Introducing Spin-Orbit Interaction in Graphene

by

Jun Yong Khoo

B.A. Physics, University of Oxford (2011)
M.Sc. Physics, University of Waterloo (2013)

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2018

© Massachusetts Institute of Technology 2018. All rights reserved.

Signature redacted

Author .............

Department of Physics
August 31, 2018

Signature redacted

Certified by.......

Leonid S. Levitov
Professor of Physics
Thesis Supervisor

Signature redacted

Accepted by ......

Nergis Mavalvala
Associate Department Head, Physics
Introducing Spin-Orbit Interaction in Graphene

by

Jun Yong Khoo

Submitted to the Department of Physics
on August 31, 2018, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The excellent electron properties of graphene, an atomically-thin material with record-high carrier mobility and gate tunability, make it central to modern nanoscience. However, the spin-orbit interaction (SOI) naturally present in graphene is extremely weak and has yet to be observed experimentally. This presents an obstacle for accessing novel phenomena in transport and optics, in particular those related to topological properties. This thesis seeks to address this limitation by artificially introducing SOI in graphene sandwiched between other atomically-thin materials that can produce an interfacial SOI in graphene. In particular, it is demonstrated that a strong SOI, naturally present in the two-dimensional materials such as transition metal dichalcogenides (TMD), can be partially transferred to graphene via the proximity effect. We predict a range of novel phenomena arising in graphene bilayers with layer-asymmetric SOI induced by a proximal TMD layer. These include a gate-tunable SOI, a gate-tunable intrinsic valley-Hall conductivity, as well as a gate-tunable edge conductivity, to name just a few. These findings will facilitate exploring previously inaccessible spin-related phenomena in graphene and other van der Waals heterostructures.

Thesis Supervisor: Leonid S. Levitov
Title: Professor of Physics
Acknowledgments

MIT and Cambridge in general, is a place filled with opportunities. While there are countless world class programs and classes, it is really the people who I have had the chance to interact with and learn from through these opportunities that have made me grow in more ways than I could have ever imagined.

I am deeply indebted to my advisor and mentor Leonid Levitov. First of all I am thankful for the physics that he has taught me, either through his classes or meetings in person and over Skype calls. He has helped me mature as a physicist and has always pointed me towards thinking beyond just solving the problem at hand and focusing on more important questions. These include the potential novel applications and phenomena that the solutions could lead to, and motivations or reasons to why anyone should be interested in the problem. His clarity of thought and emphasis on clear communication of ideas has helped me improve my scientific writing and presentation skills. Through his guidance, I gradually began to appreciate the importance of clear scientific communication, to write with a strong motivation and develop ideas coherently. I am appreciative of his care towards me as a mentor, always looking out for opportunities and collaborations that would benefit me. It is through these collaborations that I got exposed to different ideas and projects, as well as gotten the opportunity to meet and work with many others who have helped me along in this journey. Finally, I am most grateful for his time and support throughout my time at MIT, especially in these last few months.

I am fortunate to be a part of various experimental collaborations. I am privileged to have had the chance to work with local experimentalists as well as those from other institutions. I thank Pablo Jarillo-Herrero, Valla Fatemi, Yuan Cao for insightful discussions. I also thank Alberto Morpurgo and Zhe Wang from the University of Geneva for our many fruitful Skype discussions as well as for sharing with me their experimental data. Finally, I am thankful also to Andrea Young and Joshua Island from the University of California Santa Barbara for current ongoing discussions and sharing their data.
It was a privilege to have been a part of the Center for Integrated Quantum Materials (CIQM). It was in one of their annual meetings where I learned how to give my first five-minute talk, which was surprisingly difficult to do well. CIQM also provided the opportunity to meet and interact with many distinguished physicists and other members through group meetings, workshops and conferences. I am blessed to have met Shiang Fang from Harvard University through CIQM. I have learned a lot from discussions with him and I am thankful for his willingness and openness to share his ideas, derivations and matlab codes. The sincerity in our discussions has led to a friendship beyond academic relations which I am really grateful for.

This journey would not have been as enjoyable without my office mates. I am thankful for the fruitful discussions, oral exam preparation sessions, and other chit-chat sessions that I had with all of them. In particular I thank Jian Feng Kong for sharing his coffee and coffee knowledge and discussing problem sets. I also thank Oles Shtanko for teaching me some Russian phrases and the exchange of cultural knowledge. Finally I thank Cyprian Lewandowski for teaching me some Polish words, for the fruitful discussions in more recent collaborations, as well as for offering a helping hand on many different occasions.

The Condense Matter Theory (CMT) department has always been a friendly and open environment for discussions. I am thankful for discussions with Inti Sodemann and Brian Skinner. Discussions with Inti in particular, have led to collaborative side projects. I would also like to thank Michael Pretko, who had been proactive in organizing student seminars and oral exam practice sessions to promote discussion of ideas and learning among graduate students.

Moving slightly away from research but still staying within my experience with physics is my experience in grading exams for some undergraduate classes. These have been very memorable and interesting moments, not so much for the grading itself, but the brief interactions with some faculty members in an informal setting. In particular, it was a pleasure to have known Peter Dourmashkin, Ibrahim Cisse and Joseph Checkelsky. Their dedication, passion and enthusiasm in general could be felt and it has always been a joy to talk to them. For Peter, I sincerely hope that one
day in the future we will have the chance to meet up in Singapore. For Ibrahim, it was exciting to have taken a beginner’s lesson in salsa from him; he was an excellent and lively teacher! For Joe, I will never forget that one eventful grading session when something went wrong and my lunch order was missing. He immediately gave me some money and insisted that I have lunch first before grading. Words cannot express how touched I was by his kind gesture.

The classes here at MIT and Harvard were some of the best ones in my life. I was fortunate to have caught a glimpse of the different faculties’ mindsets from auditing and taking their classes. I am grateful to have been taught by Xiaogang Wen, Senthil Todadri, Patrick Lee, Mehran Kardar, Liang Fu, Hong Liu, Jesse Thaler, Steven Johnson, Subir Sachdev, Bertrand Halperin, Xi Yin, and of course Leonid himself.

Outside of physics, I am thankful for the buffet/hotpot group for the company and friendship – Jian Feng, Quntao Zhuang, Zi Wen Liu, Elton Zhu, Ta-wei Wang and Qingyang Wang. I am also grateful for the support and fellowship with Gerald Pho, Peng Shi, Annie Chen, David Miculescu and Samuel Perli in the initial years of my PhD, some of whom are talented musicians. I am blessed to have played music with them at several Graduate Christian Fellowship large group meetings.

Music has always brought me great joy and I am grateful for the opportunity to play with my current band Veisty on the bass guitar. I thank Joseph Sandt for bringing up my request to join the band at the end of last year, and for the fun times we spent together working on our maker’s class project. I also thank the rest of the band, Elise Strobach, Jean Sack and Daniel Preston, for giving me a chance to be a part of the band. Our various practice sessions and performances this year, as well as the upcoming recording sessions, will forever be imprinted in my memory.

Finally, I thank my lovely wife Berenice, parents and grandmother for being there for me and to have always been a source of encouragement and support all these years. In particular, I am eternally grateful to Berenice for her extraordinary help with the things at home in our last few months here in Cambridge, for being my lifelong companion and for being an endless source of love and support in every aspect of my life. Dominus Illuminatio Mea.
Contents

1 Introduction ................................................. 15
   1.1 Heterostructures in Solid State Physics ................. 17
       1.1.1 "The whole is greater than the sum of its parts."
            – Aristotle ........................................ 17
       1.1.2 Semiconductor-Based Heterostructures ............. 17
   1.2 Graphene-Based Heterostructures ......................... 22
   1.3 Spin Manipulation ...................................... 28
   1.4 Thesis Outline ......................................... 32

2 Bilayer Graphene Theory ................................... 35
   2.1 Monolayer and Bilayer Graphene Lattice .................. 35
   2.2 Monolayer Graphene Low-Energy Effective Hamiltonian .... 37
   2.3 Bilayer Graphene Effective Hamiltonian .................. 39
   2.4 Gated Bilayer Graphene .................................. 42
       2.4.1 Low-Energy Spectrum and Eigenstates ............... 44
       2.4.2 Capacitance Effects ................................ 46
   2.5 Intrinsic Spin-Orbit Interaction in Graphene ............ 47

3 Origin and Magnitude of Spin-Orbit Interaction in Graphene on
   Semiconducting Transition Metal Dichalcogenides .......... 51
   3.1 Interfacial Interactions in vdW Heterostructures ....... 52
   3.2 Device Fabrication and Characterization .................. 55
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.8 Conductances of Novel Phases</td>
<td>118</td>
</tr>
<tr>
<td>5.9 Effects of Layer-Specific Rashba SOI</td>
<td>118</td>
</tr>
<tr>
<td>5.10 Chapter Summary</td>
<td>119</td>
</tr>
<tr>
<td>5.11 Appendix: Derivation of Zeroth LL Spectrum</td>
<td>120</td>
</tr>
</tbody>
</table>

6 Conclusions 127
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Rectifying I-V characteristics of a p-n junction</td>
<td>19</td>
</tr>
<tr>
<td>1-2</td>
<td>Structure of bipolar junction transistor</td>
<td>20</td>
</tr>
<tr>
<td>1-3</td>
<td>Schematic of MOS capacitor</td>
<td>21</td>
</tr>
<tr>
<td>1-4</td>
<td>Schematic of MOSFET operation</td>
<td>22</td>
</tr>
<tr>
<td>1-5</td>
<td>2D crystals as Lego blocks</td>
<td>24</td>
</tr>
<tr>
<td>1-6</td>
<td>Graphene-BN resonant tunnelling transistor</td>
<td>25</td>
</tr>
<tr>
<td>1-7</td>
<td>Graphene-BN resonant tunnelling transistor</td>
<td>26</td>
</tr>
<tr>
<td>1-8</td>
<td>Moire superlattice</td>
<td>27</td>
</tr>
<tr>
<td>1-9</td>
<td>The spin Hall effect</td>
<td>29</td>
</tr>
<tr>
<td>1-10</td>
<td>Giant magnetoresistance and the spin-valve effect</td>
<td>30</td>
</tr>
<tr>
<td>2-1</td>
<td>Crystal structure of monolayer graphene</td>
<td>36</td>
</tr>
<tr>
<td>2-2</td>
<td>Crystal structure of Bernal-stacked bilayer graphene</td>
<td>37</td>
</tr>
<tr>
<td>2-3</td>
<td>Bilayer graphene bandstructure</td>
<td>43</td>
</tr>
<tr>
<td>2-4</td>
<td>Layer polarization of low energy states</td>
<td>46</td>
</tr>
<tr>
<td>2-5</td>
<td>Quantum spin Hall effect</td>
<td>49</td>
</tr>
<tr>
<td>3-1</td>
<td>Basic characterization of graphene on TMD substrates</td>
<td>56</td>
</tr>
<tr>
<td>3-2</td>
<td>Negative magneto-conductivity due to WAL in monolayer graphene on TMD substrates</td>
<td>59</td>
</tr>
<tr>
<td>3-3</td>
<td>Interfacially induced SOI in graphene-on-TMD is a robust phenomenon</td>
<td>60</td>
</tr>
<tr>
<td>3-4</td>
<td>Interfacially-induced SOI in graphene multilayers on TMDs</td>
<td>63</td>
</tr>
<tr>
<td>3-5</td>
<td>Comparison of WAL data measured on monolayer graphene on different TMDs</td>
<td>66</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3-6</td>
<td>Upper bound of $\tau_{so}$, extracted from WAL data, as a function of $\tau$ for monolayer graphene devices</td>
<td>67</td>
</tr>
<tr>
<td>3-7</td>
<td>Extracting SOI from the SdH resistance oscillations observed in BLG-on-WSe$_2$ devices</td>
<td>73</td>
</tr>
<tr>
<td>3-8</td>
<td>Low energy band structure of bilayer graphene on TMD</td>
<td>74</td>
</tr>
<tr>
<td>3-9</td>
<td>Magneto-conductivity measured at high temperatures (20-30 K) in the devices with different mean free path</td>
<td>82</td>
</tr>
<tr>
<td>4-1</td>
<td>Low-energy band structure of a biased bilayer graphene with induced SOI</td>
<td>87</td>
</tr>
<tr>
<td>4-2</td>
<td>Four lowest energy bands and layer polarization of corresponding eigenstates of bilayer graphene with layer-specific SOI</td>
<td>94</td>
</tr>
<tr>
<td>4-3</td>
<td>Phase diagram of the BLG-on-TMD system</td>
<td>96</td>
</tr>
<tr>
<td>5-1</td>
<td>Edge state configurations for the compensated electron-hole ortho-, para-, and ortho/para-phases</td>
<td>106</td>
</tr>
<tr>
<td>5-2</td>
<td>Single-particle zeroth LL spectrum of bilayer graphene with Ising and Rashba SOI</td>
<td>110</td>
</tr>
<tr>
<td>5-3</td>
<td>Phase diagrams showing how the regions of inverted orbital ordering and normal orbital ordering changes with Ising SOI</td>
<td>113</td>
</tr>
<tr>
<td>5-4</td>
<td>Phase diagram showing the ortho-, para-, and ortho/para- compensated electron-hole phases</td>
<td>116</td>
</tr>
<tr>
<td>5-5</td>
<td>Illustration of the extreme smallness of the level shifts induced by the Rashba SOI in the zeroth LL</td>
<td>119</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Heterostructures, or heterojunctions, is a concept in materials science that has revolutionized the fundamental and applied physics in the 20-th century. Heterostructures are formed by combining two (or more) distinct materials in close proximity, so that the chemical composition changes abruptly across the materials' interface. The interface, where the properties of both materials are present, can give rise to "synergetic" behaviors which are more rich and interesting than those of the constituting materials. The interplay between different material properties can in turn give rise to a wide variety of new phenomena, many of which are hard to realize in naturally-found or even artificially-made bulk materials. Furthermore, it was found that heterostructures often provide the unique means to achieve new functionalities.

Creating new properties and functionalities has always been fundamental to solid state physics. Consequently, heterostructures were central to many of the key developments in the physics of the last century. They have led to new physics and revolutionary technology such as two-dimensional electron gases in metal-oxide semiconductor structures, field-effect transistors, quantum Hall effect, topological electronic states, and so on. However, making perfect heterostructures has been a challenging task. High concentrations of impurities and defects, as well as deficiencies in the fabrication and purification technologies, have often hindered the realization of the desired effects as well as the observation of novel phenomena. However, gaining control over the properties of heterostructures, while being really hard, had enormous benefits.
The despair due to the difficulty of the problem, as well as the fascination by the prospects, is captured in the two famous quotes,

"God made the bulk; surfaces (and interfaces) were invented by the devil"
- Wolfgang Pauli

and

"Often, it may be said that the interface is the device.” - Herbert Kroemer

A big boost to the field came from the recent discovery of two-dimensional (2D) atomically-thin materials, such as graphene and its relatives. These materials, bound by interfacial van der Waals forces, can form atomically perfect interfaces, which are essentially defect-free. The van der Waals heterostructures, as they are now known, helped to address some of the issues faced by their three-dimensional counterparts. The ability to align pristine samples of 2D materials with atomic precision has substantially reduced the concentration of defects and impurities. In addition, in these devices, the system of interest itself has the same dimensionality (in this case, 2D) as its interface with other materials. The “bulk” and “surface” mentioned in the above quote by Pauli are one and the same system in this case. Furthermore, since electronic states of atomically-thin monolayers are fully exposed, they can be controlled and probed from the out-of-plane direction.

The focus of this thesis is on controlling the spin degree of freedom in graphene. Graphene is a material with excellent electron properties such as remarkably high carrier mobilities (which exceed 15000 cm²V⁻¹s⁻¹ at room temperature) and carrier densities that can be electron-like or hole-like and tuned continuously over a range as wide as 10ⁱ³cm⁻² [1, 2, 3, 4]. However, the intrinsic spin-orbit interaction (SOI) in graphene is too weak and has yet to be observed in experiments [5]. Many interesting phenomena predicted requires the presence of strong SOI. Because this cannot be achieved by graphene alone, we turn to graphene-based heterostructures to design high-mobility carriers with strong SOI. But before diving into these details, let us take a broader view of the development of heterostructures in semiconductor physics that pervade our life today and continue to empower our society.
1.1 Heterostructures in Solid State Physics

1.1.1 "The whole is greater than the sum of its parts."

– Aristotle

The key notion of heterostructures can be viewed in comparison with one of the fundamental working principles that mankind has adopted towards problem solving – unity makes strength. Because all of the desired features are never found all together in what is obtainable easily, humans attempt to combine different ingredients possessing different subsets of these features. This principle is very familiar to everybody through cooking, wherein the specific choice, quantity and process of introducing the different ingredients need to be carefully fine-tuned to achieve the desired outcome.

But humans have also learned ways to go beyond blending different properties in a simple mechanical fashion. Take selective breeding in agriculture as an example, where developing new strains and varieties of plants was followed, after the discovery of genetics, by cross-breeding, which allowed to create entirely new species of plants (including some very tasty ones, such as grapefruit, plumcot, and so on). Similarly, in designing new materials, the old-fashioned “additive” approach to improving properties of materials turned out to be less effective than cross-breeding of very different materials. Namely, novel phenomena often emerge at the transition region or interface of different materials. In such a case, in line with the adage supposedly going back to Aristotle, the whole system can become truly greater than the sum of its parts.

1.1.2 Semiconductor-Based Heterostructures

Turning to the subject of the thesis, I begin by surveying some of the most successful and impactful examples of heterostructures in semiconductor physics developed over the past few decades. These examples serve to illustrate two essential features of heterostructures that make them especially appealing. One is the acquisition of a new effect at the interface of a heterostructure which leads to a new functionality. The
particular heterostructure can then be thought of as the building block associated to that specific functionality. Second, each of these building blocks can then be combined to form a more complex heterostructure that seeks to synergize the functionalities of the building blocks to achieve a new and typically more complex functionality. These two recurring features establishes the background for approaching the 2D heterostructures that I will be focusing on in the rest of this thesis. In particular, it provides the context to discuss the advantages and novel aspects of the 2D heterostructures over their three dimensional (3D) counterparts.

A semiconductor is a material which has a conductivity that is between that of a metal and an insulator. This feature is due to the presence of a relatively small band gap in which its chemical potential lies. Because of the narrow band gap, its chemical potential can be tuned relatively easily into its conduction (n-doped) or valence band (p-doped) either by doping or gating. It is therefore much easier to alter the sign of charge carriers and vary its carrier densities compared to a metal or an insulator. This tunability sets it apart from a metal or insulator, which has electronic properties that are largely determined by their crystal structure.

An immediate consequence is that within the same semiconducting material, two different regions can be oppositely doped so that carriers in the different regions have opposite charges. The result is an interfacial region known as the p-n junction, which is the simplest example of an interface that exhibits novel effects. Separately, p- and n-doped semiconductors can be thought of as materials with oppositely charged free carriers. When placed together, the I-V characteristic across the interface is modified drastically from their individual I-V characteristic (see Fig. 1-1). This modified I-V characteristic endows the p-n junction with the new functionality of current-rectification, whereby electrical current is allowed to pass through the junction only in one direction. Current rectification can not be achieved simply with a p-doped or n-doped semiconductor but is born out of the interface between the two. As we shall see later, there exists a graphene-based heterostructure analogous to and richer than the p-n junction. In that case, the I-V characteristic is not merely modified, but can be done in a controlled fashion.
Combining two p-n junctions results in a p-n junction transistor [7] or what is now known as a bipolar junction transistor (BJT). In an n-p-n BJT for instance (see Fig. 1-2), a p-doped region is sandwiched between two n-doped regions. A contact is formed at the p-region (base), as well as each of the n-regions (collector and emitter). Under the correct bias configuration, it acts as a current amplifier. This allows BJTs to be used as amplifiers or switches, which has applications in many modern electronic devices today such as computers, telephones and radio transmitters. We thus find a new functionality in a heterostructure formed by two simpler ones. Both of these heterostructures can be separately treated as building blocks for even more complex semiconducting electronic devices.

One is not restricted to considering interfaces between differently doped regions of the same semiconducting material. A natural progression is to consider the interface between dissimilar crystalline semiconductors with unequal band gaps. This is known as a heterojunction. Heterojunctions allow for engineering devices with spatially distinct band gaps, which enable applications such as semiconducting lasers, solar cells and transistors. Going one step further, a semiconductor can be proximitized with a metal to form a metal-semiconductor junction. The junction can either be rectifying (Schottky barrier) or non-rectifying (ohmic contact).

Going up the complexity ladder, I now discuss the metal-oxide semiconductor (MOS) capacitor which is a heterostructure comprising all three types of materials.
discussed thus far. This heterostructure allows for the gate-control of the *depletion layer*, the region at the oxide-semiconductor interface that is essentially devoid of carriers. By changing the potential difference between the metal and semiconductor layers, the depletion layer changes into an *inversion layer*, which hosts carriers (e.g. electrons) with the opposite charge to those from the bulk semiconductor (e.g. p-doped semiconductor). The MOS capacitor therefore achieves the new important functionality of a gate-tunable carrier density.

Sandwiching the MOS capacitor between two p-n junctions gives rise to the metal-oxide-semiconductor field-effect transistor (MOSFET). In this heterostructure, both the current-rectifying feature of the p-n junctions and the gate-tunable carrier density
Figure 1-3: Schematic of metal-oxide semiconductor capacitor biased into (a) surface depletion and (b) inversion. Metal is labeled as ‘Gate’, Oxide as ‘SiO₂’ and p-type semiconductor ‘P-Si’. When gate voltage \( V_g \) is greater than the flatband voltage \( V_{fb} \), a depletion layer forms at the oxide-semiconductor interface where carriers are absent. Further increasing \( V_g \) past the threshold voltage \( V_t > V_{fb} \) causes electrons to accumulate at this interface [6].

feature of the MOS capacitor are combined to achieve the so called field-effect. The resulting field-effect transistor is a unipolar transistor which uses only one kind of charge carrier, in contrast to the BJT discussed earlier. For instance, an nMOSFET is constructed from a MOS capacitor with a p-doped semiconductor layer and sandwiched between two n-doped semiconductors (see Fig. 1-4). In the ‘off’ state, this is essentially an n-p-n junction and no current is allowed to pass through it when a potential difference is applied between the two adjacent p-doped regions. By changing the (transverse) potential difference across the MOS capacitor, an inversion layer with excess electrons forms. This inversion layer then acts as an n-channel that connects the two n-doped regions and current transmission is turned ‘on’. As we shall see later, there exists a different graphene-based heterostructure analogous to the MOSFET. In that case, the carrier density as well as the bandstructure itself can be tuned by
applying a transverse electric field.

Figure 1-4: Schematic of n-type metal-oxide semiconductor field effect transistor (nMOSFET), which is formed by a MOS capacitor with p-doped semiconductor layer sandwiched between two n-doped semiconducting regions. Shown here in (a) is the 'on' state when $V_g > V_t$ as in the case of Fig. 1-3(b). This can be thought of as forming an n-channel which connects the two n-doped regions and thus allowing current to flow [6].

With this new functionality, MOSFETs in turn become new building blocks which can be combined cleverly with other heterostructures in digital integrated circuits. Different integrated circuits are designed to perform different tasks such as basic logic operations and data storage found in microprocessors that proliferate our society today. Of course, there are many more examples of semiconductor-based heterostructures and our discussion thus far does not begin to even scratch its surface. The above serves to illustrate the ongoing process of building more complex heterostructures with new functionalities out of simpler ones. Having surveyed the capabilities of 3D heterostructures, we are now in a position to understand the novel aspects that come with forming heterostructures out of atomically-thin 2D materials.

### 1.2 Graphene-Based Heterostructures

In the previous section, the examples we looked at were 3D heterostructures. The discovery of graphene brought about the 2D revolution and with it, a new generation of 2D heterostructures. In this case, because the 2D materials are few-atoms thick,
the interface and bulk are one and the same. Furthermore, these 2D materials are atomically-flat so that they can be stacked on top of one another with relative ease. The resulting heterostructures formed in this way have interfaces that are essentially free of the impurities or crystallographic defects which were issues often found in 3D heterostructures.

In this section, I discuss several examples of graphene-based heterostructures, in each case highlighting functionalities analogous to some of those achieved by semiconductor-based heterostructures. We will also explore several novel phenomena recently discovered in graphene-based heterostructures that rely specifically on the aforementioned properties of these 2D heterostructures and therefore cannot be realized by semiconductor-based heterostructures. Let us begin with a brief history of how these atomically-thin materials became a reality.

Van der Waals (vdW) materials are a special class of layered crystalline materials. In these materials, adjacent layers are held together by interlayer vdW bonds which are much weaker (typically by an order of magnitude) than the intralayer bonds that hold the atoms within each layer. Examples of vdW materials such as graphite and TMDs have been studied over the past few decades with the primary goal of understanding the properties of their bulk (few-layer) form, which were naturally available or could be synthesized via techniques such as chemical vapor deposition. Theoretical attempts to describe such systems typically begin with constructing models for their monolayer crystals\(^1\) before proceeding to model multilayered crystals by introducing interlayer coupling [8, 9, 10, 11, 12, 13]. Therefore while interesting properties of such monolayer crystals have been discovered, they remained disconnected from reality as no one had worked out how to consistently extract them from their bulk crystals and whatever small quantities that existed were produced unintentionally.

It was not until in 2004 when graphene, the monolayer crystal of graphite, was successfully isolated by Andre Geim and Konstantin Novoselov at the University of Manchester [1]. By then, the theoretical understanding of the electronic properties of

---

\(^1\)Note that monolayer is not synonymous with one-atom thick. Monolayer TMDs such as WSe\(_2\) for instance are typically three atomic layers thick.
graphene has already been well established, so that the first observation of the integer quantum Hall effect in graphene that soon followed [3, 4, 14] confirmed the presence of the theoretically predicted massless Dirac fermions and their corresponding Berry’s phase. The success of the mechanical exfoliation technique used to isolate graphene prompted its application to other vdW materials. Subsequent successful isolation has led to the discovery of 2D materials such as hBN, MoS$_2$, WS$_2$, WSe$_2$ and MoSe$_2$.

![Figure 1-5: Depiction of 2D crystals to be analogous to Lego blocks (right panel) [15].](image)

In fact, Geim and Novoselov’s pioneering work not only produced graphene but also the first graphene-based heterostructure. The isolated graphene was transferred onto thin silicon dioxide on a silicon wafer so as to electrically isolate the graphene. It follows then to consider materials other than silicon dioxide to serve as substrates for graphene, such as the other successfully isolated 2D materials. The first novel substrate was hBN [16], which demonstrated its ability to preserve high quality electronic properties of graphene. This was followed by few-layer hBN, MoS$_2$ and WS$_2$ which served as tunnel barriers for two graphene electrodes in vertically-stacked devices that functioned as field-effect transistors [17, 18], resonant tunnelling diodes [19, 20], and
photodetectors [21, 22], paving the way for more designer vdW heterostructures to come [15].

One particular graphene-based heterostructure analogous to the p-n junction is formed by two graphene electrodes on either side of an atomically thin hBN tunnel barrier [19]. Like the p-n junction, this tunnel barrier has non-linear I-V characteristics. A novel feature of this device is the ability to tune the I-V characteristics into forward or reverse bias. This is achieved by gating the bottom graphene electrode and adjusting its chemical potential relative to the top graphene electrode (see Fig. 1-6).

![Figure 1-6: Graphene-BN resonant tunnelling transistor. (a) Schematic diagram of the heterostructure. (b) Measured I-V ($V_b$) characteristics of a particular device in which the hBN barrier is four atomic layers thick. The inset shows the relative positions of the chemical potentials of the two graphene layers at the peak of the I-V curve in forward bias with $V_g = 15$ V. (c) Theoretical simulation [19].](image)

A graphene-based heterostructure that is analogous to the MOSFET is the dual-gated bilayer graphene [23]. Through gating, the chemical potential of the system, and therefore its carrier density, can be tuned. This field-effect in graphene is much more sensitive than that in the MOS setup and can be understood from basic electrostatics. The electric field strength between the gates scales as the ratio of the potential difference of the gates to the distance between them. It follows that for atomically-thin heterostructures such as graphene, the distance between gates is now exponentially reduced compared to 3D materials. This leads to enhanced electric field strengths across these 2D heterostructures, which by Gauss' law, results in much higher carrier densities.

Additionally, an interlayer bias can be introduced between the top and bottom
layers of the dual-gated bilayer graphene. This induces a non-zero bandgap that scales with the applied bias and in this sense, the bandstructure itself can be tuned. I will discuss the physics behind this phenomenon in the next chapter. Here I emphasize that this novel functionality in a gate-tunable bandstructure cannot be realized in previous 3D semiconductor-based heterostructures.

The atomic precision of stacking these 2D materials, together with the fact that the above materials all share the same triangular Bravais lattice (albeit with different lattice constants) enabled yet another knob – crystallographic alignment. When the twist angle between the crystal axes of two different 2D crystals with similar lat-
tice constants is sufficiently small, a Moire pattern develops. The electronic states experience a superlattice potential with a lattice constant at least two orders of magnitude larger. When the alignment results in a commensurate superlattice structure, the Brillouin zone of the system shrinks to a much smaller region in momentum space. The electronic bands fold and separate to form minibands [24, 25] in this mini Brillouin zone. The low-energy physics of these Moire systems, even for the incommensurate cases, are adequately captured by the k-p model. Novel phenomena such as topological bloch bands in graphene-on-hBN superlattices [26], as well as Dirac-velocity renormalization of Moire bands in twisted bilayer graphene [27] have been predicted from their corresponding k-p models. The associated flat Moire bands of the magic angle ($\sim 1.05^\circ$) twisted bilayer graphene in particular give rise to even more exotic physics. The recent observation of Mott insulating behavior and unconventional superconductivity in these twisted bilayer graphene [28, 29] came as a surprise. Nobody had thought that such exotic phenomena could exist in graphene systems. This has in turn triggered an overwhelming response from the theoretical community [30, 31, 32, 33] in attempts to explain its underlying physics.

![Figure 1-8](image)

**Figure 1-8:** (Left panel) Moire pattern formed for example in graphene-on-hBN or bilayer graphene with small twist angle $\theta$ (shown) between the two layers forming a superlattice. (Right panel) Folding of the the Brillouin zone into a mini-Brillouin zone when the superlattice is commensurate [28].

With the discovery of various novel phenomena, graphene-based heterostructures is poised to remain as one of the hottest research areas for many years to come. In this thesis, I will focus on a particular designer feature that was recently experimentally
observed in graphene-on-WS₂/WSe₂ devices – interfacially-induced SOI. Although these semiconducting TMDs have been used as substrates to preserve high electronic quality of graphene, they have not until recently, been investigated for their ability to induce a strong SOI in the graphene Dirac band. To see why this is desirable, I discuss the importance of SOI in the next section and give a brief overview of how spin control has been achieved.

1.3 Spin Manipulation

In the previous sections, we looked at how charge properties can be manipulated through the heterostructures. In this section, we will see how heterostructures enable the manipulation of spin-dependent properties that result in new functionalities. The importance of SOI is emphasized in this context, which is responsible for the underlying spin-dependent transport properties. At the same time, the ability to generate strong and localized magnetic fields is necessary for achieving some of these spin-based functionalities. It also plays several different roles in spin-based phenomena because it interacts with carriers through both their charge and spin degrees of freedom. In the second half of this section, I discuss the developments in magnetic field technology in relation to the new phenomena and capabilities it has led to. Overall, the key concepts introduced in this section set up the foundation for discussing the novel spin-based phenomena achieved in graphene-based heterostructures in the later chapters.

Like charge, spin is an intrinsic degree of freedom that electrons (also holes) possess. The beginnings of solid state physics focused almost entirely on charge-based phenomena. Spin-based phenomena gradually gained interest with the incorporation of magnetic materials into devices as well as the theoretical study and observation of the spin-Hall effect. The former enabled the generation of spin-polarized currents through the use of ferromagnetic contacts and led to the understanding of spin-dependent conductivities. The latter explored various mechanisms that can affect or be used to manipulate the spin-polarization of currents. Together, these gave birth
to the modern subfield of solid-state physics known as *spintronics*.

![Figure 1-9: The Spin Hall Effect. An electrical current induces spin accumulation at the lateral boundaries of the sample. In a cylindrical wire the spins wind around the surface [34].](image)

Predicted in 1971 by Dyakonov and Perel, spin Hall effect is a phenomenon in which oppositely oriented spins are accumulated on opposite boundaries on the lateral surface of a current-carrying sample. Originally, the mechanisms proposed are *extrinsic*. These mechanisms showed how carriers with opposite spins scatter in opposite directions when colliding with impurities in the material [35, 36]. A fundamentally different mechanism *intrinsic* to the material was later proposed. In this case, the carriers’ trajectories are curved as a result of SOI arising from the broken crystallographic symmetries of the material [37, 38].

The next big step was the arrival of major discoveries in magnetic heterostructures in 1980s. Unlike semiconductor-based heterostructures, the field of magnetic heterostructures began from metallic multilayer studies primarily for superlattice effects [39]. In such multilayered heterostructures, physical properties due of different length scales coexist. Properties due to combining different materials are typically classified in increasing order of sample complexity: interface/proximity, magnetic coupling across non-magnetic materials, and superlattice [40]. Despite so, it is typically easier to observe these shorter length scale effects in superlattice structures. This is either due to an enhancement of the effect from the increased number of layers or because most interfaces are well protected from surface contamination.

One of the most important key discoveries first observed in a superlattice structure
was the *giant magnetoresistance* (GMR). GMR refers to the large change in electrical resistance across two ferromagnetic layers separated by a non-magnetic conductive layer depending on whether the magnetization of the ferromagnetic layers are parallel or anti-parallel 1-10. This trilayer effect is based on the spin-dependent scattering experienced by electrons in ferromagnetic metals such as Fe, Ni, Co and their alloys. Spin-dependent scattering gives rise to spin-dependent conductivities through the interplay between the electronic band dispersion and exchange interactions. Albert Fert [41] and Peter Grünberg [42] separately proposed a heterostructure design that would exploit this spin-dependent conductivity to realize the GMR phenomenon. From its functionality, this heterostructure can be thought of as the spintronics version of the p-n junction, from which various forms of spin field effect transistors [43, 44, 45, 46] and other spin-valve based devices [47, 48, 49] have been proposed. As we shall see in chapter 5, a different spin-based tunable conductivity can be achieved in graphene-based heterostructures.

![Figure 1-10: Giant magnetoresistance and the spin-valve effect. Schematic representation of the spin-valve effect in a trilayer film of two identical ferromagnetic layers F1 and F2 sandwiching a non-magnetic conductive layer M. When the two magnetic layers are magnetized parallel (left panel), the spin-up electrons (spin antiparallel to the magnetization) can travel through the sandwich nearly unscattered, providing a conductivity shortcut and a low resistance. In contrast, in the antiparallel case (right panel) both spin-up and spin-down electrons undergo collisions in either F1 or F2, giving rise to a higher overall resistance. [50].](image-url)

At the same time, the rapid progress in other fields such as cyclotron-based physics, electron paramagnetic resonance and nuclear magnetic resonance (NMR)
spectroscopy provided the impetus to generate stronger and more homogeneous magnetic fields. Often achieved via large currents in superconducting coils, strong magnetic fields have thus been added as an additional knob in solid state physics. This has enabled the study of quantum Hall physics which led to the discovery of the celebrated fractional quantum Hall effect [51], as well as the detection and characterization of SOI in 2D systems. Magnetic field strengths in the Tesla scale such as those used in NMR have enabled magnetoconductivity and quantum oscillation experiments in 2D systems. In particular, magnetoresistance measurements at the Tesla scale allow for better resolution of beating patterns that are present in the oscillating data. This is essential for precise characterization of the underlying SOI strength when the subbands are weakly spin-split [52, 53]. At larger field strengths ($\gtrsim 10 \text{ T}$), Zeeman splitting can reach the meV scale. This leads to several new developments. One is the lifting of the otherwise spin-degenerate Landau levels [54, 55]. Strong $B$ fields can also be used to probe electron correlation and interaction effects in such systems [56, 57, 58, 59, 60, 61, 62, 63, 64].

Integrating some of these spin-related phenomena into next generation devices necessarily demands their implementation in a form factor that is competitive to devices already available. Fortunately, half of the work has been undergoing constant development – the generation and control of nanoscale magnetic fields – which is of interest for a wide range of applications [65]. For example, magnetic field strengths on the order of 1 T, length scales of less than 100 nm and switched at the GHz bandwidth are essential for coherent control of spins at the nanoscale dimensions. This has applications in quantum spintronics [66, 67] nanoscale resolution sensing [68, 69, 70]. Such fields can be generated by modern hard disk drive write heads which can be traced back to the magnetic recording technology invented in 1898 by Valdemar Poulsen. Over the past decades, these write heads have acquired smaller form factors while achieving stronger magnetic fields with faster operational frequencies. As we shall see in Chapter 4, this fast switching field will enable the electric-dipole spin resonance effect in a bilayer graphene-on-TMD heterostructure.

Subsequent studies on the SOI in 2D systems eventually led to more interesting
discoveries. Notable examples include topological insulators [71, 72, 73, 74] as well as other valley-spin related phenomena in TMDs [75] such as valley-selective optoelectronics [76, 77, 78, 79]. The theoretical and experimental developments thus far puts us in a great position to investigate the strength and form of SOI that can be introduced into graphene via heterostructures. As I will discuss in subsequent chapters, such systems provide new ways to control not only spin, but also charge and valley degrees of freedom. Together with the dynamic control of localized strong magnetic fields, some of the ideas presented in this thesis may find their way into applications in quantum spintronics or nanoscale sensing devices and may be part of key components in solid state quantum computers.

1.4 Thesis Outline

In this thesis, I discuss how SOI can be introduced in graphene via proximity effect. Doing so on a bilayer graphene results in a layer-specific SOI. This can be exploited in conjunction with gate-tuneability in bilayer graphene to access novel phenomena. As we will see, the interplay between both elements is key in realizing these phenomena in heterostructures formed from bilayer graphene and TMDs.

I first introduce the unique electronic properties of graphene and bilayer graphene with an emphasis on the case of gated bilayer graphene in Chapter 2. I also discuss the different forms of SOI in graphene and bilayer graphene respectively. Unfortunately, the type of SOI allowed by their respective crystal symmetries are very weak and has not yet been measured directly in experiments. This provides the impetus to engineer SOI into graphene via heterostructures.

In Chapter 3, I present an experimental collaboration which showed that the low-energy carriers of graphene multilayers experience a substantially strong SOI which is imprinted in the magnetoconductivity measurements. Furthermore, the observed Shubnikov-de Haas oscillations show a beating pattern. This is a direct evidence that the strong SOI does not merely affect the transport properties of the low-energy carriers, but it in fact modifies the bandstructure that describes these carriers.
In Chapter 4, I discuss in greater detail the layer-specific SOI in the context of the strongly layer-polarized states in gated bilayer graphene. In particular, I show how a gate-tunable SOI is achieved in the bilayer graphene-on-TMD heterostructure. I explain how this new functionality leads to achieving electric-dipole spin resonance with the application of an in-plane magnetic field. In addition, I show that the TMD-bilayer graphene-TMD heterostructure can be gate-tuned into a new topological phase. This phase is characterized by its distinct intrinsic valley-Hall conductivity from the valley-Hall insulator phase of gated bilayer graphene.

In Chapter 5, I discuss the same bilayer-on-TMD heterostructure but now in its quantum Hall regime under a sufficiently strong transverse magnetic field. I explore how the Landau level physics in bilayer graphene is modified as a result of the interplay between the layer-specific SOI and interlayer potential. Depending on the strength of the magnetic field, spin- and orbital-ordering inversions can occur in the zeroth Landau level. I predict the existence of two new phases when moderately strong magnetic fields and interlayer bias are applied. The more interesting phase hosts a gate-tunable edge conductivity, in which switching the sign of the interlayer bias switches on or off its edge conductivity.

Finally, in Chapter 6, the main results of this thesis are summarized and potential extensions of the various ideas are discussed.
Chapter 2

Bilayer Graphene Theory

This chapter serves as an introduction to the basic theoretical description that is used to model the graphene monolayer and bilayer electronic structure. Starting from the tight-binding description, the low-energy effective Hamiltonian for each of them is derived, which will be the starting point for many of the subsequent chapters. Following which, I provide a review on the intrinsic spin-orbit interaction in graphene and explain why it is often neglected despite being allowed by symmetry. Finally, I discuss the capacitance effects in bilayer graphene and how it modifies the description provided by the single-particle effective Hamiltonian.

2.1 Monolayer and Bilayer Graphene Lattice

Graphene is a single layer of carbon atoms arranged in a honeycomb lattice, as is shown in Fig. 2-1. More precisely, the structure is understood as a triangular lattice with a basis of two atoms per unit cell that are spatially separated in the 2D plane. They are denoted as sublattice sites A and B. The primitive lattice vectors of the triangle lattice, \( \mathbf{a}_1 \) and \( \mathbf{a}_2 \), may be defined as

\[
\mathbf{a}_1 = \left( \frac{a}{2}, \frac{\sqrt{3}a}{2} \right), \quad \mathbf{a}_2 = \left( \frac{a}{2}, -\frac{\sqrt{3}a}{2} \right),
\]

(2.1)
where $a = |a_1| = |a_2| = 2.46\text{Å}$ is the lattice constant. This is necessarily distinct from the distance between nearest neighbor carbon atoms, which is the distance between adjacent A and B sublattice sites, $a_{CC} = a/\sqrt{3} = 1.42\text{Å}$. The corresponding primitive reciprocal lattice vectors $b_1$ and $b_2$ are given by

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

so that $a_i \cdot b_j = 2\pi \delta_{ij}$.

Figure 2-1: (a) Crystal structure of monolayer graphene with carbon atoms on A (B) sublattice shown as white (black) circles. The shaded rhombus is the conventional unit cell, $a_1$ and $a_2$ are primitive lattice vectors. (b) Reciprocal lattice of monolayer and bilayer graphene with lattice points indicated as crosses, $b_1$ and $b_2$ are primitive reciprocal lattice vectors. The shaded hexagon is the first Brillouin zone with $\Gamma$ indicating the centre, and $K_+$ and $K_-$ showing two non-equivalent corners [80].

There are many ways in which two graphene layers can be stacked. We will be mostly concerned with Bernal-stacked bilayer graphene. This refers specifically to the case when the B sublattice (B1) on one layer is exactly aligned to the A sublattice (A2) on the other such that the other two sublattices (A1 and B2) are respectively aligned to the center of the hexagonal unit cell of the other layer. These sublattice alignments are clearly seen from the top view, shown in Fig. 2-2. A continuous range of stacking configurations are achieved by introducing a shift or twist angle (or both) between the two layers until the two layers once again return to their Bernal-stacked configuration. In this thesis, ‘bilayer graphene’ by default therefore assumes
the reference Bernal-stacked configuration, i.e. where the twisted angle and relative shift between the two layers are both zero. Bilayer graphene therefore has the same primitive lattice vectors, reciprocal lattice vectors and Brillouin zone as monolayer graphene. It however has instead four atoms per unit cell, labeled A1 and B1 on the lower layer and A2, B2 on the upper layer.

![Diagram of Bernal-stacked bilayer graphene](image)

Figure 2-2: (a) Top and (b) side view of the crystal structure of Bernal-stacked bilayer graphene. Atoms A1 and B1 on the lower layer (layer 1) are shown as white and black circles, A2, B2 on the upper layer (layer 2) are black and grey, respectively. The shaded rhombus in (a) indicates the conventional unit cell. Various hopping parameters $\gamma_{i=0,1,3,4}$ are shown in (b) [80].

### 2.2 Monolayer Graphene Low-Energy Effective Hamiltonian

The nearest-neighbor tight-binding Hamiltonian of graphene takes the form

$$\hat{H}_{\text{MLG}} = -\gamma_0 \sum_{(i,j),s} (a_{i,s}^\dagger b_{j,s} + \text{H.c.})$$

(2.3)

where $\gamma_0 = 2.61 \text{ eV}$ denotes the nearest-neighbor hopping energy and $a_{i,s}$ ($a_{i,s}^\dagger$) annihilates (creates) an electron with spin $s$ ($s = \uparrow, \downarrow$) on sublattice A at position $R_i$. A similar notation is used for sublattice B with $a \to b$. In Fourier space, the Hamilto-
nian 2.3 is given by,
\[
\hat{H}_{\text{MLG}} = \sum_{k,s} \psi_{k,s}^\dagger H_{\text{MLG}}(k) \psi_{k,s}, \quad H_{\text{MLG}}(k) = -\gamma_0 \begin{pmatrix} 0 & f(k) \\ f^*(k) & 0 \end{pmatrix}, \quad (2.4)
\]
\[
f(k) = e^{ik_\parallel a/\sqrt{3}} + 2e^{-ik_\parallel a/2\sqrt{3}} \cos(k_xa/2), \quad (2.5)
\]
where \(\psi_{k,s}^\dagger = \begin{pmatrix} a_{k,s}^\dagger \\ b_{k,s}^\dagger \end{pmatrix}\) and \(k = (k_x, k_y)\) is the wavevector.

The energy bands given by this Hamiltonian, first derived by Wallace [8], are described by
\[
E_\pm(k) = \pm \gamma_0 |f(k)|. \quad (2.6)
\]
The \(+(-)\) sign here labels the positive (negative) energy, electron (hole)-like, conduction (valence) band. The spectrum is symmetric about zero energy and features a linear-dispersion near zero energy at the Dirac points \(K_\pm = \left(\pm \frac{4\pi}{3a}, 0\right)\) of the Brillouin zone,
\[
E_\pm(K_\xi + k) = \pm v_0 \hbar |k|. \quad (2.7)
\]
Here I introduce the valley index \(\xi = \pm\) which labels the Dirac point of interest. For their resemblance in the energy dispersion to those of massless relativistic particles, the low-energy carriers in graphene are described as massless Dirac quasiparticles but with a velocity \(v_0 = \frac{\sqrt{3}a}{2\hbar} \gamma_0 \simeq 10^6\) ms\(^{-1}\), which is about 300 times less than the speed of light. These massless Dirac quasiparticles are therefore four-fold degenerate, two from valley and two from spin.

Alternatively, the low-energy spectrum of graphene 2.7 can be solved by first expanding \(H_{\text{MLG}}\) from Eq. 2.4 about the \(K_\xi\) points to obtain the low-energy effective Hamiltonian of graphene,
\[
H_{\text{MLG}}(k) = v_0 \begin{pmatrix} 0 & \pi^\dagger \\ \pi & 0 \end{pmatrix}, \quad (2.8)
\]
\[
\pi = \hbar(\xi k_x + ik_y), \quad \pi^\dagger = \hbar(\xi k_x - ik_y).
\]
Here, \( \mathbf{k} \) is understood to be the wavevector taken with respect to the \( \mathbf{K}_\xi \) points. It is convenient to introduce this wavevector relabeling, since the low-energy carriers of the various graphene-based systems of interest are in the vicinity of these Dirac points.

### 2.3 Bilayer Graphene Effective Hamiltonian

As we saw in Fig. 2-2, bilayer graphene comprises two layers of graphene with additional interlayer hopping parameters \( \gamma_{i=1,3,4} \). The largest of these parameters is \( \gamma_1 = 0.361 \) eV, which describes the coupling between electronic orbitals on the vertically aligned A2 and B1 sites. For reasons that will become clear shortly, the A2 and B1 are known as ‘dimer sites’ and the A1 and B2 the ‘non-dimer’ sites. The parameter \( \gamma_4 = 0.138 \) eV describes the coupling between dimer and non-dimer orbitals A1 and A2 or B1 and B2, while the parameter \( \gamma_3 = 0.283 \) eV describes the coupling between the two non-dimer sites A1 and B2. Finally, the parameter \( \Delta' = 0.015 \) eV describes the energy difference between dimer and non-dimer sites. The values of these parameters are obtained from recent \textit{ab initio} calculations [81], although slightly different values have been obtained previously [80].

With these parameters, we can construct a tight-binding Hamiltonian for bilayer graphene,

\[
\hat{H}_{\text{BLG}}^0 = -\gamma_0 \sum_{\langle ij\rangle,\rho,s} (a_{\rho,i,s}^\dagger b_{\rho,j,s} + \text{H.c.}) + \gamma_1 \sum_{\langle ij\rangle,s} (a_{2,i,s}^\dagger b_{1,j,s} + \text{H.c.}) \\
+ \gamma_3 \sum_{\langle ij\rangle,s} (a_{1,i,s}^\dagger b_{2,j,s} + \text{H.c.}) + \gamma_4 \sum_{\langle ij\rangle,s} (a_{1,i,s}^\dagger a_{2,j,s} + b_{1,i,s}^\dagger b_{2,j,s} + \text{H.c.}) \\
+ \Delta' \sum_{ij,s} (a_{2,j,s}^\dagger a_{1,j,s} + b_{1,j,s}^\dagger b_{2,j,s}), \quad (2.9)
\]

where I introduced a new layer index \( \rho \) so that for instance, \( a_{\rho,i,s} (a_{\rho,i,s}^\dagger) \) annihilates (creates) an electron with spin \( s (s = \uparrow, \downarrow) \), on layer \( \rho (\rho = 1, 2) \) on sublattice A at position \( \mathbf{R}_i \).
In Fourier space, the Hamiltonian \( \hat{H}_{\text{BLG}}^0 \) is given by,

\[
\hat{H}_{\text{BLG}}^0 = \sum_{k,s} \psi_{k,s}^\dagger H_{\text{BLG}}^0(k) \psi_{k,s},
\]

\[
H_{\text{BLG}}^0(k) = \begin{pmatrix}
0 & -\gamma_0 f(k) & \gamma_4 f(k) & \gamma_3 f^*(k) \\
-\gamma_0 f^*(k) & \Delta' & \gamma_1 & \gamma_4 f(k) \\
\gamma_4 f^*(k) & \gamma_1 & \Delta' & -\gamma_0 f(k) \\
\gamma_3 f(k) & \gamma_4 f^*(k) & -\gamma_0 f^*(k) & 0
\end{pmatrix},
\]

where \( \psi_{k,s}^\dagger = \left( a_{1,k,s}^\dagger, b_{1,k,s}^\dagger, a_{2,k,s}^\dagger, b_{2,k,s}^\dagger \right) \) and \( f(k) \) is given in Eq. 2.5. Expanding \( H_{\text{BLG}}^0 \) from Eq. 2.10 about the \( K_\xi \) points, we get

\[
H_{\text{BLG}}^0(k) = \begin{pmatrix}
0 & v_0 \pi^\dagger & -v_4 \pi^\dagger & -v_3 \pi \\
v_0 \pi & \Delta' & \gamma_1 & -v_4 \pi^\dagger \\
v_4 \pi & \gamma_1 & \Delta' & v_0 \pi^\dagger \\
v_3 \pi & -v_4 \pi & v_0 \pi & 0
\end{pmatrix},
\]

where \( \pi \) and \( \pi^\dagger \) is given in Eq. 2.8 and \( k \) is understood to be the wavevector taken with respect to the \( K_\xi \) points. The various velocity parameters are given by \( v_i = 1, 3, 4 \). The bandstructure of \( H_{\text{BLG}}^0(k) \) is shown in Fig. 2-3.

Near the Dirac points, the \( k \)-independent \( \gamma_1 \) term dominates and separates the system into a high- and low-energy subspace. These correspond directly to the dimer and non-dimer orbitals respectively. The low-energy eigenstates are therefore strongly localized in the vertically-aligned, non-dimer orbitals (A1,B2). It is possible to obtain an effective 2\( \times \)2 Hamiltonian of the low-energy system by projecting \( H_{\text{BLG}} \) in Eq. 2.12 into the (A1, B2) subspace in a perturbative fashion in the limit where \( \gamma_0 \) and \( \gamma_1 \) are much larger than other energies. There are a few different but equivalent ways to do this [82] and I briefly discuss the application of one of these approaches in the appendix section of chapter 4.
It can be shown that the effective $2 \times 2$ Hamiltonian, $H_{BLG}^{2 \times 2}$, is given by [83, 84]

$$H_{BLG}^{2 \times 2} = h_0 + h_w + h_4 + h_\Delta + h_U,$$

$$h_0 = -\frac{1}{2m} \begin{pmatrix} 0 & (\pi^\dagger)^2 \\ (\pi)^2 & 0 \end{pmatrix}, \quad h_w = v_3 \begin{pmatrix} 0 & \pi \\ \pi^\dagger & 0 \end{pmatrix} - \frac{v_3a}{4\sqrt{3}h} \begin{pmatrix} 0 & (\pi^\dagger)^2 \\ (\pi)^2 & 0 \end{pmatrix},$$

$$h_4 = \frac{2v_0v_4}{\gamma_1} \begin{pmatrix} \pi^\dagger\pi & 0 \\ 0 & \pi\pi^\dagger \end{pmatrix}, \quad h_\Delta = \frac{\Delta v_0^2}{\gamma_1^2} \begin{pmatrix} \pi^\dagger\pi & 0 \\ 0 & \pi\pi^\dagger \end{pmatrix},$$

$$h_u = -\frac{u}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - \frac{2v_0^2}{\gamma_1^2} \begin{pmatrix} \pi^\dagger\pi & 0 \\ 0 & \pi\pi^\dagger \end{pmatrix}.$$

This effective Hamiltonian captures the bandstructure near the Dirac points (see inset of Fig. 2-3). The first term $h_0$ resembles the low-energy effective Hamiltonian of graphene (Eq. 2.8) but has instead a quadratic dependence of $\pi$ and $\pi^\dagger$. This can be understood as an effective hopping between the non-dimer sites involving a three step process: $A1 \rightarrow B1$ acquiring a $v_0\pi$ factor, followed by $B1 \rightarrow A2$ acquiring its ‘mass’ $\frac{1}{\gamma_1}$ and finally $A2 \rightarrow B2$, acquiring another $v_0\pi$ factor. This gives rise to an overall mass term $m = \gamma_1/2v_0^2$, so that the low-energy electrons of bilayer graphene are described as massive chiral electrons [83, 85]. This term alone captures the essential physics of the low-energy electrons in bilayer graphene. A simpler model obtained from setting the other intrinsic parameters $\gamma_3, \gamma_4$ and $\Delta'$ to 0 is often capable of explaining qualitative features observed by experiments and demonstrating novel phenomena as we will see in later chapters. It gives rise to a quadratic and isotropic dispersion of the low-energy carriers,

$$E(k) = \pm \frac{\hbar^2k^2}{2m},$$

which is not only particle-hole symmetric, but also valley and spin degenerate. The other terms break some of these symmetries which I briefly discuss.

Because of the symmetry of the lattice, at larger energies and momenta $k$ away from the Dirac points, the above isotropic dispersion relation acquires a triangular distortion. This effect occurs in graphene and graphite as well because they share the same Bravais lattice. In bilayer graphene however, an additional triangular distortion
occurs due to the first term of \( h_w \) linear in \( \pi \) and \( \pi^\dagger \). This occurs at an energy scale \(|E| < 1 \text{ meV} \) [83, 86] much smaller than what is of interest and will therefore be neglected for most of the works in this thesis. The other effect of \( h_w \) comes from its second term, which essentially renormalizes the mass \( m \) of the low-energy carriers and does not introduce qualitative changes to the physics.

The next two terms in \( H_{\text{BLG}}^{2x2} \), \( h_4 \) and \( h_\Delta \), achieve the same effect of breaking the electron-hole symmetry. When considered together with \( h_0 \), the low-energy dispersion Eq. 2.14 gets modified slightly to \( E(k) = \pm \frac{h_0^2 k^2}{2m} \left( 1 \pm \left( \frac{2 u_0}{v_0} + \frac{A^\dagger}{\gamma_1} \right) \right) \). I will be mostly neglecting these terms and only include their effects in the chapter 5.

For completeness, I have included the last term \( h_u \) here which breaks the inversion symmetry of the crystal and is in this sense, an extrinsic term. The extent of asymmetry introduced by \( h_u \) depends on the energy difference \( u \) between the two graphene layers. This is typically achieved experimentally by doping [87] or externally gating [88, 89] both layers with different potentials. As this interlayer bias will play a pivotal role in the rest of my thesis, I devote the next section to discuss its consequences and its effects on the low-energy spectrum.

### 2.4 Gated Bilayer Graphene

In monolayer graphene, applying a uniform gate potential from one side, say through the substrate, results in an overall shift in the electronic spectrum. Doing so dopes the system with electrons or holes depending on the sign of the potential. In bilayer graphene however, the two layers can be gated separately so that the potential on each layer can be written as \( u_i \). We can define new parameters \( u_0 = \frac{1}{2}(u_1 + u_2) \) and \( u = u_2 - u_1 \), which describes the overall spectrum shift and the interlayer potential respectively. The former influences the carrier density of the system while the latter breaks inversion symmetry and opens up a gap in the low-energy spectrum. As we shall see, this gate-tuneable gap provides direct control over the layer-polarization of low-energy carriers in bilayer graphene. It therefore serves as a knob to tune properties that are strongly layer-polarized.
Figure 2-3: Bilayer graphene bandstructure obtained from $H_{BLG}^0(k)$ showing the Dirac points $K_\pm$ and the Brillouin zone center $\Gamma$. The inset shows the low-energy spectrum of interest in the vicinity of the $K_+$ point (also $K_-$ point). Parameters values are $\gamma_0 = 3.16$ eV, $\gamma_1 = 0.381$ eV, $\gamma_3 = 0.38$ eV, $\gamma_4 = 0.14$ eV and $\Delta' = 0.022$ eV. Adapted from Ref. [80].
2.4.1 Low-Energy Spectrum and Eigenstates

The interlayer potential can be modeled by introducing an on-site potential term $\hat{H}_u$ to the tight-binding Hamiltonian $H^0_{BLG}$ in Eq. 2.9 (or in Fourier space Eq. 2.10),

$$
\hat{H}_{BLG} = \hat{H}^0_{BLG} + \hat{H}_u \tag{2.15}
$$

$$
\hat{H}_u = \frac{u}{2} \sum_{j,s} \left( -a^\dagger_{1,j,s}a_{1,j,s} - b^\dagger_{1,j,s}b_{1,j,s} + a^\dagger_{2,j,s}a_{2,j,s} + b^\dagger_{2,j,s}b_{1,2,s} \right) \tag{2.16}
$$

$$
H_u = \frac{u}{2} \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{pmatrix} \tag{2.17}
$$

Projecting the full bilayer graphene Hamiltonian $H_{BLG} = H^0_{BLG} + H_u$ to the low-energy subspace yields $H_{BLG}^{2\times2}$ given in Eq. 2.13. In this subspace, the constant leading term can be added directly by hand. The second term however is a $k^2$ correction term obtained from perturbative expansion. It analogous to $h_\Delta$ and therefore introduces electron-hole asymmetry.

To study the leading effects of interlayer potential on the low-energy carriers, let us consider the case where the only non-trivial parameters are $\gamma_0, \gamma_1$ and $u$. In this case, it can be shown that the spectrum of $\hat{H}_{BLG}$ is given by [83]

$$
E(k) = \pm \epsilon_\alpha(k), \tag{2.18}
$$

$$
\epsilon_\alpha^2(k) = \frac{\gamma_1^2}{2} + \frac{u^2}{4} + v_0^2 \hbar^2 k^2 + (-1)^\alpha \sqrt{\frac{\gamma_1^4}{4} + v_0^2 \hbar^2 k^2 (\gamma_1^2 + U^2)}. \tag{2.19}
$$

The low-energy bands $\alpha = -1$ has a 'Mexican hat' shape dispersion which gives rise to a band gap $u_g$ at wavevector $k_g$ from either valleys [80],

$$
u_g = \frac{|u|\gamma_1}{\sqrt{\gamma_1^2 + u^2}}, \quad k_g = \frac{|u|}{2v_0} \sqrt{\frac{2\gamma_1^2 + u^2}{\gamma_1^2 + u^2}}. \tag{2.19}
$$
In this thesis, we will be mainly concerned with the small $u$ limit for which $u \ll \gamma_1$ so that $k_g \to 0$ and $u_g \to u$. In this limit, the spectrum simplifies to

$$E(k) \simeq \pm \frac{1}{2} \sqrt{u^2 + \frac{\hbar^4 k^4}{m^2}},$$

(2.20)

which is the exact spectrum of the effective Hamiltonian obtained from adding only the leading term of $h_u$ to $h_0$,

$$H_{BLG}^{2\times 2} = -\left( \begin{array}{cc}
\frac{u}{2} & \frac{1}{2m} \left( \pi^1 \right)^2 \\
\frac{1}{2m} \left( \pi \right)^2 & -\frac{u}{2}
\end{array} \right).$$

(2.21)

The corresponding bandstructure is shown in Fig. 2-4(a). The eigenstates of this effective Hamiltonian is given by,

$$\psi(k) = \sqrt{\frac{E(k) - u/2}{2E}} \left( \begin{array}{c}
1 \\
-\frac{\hbar^2 k^2}{\gamma_1(E(k) - u/2)} e^{i2 \xi \varphi_k}
\end{array} \right) e^{i k r}, \quad \tan(\varphi_k) = \frac{k_y}{k_x},$$

(2.22)

where I use $\varphi_k$ to denote the azimuthal coordinate in $k$-space. The layer-1(2) polarization $g_{1(2)}(k)$ of an eigenstate with wavevector $k$ is given by

$$g_{1(2)}(k) = |\psi_{A1(B1)}(k)|^2 = \frac{1}{2} \mp \frac{u}{4E(k)}.$$  

(2.23)

In the vicinity of the Dirac point therefore, the eigenstates are strongly layer polarized when $u$ is finite. For instance, when $u > 0$, we find

$$\lim_{k \to 0} g_{1(2)}(k) \rightarrow \begin{cases} 
0(1), & E > 0 \\
1(0), & E < 0
\end{cases}.$$  

(2.24)

This implies that electrons and holes of gated bilayer graphene are strongly localized on opposite layers. These carriers of opposite charges are therefore asymmetrically sensitive to any layer-specific effects, e.g. due to having different materials adjacent to the opposite layers. I will explain how this special property of bilayer graphene leads to several novel phenomena in the later chapters.
Figure 2-4: Layer polarization of low energy states. (a) Low energy dispersion with non-trivial interlayer bias $u > 0$, opening up a gap separating the conduction band $E_+$ in red, and valence band $E_-$ in blue. (b) Layer-1 polarization of the corresponding low-energy bands in (a). Near the valley $p = 0$, states from the conduction band (red) are strongly polarized in layer 2 while states from the valence band (blue) are strongly polarized in layer 1.

### 2.4.2 Capacitance Effects

For a finite chemical potential, integrating the layer polarization Eq. 2.23 over the circular Fermi surface then gives the electron densities $n_{1(2)}$ on the respective layers,

$$n_{1(2)} = 4 \int \frac{d^2 k}{(2\pi)^2} g_{1(2)}(k) = \frac{2}{\pi} \int dk \frac{k}{g_{1(2)}(k)}. \quad (2.25)$$

In the above expression, a factor of four has been included to account for spin and valley degeneracies. When there is no asymmetry between the layers, an excess charge $n = n_1 + n_2$ induced necessarily splits equally between the two layers, $n_1 = n_2 = n/2$.

However, when layer asymmetry arises, this charge redistribution is no longer symmetric, $n_1 \neq n_2$. By treating each graphene layer as a conducting plate, we see that bilayer graphene acts like a capacitor, so that the interlayer potential necessarily satisfies a self-consistent equation [90]

$$u(n) = u_0 + \frac{e^2 n_2 c_0}{\epsilon_r \epsilon_0}, \quad (2.26)$$

where $\epsilon_r$ is the bilayer graphene dielectric constant and $u_0$ is the bare interlayer bias, i.e. $u(n = 0) = u_0$. This can be understood as an electron-electron interaction effect,
which is essentially the most significant Hartree potential for the bilayer graphene geometry.

The self-consistent treatment of carrier density is particularly important for accurately describing its dependence on the interlayer bias especially in the case when bilayer graphene is gated only from one side. I will discuss this case in greater detail in section 3.6 of the following chapter. When bilayer graphene is dual-gated, taking account of the capacitance effect is crucial for properly extracting the interlayer potential and chemical potential of the system from experimental inputs of top and bottom gate voltages. Furthermore, the capacitance effects of bilayer graphene can be used to probe layer-specific properties that are normally out of reach in transport measurements [91].

2.5 Intrinsic Spin-Orbit Interaction in Graphene

Spin-orbit interaction (SOI) is a relativistic effect which couples the spin and orbital degrees of freedom of electrons. When an electron travels with a velocity orthogonal to an external electric field, it experiences an effective magnetic field in its rest frame. This magnetic field points in the direction orthogonal to both its velocity and the electric field. It couples to the electron's spin, so that the coupling occurs between the electron's spin and orbital motion in the laboratory frame. Consequently, the effect is typically strong in crystals comprised of heavy ions where the average velocity of electrons is high. Carbon is a light atom, so that a crystal formed by carbon atoms alone such as graphene is expected to have weak SOI.

In the tight binding model, SOI manifest as spin-dependent and direction-dependent hopping terms. In the absence of magnetic field, the crystal symmetries and time-reversal symmetry are respected. This limits the types of spin-dependent hoppings allowed. Because these SOI terms respect these symmetries, they are said to be intrinsic.

In graphene, the operators corresponding to time reversal ($T$) and inversion
(I_{MLG}) are given by

\[ \mathcal{T} = i\tau_x s_y \mathcal{K}, \quad (2.27) \]

\[ I_{MLG} = \tau_x \sigma_x, \quad (2.28) \]

where \( \mathcal{K} \) refers to complex conjugation, \( \sigma_{x,y,z} \) and \( s_{x,y,z} \) are the Pauli matrices corresponding to the sublattice and spin degrees of freedom respectively. The intrinsic SOI at the \( K_\pm \) points take the form \([72, 92, 93]\)

\[ H_{MLG}^{ISO} = \Delta_I \tau_z \sigma_z s_z, \quad (2.29) \]

where \( \Delta_I \) is the intrinsic SOI interaction strength. This interaction acts as an effective mass that has opposite signs for opposite sublattices (\( \sigma_z \)), opposite spins (\( s_z \)) and at opposite valleys \( K_\pm \) (\( \tau_z \)). It is this form of SOI that opens a topological gap at the Dirac points so that the system at charge neutrality is quantum spin-Hall insulator. Quantum spin-Hall insulators feature counter-propagating edge modes with opposite spin-polarizations that are topologically protected from backscattering. This gives rise to quantized charge and spin conductances \([72]\).

More generally, when \( s_z \) does not commute with the Hamiltonian such as in the presence of a Rashba-type SOI,

\[ H_R = \lambda_R (\tau_z \sigma_x s_y - \sigma_y s_z) \quad (2.30) \]

the system is a topological insulator for sufficiently large values of \( \Delta_I > \lambda_R / 2\sqrt{3} \) \([94]\). While the edge modes of topological insulators no longer have spin-polarizations along the \( z \)-direction, they nonetheless enjoy the same topological protection from backscattering. Topological insulators therefore also feature quantized charge conductance.

Various approaches have been used to estimate the magnitude of \( \Delta_I \) ranging from 1 \( \mu eV \) to 100 \( \mu eV \) \([95, 96, 97]\). This is much smaller than the characteristic SOI of 2p electrons in carbon (\( \sim 6 \) meV). The strength of the SOI diminishes when projected onto the atomic \( \pi \) bands of a flat graphene layer \([92, 93]\) and is shown to be weaker
Figure 2-5: Quantum spin Hall effect. (a) Energy bands for a ribbon with “zig-zag” edges in the quantum spin Hall phase [94]. (b) Schematic diagrams showing (upper panel) a quantized charge current of $2\frac{e}{h}V$ flowing to the right in a two terminal measurement geometry and (lower panel) quantized spin current of $\frac{e}{4\pi}V$ flowing to the right. Diagrams to the right indicate the population of the edge states [72].

than curved graphene sheets or nanotubes [92, 98, 99]. Because of its small value, it has not yet been possible to directly measure $\Delta_I$ from experiments, although an upper bound value of $\sim 100 \mu eV$ has been set [5].

In bilayer graphene, the inversion operator ($I_{BLG}$) is given by

$$I_{BLG} = \tau_x \rho_x \sigma_x,$$  \hspace{1cm} (2.31)

where $\rho_{x,y,z}$ are the Pauli matrices corresponding to the layer degree of freedom. Consequently, the intrinsic SOI at the $K^\pm$ points takes the form

$$H_{BLG}^{ISO} = \lambda_1 \tau_z \sigma_z s_z + \lambda_2 \tau_z \rho_z s_z + \lambda_3 \rho_z (\sigma_y s_x - \tau_z \sigma_x s_y) + \lambda_4 \sigma_z (\rho_y s_x + \tau_z \rho_z s_y).$$ \hspace{1cm} (2.32)

In particular, the first term is the intrinsic SOI of graphene applied to both layers; the second and third terms have opposite signs in the two layers and correspond to an Ising-like valley-dependent SOI and a Rashba-type SOI respectively; the last term is a layer-spin analogue of the Rashba-type SOI with opposite signs on different sublattice
sites. Likewise, the magnitudes of these parameters $\lambda_{i=1,2,3,4}$ have been estimated to be roughly in the range of 1 to 100 $\mu$eV [100, 101, 102], so that just like in monolayer graphene, SOI in bilayer graphene is essentially absent.

Applying a transverse electric field through gating breaks the inversion symmetry. Although this generates new extrinsic SOI terms, the strengths of these terms are strongly suppressed due to the relativistic nature of SOI. For instance, typical electric fields strengths of 1V/nm are estimated to generate Rashba-type spin splittings of $\sim$10 to 50 $\mu$eV in graphene [93, 102], comparable to that of the intrinsic SOI strengths.

Inversion symmetry can alternatively be broken by introducing a substrate. If the substrate material is made up of light atoms such as hBN, any extrinsic SOI terms generated will once again be weak. But as we shall see in the next chapter, by choosing substrates known to have strong SOI such as WSe$_2$ or other TMDs comprising heavy atoms, a substantially strong SOI in the meV scale can be interfacially induced in graphene. Although these terms are different from the intrinsic SOI that is necessary to realize the topological insulating phase [72, 103, 104, 105], I will show that the SOI induced from TMD-substrates gives rise to interesting topological properties that are gate-tunable in the case of bilayer graphene (see chapters 4 and 5).
Chapter 3

Origin and Magnitude of Spin-Orbit Interaction in Graphene on Semiconducting Transition Metal Dichalcogenides

In this chapter, I describe a collaborative work with the University of Geneva in which a combination of experimental techniques were used to demonstrate a general occurrence of spin-orbit interaction (SOI) in graphene on transition metal dichalcogenide (TMD) substrates. Their measurements indicate that SOI is ultra-strong and extremely robust, despite it being merely interfacially-induced, with neither graphene nor the TMD substrates changing their structure. This is found to be the case irrespective of the TMD material used, of the transport regime, of the carrier type in the graphene band, and of the thickness of the graphene multilayer. Specifically, they performed weak antilocalization (WAL) measurements as the simplest and most general diagnostic of SOI, and show that the spin relaxation time is very short (approximately 0.2 ps or less) in all cases regardless of the elastic scattering time, whose

---

This chapter is reproduced in part with permission from Z. Wang, et al., Origin and magnitude of 'designer' spin-orbit interaction in graphene on semiconducting transition metal dichalcogenides, Phys. Rev. X 6, 041020 Copyright (2016) by the American Physical Society.
value varies over nearly two orders of magnitude. Such a short spin-relaxation time strongly suggests that the SOI originates from a modification of graphene band structure. This expectation was confirmed through the observation of a gate-dependent beating, and a corresponding frequency splitting, in the low-field Shubnikov-de Haas magneto-resistance oscillations in high quality bilayer graphene devices on WSe$_2$. These measurements provide an unambiguous diagnostic of a SOI-induced splitting in the electronic band structure, and their analysis allows us to determine the SOI coupling constants for the Rashba term and the so-called spin-valley coupling term, i.e., the terms that were recently predicted theoretically for interface-induced SOI in graphene. The magnitude of the SOI splitting is found to be on the order of 10 meV, more than 100 times greater than the SOI intrinsic to graphene. Both the band character of the interfacially induced SOI, as well as its robustness and large magnitude make graphene-on-TMD a promising system to realize and explore a variety of spin-dependent transport phenomena, such as, in particular, spin-Hall and valley-Hall topological insulating states.

3.1 Interfacial Interactions in vdW Heterostructures

Van der Waals (vdW) heterostructures formed by vertical stacks of different two-dimensional (2D) materials have emerged recently as designer systems, providing a new paradigm for engineering novel electronic media with widely tunable parameters [15]. Stacked vdW heterostructures nicely combine the ability to tailor interfacial interactions at the atomic scale, while at the same time preserving the integrity of individual layers. This ‘designer’ approach is epitomized by recent work on graphene paired with hexagonal boron-nitride (hBN) [106, 107, 108]. In this system, a dramatic change in the graphene band structure occurs when the crystal axes of graphene and hBN layers are nearly aligned, in the total absence of any reorganization of chemical bonding or any change in the atomic order of individual layers. The transformed band structure manifests itself in striking and robust transport phenomena —such as the appearance of so-called satellite Dirac points— that are readily observable experimentally.
These unexpected findings are opening up a wide avenue of research exploring vdW heterostructures based on many different 2D materials [22, 21, 109, 110, 111, 112]. A key goal at this stage is to identify the interfacial interactions that can alter specific electronic properties of interest, and to understand the microscopic physical processes responsible for their origin.

One fascinating question in this vein is whether vdW heterostructures can be used to control not only the orbital dynamics of electrons in graphene, but also their spin, i.e., whether vdW heterostructures can be employed to generate a strong spin-orbit interaction (SOI) in the graphene Dirac band. To this end, combining graphene with large-gap semiconducting transition metal dichalcogenides (TMDs) appears to be a promising route [113], because semiconducting TMDs exhibit an extremely strong SOI [114, 75, 115, 116], and because they are known to preserve the high electronic quality of graphene when used as substrates [117]. Recent magneto-transport measurements performed on graphene-on-WS$_2$ [118, 119]—and in particular the observation of a pronounced weak antilocalization (WAL) contribution to the conductivity of graphene [119]—confirm these expectations. Indeed, the analysis of the experimental results indicated that the spin relaxation time $\tau_{so}$ in graphene-on-WS$_2$ is between 100 and 1000 times shorter than in graphene on SiO$_2$ [120] or on hBN [121]. This is broadly consistent with ab-initio calculations, which predict the strength of SOI in graphene-on-WS$_2$ to be at least 100 times larger than the SOI intrinsic to graphene [119, 92, 93, 122].

These results pose a number of interesting and challenging questions, which are central for our understanding of the new phenomenon of ‘designer’ SOI. In particular, the physical process by which strong SOI can be imprinted by one layer on an adjacent layer without any changes in their structure remains puzzling. So far, the experiments were unable to elucidate the microscopic mechanism responsible for the strong SOI in graphene, nor did they provide any reliable insight into the functional form of the induced SOI as well as its strength. They also did not establish whether a strong interfacially-induced SOI is unique to graphene-on-WS$_2$ or whether it is a robust, generic property of all graphene-on-TMD heterostructures. Last but not least,
perhaps the most tantalizing question of all is whether SOI is dominated by disorder scattering or by a band structure modification. If the latter happens to be the case, the strong SOI present in graphene-on-TMD system can be employed to create and explore a variety of electronic media with novel properties [94, 72, 123].

A variety of graphene-on-TMD heterostructures were exploited to tackle these questions in a comprehensive way. Large part of the work focuses on the study of WAL in heterostructures formed by graphene and one of the semiconducting TMDs: WSe$_2$, MoS$_2$, and WS$_2$. For all TMDs used, irrespective of carrier mobility ($\mu$), position of the Fermi level in graphene, and thickness of the graphene layer (up to trilayer), a pronounced WAL signal is observed. This finding shows that interfacially induced SOI in graphene-on-TMDs is an extremely robust phenomenon, insensitive to virtually all details of the vdW heterostructure considered. A quantitative analysis of the WAL data allows us to establish an upper bound of approximately 0.2 ps for the spin-relaxation time $\tau_{so}$, irrespective of the carrier mobility (which was varied by nearly two orders of magnitude). Such a short $\tau_{so}$ value appears to be physically compatible only with SOI originating from a modification of the graphene band structure. This conclusion was validated from measurements of Shubnikov-de Haas (SdH) conductance oscillations, which exhibited a beating due to a splitting in their frequency [124, 125, 53]. The size of the splitting and its dependence on carrier density show that the beating originates from SOI, and its quantitative analysis allows us to establish the SOI magnitude. We find that the dominant SOI term is of the Rashba type, and that its characteristic energy is approximately $\lambda_R \approx 10 - 15$ meV; the strength of the other SOI term expected to be induced by interfacial interactions [119] –i.e., the one that couples spin and valley– ranges between $\lambda = 0$ and $\lambda \approx 5-6$ meV (i.e., experimental data are compatible with $\lambda_R = 15$ meV and $\lambda = 0$ meV or with $\lambda_R = 10$ meV and $\lambda = 5 - 6$ meV, as well as different choices in these intervals).

Besides elucidating most aspects of interfacially induced SOI in graphene, the results presented here clearly illustrate the experimental flexibility of graphene-on-TMDs heterostructures. These heterostructures allow comparative studies by vary-
ing the specific TMD material used, the thickness of the graphene layer, the position of the Fermi level, and the scattering time ($\tau$). Even more flexibility could be introduced using double gated devices, to tune the graphene band structure (e.g., in bilayers) [88, 58, 126, 127] and further extend the range of carrier densities accessible experimentally. We anticipate that this unrivaled experimental flexibility will prove useful in future experiments aiming at exploring other aspects of interfacial interactions in vdW heterostructures.

3.2 Device Fabrication and Characterization

Van der Waals heterostructures were assembled by transferring graphene layers of different thickness (mono, bi, or tri layers) onto thin exfoliated flakes of TMDs (WSe$_2$, WS$_2$, and MoS$_2$), resulting in devices whose cross section is schematically shown in Fig. 3-1(a). For the assembly, a commonly used dry transfer technique was employed [16, 128]. Conventional electron-beam lithography, lift-off, and oxygen plasma etching techniques were employed to pattern and contact multi-terminal Hall-bar devices. The contacts consist of an evaporated Ti/Au thin film (10/70 nm). All structures were realized on substrates consisting of degenerately doped Silicon covered with a 285 nm thick layer of thermally grown SiO$_2$. The charge density ($n$) of graphene is tuned by operating the doped Silicon substrate as a gate electrode. In this configuration, for sufficiently large gate voltage ($V_g$), carriers are accumulated at the surface of the TMD flake at the interface with the SiO$_2$ layer [119, 129]. When that happens, the carrier density in graphene cannot be tuned anymore by $V_g$, and the conductivity ($\sigma$) of graphene saturates.
Figure 3-1: Basic characterization of graphene on TMD substrates. (a) Schematic cross section of the devices. The graphene layer (green) is transferred onto a TMD crystal (pink) that has been previously exfoliated on a substrate consisting of highly doped Silicon (gray) covered by 285-nm-thick SiO$_2$ (light blue). The silicon substrate is operated as a back-gate. (b-c) AFM images of a graphene Hall-bar device before and after the AFM-ironing process (scale bar is 2 μm long). (d-f) Gate-voltage ($V_g$) dependence of the resistance ($R$; red curves) and the conductivity ($\sigma$; blue curves) of monolayer graphene on WSe$_2$ (d), MoS$_2$ (e) and WS$_2$ (f), measured at 4.2 K. The carrier mobility in the three cases is 110,000, 33,000, and 23,400 cm$^2$/Vs, respectively. In all devices, the conductance saturates at large enough $V_g$ away from charge neutrality point, when charges start to be accumulated at the SiO$_2$-TMD interface. The black line in (f) represents the $V_g$-dependent conductivity $\sigma(V_g)$ before AFM-ironing (the corresponding mobility is approximately two times smaller than for the blue curve). (g-h) Integer quantum-Hall effect (QHE) observed in high-quality graphene-on-WSe$_2$ at $T = 250$ mK whose basic characterization is shown in (d). The color map of the longitudinal resistance ($R_{xx}$) versus $V_g$ and $B$ (g) and the $V_g$-dependence of $R_{xx}$ (red curves) and the Hall conductance ($\sigma_{xy}$; blue curves) measured at $B = 12$ T (h) clearly confirm the occurrence of the vanishing $R_{xx}$ and concomitantly quantized $\sigma_{xy} = \nu \times e^2/h$ at integer values of filling factor ($\nu = nh/eB$). In panel (h), the QH plateaus at $\nu = \pm 1$, -3, -4, -5 due to the full degeneracy lifting of the $N = 0, 1$ Landau levels in monolayer are clearly visible.
The device quality—i.e., the carrier mobility and the inhomogeneity in carrier
density—depends on the details of the assembly process, and on the procedure to
“clean” the device structure at the end of the fabrication. Two key elements intro-
duce disorder: structural defects resulting from the transfer process, such as “bubbles”
and “wrinkles” in the graphene layer [130], and adsorbates adhering onto graphene
(mostly polymer residues remaining at the end of the fabrication process). The in-
fluence of both elements can be controlled in different ways. Structural defects can
be eliminated by defining the graphene Hall-bar in regions in which these defects are
absent. Selecting these areas, which typically have linear dimensions of the order of
5-6 μm, usually results in very high carrier mobility: low-temperature mobility values
as large as 160,000 cm²/Vs have been observed, which are comparable to (or possibly
even slightly better than) the best values observed in graphene-on-hBN structures as-
sembled by the same dry-transfer technique [106, 107, 108, 131, 132]. Selecting larger
areas is also possible, but this unavoidably prevents the full exclusion of structural
defects, resulting in lower mobility.

Adsorbates can be eliminated in a rather controlled way by a so-called atomic force
microscope (AFM) “ironing” process [133], essential to realize high mobility devices.
AFM ironing consists of scanning the graphene flake with an AFM tip in contact
mode, applying only a moderate force, in such a way to pile up all the adsorbates just
outside the edges of the graphene flake. The effectiveness of the process is illustrated
in Figs. 1(b-c). Fig. 3-1(b) shows an AFM image of a device at the end of fabrication
and Fig. 3-1(c) another image of the same device taken after the “ironing” step. The
difference—the extremely small corrugation that is measured on graphene after the
“ironing” process—is clear. In this study 16 different devices were measured, in which
—depending on their area, cleaning procedure adopted, density of “bubbles”, etc.—the
carrier mobility extracted from measurements of the conductivity and of Hall effect
ranged between 3,000 and 160,000 cm²/Vs.

Figs. 3-1(d-f) shows the gate voltage dependence of the resistance (red curves)
and of the corresponding conductivity (blue curves) measured on three representative
devices, respectively on WSe₂, MoS₂, and WS₂. As compared to an earlier work on
graphene-on-WS₂, in which no AFM ironing was done [119], in the current generation of higher-quality devices the charge neutrality point is “exposed” in all cases: it is possible to shift the Fermi level \( (E_F) \) both in the valence and in the conduction band by acting on the back gate. Whereas for graphene on WS₂ and MoS₂ only a small range of energies in the conduction band can be accessed, for WSe₂, \( E_F \) can be shifted over a rather large interval in both the valence and the conduction band. Hence, WSe₂ allows the systematic investigation of SOI in the conduction band without the need to use a top gate electrode, something that could not be done in previous work. Finally, Figs. 1(g-h) show that all integer quantum Hall effect (QHE) states are visible, including the symmetry broken states caused by the presence of electron-electron interactions [131], which is indicative of the high quality of the devices (in the best cases, symmetry broken states become already visible for applied magnetic fields as low as approximately 1 Tesla).

3.3 Extracting SOI from Weak Antilocalization (WAL) in Monolayer Graphene on Different TMDs

Weak antilocalization is a striking quantum interference effect originating from spin-orbit coupling that has long served as a direct probe of SOI in conductors [134, 135]. Also for graphene-on-TMD heterostructures, the observation of the WAL correction to the low-temperature magneto-conductivity provides the simplest and most general diagnostic of the presence of SOI [119]. Extracting the WAL contribution requires suppressing the effect of the so-called phase-coherent universal conductance fluctuations (UCF) originating from random interference of electronic waves [136]. Indeed, since the dimensions of the graphene-on-TMD devices studied are typically comparable to (or even smaller than) the phase coherence length \( L_\phi \), the WAL contribution in any individual measurement is normally eclipsed by the presence of UCF. To make the WAL contribution stand out, the magnitude of the UCF was suppressed by looking at the ensemble averaged conductivity, obtained by averaging many (typically 50)
Figure 3-2: Negative magneto-conductivity due to WAL in monolayer graphene on TMD substrates. (a-c) Ensemble-averaged magneto-conductivity measured on monolayer graphene on WSe₂, MoS₂ and WS₂, for $T = 250$ mK (black curve), $T = 1.6$ K (red curves), and $T = 20$ or 30 K (blue curves), with the Fermi level gate-tuned to be in the graphene valence band. The data are measured on the same devices whose $V_g$-dependent transport curves are shown Figs. 1(d-f). The characteristic peak due to WAL around $B = 0$ T is clearly visible at low temperature and disappears at higher temperature. (d-f) Quantum corrections to magneto-conductivity of monolayer graphene on WSe₂, MoS₂, and WS₂, obtained by subtracting the classical contribution (corresponding to the magneto-conductivity measured at $T = 20$ or 30 K) from the magneto-conductivity measured at $T = 250$ mK. Note that up to the highest magnetic field investigated, no signatures of weak-localization are visible.

magneto-conductance traces measured at slightly different values of $V_g$ (the procedure is identical to that described in Ref. [119]).

The procedure described above, performed at different temperatures, leads to the results shown in Figs. 2(a-c) for graphene on WSe₂, MoS₂, and WS₂ respectively. In all cases, a negative magneto-conductivity of order $e^2/h$ is clearly apparent at the lowest temperature investigated, $T = 250$ mK, upon the application of a magnetic field ($B$) of a few milliTesla. The magnitude of the negative magneto-conductivity decreases upon warming up the devices and the effect disappears entirely at $T \approx 20 - 30$ K, as expected for quantum interference effects [135]. It is clear that, irrespective
of the TMD used to realize the heterostructures, the presence of a pronounced WAL signal in magneto-transport demonstrates that in all cases SOI is induced in graphene.

In extracting the strength of SOI from magnetotransport measurements a special care should be taken in accounting for the interplay between WAL and weak localization, since the two effects contribute to magnetoconductivity with opposite signs [137]. Weak localization gives rise to a positive magneto-conductivity, which becomes visible when during the phase-coherent propagation of electrons their spins do not rotate or rotate by a small enough angle [134]. Despite the presence of SOI, a small positive magneto-conductivity due to weak localization may still be observed, because at sufficiently high magnetic field only the shortest trajectories give a non-negligible contribution to the interference effects probed by the ensemble-averaged conductivity. Unless SOI is extremely strong, the electron spin may not have time to rotate by a sufficient amount along these short trajectories, and signatures of weak localization may then become visible at large $B$.

The issue is relevant because—as it is clear from Figs. 2(a-c)—a positive magneto-conductivity is visible in the measurements, and it is important to establish whether
this is a manifestation of weak localization. To this end, we recall that WAL and weak localization are quantum corrections to the conductivity, i.e. they correspond to the difference between the total magneto-conductivity that is actually measured and the classical contribution. The classical contribution is straightforward to determine, since it corresponds to the magneto-conductivity measured at sufficiently high temperatures, where phase coherent effects have been suppressed because of the thermally induced shortening of $L_\phi$. Since at 20-30 K the effect of WAL has entirely disappeared and the remaining low-field (positive) magneto-conductivity does not exhibit any significant temperature dependence, we can take the magneto-conductivity measured at these temperatures to be a good approximation of the classical contribution (see Section 3.11 for more details on the classical magneto-conductivity background). The resulting quantum correction to the magneto-conductivity for the different devices is shown in Figs. 2(d-f). Within the precision of the measurements, determined by the remnant amplitude of UCF fluctuations, no positive magneto-conductivity is visible in Figs. 2(d-f). We conclude that irrespective of the TMD material used in the heterostructure, a clear WAL signal is always present with no detectable weak localization contribution. This observation provides a first clear indication that in all heterostructures investigated, the SOI induced in graphene has a very strong intensity.

Data measured on other monolayer graphene devices confirm that WAL always occurs with no detectable weak localization signal, irrespective of whether carriers are electrons or holes, and of their mobility (or, equivalently, scattering time $\tau$), which was varied over a range of nearly two orders of magnitude. Both aspects had not been addressed in a previous work on WS$_2$, in which the Fermi level could not be shifted into the conduction band and the scattering time was only varied by a limited amount [119]. The occurrence of WAL for electron transport is best illustrated with data measured on graphene-on-WSe$_2$, shown in Fig. 3-3(a), in which a fully developed WAL signal is clearly visible. The effect of the mobility can be appreciated by looking at Figs. 3-3(b) and 3-3(c). Fig. 3-3(b) shows the quantum correction to the magneto-conductivity measured on a large-area graphene-on-WSe$_2$
device, in which a high-density of “bubbles” led to a low-temperature mobility of 3,000 cm²/Vs (whereas all devices shown in Figs. 2(a-c) had mobility larger than 25,000 cm²/Vs). We find that the magnitude of the WAL correction is comparable in all cases, but in the lower mobility devices the magnetic field required to observe the negative magneto-conductivity is larger. This is expected, since when the mobility is lower a larger magnetic field is needed to pierce a flux of $\Phi_0 = h/e$ through the area in which the electronic waves propagate phase coherently and interfere. The same conclusions can be drawn by looking at Fig. 3-3(c), which shows the WAL magneto-conductivity in a same graphene-on-WS₂ device measured before (red curve) and after (black curve) performing an AFM “ironing” step, resulting in a mobility increase. It is apparent that in this case as well, a higher mobility leads to a decrease of the magnetic field scale needed to suppress the effect of WAL.
Figure 3-4: Interfacially-induced SOI in graphene multilayers on TMDs. Panels (a-c) and (d-f) refer respectively to a bilayer and a trilayer device, both realized on WSe$_2$. Panels (a) and (d) show that the resistance peak around charge neutrality is very narrow, indicative of a high uniformity of carrier density (in the trilayer, the magnitude of inhomogeneity is smaller than $2 \times 10^{10}$ cm$^{-2}$, comparable to the best devices on hBN). High quality integer quantum Hall effect is observed in both devices, as shown by data in panels (b) and (e) measured at $T = 4.2$ K. In panel (b), in which the characteristic quantization sequence of graphene bilayer is observed, the applied magnetic field is $B = 4$ T. Panel (e) shows a color plot of the longitudinal resistance of trilayer graphene as a function of filling factor and magnetic field. Full breaking of the degeneracy of Landau levels due to electron-electron interactions start to be clearly visible already at $B = 1$ T. Panels (c) and (f) show the ensemble averaged magneto-conductivity measured at $T = 250$ mK, exhibiting the characteristic signature of WAL, with no positive magneto-conductivity due to weak localization (in panel (c) the different curves show data at different gate voltages, corresponding to shifting the Fermi level from the bilayer graphene valence band to the conduction band).
3.4 SOI in Bilayer Graphene and Thicker Multi-Layers

When monolayer graphene is substituted by thicker multilayers, a strong SOI remains present in all devices (irrespective of the TMD used for the device realization). Fig. 3-4 illustrates this conclusion with data measured on two different graphene-on-WSe$_2$ devices: Figs. 3-4(a-c) refer to a bilayer graphene device with low-temperature mobility $\mu = 33,000$ cm$^2$/Vs and Figs. 4(d-f) to a trilayer having mobility $\mu \approx 110,000$ cm$^2$/Vs. Basic transport characterization show that—as for monolayer devices—heterostructures based on thicker multilayers exhibit an excellent electronic quality. The resistance peak around the charge neutrality point is extremely sharp in both cases (see Figs. 4(a) and 4(d)); in the trilayer device the measured width corresponds to a charge inhomogeneity as low as $1.8 \times 10^{10}$ cm$^{-2}$, comparable to the best ever reported for non-suspended graphene devices. In the presence of a perpendicular magnetic field $B = 4$ T the expected Hall effect quantization sequence is observed in the bilayer device, with plateaus in the Hall conductance occurring at $\sigma_{xy} = 4N e^2/h$ ($N = \pm 1, \pm 2, ...$) [85, 83], concomitantly with the vanishing of the longitudinal resistance. In the thicker multilayer, the plot of the longitudinal resistance versus filling factor $\nu \equiv nh/eB$ and $B$ (Fig. 3-4(e)) shows the appearance of broken symmetry quantum Hall states already at $B$ as low as approximately 1 T. In short, excellent quality bi and tri-layer graphene devices can be realized on TMD substrates, comparable to the very best devices realized on hBN by means of the same technique.

Figs. 4(c) and 4(f) show that a pronounced low temperature negative magnetoconductivity due to WAL is clearly visible in both the bilayer and the trilayer device. In these devices as well, no background due to weak localization is observed, indicative of the large SOI strength. This is remarkable, because interfacial interactions are expected to modify only the properties of the bottom graphene layer, the one in direct contact with the TMD crystal [119]. Under normal conditions, the eigenfunctions in the different bands of the multilayer are such that electrons have a finite amplitude of probability to be in that layer. As a result, all bands are affected and that is why thicker multilayers exhibit a pronounced WAL. Nevertheless, it is
clear that the effect of interfacially induced SOI should decrease in intensity as the thickness of the multilayer increases, since the probability for electrons to enter in contact with the TMD crystal decreases upon increasing thickness (or, equivalently, the amplitude of the electron wavefunctions in the bottom layer –the one in contact with the the TMD– decreases for thicker multilayers). The data shown in Fig. 3-4(c) –and especially those shown in Fig. 3-4(f)– indicates that despite the larger thickness, at least up to trilayer graphene the observed behavior of the WAL correction is the one typical of very strong SOI. Significantly thicker multilayers are needed to “dilute” the effect of SOI induced by interfacial interactions.

The observation of WAL in graphene bilayers (Fig. 3-4(c)) is worth an additional comment. In monolayers, WAL can occur due only to the Dirac nature of electrons (i.e., in the absence of SOI), as a consequence of the $\pi$ Berry phase picked up by the electron wave functions that undergo backscattering processes while staying in a same valley. WAL due to this effect has been seen experimentally, albeit only at elevated temperatures (typically $T \geq 10$ K) [138], since only then $L_\phi$ is sufficiently short (such a dependence on temperature allow the phenomenon to be discriminated from WAL due to SOI, which increases in amplitude upon cooling). The effect is absent in bilayer graphene, since in bilayers a $2\pi$ Berry phase is acquired by the electron wavefunction upon back-scattering, which does not lead to WAL [139]. As such, the occurrence of WAL in bilayers illustrated by the data shown in Fig. 3-4(c) provides unambiguous and more direct evidence of the presence of interfacially induced SOI.

In concluding this Section, I emphasize that the possibility of using interfacial interactions with a TMD substrate to induce strong SOI in different graphene multilayers –and not only in monolayers– adds to the flexibility of this experimental system. As I will discuss in Section 3.6, this flexibility was exploited in the present work to determine quantitatively the type and magnitude of the interfacially induced SOI.
Figure 3-5: Comparison of WAL data measured on monolayer graphene on different TMDs with the theoretical predictions of Eq. (1). In panels (a-d), the black dots represent the data; the continuous lines of different colors represent the predictions of Eq. (1) for different values of spin-relaxation time $\tau_{so}$. In panel (a), data from a low-mobility device ($\mu = 3,000$ cm$^2$/Vs) on WSe$_2$ are shown; Panels (b-d) show data on higher mobility devices (respectively, 23,400, 33,000, and 110,000 cm$^2$/Vs) on WS$_2$ (b), MoS$_2$ (c), and WSe$_2$ (d). Note in all cases, inserting values of $\tau_{so} > 0.5$ ps in Eq. (1) leads to the appearance of a positive magneto-conductivity due to weak localization at higher $B$, which is not seen in the experiments. This allows us to determine an upper bound for $\tau_{so}$ for all the devices investigated. All data in this figure have been measured at $T = 250$ mK.
Figure 3-6: Upper bound of $\tau_{so}$, extracted from WAL data, as a function of $\tau$ for monolayer graphene devices. The shape of the different symbols refers to devices realized on different TMD substrates (triangles for WSe$_2$, circles for MoS$_2$, and squares for WS$_2$). For each different symbol, data points represented with the same color refer to the same device measured at different gate voltages. The broken line corresponds to $\tau_{so} = \tau$. 
3.5 Quantitative Analysis of the WAL Data

The pronounced WAL observed in the measurements presented in the previous sections, together with the absence of any signature of weak localization, illustrate that strong SOI is a general property of graphene-TMD heterostructures. For a quantitative analysis, we confine ourselves to devices realized on monolayers, for which an explicit theoretical expression for the magneto-conductivity due to WAL is available \cite{140}. The expression, which takes into account the effect of all possible symmetry-allowed SOI terms, reads:

$$\Delta\sigma(B) = -\frac{e^2}{\pi h} \left[ F\left(\frac{\tau^{-1}_B}{\tau^{-1}_\phi}\right) - F\left(\frac{\tau^{-1}_B}{\tau^{-1}_\phi + 2\tau^{-1}_{asy}}\right) - 2F\left(\frac{\tau^{-1}_B}{\tau^{-1}_\phi + \tau^{-1}_{so}}\right)\right], \quad (3.1)$$

where $F(x) = \ln(x) + \psi(1/2 + 1/x)$ with $\psi(x)$ the digamma function. Here $\tau^{-1}_B = 4DeB/h$ ($D$ is the carrier diffusion constant), $\tau^{-1}_\phi$ is the dephasing rate, $\tau^{-1}_{asy}$ represents the spin relaxation rate due to the SOI terms that break $z \rightarrow -z$ symmetry ($z$ is the direction normal to the graphene plane) and $\tau^{-1}_{so}$ is the total spin-relaxation rate, including the effect of all SOI terms.

Since weak (anti)localization theory is developed having the fully diffusive transport regime in mind \cite{140}, we started with the analysis of the magneto-conductivity measured in the lowest mobility device, i.e. the graphene-on-WSe$_2$ device whose data are shown in Fig. 3-3(b) (carrier mobility $\mu \approx 3,000$ cm$^2$/Vs; elastic scattering time $\tau \approx 0.04$ ps). Fig. 3-5(a) shows that Eq. (1) reproduces the data well with $\tau_{so} \approx 0.2$ ps, with all other parameters satisfying the conditions of validity of the theory: $\tau_{\phi} > \tau_{asy} > \tau_{so} > \tau$. Some more considerations are however needed to understand physically the meaning of the good agreement between Eq. (1) and the experimental data.

\footnote{When fitting the data from different devices, we found that –in agreement with the assumptions made by theory– the condition, $\tau_{\phi} > \tau_{asy} > \tau_{so}$, is always satisfied. However, whereas the data analysis gives an upper bound for $\tau_{so} \approx 0.2 - 0.4$ ps for all devices, the spread in the values extracted for $\tau_{asy}$ is larger. For different devices $\tau_{asy}$ was found to vary between 2 and 10 ps. Such a spread is likely to be caused –at least in part– by uncertainties in extracting the magneto-conductivity ($\Delta\sigma$) from the measured magneto-resistance, due to the limited precisions in determining the device geometry (indeed, the value of $\tau_{asy}$ extracted from the fit is particularly sensitive to the absolute magnitude of the negative magneto-conductivity).}
To this end, we note that, for $\tau_{so}$ values longer than 0.2 ps, theory predicts that weak localization should become visible in the magnetic field range explored in the measurements (see, for instance, the green curve in Fig. 3-5(a), which represents Eq. (1) with $\tau_{so} = 0.78$ ps), contrary to what is observed experimentally (no positive magneto-conductivity is observed up to the largest magnetic field $B = 40$ mT applied in the experiments). $\tau_{so}$ values significantly larger than 0.2 ps are therefore not compatible with these observations and can be excluded. $\tau_{so}$ values shorter than 0.2 ps, however, are compatible with the measurements. This is illustrated by the red curve in Fig. 3-5(a), which is a plot of Eq. (1) with $\tau_{so} = 0.04$ ps (and with all other parameters to be the same). This choice for $\tau_{so}$ also perfectly reproduces the data and leads to a magneto-conductivity that is indistinguishable from the case $\tau_{so} = 0.2$ ps. We therefore conclude that our analysis of WAL can only provide an upper value for $\tau_{so} \approx 0.2$ ps.

These considerations make it clear that measuring WAL up to sufficiently high magnetic field is important, because the absence of a positive magneto-conductivity due to weak localization is what allows a more precise quantitative determination of the upper bound for $\tau_{so}$. In an earlier work on graphene-on-WS$_2$ [119], in which the analysis of WAL was also used to obtain $\tau_{so}$, the magnetic field range had not been extended sufficiently in the measurements. As a result, the estimates of $\tau_{so}$ reported there are approximately one order of magnitude larger than what was found here. In other words, from the analysis reported in a previous work [119], the intensity of SOI—albeit already very strong—appeared to be weaker than what it actually is.

For devices with higher mobility, in which the electron mean free path $l = v_F T$ becomes longer, the analysis of WAL requires more critical thinking. For instance, the use of Eq. (1) is meaningful only up to magnetic field values $B \approx \Phi_0 / l^2$ (where $\Phi_0 = \hbar/e$ is the quantum of flux), because in the diffusive regime described by Eq. (1) the minimum area of a time reversed trajectory giving rise to WAL or weak localization is of the order of $l^2$ [137]. For devices in which $\mu \approx 100,000$ cm$^2$/Vs, the corresponding value of $B$ is only approximately 5 mT. This does not pose problems to observe the characteristic peak in WAL, which for such high $\mu$ values becomes
extremely narrow (see, e.g., Fig. 3-5(c)), but limits the range of magnetic field that can be meaningfully used in the quantitative analysis of the data. Nevertheless, in practice, Eq. (1) does reproduce satisfactorily the measured magneto-conductivity in all cases, as illustrated in Figs. 3-5(b-d). Therefore, as per the low mobility case, the upper bound for $\tau_{so}$ for each of the devices analyzed was determined. This upper bound is such that for larger $\tau_{so}$ values the predictions of Eq. (1) shows the presence of a positive magneto-conductivity at higher magnetic field and are incompatible with the experimental observations, whereas for smaller values of $\tau_{so}$ the magneto-conductivity predicted by Eq. (1) does not change significantly and reproduces the behavior of the experimental data.

The result of this analysis is summarized in Fig. 3-6, in which the upper bound for $\tau_{so}$ extracted for all monolayer devices analyzed is plotted as a function of the elastic scattering time $\tau$. For each device, the result of the analysis performed for different applied gate voltage is also shown. The upper bounds for $\tau_{so}$ determined from WAL cluster between 0.1 and 0.4 ps, and in most cases they are close to 0.2 ps. This is a remarkably systematic behavior, especially considering the large range in carrier mobility investigated (from 3,000 to 110,000 cm$^2$/Vs).

Care is needed in interpreting this result, because for sufficiently high mobility devices (e.g., in all cases in which $\tau > 0.5$ ps in Fig. 3-6) $\tau_{so} < \tau$, which appears to be beyond the regime of validity of Eq. (1) (the assumption that motion is diffusive implies that $\tau$ is the shortest time scale). This situation is not new. It has already been encountered in the analysis of WAL in different two-dimensional systems in which SOI is known to be extremely strong, such as 2D hole gases in GaAs-heterostructures [141, 53]. As discussed in detail in Ref. [53], also in these systems a pronounced signal due to WAL is observed without any positive magneto-conductivity due to weak localization. The quantitative analysis of the magneto-conductivity gives an upper limit for $\tau_{so}$ ($\tau_{so} \approx 3$ ps in that case), such that $\tau > \tau_{so}$ ($\tau \approx 25$ ps for those systems), in complete analogy to what was found in previous high mobility graphene-on-TMD devices. For 2D holes in GaAs-heterostructures, the very short $\tau_{so}$ values extracted from the analysis of WAL were taken as a signature of a strong SOI originating from
the band structure (i.e., not from impurity scattering) [53]. This conclusion was validated through a study of Shubnikov-de Haas (SdH) resistance oscillations, that exhibit a gate-voltage dependent beating. The beating is due to SOI that splits the hole Fermi surface, causing SdH oscillations to occur with two distinct frequencies. By analyzing the frequency splitting as a function of carrier density, the precise nature of the SOI term present in the Hamiltonian could be established.

In view of the similarity in the behavior of WAL in graphene-on-TMDs and in GaAs-based 2D hole gases it is tempting to draw analogous conclusions. Namely, the observed behavior of WAL appears to indicate a band origin of SOI in graphene-on-TMD. Indeed, the SOI-induced splitting in the electronic band structure can generate WAL which is at most weakly dependent on the amount of disorder in the system. This is consistent with the observation of WAL which is strong and robust for a wide range of carrier mobilities in a variety of different samples. To confirm the band origin of interfacially induced SOI, a search for the occurrence of a beating in the low-field Shubnikov-de Haas (SdH) resistance oscillations was performed.

3.6 Spin-Orbit Band Structure Splitting and SdH Oscillations in High-Mobility Graphene-on-TMD Devices

The SdH oscillations of transport coefficients in non-quantizing magnetic fields arise due to cyclotron motion of carrier states at the Fermi level. The periodic dependence on the inverse field $1/B$ provides a convenient way to measure the Fermi surface size. In the presence of spin-orbital splitting, the electronic band structure gives rise to split Fermi surfaces with different spin polarization. In this regime, the SdH oscillations exhibit a characteristic beating pattern that provides an unambiguous diagnostic of the split Fermi surface, allowing to directly measure the spin splitting value.

Measurement of the beating patterns in SdH oscillations relies on resolving a large number of Landau levels at moderate-to-low magnetic fields. Achieving this regime
requires devices of exceptional quality. One constraint arises from carrier mobility which must be high enough to prevent the washing out of Landau levels of high order. Another constraint, which is equally important, is the absence of significant inhomogeneity in carrier density across the device. Indeed, an inhomogeneous density would result in washing out of the SdH oscillations due to different parts of the device contributing to the SdH oscillations with different frequencies. If the spread in frequencies originating from the carrier inhomogeneity is comparable to (or larger than) the SOI-induced frequency splitting, no splitting can be detected experimentally.

The high mobility values that can be achieved in these graphene-on-TMD devices are comfortably in the range needed for detecting a SOI-induced beating in the SdH oscillations. Charge inhomogeneity, on the contrary, poses certain challenges. Previous work indicates that, in the density range of the experiments in this study, the inhomogeneity effects are less prominent in graphene bilayer (BLG) as compared to graphene monolayer [142]. This is so because the density of quasiparticle states is higher in the bilayer, where quasiparticle dispersion is quadratic, and lower in the monolayer, where the dispersion is linear [143]. High-mobility BLG devices were therefore employed for this part of the experiments.

As shown in Fig. 3-7(a), magneto-resistance measurements performed on high-quality BLG-on-WSe$_2$ indeed exhibit beating in the SdH oscillations. The node of the beating pattern, marked by arrows, shifts towards higher magnetic field values when a more negative gate voltage is applied. Accordingly, the SdH oscillations Fourier spectrum (Fig. 3-7(b)) exhibits a pair of peaks with the splitting that increases upon shifting $V_g$ further away from charge neutrality. The full behavior is illustrated in Fig. 3-7(c) which shows the Fourier spectrum as a function of $V_g$ and oscillation frequency $f$. The frequencies corresponding to the maxima of the split peaks in the Fourier spectrum are plotted in Fig. 3-7(d) as a function of $V_g$. The observed splitting in the SdH oscillation frequency is a direct manifestation of the SOI splitting of the Fermi surface of BLG-on-WSe$_2$.
Figure 3-7: Extracting SOI from the SdH resistance oscillations observed in BLG-on-WSe\textsubscript{2} devices. (a) SdH resistance oscillations exhibit a beating pattern whose node position shifts with the applied gate voltage $V_g$ (curves are vertically offset for clarity; the charge neutrality point in this device is at $V_g = -2$ V). (b) Peak splitting in the Fourier spectra of the data shown in (a) is used to determine the SdH frequency splitting. (c) Color-coded Fourier spectrum plotted vs. frequency $f$ and gate voltage $V_g$. (d) The position of the two peaks in the Fourier spectra shown in (c) plotted vs. $V_g$ (black and red circles represent the lower and higher frequency peaks, respectively). The dependence on $V_g$ indicates that the SdH frequency splitting increases as the carrier density increases.
Figure 3-8: (a) Low energy band structure of bilayer graphene on TMD, obtained from the Hamiltonian in Eq. (3), with $U = -6$ meV, $\lambda_R = 13$ meV, $\lambda = 5$ meV. See the inset of panel (b) for a numerically accurate band cross section, with the red line indicating the position of the Fermi energy calculated self-consistently as described in text. (b) The circles represent the measured splitting $\delta f$ of the frequency of the SdH resistance oscillations plotted versus the average frequency peak $\bar{f}$. The values of $\delta f$ and $\bar{f}$ are obtained from the data shown in Fig. 3-7(d) as the Fermi energy is swept through the valence band ($V_g < -2$ V). The colored lines correspond to theoretical predictions calculated using the band structure obtained with the Hamiltonian in Eq. (3) for different values of $\lambda_R$ and $\lambda$. Three possible best-fit plots with $(\lambda_R, \lambda) = (15, 0)$ meV (red solid line), $(\lambda_R, \lambda) = (13, -5)$ meV (green, dashed line), and $(\lambda_R, \lambda) = (13, 5)$ meV (blue, dashed-dotted line) are shown.

These findings indicate that the effect of the interfacially induced SOI is dominated by a modification of the band structure of graphene rather than a spin-dependent disorder potential. For each value of $V_g$, the two peak frequencies are proportional to the areas of the two split Fermi surfaces. The measured dependence of these areas on $V_g$ can be used to reconstruct the band structure as described below.
However, before getting into the details of this discussion, it is useful to note that the general properties of the interfacial SOI, such as its magnitude and momentum dependence, can be inferred directly from the observed SdH beating patterns without invoking detailed modeling. Indeed, since for parabolic bands the Fermi energy is proportional to carrier density, the observed 10-15% splitting in density translates into a 10-15% splitting in the Fermi energy. This predicts the SOI splitting on the order of 10 meV or higher. Second, the monotonic increase of the splitting vs. density indicates that the SOI has a strong momentum dependence. This behavior is consistent with the dominant SOI of a Rashba type. Both conclusions are confirmed below by a detailed modeling.

The general Hamiltonian for BLG-on-TMD takes on the form of a low-energy BLG Hamiltonian with an SOI term added to describe interfacial coupling to TMD. Literature describes BLG in terms of two Dirac Hamiltonians of the two constituent monolayers coupled by the \( \gamma_1 \) term describing inter-layer nearest neighbor hopping [143, 83, 80]. The effect of interfacially induced SOI can be modeled by an effective Hamiltonian which was determined through \textit{ab-initio} calculations in a previous study of monolayer graphene on WS\(_2\) [119]. The expression for the spin-orbit Hamiltonian obtained in Ref. [119] is

\[
H_{SOI} = \frac{\lambda}{2} \tau_z s_z 1_\sigma + \frac{\lambda_R}{2} (\tau_x \sigma_x s_y - \sigma_y s_x),
\]

where Pauli matrices \( s_{x,y,z} \) and \( \sigma_{x,y,z} \) represent electron spin-1/2 and pseudo-spin (sublattice A-B wavefunction components), respectively. The parameters \( \lambda \) and \( \lambda_R \) represent the strengths of the two types of SOI induced by TMD substrate, hereafter referred to as “spin-valley SOI” and Rashba SOI, respectively [119].

Here the Hamiltonian given in Eq. (3.2) was incorporated in the part of the full BLG-on-TMD Hamiltonian describing the graphene layer in direct contact with the WS\(_2\) substrate. The minimal single-valley Hamiltonian describing BLG-on-WSe\(_2\)

\begin{footnote}{Note that in Ref. [119] the Hamiltonian in Eq.(3.2) was discussed as part of a larger Hamiltonian that contained a potential term describing sublattice A-B modulation, \( H' = \Delta \sigma_z 1_\sigma \). Under realistic experimental conditions the value \( \Delta \) was estimated to be very small, which makes the term \( H' \) entirely negligible.}

75
reads:

\[ H = v(\tau_z \sigma_x k_x + \sigma_y k_y)1_s 1_\rho + \frac{1}{2} \gamma_1 1_s (\sigma_x \rho_z + \sigma_y \rho_y) + \frac{1}{2} U 1_s 1_\sigma \rho_z + \left( \frac{1_\rho + \rho_z}{2} \right) H_{SOI}, \]

(3.3)

where \( \rho_{x,y,z} \) are Pauli matrices corresponding to the BLG layer index, the quantities \( \sigma_{x,y,z} \) and \( s_{x,y,z} \) were defined above, the quantities \( 1_\rho, 1_s, 1_\sigma \) denote the corresponding 2x2 identity matrices, and \( \tau_z = \pm 1 \) is the valley index (valley degeneracy persists in the presence of SOI). The Hamiltonian \( H \) includes the inter-layer potential difference \( U \).

In the single-gated configuration of the devices used in this study, the value \( U \) is finite in the presence of an applied gate voltage \( V_g \).

The Hamiltonian given in Eq. (3.3) was used for the quantitative estimate of the SOI parameters \( \lambda \) and \( \lambda_R \). This is done by determining the SdH oscillation frequency \( f \) of each spin-split band from the area of the corresponding Fermi surface and comparing the resulting values with the experimental data. To this end, I employed the approach of Ref. [90] for obtaining \( E_F \) self-consistently. For a given \( \lambda, \lambda_R \) and \( U \), the band structure can be easily computed, see Fig. 3-8(a) as an example. Due to the layer asymmetry, eigenstates of the Hamiltonian Eq. (3.3) are not equally partitioned between the two graphene layers of the BLG. Thus, just as in the case of the spin-degenerate gapped BLG, the total carrier density \( n = n_1 + n_2 \) fixed by a given \( E_F \) is unequally split between the two layers. The unequal layer carrier densities give rise to interlayer potential difference \( U(n) \) that depends on the total carrier density \( n \).

By varying \( E_F \) (or \( n \)), the self-consistent value for which \( U(n) = U \) can be found. The self-consistent solution corresponds to the experimental situation in which the gate voltage \( V_g \) is given by \( C_D(V_g - V_{g0}) = e n \), where \( V_{g0} \) corresponds to the charge neutrality point and \( C_D \) is the capacitance per unit area between the device and gate.

\[ ^4 \text{More specifically, for a given } E_F, \text{ we compute the total excess (i.e. relative to the background density of positive charges) electron density on each layer, including those from the completely filled bands. Due to the layer asymmetry of the system, eigenstates from the completely filled bands are also partitioned unequally between the two layers and thus also contribute (albeit independent of } E_F, \text{) to the inter-layer potential } U_n. \text{ } U_n \text{ is then given by } C_{BLG} U_n = e \delta n_{total,2}, \text{ where } C_{BLG} \text{ is the capacitance per unit area of the bilayer graphene (assuming vacuum dielectric constant } \epsilon_0) \text{ and } \delta n_{total,2} \text{ is the total excess electron density on the layer further away from the substrate.} \]
electrode. As an illustration, the red line in inset of Fig. 3-8(b) indicates the position of the self-consistent value of $E_F$ for $U = -6$ meV, $\lambda_R = 13$ meV, and $\lambda = 5$ meV.

Knowing the self-consistent $E_F$ for a given $U$, the Fermi momenta $k_{F+}$ and $k_{F-}$ of the two spin-split bands can then be determined, or -equivalently- the areas of the corresponding Fermi surfaces and the density of carriers $n_+$ and $n_-$ in these two bands (note that $n = n_+ + n_-\). For instance in the inset of Fig. 3-8(b), $k_{F+}$ and $k_{F-}$ correspond to the intersection of the red line with the two blue lines. This procedure is then repeated so that the self-consistent solutions of $n_+$ and $n_-$ is obtained as $U$ (and hence $E_F$) is swept through the valence band, just as is done in experiments. The values of the frequency peaks are related to the carrier densities through re-scaling by a factor of $\frac{\hbar}{2e}$.

Having determined self-consistently the relation between $n_\pm$ and $U$, we proceed to compare theory with experimental results. This is done by plotting the frequency splitting $\delta f$ as a function of the average frequency $\bar{f}$. Here both quantities can be extracted directly and independently from the measured data in Fig. 3-7(d) (note that $\bar{f} \propto V_g$). In Fig. 3-8(b), the empty circles correspond to the experimental data measured as $E_F$ is swept through the valence band while the colored lines represent three possible theoretical best-fit plots obtained for different combinations of SOI values $(\lambda_R, \lambda) = (15, 0)$ meV (red solid line), $(\lambda_R, \lambda) = (13, -5)$ meV (green, dashed line), and $(\lambda_R, \lambda) = (13, 5)$ (blue, dashed-dotted line). In this way a very satisfactory fit to the data can be obtained. Upon varying $\lambda_R$ and $\lambda$ over a broad interval, the range of values for which a good agreement is found is $\lambda_R \simeq 10-15$ meV and $\lambda \simeq 0-6$ meV (larger values of $\lambda_R$ constrain $\lambda$ to smaller values to fit the data).

Theory can also reproduce the data obtained when $E_F$ is swept through the conduction band, but in that case the range of carrier density for which a beating is observed experimentally is smaller (see Fig. 3-7(d)), and thus only a few data points are present in the $\delta f$-vs-$\bar{f}$ plot (again $\bar{f} \propto V_g$). In the conduction band, the best-fit values of $\lambda_R$ and $\lambda$ are in the range between 5 and 8 meV. Although in this case SOI appears to be slightly weaker than in the valence band, the smaller amount of data makes it more difficult to determine the two parameters accurately.
The above analysis of the beating patterns observed in the SdH measurements therefore indicates that the strong SOI induced in graphene by proximity with TMD is of a band origin. The SOI strength extracted from the comparison between theory and experiments is in the same ballpark, although somewhat larger than the values estimated in Ref. [119] from ab-initio calculations. The similarity of the SOI strength estimated from the data and that found from ab-initio calculations supports the consistency of our analysis. We therefore conclude that the characteristic magnitude of the interfacially-induced SOI in graphene is about 10 meV. This value is more than a hundred times larger than the SOI intrinsically present in pristine graphene [92, 93, 122].

This is also in line with the conclusion drawn from the behavior of the spin-relaxation time $\tau_{so}$ obtained from the analysis of WAL. In this regard we note that a $\sim 10$ meV spin splitting in the band structure, such as the one found here, would give rise to a characteristic "Rabi" spin precession frequency of $\lambda_R/\hbar \approx 10$ THz, which is compatible with values of the spin relaxation time $\tau_{so} \approx 0.2$ ps inferred from WAL. Indeed, it would have been difficult to reconcile this fast spin relaxation time with a much smaller SOI strength — say, a spin splitting of 1 meV or less. Specifically, a 1 meV spin-splitting corresponds to a 1 THz Rabi frequency, which is insufficient to randomize the spin direction over a time of 0.1-0.2 ps. These considerations indicate that two very different phenomena — the quantum correction to the conductivity due to electron interference and the splitting in the frequency of the SdH oscillations — lead to conclusions on the estimated strength of the SOI that are internally consistent, providing an independent confirmation of the validity of our analysis.

3.7 Chapter Summary

The main conclusion that can be drawn from the measurements presented above is that the interfacially induced SOI is dominated by spin-orbital splitting in the graphene band structure. This generalizes to spin-dependent phenomena the results obtained in graphene-on-hBN moire superlattices, where interfacial interactions alter
the graphene band structure by producing secondary Dirac points and creating a gap at the main Dirac point. Our second conclusion is that the interfacially induced SOI is extremely robust. A strong SOI is induced irrespective of the specific TMD material used, of the graphene and TMD lattice alignment angle, of the thickness of the graphene multilayer (up to three layers were tested in this study), or of the position of the Fermi level in the graphene band.

Most tellingly, a pronounced magneto-conductivity due to weak antilocalization was observed in all devices that were measured at low temperatures, where the transport regime varied from fully diffusive up to nearly ballistic. None of the devices exhibited a positive magneto-conductivity due to weak-localization. This indicates that SOI in graphene was –in all regimes investigated– sufficiently strong to cause a full precession of the electron spin even for the shortest trajectories that contribute to electron interference. These observations appear to be only compatible with a band origin of strong SOI, a conclusion that is confirmed by the experimental observation of beating patterns and a splitting in the frequency of the Shubnikov-de-Hass oscillations in high-quality bilayer devices. The evolution in the magnitude of the splitting that is observed upon varying carrier density indicates that the dominant contribution to the interfacially induced SOI is of the Rashba type. A quantitative analysis of the data indicates that the interaction coupling constant for the Rashba term is as large as 10-15 meV, whereas the strength of the other SOI term that couples spin and valley degrees of freedom is about 5-6 meV or smaller. These values correspond to a remarkably strong SOI, especially in comparison to the minute values of the intrinsic SOI in graphene which are only 20-40 μeV [92, 93, 122].

It is remarkable that such a large interfacial SOI can be induced without causing any damage to the electronic properties of graphene. The highest quality devices studied here exhibited carrier mobility reaching up to 160,000 cm²/Vs and carrier density inhomogeneity of only \( \simeq 2 \times 10^{10} \) cm\(^{-2}\), which is comparable to best graphene-on-hBN devices [106, 107, 108, 131, 132]. The possibility to achieve such a high quality both in terms of carrier mobility and density homogeneity will be crucial for probing the predicted topologically insulating states that may be realized in graphene-based
In that regard we also note that, since Rashba turns out to be the dominant SOI coupling, a gap opening between valence and conduction bands in charge neutral graphene-on-TMD (leading to a topologically insulating state) is not expected to occur [119]. However, diverse strategies are available to change the situation. For instance, encapsulating graphene in between two TMD crystals may result in a smaller asymmetry of the device structure, causing a decrease in the intensity of the Rashba term, with other spin-valley SOI contribution becoming stronger. Under these conditions, a topological insulating state may be engineered in graphene with a band gap of several meV [119] and, if so, the ability to achieve very high carrier mobility and small density inhomogeneity demonstrated in this work will be essential for probing the occurrence of edge transport in the presence of an insulating bulk.

Another interesting possibility opened up by the results presented here is achieving gate control of SOI in graphene based system. A simple strategy in this regard is to employ dual-gated BLG devices, using a TMD layer as gate insulator on one side and a hBN layer on the other side. The application of a perpendicular electric field in such a structure will lead to a band gap opening at the charge neutrality point [88, 58, 126]. In this regime, electronic states at the top of the valence band and at the bottom of the conduction band have their wave-functions localized on one of the two BLG layers (depending on the sign of the perpendicular electric field) [91]. Since the interfacially induced SOI is present only in the BLG layer which is in direct contact with the TMD, such a configuration will allow switching SOI on and off by tuning the position of the Fermi level and the polarity of the electric field responsible for the gap opening.

Finally, strong interfacially induced SOI in graphene-on-TMDs opens countless opportunities for investigating novel physical phenomena under controlled conditions. For instance, devices can be realized by employing ferromagnetic electrical contacts that enable injection and detection of spins in graphene [120, 144]. In these systems the dynamics of spin polarized carriers will be controlled, and altered in an interesting way, by the interfacially induced SOI. These systems will also help to gain new insight into the subtle phenomena originating from spin-Hall effect and inverse spin-Hall ef-
fect [145]. These examples illustrate that the ability to engineer the properties of electronic systems in van der Waals heterostructures through a layer-by-layer assembly—demonstrated here for the case of SOI in graphene—opens up a wide range of exciting opportunities for realizing and exploring new physical phenomena.

3.8 Appendix: Classical Magneto-Conductivity Background

As discussed in Section 3.5, for the quantitative analysis of the WAL data, the quantum correction to the magneto-conductivity is obtained by subtracting the classical contribution from the total measured magneto-conductivity. The classical contribution is determined by taking the magneto-conductivity measured at a temperature that is sufficiently high to reduce significantly the phase coherence time $\tau_\phi$. In the case of the devices in this study, this happens at 20-30 K, above which quantum corrections become negligible (as indicated, for instance, by the complete disappearance of the negative magneto-conductivity due to WAL).

Here we briefly discuss the behavior of the classical contribution that was subtracted, which is shown in Fig. 3-9 for three devices with different mobility and mean free path $(l)$. It is seen that the classical magneto-conductivity is small and negative for the low-mobility device (green line in Fig.3-9, $\mu \simeq 3,000$ cm$^2$/Vs) as expected from diffusive Drude transport. For these low-mobility devices, the background is so small that has virtually no effect on the quantitative analysis of the data, i.e., removing it or not leads to virtually identical estimates of $\tau_{so}$. For devices with larger mobility, it was systematically found that the classical magneto-conductivity tends to become positive, growing in magnitude with increasing mean free path. This behavior can be understood by recalling that, while the classical magnetoconductivity is negative in the diffusive regime, it changes sign and becomes positive in a narrow channel in which the mean free path is larger than the channel width. In the crossover regime, occurring when the mean free path increases and approaches the channel width, the
magneto-conductivity is positive but small, growing larger for larger mean free path values. Although the devices studied here are never fully in this regime, in some cases the mean free path is as large as half the channel width (black line in Fig. 3-9).

It might appear that the positive background magneto-conductivity that was subtracted is in fact due to weak localization. The data, however, show that this is not the case. Not only is there an absence of temperature dependence in the background (see Figs. 2a-c), but also the magnitude of the measured positive magneto-conductivity typically far exceeds $e^2/h$ (see for instance, the black curve in Fig. 3-9). Both of these features are incompatible with the weak localization interpretation. Additionally, in high-mobility devices with a long mean free path the weak localization signal should appear only in at low magnetic fields (within a few mT) when the magnetic length
is longer than (or at most comparable to) the mean free path, whereas the observed positive magneto-conductivity continues to increase up to much higher magnetic field values.
Chapter 4

On-demand Spin-Orbit Interaction from Which-Layer Tunability in Bilayer Graphene

Spin-orbit interaction (SOI) that is gate-tunable over a broad range is essential to exploiting novel spin phenomena. Achieving this regime has remained elusive because of the weakness of the underlying relativistic coupling and lack of its tunability in solids. In this chapter, I outline a general strategy that enables exceptionally high tunability of SOI through creating a which-layer spin-orbit field inhomogeneity in graphene multilayers. An external transverse electric field is applied to shift carriers between the layers with strong and weak SOI. Because graphene layers are separated by sub-nm scales, exceptionally high tunability of SOI can be achieved through a minute carrier displacement. A detailed analysis of the experimentally relevant case of bilayer graphene on a semiconducting transition metal dichalcogenide substrate is presented. In this system, a complete tunability of SOI amounting to its ON/OFF switching can be achieved. New opportunities for spin control are exemplified with electrically driven spin resonance and topological phases with different quantized

---

1This chapter is reproduced in part with permission from J. Y. Khoo, L. S. Levitov, On-demand Spin-Orbit Interaction from Which-Layer Tunability in Bilayer Graphene, *NanoLetters* 17, 7003-7008 Copyright (2017) by the American Chemical Society.
intrinsic valley Hall conductivities.

4.1 Tuning Spin-Orbit Interaction in Bilayer Graphene

Spin-orbit interaction (SOI), tunable on demand over a wide range of values can provide access to a wide variety of interesting spin transport phenomena. One popular strategy of achieving tunable SOI relies on directly tuning the SOI using an applied electric field. This approach proved successful in various instances such as tuning Rashba-type SOI in two-dimensional semiconducting systems[146, 147, 148, 149] and Ising-type SOI in transition metal dichalcogenides (TMDs)[76]. However, in all these cases the range of values in which SOI could be tuned has been relatively small because of the relativistic nature of SOI.

I propose graphene multilayers as a vehicle to achieve an on-demand SOI that is free from these limitations. The first step involves engineering an environment with a spatially inhomogeneous spin-orbit field[150, 151], which is e.g. high on one layer and low on the adjacent layer. In such a system, through applying transverse electric field, carriers can be shifted between layers with strong and weak SOI. This renders the SOI strength felt by these carriers strongly dependent on the which-layer charge polarization. Some aspects of this scheme resemble gate-tunable Zeeman coupling demonstrated in Ref. 152. The advantage of such an indirect approach to tuning SOI is that it disassociates the applied electric field from the spin-orbit field. The atomic scale separation between graphene layers then ensures an exceptionally high tunability that is achieved through a minute carrier displacement.

I illustrate this idea in the specific context of bilayer graphene (BLG) on a TMD substrate such as WS$_2$. Implementing a strongly tunable SOI in such graphene-based systems is highly desirable due to the high mobility of carriers in graphene that is preserved by these atomically flat and chemically inert substrates[153]. In this configuration, the spatially inhomogeneous spin-orbit field simplifies to an ON/OFF which-layer field – only the layer adjacent to the TMD acquires from it an interfacially-induced Rashba SOI and Ising SOI. Our proposal builds on previous work, which
Figure 4-1: Low-energy band structure of a biased bilayer graphene (near $K_+$ point) with interlayer bias $U = 10$ meV including induced SOI $H_{SO}^{\text{eff}} = \frac{1}{2} g_1^{(c)} \lambda s_z$ with $\lambda = 2$ meV. The color of the lines indicates the layer occupancy $g_{1,2}^{(c)}$ of their corresponding eigenstates given by Eq. 4.3 in the main text. A similar situation occurs near the $K_-$ point. Different sizes of arrows in the insets depict the difference in strength of the induced SOI experienced by carriers in layer 1 and 2 as a result of their proximity from the substrate.
established that strong interfacial SOI in the meV range can be induced in individual graphene layers[119, 154, 118, 155, 113]. As I will show, the low-energy carriers experience an effective SOI that has an enhanced gate-tunability to the extent of complete gate-tunability, i.e. it can be switched on and off by applying a transverse electric field of moderate strength (~ mV/Å).

Furthermore, the robust high-frequency response of graphene extending up to ~ 100 GHz[156] can enable a range of novel time-dependent spin phenomena. Indeed, because applying a transverse field in BLG directly alters the wavefunctions of its carriers, gate-tunable SOI possesses full quantum coherence. Quantum-coherent tunability enables coherent manipulation of carrier spin degrees of freedom, becoming particularly interesting if the SOI Hamiltonian can be modulated on the carrier transport time scales. As an illustration of this new capability, I discuss the electric-dipole spin resonance (EDSR) that can be driven though an application of a transverse AC electric field[157]. Quantum-coherent tunability also enables direct control of the electron Bloch Hamiltonian and Bloch bands, giving access to gate-tunable Berry phase and band topology of Bloch electrons. I illustrate these new opportunities by considering BLG sandwiched between TMD layers, a system that provides gate-switching between topologically distinct phases with different values of the intrinsic valley Hall conductivity.

4.2 Which-Layer Tunability in Bilayer Graphene

The essential aspects of the which-layer approach can be illustrated by a model of a Bernal-stacked BLG in a transverse electric field, which for the sake of simplicity only accounts for the low-energy subspace of electronic states. Microscopically, the interlayer bias potential $U$ introduces an asymmetry between the A sublattice of layer 1 and B sublattice of layer 2, denoted below as A1 and B2. Crucially, the interlayer bias breaks the layer-occupation symmetry[91]. This behavior is captured by the
two-band (spin-degenerate) Hamiltonian describing the low-energy subspace,

\[ H_{\text{BLG}}^{2 \times 2} = \begin{pmatrix} -\frac{1}{2}U & -\frac{1}{2m} \pi^2 \\ -\frac{1}{2m} \pi^2 & \frac{1}{2}U \end{pmatrix}, \quad \pi = \tau_2 p_x + i p_y. \]  

(4.1)

Here \( p \) is the momentum measured relative to the \( K \) and \( K' \) points of the Brillouin zone, which I will henceforth refer to as \( K_+ \) and \( K_- \) (\( \tau_z = \pm 1 \)) respectively, and \(-\frac{1}{2}U\) and \(\frac{1}{2}U\) are the potentials on layers 1 and 2. The spectrum of \( H_{\text{BLG}}^{2 \times 2} \) is given by

\[ \varepsilon_\zeta(p) = \frac{\zeta}{2} \sqrt{U^2 + \frac{p^4}{m^2}}. \]  

(4.2)

Here \( \zeta = \pm 1 \) refers to the conduction or valence band respectively, and the interlayer bias \( U \) incorporates the capacitance corrections[91]. When \( U \neq 0 \), the wavefunctions of electronic states are asymmetric in the layer occupancy:

\[ g^{(c)}_{1,2} = |\psi^{A_1,B_2}_\zeta(p)|^2 = \frac{1}{2} \mp \frac{U}{4 \varepsilon_\zeta(p)}, \]  

(4.3)

where the minus and plus signs correspond to the layers 1 and 2, respectively. The extent of layer polarization for each of these states is therefore directly controlled by \( U \). The carriers with specific layer polarization can be accessed in an energy-resolved manner through doping[158].

### 4.3 Interfacially-Induced Layer-Dependent SOI

Next, I discuss how a layer-dependent SOI is engineered using a proximal TMD layer, e.g. a TMD multilayer with strong SOI such as WS\(_2\) which serves as the substrate for the BLG. We expect carriers in layer 1 (blue) that is adjacent to the TMD substrate to acquire an interfacially-induced SOI (see Fig.4-1). Carriers in non-adjacent layer 2 (red) are coupled to substrate only indirectly, through interlayer hopping. This phenomenology of interfacially-induced SOI is supported by recent studies[119, 113] on monolayer graphene (MLG) on TMD substrates with strong SOI. A simple model
of MLG experiencing an enhanced SOI due to the TMD substrate can be described by a low-energy Hamiltonian

\[ \delta H_{\text{MLG}} = \frac{\Delta}{2} \sigma_z + \delta H_{\text{Ising}} + \delta H_R, \]  

(4.4)
treated as a perturbation to the MLG Hamiltonian near the \( K_\pm \) points. Here \( \delta H_{\text{Ising}} = \frac{\lambda}{2} \tau_z s_z \) and \( \delta H_R = \frac{\lambda}{2} (\tau_z \sigma_x s_y - \sigma_y s_z) \), where I use \( \sigma_i \) and \( s_i \) to denote the Pauli matrices corresponding to the A and B sublattices, and to spin degrees of freedom, respectively. The term \( \delta H_{\text{Ising}} \) has the form of Ising SOI and originates from the Ising SOI inherently present in the TMD substrate. The term \( \delta H_R \) has the form of Rashba SOI in graphene. The term \( \frac{\lambda}{2} \sigma_z \) originates from sublattice asymmetry; it is comparatively small in practice and can be ignored in most cases.

The consequence of introducing layer-dependent SOI in BLG can be illustrated by considering a simple model in which the interfacial SOI described by \( \delta H_{\text{MLG}} \) only affects the carriers localized in layer 1 and is negligible for carriers localized in layer 2. To see how this modification allows for a switchable SOI, consider the limit of weak Ising SOI, \(|U| \gg |\lambda|\), and with \( \Delta = \lambda_R = 0 \). To first order in perturbation theory, I neglect the influence of \( \delta H_{\text{MLG}} \) on the electronic states and find the spin-split low-energy bands

\[ \delta \varepsilon_{\zeta,s=\uparrow,\downarrow} = \pm \frac{\lambda}{2} \tau_z g_1^{(\zeta)}, \]  

(4.5)
where the energy shifts of sign plus and minus correspond to the \( s = \uparrow \) and \( s = \downarrow \) states respectively. The spin splitting in Eq. 4.5 is of opposite sign for different valleys, \( \tau_z = \pm 1 \), as required by time reversal symmetry.

Bands with different spin splitting can be accessed in a dual-gated system in which there is independent control over interlayer bias and doping: the induced SOI is turned on by hole doping and turned off by electron doping. Indeed, at small dopings, since the Fermi momentum is small, there is a correlation between which band a carrier is from and which layer of the BLG the carrier predominantly occupies. For positive interlayer bias \( U > 0 \), carriers from the conduction band are fully localized on layer 2, \( g_1^{(+1)} \approx 0 \), while carriers from the valence band are fully localized on layer 1,
It follows from Eq. 4.5 that in this case the Ising SOI is only present for holes \( (\delta \varepsilon_{-1,s=\uparrow, \downarrow} = \pm \lambda \tau_z) \) and is absent for electrons \( (\delta \varepsilon_{+1,s} \approx 0) \), as illustrated in Fig.4-1; the situation is reversed when \( U < 0 \) so that the Ising SOI is only present for electrons and is absent for holes (see Fig.4-2(a) and (c)). The contrast between the ON and OFF states fades away quickly as doping increases, since at large momenta the electron wavefunctions are split nearly equally between both layers (see Fig.4-2 right panel).

One interesting consequence of which-layer tunability is that the spin splitting (Eq. 4.5) acquires a dependence on the interlayer bias \( U \). Crucially, the states in Eq. 4.5, while having opposite spin projections, have identical orbital structure. It is therefore possible to view the spin splitting in Eq. 4.5 as being due to an effective magnetic field applied transverse to the BLG plane. Because of the dependence of the layer occupancy on the interlayer bias \( U \) (Eq. 4.3), this effective \( B \) field is gate-tunable and therefore defines a new form of spin-electric coupling.

### 4.4 Electric-Dipole Spin Resonance

As an illustration of the new capabilities endowed on the system by such spin-electric coupling, I discuss spin resonance of an electric-dipole spin resonance (EDSR) type driven by a time-dependent gate voltage \( U(t) \). I consider an external static magnetic field of strength such that the Zeeman energy exceeds the interfacially induced spin splitting, Eq. 4.3, for the sake of simplicity taking the field to be applied parallel to the BLG plane. The carrier spin dynamics is then governed by an effective Hamiltonian

\[
H_{EDSR} = \frac{1}{2} \varepsilon_z s_x + \frac{1}{2} \lambda(t) \tau_z s_z, \tag{4.6}
\]

where the time-dependence \( \lambda(t) = \lambda g_{1}^{(c)}(t) \) originates from gate-tunable Ising SOI. Here \( \varepsilon_z = g \mu_B B \) is the Zeeman energy, \( \mu_B \) is Bohr's magneton and, without loss of generality I consider the static magnetic field applied along the \( x \) direction, \( B = B \hat{x} \).

To achieve EDSR, the interlayer bias \( U(t) \) should not at any point in time close
the gap between the valence and conduction bands so that the carrier orbital wavefunctions remain unchanged. Without loss of generality, I consider the case for which \( \lambda > 0 \) and \( U(t) = U_0 + U_1 \cos \omega t \) with \( U_0 > \lambda \) and \( U_1 \ll U_0 \). The resulting time-dependent Ising SOI experienced by the conduction band carriers with momentum \( p \) varies with time as \( \lambda(t) \simeq \frac{\Delta p^4}{4m^2|\gamma|^2(t)} \). This time-dependent Ising SOI will thus act as an oscillating field which induces transitions between the Zeeman states \( s = | \rightarrow \rangle \) and \( s = | \leftarrow \rangle \). Consequently, EDSR is achieved by matching the frequency of the time-dependent Ising SOI to the Zeeman energy, \( \hbar \omega = \epsilon_Z \). Note that while the Ising SOI has opposite signs at the \( K_\pm \) valleys, both SOI couplings cause the spin projection on the \( x \)-axis to evolve with the same time dependence. In the absence of intervalley coupling, the effects of EDSR originating from both valleys add up constructively, resulting in doubling of the spin polarization signal.

4.5 Low-Energy Spectrum of Bilayer Graphene with Layer-Specific SOI

While the simple analysis of interfacially induced SOI presented above qualitatively captures the essential physics, it is instructive to develop a more precise and complete description of the system near the \( K_\pm \) points. That can be done by directly adding \( \delta H_{\text{MLG}}, \text{Eq. 4.4,} \) to the layer-1 subspace of the BLG tight-binding Hamiltonian,

\[
H_{\text{eff}} = H_{\text{BLG}} \otimes 1^{(s)} + P_1 \delta H_{\text{MLG}} P_1.
\]  

(4.7)

Here \( P_{i=1,2} \) is the operator that projects onto the layer-\( i \) subspace and \( 1^{(s)} \) is the \( 2 \times 2 \) identity matrix of the spin degrees of freedom. The Hamiltonian \( H_{\text{BLG}} \) describes Bernal-stacked BLG, in which two MLG layers are stacked such that the B sublattice in one layer (B1) is vertically aligned with the A sublattice on the other (A2). As is well known, the strongest interlayer coupling \( \gamma_1 \) in this stacking configuration occurs between the B1 and A2 sites. As a result the low-energy states near the \( K_\pm \) points are strongly localized over the A1 and B2 sublattices, as described by the low-energy
Hamiltonian considered above, Eq. 4.1. The Hamiltonian $H_{\text{eff}}$ can be numerically solved to obtain the band structure. The four low-energy bands obtained in this way are shown in Fig.4-2 for three different values of interlayer bias $U$ corresponding to three different phases discussed below.

To gain insight in the different regimes accessible though varying $U$, I derive the low-energy Hamiltonian in the A1/B2 subspace perturbatively in $1/\gamma_1$ (see Supporting Information). Since $[H_{\text{eff}}, s_z] = 0$ at $p = 0$, the quantity $s_z$ is a good quantum number and can be used to label states and associated bands. To order $1/\gamma_1$, the energy levels at $p = 0$ are

$$E_{\zeta=+1,s}|_{p=0} \approx \frac{U}{2} \quad \text{(doubly degenerate, } s = |\uparrow\rangle, |\downarrow\rangle),$$

$$E_{\zeta=-1,s}|_{p=0} \approx -\frac{U}{2} + \frac{\Delta}{2} + \delta\lambda^* \pm \left(\frac{\lambda}{2} - \delta\lambda^*\right)\tau_z, \quad \delta\lambda^* = \frac{\lambda^2}{\gamma_1^2} (U - \frac{\Delta}{2} + \frac{\lambda}{2}). \tag{4.8}$$

Here, the sign in front of the $U/2$ term matches the value of $\zeta = \pm 1$ introduced in Eq. 4.2. In Eq. 4.8 the plus and minus sign corresponds to the $s = |\uparrow\rangle$ and $s = |\downarrow\rangle$ states respectively. This result extends Eq. 4.5 by including the effects of sublattice asymmetry $\Delta$ and the leading correction at order $1/\gamma_1$ given by $\delta\lambda^*$. We see that the effect of $\Delta$ is to uniformly shift both $E_{-1,s}$ bands and renormalize the bias $U$. The quantity $\delta\lambda^*$ produces a similar effect, and can also generate a spin splitting between the $E_{-1,s}$ bands. However, so long as the values $\lambda$ and $\lambda_R$ are comparable, the quantity $\delta\lambda^*$, which is suppressed by a large factor $1/\gamma_1$ compared to $\lambda$, only matters as a constant energy shift to the $E_{-1,s}$ bands but not a source of spin splitting.

In this case, upon tuning interlayer bias $U$, the $E_{+1,s}$ bands (red) shift across the $E_{-1,s}$ bands (blue) so that the system undergoes transitions from an insulator to a semi-metal and then again to an insulator state. The corresponding phase diagram is shown in Fig.4-3. At large $U$ (Fig.4-2(a, c)), the system is insulating at charge neutrality and allows for gate-switching of Ising SOI as discussed above. The sign of $U$ determines which charge carriers, electrons or holes, experience the effective Ising SOI. At not-too-large $U$ values such that $\Delta - \lambda \lesssim 2U \lesssim \Delta + \lambda$, the system is semi-metallic at charge neutrality. As shown in Fig.4-2(b), in this case the $E_{+1,s}$
Figure 4-2: Four lowest energy bands (left panel) and layer polarization of corresponding eigenstates (right panel) obtained from Eq. 4.7. Upon increasing $U$, the system first transitions from an insulating phase (a) to a semi-metallic phase (b) and then to a different insulating phase (c) that has the order of the bands inverted compared to (a); labels correspond to the phase labels in Fig. 4-3. Values used are (a) $U = -2 \text{ meV}$, (b) $U = 2 \text{ meV}$, (c) $U = 5 \text{ meV}$ with other parameters fixed as $\Delta = 3 \text{ meV}$, $\lambda = 4 \text{ meV}$, and $\lambda_R = 2 \text{ meV}$. 
bands lie between the \( E_{-1,s} \) bands so that the Ising SOI gap is partitioned between the electrons and holes: \( \lambda \simeq \Delta E_e + \Delta E_h \). This partitioning can be tuned from 0\% to 100\% by varying \( U \), and thus in this regime both the electron and hole spin splittings are gate-tunable, albeit in a correlated fashion.

As \( \lambda \) decreases to zero, its effects at \( p = 0 \) eventually become subleading compared to that of \( \delta \lambda^* \) when \( \lambda_R \gg \gamma_1 \sqrt{\frac{\lambda}{U}} \). At \( \lambda = 0 \), we find from Eq. 4.8 that \( E_{-1,s}|_{p=0} \simeq -\frac{U}{2} + \frac{A}{2} + \delta \lambda^*(1 \mp \tau_z) \). The term \( \delta \lambda^*(1 \mp \tau_z) \) introduces a spin splitting which varies linearly with \( U \). In practice, which contribution dominates depends on the actual values of \( \lambda \) and \( \lambda_R \); so far, experiments in BLG-on-WS\(_2\) indicate that \( \delta \lambda^* \) is indeed dominant with \( \lambda_R \sim 10 \text{ meV} \)[154].

The \( \frac{1}{\gamma_1} \) suppression in \( \delta \lambda^* \) is a consequence of the specific form of \( \delta H_R \propto \tau_z \sigma_x s_y - \sigma_y s_x \). It couples the A1/B2 polarized low-energy states to the high-energy states which are strongly A2-B1 mixed by the interlayer coupling \( \gamma_1 \). At \( p = 0 \), the wavefunctions of the low-energy states remain layer-polarized such that the SOI is completely tunable. However, because \( \delta H_R \) introduces a substantial interlayer mixing in the low-energy subspace that increases with \( p \), the which-layer tunability of the SOI becomes increasingly suppressed away from the valleys as is evident in Fig.4-1.

### 4.6 Gate-Tunable Topological Phase Transition

An even more interesting behavior is found when BLG is encapsulated between two TMD layers. In this case, carriers in each of the two graphene layers experience an interfacially-induced SOI from the TMD layers above and below, respectively. The low-energy effective Hamiltonian near the \( K_\pm \) points now reads

\[
H_{\text{eff}} = H_{\text{BLG}} \otimes 1^{(s)} + \sum_{i=1,2} P_i \delta H_{\text{MLG},i} P_i,
\]

in which a layer index \( i \) is introduced to allow for distinct phenomenological parameters for the different layers: \( \Delta_i, \lambda_i, \lambda_{R,i} \).

A consequence of adding \( \delta H_{\text{MLG},2} \) is to open up a gap at \( p = 0 \) between the
Figure 4-3: Phase diagram of the BLG-on-TMD system described by Eq. 4.7 at charge neutrality when $\lambda \gtrsim \lambda_R$. Phases (a)-(c) have band structures with corresponding labels in Fig.4-2.

$E_{\pm 1,s}$ bands. It follows that for arbitrary values of $\Delta_i, \lambda_i, \lambda_{R,i}$, the system at charge neutrality has up to five different insulating phases with phase transitions occurring at values of $U$ for which the band gap at $p = 0$ closes. While the overall Chern number for any of these bands is guaranteed to vanish because of time reversal symmetry, the valley Chern number is unconstrained and can take non-zero values. In fact, the valley Chern numbers for some of the bands changes across a phase transition, such that these insulating phases are topologically distinct. This suggests that the intrinsic valley Hall conductivity at charge neutrality, $\sigma_{xy}^{\text{VH}}(0)$, is a suitable quantity to distinguish these phases.

To compute $\sigma_{xy}^{\text{VH}}(0)$, I first obtain the Berry curvature of each energy band near either valley, $\Omega_{\pm}^{(n)}(p)$. Here I introduced a generalized index $n$ that labels bands, including both the four low-energy bands (previously labeled by $\{\zeta, s\}$) as well as the other four high-energy bands that I have excluded from our discussion thus far. Because the Berry curvature is strongly peaked at the valleys, the corresponding valley Chern number $N_{\pm}^{(n)}$ can be obtained by numerically integrating $\Omega_{\pm}^{(n)}(p)$ over
an individual valley[159]. The intrinsic valley Hall conductivity at charge neutrality is then obtained from adding up the contributions from all the filled bands, \( \sigma_{xy}^{VH}(0) = \sum_n \left( N_+^{(n)} - N_-^{(n)} \right) \frac{e^2}{h} = 2 \sum N_+^{(n)} \frac{e^2}{h} \). Here I used the relation \( N_+^{(n)} = -N_-^{(n)} \), valid because of time reversal symmetry.

A detailed characterization of the various phases for arbitrary values of \( \Delta_i, \lambda_i, \lambda_{R,i} \) lies outside the scope of this work. Here I highlight a generic aspect which can be illustrated by considering the simplest case of \( \Delta_1 = \Delta_2 = 0 \) (this choice of values is consistent with ab initio studies[119]), \( |\lambda_i| = \lambda \), and \( |\lambda_{R,i}| = \lambda_R \). In this case, the system hosts two topologically distinct phases at charge neutrality – the ordinary valley Hall phase, in which \( \sigma_{xy}^{VH}(0) = -4\text{sgn}(U) \frac{e^2}{h} \), and the anomalous valley Hall phase, in which \( \sigma_{xy}^{VH}(0) \neq -4\text{sgn}(U) \frac{e^2}{h} \). I will denote these phases as VH0 and VH1 respectively.

A simple way to understand the VH0-VH1 phase transition is as follows. The well-studied case of dual-gated BLG in the absence of SOI, which has \( \sigma_{xy}^{VH}(0) = -4\text{sgn}(U) \frac{e^2}{h} \) [160, 88, 161, 162, 163, 164], is in fact a specific example of the VH0 phase for which \( \lambda = \lambda_R = 0 \). The system remains in the same topological phase VH0 in the presence of relatively weak SOI, i.e. when \( \lambda \lesssim |U| \), since the SOI-induced splitting of the low-energy bands is insufficient to cause the band gap to close at either valley. In agreement with the above picture, independent of the relative signs of \( \lambda_i \) and \( \lambda_{R,i} \), it can be shown that \( \sigma_{xy}^{VH0}(0) = -4\text{sgn}(U) \frac{e^2}{h} \).

Upon tuning down the interlayer-bias \( U \) such that \( |U| \lesssim \lambda \), the SOI starts to dominate and band inversion occurs between the low-energy bands at both valleys. The system undergoes a topological phase transition into the VH1 phase through the closing and re-opening of the band gap analogous to phase transitions in Chern and topological insulators. We therefore expect the valley Chern numbers of the low-energy bands to change across the VH0-VH1 phase transition such that \( \sigma_{xy}^{VH0}(0) \neq \sigma_{xy}^{VH1}(0) = 2M \frac{e^2}{h} \), where possible values of \( M \) equal 0, ±1, ±3, ...; the results from our preliminary studies are consistent with this expectation. Unlike \( \sigma_{xy}^{VH0}(0) \), the value of \( \sigma_{xy}^{VH1}(0) \) depends on the relative signs of \( \lambda_i \) and \( \lambda_{R,i} \). This dependence is, however, not yet well understood. This interplay between the induced SOI and the interlayer
bias resembles the behavior for the quantum spin Hall effect in graphene[72] and the electrically tunable topological insulator[165].

In fact, this sharp change in $\sigma_{xy}^{\text{VH}}(0)$ is independent of the specifics of the induced SOI and occurs in the generic case when $|\lambda_1| \neq |\lambda_2|$ and $|\lambda_{R,1}| \neq |\lambda_{R,2}|$ as well despite there potentially being a more elaborate phase characterization scheme. This change in $\sigma_{xy}^{\text{VH}}(0)$ no longer happens at $|U| \simeq \lambda$ but at a different critical value $U_c \lesssim \frac{1}{2}(|\lambda_1| + |\lambda_2|)$. As the experimental study of valley-based transport is still in its infancy, these predictions therefore serve as robust experimental signatures that could be used to advance our understanding of the valley Hall effect. We find ourselves with a system whose topological nature is not completely determined by the material itself, but can instead be gate-controlled, in situ.

### 4.7 Chapter Summary

In summary, graphene-based heterostructures with on-demand SOI grant access to tunable topological properties. In particular, gate-controlled intrinsic valley Hall conductivity can be achieved in these systems through combining interlayer coupling, gating and various types of interfacially-induced SOI. Further, the robust broad-band response of graphene[156] turns graphene-on-TMD heterostructures into a unique platform to realize and explore novel time-dependent spin phenomena such as the electrically driven spin resonance. It can also help to extend the optoelectronics and valleytronics phenomena of current interest[77, 78, 79, 166, 167] into the time domain.

### 4.8 Appendix: Derivation of Low-Energy Effective Hamiltonian and Band Structure

Here I present the details of our derivation of the low-energy effective Hamiltonian of the BLG-on-TMD system. Let us first consider the tight-binding Hamiltonian of the
(spin-degenerate) bernal-stacked BLG near the $K_\pm$ points. This is given by [80]

$$H_{BLG} = \begin{pmatrix}
\epsilon_{A1} & v_\pi^\dagger & -v_4^\pi & v_3^\pi \\
v_\pi & \epsilon_{B1} & \gamma_1 & -v_4^\pi \\
-v_4^\pi & \gamma_1 & \epsilon_{A2} & v_\pi \\
v_3^\pi & -v_4^\pi & v_\pi & \epsilon_{B2}
\end{pmatrix}, \quad (4.10)$$

where $\gamma_1$ is the interlayer (A2-B1) hopping, $v$ is the MLG band velocity, $v_3$ is the velocity associated with trigonal warping and $v_4$ the velocity associated with skew interlayer coupling. $\epsilon_{\alpha i}$ is the on-site energy for sublattice $\alpha = A, B$ and on layer $i = 1, 2$. Note that $H_{BLG}$ is expressed in the $(A1, B1, A2, B2)$ basis. For simplicity, I have neglected the terms associated with $v_3$ and $v_4$ and consider the case with no intrinsic sublattice asymmetry so that $\epsilon_{\alpha i} = \epsilon_i$ with $U = \epsilon_2 - \epsilon_1$. I use this simplified form of $H_{BLG}$ in $H_{eff}$ defined by Eq. 4.7.

Since we are only interested in the 4 low-energy bands which have energy scales much smaller than $\gamma_1$, i.e. $|\gamma_1| \gg |E|, |\lambda|, |\lambda_R|, |U|, |\Delta|$, we can project $H_{eff}$ to the low-energy subspace and obtain a low-energy effective Hamiltonian for our system. To proceed, we first rewrite it as $H_{eff} = H_{\gamma_1} + H'$, with

$$H_{\gamma_1} = \begin{pmatrix} 0 & 0 & 0 & 0 \\
0 & 0 & \gamma_1 & 0 \\
0 & \gamma_1 & 0 & 0 \\
0 & 0 & 0 & 0 \end{pmatrix} \otimes 1_2^{(s)}, \quad (4.11)$$

and re-express it in the basis that diagonalizes $H_{\gamma_1}$:

$$\tilde{H}_{eff} = \tilde{H}_{\gamma_1} + \tilde{H}', \quad \tilde{H}_{\gamma_1} = \begin{pmatrix} 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & -\gamma_1 & 0 \\
0 & 0 & \gamma_1 \end{pmatrix} \otimes 1_2^{(s)}. \quad (4.12)$$

In this basis, we can label each eigenstate by a quantum number $m = 0, \pm 1$ such
that \( \tilde{H}_{\gamma_1} |m\rangle = m \gamma_1 |m\rangle \). The low-energy subspace that we are interested in is spanned by the \( m = 0 \) eigenstates. We can then decompose \( \tilde{H}' \) as

\[
\tilde{H}' = T_0 + T_1 + T_{-1} + T_2 + T_{-2},
\]

where the operators (Unitary but not necessarily Hermitian) \( T_i = T_i^\dagger \) can be understood as ladder operators \(- T_i |m\rangle \propto |m + i\rangle \) – and are given by

\[
T_0 \equiv \frac{1}{2} \begin{pmatrix}
-U_2^{(s)} + 2w_+ & 0 & 0 & 0 \\
0 & U_2^{(s)} & 0 & 0 \\
0 & 0 & w_- & 0 \\
0 & 0 & 0 & w_-
\end{pmatrix}, \quad
T_2 \equiv \frac{1}{2} \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & -U_2^{(s)} + w_- & 0 \\
0 & 0 & 0 & 0
\end{pmatrix},
\]

\[
T_1 \equiv \frac{1}{\sqrt{2}} \begin{pmatrix}
0 & 0 & v\pi T_2^{(s)} + a_R & 0 \\
0 & 0 & -v\pi T_2^{(s)} & 0 \\
v\pi T_2^{(s)} + a_R & 0 & 0 & 0 \\
v\pi T_2^{(s)} & -v\pi T_2^{(s)} & 0 & 0
\end{pmatrix},
\]

where \( w_\pm = \frac{1}{2}(\pm \Delta T_2^{(s)} + \lambda \tau_z s_z) \), \( a_R = \frac{1}{2} \lambda_R (\tau_z s_y - i s_x) \).

With this framework in place, let us proceed with the Brillouin-Wigner method outlined in Section IV of Ref. 82 to obtain the effective Hamiltonian to arbitrary orders in \( \frac{A}{\hbar} \), where \( A \) refers to any energy scale set by the other parameters (including \( \nu p \)). In this expansion scheme, at second order, I recover the usual spin degenerate effective 2-band BLG Hamiltonian \( H_{BLG}^{2\times 2} \). The correction due to SOI at the same order \( (H_{SO,1}^{4\times 4}) \) is off-diagonal, the energy corrections of which only enter at the next order. To consistently account for the energy corrections to the \( H_{BLG}^{2\times 2} \) spectrum, I therefore go to third order in the expansion scheme and obtain
Here, I have explicitly separated the resulting effective Hamiltonian to the usual spin degenerate effective 2-band BLG Hamiltonian $H_{\text{BLG}}^{2\times 2}$ and the corrections due to substrate induced SOI $H_{\text{SO},j=0,1,2}^{4\times 4}$. These corrections are further organized into powers of $\gamma_1^{-1}$, so that the $(j-1)^{\text{th}}$ order correction $H_{\text{SO},j-1}^{4\times 4} \propto \gamma_1^{-j+1}$.

Solving for the eigenvalues of $H_{\text{eff}}^{4\times 4}$ perturbatively, I obtain the low-energy band structure to order $O\left(\frac{\Delta^4}{\gamma_1^3}\right)$ for $\tau_z = 1$:

$$
E_{+1,\downarrow} = \frac{U}{2} + \frac{v^2p^2}{\gamma_1^2} \left( -U - \frac{\Delta}{2} + \frac{\lambda}{2} + \frac{2\lambda_R^2}{2U - \Delta + \lambda} + \frac{2v^2p^2}{2U - \Delta - \lambda} \right) + O\left(\frac{\Delta^4}{\gamma_1^3}\right),
$$

$$
E_{+1,\uparrow} = \frac{U}{2} + \frac{v^2p^2}{\gamma_1^2} \left( -U - \frac{\Delta}{2} - \frac{\lambda}{2} + \frac{2\lambda_R^2}{2U - \Delta - \lambda} + \frac{2v^2p^2}{2U - \Delta + \lambda} \right) + O\left(\frac{\Delta^4}{\gamma_1^3}\right),
$$

$$
E_{-1,\downarrow} = -\frac{U}{2} + \frac{\Delta + \lambda}{2} + \frac{v^2p^2}{\gamma_1^2} \left( U - \frac{\Delta}{2} - \frac{\lambda}{2} - \frac{2\lambda_R^2}{2U - \Delta - \lambda} + 2v^2p^2 \right) + O\left(\frac{\Delta^4}{\gamma_1^3}\right),
$$

$$
E_{-1,\uparrow} = -\frac{U}{2} + \frac{\Delta - \lambda}{2} + \frac{1}{\gamma_1^2} \left\{ \lambda_R \left( U - \frac{\Delta}{2} + \frac{\lambda}{2} \right) \right\}
+ \frac{v^2p^2}{2} \left( U - \frac{\Delta}{2} + \frac{\lambda}{2} - \frac{2\lambda_R^2}{2U - \Delta + \lambda} \right) - v^4p^4 \left( \frac{2}{2U - \Delta + \lambda} \right) + O\left(\frac{\Delta^4}{\gamma_1^3}\right).
$$
Chapter 5

Tunable Quantum Hall Edge
Conduction in Bilayer Graphene through Spin-Orbit Interaction

In this chapter\(^1\), I solve for the single-particle Landau levels of bilayer graphene with layer-specific spin-orbit interaction (SOI) to highlight its deviation from the established Landau level spectrum of bilayer graphene. Bilayer graphene, in the presence of a one-sided spin-orbit interaction (SOI) induced by a suitably chosen substrate, is predicted to exhibit unconventional Quantum Hall states. The new states arise because the one-sided SOI creates a strong splitting of the eight zeroth Landau levels, which are strongly layer-polarized, residing fully or partially on one of the two graphene layers. In particular, an Ising SOI in the meV scale is sufficient to invert the Landau level order between the \(n = 0\) and \(n = 1\) orbital levels under moderately weak magnetic fields \(B \lesssim 10\) T. Furthermore, when the Ising field opposes the \(B\) field, the order of the spin-polarized levels can also be inverted. We show that, under these conditions, three different compensated electron-hole phases can occur at \(\nu = 0\) filling, in which equal concentrations of electrons and holes are present. The

\(^1\)This chapter is reproduced with permission from J. Y. Khoo, L. S. Levitov, Tunable Quantum Hall Edge Conduction in Graphene Bilayer through Spin-Orbit Interaction, submitted for publication in Physical Review B.
three phases have distinct edge conductivity values. One of the phases is especially interesting, since its edge conduction can be turned on and off by switching the sign of the interlayer bias.

5.1 Absence of Spin-Orbit Interaction (SOI) in Graphene Landau Levels (LLs)

The effect of spin-orbit interaction (SOI) on the Landau levels (LLs) in graphene has been largely unexplored experimentally. The main obstacle has been the extreme weakness of the intrinsic SOI, corresponding to spin splittings as small as 1 to 100 $\mu$eV in graphene monolayer [92, 5, 97] and bilayer [100, 101, 102]. Consequently, the SOI splitting is smaller than the Zeeman splitting $E_Z \sim 0.1B(T)$ meV even when magnetic fields $B$ are relatively weak. However, the situation has changed with the advent of graphene-based heterostructures. An SOI of 1 to 10 meV has been interfacially-induced in graphene by transition-metal dichalcogenide substrates with strong SOI such as MoS$_2$, MoSe$_2$, WS$_2$ and WSe$_2$ [154, 119, 118, 168, 169, 170, 171] while an SOI of $\lesssim$ 100 meV was achieved at the graphene-Ni interface by introducing Au intercalation [172]. These developments have therefore opened the door to realizing SOI-based phenomena in graphene.

It is particularly interesting to study how the LLs of (Bernal-stacked) bilayer graphene are modified by interfacially-induced SOI. Apart from having spin and valley degrees of freedom, the low-energy carriers in bilayer graphene are sensitive to the potential difference between the two graphene layers. These properties collectively give rise to a gate-tunable single-particle LL spectrum [173, 174]. By including the effects of electron-electron interactions, we obtain a very rich phase diagram which hosts gate-tunable phase transitions [175, 176, 177] and can be directly probed by experiments [63, 56, 64]. This gate-tunability is therefore expected to provide a means to probe the effects of interfacially-induced SOI on the LLs.
5.2 Zeroth LLs of Bilayer Graphene with Layer-Specific SOI

In our discussion, special attention will be given to the zeroth LLs of bilayer graphene, the set of eight-fold nearly degenerate lowest energy bands. The states of the zeroth LLs belonging to different valleys are strongly localized on different layers. This has two consequences. First, the valley degeneracy can be lifted by introducing layer-asymmetry to the system [174]. This can be achieved by applying an interlayer bias or by constructing an inversion-asymmetric heterostructure. Second, layer asymmetric effects will be most noticeable in the zeroth LLs. Motivated by these observations, and in departure from previous treatments which considered SOI of equal strength for both layers [178, 179], we shall focus on the case of bilayer graphene with a layer-specific SOI. Recent experimental progress has led to a better understanding of the bilayer graphene zeroth LLs at various integer [59, 55, 61, 64, 56] and fractional fillings [62, 63, 64], as well as to the recently demonstrated artificial SOI enhancement in graphene [154, 119, 118, 168, 169, 170, 171, 172]. We thus find ourselves in the opportune moment to investigate the novel valley-asymmetric effects on the zeroth LLs due to substrate-induced SOI.

To this end, here we analyze the single-particle LL spectrum of bilayer graphene with a layer-specific SOI of both the Ising and Rashba types. We highlight several interesting features that arise already at the non-interacting level. Some of these features are expected to remain robust in the presence of interactions. In particular, an Ising SOI $\lambda$ at the meV scale is strong enough to significantly change the zeroth LL spectrum. In contrast, Rashba SOI is of an off-diagonal character, and thus its effect is small even at values as large as $\lambda_R \sim 15\text{meV}$. As a result, the energy ordering of the zeroth LL states is essentially determined by the competition between the layer-asymmetric Ising splitting and the Zeeman as well as orbital splittings. In particular, the orbital and spin order inversions occur at relatively weak and moderate $B$ field values, respectively, as we discuss in detail below.

We predict three compensated electron-hole phases in this system, one of which is
Figure 5-1: Edge state configurations for a) para-phase, b) ortho/para-phase and c) ortho-phase. Edge conduction is reduced due to backscattering between the counter-propagating modes with equal spin polarization. Backscattering between the modes with opposite spins cannot occur because of the orthogonality of the spin wavefunctions, leading to quantized edge conduction. In the phase b), spin polarization of one of the edge modes can be reversed by a transverse electric field, giving rise to a switchable edge conduction.

A conventional phase, whereas the other two phases arise due to the layer-asymmetric nature of the Ising SOI. These three phases occur at the $\nu = 0$ filling and when the interlayer bias $u$ is moderately large so that one layer becomes electron-doped while the other becomes hole-doped by the same amount. In the absence of SOI, there is only one compensated electron-hole phase at $\nu = 0$, which is expected to host helical edge modes with opposite chiralities and spin polarizations in each layer [56]. Spin wavefunctions of these edge modes are orthogonal, forbidding interlayer tunneling processes, and thereby protecting the edge states from backscattering. The expected edge conductance in this phase is therefore $2e^2/h$ [180].

Introducing a layer-specific Ising SOI can invert the energy ordering of the spin-polarized zeroth LLs set by the Zeeman splitting. This inversion occurs only within the zeroth LLs of the corresponding valley, and therefore requires that the Ising splitting at that valley be opposite in sign to the Zeeman splitting, and dominate over it. In this case, the corresponding compensated electron-hole phase will have helical edge modes with the same spin polarization. Backscattering processes are now allowed, so that the edge conductivity of this new compensated electron-hole phase is expected to be suppressed compared to the $2e^2/h$ value.

Increasing $B$ eventually restores the energy ordering so that we recover the com-
pensated electron-hole phase equivalent to that in the system without SOI. However, the magnitude of Ising splitting is somewhat larger for the more strongly layer-polarized \( n = 0 \) LLs as compared to the \( n = 1 \) LLs. As a result, in the presence of moderately large \( B \) fields only the ordering between the \( n = 1 \) spin-polarized levels is restored while that of the \( n = 0 \) remains inverted. Therefore, the transition between the two phases, which are dominated by the spin-orbital and Zeeman interactions, occurs via a third intermediate phase. This intermediate phase has edge modes with the same spin polarization for positive interlayer bias, but with opposite spin polarization for negative interlayer bias.

There are therefore a total of three different compensated electron-hole phases distinguished by whether their edge modes have the same or opposite spin polarizations for the cases of positive and negative interlayer bias, as illustrated in Fig. 5-1. The intermediate phase is particularly interesting because it hosts an edge conductivity that can be turned on or off by switching the sign of the interlayer bias.

### 5.3 Low-Energy Effective Hamiltonian

The low-energy states of bilayer graphene near the Dirac points (the \( K_+ \) and \( K_- \) valleys) can be modeled by an effective Hamiltonian, expressed in the \(( A1, B1, A2, B2)\) basis as \([80, 56? ]\)

\[
H_0^{B=0} = \begin{pmatrix}
\frac{\pi}{2} & v_0\pi \dagger & -v_4\pi \dagger & 0 \\
v_0\pi & \frac{\pi}{2} + \Delta' & \gamma_1 & -v_4\pi \dagger \\
-v_4\pi & \gamma_1 & -\frac{\pi}{2} + \Delta' & v_0\pi \dagger \\
0 & -v_4\pi & v_0\pi & -\frac{\pi}{2}
\end{pmatrix},
\tag{5.1}
\]

\[\pi = \hbar(\xi k_x + ik_y), \quad \pi \dagger = \hbar(\xi k_x - ik_y), \quad v_{0,4} = \frac{\sqrt{3a}}{2\hbar} \gamma_{0,4}.\]

Here, \( a = 2.46 \text{ Å} \) is the monolayer graphene lattice constant, the sign factor \( \xi = \pm 1 \) serves as the valley index corresponding to the valley wavevectors \( \mathbf{K}_\pm = (\pm \frac{4\pi}{3a}, 0) \). The wavevector \( \mathbf{k} = (k_x, k_y) \) is measured relative to \( \mathbf{K}_\pm \). The hopping parameters are
denoted by: $\gamma_0 = 2.61 \text{ eV}$ for the intralayer nearest neighbor hopping, $\gamma_1 = 0.361 \text{ eV}$ for the interlayer coupling between orbitals on the dimer sites B1 and A2, and $\gamma_4 = 0.138 \text{ eV}$ for the interlayer coupling between dimer and non-dimer orbitals A1 and A2 or B1 and B2. The parameter $\Delta' = 0.015 \text{ eV}$ describes the energy difference between dimer and non-dimer sites. The interlayer bias is given by $u = V_2 - V_1$ where $V_i$ is the potential on layer $i = 1, 2$.

We model the interfacially-induced SOI by introducing a layer-specific spin-orbit Hamiltonian to the monolayer subspace of the system [181], described by the components $A1$ and $B1$ of Eq.(5.1):

$$
\delta H_{SO}^{(1)} = \delta H_{Ising} + \delta H_{R},
\delta H_{Ising} = \frac{\lambda}{2} \xi s_z, \quad \delta H_{R} = \frac{\lambda_R}{2} (\xi \sigma_x s_y - \sigma_y s_x),
$$

where $s_i$ and $\sigma_i$ are the Pauli matrices corresponding to the spin and A/B sublattice degrees of freedom respectively. Under time reversal, the spin variables $s_i$, as well as the valley index $\xi$, change sign, whereas the sublattices $A$ and $B$ are not interchanged. The SOI Hamiltonian, Eq.(5.2), is therefore invariant under time reversal. However, it is not invariant under inversion.

Indeed, our interfacial SOI interaction is distinct in its symmetry properties from the intrinsic SOI for graphene monolayer analyzed by Kane and Mele [?]. Both of the SOI terms in Eq.(5.2) are extrinsic, i.e. they are allowed by symmetry only because of the presence of the transition metal dichalcogenide substrate [182]. In particular, the Ising term $\delta H_{Ising}$ is of the same form as the intrinsic Ising SOI of transition metal dichalcogenides with broken inversion symmetry [? ? ?]. Thus, the interfacial SOI induced in the graphene monolayer (and hence in the bilayer) also breaks the inversion symmetry. Likewise, the term $\delta H_{R}$, which has the standard low-energy form of Rashba SOI in graphene subject to a transverse electric field at the substrate/graphene interface [? ], also breaks the inversion symmetry.

We neglect the small intrinsic SOI terms of bilayer graphene, as well as Rashba SOI terms generated by the transverse electric field due to the interlayer potential between
the graphene layers. Spin splittings due to these effects have been estimated to be in the range of 1 to 100 $\mu$eV [100, 101, 102], which are much smaller than those arising from the interfacially-induced SOI in the meV range [154, 119, 118, 168, 169, 170, 171].

We introduce a perpendicular magnetic field $\mathbf{B} = B \hat{z}$ via the usual replacement of $k_i$ with $q_i = k_i - \frac{e}{\hbar} A_i$ ($i = x, y$) where $\mathbf{A} = (A_x, A_y)$ is the vector potential, $\mathbf{B} = \nabla \times \mathbf{A}$. We construct the magnetic ladder operators,

$$\hat{a} = \frac{l_B}{\sqrt{2}} (q_x + iq_y), \quad \hat{a}^\dagger = \frac{l_B}{\sqrt{2}} (q_x - iq_y),$$

which satisfy $[\hat{a}, \hat{a}^\dagger] = 1$, where $l_B = \sqrt{\frac{\hbar}{eB}}$ is the magnetic length. Substituting these quantities into the full Hamiltonian $H = H_0 + \delta H_{SO}^{(1)}$ of the system, we perform matrix diagonalization to solve for the LLs (see Section 5.11 for detailed derivation of the zero Landau levels of bilayer graphene with layer-specific spin-orbit interaction).

In particular, the eigenstates of $H$ corresponding to the zeroth LLs are well parameterized by $|\xi n s_z\rangle$,

$$H|\xi n s_z\rangle = E_{\xi n s_z}|\xi n s_z\rangle$$

Here $\xi = \pm 1$ is the valley index, $n = 0, 1$ is the orbital LL index, and $s_z = \pm 1$ is the out-of-plane spin polarization. To leading order in $u/\hbar \omega_c$, where $\omega_c$ is the cyclotron frequency, the corresponding single-particle energies are

$$E_{\xi n s_z} \simeq -E_Z s_z + n \Delta_{10} + \frac{u}{2} \alpha_{\xi n s_z} + \frac{\xi \lambda}{2} \zeta_{1, \xi n},$$

$$\Delta_{10} \approx \hbar \omega_c \left( \frac{\Delta'}{\gamma_1} + 2 \frac{\gamma_4}{\gamma_0} \right), \quad \hbar \omega_c = \frac{3a^2 \gamma_0^2}{2l_B^2 \gamma_1},$$

where $E_Z = \mu_B B$ is the Zeeman splitting, $\alpha_{\xi n s_z}$ is the layer polarization, $\zeta_{1, \xi n}$ is the spin polarization on layer 1 and $\Delta_{10}$ is the orbital splitting. These leading order energy corrections already account for most of the features given by the exact solutions, which are shown in Fig. 5-2 and Fig. 5-5 for typical values of $\lambda$, $\lambda_R$ and selected values of $B$. The effects of the layer-specific SOI can be understood by contrasting Fig. 5-2 and Fig. 5-5(b) to Fig. 5-5(a), which shows the zeroth LL spectrum in the absence
of SOI ($\lambda = \lambda_R = 0$). In what follows, we discuss some of these effects and their implications.

5.4 Effects of Layer-Specific Ising SOI

As we will see, the SOI-induced changes to the zeroth LL spectrum arise mainly due to the Ising SOI $\lambda$. This is illustrated by comparing the zeroth LL spectrum for $\lambda = \lambda_R = 5$ meV and $B = 5$ T, shown in Fig. 5-2(a), to Fig. 5-5(b), which shows the changes in the spectrum solely due to the $\lambda_R$ coupling for the same $B$ field strength as in Fig. 5-2(a). In comparison to the Ising SOI, the effect of the Rashba SOI $\lambda_R$
on the zeroth LLs is negligible and will not play a significant role in our analysis. A more detailed discussion of the reasons that effects due to $\lambda_R$ are small can be found below (see the penultimate paragraph).

The Ising term $\lambda$ generates a valley-antisymmetric Zeeman-like splitting with opposite signs at the two valleys. Depending on the relative sign between $\lambda$ and $B$ at a given valley, this Ising field induces a splitting that either assists or counteracts the Zeeman splitting for a given valley. Importantly, because this Ising field is layer-specific, its splitting is directly proportional to the layer-1 polarisation of the state. For the zeroth LL states, the valley polarization is essentially in one-to-one correspondence with the layer polarization of the state. Consequently, the layer-specific Ising field influences the zeroth LLs in a valley-asymmetric fashion, whereby it only modifies the spectrum of the LL states in the $K_+$ valley (in our convention) and not those in the $K_-$ valley.

5.5 Inverted Orbital Ordering

The most noticeable feature seen in Fig. 5-2 is the evolution of the $| + 0 \uparrow \rangle$ (black solid line) and $| + 1 \downarrow \rangle$ (blue solid line) energy levels with increasing $B$ field from $5 \, T$ in Fig. 5-2(a) to $31 \, T$ in Fig. 5-2(d). From Eq. (5.5), we see that this is a direct consequence of the competition between the orbital splitting $\Delta_{10}$ and the Ising splitting $\lambda$. When the Ising splitting dominates in relatively weak $B$ fields, it changes the ordering between the more energetic $n = 0$ and less energetic $n = 1$ states in the $K_+$ valley. Consequently, at filling level $\nu = 2$, one of the $n = 0$ and $n = 1$ states are filled instead of both $n = 0$ states.

The onset of this inverted orbital ordering depends on the relative orientation between the Ising field and the external $B$ field. This inversion occurs when $\delta E_{\lambda\Delta}^+ < 0$,

$$\delta E_{\lambda\Delta}^+ = E_{+1\uparrow} - E_{+0\uparrow}, \quad \lambda > 0, \quad (5.6)$$

$$\delta E_{\lambda\Delta}^- = E_{+1\uparrow} - E_{+0\downarrow}, \quad \lambda < 0. \quad (5.7)$$
which at \( u = 0 \) is approximately given by,

\[
|\lambda| \gtrsim \Delta_{10} + \text{sgn}(\lambda)2E_z \tag{5.8}
\]

\[
\approx \begin{cases} 
(0.381 + 0.116)B(T) \text{ meV}, & \lambda > 0 \\
(0.381 - 0.116)B(T) \text{ meV}, & \lambda < 0
\end{cases}
\]

so that it occurs over a larger range of \( B \) when the Ising and magnetic fields are aligned \((\lambda < 0)\) than when they are anti-aligned \((\lambda > 0)\).

In Fig. 5-3, we include the effects of the interlayer bias \( u \) and map out the phase diagrams for several values of \( \lambda = \pm 1, \pm 3, \pm 5 \text{ meV} \). This serves to assist visualizing the region in the three-dimensional \((u, B, \lambda)\) phase space in which orbital inversion occurs, i.e. when \( \delta E_{A}\lambda \Delta < 0 \) is satisfied. For \( \lambda > 0 \), the ordering inversion occurs between \( | + 1 \downarrow \rangle \) and \( | + 0 \uparrow \rangle \) while for \( \lambda < 0 \), ordering inversion occurs between \( | + 1 \uparrow \rangle \) and \( | + 0 \downarrow \rangle \). The occurrence of inverted orbital ordering at the non-interacting level will likely lead to novel phases near the \( \nu = 2 \) filling when interaction effects are included.

In particular, two observations can be made from Fig.5-3. First, the region in the \( B - u \) phase space with orbital ordering inversion increases with the magnitude of \( \lambda \). Second, these inversion regions are larger for \( \lambda < 0 \) than for \( \lambda > 0 \) of the same magnitude. These observations are consistent with what we have discussed above and are accounted for by Eq. (5.8).

### 5.6 Novel Compensated Electron-Hole Phases

Another interesting feature occurs at larger \( B \) fields when orbital ordering of energy levels is restored. The occurrence of this feature also requires that the Ising splitting opposes the Zeeman splitting. Applying a moderately large interlayer potential \(|u|\) dopes one layer into the electron band and the other into the hole band \([180]\). Three different compensated electron-hole phases exist as a result of the competition between the Ising splitting and Zeeman splitting, which we denote as para-phase, ortho-phase and ortho/para-phase. They are characterized by the alignment between the spin-polarization of their respective edge modes under positive and negative interlayer bias.
Figure 5-3: Phase diagrams in the $B-u$ plane showing how the regions of normal and inverted orbital ordering change with the Ising SOI $\lambda$ magnitude and sign. The panels a), b) and c) correspond to $\lambda > 0$ and normal/inverted ordering refers to the ordering between the levels $| + 1 \downarrow \rangle$ and $| + 0 \uparrow \rangle$: a) $\lambda = 5$ meV, b) $\lambda = 3$ meV, c) $\lambda = 1$ meV. The panels d), e) and f) correspond to $\lambda < 0$ and normal/inverted ordering refers to the ordering between the levels $| + 1 \uparrow \rangle$ and $| + 0 \downarrow \rangle$ (see Eq.(5.4) for notation): d) $\lambda = -5$ meV, e) $\lambda = -3$ meV, f) $\lambda = -1$ meV. Dashed lines indicate the phase boundary along which $\delta E_{\lambda\Delta} = 0$ [panels a)-c)], and $\delta E_{\lambda\Delta} = 0$ [panels d)-f)]. The energy difference $\delta E_{\lambda\Delta}$ (in meV) is indicated in Fig. 5-2(c).

$u$ (see Fig. 5-1):

Para-phase:
$$
\begin{align*}
|\text{electron, } \uparrow \rangle \otimes |\text{hole, } \downarrow \rangle, & \quad u > 0 \\
|\text{electron, } \uparrow \rangle \otimes |\text{hole, } \downarrow \rangle, & \quad u < 0
\end{align*}
$$

Ortho-phase:
$$
\begin{align*}
|\text{electron, } \uparrow \rangle \otimes |\text{hole, } \uparrow \rangle, & \quad u > 0 \\
|\text{electron, } \downarrow \rangle \otimes |\text{hole, } \downarrow \rangle, & \quad u < 0
\end{align*}
$$

Ortho/para-phase:
$$
\begin{align*}
|\text{electron, } \uparrow \rangle \otimes |\text{hole, } \uparrow \rangle, & \quad u > 0 \\
|\text{electron, } \uparrow \rangle \otimes |\text{hole, } \downarrow \rangle, & \quad u < 0
\end{align*}
$$

(5.9)

When the Ising field is absent or when the Zeeman splitting dominates, the com-
pensated electron-hole phase is in the para-phase, which corresponds to 'phase III' in Ref. [56]. In this phase, the filled hole and electron bands residing on opposite layers have opposite spin-polarization. When the substrate-induced Ising splitting dominates ($\lambda \gg 2E_z \simeq 0.116B(T)$ meV), the ordering of the spin-polarized states localized on the layer nearer to the substrate is now reversed. The corresponding compensated electron-hole phase is in the ortho-phase, in which case the filled electron band has the same spin-orientation as the filled hole band.

Unlike the para-phase, the ortho-phase is overall spin-neutral. In addition, having filled electron and hole bands with the same spin polarization means that their corresponding helical edge modes do not have protection from backscattering, unlike those of the para-phase. The edge conductivity of the ortho-phase is therefore expected to be strongly suppressed compared to that of the para-phase.

The mixed ortho/para-phase occurs at moderate $B$ fields when the Ising and Zeeman splittings are comparable. Because the $|+0s_z\rangle$ states are more strongly polarized on layer 1, they experience a stronger substrate-induced Ising field compared to the $|+1s_z\rangle$ states. The Ising splitting between the $|+0s_z\rangle$ states is therefore slightly larger than that between the $|+1s_z\rangle$ states, so that they do not necessarily have the same spin ordering for a given value of $B$. This can be seen from the larger spin splitting between the $|+0s_z\rangle$ states compared to that between the $|+1s_z\rangle$ states in Fig. 5-2(d). Consequently, the para-phase to ortho-phase transition occurs at different values of $B$ when $u > 0$ and when $u < 0$. To describe this ordering, we define the following parameters

$$
\delta E_{\lambda B}^+ = E_{+0\downarrow} - E_{+0\uparrow}, \quad u > 0, \\
\delta E_{\lambda B}^- = E_{+1\downarrow} - E_{+1\uparrow}, \quad u < 0.
$$

114
The conditions for the different phases are then given by

\begin{align*}
\text{Para-phase:} & \quad \delta E^+_{\lambda B} > 0, \quad \delta E^-_{\lambda B} > 0, \\
\text{Ortho-phase:} & \quad \delta E^+_{\lambda B} < 0, \quad \delta E^-_{\lambda B} < 0, \\
\text{Ortho/para-phase:} & \quad \delta E^+_{\lambda B} < 0, \quad \delta E^-_{\lambda B} > 0.
\end{align*}

(5.11)

The orbital splitting gives rise to a mixed ortho/para-phase in the $\lambda - B$ plane (see Fig. 5-4), in which the system can be thought of as being in the ortho-phase for $u > 0$ and being in the para-phase for $u < 0$. This mixed phase is particularly interesting because its edge conductivity can be switched on or off via switching the sign of the interlayer bias $u$. 
Figure 5-4: Phase diagram in the $\lambda - B$ plane, showing the different regions in which the system is found in one of the three different compensated electron-hole phases: the para-phase, the ortho-phase, and the ortho/para-phase. The notation for the energy differences $\delta E_{\lambda B}^\pm$ is given in Eq.(5.10).
5.7 Effects of Electron-Electron Interactions

While interaction effects are not included in this work, we expect the above discussion to remain qualitatively unchanged for $B \gtrsim 15\, \text{T}$. To understand why it is so, let us consider the $\lambda = 0$ case, in which the compensated electron-hole phase only exists in the para-phase (phase III in Ref. [56]). The $\nu = 0$ phase diagram mapped out in Ref. [56] suggests that the net effect of electron-electron interactions is to reduce the single-particle orbital splitting $\Delta_{10} \propto B$ which stabilizes the compensated electron-hole phase. In weak $B$ fields, orbital splitting is unable to overcome the interaction energy between the filled LLs on the same layer. In this case, the total energy of the system is minimized by one of two phases depending on the potential difference between the two layers. The ground state is in the layer-unpolarized canted antiferromagnetic phase (phase I in Ref. [56]) when the potential on both layers are comparable. Otherwise, when the potential on one layer is much larger than the other, the ground state is in the completely layer-polarized, spin-neutral phase (phase II in Ref. [56]). The compensated electron-hole phase does not exist as a ground state of the system in weak $B$ fields. On the other hand, when $B \gtrsim 15\, \text{T}$, $\Delta_{10}$ is sufficiently large and can exceed the interaction energy between the filled LLs on the same layer. In this case, the physics is qualitatively captured by the single-particle picture. At moderately large values of $u$, the energy of the $|-1\uparrow\rangle$ state becomes lower than that of the $|+0\downarrow\rangle$ state (see Fig.5-5(a)). The total energy of the system is therefore minimized by filling three LLs on one layer and one LL on the other – i.e. the compensated electron-hole phase.

The novel phases, ortho and ortho/para, predicted in this work, arise due to the competition between the Zeeman splitting and the layer-specific Ising SOI splitting. This competition gives rise to the normal and inverted spin-ordering between LLs of the same orbital number $n$ and is therefore independent of the orbital splitting. We therefore do not expect the electron-electron interactions to have a big effect on the spin ordering, since they mainly affect the orbital splitting and, through that, may alter the effective interlayer potential. This effect, however, will have no direct
impact on the competition between the SOI and Zeeman interactions that govern spin ordering. Therefore, so long as $B > 15 \text{T}$, the $\nu = 0$ ground state of the interacting system will be found in one of the three compensated electron-hole phases. This means that for an Ising field strength of $\lambda \gtrsim 2 \text{meV}$, all three phases are expected to be accessible even when the interaction effects are included, since the transitions between the different phases occur at $B \gtrsim 15$ (see Fig. 5-4).

Following this reasoning, we expect that for these moderately large values of $B$, the $\nu = 0$ layer-unpolarized canted antiferromagnetic phase discussed in Ref. [56] is likely to remain unchanged in the presence of a strong SOI substrate. However, this also means that the $\nu = 0$ phase diagram may change at weaker $B$ fields because of the inverted orbital ordering (see Fig. 5-3). A more detailed study including interaction effects is required to map out the phase diagram in this regime.

### 5.8 Conductances of Novel Phases

An immediate consequence of these three different phases is the difference between their two-terminal conductances $G$. The para-phase conductance, as discussed earlier, is expected to take a quantized value $G_{\text{para}} = 2e^2/h$ for both positive and negative values of $u$ (each edge contributes $e^2/h$ in parallel). For the ortho-phase, backscattering between the counterpropagating edge states is allowed. This causes the two terminal conductance to decay exponentially from $2e^2/h$ to 0 as the sample dimension increases. Therefore, for sufficiently large samples, we expect $G_{\text{ortho}} \approx 0$ for both positive and negative values of $u$. Finally, we expect the mixed ortho/para-phase to have a conductance that is gate-tunable $- G_{\text{ortho/para}}(u > 0) \approx 0$ and $G_{\text{ortho/para}}(u < 0) = 2e^2/h$.

### 5.9 Effects of Layer-Specific Rashba SOI

Finally, we comment on the effects of Rashba SOI $\lambda_R$, which are important at high LLs but are negligible at the zeroth LL. The smallness of the Rashba SOI $\lambda_R$ for the zeroth LL is illustrated in Fig.5-5. Indeed, by comparing the zeroth LL spectra
Figure 5-5: Illustration of the extreme smallness of the level shifts induced by the Rashba SOI $\lambda_R$ in the zeroth LL, as compared to those induced by the Ising SOI $\lambda$ illustrated in Fig. 5-2. The spectrum found in the absence of SOI interaction remains essentially unchanged after adding a relatively large Rashba SOI: a) $\lambda = \lambda_R = 0 \text{ meV}$ and b) $\lambda = 0$, $\lambda_R = 15 \text{ meV}$. Magnetic field is $B = 5 \text{ T}$ in both cases. Labeling of Landau levels (color and linestyle) is the same as that in Fig. 5-2.

shown in Fig. 5-5(a) to those in Fig. 5-5(b)($\lambda = 0$, $\lambda_R = 15 \text{ meV}$), it is evident that the effects of the Rashba SOI are strongly suppressed even at values of $\lambda_R$ as large as 15 meV. It gives rise to an energy correction $\approx 10^{-2} \text{ meV}$, which is comparable to the Zeeman energy at $B = 1 \text{ T}$, but quickly becomes negligible at larger field strengths ($B \gtrsim 5 \text{ T}$). It is therefore justified to ignore the correction due to $\lambda_R$ at leading order, which was done in Eq. (5.5). The physical reason for this smallness is as follows. The matrices $\sigma_i$ in the Rashba term generate $A1-B1$ couplings, which mix the zeroth LL state $|+,0,\downarrow\rangle = |A1,\downarrow,0\rangle$ with the dimer states at relatively high energies $\pm \gamma_1$. As a result, the $\lambda_R$-dependent corrections scale as $\delta \propto \lambda_R \left( \frac{\lambda_R}{\gamma_1} \right)$. Furthermore, because the high-energy states $\pm \gamma_1$ are particle-hole symmetric, their contributions cancel out at the lowest order. The $\lambda_R$-dependent corrections survive only at the next order, giving a small contribution to the level shifts of the order $\delta \propto \lambda_R \left( \frac{\lambda_R}{\gamma_1} \right)^2$.

5.10 Chapter Summary

In summary, the layer-specific SOI adds some unique features to the bilayer graphene single-particle zeroth LL spectrum. Those include the occurrence of an interesting pattern of orbital and spin inversions in the LL energy ordering. In contrast to the
SOI-induced splittings of higher LLs, which are dominated by the Rashba SOI, the zeroth LL splittings are dominated by the Ising SOI. Furthermore, the states with spin inversion give rise to novel compensated electron-hole ortho- and ortho/para-phases with a unique set of edge modes and gate-tunable edge conduction. While we anticipate the discovery of other novel phases when electron-electron interaction effects are taken into account, the compensated electron-hole ortho-phase and ortho/para-phase predicted here are expected to be robust to electron-electron interactions when $B \gtrsim 15$ T. The ortho/para-phase in particular provides a new knob that controls the edge-conductivity of quantum Hall systems. The gate tunability enables a field-effect transistor-like behavior of the edge states, a property that can lead to interesting applications of graphene-on-transition metal dichalcogenide heterostructures.

5.11 Appendix: Derivation of Zeroth LL Spectrum

In this section, I provide a more detailed derivation of the zeroth LL spectrum of BLG with layer-specific SOI. I follow the convention given in Ref. [80] to describe the spin-degenerate, low-energy effective Hamiltonian of bernal-stacked bilayer graphene including the hopping parameters $\gamma_0, \gamma_1, \gamma_4, \Delta'$ as per Ref. [56]. The effective Hamiltonian is given in the main text Eq. 5.1. The monolayer SOI Hamiltonian Eq. 5.2 can be written explicitly as

$$\delta H^{(1)}_{SO} = \begin{pmatrix} \frac{\lambda}{2} \xi s_z & \frac{\lambda R}{2} (\xi s_y + i s_x) & 0 & 0 \\ \frac{\lambda R}{2} (\xi s_y - i s_x) & \frac{\lambda}{2} \xi s_z & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (5.12)$$

and will be added to the layer-1 subspace of the Hamiltonian. Following the main text, I then introduce a perpendicular magnetic field and construct Landau level creation and annihilation operators $\hat{a}$ and $\hat{a}^\dagger$. Defining these operators as in Eq. 5.3 gives rise to the following valley-specific transformations,
\[ \xi = +: \begin{cases} \pi \rightarrow \frac{\sqrt{2\hbar}}{l_B} \hat{a} \\ \pi^\dagger \rightarrow \frac{\sqrt{2\hbar}}{l_B} \hat{a}^\dagger \end{cases} \quad (5.13) \]

\[ \xi = -: \begin{cases} \pi \rightarrow -\frac{\sqrt{2\hbar}}{l_B} \hat{a}^\dagger \\ \pi^\dagger \rightarrow -\frac{\sqrt{2\hbar}}{l_B} \hat{a} \end{cases} \quad (5.14) \]

which we can substitute directly into the full Hamiltonian of the system \[ H = H_0 + \delta H_{SO}^{(1)}, \]
and write the following valley-specific Hamiltonians in the \( (A1 \uparrow, A1 \downarrow, B1 \uparrow, B1 \downarrow, A2 \uparrow, A2 \downarrow, B2 \uparrow, B2 \downarrow) \) basis,

\[ H_+ = \hbar \omega_0 \begin{pmatrix} \frac{1}{2}(\hat{u}+\lambda) & 0 & \hat{a}^\dagger & 0 & -\frac{24}{\gamma_0} \hat{a}^\dagger & 0 & 0 & 0 \\ 0 & \frac{1}{2}(\hat{u}-\lambda) & i\lambda_R & \hat{a}^\dagger & 0 & -\frac{24}{\gamma_0} \hat{a}^\dagger & 0 & 0 \\ \hat{a} & -i\lambda_R & \frac{1}{2}(\hat{u}+\lambda) + \Delta' & 0 & \gamma_1 & 0 & -\frac{24}{\gamma_0} \hat{a}^\dagger & 0 \\ 0 & \hat{a} & 0 & \frac{1}{2}(\hat{u}-\lambda) + \Delta' & 0 & \gamma_1 & 0 & -\frac{24}{\gamma_0} \hat{a}^\dagger \\ -\frac{24}{\gamma_0} \hat{a} & 0 & \gamma_1 & 0 & -\frac{1}{2} \hat{u} + \Delta' & 0 & \hat{a}^\dagger & 0 \\ 0 & -\frac{24}{\gamma_0} \hat{a} & 0 & \gamma_1 & 0 & -\frac{1}{2} \hat{u} + \Delta' & 0 & \hat{a}^\dagger \\ 0 & 0 & -\frac{24}{\gamma_0} \hat{a} & 0 & \hat{a} & 0 & -\frac{1}{2} \hat{u} & 0 \\ 0 & 0 & 0 & -\frac{24}{\gamma_0} \hat{a} & 0 & \hat{a} & 0 & -\frac{1}{2} \hat{u} \end{pmatrix} \quad (5.15) \]

\[ H_- = \hbar \omega_0 \begin{pmatrix} \frac{1}{2}(\hat{u}-\lambda) & 0 & -\hat{a} & i\lambda_R & \frac{24}{\gamma_0} \hat{a} & 0 & 0 & 0 \\ 0 & \frac{1}{2}(\hat{u}+\lambda) & 0 & -\hat{a} & 0 & \frac{24}{\gamma_0} \hat{a} & 0 & 0 \\ -\hat{a}^\dagger & 0 & \frac{1}{2}(\hat{u}-\lambda) + \Delta' & 0 & \gamma_1 & 0 & \frac{24}{\gamma_0} \hat{a} & 0 \\ -i\lambda_R & -\hat{a}^\dagger & 0 & \frac{1}{2}(\hat{u}+\lambda) + \Delta' & 0 & \gamma_1 & 0 & \frac{24}{\gamma_0} \hat{a} \\ \frac{24}{\gamma_0} \hat{a}^\dagger & 0 & \gamma_1 & 0 & -\frac{1}{2} \hat{u} + \Delta' & 0 & -\hat{a} & 0 \\ 0 & \frac{24}{\gamma_0} \hat{a}^\dagger & 0 & \gamma_1 & 0 & -\frac{1}{2} \hat{u} + \Delta' & 0 & -\hat{a} \\ 0 & 0 & \frac{24}{\gamma_0} \hat{a}^\dagger & 0 & -\hat{a}^\dagger & 0 & -\frac{1}{2} \hat{u} & 0 \\ 0 & 0 & 0 & \frac{24}{\gamma_0} \hat{a}^\dagger & 0 & -\hat{a}^\dagger & 0 & -\frac{1}{2} \hat{u} \end{pmatrix} \quad (5.16) \]
For brevity, I introduce $\hbar \omega_0 = \hbar \sqrt{2m \nu} = \sqrt{\frac{3}{2} \frac{\sigma_0}{\nu}}$ and various barred quantities which are related to their unbarred counterparts via $x = \hbar \omega_0 \bar{x}$.

The creation and annihilation operators act on the site-specific Landau level wavefunctions $|n\rangle$ in the usual sense, $\hat{a}|n\rangle = \sqrt{n}|n-1\rangle$ and $\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$ so that by considering the following valley-specific ansatz for $n \geq 3$,

\[
|+, n, i\rangle = (c_{i,A1\uparrow}|n\rangle, c_{i,A1\downarrow}|n-1\rangle, c_{i,B1\uparrow}|n-1\rangle, c_{i,B1\downarrow}|n-2\rangle, c_{i,A2\uparrow}|n-2\rangle, c_{i,A2\downarrow}|n-1\rangle, c_{i,A2\uparrow}|n-1\rangle, c_{i,A2\downarrow}|n-1\rangle, c_{i,A2\uparrow}|n-1\rangle, c_{i,A2\downarrow}|n-1\rangle)^T,
\]

and

\[
|-, n, i\rangle = (c_{i,A1\uparrow}|n-2\rangle, c_{i,A1\downarrow}|n-3\rangle, c_{i,B1\uparrow}|n-1\rangle, c_{i,B1\downarrow}|n-2\rangle, c_{i,A2\uparrow}|n-2\rangle, c_{i,A2\downarrow}|n-1\rangle, c_{i,A2\uparrow}|n-1\rangle, c_{i,A2\downarrow}|n-1\rangle, c_{i,A2\uparrow}|n-1\rangle, c_{i,A2\downarrow}|n-1\rangle)^T,
\]

the Landau levels are given by the eigenvalues and eigenstates of the valley-specific matrices,

\[
H_{+,n \geq 3} = \hbar \omega_0 \\
\begin{pmatrix}
\frac{1}{2} \sigma_{i} & 0 & \sqrt{n} & 0 & -\frac{24}{\sigma_0} \sqrt{n} & 0 & 0 & 0 \\
0 & \frac{1}{2} \sigma_{i} & \sqrt{n-1} & 0 & -\frac{24}{\sigma_0} \sqrt{n-1} & 0 & 0 \\
\sqrt{n} & -i\lambda_R & \frac{1}{2} \sigma_i + \Delta' & 0 & \gamma_1 & 0 & -\frac{24}{\sigma_0} \sqrt{n-1} & 0 \\
\sqrt{n-1} & 0 & \frac{1}{2} \sigma_i + \Delta' & 0 & \gamma_1 & 0 & -\frac{24}{\sigma_0} \sqrt{n-2} \\
-\frac{24}{\sigma_0} \sqrt{n} & 0 & \gamma_1 & 0 & -\frac{1}{2} \sigma_i + \Delta' & 0 & \sqrt{n-1} & 0 \\
0 & -\frac{24}{\sigma_0} \sqrt{n-1} & 0 & \gamma_1 & 0 & -\frac{1}{2} \sigma_i + \Delta' & 0 & \sqrt{n-2} \\
0 & 0 & -\frac{24}{\sigma_0} \sqrt{n-1} & 0 & \sqrt{n-1} & 0 & -\frac{1}{2} \sigma_i & 0 \\
0 & 0 & 0 & -\frac{24}{\sigma_0} \sqrt{n-2} & 0 & \sqrt{n-2} & 0 & -\frac{1}{2} \sigma_i
\end{pmatrix}
\]

(5.17)

\[
H_{-,n \geq 3} = \hbar \omega_0 \\
\begin{pmatrix}
\frac{1}{2} \sigma_{i} & 0 & -\sqrt{n-1} & i\lambda_R & \frac{24}{\sigma_0} \sqrt{n-1} & 0 & 0 & 0 \\
0 & \frac{1}{2} \sigma_{i} & 0 & -\sqrt{n-2} & 0 & \frac{24}{\sigma_0} \sqrt{n-2} & 0 & 0 \\
-\sqrt{n-1} & 0 & \frac{1}{2} \sigma_i - \Delta' & 0 & \gamma_1 & 0 & \frac{24}{\sigma_0} \sqrt{n} & 0 \\
-i\lambda_R & -\sqrt{n-2} & 0 & \frac{1}{2} \sigma_i - \Delta' & 0 & \gamma_1 & 0 & \frac{24}{\sigma_0} \sqrt{n-1} \\
\frac{24}{\sigma_0} \sqrt{n-1} & 0 & \gamma_1 & 0 & -\frac{1}{2} \sigma_i - \Delta' & 0 & -\sqrt{n} & 0 \\
0 & \frac{24}{\sigma_0} \sqrt{n-2} & 0 & \gamma_1 & 0 & -\frac{1}{2} \sigma_i + \Delta' & 0 & -\sqrt{n-1} \\
0 & 0 & \frac{24}{\sigma_0} \sqrt{n} & 0 & -\sqrt{n} & 0 & -\frac{1}{2} \sigma_i & 0 \\
0 & 0 & 0 & \frac{24}{\sigma_0} \sqrt{n-1} & 0 & -\sqrt{n-1} & 0 & -\frac{1}{2} \sigma_i
\end{pmatrix}
\]

(5.18)
By setting to zero all other energy parameters in Eq. 5.1 except $\gamma_0$ and $\gamma_1$, we recover the eight-fold degenerate (2 spin, 2 valley, 2 orbital) zeroth LLs states with zero energy, $|\xi, 0, s_z\rangle$ and $|\xi, 1, s_z\rangle$. By considering $H_{\pm,n=0,1,2}$ in this limit, we can understand how the zeroth LL degeneracy is lifted (i.e. how these states mix with the higher Landau levels and the terms responsible for the mixing).

At $n = 0$, the solutions are immediately given,

$$
H_{+,0} = \frac{1}{2}(u + \lambda), \quad \text{eigenstate: } |+, 0, \uparrow\rangle = |A1 \uparrow, 0\rangle
$$

(5.19)

$$
H_{-,0} = -\frac{1}{2}u, \quad \text{eigenstate: } |-, 0, \uparrow\rangle = |B2 \uparrow, 0\rangle.
$$

(5.20)

The eigenstates and energies are exact. At $n = 1$, we have

$$
H_{+,1} = \hbar \omega_0 \begin{pmatrix}
\frac{1}{2}(\bar{u} + \bar{\lambda}) & 0 & 1 & -\frac{2\Delta}{\gamma_0} \\
0 & \frac{1}{2}(\bar{u} - \bar{\lambda}) & i\bar{\lambda}_R & 0 \\
1 & -i\bar{\lambda}_R & \frac{1}{2}(\bar{u} + \bar{\lambda}) + \bar{\Delta}' & \bar{\gamma}_1 \\
-\frac{2\Delta}{\gamma_0} & 0 & \bar{\gamma}_1 & -\frac{1}{2}\bar{u} + \bar{\Delta}'
\end{pmatrix},
$$

basis: $(|A1 \uparrow, 1\rangle, |A1 \downarrow, 0\rangle, |B1 \uparrow, 0\rangle, |A2 \uparrow, 0\rangle)$

(5.21)

$$
H_{-,1} = \hbar \omega_0 \begin{pmatrix}
\frac{1}{2}(\bar{u} - \bar{\lambda}) + \bar{\Delta}' & \bar{\gamma}_1 & \frac{2\Delta}{\gamma_0} & 0 \\
\bar{\gamma}_1 & -\frac{1}{2}\bar{u} + \bar{\Delta}' & -1 & 0 \\
\frac{2\Delta}{\gamma_0} & -1 & -\frac{1}{2}\bar{u} & 0 \\
0 & 0 & 0 & -\frac{1}{2}\bar{u}
\end{pmatrix},
$$

basis: $(|B1 \uparrow, 0\rangle, |A2 \uparrow, 0\rangle, |B2 \uparrow, 1\rangle, |B2 \downarrow, 0\rangle)$.

(5.22)

Keeping only $\gamma_0$ and $\gamma_1$ non-trivial and setting the other parameters to 0, we find
that the zeroth LL eigenstates are

\[
|+, 1, \uparrow\rangle_0 = \frac{1}{\sqrt{\bar{\gamma}_1^2 + 1}} (-\bar{\gamma}_1 |A1 \uparrow, 1\rangle + |A2 \uparrow, 0\rangle),
\]

(5.23)

\[
|+, 0, \downarrow\rangle_0 = |A1 \downarrow, 0\rangle,
\]

(5.24)

\[
|-, 1, \uparrow\rangle_0 = \frac{1}{\sqrt{\bar{\gamma}_1^2 + 1}} (|B1 \uparrow, 0\rangle + \bar{\gamma}_1 |B2 \uparrow, 1\rangle),
\]

(5.25)

\[
|-, 0, \downarrow\rangle_0 = |B2 \downarrow, 0\rangle,
\]

(5.26)

With the above information, we can now discuss the effects due to the various parameters in the Hamiltonian as corrections which are justified in the strong field limit \(\hbar \omega_0 \approx 31 \text{ meV} \sqrt{B(T)} \gg \gamma_4, \Delta', u, \lambda, \lambda_R\) but constrained to \(\hbar \omega_0 \lesssim \gamma_1\) or equivalently \(\bar{\gamma}_1 \gtrsim 1\). The effects from SOI are already discussed in the main text and here I include a short discussion on the effects of the non-SOI parameters for completeness. The interlayer bias \(u\) shifts the energies by an amount that measures the layer polarization of the state as is expected, \(+\frac{u}{2}\) if the state is completely polarized in layer 1 and \(-\frac{u}{2}\) if it is completely polarized in layer 2. Similarly, the \(\Delta'\) term gives rise to an energy shift proportional to the state’s polarization on the \(A2\) or \(B1\) sites. Finally, \(\gamma_4\) introduces mixing between the different sublattice components of the \(|+, 1, \uparrow\rangle_0\) and \(|-, 1, \uparrow\rangle_0\) respectively and shifts the state’s energy by the difference between the amount of symmetric and anti-symmetric superposition of the sublattice components.
At \( n = 2 \), we have

\[
H_{+, 2} = \hbar \omega_0 \begin{pmatrix}
\frac{1}{2} (\bar{u} + \lambda) & 0 & \sqrt{2} & 0 & -\frac{24}{\gamma_0} \sqrt{2} & 0 & 0 \\
0 & \frac{1}{2} (\bar{u} - \lambda) & i\lambda_R & 1 & 0 & -\frac{24}{\gamma_0} & 0 \\
\sqrt{2} & -i\lambda_R & \frac{1}{2} (\bar{u} + \lambda) + \Delta' & 0 & \gamma_1 & 0 & -\frac{24}{\gamma_0} \\
0 & 1 & 0 & \frac{1}{2} (\bar{u} - \lambda) + \Delta' & 0 & \gamma_1 & 0 \\
-\frac{24}{\gamma_0} \sqrt{2} & 0 & \gamma_1 & 0 & -\frac{1}{2} \bar{u} + \Delta' & 0 & 1 \\
0 & -\frac{24}{\gamma_0} & 0 & \gamma_1 & 0 & -\frac{1}{2} \bar{u} + \Delta' & 0 \\
0 & 0 & -\frac{24}{\gamma_0} & 0 & 1 & 0 & -\frac{1}{2} \bar{u}
\end{pmatrix},
\]

basis: \( (|A1 \uparrow, 2\rangle, |A1 \downarrow, 1\rangle, |B1 \uparrow, 1\rangle, |B1 \downarrow, 0\rangle, |A2 \uparrow, 1\rangle, |A2 \downarrow, 0\rangle, |B2 \uparrow, 0\rangle) \) \hspace{1cm} (5.27)

\[
H_{-, 2} = \hbar \omega_0 \begin{pmatrix}
\frac{1}{2} (\bar{u} - \lambda) & -1 & i\lambda_R & \frac{24}{\gamma_0} & 0 & 0 & 0 \\
-1 & \frac{1}{2} (\bar{u} - \lambda) + \Delta' & 0 & \gamma_1 & 0 & \frac{24}{\gamma_0} \sqrt{2} & 0 \\
-i\lambda_R & 0 & \frac{1}{2} (\bar{u} + \lambda) + \Delta' & 0 & \gamma_1 & 0 & \frac{24}{\gamma_0} \\
\frac{24}{\gamma_0} & \gamma_1 & 0 & -\frac{1}{2} \bar{u} + \Delta' & 0 & -\sqrt{2} & 0 \\
0 & 0 & \gamma_1 & 0 & -\frac{1}{2} \bar{u} + \Delta' & 0 & -1 \\
0 & -\frac{24}{\gamma_0} \sqrt{2} & 0 & -\sqrt{2} & 0 & -\frac{1}{2} \bar{u} & 0 \\
0 & 0 & -\frac{24}{\gamma_0} \sqrt{2} & 0 & -1 & 0 & -\frac{1}{2} \bar{u}
\end{pmatrix},
\]

basis: \( (|A1 \uparrow, 0\rangle, |B1 \uparrow, 1\rangle, |B1 \downarrow, 0\rangle, |A2 \uparrow, 1\rangle, |A2 \downarrow, 0\rangle, |B2 \uparrow, 2\rangle, |B2 \downarrow, 1\rangle) \). \hspace{1cm} (5.28)

Once again, turning off all parameters except \( \gamma_0 \) and \( \gamma_1 \) allows us to recover the remaining 2 zeroth LL states:

\[
|+, 1, \downarrow \rangle_0 = \frac{1}{\sqrt{\gamma_1^2 + 1}} (\mp \gamma_1 |A1 \downarrow, 1\rangle + |A2 \downarrow, 0\rangle), \hspace{1cm} (5.29)
\]

\[
|-, 1, \downarrow \rangle_0 = \frac{1}{\sqrt{\gamma_1^2 + 1}} ((|B1 \downarrow, 0\rangle + \gamma_1 |B2 \downarrow, 1\rangle), \hspace{1cm} (5.30)
\]

125
The qualitative nature of how the various parameters affect these two zeroth LL states is similar to that of the $n = 1$ case discussed earlier.

Consistent with the above discussion, the single-particle energies of the zeroth LL to leading order in $u/\hbar \omega_c$ is given by main text Eq. 5.5, and here I give the explicit expressions for the layer and spin polarizations that were omitted in the main text,

$$\alpha_{\zeta,n,s_z} = |c_{A1,s_z}|^2 + |c_{B1,s_z}|^2 - |c_{A2,s_z}|^2 - |c_{B2,s_z}|^2$$  \hspace{1cm} (5.31)

$$\zeta_{1,n,s_z} = s_z \left( |c_{A1,s_z}|^2 + |c_{B1,s_z}|^2 \right).$$  \hspace{1cm} (5.32)
Chapter 6

Conclusions

A wide range of interesting phenomena have been discovered in graphene. The vdW heterostructures certainly add a lot of excitement to graphene research by introducing novel interfacial physics. In this thesis, I have discussed how heterostructures help to introduce interfacial SOI in graphene. This opens up a plethora of new phenomena that are awaiting to be discovered. I explored some interesting effects that arise when SOI is interfacially-induced in bilayer graphene via a TMD substrate. Special emphasis was given to gate-tunable properties of the bilayer graphene-on-TMD heterostructure.

The experimental verification of the change to the graphene bandstructure as a result of interfacially-induced SOI from a TMD substrate was reviewed in chapter 3. It is especially important that TMD substrates are able to directly modify the band structure of graphene and give rise to carriers that are spin-split in energy. Analogous to the intrinsic spin-Hall effect, a spin-split band structure enables robust control over the spin degree of freedom (chapter 4). It can also modify the system’s topological properties (chapter 4) and lead to novel phases (chapter 5). Several extensions to the ideas presented in this thesis are discussed below.

The Magnitude of Induced SOI. The experiments performed at relatively high doping clearly confirm the presence of spin-split bands (chapter 3). However, at present there is insufficient data at low carrier densities and thus it is impossible to accurately determine the relative strength of the Rashba and Ising-type SOI. These
values can be precisely determined by measurements at the energies near the Dirac point, in which case the spin-splitting (if present) will be completely due to the Ising SOI. An alternative approach is to directly probe the zeroth LLs and look for effects due to a layer-specific SOI. As shown in chapter 5, the effects of Rashba SOI is strongly suppressed in the zeroth LL compared to the Ising SOI. In particular, the Zeeman splitting competes against the Ising SOI responsible for spin-ordering inversion within the zeroth LLs. An observable that is sensitive to the spin-ordering will therefore show a dependence on the strength of the applied transverse magnetic field. This effect is expected to be suppressed by applying a strong in-plane magnetic field in conjunction with a relatively weaker transverse magnetic field. The strong in-plane field acts to overwhelm the system’s spin-ordering while not introducing other field dependent energy splittings in the LL spectrum. At the time of writing this thesis, an ongoing collaboration with the University of Santa Barbara, where these measurements are being made, indicates that the interfacially induced Ising SOI can have strength of up to 1.5 to 2 meV.

*Realizing topological materials.* Achieving a sufficiently strong Ising SOI is desirable as it can produce band inversion and give rise to topologically non-trivial states. With this in mind, a few interesting heterostructure configurations can be immediately proposed. The Ising SOI breaks inversion symmetry, with the symmetry breaking occurring due to the TMD layer on one side of the system. The symmetry can be restored by introducing TMD layers on both sides of the system. The opposing Ising SOI induced by the different TMD layers will likely cancel each other out in the case of monolayer graphene when the heterostructure is inversion symmetric. However, in the bilayer or trilayer graphene, for example, these interfacially induced SOI act on different layers and, therefore, on different low-energy states of the system. The resulting heterostructures are expected to exhibit a rich topological phase diagrams. A more thorough investigation, which goes beyond what was presented in chapter 4, is required for a complete characterization of the various phases in the case of TMD-bilayer-graphene-TMD heterostructures.

*Interfacially-induced SOI on twisted bilayer graphene.* Another interesting plat-
form to introduce interfacially-induced SOI is the twisted bilayer graphene, which has electron properties very different from those of the Bernal-stacked bilayer graphene considered in this thesis. Likewise, the cases in which SOI is introduced on one or both graphene layers can be separately investigated. In light of the Mott-insulating and superconducting phases that were recently discovered in magic angle twisted bilayer graphene, introducing SOI can give rise to novel phases. These include novel classes of Mott-insulating phases or topological superconductors.

*Gate-tuning other layer-specific effects.* Besides the interfacially-induced SOI considered above, it is interesting to consider introducing other layer-specific effects. Such effects can modify the bandstructure of bilayer graphene in a variety of different ways. Analogous to our findings for interfacially-induced SOI, such modifications of the bilayer graphene bandstructure will be gate-tunable. Examples of such gate-tunable effects include superconductivity, disorder physics, as well as superlattice potentials/Moire physics. Alternatively, tunability beyond electrostatic gating or based on different underlying physics can also be explored. With the improved ability to construct layered 2D heterostructures, we find ourselves in the perfect moment to pursue these questions not only theoretically, but experimentally as well.

The work undertaken in this thesis has barely scratched the surface of the rapidly growing field of 2D van der Waals heterostructures. It follows the notion 'design and build' in an attempt to extend the success of the semiconducting heterostructures. The ultimate goal therefore is to have a comprehensive set of tools at our disposal to create heterostructures with any desired functionality. As we are in the stage of discovering and building up our toolbox, collaborations between theory and experiment are ever more so important. I am therefore grateful to have had the opportunity to experience this firsthand and be a part of this grand endeavor.
Bibliography


140


[125] Junsaku Nitta, Tatsushi Akazaki, Hideaki Takayanagi, and Takatomo Enoki. Gate control of spin-orbit interaction in an inverted in$_{0.53}$ga$_{0.47}$as/in$_{0.52}$al$_{0.48}$as heterostructure. Phys. Rev. Lett., 78:1335–1338, Feb 1997.


