

where $H_{\text{leads}}$ is a semi-infinite matrix describing all of the leads, $H_{\text{channel}}$ is a finite matrix describing the channel, and $V$ is a matrix which couples the leads and channel together. One
assumption which we require is that each lead is made up of a series of periodic unit cells, this lets us expand $H_{\text{leads}}$ such that

$$H_{\text{leads}} = \begin{bmatrix} \cdots & V_L & \vdots & \ldots \\ V_L^\dagger & H_L & V_L & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ V_L^\dagger & H_L & V_L & \cdots \end{bmatrix} \tag{2.27}$$

where $H_L$ and $V_L$ are square blocks which describe a unit cell of each lead and the coupling between unit cells, respectively. In general, $H_L$ and $V_L$ are the Hamiltonians of all of the leads in the system lumped together, for simplicity.

The wave function of this system will have the form $(\ldots, \psi^L(2), \psi^L(1), \psi^S)$, where $\psi^L(i)$ is the wave function of the $i^{th}$ unit cell of the leads and $\psi^S$ is the wave function of the central scattering region. From the periodicity of the leads, the wave function in the leads is a superposition of planewaves, therefore we can write down Schrödinger’s equation with a Bloch Hamiltonian as

$$(H_L + V_L e^{-ik_n} + V_L^\dagger e^{ik_n}) \chi_n = E \chi_n. \tag{2.28}$$

where the $n^{th}$ eigenstate (or mode) of the translation operator for the $j^{th}$ lead unit cell is given by $\phi_n(j) = (\lambda_n)^j \chi_n$ with $\lambda_n = e^{ik_n}$. Modes for which $|\lambda_n| = 1$ are referred to as propagating modes and modes for which $|\lambda_n| < 1$ are evanescent.

Furthermore, all of the propagating modes can be sorted into ingoing or outgoing modes. Groth et al do this by normalizing the modes to the expectation value of the particle current, $\langle I \rangle \equiv 2 \text{Im} \langle \phi_n(j)|V_L|\phi_n(j-1) \rangle = \pm 1$. When $\langle I \rangle = 1$, the mode is incoming, with eigenfunction $\phi_n^{\text{in}}()$, and when $\langle I \rangle = -1$, the mode is outgoing, with eigenfunction $\phi_n^{\text{out}}()$. Additionally, the evanescent mode eigenfunctions are denoted as $\phi_p^{\text{ev}}$ with index $p$.

From these equations, Groth et al define the scattering state of the $n^{th}$ mode in the $i^{th}$ unit cell of the leads by

$$\psi_n(i) = \phi_n^{\text{in}}(i) + \sum_m S_{mn} \phi_m^{\text{out}}(i) + \sum_p S_{pn} \phi_p^{\text{ev}}(i) \tag{2.29}$$

where the summation of $m$ takes place over the other propagating modes of all the leads.
and $S_{mn}$ is the scattering matrix element which tells us how modes $n$ and $m$ scatter into one another. Finally, the wave function inside the scattering region is given by

$$\psi_n(0) = \phi_n^S.$$ \hspace{1cm} (2.30)

With the scattering matrix, which tells us the connection between the propagating modes in each of the leads, we can learn about the terminal characteristics of the channel and with the wave function inside the channel, $\psi_n(0) = \phi_n^S$, we can learn about spatially resolved properties inside the finite scattering region. With both of these properties, we have the building blocks for describing transport in our problem. The Kwant package’s main computational tasks are solving for these two quantities, which is done via wave function matching between the leads and the scattering region. Kwant also performs the calculations of the propagating lead modes from above, however, to date the exact algorithms of this process have not been published, although some details have been provided in the literature [27][28].

**Observables**

In order to make our tight-binding transport calculations useful, we need to eventually calculate some parameters which may be connected to experimentally observable variables from the lab. In this dissertation, our general strategy will be to first connect to experimental measurements of current or resistance- in other words, we will benchmark against the terminal characteristics (stemming from the scattering matrix). From there, we will look at spatially resolved properties of the system, which we determine using the wave function inside the channel.

In the Landauer picture of transport, as we have already discussed, the central quantity of interest is the transmission function. From the transmission function we can calculate the ballistic conductance and current. We will also see in chapter 6 how the transmission function can be used within the Landauer-Büttiker formula to calculate non-local resistance. To determine the transmission function from lead $p$ to lead $q$, we use the scattering matrix

$$T_{qp}(E) = \sum_{n \in p, m \in q} | S_{nm}(E) |^2.$$ \hspace{1cm} (2.31)
where the summations take place over the $n$ modes in lead $p$ and the $m$ modes in lead $q$ and $S_{nm}(E)$ is the scattering matrix element we saw before, written as a function of energy $E$.

We will also look at spatially resolved properties within the scattering region. For example, the local density of states (LDOS) at site $i$ for a system with $j$ leads is

$$N(r_i, E) = \frac{1}{\pi} \sum_j \psi_j(i, E)\psi_j(i, E)\dagger.$$  \hspace{2em} (2.32)

The summation takes place over all of the leads in the system, however practically we will usually just look at the LDOS which is contributed from one lead at a time, since the properties become more challenging to interpret as additional leads are summed.

Additionally, we can calculate the spatially resolved current density between two sites at $r_i$ and $r_j$, injected by lead $p$, by using

$$J_{r_i, r_j}(E) = -2 \sum_{n \in p} \text{Im} \left[ \psi_n(r_i, E)\dagger \hat{H}_{i,j}\psi_n(r_j, E) \right]$$  \hspace{2em} (2.33)

where the summation takes place over all of the modes in the lead. Similarly to the LDOS, we prefer to plot this quantity for just a single lead at a time. It is also interesting to look at the current which is due to different modes which are injected into the channel, as we do in chapter 6. Also, if we do a vector sum of $J_{r_i, r_j}(E)$ at each site, we produce a vector field which leads to informative steamplot maps of carrier transport.

Other packages used in this dissertation

In addition to Kwant, which is used heavily throughout the dissertation, several other Python packages are worth mentioning. Except for chapter 3, every chapter in this dissertation has relied heavily on open-source Python packages. Nearly all of the numerical analysis, including fitting, was performed using the fantastically documented NumPy [29] and SciPy [30] packages. Most of the figures throughout the dissertation relied on the late John D. Hunter’s Matplotlib library [31]. Additionally, analysis of all of the experimental data included in the second part of this dissertation was performed within the IPython notebook framework [32].
Chapter 3
Quantum transport modeling of carbon nanotubes under proton irradiation

In this chapter we present an atomistic, quantum transport-based approach for studying the effect of radiation-induced ions on transport across carbon nanotube transistors. We will use the non-equilibrium Green’s function technique to study carrier transport in CNTs which are subjected to the Coulomb potential due to ions at various distances. We will investigate both the local density of states and transfer characteristics (both current and transmission functions) to provide a detailed understanding of transport in the CNT. We will begin with a review of radiation effects literature of CNT-based electronics, then introduce our NEGF-based model, and finally present and analyze our simulation results.

3.1 Introduction
The one-dimensional nature of CNTs makes them of great interest for fundamental studies of quantum transport. CNTs are capable of carrying extremely high current density 33[34], which makes them promising for applications in nanoelectronics 35. In particular, semiconducting CNTs have received much attention as a post-silicon channel material in field effect transistors (FETs) 363738394041424344.
In addition to their terrestrial applications, the very small radiation cross-section (due to their extremely small diameters) makes CNTs of interest in the development of radiation-hard electronics [45, 46]. The substrate tends to dominate the radiation response of CNTs [45, 47, 48, 49, 50], similar to silicon-based FETs [51], however dielectric engineering has been shown as one option to combat this problem [52].

One way to study intrinsic properties of CNTs is to suspend them off a substrate. Not only does suspending the CNT result in drastically improved electrical properties, it also produces a significant change in the samples radiation response [53]. Indeed, even at very high fluences of protons, direct damage of the CNT is not observed. Instead, as has been supported by separate molecular dynamics calculations [54], the primary radiation induced-damage is in the substrate and substrate-supported regions of the CNT [53].

Suspended CNTs irradiated with protons and measured in situ in an ambient environment, however, show a striking radiation response [55]. Bushmaker et al observed that during proton irradiation, the transfer characteristics of their CNT FETs would spontaneously degrade and then recover. Originally, this effect was observed to take place on the scale of tens to hundreds of milliseconds [55]. Given this timescale, we reasoned that the characteristic was due to ionized gasseous molecules nearby the nanotube [1]. However, future results revealed that the degradation could last for hundreds of seconds before recovering [56], suggesting ions actually physically adsorb to the CNT to cause this effect.

Despite this updated mechanism, our simulation result, which only considered static ions, are still relevant to understanding the two studies by Bushmaker et al [55, 56]. We will also see how our model can be simply extended to model CNT FETs which are embedded in gate dielectrics, as would be the case in more realistic nanoelectronic applications.

### 3.2 Transport Model

#### 3.2.1 Non-equilibrium Green’s Function method

In this chapter we will use the NEGF method [57] to model transport across CNTs. NEGF has emerged as a useful technique for modeling electrical transport in CNTs [58, 59] and
has been used successfully to understand transport in gate-all-around CNT FETs [60]. The NEGF technique yields essentially the same numerical results as the scattering matrix/tight-binding transport method (used in the other chapters of this dissertation) we discussed, but is less efficient (in terms of both memory and computation time). Despite this, the results in this chapter were all produced using an entirely home-built MATLAB code and were calculated using a standard laptop with about 4 GB of memory.

The central quantity of interest in the NEGF is the time-retarded Green’s function,

\[
G^R = [(E + i\delta) I - H - \Sigma_L - \Sigma_R]^{-1}
\]  

(3.1)

which is used to calculate all of the other quantities of interest. \( E \) is the energy of the simulation, \( I \) is the identity matrix, \( H \) is the channel Hamiltonian, and \( \Sigma_L, \Sigma_R \) are the self-energies accounting for the left and right device contacts. An infinitesimally small imaginary quantity, \( i\delta \) is added to the diagonal of the matrix, shifting the poles of eq. (3.1) to the lower half of the complex plane, which prevents problems when numerically inverting the \((E + i\delta) I - H - \Sigma_L - \Sigma_R\) term.

Direct examination of \( G^R \) does not offer much insight into the physics of a problem, however, it may be used in calculating a number of useful quantities which do. For example, the LDOS of the device channel is given by

\[
N(E) = -\frac{1}{\pi} \text{Im} \left[ G^R \right],
\]  

(3.2)

and transmission function is given by [61]

\[
T(E) = \text{Tr} \left[ G^R \Gamma^L G^A \Gamma^R \right].
\]  

(3.3)

\( G^A \) is referred to as the time advanced Green’s function and is just \( G^A = G^{R\dagger} \) (therefore by calculating \( G^R \), we essentially get \( G^A \) at the same time). The terms \( \Gamma^{L,R} \) are defined as

\[
\Gamma^{L,R} = i \left[ \Sigma^{L,R} - \Sigma^{L,R\dagger} \right],
\]  

(3.4)
the anti-Hermitian components of the contact self energies. These terms represent carrier injection from the leads and result in a broadening of the states inside the channel.

The transmission function above is the same quantity we saw previously (calculated via the scattering matrix) and may be used in the Landauer equation to calculate current through the channel with
\[ I(E_F) = \frac{2e}{h} \int_{-\infty}^{\infty} T(E) [f_L(E, \mu_L) - f_R(E, \mu_R)] dE. \]

In this chapter, the chemical potentials of the source and drain (or left and right) contacts are given by
\[ \mu_{L,R} = E_F \pm \frac{eV_{DS}}{2}, \]

where \( E_F \) would be the Fermi level resulting from the back gate bias and \( V_{DS} \) is the source-drain bias (the plus term above is for the left contact and the minus term is assigned to the right contact).

### 3.2.2 CNT Hamiltonian

The Hamiltonian in eq. (3.1) is general to any system we would want to study. In this case, we will use a nearest-neighbor tight-binding model,
\[ \hat{H} = \sum_{n=m} \epsilon_n \hat{c}_n^\dagger \hat{c}_n + \sum_{n \neq m} t_0 \hat{c}_{n/m}^\dagger \hat{c}_{m}, \]

which is effectively identical to the one we use for graphene, except in this case the lattice coordinates correspond to a CNT. The terms are the same as before, with \( \epsilon_n \) being the on-site lattice energy and \( \hat{c}_{n/m}^\dagger / \hat{c}_{n/m} \) are fermionic annihilation/creation operators, respectively. While the Kwant code orders their matrices using very efficient methods [27], the MATLAB-based NEGF code used in this section generates a block-tridiagonal matrix. In this case, the diagonal blocks represent the Hamiltonians for each unit cell of the CNT and the off-diagonal blocks couple the unit cells together.

In these simulations, the hopping parameter is set to \( t_0 = -3.0 \) eV, representing the \( \pi \)-bond overlap between carbon atoms. In CNTs, the \( \pi \)-bonds are the only bonds which contribute to states close to the intrinsic Fermi energy, and thus we can neglect any others. In contrast to graphene which is gapless, depending on the way a CNT is rolled up, it can be
semiconducting or metallic [62]. In all the simulations in this chapter we will look at a (10,0) CNT, which is an zigzag edge tube with $E_G \approx 1.05$ eV. The other term in the Hamiltonian, the on-site energy (for site $n$), is described by

$$\epsilon_n = U_n(r_n) + eV_{ion}(r_n, r_i)$$  \hspace{1cm} (3.7)$$

where $U_n(r_n)$ is the potential induced by the back gate which shifts the Fermi energy of the CNT and $V_{ion}(r_n, r_i)$ is the potential due to an ion nearby the nanotube.

$V_{ion}(r_n, r_i)$ given by a simple Coulomb potential,

$$V_{ion}(r_n, r_i) = \frac{q_i}{4\pi\epsilon_0\epsilon_r|r_n - r_i|}$$  \hspace{1cm} (3.8)$$

where $r_n$ is the position of the $n^{th}$ lattice site and $r_i$ is the position of the ion. $q_i$ is the charge of the ion, $\epsilon_0$ is the vacuum permittivity, and $\epsilon_r$ is the relative permittivity. In our simulations we assume the CNT has the same relative permittivity as the surrounding environment (either air or SiO2), as doing otherwise would require significantly more computationally intensive methods, such as density functional theory calculations.

### 3.2.3 Computational details

There are two computationally intensive steps involved in the NEGF calculation. The first is in the calculation of the lead self-energy terms and the second is in the calculation of the Green’s function itself.

Treatment of contacts is particularly important in quantum transport simulations. In this chapter we assume the contacts are ideal, meaning that they are semi-infinite reservoirs of carriers which are reflection-less. The simple approach to this problem is to assume the contacts are simply extensions of the channel, in this case, more nanotube. To make the contact reflection-less, we match the potential of each contact to the on-site energy of the very last unit cell of the CNT on each end of the simulation cell. Otherwise, truncating the Coulomb potential at the edges of the channel would lead to steps in the on-site energy, resulting in reflections at the contacts.
The self-energies, $\Sigma^{L,R}$, account for the connection of the semi-infinite leads to the channel. The most straightforward method to calculate the self-energies is extremely inefficient and requires many matrix inversion steps for matrices the size of the number of sites in the device channel [18]. Sancho and Rubio, thankfully, developed an iterative method for calculating self-energies of layered materials [63]. The Sancho-Rubio method is based on imposing a Bloch boundary condition on the semi-infinite leads. Rather than working with matrices the size of the entire channel, with this method we consider the leads as a periodic series of unit cells with some coupling between each.

In fact, we never want to perform the full matrix inversion described in eq. (3.1). As matrix inversion is an $O(n^3)$ process, increasing the dimensions of our channel can rapidly make the calculation computationally intractable, both in terms of time and computer memory. Additionally, our observables do not actually require the full Green’s function— the relevant LDOS only come from the diagonal elements of the matrix and the matrix multiplication for transmission only requires the diagonal blocks. For this reason, we use the recursive Green’s function algorithm (RGFA) [64], a clever implementation similar to the Sancho-Rubio method which lets us simply target the diagonal blocks of the Green’s function, drastically reducing the computation time and memory requirements.

3.3 Results and Discussion

In this section we will consider a 50 nm long (10,0) CNT channel with a singly-charged ion centered along the transport direction, at various distances from the CNT. A schematic depicting the simulated device is presented in Fig. 3-1. While typical suspended CNT devices are on the order of micrometers long, the relatively small range of effect due to the Coulomb potential described earlier only requires us to investigate a small channel length.

We will begin by looking at the variation in LDOS across the device channel due to the Coulomb potential of the ion. From there we will see how the same potential affects transmission and current across through the nanotube channel. Finally, we will look at three special cases— a more realistic simulation guided by the recent results of Bushmaker et al [56], a CNT embedded in a dielectric (SiO$_2$) with a trapped charge, and a special case
Figure 3-1: 3D model summarizing the simulations performed in this chapter
This 3D model depicts a CNT FET suspended between source and drain contacts over a
gate/dielectric stack. The blue sphere represents an ionized gas molecule as the result of
proton irradiation, with vectors describing the location of the nanotube sites and the ion
position which are used in calculating the Coulomb potential throughout this chapter. The
nanotube studied in this chapter is a (10,0) zigzag CNT, which is a semiconductor with a
band gap around 1 eV. This figure is reproduced from ref. [1], copyright IEEE Transactions
on Nuclear Science, 2015.
3.3.1 Local Density of States

Fig. 3-2 shows NEGF-simulated LDOS for CNTs with ions placed at 2, 4, and 6 nm away from the center of the primary CNT axis. The LDOS is extracted along the transport direction (x-direction) of the CNT, where we sum the DOS around each unit cell of the nanotube as a function of Fermi energy. Additionally, the edges of the valence and conduction bands with the contribution from the Coulomb potential are overlayed in yellow-green in each plot. While it is very difficult move the Fermi energy beyond the first sub-band level (marked $E_{V1}$) in experiment, we show the second and edge of the third valence sub-bands as well.

The Coulomb potential bends down both the valence and conduction bands, creating a barrier for holes in the valence band and locally enhancing the LDOS in the conduction band. Since the enhancement of the conducton band LDOS is localized in the middle of the sample, it does not have a significant effect on the terminal characteristics of the CNT.

While the conduction band sees little effect due to the positively charged ion, the valence band sees a large effect. The Coulomb potential induces a barrier which decays as the ion is positioned further away from the CNT. In these maps we clearly see the wave-like nature of the LDOS, with dramatic interference patterns due interference of incoming and outgoing waves scattering off the barrier, similar to what has been observed in scanning tunneling spectroscopy measurements of systems with potential barriers [65],[66],[67]. Additionally, even for holes with energies above the barrier (more negative energies, for holes) we see significant modulation in the LDOS, another manifestation of quantum mechanical behavior.

3.3.2 Transmission

Fig. 3-3 shows the NEGF-simulated transmission spectrum through the nanotube as a function of energy for ions placed at a range of distances from extremely far away (where the CNT is pristine) up to 2 nm away (where there Coulomb potential at the nanotube is strong). As the effect we are concerned with is for p-type CNT FETs, we focus just on the
Figure 3-2: CNT local density of states under the influence of an ion at different distances
NEG F calculation of the LDOS along the transport direction of a CNT FET for an ion located at the middle of the tube at distances $D = 2$, $4$, and $6$ nm away. The units of the map are number of states per unit cell per unit energy. The band-bending of the valence and conduction bands due to the Coulomb potential is denoted as green lines in each panel. This figure is reproduced from ref. [1], copyright IEEE Transactions on Nuclear Science, 2015.
When the CNT is pristine, we see sharp steps in the transmission corresponding to the valence band edge and then the second sub-band edge. As the ion is brought closer to the CNT, the barrier which was observed in the LDOS (Fig. 3-2) begins to form, blocking transmission. As the potential barrier is fairly broad, tunneling is not a strong effect and transmission drops to nearly zero for any energies which are below the barrier height.

Once the carrier energy exceeds the barrier height, transmission eventually recovers to the expected pristine level. However, this is not an abrupt transition as we see at the band edges for the pristine case, but more gradual. This gradual transition is more noticeable the closer the ion is to the nanotube, which may also be seen in the LDOS in Fig. 3-2, where the ion at 2 nm shows more perturbation in the LDOS above the barrier than the ion at 4 or 6 nm.
3.3.3 Drain Current

Fig. 3-4a shows simulated current (using Landauer’s equation) as a function of Fermi energy assuming $T = 300$ K and $V_{DS} = -0.1$ V. As we are assuming the total CNT length is actually significantly longer than our channel (on the scale of micrometers), the overall electric field for a 0.1 V bias is quite small. This would cause only a slight perturbation to the energy bands shown in Fig. 3-2 and thus we will use the transmission from Fig. 3-3 without any more complex treatment of the band structure of the channel. Again, since we are mostly concerned with hole transport, we only show the valence band currents. We also restrict the range of Fermi energies such that we only see the current through the first valence sub-band. For ions placed up to 20 nm away from the CNT, we observe a strong reduction in the simulated drain current. Ions which are closer than 10 nm away cause very large drops in current, effectively switching off the current.

We reproduce the measurements of Bushmaker et al [55] in Fig. 3-4b for comparison. The strong reduction of current in their measurement (the portion labeled as major SEEs) is similar to the reduction we simulate for ions closer than 10 nm to the CNT. Given this similarly, it was not surprising when further experimental measurements indicated that the ion was actually sitting directly on the CNT itself [56]. Indeed, DFT simulations performed in the work by Bushmaker et al indicate that partial charge transfer between adsorbed ions and the CNT yields a very similar potential profile as our simulations for the ion at 4 nm away from the CNT. While these mechanisms for producing the Coulomb potential are quite different, the potential profile is what influences the transport, and thus we are still able to gain insight into the experiment.

3.3.4 Ion adsorbates

In light of the more recent results from Bushmaker et al [56], in Fig. 3-5 we show a modified LDOS and transmission simulation for an ion adsorbed onto the CNT surface. Using the parameters provided by Bushmaker et al’s DFT calculations, the ion is placed 7 Å away from the surface of the nanotube and we assume there is charge transfer from the ion to the CNT, such that $q = 0.18e$. While the adsorbed ion produces a slightly different potential
Figure 3-4: Comparison of the simulated drain current for different ion distances with the results of Bushmaker et al (A) NEGF calculation of the drain current corresponding to the transmission calculation shown in Fig. 3-3. A 0.1 V bias and a temperature of 300 K are used in the calculation. Panel (A) is reproduced from ref. [1], copyright *IEEE Transactions on Nuclear Science*, 2015. (B) Experimental measurements on a suspended CNT FET by Bushmaker et al [55]. Panel (B) is reproduced from the work of Bushmaker et al, copyright *IEEE Transactions on Nuclear Science*, 2014.
Figure 3-5: LDOS and transmission calculation assuming an ion adsorbate with charge transfer to the CNT

In light of the more recent results from Bushmaker et al. [56], we include a simulation assuming an ion adsorbate directly on the nanotube. We assume a partial charge transfer from the ion to the CNT, such that the ion has $q = 0.18e$, which reduces the effect of the ion being so close to the surface of the CNT.

In Fig. 3-6 we consider an interesting thought experiment of a rather extreme version of the ion adsorbed onto the CNT. Ions are placed at the same position as Fig. 3-5, except this time the charge transfer is less pronounced such that $q = 0.5e$ and $q = e$. In each case, the band bending due to the ion is strong enough such that confined states form in the potential well. For $q = 0.5e$ there are two confined states which form in between the valence band edges. For $q = e$ the effect is so pronounced that the conduction band edge is actually brought below the level of the pristine valence band edge.

For each case we plot the transmission spectra, showing dramatically reduced transmission, except for energies close to the confined levels from the Coulomb potential. At these specific energies, carriers are able to resonantly tunnel through the confined states within the band gap. These transmission spectra are quite similar to what is observed from resonant tunneling diodes formed in semiconducting 2DEGs [18].

While these results do not consider the complex nature of quantum capacitance which would be present in a case with such strong band bending, these results present an intriguing goal for experimentalists to realize. If one could produce a potential similar to the ones shown in Fig. 3-6 it would provide a method for creating ultimately-scaled one-dimensional tunnel
**Figure 3-6: Resonant tunneling effects resulting from ion adsorbates** Here we show the LDOS and corresponding transmission function for ions adsorbed to the CNT with reduced levels of charge transfer compared to Fig. 3-5. These results present a potential goal for future experiments, suggesting that reducing the charge transfer from the adsorbed ion could produce a strong enough potential to result in confined states within the potential well. These are seen as bright spots in the LDOS within the band gap of the CNT, which result in resonant tunneling similar to what is seen in resonant tunneling diodes, FETs and tunnel diodes.

One obvious place to start would be in engineering the amount of charge transfer between the adsorbed ions and the CNT. This could be achieved by using ultra-thin high-κ dielectric barriers wrapping the CNT. Another possible route would be to look at different ion adsorbates. For example, negatively charged (so the band bending would enter the conduction band instead) fluorine ions have been predicted to retain nearly 40% of their charge on the surface of graphene, similar to the first case presented in Fig. 3-6. Additionally, adsorbed potassium has been ionized by heating on the surface of graphene and may be a good candidate for study.

68
3.3.5 Encapsulated CNTs

While interesting in the laboratory, generally nanoelectronics based on CNTs require that the CNT be encapsulated in some dielectric. Therefore, to complete this chapter, we will consider the effect of an ion trapped in a dielectric, in this case a nanotube encapsulated in SiO₂. In Fig. 3-7 we show results for a CNT in a SiO₂ dielectric with an ion placed 2 nm away from the CNT. Fig. 3-7a shows the LDOS, with the band bending marked with a solid line (a dotted line is given for comparison for the suspended case). Fig. 3-7b shows the transmission spectrum comparing the pristine CNT, the suspended CNT, and the CNT in SiO₂. Whilst the effect of the trapped ion is diminished due to the enhanced dielectric screening, we note that there is a high likelihood of multiple charges being trapped in the dielectric. Given the one-dimensional nature of transport in CNTs, this would be extremely detrimental to the electronic properties of an encapsulated CNT FET. Thus, radiation-hardened dielectrics would be particularly important in CNT FETs subjected to harsh radiation environments.

3.4 Chapter conclusions

In this chapter we have presented NEGF-based simulation results for CNTs in the presence of Coulomb potentials due to nearby positively charged ions. We saw how the Coulomb potential induced band bending in the CNT, leading to perturbations in the LDOS along the tube. Positive ions produced local potential barriers for holes, which drastically reduced the quantum mechanical transmission through the nanotube. Additionally, we showed simulation results suggesting that ions adsorbed to the CNT may be exploited to study resonant tunneling effects in CNTs. Finally, the effect of trapped charges in an SiO₂ dielectric encapsulating a CNT FET was studied. These results provide a model for future analysis of CNTs subjected to radiation, as well as a pathway to interesting experiments involving gating CNTs with single ions.
Figure 3-7: Comparison of the effect of different dielectric environments. This figure shows an extension to the initial model presented in the main text, now assuming the CNT is encapsulated in an SiO$_2$ dielectric with $\epsilon_r = 3.9$. (A) shows the LDOS calculation for a trapped charge at $D = 2$ nm within the SiO$_2$, with dotted lines super-imposed for comparison to the CNT in air. (B) The corresponding transmission function calculation comparing the pristine CNT, a trapped charge as in (A), and an ion at $D = 2$ nm in air. This figure was reproduced from ref. [1], copyright IEEE Transactions on Nuclear Science, 2015.
Chapter 4

Quantum transport modeling of graphene devices with radiation induced Coulomb scatterers

Graphene’s high electron mobility [69] and thermal conductivity [70] have made it a candidate for post-silicon electronics [71]. Given graphene’s potential, it is important to exhaustively understand its properties. One area of interest is within the field of radiation effects, where we study how various forms of radiation affect properties of materials. In this chapter we will work to model a recent radiation effects study of graphene [72]. We will see how quantum transport methods may be used to shed light on the underlying electronic transport mechanisms in graphene in the presence of radiation induced Coulomb scatterers. First, we will work to model the pristine characteristics of a graphene-based test structure. From there, we will introduce Coulomb scatterers and study their effect on charge transport in the graphene.

4.1 Introduction

A typical graphene electronic test structure (depicted in Fig. 4-1a) will consist of a graphene ribbon with metal contacts placed upon some insulating substrate which is coupled to a conductive gate. The prototypical substrate for graphene has been a heavily doped silicon wafer.
Figure 4-1: Visual summary of the graphene TID problem (A) Schematic of a basic graphene test structure. The test structure consists of a graphene ribbon with source and drain contacts placed upon a gate dielectric ($\text{SiO}_2$) which has been grown upon a heavily-doped silicon wafer (which is used as a back gate). In this image, the green hue depicted within the gate dielectric represent the charging effects due to TID exposure which will be studied within this chapter. (B) Energy band diagram summarizing the transport model developed in this chapter depicting a doping offset in the source and drain contacts as well as disorder due to the Coulomb potential from charged impurities produced by TID. This figure is reproduced from [2], copyright *IEEE Transactions on Nuclear Science*, 2017.
with native silicon dioxide dielectric grown on top the silicon. In this case, the heavily-
doped wafer may act as the gate, letting one sample the electronic structure of the graphene
by varying the gate voltage and measuring the resistance of the graphene ribbon. This
measurement, which is referred to as a “transfer curve,” provides a complete electrical char-
acterization of the graphene test structure, giving us insight into the quality of the graphene,
the contacts, and the substrate itself.

The substrate of the sample plays a major role in determining graphene’s response to
radiation [52]. Several studies have been presented in the literature comparing the radiation
response of graphene on substrates like SiO$_2$ [73], SiO$_2$ passivated with trimethylsiloxyl (TMS)
[72], and $h$-BN [74]. Not only do these various substrates affect the electronic transport
properties of the pristine graphene test structure, they also dictate the mechanism of the
device response to being irradiated.

In this chapter, we will focus on modeling a recent total ionizing dose (TID) study
performed by Cress et al on graphene devices on SiO$_2$/TMS substrates [72]. In the study by
Cress et al, the graphene was subjected to varying levels of gamma ray exposure from a $^{60}$Co
source. Their measurements indicated that the TID introduced trapped positive charges in
the gate dielectric. It is well known that such trapped charges as responsible for degrading
the electrical performances of graphene devices [75, 76, 77].

Modeling the work by Cress et al [72] presents two challenges. First, we need to capture
the characteristics of the pristine graphene device which was measured. We will build off
of our standard tight-binding model, including effects which are specific to the experiment.
Commonly in graphene devices, especially on SiO$_2$ substrates, there is asymmetry between
the electron and hole conduction in graphene. In this case, the electron branch is weaker
than the hole branch. To account for this effect, we will develop a model for the Fermi level
pinning effects introduced by the metal-graphene contact interfaces.

Once we have captured the characteristics of the pristine sample, we will introduce
trapped charges in the dielectric, mimicking the effect of TID due to gamma irradiation.
Finally, we will explain the modeled current by investigating quantum mechanical transmis-
sion functions, local density of states, and spatially-resolved current density in the presence
of trapped charges. This modeling effort will help us further understand the mechanism of
how gamma ray-induced TID affects transport in graphene. Additionally, this work will serve as a guide for future work in modeling radiation effects in two-dimensional material-based devices.

4.2 Transport model

Our first goal is to model the pristine characteristics of the graphene sample studied by Cress et al [72]. We will begin with the basic nearest neighbor tight-binding model for graphene as we have used before. However, this model approximates “perfect” graphene, which is challenging to achieve in the laboratory. Without any special modifications, the standard tight-binding transport model predicts a perfectly symmetric transfer curve, as is shown with a dashed line in Fig. 4-3a. To match the data from Cress et al [72], reproduced in Fig. 4-3b, we will need to add some extrinsic effect to the model.

4.2.1 Contact model

The metal contacts attached to the channel can have a significant effect on transport through the device channel. An easy first-order assumption is that the metal contact on top of graphene will completely pin the Fermi energy of the covered areas, however that is not the case. Studies on CNTs have shown that, while metal contacts interposed between the gate and the channel can shield the channel, metal on top of the channel does not fully control the electrostatics [78, 79, 41]. Indeed, the back gate of the sample can compete with the contact for electrostatic control over the channel, modulating the Fermi energy of the nanotubes even when covered in metal.

Similar effects to the above studies on CNTs have been demonstrated in graphene. First principles calculations have also predicted that metals on the surface of graphene can preferentially dope the graphene $p$ or $n$-type [80]. These predictions have been supported by experimental evidence, where it is possible to approximate the contact-graphene-back gate system as a competing set of parallel plate capacitors [81]. Additional experiments have also suggested that there is not a complete pinning of the Fermi level of the graphene beneath the metal contacts [82, 83].
It is well known that when the metal-induced doping effect is strong that it can cause the conduction asymmetry as in the device we are trying to model [84]. In this case, the shift in the Fermi energy underneath the contact is sufficient to cause a significant offset in the energy bands around the perimeter of the metal overlapping graphene. This effect appears in transport measurements as a pronounced secondary Dirac point.

Due to the large conduction asymmetry demonstrated in Fig. 4-3b, it is apparent that we need to account for metal-induced doping of the graphene. We accomplish this by introducing an offset in the energy bands at the contact edges. We estimate that the contacts shift the energy of the graphene underneath the metal contacts $p$-type by 75 meV, which produces asymmetry between the hole and electron branches of the transfer curve. This contact-doping effect which we are including in the model is illustrated in Fig. 4-1b.

4.2.2 Coulomb scatterer model

To account for the effect of TID, we introduce Coulomb potentials in the simulated graphene channel. These potentials act as scattering sites and are induced by trapped charges created in the gate dielectric by irradiation. Each lattice site in our simulation is subjected to the potential of the ensemble of trapped charges, given by:

$$V_{\text{charges}} (r_i) = \sum_n \frac{z_n e}{4 \pi \varepsilon_0 \varepsilon_r |r_i - r_n|}, \quad (4.1)$$

where $n$ is the number of trapped charges, $z_n$ is the charge state for each trap (here we always assume $z_n = +1$), $e$ is the electron charge, $\varepsilon_0$ is the vacuum permittivity, and $\varepsilon_r = 3.9$ is the dielectric constant of SiO$_2$. The term $r_i$ is the position of the $i^{th}$ lattice site and $r_n$ is the position of the $n^{th}$ trapped charge.

We include the potential described by eq. (4.1) as a perturbation to the on-site energy of each site $i$ in our simulation cell:

$$\epsilon_i = eV_{\text{charges}} (r_i) + U_{\text{BG}} (r_i) \quad (4.2)$$

$U_{\text{BG}} (r_i)$ is an additional shift in the on-site energy which is due to the capacitive coupling to
Figure 4-2: Electric potential due to trapped charges at different densities

The electric potential is calculated in each panel for three different randomly generated trapped charge densities ($1 \times 10^{15}$ cm$^{-3}$, $2 \times 10^{15}$ cm$^{-3}$, and $3 \times 10^{15}$ cm$^{-3}$), discussed within the main text. This figure is reproduced from [2], copyright IEEE Transactions on Nuclear Science, 2017.

When there are no trapped charges in the dielectric, we will consider the device to be pristine. From there, we will add increasing numbers of charge, up to a density of $3.25 \times 10^{15}$ cm$^{-3}$. The range of volumetric trapped charge density we will employ is comparable to the areal trapped charge densities which were predicted in the work by Cress et al [72].

In our simulations we randomly place trapped charges according to a uniform distribution to achieve different charge densities. When simulating the transfer characteristics of the device, we average over several different ensembles of charges at each density. This is important since our simulation cells are significantly smaller than the measured devices (due to computational constraints). Whereas a large device has many trapped charges, statistically averaging out the influence of sites at the extremes, ours just has a few charges. Simulating several ensembles of charge distributions lets us work around this problem without having to simulate extremely large sample sizes. Examples of various potentials $V_{\text{charges}}(r_i)$ for typical trap densities are shown in Fig. 4-2.

### 4.2.3 Gate voltage considerations

When comparing our simulations to experimental data it is preferable to be able to tie directly to the applied gate voltage used in measurements. For this reason, we would like...
to convert the Fermi energy of the sample to the applied gate voltage. For our purposes, it is sufficient to neglect quantum capacitance and only consider the geometric capacitance of the back gate/graphene system (a parallel plate capacitor).

The relationship of the gate voltage to the total charge density in the channel and the geometric capacitance is given by $V_G = \frac{\rho_g}{C_{geom}}$. The carrier density is given by $\rho_g = e \int_{-\infty}^{+\infty} dE \left[ D(E) f(E, U_{BG}) - D(E) ((1 - f(E, U_{BG})) \right]$. $D(E)$ is the density of states of graphene which is given by $D(E) = \frac{g_s g_v |E|}{2\pi(\hbar \nu)^2}$ as we have described earlier, $f(E, U_{BG})$ is the Fermi-Dirac function, with $U_{BG}$ being the shift in the Fermi energy in response to the gate. The geometric capacitance is given by $C_{geom} = \frac{\epsilon_0 \epsilon_r t_{ox}}{t_{ox}}$. Since the device measured by Cress et al is quite large (on the scale of micrometers), we neglect the effect of quantum capacitance and contact-channel capacitances.

In practice, we perform our simulations by inputing a range of different Fermi energies directly, corresponding to a gate voltage sweep. After the simulation has finished, we re-scale these Fermi energies to gate voltages by numerically solving the integral describing $\rho_g$. This basic conversion enables easy comparison with experimental measurements.

### 4.3 Results and Discussion

In this section we will present the results of the model described above. We will begin by comparing our simulated transfer curves with the work by Cress et al. First we will show show the effect of our doped-contact model and then we will introduce varying levels of trapped charges to account for the effect of TID. To explain the simulated transfer characteristics we will study the gate-dependent transmission function for different trapped charge densities, which will show the mechanism of transport degradation. Finally, to learn about the effect of Coulomb scatterers on charge transport, we will look at visualizations of both the LDOS and spatially resolved current density.

#### 4.3.1 Transfer Characteristics

Fig. 4-3a shows our tight-binding simulated transfer curves for a 200 × 200 nm graphene ribbon assuming a gate SiO$_2$ thickness of 100 nm and a temperature of 300 K. We include a
small source-drain bias of 0.1 V and simulate the current for different gate voltages. These conditions are similar to those used in the measurements by Cress et al. [72], reproduced in Fig. 4-3b, except their device had a significantly larger area, which we will discuss further.

Without our contact model, the dashed line in Fig. 4-3a marked as “No c.m.”, we see a completely symmetric transfer curve. When we repeat the same simulation, now including our +75 meV shift in the on-site energy in each device contact, we achieve the desired transport asymmetry. Our contact doping model makes the contacts preferentially inject holes, therefore for the same gate voltage magnitude we see reduced electron conduction. In fact, this on-site energy shift introduces scattering at the ends of the device for both types of carriers, which reduces the total current through the sample.

While our contact model makes our simulated transfer curve quite close to the measurement by Cress et al., it is necessary to explain a few discrepancies before we continue. We assume a fixed shift in the Fermi energy of the graphene underneath the metal contact, as suggested by the work of Giovannetti et al. [80]. However, it has been demonstrated that this contact-incuded doping can be modulated with a gate by Xia et al. [81], an effect which we do not consider. To include this effect, it is necessary to estimate several parameters about the system, such as charge transfer between the metal contact and graphene and the metal-graphene interlayer difference. As these parameters are difficult to estimate and overall the resulting effect is minimal, we have opted to use our more simple model.

Additionally, it is necessary to point out that our model overestimates the current passed through the graphene. Our simulated sample does not include the effect of any trapped charge which is intrinsic to the SiO2 gate dielectric. While the TMS treatment performed by Cress et al. compensates for some of the most egregious traps, the measurements still indicate significant disorder present in the channel, reducing the graphene mobility. We also do not consider phonon scattering which creates an upper limit for the mobility of graphene at room temperature [85]. These effects both act to reduce the current levels in the experimental measurements.

As we are mostly concerned about capturing the effect of TID in the sample, it is most important to match the functional form of the experimentally measured transfer curve. Our contact model introduces sufficient asymmetry to the transfer curve such that our simulation
captures the essential aspects of the experiment. Therefore, we continue with discussing the
effect of TID on the sample. However, there is an opening for future studies aimed at more
closely tying to the pristine characteristics of the sample.

Now we will look at how adding trapped charges in the gate dielectric affects the transfer
curve. In Fig 4-3a we also show transfer curve simulations of ten different disorder con-
figurations averaged together for trapped charge density ranging from $2.5 \times 10^{14}$ cm$^{-3}$ to
$3.25 \times 10^{15}$ cm$^{-3}$. The trapped charges introduce a disorder potential in the graphene ac-
cording to eq. (4.1), which perturbs carrier flow through the sample. This disorder causes
Coulomb scattering in the channel and reduces the current flow. Importantly, we see that
the hole branch degrades at a faster rate than the electron branch. This result is congruent
with what was observed by Cress et al [72] (Fig. 4-3b) and also what we have previously
simulated for the effect of trapped charges in CNTs [1] (discussed in Chapter 3).

We also observe a slight shift in the minimum conductivity point towards more negative
gate voltages. This is also present, though more pronounced, in the experimental data in
Fig. 4-3b. It is not surprising that introducing positive trapped charges in the gate dielectric
should shift the transfer curve towards more negative gate voltages. We expect that the
reason this effect is somewhat weaker in our simulation is that the effect of the Coulomb
potentials are cut off at the edges of our simulation cell, where the contacts connect. The
200 nm channel length also exacerbates this effect, as some of the long range effects of the
Coulomb potentials are not captured in such a small dimension. Future experimental studies
isolating the effect of irradiation on the device contacts and channels would be useful for
improving this aspect of the model.

In analyzing our simulation in Fig 4-3a it is useful to look at the tight-binding transmis-
sion function which was used to give the final current result. This transmission function,
which is shown in Fig. 4-4 is convoluted with the Fermi window, $f_S(E, E_{FS}) - f_D(E, E_{FD})$,
according to Landauer’s equation. The Fermi window, whose size is determined by both
the applied bias and temperature, smears the transmission function. The Fermi window is
also a general concept, not exclusive to the problem we are studying, while the transmission
function is specific to our particular problem. In this sense, the transmission function gives
us a more fundamental picture of carrier transport and lends us more physical insight than
Indeed, the features in Fig. 4-3a result from a dramatic smearing of the transmission function in Fig. 4-4. In the transmission function we see several interesting features which require analysis. Starting with the pristine characteristic, we note that there are two minimum transmission points. The first minimum transmission, at $V_G = 0$ V, is due to the primary Dirac point of the channel and is normal for graphene. The second minimum transmission, at $V_G \approx 7$ V, is a result of our contact doping model. This secondary Dirac point, once convoluted with the Fermi window, is what serves to bring down the electron branch of the transfer curve.

As a result of our contact doping model, there are three transport regimes as the back gate voltage sweeps from negative to positive. As the contacts are fixed more $p$-type than the channel, for negative gate voltages the sample is always configured as a $p^+-p-p^+$ junction. Between $V_G = 0$ V and $V_G \approx 7$ V the device is configured as $p-n-p$. Finally, above $V_G \approx 7$ V the device is configured as $n-n^+-n$. If the contact doping is effect is larger than the size of the Fermi window, the transition between $p-n-p$ and $n-n^+-n$ results in a secondary minimum current point, as has been shown experimentally by DiBartolomeo et al [84].

With increasing trapped charge density, the transmission function shows a shift in the primary Dirac point towards more negative gate voltages. Normally, the primary Dirac point corresponds to the entire channel being un-doped, however introduction of disorder makes it so this condition is never met. As the positive charges collectively gate the channel, the intrinsic point shifts more and more towards negative gate voltages.

As the transmission functions for non-zero trapped charge density are the average of several disorder realizations, we can gain some insight by looking at the variation in the calculated transmission. We do this by shading the standard deviation of the constituent simulations around each transmission function in Fig. 4-4. For $V_G > 0$, the variation between disorder realizations is imperceptible regardless of level of trapped charge density. This indicates that the mechanism of current degradation for the electron branch is insensitive to any outliers in the trapped charge distribution.

For the hole branch we see some variation between disorder realizations, which increases with higher amounts of trapped charges. For the smallest trap density we study, $2.5 \times 10^{14}$
cm$^{-3}$, there is a very small amount of variation. This is not surprising given the small number of trapped charges in this configuration. However, it tells us that the degradation is not heavily influenced by the position any single individual trapped charge, a trait we will focus on more in looking at transport visualizations.

4.3.2 Transport visualization

The variance between different disorder configurations in Fig. 4-4 is quite small, indicating that the trapped charge density is more significant in determining the level of transport degradation than any given charge distribution. This is evident in the potential profiles shown in Fig. 4-2. For comparison, we show each potential map using the same color scale. We observe that except for a few outliers, the potential in each map is more or less uniform for each trap density. Therefore, we reason that the ensemble of charges act as a global gate on the channel, shifting the carrier-type more towards electrons, reducing the number of hole-like modes. This reduction in the number of conductive modes is the primary mechanism of transport degradation for the hole branch.

To help us understand these effects more, it is useful to look at the influence of a single trapped charge on the LDOS in the channel. In Fig. 4-5, we show calculated LDOS (units: number of states/eV/unit cell area) spectra for a single trapped charge at $[0, 100 \text{ nm}, z]$ (where $z = 5$ or $20 \text{ nm}$). In these plots, we take a slice along the $x$-direction at $y = 100 \text{ nm}$ (the middle of the ribbon) for different Fermi energies. As our goal is to understand the effect of the trapped charge, we turn off the contact model for the LDOS plots.

Previously, we saw how in materials with a band gap, trapped charges produce a hard wall potential which blocks transport. In graphene, the effect is much more subtle. Rather than a hard wall potential, the trapped charge produces continuous variations in the LDOS. A positive trapped charge will enhance the LDOS of electrons whilst decreasing the LDOS for holes.

For the trap at $z = 5 \text{ nm}$ there is a significant modulation in the LDOS localized at $x = 0 \text{ nm}$, as well as a pronounced asymmetry in the LDOS between electrons and holes for Fermi energies close to the Dirac point (between $\approx \pm 50 \text{ meV}$). For the trap at $z = 20 \text{ nm}$, we observe only a slight bending in the LDOS, with very little localized effect (as shown for
Figure 4-3: Benchmarking tight-binding simulation against the experimental data of Cress et al (A) Simulated transfer curves comparing a graphene device without the contact model, the pristine device with our contact model included, and then the device with contact model and trapped charges in increasing densities (in units of $1/cm^3$). The simulated graphene device is $200 \times 200$ nm and assumes a gate dielectric of 100 nm thick SiO$_2$. This panel is reproduced from [2], copyright IEEE Transactions on Nuclear Science, 2017. B Experimental measurements on Co-60 irradiated graphene for TID between 0 and 2000 krad(Si), reproduced from the work of Cress et al [72], copyright IEEE Transactions on Nuclear Science, 2012.
Figure 4-4: Simulated transmission function corresponding to the current calculation in Fig. 4-3a. Each curve is the result of an average of different transmission simulations for ten realizations of trapped charge distributions for each density. The variance of the different simulations is shaded around each line, as is discussed within the main text. This figure is reproduced from [2], copyright IEEE Transactions on Nuclear Science, 2017.
the $z = 5 \text{ nm}$ trap). However, we observe nearly the same amount of asymmetry between electron and hole LDOS, demonstrating the long range effect of the Coulomb potential.

There are many interesting resonant states in each pane of Fig. 4-5, however they do not have a significant contribution in the final transfer characteristics of the sample. These states are due to standing waves, relating to the waveguide-like wavefunctions of each propagating mode. At extremely low temperature, it may be possible to detect these fluctuations, however at room temperature the Fermi window completely smooths out these effects.

Now that we have analyzed the influence of a single trapped charge on the LDOS inside the channel, we will investigate the spatially resolved current density in the sample. The figures are oriented such that contacts are on the top and bottom of the square channel. For clarity, we only plot the injection from the bottom contact. In Fig. 4-6 we show the current density for 100 meV electrons (Fig. 4-6a) and -100 meV holes (Fig. 4-6b) in the same distribution of charges for a trap density of $2.25 \times 10^{15} \text{ cm}^{-3}$. These carrier energies correspond to gate voltages of about $\pm 10 \text{ V}$. The position of each trapped charge is marked by a black “×” and larger markers indicate traps which are closer to the graphene.

The maps in Fig. 4-6 include the contact model described earlier. In the case of electrons (Fig. 4-6a), the channel is more heavily $n$-type than the contacts, which creates a resonant cavity, resulting in a more active interference pattern than for holes. Additionally, with the full contact model we can calculate the transmission, giving further context to the maps, which is included as two scatter points in Fig. 4-4.

While there are chaotic interference patterns in Fig. 4-6a, this current density map only corresponds to a small reduction in the overall transmission in the channel. As we have seen, the positive trapped charges act to enhance the LDOS inside the channel. This effect introduces more scattering, but in graphene this scattering simply redirects the current slightly, rather than backscattering the carriers entirely. This weak scattering effect has some effect on reducing the transmission through the channel, however it is not nearly as strong as the effect on hole transport.

In contrast, for holes (in Fig. 4-6b), there is a significant reduction in the local current density around the trapped charges. As the contacts do not produce a resonant cavity in this
Figure 4-5: Energy-resolved LDOS in the presence of a trapped charge. Slice of tight-binding simulation of local density of states along central axis of the transport direction for a single positively charged trap at (0, 100 nm, 5 nm) (left panel) and (0, 100 nm, 20 nm) (right panel). The LDOS is calculated as a function of Fermi energy and a single slice is plotted along the transport direction, taken from the middle of the sheet at $y = 100$ nm. This figure is reproduced from [2], copyright IEEE Transactions on Nuclear Science, 2017.
case, the maps are significantly more simple. We note a clear reduction in the local current density around each trapped charge. The areas surrounding each trapped charge have a reduced carrier density and scatter the electrons. This scattering is not as pronounced as for a gapped material, but is still present in graphene. In this case, the scattering off the Coulomb potentials is responsible for the variation we see in Fig. 4-4 while the primary reduction in current is from the reduction in conductive modes we discussed earlier.

### 4.4 Chapter conclusions

In this chapter we have seen how quantum transport modeling approaches may be used to understand the effect of total ionizing dose on graphene-based devices.

First we developed a basic model to capture the pristine device characteristics observed by Cress et al. [72], where we had to introduce a contact doping effect resulting in asymmetry between the electron and hole branches of the device transfer curve. Next, we introduced increasing levels of trapped positive charges, simulating the effect of gamma irradiation-induced TID. The Coulomb potential due to the trapped charges degraded transport through the channel. Similar to the experimental results, we saw that hole transport was more heavily degraded than electron transport for equal and opposite gate voltages.

We analyzed our simulated transfer curve by studying their constituent transmission functions, giving us a more fundamental picture of the carrier transport. From there, we looked at several visualizations such as representative Coulomb potentials, local density of states spectra, and local current density maps. From this analysis, we concluded that the primary effect which reduces current for the hole branch is the reduction of conductive modes resulting from the long range effects of the Coulomb potential. We also saw how trapped charges close to the graphene could introduce significant scattering, resulting in the variation between random disorder realizations.

Our results represent one of the first times tight-binding transport has been used to study radiation effects in nanoelectronic devices. They will serve as a template for future work in devices where quantum effects and ballistic transport are present. In the future, this work may be improved in two major ways. First, a more complex electrostatic model for the
Figure 4-6: Tight-binding calculation of local current density in the presence of trapped charges. Each map shows spatially resolved current density injected from a contact at \( y = 0 \) nm (transporting towards another contact at \( y = 200 \) nm). Trapped charges are included, which are marked on the maps by black “x” marks in each plot, with charges closer to the graphene denoted by larger marks. Each figure uses the same distribution of trapped charges, but (A) is for 100 meV electrons and (B) is for -100 meV holes (the trapped charge labels are easier to see in (B)). The corresponding transmission for each of these maps is marked as scatter points in Fig. 4-4. This figure is reproduced from [2], copyright IEEE Transactions on Nuclear Science, 2017.
pristine device could be developed. Second, the effects of radiation on the device channel and the device contacts should be decoupled. Our results also suggest that experimentally measuring irradiated graphene at very low temperatures would be enlightening, as we observe some interesting features near the minimum transmission points which are obscured in the room temperature transfer curve.
Chapter 5

Landau level mixing in heavily graded graphene $p$-$n$ junctions

Quantum Hall effects are amongst the most heavily studied phenomena in condensed matter physics. Indeed, we have already seen how von Klitzing’s seminal experiment pushed forward the entire community. 2DEGs formed from semiconductor heterostructures offer an excellent platform for studying quantum Hall systems, however they have one major limitation—generally they are restricted to unipolar transport. In graphene, it is possible to form an ambipolar quantum Hall state at a $p$-$n$ junction. The Landau levels in graphene form interesting snake states at the junction, and the electron/hole states have a chance of tunneling into one another, or mixing, giving rise to new quantum Hall plateaus. In this chapter, we will use tight-binding transport modeling to investigate a recent experimental work in which Landau level mixing is suppressed at the GPNJ.

5.1 Introduction- Quantum Hall effect in graphene $p$-$n$ junctions

The trajectory Quantum Hall edge states in a 2DEG is classically determined by the position of guiding centers, which steer the carriers [17]. Simply, the Landau level will track along
an equipotential line to satisfy the criteria

\[ E_F - E_{LLn} = qV_{\text{Local}}(x, y), \]  

(5.1)

where \( V_{\text{Local}}(x, y) \) is the local electrical potential in the 2DEG. Generally, this forces the Landau levels to the edges of the 2DEG, hence the name "edge state". The ability to easily spatially vary the carrier density in a graphene sheet, however, opens up the interesting possibility of electrically steering the Landau levels.

Beyond simply steering the Landau level, the ambipolar capability of graphene lets us study Landau level behavior at the domain wall between \( n \)-type and \( p \)-type graphene. The electron- and hole-like Landau levels in a GPNJ will travel up the junction and counter-propagate at the junction interface, forming what we refer to as a "snake state". The conductance across this junction will be determined by the degree of Landau level mixing at the junction, which is where carriers will transition from occupying one Landau level to another. Abanin et al.\[86\] first predicted the conductance given in the full Landau level mixing condition:

\[ G_{PN} = \frac{|\nu_1||\nu_2|}{|\nu_1| + |\nu_2|} \]  

where \( \nu_{1/2} = \pm 2, \pm 6, \pm 10... \) is the filling factor (essentially, the number of occupied Landau levels) of the left/right side of the junction. They assumed an equal probability of any mode at the junction transmitting into any mode, yielding eq. \(5.2\).

Landau levels in graphene can be steered with electrostatic gating. As shown in Fig. 5-1A, a bottom and top gate configuration was successfully used by Williams et al to form a reconfigurable graphene \( p-n \) junction where the two halves of their device could be continuously and independently tuned from \( n \) to \( p \) type \[20\]. The top/bottom gate scheme tends to produce an asymmetric potential profile whose junction width \( D_W \) is highly dependent on the relative voltages of the gates.

Williams et al verified eq. \(5.2\) in their device by fixing the magnetic field applied to their sample to \( B = 4 \) T and measuring the conductance. By separately varying the bottom and top gate voltages of their device they created all configurations of junctions, such as \( n-n' \), \( p-p' \), \( p-n \), and \( n-p \). In the ambipolar junction configuration, nearly complete mixing of
the Landau levels was observed, showing the 1, 3/2, and 3 plateaus predicted by eq. (5.2).

Another recent experiment studying quantum Hall effect in graphene was performed by Klimov et al [87]. Their devices were built upon buried split-gate structures fabricated at CNSE’s 300 mm chip fabrication plant. These split-gates are particularly useful for studying quantum Hall effect in graphene due to the fact that they give nearly perfectly symmetrical control over the carrier density on each side of the device. This effect is clearly observed when comparing the skewing (or lack thereof) of the plateaus in Fig. 5-1B and 5-1D. Top and bottom gates have asymmetrical control over the channel, leading to the skewed plateaus observed in the conductance in Fig. 5-1B. However, in devices fabricated from our lab’s split-gates, the different configurations are clearly defined, with virtually no skewing.

Furthermore, the junction width correlates more directly with the split-gate gap, especially when compared to a top and bottom gate-scheme. Klimov et al utilized this feature to create junctions with significantly larger values of $D_W$ compared to Williams et al, which has a profound effect on the observed conductance. Indeed, in the $p-n$ regime, they observe a conductance plateau of $1 \frac{e^2}{h}$ (the red plateau in Fig. 5-1D) over a wide range of filling factors.

In Fig. 5-1D as the filling factor increases, the corresponding conductance does not, pinning at $1 \frac{e^2}{h}$. Klimov et al attributed this effect to the suppression of Landau level mixing at the $p-n$ junction. The 150 nm split-gate spacing used in their work produces a junction profile that is at least twice as wide as Williams et al [87]. Thus, this effect was attributed to spatially delocalizing the Landau levels, which only ever allowed the lowest Landau level to mix in their sample.

There have been multiple studies which presented modeling results for the work by Williams et al. The work of Tworzydlo et al demonstrated the effect of valley-isospin and intervalley scattering [88] which is important to attaining the conductance plateaus given by eq. (5.2). Both Long et al [89] and Li et al [90] numerically modeled Landau level mixing at graphene $p-n$ junctions by including randomized site disorder at the junction. Low et al [91] added an alternative approach to modeling this effect, introducing a model which included roughness at the $p-n$ interface, edge roughness, and strong localized scattering centers.

From all these theoretical studies, one detail is clear- in order for modeling techniques
Figure 5-1: Experimental studies on QHE in graphene $p$-$n$ junctions: (A) Schematic of the combination bottom and top gate scheme used by Williams et al to first probe the QHE in graphene $p$-$n$ junctions. (B) QHE measurement performed by Williams et al resulting in QH plateaus. In their measurement, the Landau levels fully mixed, giving conductance values described by Eq. (Insert equation here!!!). (C) Buried split-gate structure used by Klimov et al (top panel) and an optical image of the finished graphene Hall bar (bottom panel) with the buried split gates visible. D Experimental result from Klimov et al where for the upper left quadrant of the $V_{G1}/V_{G2}$ resistance map Landau level mixing is suppressed, giving rise to a resistance of $e^2/h$. Panels (A) and (B) are reproduced from Williams et al [20], copyright AAAS 2007. Panels (C) and (D) are reproduced from Klimov et al [87], copyright American Physical Society, 2014.
to accurately reproduce the predictions of eq. \(5.2\), some extrinsic effect must be included in a simulation. These effects are typically some sort of disorder, such as roughness or site disorder, however, recently electron dephasing effects have been shown to be a particularly important mechanism \[92, 93\] which can be added by including Büttiker \[94\] probes along the junction interface.

In the rest of this chapter, we will see how quantum transport modeling may be used to understand the importance of the \(p-n\) junction width on carrier transport in the quantum Hall regime. First, we will look at the extrinsic effects which were added to the tight-binding transport model to accurately describe the work of Williams \textit{et al} and Klimov \textit{et al}. From there, visualizations of transport across the junction, which clearly explain the characteristics of the junction, will be presented.

### 5.2 Quantum transport model

From the work of others, it is evident that the standard tight-binding model for graphene is not sufficient for describing transport of Landau levels in a graphene \(p-n\) junction. In order to accurately describe the junction conductance, some extrinsic effect must be included on top of the tight-binding model, such as roughness, site disorder, or dephasing effects. In this section, we will go over the extrinsic effects which were included in our study. In short, our model consists of a broad \(p-n\) junction with a rough profile and on-site energy disorder in the junction transition region.

One of the first effects which we considered was the smoothness of the \(p-n\) junction. While Low \textit{et al} did consider a junction with a non-zero junction width \[91\], the junction was still quite abrupt with \(D_W = 25\) nm, at most. While such a width is sufficient for describing transport in junctions which display full equilibration of the Landau levels, it is not enough for the broad junction described in the work of Klimov \textit{et al} \[87\].

In order to include a broad \(p-n\) junction into the graphene tight-binding model, we must modify the on-site energy in the tight-binding Hamiltonian. This allows us to model the effect due to electrostatic gates, in this case for a device geometry as is shown in Fig. \[5-2\]. We chose to include the junction width in the most simple manner possible, by assuming a
linear transition between the heavily gated ends of the junction. In our graded \(p-n\) junction model, for a sample oriented with the transport direction parallel to the x-axis, the on-site energy is given by a piece-wise function:

\[
\begin{cases}
E_1, & x \leq I(y) - D_W/2 \\
E_2 - \frac{E_1}{D_W} x + \delta_i, & I(y) - D_W/2 \leq x \leq I(y) + D_W/2 \\
E_2, & x \geq I(y) + D_W/2
\end{cases}
\]

In our simulations it is assumed that the junction is centered around \(x = 0\) nm and has a transition width \(D_W\). \(E_{1,2}\) are the on-site energies on the left and right sides of the junction and represent the effect due to capacitively coupled gates as in Fig. 5-2. When \(E_{1,2}\) is positive, the area is \(p\)-type and when it is negative, the area is \(n\)-type. The term \(\delta_i\) is a delta function disorder potential and \(I(y)\) is a function which describes the roughness of the junction interface, both of which will be described further, later.

We use the same model presented by Low et al. [91] to describe the roughness of the \(p-n\) interface. The model, unchanged, will be repeated here for clarity. \(I(y)\) is represented as a Fourier series,

\[
I(y) = \sum_{n} A_n \sin \left( \frac{n \pi y}{W} \right),
\]

whose \(n^{th}\) Fourier amplitude is:

\[
A_n = R(D_1) e^{-\frac{n}{D_2}}.
\]

The function \(R(D_1)\) gives a uniformly distributed random number around \(\pm D_1\). \(D_{1,2}\) and \(N\) (an integer) determine the general shape of the roughness profile. Here, \(D_{1,2} = 13\) and \(N = 30\), which generally gives an RMS roughness of about 12 nm.

In our simulations, the roughness of the junction serves to enhance scattering, or mixing, between different Landau levels. Since the Landau level LDOS will track along the rough path of the junction, sharp kinks in the junction roughness can bring two different states together, enhancing tunneling between them. However, it should be noted that when \(D_W\) is much larger than the RMS roughness of the junction, this effect is minimized.

While junction roughness is an important effect for simulating Landau level mixing, alone,
it is not enough to contact experiment. It is a good start for our model and represents a physically relevant quantity (i.e. line edge roughness of metallic gates), but it is necessary to go further. Here we depart from the model of Low et al. [91] and instead add in aspects of the disorder potential used by Long et al. [89] and Li et al. [90].

In our model, an on-site energy disorder term, $\delta_i$, is included in the Hamiltonian. This disorder is only included in the transition region between $n$- and $p$-type graphene. This restriction on the area of disorder is included for multiple reasons. Practically, random site disorder can very easily transition the channel of the system into an Anderson insulator, via a phenomenon known as Anderson localization. The level of perfection in tight-binding simulations makes them particularly vulnerable to this effect. Thankfully, by restricting the area of the disorder, this effect can be easily avoided.

The second reason for placing disorder just in the junction transition region has particular physical significance. It is well known that graphene is most sensitive to disorder when the Fermi energy is near the Dirac point. This disorder could come from surface roughness, trapped charges in the substrate, or adsorbates on the surface of graphene. Since both devices modeled in this work were made directly on SiO$_2$, these effects are of high significance for our model to contact experiment.

The disorder included in the model is randomly generated using a Gaussian distribution function centered at 0 eV with a standard deviation of ±0.15 eV. Most importantly, this disorder is sufficient to induce intervalley scattering at the junction, which allows Landau levels of different polarity to mix [88]. In our model, most of the disorder is significantly smaller than what was used by Li et al. [90]. This is important because if the disorder energy is comparable to $E_{1/2}$, it will erase the effect of $D_W$ on the junction conductance. In our model, the added potential will help mix only Landau levels which are in close proximity to one another, not those which are spatially separated.

5.3 Results and Discussion

Now that the important aspects of the model have been laid out, we can begin to probe the transport mechanisms of the system. No discussion of theoretical modeling would be
Figure 5-2: Schematic of quantum Hall transport in a graphene p-n junction

This figure depicts a three-dimensional schematic of the graphene sample simulated throughout this chapter. The sample modeled consists of a graphene ribbon which is placed upon buried split-gates, similar to the style of the sample measured by Klimov et al. Buried split-gates are depicted, which are used to modulate the carrier density independently on each side of the sample. When subjected to a large magnetic field, a graphene p-n junction will exhibit snake-state-like current flow, as is shown with the gold arrow in the figure. This figure is reproduced from LaGasse et al., copyright American Physical Society, 2016.
complete without comparison with experiment, so we will begin by benchmarking our model against the experimental results of Williams et al\cite{20} and Klimov et al\cite{87}. From there we will break down the simulated conductance by performing maps of Landau level transport in graded graphene $p$-$n$ junctions. Having built our intuition from transport visualizations, we will develop a some guidelines for determining Landau level mixing in the $p$-$n$ junction system. Finally, we will investigate the fine characteristic of Landau levels in graphene and comment on the potential experimental implications.

### 5.3.1 Benchmarking against experiment

It is critically important to benchmark theoretical modeling against experimental results. This is particularly important when using tools such as tight-binding transport simulation, where non-physical parameters are deceptively easy to program! For this reason, benchmark results will be presented prior to any other discussion.

In a typical $p$-$n$ junction experiment in graphene, two gates will be used to independently control the carrier density of each side of the junction. An example device similar to that studied by Klimov et al is shown in Fig. 5-1. Generally, the two gates are varied independently whilst measuring the device conductance, generating a four-quadrant map consisting of $p$-$n$, $n$-$p$, $p$-$p'$, and $n$-$n'$ configurations. Since unipolar configurations are simply captured by tight-binding transport without any special model, and ambipolar transport is what makes graphene special compared to typical 2DEGs, we will focus on the ambipolar configurations.

Fig. 5-3 shows simulated conductance for a 200 nm wide graphene ribbon in a perpendicular magnetic field of 4 Tesla. As the important experiments are performed at extremely low temperatures, below 4.2 K, we make the assumption of zero temperature in our calculations. Additionally, quantum Hall measurements are typically performed at either extremely small DC bias, or with effectively no bias (using lock-in amplifiers with excitation voltages less than $k_B T$). As a result, to capture experimental measurements, it is sufficient to simulate transport just as the Fermi energy.

In these simulations, a 200 nm wide ribbon is sufficient for capturing the behavior of the experimental measurements. Importantly, this ribbon width completely spatially separates
the Landau levels on opposite sides of the sample, giving rise to well defined quantum Hall plateaus. This is, of course, linked with the magnetic field applied, which confines the Landau levels. A magnetic field of 4 Tesla is sufficient to give good separation of the Landau levels on each side of the ribbon and is similar to what is typically used in quantum Hall measurements by experimentalists.

The full model, as described in the previous section, was used in these simulations. Rather than simulating the entire four-quadrant map, slices are taken through the ambipolar region at key points. These are the key slices to benchmarking the experimental data. Despite reducing the simulations down to just a few slices, the averaging procedure used means that 400 simulations needed to be generated for every single point in Fig. 5-3. As a result, each slice took about one day to simulate.

Fig. 5-3a shows simulations corresponding to a diagonal slice through one ambipolar quadrant, where the opposite sides of the junction have equal and opposite carrier densities. Fig. 5-3b and Fig. 5-3c show horizontal slices through the sample with fixed filling factors (on the left side of the simulation) $\nu_1 = 2$ and $\nu_1 = 6$, respectively. Each simulation is also repeated for different junction widths, of $D_W = 30, 100, 125$ nm.

In Fig. 5-3, the simulations for $D_W = 30$ nm match with eq. (5.2), displaying the predicted conductance plateaus each time one of the junction filling factors jumps from 2 to 6 to 10. This behavior is congruent with what was observed in Ref. [20], shown in Fig. 5-1b, corresponding to full mixing of the Landau levels at the junction. This simulation result is particularly important, as it shows that the model we have described so far is able to capture the abrupt junction characteristics.

As the junction width is increased from 30 nm to 100 nm and then 125 nm, the simulated conductance for all energies (or filling factors) begins to converge towards a plateau of $1 e^2/h$. This indicates that increasing the junction width gradually suppresses the mixing of all the Landau levels except for the zeroth level. This result matches with what was observed experimentally in Ref. [87], where the measurements were performed using split-gates producing junction widths greater than 100 nm.
Figure 5-3: Quantum Hall simulation benchmarking
This figure shows simulation results benchmarking our p-n junction model for different junction configurations at varying junction widths. All of the simulations were performed for a simulation cell of $L = 320$ nm and $W = 200$ nm with $B = 4$ T. The points of each curve are the result of an ensemble average over different disorder configurations, as discussed in the main text. Panel (A) shows the results for a symmetric n-p junction and the other panels show results for fixed values of $E_1$ at (B) $E_1 = -0.05$ eV and (C) $E_1 = -0.09$ eV as a function of $E_2$. Similar to the experimental works we are modeling, we see that for small junction widths, the Landau levels mix, giving rise to plateaus described by eq. (5.2). When the junction width is large, the higher order Landau levels become spatially separated, which inhibits mixing of all but the lowest Landau level. This figure is reproduced from LaGasse et al [3], copyright American Physical Society, 2016.
5.3.2 Mapping Landau level transport

Now that we have demonstrated good agreement between our quantum transport model and experiment, we will work to understand the results more closely. Clearly, the width of the $p$-$n$ junction transition region is a key parameter in determining the degree of Landau level mixing. To further understand this phenomenon, we will investigate the LDOS of Landau levels transporting at the $p$-$n$ junction for varying junction widths.

In Fig. 5-4 we show two different simulations of the LDOS at the Fermi energy injected from the left contact of the simulated graphene device for two different junction widths (30 nm and 125 nm). The on-site energy on each side of the sample is fixed, $E_1 = -E_2 = 0.0875\text{ eV}$, such that with an applied magnetic field of 4 T, the filling factor in the sample is $\nu_1 = -\nu_2 = 6$. This means that there are two Landau levels which transport at the junction, which we call the 0$^{\text{th}}$ and 1$^{\text{st}}$ Landau levels.

It is important to note that we have made the decision to plot just the Landau levels which are sourced from the left contact in these maps. Carriers which are injected from the opposite contact have a nearly identical behavior, except that they counter-propogate with the right-moving carriers, due to the magnetic field. In this symmetric doping configuration, the Landau levels sourced from the right side of the channel produce a transport map which is a mirror image of what is shown in Fig. 5-4. When the total LDOS is shown in the channel, it becomes quite difficult to understand the maps, thus we prefer to just look at one contact at a time.

The Landau levels from the left contact enter the sample on the lower left of the device and trace along the bottom edge until they hit the $p$-$n$ junction transition. In Fig. 5-4a, for the 30 nm junction width, the junction transition is quite sharp. In this case, both Landau levels track up the junction at nearly the same position. The rough junction does produce some perturbation in the Landau levels, which may be seen on the right side of the plot around $y = 100$ nm, where the 0$^{\text{th}}$ level skips over the tight-kink in the junction.

When the junction width is increased to 125 nm (Fig. 5-4b), the Landau levels completely separate at the junction. This is a direct result of eq. (5.1), which predicts how each Landau level will flow in an external electrostatic potential. In the nearly-abrupt case,
the spacing between Landau levels is almost non-existent. However, the broad junction completely spatially separates the 0\textsuperscript{th} and 1\textsuperscript{st} levels.

The spatial separation of the Landau levels at the junction is the most important factor in determining Landau level mixing. In the nearly-abrupt case, where all the Landau levels are essentially co-located at the junction, there is ample chance for carriers in one level to transmit into any other level. Indeed, as evidenced by agreement with eq. (5.2), there must be equal probability of transmission between each Landau level.

As the 0\textsuperscript{th} and 1\textsuperscript{st} Landau levels are spatially separated, the transmission probability between them will rapidly fall off. In order for the 1\textsuperscript{st} Landau level to transmit to the other side of the device, giving rise to an increase in conductance, it must transmit into a right moving state at the junction. This means it must transmit into the 0\textsuperscript{th} Landau level or one of the higher order Landau levels entering from the opposite side of the device. In the graded junction, this effect is not possible, resulting in the Landau level turning around at the junction and exiting the device through the same contact.

The 0\textsuperscript{th} Landau levels coming from each side of the sample are particularly interesting, as at the junction they are always located exactly on top of one another. This produces an extremely strong mixing of just the 0\textsuperscript{th} Landau levels which is fundamentally linked to the band structure of graphene in the quantum Hall regime. As such, the Hall conductance of graphene p-n junctions, assuming there is no splitting of the 0\textsuperscript{th} Landau level energy, will always be at least $1 e^2/h$ regardless of the size of $D_W$.

To understand this effect further, we will map transport across a graded junction for various filling factors. In Fig. 5-5, we show three simulations for the same junction with various filling factors (by changing the Fermi energy) on the left side of the device. Below each panel in Fig. 5-5 the on-site energy as a function of $x$-position is shown. Whenever $E_F - E_{LL} > E_{\text{on-site}(x)}$, where $E_{LL}$ is the energy of a particular Landau level, the Landau level begins to transport. The scatter points are given by eq. (5.1) and show good agreement with the turning point of each Landau level at the junction predicted by tight-binding simulations.

At this point, it is worthwhile to discuss the partial-Landau level mixing which is observed in Fig. 5-3. Eq. (5.1) predicts that increasing the doping of either side of the junction other will cause the Landau levels to move closer together. When the Landau levels move closer
Figure 5-4: Maps of quantum Hall transport in a disordered graphene $p$-$n$ junction at small and large junction widths. Maps of the local density of states injected by the left contact for the simulation configuration shown in Fig. 5-3. Each simulation is performed for a symmetric junction at $E_1 = -E_2 = 0.0875$ eV at $B = 4$ T and includes junction roughness (marked by dashed lines) and the delta disorder described in the main text is included between the uniform $n$- and $p$-regions. (A) is for a junction width of 30 nm, where the Landau levels are kept close together and allowed to mix. (B) is for a junction width of 125 nm, where the Landau levels are spatially separated and only the zeroth levels are able to mix at the junction. This figure is reproduced from LaGasse et al [3], copyright American Physical Society, 2016.

together, the carriers in each level will begin to tunnel into neighboring levels. However, this effect does not result in full mixing of the Landau levels, yielding conductance values smaller than what would be predicted for the filling factors on each side of the junction. This effect is most noticeable in the intermediate junction width of $D_W = 100$ nm.

5.3.3 Factors influencing Landau level mixing

To provide a more detailed picture of Landau level mixing we will look at the effect of $D_W$ and the applied magnetic field. We perform simulations (Fig. 5-6) for a fixed filling factor configuration of $\nu_1 = -\nu_2 = 6$ as a function of junction width for two different magnetic
Figure 5-5: Mapping Landau level transport at different Fermi energies The top panels depict Landau level transport for a fixed energetic landscape with $D_W = 100$ nm and $B = 10$ T, for increasing Fermi energy. The local density of states is plotted for carriers injected from the left contact for (A) $\nu_n = 2$, (B) $\nu_n = 6$, and (C) $\nu_n = 10$. The lower panels show the energy band diagram of the sample, with scatter points marking the predicted turning point of each Landau level according to eq. (5.1). This figure is reproduced from LaGasse et al [3], copyright American Physical Society, 2016.
fields (4 and 14 Tesla). When the junction is abrupt, $D_W \leq 30$ nm, full mixing of the Landau levels is observed, resulting in a conductance plateau of $3 e^2/h$ (full mixing). When the junction width is increased, $D_W > 100$ nm the conductance plateau begins to converge to $1 e^2/h$ (only the lowest Landau level mixes).

The magnetic field also plays an important role in dictating the degree of Landau level mixing in the simulation. We already know that by increasing the junction width it is possible to spatially separate the Landau levels at the junction. Another way to control this separation is the magnetic field. The stronger the magnetic field, the more highly confined each Landau level will be. This increased confinement will reduce the spatial extent of each Landau level (being roughly proportional to the magnetic length, $l_B = \sqrt{c/eB}$ [11]) and cause them to more accurately track the profile of the junction. For very abrupt or very graded junctions, the effect of the magnetic field is not noticeable. However, in the partial mixing regime, stronger magnetic fields will generally reduce the degree of Landau level mixing.

Other parameters which influence the degree of Landau level mixing in our model are the junction roughness and the site disorder at the junction. These values are both difficult to control and measure in experiments, however, we can tell from our simulations that each play an important role in controlling the degree of Landau level mixing. Generally, increasing either of these effects will also increase the degree of Landau level mixing.

Disorder is particularly important parameter and in the extreme case will result in Landau level mixing regardless of all the other model parameters we have included. While interesting to note, in this work we have strived to keep these parameters as physically relevant to the works of Ref. [20] and [87] as possible, and thus their overall contributions are fairly low. Therefore, in the absence of extremely large disorder potentials or junction roughness, junction width will have the strongest influence on the degree of Landau level mixing.

### 5.3.4 Landau level shape

Carefully examining Fig. 5-5 reveals that higher order Landau levels have an interesting substructure. Rather than forming just a simple uniform beam, the higher order Landau levels clearly have some more complex shape. Indeed, this may even be seen in Fig. 5-4.
Figure 5-6: Junction width dependence on the quantum Hall conductance of a graphene p-n junction. Simulated Hall conductance for a symmetric p-n junction sample configured such that \( E_1 = -E_2 = 0.0875 \text{ eV}, \) or \( \nu_n = -\nu_p = 6. \) The junction width is varied between 10 and 130 nm and the conductance is calculated at 4 and 14 T. For \( D_W < 30 \text{ nm}, \) the junction is able to equilibrate all of the Landau levels, leading to the conductance predicted by eq. 5.2. For \( D_W > 30 \text{ nm}, \) the spatial separation of the Landau higher order Landau levels begins to suppress Landau level mixing. At \( D_W > 100 \text{ nm}, \) for \( B = 4 \text{ T}, \) the higher order Landau level mixing is mostly suppressed, and for \( B = 14 \text{ T}, \) the higher order Landau level mixing is completely suppressed. Generally, our simulations tell us that moving to higher junction widths will suppress mixing of higher order Landau levels and that increasing the magnetic field (which further constrains the Landau levels) will enhance this effect. This figure is reproduced from LaGasse et al. 3, copyright American Physical Society, 2016.
where the 1\textsuperscript{st} Landau level has a small dip in the middle, however the disorder present makes the effect more subtle. This effect is not noticeable when the Landau levels are all confined to the edge of the sample, as in standard quantum Hall measurements. However, in a graphene p-n junction with a large junction width, the Landau levels may each be spatially separated at the junction, presenting an interesting opportunity to study their structure.

Fig. 5-7\textsuperscript{a} shows simulated LDOS results (similar to previous plots) for an extremely graded junction with $D_W = 500 \text{ nm}$. The applied magnetic field is 4 Tesla and $E_1 = -E_2 = 0.125 \text{ eV}$, which results in a filling factor of $\nu_1 = -\nu_2 = 14$. Whilst such a graded junction would be difficult to achieve in the laboratory, it is trivial to include in our simulations, and the results make for an interesting thought experiment which should motivate future experiments.

In Ref. [95], Lukose \textit{et al} reached an analytical solution for the wave function of a graphene sheet subjected to crossed electric and magnetic fields. The result of Lukose \textit{et al} assumed a uniform electric field (an infinite system) which is essentially what we have when the junction width is very large in our simulation. The wave function along the junction (adopted directly from Ref. [95]), away from the edges of the graphene ribbon, is given by

$$
\Psi_{n,k_y}(x,y) \propto e^{ik_y y} e^{-\frac{\varphi|n|-1}{2}} \left[ \frac{\text{sgn}(n)\phi_{|n|-1}(\xi)}{\phi_{|n|}(\xi)} \right],
$$

with

$$
\xi = \frac{(1 - \beta^2)^{\frac{1}{4}}}{l_b} \left( x + l_b^2 k_y + \text{sgn}(n) \sqrt{\frac{2 l_b}{n(1 - \beta^2)^{\frac{1}{4}}}} \right),
$$

$$
\beta = \frac{E}{\nu_F B}, \text{ with } \mathcal{E} \text{ being the applied electric field.}
$$

$\phi_{|n|}(\xi)$ are quantum harmonic oscillator wave functions, with one graphene sub-lattice contributing the $n$\textsuperscript{th} harmonic oscillator function and the other sub-lattice giving $n - 1$ term, where $n$ is the Landau level index. Interestingly, the 0\textsuperscript{th} Landau level only receives a contribution from one sub-lattice.

Fig. 5-7\textsuperscript{b} shows a slice taken from the simulated map in Fig. 5-7\textsuperscript{a} at $y = 100 \text{ nm}$. To compare with the result from Ref. [95], we also plot $P(x) = M_n \Psi_{n,k_y}(x) \Psi_{n,k_y}^*(x)$, where $M_n$ is a normalization function which accounts for the number of modes in the calculation. Since
our simulation does not consider an infinite system, we define the electric field as $E = \frac{E_2 - E_1}{eD_W}$.

In Fig. 5-7b we note good agreement between our simulations and the analytical solution for the 0th, 1st, and 2nd Landau levels. The 3rd Landau level transmits very close to the left edge of the junction where the electric field becomes zero, which results in some skewing of the simulated level, however it is clear that the fundamental physics which govern both results are the same.

5.3.5 Chapter conclusions

In this chapter we have gone through a detailed study of quantum Hall effect in heavily graded graphene $p$-$n$ junctions. We developed a tight-binding quantum transport model which included multiple extrinsic effects (smooth junction transitions, junction roughness, and averaging over site-disorder at the junction) which enabled us to capture results from two key experimental results in the literature on abrupt and graded junctions [20][87].

Our model demonstrated the transition between full-mixing of Landau levels (in the abrupt case) and mixing of only the lowest Landau level (in the graded case). After benchmarking against experiment, transport visualizations revealed the primary contribution to determining Landau level mixing is the junction width, $D_W$. We also gave a simple model for determining the position of Landau levels across a $p$-$n$ junction, which was verified with tight-binding calculations. Finally, we compared our tight-binding calculations with an analytical calculation for the Landau level probability density from the literature [95], revealing an interesting spinor structure which the graded junction lets us study.

The results of this chapter are of broad interest to both experimentalists and theorists working in the area of quantum Hall effect in graphene. Our model provides a basis for studying Landau level transport in more complex graphene systems including $p$-$n$ junctions. Our model is particularly relevant to comparing with experimental results because it predominantly relies on experiment-relevant parameters, such as $D_W$. One major application of the model would be in the study of multi-quantum Hall resistance standards, which have been proposed in graphene devices consisting of many back-to-back $p$-$n$ junctions [96].
Figure 5-7: Comparison of individual Landau levels separated at an extremely graded p-n junction with analytical expression derived by Lukose et al. (A) Local density of states injected from the left contact for a simulation with $B = 4$ T and $E_1 = -E_2 = 0.125$ eV ($\nu_n = -\nu_p = 14$). The junction width is 500 nm, which completely spatially separates the Landau levels at the junction, revealing the interesting sub-structure in each Landau level, discussed in the main text. (B) Comparison of the expression derived by Lukose et al. [95], eq. (5.5), with a slice taken from (A) at $y = 100$ nm. This figure is reproduced from LaGasse et al. [3], copyright American Physical Society, 2016.
Chapter 6

Ballistic electron focusing in graphene

*p-n* junctions

Transverse electron focusing (TEF) experiments have been used to study electron transport for many years, first starting with metals [97]. Eventually, the invention of quantum point contacts enabled the study of TEF in 2DEGs, using QPCs as both sources and probes of ballistic electrons [98]. Just as in the previous chapter discussing quantum Hall effect, graphene presents a fantastic platform for studying ballistic electron focusing. Not only does graphene display long mean free paths, a requirement for TEF, it is the first 2DEG platform in which a *p-n* junction can be formed laterally across the 2DEG. This special capability of graphene enabled a recent experiment by Chen *et al*, in which a graphene *p-n* junction was used to turn typical TEF experiments on their heads [21]. In this chapter, we will look at a detailed study in which tight-binding transport methods are used to build enhanced understanding of TEF in graphene *p-n* junctions.

6.1 Introduction

TEF experiments utilize cyclotron motion of electrons (or holes) induced by a magnetic field to probe ballistic carrier transport. In a phase-coherent sample, in a small magnetic field, carriers will move in semi-classical skipping orbits. This effect was first observed in metals [97], but gained major popularity when studied in 2DEGs [98].
TEF experiments in 2DEGs were first performed by van Houten et al using QPCs as both electron sources and as voltage probes \[98\] (Fig. 6-1). In their GaAs-based 2DEGs, ballistic electron transport was readily studied. This experiment is particularly exciting as it motivated a slew of electron-optics experiments in 2DEGs which built upon the techniques developed in the paper.

Simply, electrons are injected via one QPC and an applied magnetic field, via the Lorentz force, causes the electrons to exhibit cyclotron motion with a diameter \(D_c = \frac{\hbar k_f}{eB}\) \[98\]. The cyclotron diameter is inversely proportional to the applied magnetic field, which is a continuously variable experimental parameter. By measuring the voltage of the neighboring QPC, it is possible to detect whenever electrons are focused into the QPC. This occurs whenever the inter-QPC distance is an integer multiple of the cyclotron diameter, described by the resonance condition \(D_{QPC} = nD_c\).

When the QPCs are used in a non-local resistance measurement, a peak in the resistance may be observed whenever this criterion is met (Fig 6-1c), enabling spectroscopy of these ballistic electrons. TEF experiments require nearly-ballistic electron transport, which is enabled by extremely clean samples. Electrostatic disorder throughout a sample will act as scattering sites, re-directing the orbits of electrons and diluting any measured resonances.

As the sample quality of graphene-based devices has improved, TEF measurements have emerged as an extremely powerful technique for probing ballistic transport. The ability to continuously tune the carrier density of graphene has led to interesting measurements \[99\] and scanning gate microscopy has been used to image the cyclotron orbits \[100\]. In one very interesting experiment, deformed cyclotron orbits due to mini-band formation between graphene and \(h\)-BN led to new TEF resonances \[101\]. Additionally, absorptive pinhole collimators have been developed as an alternative to the early experiments which used QPC-based probes \[102\].

In a light magnetic field, carriers which transport across a \(p-n\) junction in graphene will exhibit a special type of “snake-state” motion, where the carriers arc back and forth across the \(p-n\) junction, following the cyclotron radius. While similar to what we saw in the previous chapter, in this case the electrons follow cyclotron orbits, rather than propagating as well-defined Landau levels. In graphene, this phenomena has been studied both by experiment.
Figure 6-1: Early transverse electron focusing experiment in GaAs 2DEGs [98]

(A) Schematic of the experimental setup used by van Houten et al. QPCs are used as extremely sensitive sources and probes for carriers. (B) SEM image of split-gates used to define the two QPCs used in the measurement. (C) Example measurement from van Houten et al showing TEF resonances whenever electrons are focused from one QPC to the other, corresponding to cyclotron diameters in fractions of the inter-QPC distance. All of these images are reproduced from Ref. [98], copyright *American Physical Society*, 1989.
Graphene’s ability to support ambipolar transport enables a new type of TEF experiment which was not possible in 2DEGs formed in standard heterostructures. Typical TEF experiments have been performed only on homogeneous, unipolar channels. Chen et al, however, recently studied TEF in a graphene sample (Fig. 6-2) with a tunable p-n junction inserted in the path of the focused electrons. In contrary to previous TEF experiments, where charged carriers were restricted to skipping along one of the edges of the sample, inserting a p-n junction let Chen et al redirect the beam of carriers across their sample [21].

The experiment of Chen et al [21] probes a special case of snake-state motion where the device geometry was carefully designed to enable the to study of angle-dependent transmission across the p-n junction. In the experiment, a Hall bar geometry was used (Fig. 6-2a) where the distance, $D_C$ between neighboring Hall probes was matched exactly to the distance of the Hall probes on the opposite sides of the sample. In this special geometry, by tuning either the magnetic field or the gates (which define the type of junction), it is possible to steer carriers across the junction.

This chapter will be dedicated to a quantum transport study [4] aimed to further understand the effect studied by Chen et al [21]. In the original work by Chen et al, a semi-classical ray tracing model was used to interpret the data. Here, we will see how quantum transport methods based on tight-binding theory, whilst requiring some special care, can achieve superb agreement with experiment. Detailed transport visualizations will shed light on the nuances of the TEF problem in graphene p-n junctions.

### 6.2 Transport model

In this chapter, we will use the same basic tight-binding model for graphene which we have used before. Once again, the key to capturing experiment is in how we apply the model, which requires a deep consideration of the physical details of the experiment. The tight-binding model intrinsically captures most of the fundamental physics of graphene, such as
Figure 6-2: Device for ballistic electron focusing by Chen et al. (A) Optical image of the device measured by Chen et al. (B) Schematic of the device of Chen et al depicting h-BN encapsulated graphene assembled on a graphite split-gate which was used to form reconfigurable p-n junctions in the graphene channel. Both of these images are from Ref. [21], copyright AAAS, 2016.
angle-dependent transport across the graphene $p$-$n$ junction. However, to accurately model the experiment, we will need to look deeply into the nature of phase-coherent transport in graphene.

### 6.2.1 Experimental details

Before discussing the details of our model, it is necessary to set the stage by describing a few of the experimental details and the sample geometry. At its heart, the device measured by Chen et al is a graphene Hall bar in which a tunable $p$-$n$ junction is formed using a double bottom gate scheme. The measurements which we seek to model were performed at extremely low temperatures, down to 1.7 K, and used low-bias lock-in techniques. Therefore, all of our discussions will assume zero temperature and zero bias, which simplifies the calculations significantly.

If we consider carrier injection in such a Hall bar from one Hall probe, as is shown in Fig. 6-3a, in a small magnetic field, the carrier will arc across the sample. In a uniformly gated sample, for example if the entire graphene is $p$-type, the magnetic field can be tuned such that the carriers simply arc from the source hall probe into the neighboring Hall probe. If the gates are configured to form a $p$-$n$ junction, the carriers will instead be redirected into the Hall probe on the opposite side.

The major question is this: how should one set up their measurement to electrically detect these TEF trajectories? The answer is fairly nuanced. A simple two-point electrical measurement will not be sufficient to capture all of the possible configurations enabled by a dual-gated $p$-$n$ junction. Instead, in such a Hall bar configuration, it is possible to perform what is known as a “non-local resistance” measurement. In such a measurement, a current is driven between two contacts and then a voltage is measured between two others.

In this case, as is shown in Fig. 6-3b, a current is driven between the two opposite-facing Hall probes (labeled 1 and 5) on the same side of the $p$-$n$ junction. At the same time, the voltage picked up by the opposite-facing Hall probes (labeled 2 and 4) on the opposite side of the junction is measured. In this measurement, we can define the non-local resistance in
the sample as
\[ R_{15,24} = \frac{V_2 - V_4}{I_1}. \] (6.1)

As such, the measured voltages \( V_2 \) and \( V_4 \) will let us detect whenever carriers are focused from contact one into contact two or four.

### 6.2.2 Phase coherence

One of the critical assumptions used in quantum transport modeling is that reservoirs of charged carriers, electrical contacts, are sources of phase-coherent carriers. In typical samples, there are scattering events present in the channel which dilute the phase of these injected waves, randomizing their velocities and putting them out of phase with the source they originated from. However, in modern graphene samples the phase coherence length, especially when measured at low temperatures, tends to be comparable to the sample size or larger. In this case, special care must be taken to handle dephasing in our model.

There are also two red probes (labeled 0 and 3) shown in our schematic in Fig. 6-3b, these probes are included as *dephasing probes*, which are critical to the success of our model. Similar sets of probes were included on the sides of the Hall bar studied by Chen et al [21], however it is not clear whether or not they were included purposefully or for un-published experiments. Nevertheless, they likely fulfilled an extremely important role in the experiment by acting to dephase any stray carriers which do not resonantly focus into any of the voltage probes.

One of the basic properties of these phase coherent waves of carriers is that they may only interfere with waves they are phase coherent with. Carriers are injected from the contact and into the channel with a particular energy and will scatter within the channel until they transmit through any of the attached contacts. Thus, in the absence of any coherence-breaking mechanisms in the channel, a wave injected into a the ballistic channel may reverberate for a significant amount of time before reaching steady state.

This has significant implications on the conduction through the channel, as at steady state the wave will have produced chaotic interference patterns in the channel. This is particularly evident when studying TEF in p-n junction systems. For example, if we simulate a graphene
Figure 6-3: Schematics for transverse electron focusing modeling (A) 3D model of the graphene Hall bar used to study the TEF problem. The red arrows depict the motion of the first TEF resonance for a p-p and p-n junction. (B) Schematic of the Hall bar simulated in this chapter, denoting the contact separation, $D_C$, the junction width, $D_W$, and the Hall probe width, $L_C$. Importantly, $D_C = W$, which enables a symmetric probing of the unipolar and ambipolar characteristics of the sample. Yellow contacts denote the Hall probes and salmon-colored contacts denote the dephasing probes which are described in the main text. (C) Energy band diagram depicting the conventions for forming p-n junctions in this chapter. This figure is reproduced from Ref. [4], copyright American Physical Society, 2017.
channel with four Hall probes without adding extra dephasing, and study the current density injected from one contact, we yield the chaotic interference pattern shown in Fig. 6-4b. This is a result of the injected wave scattering within the channel and interfering with itself. This process continues for some time (which is intrinsically included in energy-domain tight-binding calculations) until steady state is reached, by which any interesting focusing effects have been destroyed.

However, by adding extra side-contacts to the Hall bar, which act as dephasing probes, the same current density leads to a striking result in Fig. 6-4a. Even though only a small portion of the injected wave does not focus directly into contact five, that is sufficient (without dephasing) to disrupt the transport in the channel. The particular parameters for this simulation setup will be explained later, for now we note, as have others for standard TEF simulations [115, 116, 117], that dephasing probes are critical to capturing the experimentally observed results.

### 6.3 Landauer-Büttiker methods

One might assume that the added dephasing probes in depicted in Fig. 6-3b and demonstrated in Fig. 6-4a are absorbing or current sinking contacts. However, this is not the case. By attaching a lead to a channel and not biasing it, it is allowed to float. When a contact is floating, it must carry no net current. That is, when a current does flow into the contact, the voltage of the contact will float up such that the same level of current is re-injected into the channel (leading to no net current flow at the specified contact). Importantly, the re-injected current comes back into the channel with a different phase! In this way, the chaotic interference in Fig. 6-4b may be avoided.

This process is critically important to understanding the results of Chen et al and is governed by the Landauer-Büttiker equation [11]:

\[
I_p = \frac{2e^2}{h} \sum_q [T_{qp}V_p - T_{pq}V_q].
\]  

It relates the steady state current of a lead \( p \) to the transmission from lead \( p \) into and out of,
Figure 6-4: Comparison of particle current density with and without dephasing edge contacts Local particle current density injected from contact one both (A) with dephasing edge contacts and (B) without. Both simulations are configured as symmetric p-n junctions, the same as will be further studied in detail in Fig. 6-8a. Without the dephasing contacts, the wave injected from contact one scatters repeatedly throughout the channel, interfering with itself and destroying any signature of the TEF resonance. The scale bars are 60 nm. The magnetic field in each panel is shown pointing out of the page. This figure is reproduced from Ref. [4], copyright American Physical Society, 2017.
T_{qp/pq}, and the voltages of, V_q, every other lead q. The Landauer-Büttiker equation describes the balancing act which occurs in a multi-terminal system to determine the steady state current in each lead.

Eq. (6.2) defines a system of linear equations relating the current and voltage of each lead to the transmissions between them,

\[ \mathbf{I} = \frac{2e^2}{h} \mathbf{T} \mathbf{V}. \] (6.3)

Here, \( \mathbf{I} \) and \( \mathbf{V} \) are column vectors containing the different lead currents and voltages and \( \mathbf{T} \) is a matrix of transmission functions.

The Landauer-Büttiker equation tells us the rules which govern multi-terminal transport in a ballistic device, however it is up to the user to set it up properly. Again, we must consider the exact measurement which are attempting to model. Our goal is to calculate the non-local resistance, \( R_{15,24} \), defined in eq. (6.1). By remembering the experimental parameters used in attaining eq. (6.1), we can drastically simplify eq. (6.3).

In the measurement by Chen et al, a current is driven between contacts 1 and 5 and all the other contacts are left floating, i.e. they carry no current. This equates to saying that \( I_1 = -I_5 \) and \( I_0 = I_2 = I_3 = I_4 = 0 \). We also need to choose one contact to reference all the other voltages in the system to; it ends up being convenient to ground contact 1, \( V_1 = 0 \) V. This results in a series of simplifications in eq. (6.3) that let us solve for the \( V_1 \), \( V_4 \), \( I_1 \), and \( I_5 \), yielding

\[ R_{15,24} = \frac{h}{2e^2} \frac{1}{2} (R_{45} - R_{25}). \] (6.4)

\( R_{45} \) and \( R_{25} \) are elements of the \( R \)-matrix, which we define as \( \mathbf{R} = \mathbf{T}^{-1} \). Therefore, calculating the non-local resistance turns into a problem solely consisting of calculating the transmission functions between each contact.

The analytical solution for \( R_{45} \) and \( R_{25} \) in terms of transmission functions is quite enormous, spanning about forty pages of text when written out fully. Thankfully, interpreting the non-local resistance measurements and calculations does not require using the full equation. Instead, we will see how by looking at a few key transmission functions between certain contacts, it is possible to gain a great amount of physical insight into the TEF problem.
6.3.1 Model summary, final details

In studying the TEF problem in reconfigurable graphene p-n junctions, our standard first nearest neighbor tight-binding model intrinsically captures many of the important transport mechanisms in play. Aggressive improvements in sample quality have brought experimental devices significantly closer to the simple tight-binding model. As a result, the effects we include in our model have become much more subtle. The most important effect which we engineer into our model is additional dephasing contacts, as governed by the Landauer-Büttiker equation, described above.

All of the simulations shown in the rest of this chapter will be based on the schematic shown in Fig. 6-3b. We will simulate a six-terminal Hall bar, where the large contacts on the end serve as virtual dephasing contacts, as described in the previous section. The width, \( W \), of the Hall bar is 200 nm. Accordingly, the distance between the Hall probes, \( D_C \), is matched to the width, at 200 nm. These dimensions represent a scaled-down version of the experimental device measured by Chen et al.\[21\]. The major difference induced by shrinking the device size is the overall resistance values will be different and larger magnetic fields are used in the simulation, however the fundamental underlying physics between the experiment and physics are the same.

We adopt the same on-site energy scheme, accounting for the effect of electrostatic gating, as we did previously for studying quantum Hall effect in graphene p-n junctions. Repeated in Fig. 6-3b, for clarity, the left and right sides of the device have their on-site energy shifted by \( E_{1,2} \), respectively. The Fermi energy is always assumed to be at \( E = 0 \) eV. While the scheme for setting up the energies in the sample is arbitrary, we find that this provides the best platform for comparing to experimental results. Finally, a simple linearly graded junction with a junction width \( D_W = 50 \) nm is assumed, to account for the finite potential grade which is present in the measurements.

6.4 Results and Discussion

Now that we have finished laying out the important aspects of our model, we will apply it to the TEF problem in the presence of a reconfigurable p-n junction. As before, we will
Figure 6-5: Off-resonance characteristics of the symmetric graphene $p$-$n$ junction
This figure depicts the off-resonance transport for the graphene $p$-$n$ junction configured as $E_1 = -E_2 = 50$ meV. In (A), $B = 0.11$ T, where the primary beam injected from contact one across the junction hits to the right of the contact four, skipping out the right dephasing contact. In (B) the magnetic field is on the opposite side of the first TEF resonance, with $B = 0.16$ T, this time, causing the beam of carriers to hit to the left of contact four. This figure is reproduced from Ref. [4], copyright *American Physical Society*, 2017.
begin by benchmarking our results against experiment. Using our model for calculating the non-local resistance of the device, we will systematically approach the experimental results of Chen et al [21]. We will then bring new understanding to the problem by performing transport visualizations of local particle current density for different sample configurations. Utilizing the mode-picture of tight-binding transport, we will further break down the important results, revealing an interesting interplay between semi-classical billiards and quantum mechanics.

6.4.1 Benchmarking with experiment

Fig. 6-6 shows a comparison of the non-local resistance (6.1) calculated using our tight-binding transport model and the experimental results of Chen et al [21]. With the addition of our dephasing probes, we see a striking similarity between our simulation and experiment. Indeed, our simulations capture all of the major features observed in experiment, including all observed focusing resonances, for both positive and negative values of resistance. We also observe several subtle features, where the junction configuration goes through different doping configurations, which we will explain. In this section, we will lay out the important details for comparing these two results.

To compare with experiment, we convert the applied $E_1$ and $E_2$ values to the corresponding carrier densities using a simple geometric capacitance model. Carrier density has the same functional form as gate voltage, which gives us a way of qualitatively comparing the results. Note that we neglect any effect the magnetic field may have on the carrier density, which would only be a second order effect, given the low magnetic field values.

In Fig. 6-6a, We fix the carrier density of the left side of the junction, with $E_1 = 50$ meV (making it $p$-type). Then, the carrier density of the right side of the simulation and the applied magnetic are varied independently to generate a map of the resistance. This mimics the measurement done by Chen et al, where they applied a fixed graphite gate voltage and varied the back gate voltage and magnetic field.

This series of configurations results in a four-quadrant map, where for negative carrier densities (or negative gate voltages), we see a $p-p'$ junction configuration and for positive carrier densities (or positive gate voltages) we see a $p-n$ junction configurations. We note that the
transport is essentially symmetric for positive and negative magnetic field values. For positive magnetic field values it is convenient to focus on contact 1. For negative values, all of the same arguments may be repeated, except for contact 5. While this is a given for simulated values, the fact that it is also observed in experiment is a strong indication that the sample is of very high quality.

Additionally, as the carrier density moves from negative to positive there are several different regimes. Working from left to right in Fig. 6-6a, the system starts off as $p-p^+$, where the right side of the sample is more highly $p$-type than the left. This corresponds to a vanishing non-local resistance for the second negative resonance at around $B = \pm 0.4$ T. When the junction transitions to uniformly $p$-type, this second negative resonance has a resurgence, continuing across the $p-p^-$ configuration. A similar effect occurs for the $p-n^-$ to $p-n$ to $p-n^+$ transitions on the other side of the map, except this time we only see the second resonance in the $p-n^+$ configuration. This effect, which was not captured by a semi-classical billiard model, is clearly seen in both the experimental results and our quantum transport simulation.

Moving forward, we will break down the complexities of the maps in Fig. 6-6 by taking slices of the resistance and probing them with key current density visualizations.

### 6.4.2 Understanding the non-local resistance

We begin by extracting constant-carrier density slices out of Fig. 6-6. These slices, shown in Fig. 6-7, are representative to the results shown in Figure 4 of Ref. [21].

Fig. 6-7a shows the simulation configured as a $p-p'$ junction with the carrier density on the left side of the device fixed to $-0.18 \times 10^{12}$ cm$^{-2}$ and the right side of the device is fixed to $-0.41 \times 10^{12}$ cm$^{-2}$. In our simulation we observe four positive resistance TEF resonances and three negative resistance resonances. At the highest magnetic fields, beyond 0.6 T, the resistance flattens to zero, which corresponds with the onset of quantum Hall effect.

We have two important tools to analyze this data, transmission function calculations and current density mapping. Superimposed on the resistance, Fig. 6-7a also shows two relevant transmission functions, $T_{21}$ and $T_{31}$. Recall, by our definition of the transmission function, these two transmissions are read as transmission from contact one into contact two.
Figure 6-6: Benchmarking tight-binding transport model against results of Chen et al (A) Non-local resistance as a function of magnetic field for a fixed $p$-type ($E_1 = 50$ meV) graphene on the left side of the Hall bar and a varying carrier density on the right side. The scatter points correspond to the current density maps in Fig. 6-8. These results show fantastic qualitative agreement with (B) the results of Chen et al [21]. (A) is reproduced from Ref. [4], copyright American Physical Society, 2017. (B) is reproduced from ref. [21], copyright AAAS, 2016.
At first glance, we can note that each positive resonance in Fig. 6-7a corresponds to a local peak in the $T_{21}$ and a local valley in $T_{31}$. This indicates that in this resonant state, the carriers injected from contact one are being focused into contact two. By simulating the local particle current density injected by contact one in the first two positive resistance resonances, in Fig. 6-8a and c, we are able to confirm this conclusion. In the definition of the non-local resistance, eq. (6.1), a positive voltage in contact two should make the resistance positive. Thus, for the $p-p'$ junction the TEF resonances appear as positive resistances.

Conversely, for the negative resistance resonances in Fig. 6-7a, we see peaks in $T_{31}$ and a valleys in $T_{21}$. This case corresponds to when the current density skips over contact two and ejects out contact three. Interestingly, this results in a negative resistance. We will use the Landauer-Büttiker picture to interpret this result. When the carriers are transmitted from contact one into contact three, contact three will correspondingly float up in voltage. This will more heavily weight the transmission from contact three into other contacts when calculating the final resistance. Subsequently, carriers re-injected from contact three will skip along the top edge of the device, with a significant portion exiting contact four. Just as charging up contact two made the resistance positive, charging contact four will make it negative, giving rise to the off-resonant characteristics we observe.

We can perform similar analyses on the $p-n$ junction configuration, with the carrier density of the left and right sides of the sample set to $\pm 0.18 \times 10^{12}$ cm$^{-2}$. In this case, the $p-n$ junction is expected to redirect the flow of electrons at the junction towards the opposite edge of the sample. Therefore, this time, we will focus on $T_{41}$ instead of $T_{21}$. Fig. 6-7b shows the simulation results for the $p-n$ configuration. We observe significantly fewer TEF resonances in the $p-n$ configuration, which we attribute to the reduced transmission through the $p-n$ junction compared to the $p-p'$ junction, especially in the presence of a non-zero junction width, $D_W$, which enhances the angular filtering of the junction. This time we note two major positive resistance resonances and one major negative resistance resonance.

For the $p-n$ configuration, the negative resistance resonance is the true TEF resonance, corresponding to a peak in $T_{41}$. Again, we can look at the local particle current density injected by contact one corresponding to the TEF resonance, as is shown in Fig. 6-8b.
Figure 6-7: Slices through the simulated non-local resistance map. The panels of this figure show the non-local resistance as a function of applied magnetic field for (A) a p-p’ junction and (B) a p-n junction. In each panel the left side of the device is fixed to a carrier density of $-0.18 \times 10^{12} \text{ cm}^{-2}$. In (A), the right side of the sample is fixed to $-0.41 \times 10^{12} \text{ cm}^{-2}$, forming a p-p’ junction. In (B) the right side of the junction is fixed to $+0.18 \times 10^{12} \text{ cm}^{-2}$, giving a symmetric p-n junction. In each panel, the important transmission functions are also included, as is discussed in the main text. This figure is reproduced from Ref. [4], copyright American Physical Society, 2017.
Clearly, the carriers are re-directed by the junction and focused into contact four. This is accompanied by a noticeable caustic where the carriers interfere with one another and enter contact four.

Similar to the $p-p'$ configuration, carriers which are just outside of being resonantly focused give rise to an off-resonant signal. This time, the off-resonance state corresponds to the two positive resistance peaks in Fig. 6-7b to the left and right of the primary TEF resonance. Again, these peaks are accompanied by a peak in $T_{31}$. The same argument used for the $p-p'$ configuration holds here, except this time carriers are re-injected into contact two, on the bottom half of the Hall bar. This effect results in a positive resistance when the $p-n$ system is just out of resonant focusing into contact four.

Current density maps corresponding to the off-resonance characteristics of the $p-n$ configuration in Fig. 6-7b are shown in Fig. 6-5. These two off-resonant states correspond to when the magnetic field is either too weak, and the focused current ends up skipping just to the right of contact four (Fig. 6-5a), or when the magnetic field is too strong, bending the current flow such that it hits just to the left of contact four (Fig. 6-5b).

The maps of local particle current density in Fig. 6-8 are particularly useful in understanding the characteristics which are observed in the simulated non-local resistance in Fig. 6-6. Each map, (a, b, c, and d) in Fig. 6-8 corresponds to a key point marked in Fig. 6-6a. In the next several paragraphs, in an effort to build our physical intuition, we will look at key details of these maps.

Starting with the $p-p'$ configurations, in Fig. 6-8a and c, we look at the first and second resonant TEF states. The first resonant $p-p'$ state corresponds to a single skipping orbit. At first glance, the effect of the junction is not obvious, however deeper inspection shows that the current flow is in fact bent slightly at the junction. This results in an oblong, somewhat elliptical path, centered around the junction interface. Nevertheless, the current forms a strong caustic, exiting predominantly out contact two.

The second $p-p'$ resonance (Fig. 6-8c) corresponds to two skipping orbits between contact one and two. In this case, the effect of the junction is similarly as subtle. The bulk of the injected wave ends up arcing such that it skips directly in between the two contacts, right at the junction interface, slightly obscuring the junction effect. However, a portion of the wave
Figure 6-8: Spatially resolved maps of current density in the first and second TEF resonance configurations
(A) and (C) are both configured as $E_1 = 50$ meV and $E_2 = 75$ meV ($p-p'$ configuration) with the magnetic field tuned to be on the first and second TEF resonance, respectively. (B) is configured as a symmetric $p-n$ junction with $E_1 = -E_2 = 50$ meV, with the magnetic field placing it in the first TEF resonance and (D) is configured as $E_1 = 50$ meV and $E_2 = -100$ meV, with the magnetic field placing it in the second TEF resonance for the $p-n$ configuration. Brighter colors indicate large levels of current density and the scale bars are all 60 nm. The magnetic field in each panel is shown pointing out of the page. This figure is reproduced from Ref. [4], copyright American Physical Society, 2017.
Figure 6-9: Mode resolved current density This table depicts the mode-resolved current density for each of the junction configurations depicted in Fig. 6-8. Each column is summed in the final calculation of the maps shown in Fig. 6-8. The magnetic field in each panel is shown pointing out of the page. This figure is reproduced from Ref. [4], copyright American Physical Society, 2017.

is refracted at the junction, resulting in a secondary beam of carriers with a lower frequency skipping trajectory. This portion ends up completely missing contact two, transmitting into contact three. This subtle effect has a significant contribution to the final resistance, diluting the TEF signal and resulting in the reduced magnitude of the non-local resistance for higher order TEF peaks.

Fig. 6-8b shows the local particle current density for the first p-n resonance, with the junction configured symmetrically, \( E_1 = -E_2 = 50 \text{ meV} \). In the symmetric configuration, the carriers injected from contact one are bent towards the junction such that by the time they hit the p-n interface, they are largely transporting normal to the junction. Therefore, despite the large junction width, which enhances the filtering effect of the p-n junction, the bulk of the carriers transmit to the n-type side. After crossing the junction, the carriers are focused into a caustic, exiting contact four.
The second TEF resonant in the $p-n$ configuration is quite peculiar. In contrast to the first resonance, which is present for all $p-n$ configurations, the second resonance only appears for the $p-n^+$ configuration. This is due to the fact that around the second $p-n$ resonance, the carriers impinge on the junction at more extreme angles than the first resonance, enhancing the reflection off the junction.

When the right side of the sample is lightly $n$-type, the reflection is too extreme to observe a resonance. However, when the carrier density for the $n$-type side exceeds that of the $p$-type side of the junction, we do see a second TEF resonance. This corresponds to the current density map which is observed in Fig. 6-8d. This map is by far the most complex of all the current density maps we have observed thus far.

The left side of the map in Fig. 6-8d is actually quite similar to Fig. 6-8c, except for in the $p-n$ configuration the junction is significantly more reflective. This produces a much more chaotic interference pattern with a significant amount of carriers skipping across the junction and out contact five. However, at two points we see beams of carriers are able to tunnel across the junction, being focused into contact four and giving us a resistance resonance. This case particularly shows the power of quantum transport calculations for approaching the TEF problem. The secondary $p-n$ resonance is heavily influenced by quantum interference and is unlikely to be captured with a simple billiard model.

### 6.4.3 Mode-by-mode analysis

Another useful way to analyze the current density maps in Fig. 6-8 is to separate them by transverse modes. Recall that the lead, $p$, injecting carriers has an in-scattering wave function $\psi_p^{in}(x,y) = \sum_{n\in p} \phi_n(x) e^{ik_ny}$ with some number of transporting modes, $n$ (determined by the energy). The maps in Fig. 6-8 are each configured such that there are five modes injected by lead one. In Fig. 6-9 we separate the current density maps into the constituent transverse modes.

Separating the current density maps of Fig. 6-8 into different modes reveals an interesting interplay between quantum mechanical and semiclassical effects. Generally, the first transporting mode most closely resembles simple cyclotronic motion. The first modes inject straight into the sample, perpendicular to the lead, giving rise to the most simple transport.
Higher order modes enter the channel at oblique angles, which strongly affects their behavior.

Isolating the higher order modes for the first $p-p'$ resonance (left most plot of Fig. 6-9) gives a clear picture of the skewing effect we described earlier. The second and fourth modes have a pronounced kink at the top of the skipping orbit, whilst the third and fifth modes have more elliptical shapes. Similarly, starting at the third mode, in the second $p-p'$ resonance, we can see the small portion of carriers which are refracted near the top of their orbit, resulting in a secondary, lower frequency skipping pattern.

The third column of Fig. 6-8 shows the primary $p-n$ resonance. In this case the first two modes are quite simple, with only a small amount of interference fringes in their trajectory. The third and higher modes, however, display an interesting effect which was quite subtle for the full current density picture. Each mode propagates as a pair of beams which are injected at equal but opposite sign angles into the channel. Starting at the third mode, in this case the component which injects moving to the left of the channel follows an arc which results in it reflecting off the top edge of the Hall bar. From there, it skips towards the $p-n$ interface with a nearly normal trajectory. These higher-order modes have significant levels of interference effects present and are critical to obtaining the correct simulated resistance, highlighting the benefits of the tight-binding transport approach.

Finally, the second $p-n$ resonance is particularly rich when broken into separate modes. The first few modes begin transporting similarly to the second $p-p'$ resonance, but when they are reflected by the $p-n$ junction, their trajectory results in the carriers traveling straight up the junction. The carriers in these modes follow paths which are reminiscent of the Landau level transport which we studied in Chapter 4, creating an interesting parallel between low magnetic field and high magnetic field transport. Additionally, by isolating the higher order modes we are able to more clearly see the skipping orbit patterns which form, with the bulk of each wave being reflected off the junction, transmitting only at a few points to make up the final caustic which yields the TEF resonance.
6.5 Chapter Conclusions

In this chapter we have used tight-binding transport theory to accurately model the TEF experiment in graphene p-n junctions presented by Chen et al [21]. We demonstrated that dephasing is a key effect for modeling magnetotransport in today’s high quality graphene samples. Using the Landauer-Büttiker approach we were able to model non-local resistance measurements with a high degree of success.

We were able to model both the positive and negative resistance TEF resonances, which were then explained using transmission function calculations and transport visualizations. Key features of the maps in Fig. 6-6 were pinpointed and studied, building significant physical intuition for the problem. Mode-by-mode analysis of the transport also let us appreciate the subtle effects which built to give the final results.

These results are particularly relevant for understanding other modern devices where the phase coherence length is long. Our results show that when uncontrolled, phase coherence can result in a chaotic sample. The techniques we have presented give a platform for others modeling modern devices, but also serve as a guide to experimentalists who made need to engineer dephasing into their samples. The insights we have provided here may prove particularly useful to those who are trying to build different quantum interferometers, such as Fabry-Perot or Mach-Zender interferometers, where control over the mechanism of dephasing in the sample is is critical.
Part II

Nanoelectronics in van der Waals heterostructures
Chapter 7

Fabrication and measurement of high-quality van der Waals heterostructures

This chapter will serve as a sort of secondary introduction—whereas the first part of this thesis was heavily based in theoretical calculations, here we will focus on the practical aspects of experiments in nanoelectronics. We will start by going through an introduction to van der Waals heterostructures, stacked layers of two-dimensional (2D) materials. While we have already met one of the major materials used in these heterostructures, graphene, we will see how there are several other very interesting 2D materials to work with. From there, we will delve deeply into the fabrication techniques which have been developed over the last year and a half, including multiple versions of the pick-and-place technique. At the end of this chapter we will also see results from a very high quality graphene p-n junction sample and also measurements from an electrically reconfigurable WSe$_2$ device with tunable graphene contacts.
7.1 Two-dimensional materials and van der Waals heterostructures

2D materials are not limited to the Dirac semimetal graphene, which we have discussed at length up to this point. In fact, there is an extremely wide range of crystals populating the growing field of 2D materials [13]. 2D materials cover essentially the entire cross-section of electronic materials, including semiconductors [118] [119], insulators [120] [121], topological insulators [122], and other even more exotic systems [123] [124].

2D materials share a common trait— their crystal structures consist of layers of atoms which are held together by van der Waals forces. In many of these materials, the van der Waals force between the crystalline layers are weak enough such that they may be mechanically exfoliated (with household adhesive tape) down to few-layers or even monolayers [13]. Recently, it has become possible to assemble stacks of 2D materials [125], forming what is known as a van der Waals heterostructure [123] [124], the major focus of the rest of this dissertation.

7.1.1 Transition metal dichalcogenides

Aside from graphene, one of the most highly studied 2D materials have been transition metal dichalcogenides (TMDs). Indeed, mechanical exfoliation of TMDs was performed as far back as the 1960s [126] and the first monolayers of TMDs (MoS$_2$, specifically) were observed in 1986 [127], well before the discovery of monolayer graphene. The makeup of TMDs follow the formula MX$_2$ with M being a transition metal element and X being a chalcogen element.

While TMDs host a number of electronic properties (including insulators, metals, and semimetals), in this dissertation we are primarily interested in TMDs which are semiconducting, including MoS$_2$, WS$_2$, MoSe$_2$, and WSe$_2$. These semiconducting TMDs have a trigonal prismatic structure, referred to in shorthand as the 2H phase, which tends to be the stable phase for these materials [128]. The monolayers of the 2H phase follow an ABA stacking order, where chalcogen atoms sandwich the transition metal atoms to form one monolayer of the crystal. Viewed from the top down, the 2H phase forms a honeycomb-like crystal struc-
Figure 7-1: Crystal structure of MX$_2$ transition metal dichalcogenides This figure depicts the crystal structure of three phases of TMD materials, reproduced from a review by Manzeli et al. [128], copyright Springer Nature, 2017. In this dissertation, we will be exclusively working with the 2H phase, which results in a TMD with semiconducting properties.

There is also a 1T phase which follows an octahedral structure with ABC stacking. While this thesis primarily focuses on TMDs which follow the 2H phase, it is interesting to note that phase engineering—transforming a crystal from the 2H to the 1T phase—has emerged as a powerful technique [129]. These two phases are shown in Fig. 7-1 reproduced from the review by Manzeli et al [128] (copyright Springer Nature, 2017).

The electronic structure of semiconducting TMDs is particularly interesting. Bulk TMDs, for our purposes thicknesses greater than five layers or so, have indirect band gaps between $\approx 0.6$ and $1.2$ eV [130]. As the thickness is reduced, the band gap of TMDs increases and finally at the monolayer form it even becomes direct (at the K and K’ points of the hexagonal Brillouin zone) with band gaps around $2$ eV [130] and even up to $2.4$ eV for WS$_2$ [131].

The ability to isolate atomically thin layers of TMDs has made them a major candidate in post-silicon electronics [118][119][132]. Recently, the shortest field effect transistor in the world was demonstrated in a MoS$_2$ device which used a carbon nanotube as a gate, reaching a $1$ nm channel length. Additionally, the indirect to direct band gap transition, as well as the tunable band gap as TMDs are thinned to a monolayer, makes TMDs particularly interesting in optoelectronic studies [133].
7.1.2 van der Waals heterostructures

While on their own, 2D materials present a great deal of interesting physics to study, their most exciting properties may be found when combined together. When two different two-dimensional materials are stacked together, joined by the van der Waals force, we call them a van der Waals heterostructure [123]. The plethora of different types of 2D materials we have mentioned above (metals, semimetals, semiconductors, insulators, and others) can be stacked on top of one another to form a heterostructure with new and special properties. An excellent summary of van der Waals heterostructures may be found in Fig. 7-2, reproduced from an excellent perspective written by Geim et al (copyright, Springer Nature, 2013).

Perhaps the most basic form of van der Waals heterostructure which has been imple-
mented is where hexagonal boron nitride (h-BN) is used as a substrate for other 2D materials. h-BN was demonstrated as a substrate for graphene, which allowed Dean et al to observe pristine, phonon-limited electronic properties in their graphene device \[120\]. Since then, h-BN has become ubiquitous in the world of van der Waals heterostructures, proving to be the best substrate for 2D materials so far.

In addition to the simple graphene-BN heterostructures, much more complex devices may be created. 2D versions of conventional electronics have been realized, such as atomically thin p-n junctions which exploit band alignment of different stacked TMDs \[134\]. Also, more exotic physics have been observed, such as Hofstadter’s butterfly \[135\] and ballistic minibands resulting from Moiré potentials of misaligned of h-BN and graphene \[136\]. Van der Waals heterostructures give us many variables to control in the lab, such as layer thickness, interlayer spacing, and interlayer twist angle \[137\].

7.2 Challenges in van der Waals heterostructures

7.2.1 Fabrication and assembly

Likely the largest challenge in studying van der Waals heterostructures has been their fabrication and assembly. Early heterostructures were formed using a deterministic transfer technique in which 2D materials were mechanically exfoliated onto polydimethylsiloxane (PDMS) stamps and transferred in a layer-by-layer fashion \[138\]. CVD-grown heterostructures, such as direct growth of MoS\(_2\) onto h-BN, have also been demonstrated \[139\]. Each of these techniques, unfortunately, are not capable of forming ultra-high quality heterostructures. The layer-by-layer mechanical transfer, whilst relatively easy to perform, leaves unavoidable polymer residue in between the heterostructure layers. Additonally, CVD growth of 2D materials have shown a great deal of progress, but the crystals still do not reach the level of quality possible with mechanically exfoliated materials \[140\].

Our main interest here is attaining high quality van der Waals heterostructures for studies of fundamental physics and nanoelectronics. By far, the highest quality van der Waals heterostructures to date have been the result of of the van der Waals assembly process \[125\].
Demonstrated first by Wang et al., the van der Waals assembly process provides a way to form extremely clean van der Waals interfaces between 2D materials [125]. This technique has received a number of names, such as the van der Waals transfer, polymer-free transfer, pick-and-place (our lab’s favorite), and others.

While the van der Waals assembly techniques have come under many names (with variations in techniques, as we will discuss later), they all are based off a common method. In its most basic form, using graphene as an example, a thermally-sensitive polymer is used to first pick up $h$-BN, then the $h$-BN is used to pick up a piece of graphene, and finally the $h$-BN/graphene stack is placed onto a bottom $h$-BN layer. With this technique, it is possible to form heterostructures with an arbitrary number of layers with very clean interfaces. This is mainly due to the fact that only the $h$-BN is ever exposed to polymer, keeping the interior layers pristine.

While the van der Waals assembly process does not expose the interior of the heterostructure to polymer residue, it can still be challenging to achieve clean interfaces between the layers. During the assembly process, if the layers are not brought together in a very slow way, hydrocarbons on the surface of the 2D materials can coalesce, forming sites of trapped contamination which are commonly referred to as “bubbles” or “blisters”. These interfacial bubbles may become so large that they are even thicker that of the total constituent crystals of the heterostructures [141].

Generally, interfacial bubbles are extremely undesirable and can produce electrical inhomogeneity inside the heterostructure, thus, techniques which allow us to avoid the contamination are very useful. While Wang et al. were able to demonstrate very clean interfaces [125], most users will notice that directly replicating the process is quite challenging (although not impossible). This is likely the reason that several publications have emerged since Wang et al. presented their work, demonstrating similar but slightly modified assembly techniques. These include the so-called “fast pick-up method” [142], the “hot-pick up technique” [141], and an “interfacial cleaning technique” [143]. These processes all share common traits in that they present methods by which interfacial bubbles may be reduced or avoided, usually by performing the transfers at elevated temperatures beyond the 40 °C suggested by Wang et al. Later in this chapter, we will go through a detailed description of the trials which led
to creating bubble-free heterostructures.

### 7.2.2 Electrical contact

Encapsulation in $h$-BN has emerged as the superior technique for creating pristine, extremely high quality van der Waals heterostructures. Unfortunately, sealing the electrical channel of a device inside of $h$-BN presents a challenging problem: how do we create electrical contact to the material? In encapsulated samples, it is no longer possible to metallize just the top side of the channel, as was done in the early days of work on 2D materials.

Wang et al presented a surprisingly simple solution to the problem for encapsulated graphene samples, which turned out to produce extremely high quality contacts [125]. They found that by using reactive ion etching (RIE) to slice vertically through the stack, it is possible to create one-dimensional edge contacts to the graphene. While this technique works very well for graphene, making edge contacts to TMDs have presented a significant hurdle. While some progress has been made very recently in two papers [144][145], the results have been exclusive to MoS$_2$.

One promising method for creating contacts to encapsulated TMDs has been to create van der Waals interfaces between the TMD and a metal or semimetal. Recently, James Hone’s lab has published two papers describing their via method, which yields very high quality contacts by pre-patterning holes into the top $h$-BN which are then metallized and used for the subsequent van der Waals assembly [146][147]. Using 2D metals, such as NbSe$_2$, has also been presented as an alternative method [148][149]. Another method is to use graphene as an intermediate contact material to the encapsulated 2D, which lets one exploit the ability to create edge contact to the graphene [150].

Indeed, the van der Waals interface-based contact technique has become quite useful in studying the electrical properties of encapsulated TMDs. Additionally, the van der Waals metal-semiconductor interface itself turns out to be quite interesting. Recently, a number of papers have emerged showing a lack of Fermi level pinning in peeled/transferred metals [151][152][147]. We will discuss this topic in detail in chapter 8, where we will see results on graphene-WSe$_2$ van der Waals junctions which exhibit a complete lack of Fermi level pinning [5].

122
7.3 Device fabrication

All of the modeling and simulation work I performed in the first portion of graduate school heavily focused on connecting quantum transport theory and experimental results. To do this successfully, I found it was necessary to put myself in the mindset of an experimental physicist- I had to consider not just the fundamental physics of transport, but understand how the experiments were performed. Particularly, modeling the transverse electron focusing problem piqued my interest in measurements on very high quality samples. Near the end of the summer of 2017, Professor Lee acquired a new electron beam lithography tool (the Raith Voyager) for our lab, which gave me the opportunity to begin working on experiments. To my surprise, I found the aspects of advanced device fabrication and measurement to be just as, if not more interesting, than the theoretical work I had done in the past.

Fairly early on throughout this effort, I became dissatisfied with the quality of devices I could fabricate using our lab’s older techniques (based on graphene grown via chemical vapor deposition (CVD)). Seeing the quality of graphene devices others were making based on the van der Waals heterostructure platform, I was inspired to move towards that area. Learning to assemble van der Waals heterostructures, using a technique which we have come to refer to in the lab as “pick-and-place”, ended up being a major challenge, however I believe the results have been worth it.

In this section I will be switching to the first-person voice to discuss the details of the device fabrication steps I have followed for making the samples measured throughout the rest of the dissertation. I have found device fabrication to be an extremely challenging task, but the results have been worthwhile. Furthermore, I have found the entire process to be like an art- a number of the steps, especially in the pick-and-place transfer, are hard to communicate directly with words and still images. However, I will emphasize what I have found to be the most important details here, which should prove useful to anyone who wishes to replicate these methods in their own work.
7.3.1 Mechanical exfoliation techniques

Early on in my efforts to learn how to fabricate devices from 2D materials, I opted to switch from materials grown from CVD to working on van der Waals heterostructures—fully encapsulating the samples in \( h \)-BN. As part of the move to encapsulated samples, I began working with mechanically exfoliated materials. I consider device fabrication to be an art, and the exfoliation step is no different. There are many variables to consider for achieving optimal exfoliations—the substrate type, substrate cleanliness, substrate functionalization, tape selection, crystal selection, tape handling, sample heating, and others. Additionally, the optimal techniques and parameters seem to vary wildly for different crystals. A good exfoliation makes the subsequent heterostructure assembly significantly easier, so it deserves a careful documentation.

Proper preparation of the substrate is extremely important prior to exfoliation. One major consideration is ease of identification of exfoliated crystals based on contrast \[153\]. Generally, I prefer to exfoliate crystals onto silicon wafers with 90 nm of thermal oxide grown on them. In our group, we have gotten these wafers in cassettes of 6 300 mm wafers from CNSE’s chip fab, and I have found that they work quite well for both exfoliation and as a final device substrate. On 90 nm oxide, thin layers of graphite and TMD materials are quite easy to identify, turning dark shades of grey and then becoming translucent as they reach monolayer form. Boron nitride is also quite easy to identify, with different thicknesses going through a rainbow of colors \[154\]. Additionally, 300 nm thermal oxide seems work just as well, but finding thin flakes on this thickness of SiO\(_2\) is more challenging.

If the wafer is not clean prior to exfoliation, the results will likely be quite poor. Early on in learning exfoliation, I would clean every single piece of wafer in boiling acetone followed by IPA, usually submerging the wafers for several minutes in each, followed by drying with nitrogen. Often, ultrasonication would even be included in an effort to further clean the substrate. After some experience, however, I believe these cleans are not necessary on fresh samples. It seems that wafers which come out of the fab are already ultimately clean, and further wet cleaning with common solvents only has the chance of introducing contamination to the wafers. For this reason, the thermal oxide wafers from the fab should always be handled
very carefully and with fresh nitrile gloves, such that they remain in their pristine state until being used.

There is one cleaning technique, though, that is quite useful. Although, while we call it a *clean*, I believe it should be considered as a *functionalization* step. Subjecting the wafer to an oxygen plasma clean prior to exfoliation is one of the best ways to increase both the yield of exfoliated flakes on a sample as well as the size. Huang *et al* showed how oxygen plasma cleaning may be used to produce up to millimeter-sized monolayer graphene [155]. In practice, millimeter-sized graphene is difficult to achieve and perhaps even undesirable (if the process which yields such graphene makes it impossible to pick up), but graphene areas up to hundreds of micrometers across are routinely observed. Additionally, with an oxygen plasma treatment, TMD monolayers (which are much more challenging to obtain than graphene) between 10 and 100 micrometers across have become possible.

My experience has been that the particular oxygen plasma recipe is quite important. All of the oxygen plasma treatments I have performed have been done in a Technics 800 MicroRIE- an *etch* tool, not a dedicated plasma cleaning tool. The MicroRIE lets us tune the flow rate of the oxygen, the RF power generating the plasma, and the time. For graphene, most basic configurations seem to work, although going too high on the flow rates (greater than 30 sccm of oxygen) or time (greater than two minutes), while yielding the largest area graphene, can make the graphene *impossible* to pick up during the subsequent assembly process.

Recently, I have settled on doing very short plasma treatments (20 seconds or less) for graphene, usually with about 20 sccm of O$_2$, which works quite well. For TMDs, it seems a more aggressive plasma treatment is necessary, usually with higher O$_2$ flow rates for several minutes. While these aggressive cleans can make graphene difficult to pick up, I have not found the same problem with TMDs. I have found one major caveat for this process- raising the RF power too high (greater than 50 watts) will generally ruin any exfoliation if the plasma is applied for more than a few seconds, likely due to roughening of the SiO$_2$. Recently, Professor Lee has attained a dedicated plasma cleaning tool for the lab, which should hopefully simplify this process greatly.

Crystal selection is another important factor for achieving good exfoliation results. For
graphene, I have experimented with a number of different types of graphite sources. These include Naturagrapht Systems (NGS) graphenium natural graphite, NGS “fishy flakes”, HQgraphene natural graphite, kish graphite, and highly oriented pyrolitic graphite (HOPG, ZYA and ZYH grade). Of these crystals, I have had little success with HOPG, which yields many thin layers (including graphene), but the exfoliated crystals come off in many very fine strips which are not usable for devices. The HQgraphene natural graphite has also yielded poor results.

Kish graphite and both types of NGS-brand graphites work very well, routinely yielding graphene areas over 100 micrometers across. For final devices, both Kish graphite and graphenium graphite have yielded high electron mobility samples, as well. Given the volume of graphenium graphite we have, I have gravitated towards using it the most, and all of the devices in the rest of this dissertation will use graphene exfoliated from those crystals.

For TMD crystals, I have primarily used flux-grown WSe$_2$ and WS$_2$ from the 2D semiconductors company. Flux grown crystals appear to be much higher quality when compared with TMDs grown from the more common chemical vapor transport (CVT) method [156]. While we do not have a good comparison of the electrical characteristics of the 2D semiconductors brand crystals and the previous CVT grown crystals our lab used to use (as the overall device structures are quite different, with newer samples being encapsulated in $h$-BN), the flux grown crystals yielded drastically superior exfoliations. From a practical standpoint, larger area crystals are much easier to work with during the transfer. Additionally, the ability to exfoliate larger area crystals likely indicates a higher quality crystal, overall.

None of the work which will be presented in the following chapters would be possible without high quality $h$-BN crystals. We obtain high-quality bulk $h$-BN crystals provided from our co-workers from the NIMS institute, Takashi Taniguchi and Kenji Watanabe. These $h$-BN crystals, when handled properly, yield large and clean areas of exfoliated material, which can extend up to nearly a millimeter across. The $h$-BN which we receive comes in different varieties, which indicate different growth conditions [157]. My experience has been that the “694” and “695” varieties of $h$-BN have both yielded high quality devices without gate hysteresis, although I expect other varieties will work too.

Tape selection has a large influence on the exfoliation process. In all of the tapes I have
worked with, there is a trade-off between the level of adhesion of the tape (its ability to cleave crystals) and the amount of adhesive residue which is left behind. Tapes which are very sticky (such as Scotch magic tape) readily cleave most crystals, but tend to leave a large amount of residue behind on the final substrate. For larger area bulk crystals, such as our graphite and TMD sources, this tends not to be a large problem, as the exfoliated crystals will not end up near the adhesive residue. For bulk $h$-BN crystals, which have relatively small area, this can be a major issue.

Over time, I have found that Scotch magic tape consistently works well for both graphite and TMDs. For $h$-BN, or when low residue exfoliations are desired, the Nitto-brand pressure sensitive tape works extremely well. While the Nitto-brand tape is not very sticky normally, it appears to have some viscoelastic behavior, whereby ripping the tape off of itself or a surface very quickly readily cleaves the bulk crystals.

Now that we have discussed the sample preparation, the crystal selection, and tape selection, we can discuss how the actual exfoliation is performed. One of the most important things to remember is one does not obtain monolayers of material directly on the surface of their tape. Rather, cleaving bulk crystals on a piece of tape exposes fresh crystalline surfaces which can then be adhered to the final substrate. With some luck, these fresh surfaces will adhere to the substrate and cleave once more when the tape is removed, leaving behind a monolayer or thin film.

Understanding this mechanism of exfoliation is key to developing the proper technique. Our goal in exfoliation is to cover the tape with as many freshly cleaved surfaces as possible, at the same time maximizing the area of these surfaces which can adhere to the substrate. To achieve this, I generally will apply a single bulk crystal to one tape, and then use a second tape to repeatedly cleave the original crystal and tessellate it around in a regular pattern. Special care is given to simultaneously achieve good coverage over the tape (reducing the residue left behind) but also minimize overlap between cleaved crystals (overlapping crystals reduce the area of the clean surfaces which may adhere to the substrate). It is also important to be as gentle as possible with the tape, as the crystals can be easily pulverized under too much force. An example of this technique, as well as an example of an excessively exfoliated tape, is demonstrated using graphenium graphite in Fig. 7-3.
Figure 7-3: Demonstration of the mechanical exfoliation technique. This figure depicts the process of mechanical exfoliation of graphite using Scotch tape. (A) We start with two pieces of tape with the ends folded to create handles. (B) A piece of natural graphite is applied to the corner of one piece of tape and (C) the other tape is used to carefully cleave the crystal in half and tessellate it across the two tapes, with little overlap. (D) An example of a finished graphite tape ready to be adhered to a substrate to isolate graphene. (E) An example of a tape which has been exfoliated too many times, which will result in small area graphene and a messy exfoliation.
As a corollary to this method, I have found that folding the tape with a sharp kink at any time is detrimental to the entire process. Additionally, the number of times a given crystal is peeled must be considered. Cleaving a crystal too many times seems to reduce the maximum area of exfoliated materials quite drastically, especially for synthetic crystals like $h$-BN and TMDs (natural crystals are much more forgiving). Usually, I find peeling the initial source crystal more than five times is excessive.

Usually, I will exfoliate my crystals across the tapes within a few minutes of adhering them to the plasma-treated samples, with the final peel of the tape being performed right beforehand. Leaving the tapes adhered together for too long (especially with Scotch tape) tends to produce greater levels of residue with fewer usable crystals. Some care should be used when placing the tape on the substrate, with as few bubbles as possible being trapped between the two. If desired, a flat tweezer can be used to gently increase the adhesion between the tape and the substrate. From there, usually I will apply pressure to the tape using my thumb. For graphite, the best results have come when I press the tape extremely hard, which appears to be a very important step in the entire process. For TMDs and $h$-BN, pressing the tape hard appears to be less desirable.

I have encountered many differing opinions on the topic of removing the tape from the substrate. Some say that the tapes should be removed quite slowly, while others encourage “ripping” the tape off the substrate. I have found the answer to be quite nuanced- with Scotch tape (and likely other high-adhesive tapes), removing the tape slowly leaves a great deal of glue residue compared to very quickly. Over time, I have adopted a medium pace, where the tape is usually removed over the course of a few seconds and as smoothly as possible. For $h$-BN exfoliation, very slowly removing the Nitto-brand pressure sensitive tape will generally leave behind more crystals on the surface, but I have found ripping that tape very fast tends to leave behind more freshly cleaved (and clean) crystals.

After exfoliation, it can be very tempting to clean the sample in acetone and IPA to remove any tape residue, however this should not be done. The crystals which are exfoliated are usually coming from the interior of the freshly cleaved bulk, and will be about as clean as they will ever be. Occasionally, even thin exfoliated crystals will have glue residue on them, and these should be discarded in favor of fresh samples, rather than cleaned. Introducing
solvents will simply add contamination to otherwise clean crystals, and should be avoided. It is possible that various anneals in inert gas (such as H₂/Ar anneals) may further clean the surface without contaminating the exfoliated crystals, although I have found this is not useful if the intention is to eventually transfer any of the crystals to a different final substrate (in the van der Waals assembly process).

As a final word, heating the substrate on a hot-plate prior to removing the tape can be used as a last resort to improve the yield of thin flakes and monolayers [155]. While this step has shown some improvement in certain crystals (graphene and TMDs, in particular) it simply seems to skew the probability in favor of getting decent results, not enable them. Of course, this is at the expense that the residue is markedly increased on the sample, which can make the transfer more challenging. Additionally, for graphene, this heating step has shown to make the graphene extremely hard or impossible to pick up with h-BN, so I have largely abandoned it in my processes.

7.3.2 van der Waals assembly

While the van der Waals assembly method proposed by Wang et al enables the creation of layered heterostructures with very clean interfaces between the different materials [125], it requires a great deal of care to perform correctly. While polymer residue is no longer a problem with this technique, trapped interfacial bubbles of coalesced hydrocarbons are quite problematic when making electrical devices. An example of an early sandwiched graphene sample is shown in Fig. 7-4, where numerous interacial bubbles are seen both optically and with atomic force microscopy.

When assembling a heterostructure using this method, a heat-sensitive polymer stamp is used to pick up and h-BN flake which is then subsequently used to pick up other 2D crystals and finally placed onto a bottom h-BN. At each step, there is a risk of introducing interfacial contamination between the layers of 2D materials. Primarily, this tends to be the result of bringing the crystals together too quickly, which does not allow the contamination to escape during the assembly. The simple answer to this problem is to bring the crystals together as slowly as possible- practically, it is not that easy. For this reason, Purdie et al developed a method which is used to clean trapped contamination at the end of the assembly process.
Figure 7-4: Example of an early encapsulated graphene sample. An early graphene sample encapsulated in $h$-BN placed on buried split-gates. Early samples were plagued by numerous bubbles which result in poor transport characteristics. The top image is an optical microscope image and the bottom image is an AFM topography scan, which clearly highlights numerous interfacial bubbles.
While I have not directly replicated their results, the final assembly process which I have developed takes a great deal of influence from their method, mixed closely with the original technique of Wang et al.

As part of learning how to do the van der Waals assembly process, I had to develop a transfer stage to perform the process. The final transfer setup is shown in Fig. 7-5. The setup consists of an inspection microscope with long working distance objectives and a global XY-stage. On top of the global stage I placed an optics breadboard which contains a secondary sample stage and stamp stage. The sample stage consists of an $XY\theta$ stage and the stamp stage consists of an $XYZ$ stage with an extra pitch/yaw stage (for controlling the glass slide angle) and special mount for attaching a glass slide. I have found that the freedom of having many redundancies for XY motion, particularly the global XY stage, to be very useful during the assembly process.

On top of the sample $XY\theta$ stage I have mounted a custom aluminum heater and vacuum...
chuck on ceramic stand-offs. The heater stage was designed and machined with the help of Brian Taylor and Kyle Unser (of the CNSE academic support group). The heater is controlled with an Omega PID controller and can reach a temperature up to 300 °C, although the most recent transfer process steps rarely exceed 150 °C. During the heating and cooling of the stage, a great deal of thermal expansion can be observed, which acts as a makeshift sample z-axis motion and has even been exploited in some transfer methods [142].

Additionally, I added a ThorLabs Z-series compact stepper motor to the Z-axis of the stamp stage, which assists in fine-control of the pick-up and drop-down steps of the assembly. The benefits of the stepper motor have been two-fold. First, the addition of the stepper motor was a major step forward in the usability of the transfer system, as it removed the requirement of physically touching the stages during the critical transfer steps (which would induce lots of vibrations). Second, the stepper can be programmed to move in fixed step sizes or speeds, which improves reliability of the entire process. Unfortunately, the stepper does have a noticeable amount of backlash in its motion, but that should not be a problem for a practiced user.

During assembly of heterostructures, the pitch and yaw stage will be used to slightly angle the transfer stamp down relative to the stample, such that when pressed down, the end of the stamp will contact the sample first. As the stamp is a soft material, it will deform, creating a contact front which will proceed across the sample from one side to the other until the stamp is fully in contact with the sample. This allows us to carefully and gradually pick up and set down flakes of 2D materials, rather than simply ripping them off the surface all at once (or slamming them down). This control over the motion is very important as it can allow us to manually manipulate interface contamination out of our heterostructures, given the right conditions [143].

Construction of the transfer stamp is an integral part of this entire technique. The basic stamp consists of a PDMS block mounted on a glass slide which is then covered in a thermally-sensitive polymer, either polypropylene carbonate (PPC) or polycarbonate (PC). I have used both polymers with some success throughout my thesis work, although I have recently settled on PPC as the preferred polymer. Despite that, PC can work quite well- for example, the samples fabricated for chapter 8 were assembled using PC.
Figure 7-6: Example process flow for the creation of a stamp used in the transfer process. This figure represents the basic process flow for creating the stamps used in the transfer methods described within the main text. While the figure denotes PPC as the transfer polymer, the steps are essentially the same for PC-based stamps.
Aside from the choice of polymer, the assembly of the stamp for each method is essentially the same, shown in Fig. 7-6. Prior to assembling the stamp, Sylgard 184 PDMS is mixed in a ratio between 9:1 to 11:1 base to binder, where 9:1 produces a stiffer stamp. The PDMS is cast onto clean silicon wafers in petri dishes, at a thickness between 1 and 3 mm, followed by curing overnight on a hotplate set to 70 °C. In all of these steps it is important that the PDMS is kept very clean and that during the curing step, the petri dish is kept level.

First, a double-sided piece of kapton tape is hole punched and applied to the end of a glass slide. As PC or PPC will not adhere well directly to the glass slide, the kapton ensures our final stamp will be well bonded together. With a razor blade, a small block (a few millimeters across) is cut from the previously cast PDMS and placed in the middle of the kapton tape hole. As long as the glass slide and PDMS are clean, they should adhere together well enough. I prefer to turn the PDMS block as shown in Fig. 7-6 since the portion of the stamp towards the end of the slide will touch first, having a sharp corner on the PDMS there gives better precision when transferring multiple heterostructures to a sample.

At this stage, the stamp will usually receive a short O₂ plasma treatment, which promotes adhesion between the PDMS and the PPC or PC. I find that for PPC, this step is essential to prevent delamination of the layers in subsequent use. For very freshly mixed PC, the adhesion will be much better, thus the plasma treatment is not required (and may even be harmful to the entire process).

Separately, we must prepare either a PPC or PC film to cover our PDMS block. For PPC, we follow the typical suggestion of mixing PPC pellets 15 % by weight in anisole by stirring at 100 °C. Once mixed, this solution may be spin-coated at 1500 rpm onto clean silicon wafers. I have experimented with different times, and settled on a short spin coat (less than 30 seconds) with a slow ramp up speed (approximately 250 rpm/second) which produces a smooth center and significant edge bead for easier handling.

For PC, I have found the best success with a 10 % solution mixed in chloroform. Usually, this mixing will be done by ultrasonication. With PC I have noticed one major drawback, which is that the PC solution must be used within a week or so of mixing, otherwise its ability to pick up crystals greatly reduces. As a rule of thumb, if the PC delaminates from the PDMS block during the transfer steps, it likely is not fresh enough.
PC in chloroform can be quite hard to spin-coat, as the chloroform will flash dry, resulting in spinning a dry film which becomes very rough. Very short spins have produced reasonably smooth films, however they are hard to achieve consistently. As a result, I prefer the method of casting PC between two glass slides \[142\] \[143\]. To perform this method, a few drops of PC solution are dropped onto a nitrogen cleaned glass slide. Quickly, another glass slide is pressed on top, spreading the PC, and the two slides are smoothly slid apart and dried for at least five minutes. With some practice, this can produce a fairly smooth film. As the PC solution will almost certainly get on ones gloves in this process, I use the same heavy duty gloves which would be used for handling acids and bases in the cleanroom- chloroform will immediately penetrate nitrile!

Now that we have a slide with a PDMS block, as well as our cast PPC or PC, the final assembly of the stamp can begin. A piece of scotch tape is carefully hole-punched (applying the backing of the double stick kapton to the sticky side of the scotch tape can help) and then placed onto our previously cast polymer. With a clean razor blade, the polymer should be cut around the edges so that it lifts cleanly from the substrate. Slowly peeling off the scotch tape should remove the PPC or PC from the substrate, giving a supported drumhead in the scotch tape window. This may then be carefully stretched over the PDMS block (sticky side of the scotch tape down) and the entire tape can be adhered to the glass slide, with the excess removed or folded around the edges of the slide.

At this point, the stamp is essentially complete. It should be carefully inspected to make sure that the polymer membrane over the PDMS has not been ripped and is free of any gross wrinkles. If the PDMS is particularly thick, there is a risk that the membrane can tent over the corners of the block, producing a free-standing dome where the PDMS and membrane are not in contact with one another- this can be avoided by carefully pressing the scotch tape down and leaving a small region which is not completely adhered. I will usually make three to four stamps at a time in this way, with PPC stamps having remained usable for over a month after assembly. PC stamps, on the other hand, should generally be used within a day or so.

Once the stamps are assembled, there are two major techniques which I have used to create the final heterostructures. The core concepts between the two techniques are the same
as we have already described, where the stamp is used to pick up the top $h$-BN, subsequent crystals, and then the bottom $h$-BN. The exact conditions (namely, temperature) these steps are performed at and the behavior of the stamp at each step, however, differs significantly between PPC and PC. First, I will describe the methods I use for PC-based stamps, which are what were used to produce the TMD-based heterostructures in chapter 8. Then, I will describe the methods used for PPC-based stamps, which I have used to create the highest quality heterostructures I have achieved throughout my thesis work.

### 7.3.3 PC-based stamps procedure

As assembled, the PC stamp will be very sticky between 40 °C and 80 °C, and may be used as described in other publications [142][143]. I have found an alternative method that I prefer, which pushes the “sticky”-point of the stamp up to greater than 125 °C, with approximately 135 °C being the maximum temperature before the stamp will be destroyed if it is attempted to be removed from the substrate (because of the extremely high level of adhesion). To achieve this change in the usable temperature of the stamp, the finished PC stamp is annealed on a hotplate at around 200 °C for a few minutes. This is roughly 50 °C higher than the glass transition temperature for PC, which may be responsible for the changes in the viscoelastic behavior in the stamp.

This procedure produces significantly fewer interfacial bubbles than what I was able to achieve with untreated stamps, which is the result of two linked effects. Generally, the slower that we bring together two 2D flakes, the fewer interfacial bubbles will be produced. However, achieving a stamp which has smooth motion (enabling a slow assembly of the different layers) and is also able to adhere well to the 2D materials is very difficult. As it has been noted that performing the assembly at elevated temperatures reduces the likelihood of bubbles forming for the same set-down speeds [141], our heat treatment eases the requirements of extremely slow set-downs of the crystals. Additionally, the heat treatment makes it so there is a low level of adhesion on the stamp at around 120 °C, which makes the motion of the stamp significantly smoother and more controllable.

As a result, for annealed PC stamps, whenever crystals are being picked up, the stage will be set to around 125-135 °C. When crystals are being brought together, the stage will be set
between 90 °C and 120 °C, which lets us perform a significantly smoother assembly. Despite this, interfacial bubbles will still often appear in the set-down onto the bottom h-BN, which seems to be the hardest step to perform very smoothly.

For this reason, I have adopted the method proposed by Purdie et al, in which interfacial bubbles may be swept out of the heterostructure at elevated temperatures [143]. During the set down of the heterostructure onto the final substrate, I will raise the temperature of the stage over 200 °C, which will simultaneously melt the PC off of the stamp and release the heterostructure. By carefully performing this step, often interfacial bubbles may be pushed out of the stack. It is worth noting, that since the interfacial bubbles usually appear between the bottom h-BN and the rest of the stack, the interfaces between the above crystals are likely quite clean, as evidenced by the results in chapter 8.

Over time, I have noted a number of problems with this method which have caused me to move away from it. First, the time-dependent degradation of any mixed PC solution (likely due to the evaporation of the chloroform into the head-space of the vial the solution is stored in) is very problematic for variability during the assembly process. Ideally, the van der Waals assembly will require the same conditions each time, or else the statistical yield of the heterostructure assembly is quite low. Second, the elevated temperatures used in this method can often result in a superlubridicity-like motion of the graphene and TMD crystals adhered to the h-BN. Regularly, I have observed the crystals attached to the h-BN completely re-orient themselves, often ruining the process. This has even occurred during the final set-down of the heterostructure, appearing to happen more frequently at high temperatures.

Finally, the likelihood of interfacial bubble formation is still too high with this process, which introduces a large variability between heterostructures and inhibits the creation of more complex structures. For this reason, after completing the work documented in chapter 8 I went back to the drawing board to try and improve my transfer techniques. This has led me to a PPC-stamp-based technique which I believe is likely quite similar to the one used by Ref. [125], however there are some nuanced details in the process which I discovered that I will discuss.
7.3.4 PPC-based stamp procedures

Initially when performing the van der Waals transfer with PPC-based stamps, I had quite a number of issues, generally revolving around being unable to reliably pick up $h$-BN crystals. Obviously, being unable to pick up the top crystal of the heterostructure is major issue for the entire process. Recently, though, I have found that switching to mixing the PPC solution as described above seems to have made the entire process much more reliable, largely replacing the PC-based stamps in my work. Another modification I have found useful is performing a hot-plate anneal of the finished stamp prior to use at around 70-100 °C for at least a few minutes and up to an hour or more.

PPC stamps tend to be very sticky between 40 and 60 °C. Generally, I find myself doing every pick up and set down step at around 45 °C, now. At this temperature, the stamp actually yields extremely jumpy motion when used normally, due to the level of adhesion to the chip surfaces. However, I have found that by very carefully nudging the stamp downward, using the stepper motor of the stamp stage in small increments and waiting for the stamp to relax, the contact front will creep forward very slowly. In this way, it is possible to bring together the different layers of flakes over the course of tens of minutes, instead of tens of seconds, as before.

Despite this advancement in the control over the stamp, it can still be quite difficult to consistently assemble the different layers of a heterostructure without encountering any interfacial bubbles. This can be due to defects in the stamp itself, or often very thick flakes away from the crystals of interest can cause the stamp to set down in unexpected ways. Taking influence from the work of Purdie et al [143], however, I have found that it is actually possible to continuously manipulate these interfacial bubbles throughout the assembly process, both at each step of the assembly and during the final set down of the heterostructure. Interestingly, whereas Purdie et al predict the manipulation works best over 180 °C (which would destroy PPC), I have found it is possible to manipulate the bubbles even at 40 °C.

As an example, Fig. 7-7 shows manipulation of interfacial bubbles which formed in the first set-down of the top $h$-BN onto monolayer graphene. Fig. 7-7 displays the sequence
of picking up the graphene with $h$-BN, and the four images were collected over the course of about fifteen minutes. In panel one, numerous contamination bubbles can be observed, circled in red. As the contact front is pulled to the left (picking up the crystals), the edge of the contact front may be seen pushing the interfacial bubbles along. The bubbles may even be seen to coalesce into larger bubbles as the contact front progresses. When the bubbles reach the edge of the graphene-$h$-BN overlap, I have noticed that they often simply disappear optically, such as the bubble circled at the top edge of panel one which is no longer present at the end.
The ability to clean the heterostructures assembled on the stamp is a major boon to creating very high quality samples. In this way, at each step of the process the layers may be insured to be without major contamination bubbles. I have noted that it is generally easier to remove the bubbles during the picking up step, rather than the set-down, likely because the contact front is slightly more controlled. For very stubborn bubbles, it is possible to move the contact front back and forth over them, gradually moving them through the stack.

The final encapsulated graphene heterostructure formed from the in-process pictures in Fig. 7-7 is shown in Fig. 7-8. As a result of the cleaning shown in Fig. 7-7, the final heterostructure had an imperceptible level of contamination, with essentially all of the encapsulated graphene being available for device fabrication.

I have also noted that it is possible to perform the cleaning step on the final heterostructure, as is shown in Fig. 7-9 for an encapsulated graphene sample. Shortly after the assembly of this heterostructure, but prior to the final set-down step, I observed numerous interfacial bubbles which spontaneously nucleated nearly everywhere throughout the heterostructure, which can be seen in the top left panel of Fig. 7-9. When setting down this heterostructure at 90 °C onto the final substrate, I found that I was able to push a great deal of the interfacial bubbles from the stack, leaving two completely clean areas of encapsulated graphene. There were some localized interfacial bubbles left behind, which could not be removed, which I believe were due to adhesive residue from the exfoliation of the bottom h-BN. However, the clean areas areas in this case are significantly larger than what is required for making graphene devices on the split-gates shown.

An interesting note for PPC-based stamps, which I have not found reported elsewhere, is that the stamps reported here are re-usable. When heterostructures are set down between 90 and 110 °C, the PPC will often release the heterostructure onto the substrate, rather than needing to be melted down. This provides multiple advantages over the typical steps where the PPC (or PC) needs to be melted down to the substrate and subsequently removed with solvents. For example, my co-worker Prathamesh Dhakras and I have utilized this method to rapidly transfer single h-BN flakes for use in his CNT experiments. It is also very nice to re-use stamps which show good performance, allowing very consistent results across the fabrication of several heterostructures. Even after several uses, if a stamp is becoming
Figure 7-8: Example of a pristine encapsulated graphene sample resulting from the interfacial cleaning technique. This heterostructure was assembled using the PPC-based stamp method described in the main text. The cleaning procedure was used to remove any interfacial bubbles which formed throughout the assembly, resulting in a bubble-free heterostructure.
Figure 7-9: Example of interfacial cleaning of a completed heterostructure during the final set-down. Demonstration of how interfacial bubbles which occur after the final encapsulation step can be cleaned. Numerous bubbles can be observed in the upper left panel, which are swept through the heterostructure as the contact front moves across, transferring the heterostructure to a split-gate substrate. The lower panel shows an optical image of the completed heterostructure, where two clean areas are shown between the trapped bubbles (which are likely due to tape residue on the lower h-BN).
roughened (for example from being in contact with bulk graphite flakes), annealing it on a hot-plate around 90 °C seems to reflow the PPC without changing its properties significantly.

7.4 General fabrication procedures

The majority of the work involved with fabricating the devices in this thesis was spent in the assembly of the heterostructures themselves, however the processing of the heterostructures is just as important as the earlier assembly in order to get working electrical devices. Final heterostructures go through a process of patterning and etching into their final shape, as well as metallization to create electrical contacts. I have developed a number of recipes for these steps which have worked very consistently over time.

7.4.1 A note on substrates

All of the samples present in this dissertation are for heterostructures which were placed onto buried split-gate structures which were created in CNSE’s 300 mm chip fabrication line. I was extremely lucky to have access to these structures, which were created through no small effort of the graduate students prior to me in the research group. The process to create them, which have been described in detail elsewhere \cite{19}, produces polysilicon split-gates buried underneath 100 nm of SiO\(_2\) as a gate dielectric. Using chemical mechanical planarization techniques, the resulting gate SiO\(_2\) is extremely smooth, with no noticeable topography even between the edges of two split-gates.

I have fabricated samples on three different types of structures which are placed on these wafers- three-gated regions (which allow making reconfigurable FETs and other structures) and two different types of two-gated regions (2 micrometer wide interdigitated gates and 15 micron wide gates). The two-gated structures are primarily used in measuring \(p-n\) junctions, but I have also found when carefully applied, can be used to measure individual Schottky contacts (the primary focus of chapter \[8\]). The split-gates come in a variety of spacings, down to 100 nm. I have primarily used the 100 and 300 nm split-gate spacings in my work.
7.4.2 Electron beam lithography

Fabricating van der Waals heterostructures into electrical devices requires multiple different types of patterning. First, we need to pattern the structure into its final shape, and second, we need to apply metal contacts to the structure. To create the patterns for each of these steps, electron beam lithography is very convenient, as it allows us to create features down to approximately 5 nm and we are able to create completely custom patterns for each step, rather than being limited by pre-defined optical lithography masks. In late 2017, Professor Lee acquired a Raith Voyager 50 kV electron beam lithography system for the lab, which I have gotten the opportunity to use extensively in my work.

In electron beam lithography we use an electron beam to selectively expose an electron-sensitive resist, in this dissertation a positive-tone resist (PMMA 950K 4% in anisole from Microchem), which means that the exposed areas will be removed when developed. The remaining resist may be used as a mask for different purposes, depending on the particular step. For example, in this dissertation I have used PMMA as an etch mask for defining the heterostructure geometry and as a metal-lift off mask for defining metal contacts (both to be discussed, later).

The specific processing parameters used in electron beam lithography are very important. Generally, I spin-coat PMMA A4 at 3000 rpm, which gives about a 200 nm thick layer of resist on my sample. The resist is exposed in the Voyager using a beam current between \( 0.4 \) and 10 nA (smaller currents are used for smaller features) and a dose between 300 and 800 \( \mu \)C/cm\(^2\). When using PMMA as an etch mask, since the exposed areas will be developed away, we expose only the areas we wish to be removed in the etch. Since this means we are generally exposing a large area around the outside of the region we want to protect with the resist, this means we need to use relatively lower doses. For metallization, we expose where we want metal to remain in the final process—this, coupled with the fact that we want to ensure complete removal of the resist, means we will generally use higher doses for metallization steps.

To develop the PMMA resist I submerge the sample in a methyl isobutyl ketone:IPA solution in a ratio of 1:3 and agitate it for one minute. For the feature sizes typically used
in this dissertation (usually greater than 200 nm), no special development techniques are required, however consistency of the development is very important when dialing in the correct dose for different structures. After the development, the sample is rinsed with IPA and blown dry with nitrogen.

Layer-to-layer alignment is a challenging part of electron beam lithography- while the Voyager tool is able to assist a great deal in this process, it does not completely solve the problem for us. In order to do layer-to-layer alignment, there must be some pre-defined features on your sample with which to align to, called alignment marks. An important feature of good alignment marks is that they are able to be detected by the electron beam of the lithography system. Unfortunately, the “GR” variety of wafer which I generally use, has marks which are very difficult to detect inside the Voyager. The wafer primarily consists of buried polysilicon features and etched features. Underneath PMMA, the polysilicon is completely impossible to see, while certain etched features may be observed with a keen eye and careful selection of brightness/contrast settings.

The first step of alignment is bringing the heterostructures which we have transferred into the GDS file (where we define our patterns for the lithography tool) and aligning them to the gate structures created in the fab. To do this, high magnification optical images are taken of the heterostructures, with identifying features (such as the corners of the split-gates) in the field of view. These images are then imported into the GDS file and carefully re-sized and rotated such that the pre-defined features in the image line up well with the GDS file. From there, we can trace around the flake and design the final device. With this method we can reliably pattern on van der Waals heterostructures with sub-micrometer accuracy for specific features within the heterostructure. In the case of large area graphene samples, this step can be done in a rather careless way, as usually the device will be a very small area cut out of a much larger area of graphene. However, in the graphene-TMD hybrid structures which I will discuss later, this step must be done very carefully.

With the pattern properly located in the GDS file, the next step is aligning the lithography tool to the sample itself. While the Voyager has various auto-alignment features which can align to alignment marks with great accuracy, we cannot use those features on the “GR” wafers. Instead, we perform what is referred to as a “three point alignment”, where we
physically locate three features on the corner of our sample and tag them in the GDS file. When done properly, this lets us accurately drive to different locations on our sample simply by clicking on the GDS file. On the “GR” wafer, there are several etched features (such as bar code structures) on the perimeter of the sample, which work well for this step. With this method, we can confidently achieve layer-to-layer alignment with less (often much less) than 200 nm error, which is a testament to both the quality of the Voyager stage and the lithography tolerances used in manufacturing the base wafers.

7.4.3 Heterostructure etching

Etching the heterostructure does two important things in the overall process. First, it lets us shape the heterostructure however we like. For example, if we picked up a flake consisting of multiple terraces of different thickness, we can cut away any undesired regions. We can also form interesting shapes such as constrictions and Hall bars. Simultaneously, etching can be used to expose the edges of any encapsulated flakes, particularly graphene, for making one-dimensional edge contacts.

The etch is performed in a Technics 800 MicroRIE (the same as we perform plasma cleaning in) using an SF$_6$/O$_2$ plasma. On the MicroRIE, we can control the flow rate of the oxygen with a mass flow controller, but not the SF$_6$, which is connected to a needle valve. The only way to control the SF$_6$ is by measuring the pressure in the chamber, which is not ideal. Despite this, the process is quite reliable. Generally, I set the chamber pressure to approximately 220 mTorr with just SF$_6$, and then add 2 sccm of O$_2$. Recently, I have also experimented with going up to 4 sccm of O$_2$, which works about the same. The plasma strikes at 30 W on the RF power source, which works well for the whole etch.

SF$_6$ has become my preferred etch gas as it is extremely selective to etching h-BN compared to SiO$_2$ and the PMMA mask. Usually, an etch of a minute is enough to completely etch through a heterostructure, although I have gone up to two minutes of etching without any problems (such as etching through the SiO$_2$ or PMMA) and the etch could likely be run for even longer. Aside from the risk of over etching the SiO$_2$ and PMMA, another concern is hardening the PMMA, such that it cannot be easily removed with acetone.

Resist hardening was a major problem I encountered when I first performed the het-
erostructure etching using CF$_4$ instead of SF$_6$. Switching to SF$_6$ largely fixed this problem, but to ensure good removal of the resist in the case of hardening, I have also switched to stripping the resist with Microchem’s PG remover. Generally, if the resist is hardened and not removed by the first solvent you put the sample in, the hardened portion will become stuck to the sample, ruining it!

I have found the SF$_6$/O$_2$ etching process to produce high-quality edge-contacts with high yield. Nearly 100% of edge contacts to graphene are functioning in final samples. For contacts which are found to be dead, usually it is found to be that the graphene was not actually where I thought it was in the stack, not that the process itself failed. Additionally, working contacts have only died during measurements once or twice over the course of hundreds of electrical measurements. One likely contributor to this high yield is that the etch produces fairly sloped profiles on the edges of the heterostructure, for example the angle of the etch measured from the plane of the substrate is about 50° in Fig. 7-10. For comparison, failed contacts were the primary contributor to poor yield for my earlier, CVD graphene-based samples.

### 7.4.4 Metallization

For creating contacts I use a metal lift-off process. After the etch step, the resist will be stripped with PG remover and then another layer of PMMA will be applied the same as before. For metal lift-off with PMMA, wherever the resist is exposed, metal will be left behind after the final processing. Generally, I design the patterns to give as much overlap of the metal on the edges of the heterostructure as possible, to give a large tolerance for any patterning misalignment. A cross-sectional schematic depicting the edge-contact process is shown in Fig. 7-11.

The metal lift-off process is summarized in Fig. 7-12. After standard electron beam lithography as described above, a metal stack of 3/15/60 nm Cr/Pd/Au is evaporated with electron beam evaporation. The Cr/Pd/Au stack works very well for edge contacts to graphene and is quite durable for probing. I usually aim for evaporation rates of 1 Å/s or greater. Additionally, especially with chromium, I will evaporate up to 20 nm of metal prior to performing the actual deposition, to ensure (for example) the deposition is chromium and
Figure 7-10: Example of an etched $h$-BN encapsulated graphene sample. The top panel shows an AFM topography image of a graphene-based heterostructure which has been etched using the SF$_6$/O$_2$ recipe described within the main text. The etch produces a sloped side-wall, which likely helps in the edge-contact process. The lower panel shows a line cut from the top image, highlighting the slope of the sidewall.
not some chromium oxide.

Since electron beam evaporation is very anisotropic in its deposition of metal, trenches in the PMMA will be filled without coating the side-wall of the developed resist. Therefore, when we submerge the sample in boiling acetone (usually around 85 °C), the resist will readily be washed away, taking unwanted metal with it. Often, the metal will begin to remove in a nearly complete sheet almost immediately upon the sample entering the acetone, however I prefer to agitate the solution by sucking up acetone in a pipette and spraying it down onto the sample. In this way, I have found I can reliably lift off even very complex structures, such as Hall bars with many closely spaced (sub-micrometer spacing) contacts. To improve yield, though, I have found it is helpful to avoid closed loops of metal (or configurations which are close to closed loops) during the design process.

### 7.5 Electrical measurements

The electrical measurements presented in this dissertation have largely been performed inside of a Lakeshore CPX-VF cryogenic vacuum probe station. The probe station may be cooled using either liquid nitrogen or liquid helium, and is also equipped with a hot stage which
I used to perform the Schottky barrier height extraction measurements demonstrated in chapter 8. The probe station has six micromanipulation probes which can be used for both biasing gates and performing measurements.

To perform the measurements, I have primarily used an HP4156B semiconductor parameter analyzer. The HP4156B is equipped with 2 voltage source units (VSUs) which go up to 20V and 4 source/measure units (SMUs) which can source over 40V and simultaneously measure currents down to a 1 fA. The HP4156B was made prior to the days of USB cables, thus for controlling the unit and getting data from it, I drove the unit via GPIB commands sent through a home-built LabView program. This gave some added flexibility of being able to set up two parameter sweeps for an arbitrary number of data points (whereas the HP4156B usually is limited to very few points), which I used in the four-quadrant $p$-$n$ junction maps which will be shown later in this chapter.
7.6 Transport case studies

I have studied two major types of devices utilizing the van der Waals transfer process. The first device is the graphene $p$-$n$ junction (and just graphene samples in general). Graphene devices provide a relatively simple benchmark of the quality of heterostructure which can be produced with a given technique, which was useful whilst learning the van der Waals transfer. The second type of device which I have focused on is the hybrid graphene/WSe$_2$ structure where edge-contacted graphene is used as an intermediary contact to WSe$_2$. This device type can be re-configured using split-gates to behave as a $p$-$n$ diode, a Schottky diode (discussed heavily in chapter 8), and either an $n$-type or $p$-type MOSFET. Note that from this point forward, I will be switching back to the pedagogical third-person tone used throughout the rest of the dissertation.

7.6.1 Graphene $p$-$n$ junction device

As we discussed in chapter 2, graphene $p$-$n$ junctions exhibit a conductance asymmetry compared to unipolar graphene due to angle-dependent transport. For an abrupt junction, we saw that $G_{pm} = \frac{2}{3}G_{nn}$, and for a graded junction we predicted a significantly greater level of asymmetry due to enhanced filtering of higher-angled modes, as was demonstrated in Fig. 2-9. However, experimental measurements on graphene samples typically have not shown the level of conductance asymmetry predicted theoretically. Indeed, despite the fact that experimental junctions are smoothly varying, it is quite challenging to reach even the level of asymmetry predicted for the abrupt junction. As these conductance asymmetry calculations are generally based upon the assumption of ballistic transport, one likely culprit are scattering effects inducing diffusive transport in graphene. Another likely problem is contact resistance, which could obscure the intrinsic properties of a graphene $p$-$n$ junction.

Fig. 7-13 shows an optical image of an encapsulated graphene sample placed on inter-digitated buried split-gates with a 100 nm spacing. The heterostructure was assembled using a PPC stamp with the cleaning techniques described above, resulting in a large area of clean encapsulated graphene on the left side of the stack. A six-terminal hall bar was patterned onto the stack, straddling two neighboring split-gates, as is shown in the schematic on the
Figure 7-13: Heterostructure and layout of a high quality graphene p-n junction device. The left panel shows an optical image of an encapsulated graphene sample created using the PPC-based stamp technique described above. The left side of the heterostructure has a large area of clean graphene in which a six-terminal Hall bar was created, with the Hall probes straddling two split-gates. The schematic of the Hall bar is depicted on the upper right panel, which forms a $1 \times 1$ micrometer channel. The lower right panel shows an optical image of the finished sample, with the blue arrow pointing towards the device studied here.
In such a structure we can perform four terminal measurements which let us remove the resistance of our measurement leads and, most importantly, the contact resistance. Specifically, the four-terminal resistance of the area between the hall probes can be measured by driving a current between the contacts labeled 1 and 4 and measuring the voltage between contacts 2 and 3 (or 5 and 6). Furthermore, using the buried split-gates, we can independently vary the carrier density on each side of the sample, enabling the study of p-n junctions.

Fig. 7-14a shows a four-quadrant measurement of the four-terminal resistance as a function of $V_{G1}$ and $V_{G2}$, where the four quadrants indicate different types of junction configurations ($p$-$n$, $n$-$p$, $p$-$p'$, and $n$-$n'$). The measurements were all taken at $T=300$K with a small bias (50 mV) applied between contacts 1 and 4 and the current measured between contacts 2 and 3. Both gate voltages have been corrected with respect to the Dirac point voltage, such that positive gate voltages indicate $n$-type graphene and negative gate voltages are $p$-type graphene. The four-terminal resistance measurement is almost perfectly symmetric with respect to the Dirac point, and we note a large conduction asymmetry between the ambipolar and unipolar quadrants.

If we take a slice along $V_{G1} = V_{G2}$, we get the true transfer curve of the device (the red curve in Fig. 7-14b), tuning the carrier density uniformly through the entire device. Again, the resistance is very symmetric about the Dirac point, which is an indication of little electrostatic disorder or contact effects in the measurement. Similarly, if we take a slice at $V_{G1} = -V_{G2}$, we see the effect of the symmetrically doped p-n (or n-p) junction.

Primarily, though, we are interested in seeing the conduction asymmetry, which is plotted in Fig. 7-14c. We observe a very high degree of asymmetry, with $\sigma_{pn} \approx 0.175\sigma_{nn}$ for the largest gate voltages. To compare, we perform tight-binding simulations for a symmetric p-n junction with a 100 nm junction width (the black line in Fig. 7-14c). For simplicity, these simulations were performed assuming zero temperature, despite this, they agree quite well with the experimental measurements.

The agreement between the experiment and the purely-ballistic tight-binding transport modeling suggests that we observe nearly ballistic transport at room temperature. To understand this phenomena a bit further, we adopt a simple Drude model for graphene which
Figure 7-14: Transport measurements on a high quality graphene p-n junction

(A) Room-temperature four-quadrant p-n junction map of the four-terminal resistance measured on the device described in Fig. 7-13. The two split-gates are swept independently to configure the device as a p-n, n-n', p-p', and n-p junction, depicting a large asymmetry between the ambipolar and unipolar quadrants. The measurement has been Dirac point corrected, such that the maximum resistance point corresponds to zero volts on each gate.

(B) Slices through the four-quadrant map. The red slice corresponds to the channel with a uniform carrier density and the blue slice corresponds to a symmetric p-n junction configuration.

(C) Conductance asymmetry of the p-n junction compared to the unipolar junction. The red and blue slices correspond to different quadrants of the experimental measurements and the black line corresponds to a zero-temperature tight-binding simulation for a junction with $D_W = 100$ nm.
defines the mobility as $\mu = \frac{\sigma}{n_e}$ with $n$, the carrier density \cite{85}. Also, within the Drude model, we can define the mean scattering length as $l = \frac{\hbar}{2e\sigma(\frac{1}{\sqrt{n}})}$. Each of these parameters is extracted from the transfer curve slice (in Fig. 7-14b) and shown in Fig. 7-15. To estimate the carrier density as a function of gate voltage, we assume a simple geometric capacitance model for the split-gate-dielectric-graphene stack, similar to what was discussed in chapter 4. Our extracted mobility is approaching the acoustic phonon-dominated limit at room temperature \cite{85}, resulting in a mean free path nearing 450 nm at the highest carrier densities we study.

As the mean free path we extract is comparable to half the distance between our Hall probes, it is feasible that we are approaching the ballistic limit, even at room temperature. This result corroborates the high level of conductance assymmetry which is shown in Fig. 7-14, suggesting we are observing nearly-ballistic angle-dependent transport from the graded junction. Additionally, these results provide evidence that the fabrication techniques described in this chapter are capable of producing high quality samples, where study of fundamental device physics is possible.
7.6.2 Graphene-WSe\textsubscript{2} hybrid devices

The other major type of device which we study in this thesis is a hybrid device in which graphene is used as an intermediary contact to WSe\textsubscript{2}. Two different methods for fabricating this device have been used in this dissertation and are summarized in Fig. 7-16. At this point, it is also worth noting that all of the Graphene-WSe\textsubscript{2} devices in this dissertation were created using the PC-based stamp procedures described earlier.

The first method involves using $h$-BN to pick up two strips of graphene with some separation between them. Often in the graphene exfoliation, we observe areas with multiple parallel strips of graphene, which work well for this technique. Then, a WSe\textsubscript{2} crystal is picked up, stretching between the two strips of graphene, and finally the entire structure is sandwiched by a bottom $h$-BN crystal. This stack is then transferred to split-gate structures, where the key detail is that the graphene strips must run parallel to the gates and not cross them. In this way, when the ends of the graphene strips are metallized, it is possible to measure a graphene-WSe\textsubscript{2}-graphene structure across the gates.

The first method, using two graphene strips, requires a very high degree of precision for alignment over the split-gates. Often, once the sandwich is fully formed, the graphene and WSe\textsubscript{2} crystals will become nearly invisible within the stack, meaning the final alignment must be done nearly blind, using reference images from the assembly procedure. Furthermore, the gap between the graphene and the size of the WSe\textsubscript{2} crystal must be carefully matched. For this reason, a second method has been adopted, shown in the bottom half of Fig. 7-16.

In the second method, instead of picking up strips of graphene, one large graphene is picked up such that it covers half of the top $h$-BN crystal. A WSe\textsubscript{2} crystal is then picked up, such that the graphene partially overlaps it. The sandwich is then finished as before, and transferred to a split-gate sample. This time, however, the graphene does straddle the split-gates. To create the same graphene-WSe\textsubscript{2}-graphene structure which we want to study, in this case, we just etch away the portion of graphene which would short the channel.

There are pros and cons to each method of forming the Graphene-WSe\textsubscript{2} hybrid device. The multiple graphene strips method allows us to verify the quality of the graphene in the sample by measuring using the contacts on just one strip of graphene at a time. The devices
Figure 7-16: Two methods of assembling graphene-contacted WSe$_2$ devices encapsulated in h-BN Two methods are used throughout this dissertation to fabricate hybrid graphene/WSe$_2$ devices. The first method (top panel) involves picking up two separate strips of graphene which are aligned such that they straddle the split-gates in the final device. The second method (bottom panel) involves picking up a single piece of graphene, which is eventually etched into two pieces straddling the split-gates. In each process, the etch step simultaneously defines the channel and exposes the edge of the graphene for forming one-dimensional edge contacts.
Figure 7-17: Two completed graphene-WSe$_2$ hybrid devices using the two different process flows in Fig. 7-16. (A) An example device which will be studied later in this section fabricated using the two graphene strip method and placed on a three-split-gate substrate. (B) A device fabricated using the single graphene approach which will be studied heavily in chapter 8.

fabricated with multiple strips of graphene also appear to exhibit ambipolar conduction more often than the single graphene devices (which tend to be predominantly $p$-type), although underlying mechanism of this effect still requires further study. The single-graphene method, however, is significantly easier to fabricate samples with, giving higher yield of working devices and more easily reproducible results. Examples of devices fabricated using these two methods are shown in Fig. 7-17, where the left panel is using the multiple graphene strips method (for which data will be shown in the rest of this section) and the right panel uses a single graphene strip (which will be the focus of chapter 8).

The device shown in 7-17a is built on top of a triple split-gate structure, where the outer gates are 15 $\mu$m wide and the middle gate is 1 $\mu$m wide, with a 300 nm spacing between the gates. By tying all three gates to the same voltage and applying a small bias, we can measure the transfer characteristic of this device, which is shown in Fig. 7-18a. The device
displays ambipolar characteristics, with an on/off ratio of about $10^8$. The minimum current point for the sample is at around -2.5 V applied to the three gates, and actually drops below the minimum level of current which we can measure.

Additionally, the $p$- and $n$-branches of the transfer curve exhibit an interesting effect where we observe two distinct regimes for each branch. Around the minimum current point the slope of the transfer curve is very steep, around 100 mV/decade of current, and then at around a current level of 1 nA, there is a sharp kink (on a log-scale) in the current. This effect was the first indication that we were likely measuring a tunable Schottky barrier formed between the graphene and WSe$_2$.

The presence of ambipolar transport in the device transfer curve means that we can electrically reconfigure it to behave as a $p$-$n$ junction. As this sample was built on a three gate structure, to form a $p$-$n$ junction we tie the left and middle gates together to one voltage ($V_{G1} = V_{G2} = -6.25$ V) and the right gate to another voltage ($V_{G3} = 6.25$ V), using the transfer characteristics as a guide for the selection of voltages. Now, with fixed gate voltages we perform a drain voltage sweep (between the two bottom contacts straddling the gates) to measure the $I_D$-$V_D$ characteristic of the $p$-$n$ junction, shown in Fig. 7-18b.

In contrast to the graphene $p$-$n$ junction, semiconducting $p$-$n$ junctions behave as rectifying diodes. To analyze the $I_D$-$V_D$ measurement of the $p$-$n$ junction, we fit the data to the diode equation,

$$I_D = I_0 \left[ e^{qV_D/nk_BT} - 1 \right], \quad (7.1)$$

where $I_0$, the leakage current of the diode, and $n$, the diode ideality factor, are used as fitting parameters. When $n = 1$, the diode is considered ideal ($n > 1$ is non-ideal). As the basic diode equation does not include any extra effects, such as series resistance roll-off (a phenomenon that occurs at large forward biases), we only apply the fit to the exponential portion of the forward bias of our measurement, between around -0.25V and 0V on the drain. We are able to fit the forward bias current over about four decades of current, resulting in an ideality factor of $n = 1.26$, which is close to ideal. In this case the diode equation does not perfectly match the leakage current, which appears to be due to some reverse bias effect such as Zener tunneling, and is a topic of future work.
Figure 7-18: Electrical characterization of a reconfigurable graphene-WSe$_2$ device. Electrical measurements on the device depicted in Fig. 7-17a. (A) The device exhibited an ambipolar transfer curve with a steep subthreshold region near the minimum current point. For this measurement all three gates were held together. By tying the gates to separate voltages, as summarized in each panel, the device may be reconfigured as an (B) nearly-ideal p-n diode, (C) a Schottky diode, and (D) both n- and p-type MOSFETs.
It is also possible to configure this device as a Schottky diode, which will be the major topic of chapter 8. The Schottky diode configuration is similar to the p-n junction, where two neighboring gates are tied together, except this time all the gates of the device are all the same polarity. In Fig. 7-18c, we show the $I_D-V_D$ characteristic of the Schottky diode when the gates are configured as $V_{G1} = V_{G2} = 0.1$ V and $V_{G3} = 10$ V.

Again, we use the transfer curve as a guide, tying the left side of the device to a voltage along the steep portion of the n-branch of the transfer curve and the right side of the device to the quasi-saturating portion. This configuration makes the left graphene-WSe$_2$ junction into an n-type Schottky barrier and the right graphene-WSe$_2$ junction into an Ohmic contact. In this way, the $I_D-V_D$ characteristic will be limited by the Schottky contact, resulting in the diode-like behavior in Fig. 7-18c.

As the diode equation is not specific to p-n junction diodes, we use it to fit this data in a similar way, resulting in an ideality factor of $n = 1.26$. The identical ideality factor to the p-n diode happens to be purely coincidental, as the physical mechanisms which determine the ideality factor for the Schottky diode and p-n diode are not the same. In this case, however, we see fantastic agreement between the measured leakage current of the diode and the fit, giving a leakage current of $I_0 \approx 3 \times 10^{-12}$ A. The leakage current of this Schottky junction matches the diode equation in a way which is rarely observed, a result which is deeply connected with the tunability we have found the graphene-WSe$_2$ junction to exhibit, and will be explained in detail later on.

Finally, we can reconfigure the three gated structure as either an n or p-type MOSFET. To perform a transfer curve on the MOSFET, this time, the outer gates are tied to a large voltage ($V_{G1} = V_{G3} = -10$ V for the p-FET and $V_{G1} = V_{G3} = +10$ V for the n-FET) and the middle gate is swept to turn the FET on and off. FET transfer curves are shown in Fig. 7-18d. In this case, there should be a small disclaimer that the FET measurements here were performed at the very end of the lifespan of the device, after many days of measurements (as well as several temperature cycles). This degradation is responsible for the different current levels compared to the other measurements in Fig. 7-18. Despite this, Fig. 7-18d serves as a proof of concept for a WSe$_2$ FET with gate-tunable graphene contacts, displaying clear reconfigurable FET behavior. Additionally, even though the device was in a decaying state...
at this point, the subthreshold slope was around 100 mV/decade, which is beginning to approach the 60 mV/decade limit.

To wrap up this discussion, we have seen how the graphene-contacted WSe$_2$ encapsulated in $h$-BN can be used as an electrically configurable device, supporting nearly-ideal $p$-$n$ junction diodes, nearly-ideal Schottky diodes, and ambipolar MOSFETs (with decent subthreshold performance, even in a decaying state). This result is interesting because the ideality factors of the diodes and the subthreshold slope inform us about completely different aspects of the sample. The ideality factor of the $p$-$n$ junction can inform us about trap states inside the band gap of the semiconductor- an indication of the quality of the crystal. The ideality factor of the Schottky diode tells us about the quality of the metal-semiconductor interfaces in the sample. Additionally, the subthreshold slope of the MOSFET can tell us a lot about traps at the interface between the semiconductor channel and the dielectric. Together, these measurements let us fully characterize the electronic properties of the device, the quality of the WSe$_2$, and even learn about the quality of the substrate.
Chapter 8

Gate-tunable graphene-WSe$_2$ ideal Schottky diodes at the Schottky-Mott limit

8.1 Introduction

The metal-semiconductor interface is one of the most fundamental components of electronic devices. Indeed, to perform an electronic measurement on a semiconductor, inevitably, one must contact the semiconductor with some metal conductor. The metal-semiconductor interface is, simply, unavoidable. For this reason, the metal-semiconductor interface has garnered quite a bit of interest, going back to around the time of the birth of semiconductor electronics [158].

Metal-semiconductor interfaces, most fundamentally, are characterized by an energy barrier which forms between the metal and the semiconductor, referred to as a Schottky barrier. Fig. 8-1 demonstrates the energy band diagram for the classic textbook example of the formation of an $n$-type Schottky barrier. In Fig. 8-1a, we begin with a metal and an $n$-type semiconductor which are separated from one another. The metal is characterized by a work function, $\phi_m$, which denotes the separation between the highest filled level of the metal and the vacuum energy level and the semiconductor is characterized by an electron affinity $\chi$,
which is the conduction band energy relative to the vacuum level.

When we bring the metal and semiconductor together, charge transfer will occur between the two materials to equilibrate their Fermi energies to the same level. This results in local band bending in the semiconductor at the metal-semiconductor interface. As a result of this equilibration, an energy barrier forms, and in an ideal case the energy barrier will be given by

$$ q\Phi_B = q(\phi_M - \chi). $$

(8.1)

Similarly, a $p$-type Schottky barrier is described by

$$ q\Phi_B = E_G + q(\chi - \phi_M). $$

(8.2)

In an ideal case, which is referred to as the “Schottky-Mott limit”, both eq. (8.1) and eq. (8.2) hold. The Schottky-Mott rule tells us that changing the work function of the metal that there should be a one-to-one change in the Schottky barrier height. However, from the earliest days of semiconductor research, it has been well documented that interfacial traps at

Figure 8-1: Textbook-style example of the formation of an $n$-type Schottky diode which follows the Schottky-Mott rule (A) The energy band diagram of a metal and semiconductor separated from one another. (B) The metal-semiconductor junction band diagram after bringing the two materials together and allowing them to equilibrate, resulting in a Schottky barrier.
the metal-semiconductor interface cause the breakdown of this rule \[159\]. Interfacial traps act to pin the Fermi level at the metal-semiconductor interface, resulting in a Schottky barrier height which is challenging to predict \[160\]. Even in extremely high quality Schottky junctions formed on GaAs substrates \[161\], with diode ideality factors averaging \(n \approx 1.02\), very little correlation was shown between the metal work function and the Schottky barrier height \[162\]. Myburg et al went as far as fabricating 43 sets of samples with different metals, finding little relationship between the metal work function and the Schottky barrier heights they measured (a summary of their result is reproduced in Fig. 8-2 copyright Elsevier) \[162\].

2D materials offer an interesting opportunity for studying the Schottky-Mott limit. In contrast to three-dimensional materials, the surfaces of 2D materials are characterized by a lack of surface states and dangling bonds. The pristine surfaces of 2D materials have been predicted to exhibit a reduced level of Fermi level pinning effects, which could lead to observation of the Schottky-Mott limit \[151\][163\]. Recently, it has been demonstrated how using peeled metal contacts applied to 2D TMDs produces an interface which obeys the Schottky-Mott limit \[164\]. The rise of 2D metals and semimetals (like graphene) are likely candidates for attaining the Schottky-Mott limit, however experimental verification has remained illusive \[165\][166][167][168][169][170][171][150][164][172][173\].

Here, we will present measurements on graphene-WSe\(_2\) Schottky junctions which behave as nearly-ideal diodes. We have used a three-pronged approach to create a Schottky-Mott limited interface, where both the metal and the semiconductor have been replaced with 2D materials and then passivated in the 2D dielectric, \(h\)-BN. In the following sections, we will show how it is possible to rigidly tune the Schottky barrier by electrostatic gating of the WSe\(_2\).

Graphene, which has a gate-tunable work function \[174\], is a promising alternative to traditional bulk-metal contacts. Additionally, the ability to make edge-contact to \(h\)-BN encapsulated graphene \[125\] makes it ideal as an intermediary contact to encapsulated 2D materials, as we have seen in the previous chapter. As we intend to tune the Schottky barrier in our sample with a gate, rather than fabricating multiple samples with different metals, it
Figure 8-2: Demonstration of the effect of Fermi-level pinning on the Schottky barrier height from Myburg et al [162]. Plot of metal work function versus barrier height for metal-semiconductor junctions formed on GaAs substrates, reproduced from the work of Myburg et al [161], copyright Elsevier, 1998.
becomes necessary to re-define the Schottky-Mott limit. In our case, we define a new rule,

\[ \left| \frac{d\phi_B}{dV_G} \right| = S_G. \quad (8.3) \]

When the quantity \( S_G = 1 \), the gate voltage \( V_G \) rigidly tunes the Schottky barrier height, indicating that the junction is not pinned and is at the Schottky-Mott limit.

Fig. 8-3a shows the completed device which we will focus on for the majority of this chapter. It is a graphene-WSe\(_2\) device encapsulated in \( h \)-BN fabricated using the PC-based transfer technique described in the previous chapter. Specifically, this sample was created using the single-graphene pick-up method we described. An AFM image of the sample is shown in Fig. 8-4; the sample shows some interfacial contamination, but as it generally occurs away from the important interfaces and appeared in the final assembly steps, the important graphene-WSe\(_2\) interfaces are likely pristine.

The heterostructure is built on two split-gates which let us independently tune the carrier density on each side of the sample. A three-dimensional cross-sectional schematic is shown in Fig. 8-3b, depicting the different layers of the heterostructure and the poly-silicon gates buried in SiO\(_2\). Similarly to the last section of the previous chapter, the device may be put into a unipolar configuration with the two gates, where one graphene-WSe\(_2\) junction is made electrically transparent (Ohmic, with a large gate voltage), and the other gate (\( V_{G1} \)) makes the opposite graphene-WSe\(_2\) junction lightly gated, resulting in a Schottky barrier. \( V_{G1} \) may further be modulated to tune the Schottky barrier in that side of the device, described by the energy band diagram included in Fig. 8-3c.

In the following sections, we will perform a detailed electrical characterization of this sample. A series of measurements will be presented which build towards extracting the Schottky-Barrier height of the sample as a function of gate voltage, leading to realization of the Schottky-Mott limit. We will present a capacitance model which explains the relationship between the gate voltage and the Schottky barrier height. Finally, we will develop a dynamic diode tuning approach which is fundamentally built off the ability to rigidly tune the Schottky barrier height, enabling digital tuning of the diode ideality factor (including \( n < 1 \)).
Figure 8-3: Images and schematics of the Schottky barrier device studied in this chapter (A) Optical image of the graphene-WSe$_2$ device studied here, built on buried split-gates. This sample was fabricated using the single graphene strip method with PC-based stamps, described in chapter 7. (B) Cross-sectional 3D schematic of the device studied in this work. Buried poly-silicon split-gates are used to make one graphene-WSe$_2$ junction Ohmic so that we can tune and study a Schottky barrier in the other graphene-WSe$_2$ junction. (C) Energy band diagram of a gate-tunable graphene-WSe$_2$ Schottky junction. This figure is reproduced from LaGasse et al [5], copyright Wiley 2019.
Figure 8-4: AFM image of the device studied in this chapter The AFM image of this sample shows few trapped bubbles, which are generally far away from the important graphene-WSe$_2$ interface. This sample was fabricated using the PC-based method, which was used prior to the development of the PPC-based technique shown in chapter 7. This figure is reproduced from LaGasse et al. [5], copyright Wiley 2019.
8.2 Results and discussion

8.2.1 Electrical characterization

The transfer curve of the device which we will be discussing is shown in Fig. 8-5a, in which both gates are tied to the same voltage and swept together with a fixed drain bias of -0.1 V applied. The transfer curve is dual-swept (forward and back) and shows negligible levels of gate hysteresis, indicating a stable device, which is very important for the Schottky barrier extraction.

The device transfer curve shows predominantly $p$-type conduction, where the $p$-branch has a similar form to what we have seen in the device in the previous chapter. This time, however, the slope of the steep region of the $p$-branch is approaching the 60 mV/decade limit, the limit imposed by the Fermi-Dirac function for gate-switching at room temperature. Since transfer curve measurements are known to be the result of modulating the contacts of the device, this measurement strongly indicates that the 60 mV/decade slope is the result of modulating the Schottky barriers at the device contacts.

Using the transfer curve as a guide, we can study the Schottky barrier diode characteristics, shown in Fig. 8-5b. To configure a Schottky barrier in the left side of the device we tie $V_{G1}$ between 0 V and 1 V (lightly $p$-type, along the steep sub-threshold region) and $V_{G2}$ to -10 V (highly $p$-type). In this configuration, the diode shows nearly-ideal $I_D-V_D$ characteristics. As we saw in the device discussed at the end of chapter 7, the reverse bias current matches the diode equation very closely, with almost no modulation with respect to the drain bias. Furthermore, we see that the diode characteristics are strongly tuned by $V_{G1}$. By increasing $V_{G1}$, the forward bias current of the diode rigidly shifts towards larger drain voltages, where the shift is is almost exactly the increment of $V_{G1}$. As the diode characteristic shifts towards larger drain biases, the leakage current also rapidly drops below our measurement threshold.

Previously, our lab has shown that for well-behaved diodes, it is possible to fit the forward bias current of the $I_D-V_D$ curve to the diode equation and extrapolate the reverse bias leakage current, even below the minimum current we can measure \[175\]. In this case, the configurations for which we can measure the leakage current provide further verification of this technique, where just fitting the forward bias accurately predicts the leakage current.
Figure 8-5: Electrical characterization of the graphene-WSe$_2$ device (A) Transfer curve of the device from Fig. 8-3b, taken by tying both split-gates together and applying a -0.1 V drain bias. The device shows p-type conduction with a steep subthreshold region approaching the 60 mV/decade limit. Dual-sweeping the gate reveals very low hysteresis, a requirement for the battery of measurements performed in the rest of the chapter. (B) Gate-tunable diode measurements taken by tying the right gate to -10 V and stepping the left gate for voltages around the steep subthreshold region. (C) Extracted ideality factors and leakage currents from fitting the forward bias data in (B) to the diode equation. This figure is reproduced from LaGasse et al. [5], copyright Wiley 2019.
through the diode. Fig. 8-6 shows an example of the leakage current extraction method for the diode measurements in Fig. 8-5b.

Fig. 8-5c shows a summary of the diode fits for Fig. 8-5b as a function of $V_{G1}$. The diode ideality factor is around 1.1-1.2 over this range of gate voltages. What is most interesting, is that we extract over 13 decades of modulation of the leakage current for a 1 V change on $V_{G1}$, which is approaching the 60 mV/decade limit, similar to what we saw in the transfer curve. This data provided us further clues that we were performing a large modulation of the Schottky barrier in our sample, which warranted further measurement.

### 8.2.2 Schottky barrier height extraction

In order to work towards determining the Schottky barrier height as a function of the gate voltage, we adopt a 2D thermionic emission model which has been used in the past for 2D material-based Schottky junctions [166]:

$$I_D = AA^* T^{3/2} e^{-q\Phi_B/k_B T} \left[ e^{qV_D/nk_B T} - 1 \right]$$  \hfill (8.4)

with $A$, $A^*$, and $V_D$ being the effective junction area, the Richardson constant, and the drain voltage. If we reverse bias the junction, we can identify that the leakage current is simply $I_0 = AA^* T^{3/2} e^{-q\Phi_B/k_B T}$, which can be reduced to an Arrhenius-type relation from which we can determine the Schottky barrier height:

$$\ln \frac{I_0}{T^{3/2}} = \ln AA^* - \frac{q\Phi_B}{k_B T}.$$  \hfill (8.5)

Fig. 8-7a shows temperature-dependent measurements for the different values of $V_{G1}$, where the reverse bias leakage current is extracted from each measurement similar to as was done in Fig. 8-6. By doing a linear fit of $\ln \frac{I_0}{T^{3/2}}$ as a function of $1/k_B T$ to each set of measurements for a given $V_{G1}$, the Schottky barrier height may be attained directly from the slope of the curve, as is described in eq. 8.5.

The Schottky barrier heights extracted from the fitting performed in Fig. 8-7a are plotted in Fig. 8-7b as a function of $V_{G1}$. A linear fit of the data leads to $S_G \approx 1$, which indicates
Figure 8-6: Demonstration of diode fitting and leakage current extrapolation below the measurable current limit. To extract the leakage current when the diode is tuned into a large barrier configuration, we fit just the forward bias portion of the measurement. As fitting just the forward bias portion gives the fantastic agreement with the measurable leakage current configurations, we confidently extrapolate down below the range which we can measure in the lab. This figure is reproduced from LaGasse et al [5], copyright Wiley 2019.
that the Schottky barrier is being tuned at the Schottky-Mott limit, which is enabled by an absence of Fermi-level pinning effects at the graphene-WSe$_2$ interface.

Usually, graphene-WSe$_2$ Schottky barriers are tuned over a range of gate voltages which is quite large, such as over 100 volts as in Ref. [164]. In these cases, the Schottky barrier tuning is the result of modulating the work function of graphene with the gate. This modulation is actually quite a weak effect- for example, for the graphene device which we measured in the previous chapter, we only estimate about 0.1 eV change in the graphene work function. Generally, achieving a change of more than ±0.2 eV is quite difficult. Therefore, the level of tuning of the Schottky barrier we observe must not come from modulating the work function of the graphene.

8.2.3 Capacitance model

The large change in the Schottky barrier height which we observe in Fig. 8-7 cannot be explained just from modulating the graphene work function. For a 1 V change of the gate bias, the change in the graphene work function will be less than 75 meV, so we must consider other avenues by which the Schottky barrier can be tuned. When dealing with gated samples, understanding the capacitive coupling of the gate to the channel is quite important, so we will begin there.

We consider a model of the Graphene-WSe$_2$ interface as is shown in Fig. 8-8. This model is developed with the important assumption that the WSe$_2$ is at a low carrier density with its Fermi level within the band gap, similar to a conventional FET in the subthreshold regime. There are multiple voltages and capacitances in play which we must consider. The graphene contact is grounded and the opposite side of the device is tied to a drain voltage, $V_D$, which we assume rigidly couples to the potential of the channel, $\phi_S$.

We consider the capacitive coupling of gate one to the channel by the capacitance $C_{G1}$ with the gate voltage held at $V_{G1}$. Additionally, there is a parasitic capacitance, $C_{\text{Parasitic}}$, which represents any electric field lines from the gate which do not terminate themselves on the WSe$_2$ and instead extend out to the surrounding environment. One additional assumption which is key to our analysis is that there is a low interface trap density between the gate and the channel, which is reasonable given our passivation with $h$-BN.
Figure 8-7: Extraction of Schottky barrier heights as a function of gate voltage (A) Arrhenius plot for extracting the Schottky barrier height at different gate voltages. Each series of temperatures at each gate voltage is fit according to eq. (8.5), to give the Schottky barrier height. The Schottky diode $I_D-V_D$ curve was measured at different gate voltages for temperatures between 300 and 340K, and for each measurement the leakage current was extracted using the diode equation, with each measurement giving one point on the chart. (B) Schottky barrier height as a function of gate voltage, depicting perfect, Schottky-Mott limited barrier tuning by the gate voltage. This figure is reproduced from LaGasse et al [5], copyright Wiley 2019.
The drain and gate voltages compete for control over the channel potential (defined as \( \phi_S \)) and results in a voltage divider model for \( \phi_S \),

\[
\phi_S = \frac{C_G}{C_G + C_{\text{Parasitic}}} (V_{G1} - V_D) + V_D, \tag{8.6}
\]

where the \(-V_D\) term in \((V_{G1} - V_D)\) is because the effect of \(V_{G1}\) is rigidly offset relative to the channel as the drain voltage is changed. Since \(C_G >> C_{\text{Parasitic}}\), we can reduce eq. \ref{eq:8.6} to \( \phi_S = V_{G1} - V_D + V_D \), which clearly leads to

\[
\phi_S = V_{G1}. \tag{8.7}
\]

This analytical result tells us that in the low carrier density regime, \(V_{G1}\) directly determines the potential of the channel. Since the graphene-WSe\(_2\) interface is not pinned, this effect is responsible for the perfect barrier tuning presented in Fig. 8-7b.

Interestingly, it is possible that measurements with \(S_G < 1\) also represent metal-semiconductor interfaces which are at the Schottky-Mott limit, but are limited by poor control of their gate over the device channel. However, the same effects which reduce the gate efficiency seen in most graphene-WSe\(_2\) Schottky barrier measurements are likely to pin the Fermi level as well. In this sense, our rule is the most strict condition for verifying the Schottky-Mott limit for gated metal-semiconductor junctions, where \(S_G = 1\) unambiguously shows there is no Fermi level pinning at the junction.

\subsection*{8.2.4 Dynamic ideality factor tuning}

If we start from the thermionic emission model for 2D Schottky diodes, eq. \ref{eq:8.4}, and restrict our gate voltages to the Schottky-Mott-limited regime we have studied so far, based off our measurements, we can rewrite the Schottky barrier height as \( \Phi_B = \Phi_0 + V_{G1} \), with \( \Phi_0 \) being the Schottky barrier height at \( V_{G1} = 0 \) V. By substituting this into eq. \ref{eq:8.4}, we arrive at

\[
I_D = AA^*T^{3/2}e^{-q(\Phi_0+V_{G1})/k_BT} \left[ e^{qV_D/nk_BT} - 1 \right]. \tag{8.8}
\]
Figure 8-8: Schematic depiction of the capacitor model used to describe our perfect tuning of the Schottky barrier height. We adopt a capacitor model which is analogous to the capacitance seen in MOSFETs which follow the 60 mV/decade limit. The model, derived in the main text, describes how the drain and gate voltages compete for control over the potential of the WSe$_2$ channel. This figure is reproduced from LaGasse et al [5], copyright Wiley 2019.
Eq. (8.8) tells us that the gate and drain voltages essentially act at parity to determine the drain current—this presents an interesting opportunity for a new type of diode measurement.

Fig. 8-9 shows the forward bias portion of the $I_D-V_D$ measurement for several different values of $V_{G1}$ in the perfect-tuning regime. The ideality factor for each curve is essentially the same and the curves follow the relationship which we derived above in eq. (8.8). What we see is that by simultaneously changing the gate voltage with the drain voltage, it is possible to “hop the tracks” within the measurement, and change the effective diode ideality factor, which we will call $n_{\text{Eff}}$. Physically, this is achieved by changing the Schottky barrier height mid-measurement. Shrinking the Schottky barrier during the measurement, $\Delta \Phi_B < 0$, leads to $n_{\text{Eff}} < n$, where $n$ is the standard diode ideality factor. Similarly, increasing the Schottky barrier mid-measurement leads to $n_{\text{Eff}} > n$.

To analytically describe this phenomenon, we will work from eq. (8.8), starting by rewriting the gate voltage as $V_{G1} = V_0 + mV_D$, with $V_0$ being the value of $V_{G1}$ when the measurement begins and $m$ being the rate of sweeping of the gate voltage relative to the drain. Substituting this into eq. (8.8) and bringing the $mV_D$ term into the brackets gives

$$I_D = AA^*T^{3/2}e^{-q(\Phi_0 + V_0)/k_BT} \left[ e^{\frac{q}{n_{\text{Eff}}k_BT}V_D} - e^{\frac{-q}{k_BT}mV_D} \right].$$

(8.9)

We can further re-write the pre-factor as $I_0 = AA^*T^{3/2}e^{-q(\Phi_0 + V_0)/k_BT}$, which is the leakage current for the standard diode measurement with $V_{G1} = V_0$. Additionally, the first term inside the braces may be simplified by defining

$$n_{\text{Eff}} = \frac{n}{1 - nm}.$$  

(8.10)

Finally, we arrive at what we refer to as the “Super-ideal” diode equation:

$$I_D = I_0 \left[ e^{\frac{q}{n_{\text{Eff}}k_BT}V_D} - e^{\frac{-q}{k_BT}mV_D} \right].$$

(8.11)

Eq. (8.11) is essentially the same as Shockley’s diode equation, eq (7.1), except for a small correction, $-e^{\frac{-q}{k_BT}mV_D}$, which is negligible when the diode is forward biased, and now the ideality factor is digitally controlled by the $m$ parameter.
Figure 8-9: How to achieve dynamic tuning of the diode ideality factor This figure depicts how dynamic tuning of the diode ideality factor can be achieved by simultaneously sweeping the gate and the drain voltages during a measurement. Expanding the Schottky barrier during the measurement results in effective ideality factors which are greater than the standard ideality factor, while shrinking the Schottky barrier during the measurement leads to steeper diode characteristics. This figure is reproduced from LaGasse et al [5], copyright Wiley 2019.
Figure 8-10: Demonstration of dynamic control over the diode ideality factor
(A) Diode measurements with dynamically tuned ideality factors between \( n_{\text{Eff}} = n/10 \) and \( n_{\text{Eff}} = 4n \). The starting gate voltage for these measurements was \( V_{G1} = 0.4 \) V. (B) Comparison with the sweeping rate and extracted values of \( n_{\text{Eff}} \) for different starting \( V_{G1} \). Good agreement is shown between the experimentally extracted ideality factors and eq. (8.10), derived in the main text. This figure is reproduced from LaGasse et al. [5], copyright Wiley 2019.
Fig. 8-10a shows tunable Schottky diode measurements for $n_{\text{Eff}}$ between $n/10$ and $4n$ for a starting gate voltage of $V_{G1} = V_0 = 0.4$ V. Fig. 8-11 shows similar data for starting gate voltages of 0.2 and 0.6 V. In each of these measurements, sweeping the gate in the opposite direction of $V_D$ ($m < 0$) leads to ideality factors where $n_{\text{Eff}} < n$. Conversely, sweeping the gate in the same direction as $V_D$ ($m > 0$) leads to ideality factors where $n_{\text{Eff}} > n$. Again, this process is summarized in Fig. 8-9. The tunable ideality factors we show are only limited by the gate and drain voltages which we are comfortable with applying to the sample; indeed, some $m$ values can lead to very sharply increasing gate or drain voltages which could damage the sample. Generally, starting with small $V_0$ means higher ideality factors are safely achievable and starting with large $V_0$ means smaller ideality factors are safer.

To verify the equation which we derived for the effective ideality factor, we compare the experimentally extracted ideality factors with eq. (8.10) using $n = 1.15$ for the standard ideality factor (Fig. 8-10b). The experimentally extracted ideality factors show good agreement with eq. (8.10). Some small deviation is observed, which is likely due to the limited resolution of the parameter analyzer voltage sources, which cannot always accurately capture the voltages needed for some $m$ values. Another reason could be that the ideality factor, while nearly constant with gate voltage, does have some small variation.

While this type of measurement is feasible in any gated diode, our ability to perfectly tune the Schottky barrier height granted by our pristine metal-semiconductor interface leads to near perfect control over $n_{\text{Eff}}$. While so far, this technique is an interesting novelty, one possible application where this effect could be harnessed is in the area of photovoltaic (PV) devices. In our tunable diode measurement, the effective ideality factor and the leakage current are decoupled. For example, Fig. 8-12 shows extracted leakage current values for the three different starting gate voltages which we used before. In these measurements, the leakage current is fixed by the value of $V_0$, with little observable deviation between for $0.1 < n_{\text{Eff}} < 6$. The above effect is in stark contrast to conventional diodes, where increasing the ideality factor is usually accompanied with a large increase in the leakage current of the diode. This effect can potentially be leveraged to create a diode with a large PV fill factor,
Figure 8-11: More examples of dynamically tuned diode measurements. More examples of the same type of measurement demonstrated in Fig. 8-10a, for starting gate voltages of (A) \( V_{G1} = 0.2 \) V and (A) \( V_{G1} = 0.6 \) V. This figure is reproduced from LaGasse et al. [5], copyright Wiley 2019.
Figure 8-12: Demonstration of decoupling of the diode leakage current and $n_{\text{Eff}}$
Leakage currents resulting from fitting the forward bias portions of Fig. 8-10a and Fig. 8-11,
showing how the leakage current is essentially constant for a given $V_0$, for changing values
of $n_{\text{Eff}}$. This figure is reproduced from LaGasse et al [5], copyright Wiley 2019.
since the diode open circuit voltage is given by

\[ V_{OC} = n_{Eff} \frac{k_B T}{q} \ln \left( \frac{I_{PC}}{I_0} + 1 \right). \]  \hspace{1cm} (8.12)

Given our ability to increase \( n_{Eff} \) without changing \( I_0 \), it should be possible to arbitrarily increase the open circuit voltage, only limited by the bandgap of the semiconductor. At the extreme limit, tying \( V_{G1} \) and \( V_D \) together \((m = 1)\), assuming a starting ideality factor of \( n = 1 \), the effective ideality factor will approach infinity. Achieving an ideality factor of exactly one is unlikely to happen, but it should still be possible to achieve very large open circuit voltages for diodes with ideality factors close to one. Furthermore, trying \( V_{G1} \) and \( V_D \) together turns the device from a three- to a two-terminal device, making it comparable with standard PV cells.

### 8.3 Remarks on measurement stability, yield

To properly extract the Schottky barrier height as a function of gate voltage, the graphene-WSe\(_2\) device must be remarkably stable. This requires that the sample exhibit both low gate hysteresis, as is demonstrated in Fig. 8-5a, and good stability at different temperatures. For example, the earlier device studied in the previous chapter exhibited a similarly tunable Schottky barrier, but the threshold voltage of the sample would shift slightly between measurements, making it impossible to compare the same gate voltage across different temperatures.

To further show how stable our sample is, we include Fig. 8-13a. Fig. 8-13a shows the same diode configuration measured throughout different points of the day in which the Schottky barrier extraction was performed, with little variation between the three measurements. Fig. 8-13b gives a different perspective, but with the same conclusion, where this time we dual-sweep the super-ideal diode measurement. In this case, the device is remarkably stable even when simultaneously changing the gate and drain voltages.

Here we will briefly comment on the statistical yield of these types of samples. A series of measurements are typically used to classify devices which warrant further Schottky barrier
Figure 8-13: Demonstration of measurement repeatability The measurements presented here show that the sample was highly stable across different measurements. (A) Shows repeated diode measurements for the same gate configuration, taken throughout the same day as the Schottky barrier extraction study was performed. (B) Shows a double-sweep version of the dynamic diode tuning measurement, showing essentially zero hysteresis. This figure is reproduced from LaGasse et al [5], copyright Wiley 2019.
measurements. Double-sweep transfer curves are first performed on each device—any sample with noticeable hysteresis is immediately sorted out as defective. In total, seven devices have been measured with acceptable levels of hysteresis, of these devices six have shown characteristics of un-pinned Schottky barriers (primarily, gate-tunable diode characteristics similar to what is shown in Fig. 8-5). As an example, Fig. 8-14 shows measurements on another sample, demonstrating essentially the same results as were shown in Fig. 8-5b and c.

8.4 Chapter conclusions

In this chapter we have seen how it is possible to solve the problem of Fermi-level pinning at a metal-semiconductor interface by replacing the metal and the semiconductor with 2D materials, in this case, graphene and WSe$_2$. We measured the Schottky barrier on a gate-tunable graphene-WSe$_2$ device which was encapsulated in h-BN. The low level of disorder in the sample, coupled with the terminated bonds of 2D materials, resulted in a metal-semiconductor interface which was free of Fermi-level pinning effects.

This was demonstrated by extracting the Schottky barrier height at the graphene-WSe$_2$ interface as a function of the gate voltage, showing one-to-one tuning of the barrier. This result can be explained by adopting a voltage divider capacitor model, whereby the gate voltage directly couples to the potential of the semiconductor channel. The ability to create extremely clean metal-semiconductor interfaces which may be electrically tuned with a gate has applications towards creating higher quality nanoelectronic devices from 2D materials.

Finally, we demonstrated a novel technique of dynamic diode tuning enabled by the extreme level of Schottky barrier tunability exhibited by our sample. We showed how it is possible to achieve digital control over the diode ideality factor, in which the effective ideality factor is decoupled from the leakage current of the diode. This result has some profound implications for the study of PV devices, and is a topic for future study.
Figure 8-14: Measurements on an alternate Schottky diode sample These measurements are for a later generation graphene-WSe$_2$ sample which exhibited the same characteristics as the sample studied within the rest of this chapter. (A) shows a gate-tunable diode measurement, exhibiting the perfect incrementing of the diode characteristic with gate voltage, a signature of the tunable Schottky barrier. (B) Shows the extracted ideality factor and leakage currents as a function of the tuning gate voltage. Both of these results are essentially the same as what was demonstratd in Fig. 8-5b and c. This figure is reproduced from LaGasse et al [5], copyright Wiley 2019.
Final conclusions, outlook for the future

In this dissertation we have covered topics ranging from quantum transport theory and tight-binding models all the way to the isolation of 2D materials with Scotch tape and their manual assembly into van der Waals heterostructures. In this final section, I will wrap up all of these discussions and provide a road-map for the future.

In chapters 3 and 4 we saw how tight-binding transport theory can be used to model radiation effects in both carbon nanotube and graphene devices. I expect that as the ballistic limit becomes more and more readily achievable in electronic devices, as the room-temperature measurements quasi-ballistic graphene in chapter 7 suggest, that these papers will become more and more relevant. Indeed, the semiconductor industry has been pushing towards extremely small transistor channels in recent years, and as the radiation response of these types of devices begins to be quantified, tight-binding transport is likely to be a very useful modeling tool.

Chapters 5 and 6 were both examples of how with the right application of tight-binding models, excellent agreement with complex magnetotransport experiments, both for low and high magnetic fields, can be achieved. These chapters, in particular, demonstrated how a fundamental understanding of the experimental conditions can be leveraged to develop a successful tight-binding simulation. I expect that the physical intuition developed in these chapters will be useful in understanding the fundamental aspects of more complex devices formed in ballistic graphene channels in the future, such as Mach-Zender and Fabry-Perot interferometers and for $p$-$n$ junction based quantum Hall resistance standards, which will rely on the complex effects studied in these chapters.
Initially when beginning experimental measurements, my goal was to come up with more experimental data to model, similar to the experiments referenced in Chapter 6. I quickly learned that is much easier said than done—this led to spending a great deal of time developing the methods for creating high-quality samples for our lab, which I documented in Chapter 7. I am happy to say that near the culmination of my thesis work, I was finally able to obtain graphene $p$-$n$ junction samples of high enough quality that I could readily tie to them with tight-binding transport theory. While within this dissertation, those measurements were simply used to demonstrate the quality of the transfer techniques described, the possibility of readily achieving room temperature ballistic transport in graphene $p$-$n$ junctions is an enticing area of research.

Finally, in Chapter 8, we looked at gate-tunable graphene-WSe$_2$ Schottky junctions which obeyed the Schottky-Mott rule. Using a dual-gated structure, we saw how it was possible to tune the Schottky junction in one graphene-WSe$_2$ interface. Using a battery of transport measurements, the Schottky barrier was extracted over a range of gate voltages, with perfect tuning. Furthermore, the tunability of the Schottky barrier led us to developing a dynamic diode tuning technique where we demonstrated digital control over the effective diode ideality factor. The results in Chapter 8 are exciting for a number of reasons. The ability to create good electrical contact to encapsulated TMD crystals is a major challenge, and these results point towards the possibility of doing quantum transport measurements of $h$-BN encapsulated 2D materials, with high quality gate-tunable contacts. The ability to tune the diode ideality factor also has potentially profound effects on PV devices. The results presented in this dissertation provide the first steps towards achieving either of these goals.
Bibliography


195


