UNVEILING THE TOPOLOGICAL QUANTUM EFFECTS IN DEFECT-ENGINEERED TOPOLOGICAL INSULATOR FILMS

By

MARYAM SALEHI

A dissertation submitted to the
School of Graduate Studies
Rutgers, The State University of New Jersey
In partial fulfillment of the requirements
For the degree of
Doctor of Philosophy
Graduate Program in Materials Science and Engineering
Written under the direction of
Professor Seongshik Oh
And approved by

New Brunswick, New Jersey
May, 2019
ABSTRACT OF THE DISSERTATION

UNVEILING THE TOPOLOGICAL QUANTUM EFFECTS IN DEFECT-ENGINEERED TOPOLOGICAL INSULATOR FILMS

By MARYAM SALEHI

Dissertation Director: Seongshik Oh

Topological insulators (TIs) are a class of quantum materials where the bulk is insulating, while the surfaces host in-gap metallic states. The topological surface states (TSSs) are described by the Dirac Hamiltonian whose spin is locked to momentum, and as a result their metallic behavior is protected against crystal imperfections. These features of TSS make TIs a unique test bed for exploration of various new physics and applications. However, almost all the first-generation TIs suffered from a high level of defect densities which pushed the Fermi level away from the bulk energy gap (where the Dirac TSS lies) into the conduction band, thereby obscuring the electronic signature of the TSS.

This dissertation seeks to shed light on the physical origins of defects in pnictogen-chalcogenide TIs, as well as the various defect and interface engineering schemes that have been exploited to effectively suppress these defects. Here, we explain how growing TI films on virtually-grown structurally/chemically-compatible substrates led to films with 10 to 100 times lower defect
densities compared to the first-generation films, and with the Fermi level being only few tens of meV above the Dirac point. This is followed by a discussion about the important role of an effective capping layer in stabilizing the properties of low-carrier-density TI films. We further discuss how employing such optimally-designed buffer and capping layers along with proper compensation-doping resolved the previously challenging task of carrier-type tunability in TI films. We then expound on how the ultralow-carrier-density Bi$_2$Se$_3$ and Sb$_2$Te$_3$ films enabled the observation of a series of previously unseen topological quantum effects. Additionally, we present how the ultralow-carrier-density Sb$_2$Te$_3$ films allowed us to reach an extreme quantum limit and explore the previously inaccessible zeroth Landau level. Finally, we conclude by providing a summary of results as well as an outlook towards the future of TI research.
Acknowledgements

First and foremost, I would like to thank my advisor, Prof. Seongshik Oh, for his guidance and patience. This dissertation would have not been possible without his continual support and his invaluable scientific and technical insights. I am grateful for his optimism and his incredibly infectious enthusiasm for physics which boosted my energy and motivation during the low points of my research. His scientific and life advice will always remain with me.

Additionally, I would like to thank my past and present lab-mates, Matthew Brahlek, Nikesh Koirala, Jisoo Moon, Namrata Bansal, Deepti Jain, and Pavel Shibayev for their help, cooperative attitude, and overall for making my PhD years more enjoyable. In particular, I am extremely thankful to my friends Matthew Brahlek and Nikesh Koirala for guiding me, especially during my first year in the lab. Also, I thank Jisoo Moon for sharing the work-load over these years. Doing experiments with Deepti Jain and Pavel Shibayev was a joyful experience.

I would like to thank my committee members, Prof. Leonard Feldman, Prof. Eric Garfunkel, Prof. Deirdre O’Carroll, and Prof. Laura Fabris, and the late Prof. James Harrington for their time and inputs. I really enjoyed attending Prof. Harrington’s Advanced Optical Materials course and always appreciated his support and encouragement.

I am thankful to all our collaborators for the helpful discussions and fruitful collaborations. Special thanks go to Prof. Peter Armitage as well as his former graduate student, Liang Wu, his former postdoc, Mintu Mondal, and his current student, Dipanjan Chaudhuri, for carrying out THz spectroscopy on our thin-films. I thank Prof. Howard Katz and his student, Robert Ireland, for charge-transfer doping as well as charged polymer capping study on our Bi$_2$Se$_3$ samples. I thank Prof. David Goldhaber-Gordon and his student, Ilan Thomas Rosen, for fabricating gated devices on our Sb$_2$Te$_3$ thin-films and measuring them. I thank the people from the neighboring lab, Prof. Weida Wu and his group, for performing scanning tunneling microscopy on our samples. Many
thanks go to Prof. Bob Bartynski and his group members, especially the late Boris Yakshinskiy, the late Leszek Wielunski, and Ryan Thorpe for carrying out Rutherford backscattering experiments on our samples and for training me to become a user of X-ray photoemission spectroscopy facility. I thank Prof. Girsh Bloomberg and his student, Sean Kung, for performing Raman spectroscopy study on our thin-films. I thank Prof. Michael Gershenson and his student, Wen-Sen Lu, for helping us with experiments in the physics department clean room. I would also thank Prof. Dale J. Van Harlingen and his students, Can Zhang and Erik Huemiller, for fabricating superconductor-topological insulator-superconductor Josephson junctions on our thin-films. I thank Myung-Guen Han and Kim Kisslinger for performing high-angle annular dark-field scanning transmission electron microscopy and energy-dispersive X-ray spectroscopy analysis on our samples using facilities in the center for functional nanomaterial at Brookhaven National Laboratory. I thank prof. Sang-Wook Cheong’s group for performing atomic force microscopy on some of our samples. Special thanks go to William Schneider from the machine shop for helping me design and machine some of the required tools for our lab and to Marc Lesko for training me on various machine shop instruments, such as drill press, lathe, milling, saw, etc. Last but not least, I would like to thank the National High Magnetic Field laboratory in Tallahassee for giving us the opportunity to use the state-of-the-art magnet facilities for the past couple of years. I would like to particularly thank Tim Murphy, Scott Hannahs, Hongwoo Baek, Eun-Sang Choi, Scott Maier, Alexy Suslov and all the other staff at NHMFL for their support and for helping us perform our measurements in 35 T and 45 T magnet systems.

I am incredibly indebted to my beloved husband, Hassan, for his limitless encouragement, companionship, patience, and love. Immigrating to the United States and going through ups and downs of the PhD journey together helped us develop a deep mutual understanding and continual support of each other in every aspect of life. I am extremely lucky to have an intelligent and talented physicist and more importantly a good human by my side. Hassan, learning from you and
discussing physics and other general topics about science and technology with you is always joyful. I had a lot of fun collaborating with you on two of my research projects and I hope we will again share this experience in the future. You push me every day to be the better version of myself and I am utterly grateful for having you in my life.

Finally, I would like to thank my Mother, Mahnaz, and my Father, Ardeshir, and my younger siblings, Dorsa and Mohammadreza, for their love, support, and for putting up with the physical separation. Maman and Baba, I never would have come this far if it was not for your encouragement, selflessness, generosity, and open-mindedness. Words cannot express my gratitude towards you. I also thank my parents-in-law for their support and encouragement.
Dedication

To Maman, Baba, and my love, Hassan,

for their limitless support and love.
Table of Contents

ABSTRACT ......................................................................................................................... ii
Acknowledgements .............................................................................................................. iv
Dedication ............................................................................................................................ vii
List of Tables ...................................................................................................................... xi
List of Figures .................................................................................................................... xii
List of abbreviations .......................................................................................................... xviii

Chapter 1 Prologue ........................................................................................................... 1
  1.1 Topology and the classification of materials electronic structure ....................... 1
  1.2 Topological phases in 2 and 3 dimensions .............................................................. 2
  1.3 Spin-momentum locking and weak anti-localization .............................................. 8
  1.4 Experimental realization ......................................................................................... 11
  1.5 Other topological materials .................................................................................... 15
  1.6 Scope of present work .............................................................................................. 18

Chapter 2 Thin-film growth and characterization ......................................................... 19
  2.1 Short History of thin-film growth ............................................................................. 19
  2.2 Thin-film growth process ........................................................................................ 20
    2.2.1 Physisorption and chemisorption .................................................................... 21
    2.2.2 Different arrangements and growth modes ...................................................... 22
    2.2.3 Defects in crystals ........................................................................................... 26
  2.3 Different thin-film growth techniques ...................................................................... 28
    2.3.1 Chemical vapor deposition .............................................................................. 28
    2.3.2 Sputtering ......................................................................................................... 29
    2.3.3 Pulsed laser deposition ..................................................................................... 31
    2.3.4 Molecular Beam epitaxy .................................................................................. 32
  2.4 MBE system at Rutgers University .......................................................................... 34
  2.5 Atomic flux measurement techniques ...................................................................... 39
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5.1</td>
<td>Quartz crystal microbalance</td>
<td>40</td>
</tr>
<tr>
<td>2.5.2</td>
<td>Rutherford Back scattering</td>
<td>42</td>
</tr>
<tr>
<td>2.6</td>
<td>Structural characterization techniques</td>
<td>45</td>
</tr>
<tr>
<td>2.6.1</td>
<td>Reflection high energy electron diffraction</td>
<td>46</td>
</tr>
<tr>
<td>2.6.2</td>
<td>Other characterization tools</td>
<td>52</td>
</tr>
<tr>
<td>2.7</td>
<td>Growth of pnictogen chalcogenide TIs: Van der Waals epitaxy</td>
<td>55</td>
</tr>
<tr>
<td>2.7.1</td>
<td>Substrate preparation</td>
<td>55</td>
</tr>
<tr>
<td>2.7.2</td>
<td>Growth on different substrates</td>
<td>57</td>
</tr>
<tr>
<td>2.8</td>
<td>Electrical transport measurements</td>
<td>65</td>
</tr>
<tr>
<td>2.8.1</td>
<td>Drude transport</td>
<td>66</td>
</tr>
<tr>
<td>2.8.2</td>
<td>Hall effect: Determination of type and number of charge carrier density</td>
<td>69</td>
</tr>
<tr>
<td>2.9</td>
<td>Experimental details of transport measurement</td>
<td>70</td>
</tr>
<tr>
<td>2.9.1</td>
<td>Geometry: van der Pauw and Hall-bar</td>
<td>70</td>
</tr>
<tr>
<td>2.9.2</td>
<td>Preparing the sample</td>
<td>73</td>
</tr>
<tr>
<td>2.9.3</td>
<td>Reverse field reciprocity and symmetrization</td>
<td>78</td>
</tr>
<tr>
<td>2.9.4</td>
<td>Magnets</td>
<td>80</td>
</tr>
<tr>
<td>2.10</td>
<td>Measurements at the National High Magnetic Field Laboratory</td>
<td>82</td>
</tr>
</tbody>
</table>

Chapter 3 Defect and interface engineering of topological insulators thin-fims ...... 85

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>The origin of Bulk Defects</td>
<td>86</td>
</tr>
<tr>
<td>3.1.1</td>
<td>the Mott criterion</td>
<td>87</td>
</tr>
<tr>
<td>3.1.2</td>
<td>The Surface versus bulk defects: Band-bending effect</td>
<td>88</td>
</tr>
<tr>
<td>3.2</td>
<td>Conventionally grown TI films with no buffer layer: substrate matters</td>
<td>92</td>
</tr>
<tr>
<td>3.3</td>
<td>Defect suppression via interface-engineering: new generation TI films with a buffer layer</td>
<td>95</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Bi$_2$Se$_3$ films on In$_2$Se$_3$/BiInSe$_3$ buffer layer: growth and transport properties</td>
<td>95</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Sb$_2$Te$_3$ films on In$<em>2$Se$<em>3$/Sb$</em>{0.65}$In$</em>{0.35}$Te$_3$ buffer layer: growth and transport properties</td>
<td>101</td>
</tr>
<tr>
<td>3.4</td>
<td>Topological surface states-originated response</td>
<td>106</td>
</tr>
<tr>
<td>3.4.1</td>
<td>ARPES measurements</td>
<td>108</td>
</tr>
<tr>
<td>3.4.2</td>
<td>THz measurements</td>
<td>110</td>
</tr>
<tr>
<td>3.5</td>
<td>Defect suppression via impurity addition</td>
<td>114</td>
</tr>
<tr>
<td>3.5.1</td>
<td>Compensation-Doping</td>
<td>114</td>
</tr>
<tr>
<td>3.5.2</td>
<td>Ternary and quaternary compounds and isovalent alloying</td>
<td>117</td>
</tr>
<tr>
<td>3.6</td>
<td>Stabilizing the properties of TI films: capping layer matters</td>
<td>118</td>
</tr>
<tr>
<td>3.7</td>
<td>Conclusion</td>
<td>125</td>
</tr>
</tbody>
</table>
Chapter 4 Unveiling the topological quantum effects ............................................. 127

4.1 Quantum Hall effect for topological surface states .............................................. 127
   4.1.1 TSS-originated QHE in low-carrier-density Bi₂Se₃ films .............................. 130
   4.1.2 QHE and QH-to-insulator transition in low-carrier-density Sb₂Te₃ films ......... 132

4.2 Quantized Faraday and Kerr rotation ................................................................. 150

4.3 Finite-size driven topological and metal-insulator transition in (Bi₁₋ₓInₓ)₂Se₃ thin-films .......................................................... 153

4.4 Conclusion ........................................................................................................... 161

Chapter 5 Epilogue ................................................................................................... 163

5.1 Summary of the results ........................................................................................ 163

5.2 Outlook and future direction ............................................................................... 165

References .................................................................................................................. 173
List of Tables

Table 3.1 Ti vapor pressure as a function of temperature................................................................. 104

Table 3.2 Some of the Ti fluxes at lower temperatures along with corresponding vapor pressure and calculated doping used for the growth of ultralow carrier density Sb₂Te₃ films. ............... 104

Table 3.3 Optical and DC transport parameters of two 16 QL thick Bi₂Se₃ films grown on BIS-BL with Se and MoO₃ capping layers along with a 15 QL uncapped film. ................................................. 114
List of Figures

Figure 1.1 The quantum Hall state (QHS) versus the quantum spin Hall state (QSHS) ...................... 3

Figure 1.2 Schematic of the spin-orbit coupling effect.......................................................... 4

Figure 1.3 Surface state dispersion and the Kramer’s pairs at time-reversal symmetric points (k = 0 and ±π/a) of the Brillouin zone ................................................................. 5

Figure 1.4 The surface band structure of a 3D topological insulator is a Dirac cone ................. 6

Figure 1.5 Band inversion in Bi₂Se₃ ................................................................. 7

Figure 1.6 Forbidden backscattering of topological surface states .......................................... 10

Figure 1.7 Experimental observation of 2D TI ......................................................................... 12

Figure 1.8 Band structure of Bi₀.₉Sb₀.₁ versus Bi₂Se₃ ...................................................... 13

Figure 1.9 Pnictogen (group V) chalcogen (group VI)-based TIs .............................................. 14

Figure 1.10 Topological crystalline insulator ........................................................................ 15

Figure 1.11 The band structure and transport of SmB₆ .......................................................... 16

Figure 1.12 Experimental observation of 3D Dirac semi-metal by angle-resolved photoemission spectroscopy ........................................................................................................... 17

Figure 2.1 Diagram of physisorption and chemisorption process ............................................ 22

Figure 2.2 The possible arrangements of adsorbed atoms on a substrate ................................. 23

Figure 2.3 A schematic of possible epitaxial growth modes ................................................... 25
Figure 2.4 A schematic of 3 different crystallographic defects.......................................................... 27

Figure 2.5 A schematic of magnetron sputtering system. ................................................................. 30

Figure 2.6 A schematic of pulsed laser deposition. ............................................................................. 32

Figure 2.7 A schematic of an MBE system. ......................................................................................... 33

Figure 2.8 Schematic of chalcogenide and oxide-based molecular beam epitaxy systems in Rutgers University. ........................................................................................................ 36

Figure 2.9 MBE system in Oh’s lab at Rutgers. .................................................................................. 38

Figure 2.10 QCM system on the MBE in Prof. Oh’s Lab. ................................................................. 42

Figure 2.11 Rutherford backscattering spectroscopy system located at the Laboratory for Surface Modification at Rutgers University. .............................................................................. 43

Figure 2.12 RBS experiment for a (Bi$_{1-x}$In$_x$)$_2$Se$_3$ film. .......................................................... 44

Figure 2.13 Reflection high energy electron diffraction................................................................. 47

Figure 2.14 Universal curve for electrons mean free path within a solid as a function of the energy of the impinging beam. .......................................................................................... 48

Figure 2.15 Ewald circle showing the diffraction of an electron beam by a crystal lattice............ 49

Figure 2.16 RHEED images of various crystals with 2D and 3D surface structure.................. 51

Figure 2.17 X-ray diffraction on (Bi$_{1-x}$In$_x$)$_2$Se$_3$ films. .......................................................... 53

Figure 2.18 AFM images of Bi$_2$Se$_3$ grown on various substrates reveal different surface morphologies...................................................................................................................... 54
Figure 2.19 Substrate preparation and cleaning steps................................................................. 57

Figure 2.20 RHEED and TEM images of Bi$_2$Se$_3$ grown on Si (111) using the two-step growth technique................................................................. 60

Figure 2.21 RHEED images of Bi$_2$Se$_3$ grown at different temperatures on Si (111) substrate.... 61

Figure 2.22 RHEED images of Bi$_2$Se$_3$ growth procedure on Al$_2$O$_3$ substrate....................... 63

Figure 2.23 RHEED images of Bi$_2$Se$_3$ growth procedure on amorphous SiO$_2$ substrates. ....... 64

Figure 2.24 The Bi$_2$Se$_3$ film etching and transfer process to an arbitrary substrate. ............... 65

Figure 2.25 A schematic of the scattering processes of a charge carrier in a material under an applied electric field. ................................................................. 67

Figure 2.26 Schematics of a Hall effect measurement. ................................................................. 69

Figure 2.27 A schematic of different geometries used in transport measurement...................... 71

Figure 2.28 Sample cutting and Hall-bar scratch procedure....................................................... 74

Figure 2.29 Sample preparation and mounting on the Advance Research Systems cryostat and GMW magnet................................................................. 75

Figure 2.30 A schematic of the source meter and switch matrix setup. ........................................ 76

Figure 2.31 The front interface of the custom-built program using LabView................................. 77

Figure 2.32 An example of sample-sample method. ................................................................. 79

Figure 2.33 Photos of the American Magnetics, Inc. liquid helium cryostat. ............................... 81
Figure 2.34 Images of the 45 T hybrid magnet at the National High Magnetic Field Laboratory.

Figure 3.1 Band-bending dependence on the TSS carrier density and thickness.

Figure 3.2 Transport properties of the Bi₂Se₃ films grown on Al₂O₃ and Si substrates.

Figure 3.3 The growth procedure and characterization of the Bi₂Se₃ films grown on the In₂Se₃/BiInSe₃ buffer layer.

Figure 3.4 The RBS and STM on Bi₂Se₃ films grown on the In₂Se₃/BiInSe₃ buffer layer.

Figure 3.5 ARPES confirming the topological nature of Bi₂Se₃ grown on buffer layer.

Figure 3.6 Comparison of $n_{sheet}$ and $\mu$ of Bi₂Se₃ grown on Al₂O₃ (0001), Si (111), and on the In₂Se₃/BiInSe₃ buffer layer (BIS-BL).

Figure 3.7 Growth procedure and characterization of buffer layer-based Sb₂Te₃ films.

Figure 3.8 Transport Properties of various Sb₂Te₃ films.

Figure 3.9 The two-carrier model fit to the nonlinear Hall effect for Bi₂Se₃ films grown BIS-BL.

Figure 3.10 Finding the TSS wave vector using ARPES of Bi₂Se₃ grown on Al₂O₃.

Figure 3.11 Faraday rotation at different magnetic fields, Cyclotron resonance, and zero-field real conductance for two 16 QL thick Bi₂Se₃ films grown on BIS-BL and capped by 20 nm Se and 50 nm MoO₃.

Figure 3.12 Transport properties of Ca-doped Bi₂Se₃ films.

Figure 3.13 Transport properties of the Cu doped Bi₂Se₃ grown on Al₂O₃ (0001).
Figure 3.14 Aging effect on room temperature transport properties of 30 QL Bi$_2$Se$_3$ films with and without various capping layers. .......................................................... 120

Figure 3.15 Baking temperature dependence of aging effect for room temperature transport properties of 30 QL Bi$_2$Se$_3$ films with PMMA-capping. ......................................................... 122

Figure 3.16 Aging of Bi$_2$Se$_3$ films measured by time domain THz spectroscopy at 5 K. ........ 124

Figure 4.1 The edge-state picture of the QHE. ................................................................. 128

Figure 4.2 Surface quantum Hall Effect in an 8 QL thick MoO$_3$-capped Bi$_2$Se$_3$ film grown on In$_2$Se$_3$/BiInSe$_3$ buffer layer. ................................................................. 131

Figure 4.3 Transport Properties of 8QL films with various Ti-doping levels. ......................... 133

Figure 4.4 Absence of a well-defined QHE in p-type Ca-doped Bi$_2$Se$_3$ films and the absence of LLs on the p-side of Bi$_2$Se$_3$ in STS. ................................................................. 134

Figure 4.5 The conductivity tensor flow ($\sigma_{xx}$ vs. $\sigma_{xy}$) of N8-1.5% for $|B| \leq 35$ T ............. 136

Figure 4.6 A possible scenario for the insulating behavior could be a (quantum) Hall insulator. 137

Figure 4.7 The schematic of energy spectrum of LLs along with the edge modes for different 2D systems in a magnetic field with and without considering the Zeeman coupling. ................. 138

Figure 4.8 Surface LLs and edge modes on the side surfaces. ........................................... 139

Figure 4.9 The longitudinal conductivity of a gated device at zero field, gate-tuned to the charge neutrality point, shown on a log scale as a function of inverse temperature. ................. 140

Figure 4.10 Theoretical modeling of the effect of random impurity potential on LLs in the absence (a and b) and the presence (c and d) of tunneling-induced surface gap. ............................... 141
| Figure 4.11 | Scaling-invariant plots for an 8QL Sb$_2$Te$_3$ film with the lowest carrier density | 143 |
| Figure 4.12 | Scaling analysis for N8-1.3% and more detail on high-field measurements of N8-1.5%. | 144 |
| Figure 4.13 | More details on the scaling analysis for N8-1.5% | 146 |
| Figure 4.14 | Transport properties of a gated 8QL film with a 2nm (Sb$_{0.65}$In$_{0.35}$)$_2$Te$_3$ capping | 147 |
| Figure 4.15 | The raw transport data of the gated 8QL Sb$_2$Te$_3$ film | 148 |
| Figure 4.16 | $n_{\text{sheet}}$ as a function of gate-voltage the gated 8QL Sb$_2$Te$_3$ film | 149 |
| Figure 4.17 | Schematic of measurement setup for Faraday and Kerr rotation experiments | 151 |
| Figure 4.18 | Quantized Faraday and Kerr rotations | 152 |
| Figure 4.19 | Transport properties of (Bi$_{1-x}$In$_x$)$_2$Se$_3$ thin-films | 155 |
| Figure 4.20 | Metal-insulator transition and Weak anti-localization in buffer layer-based (Bi$_{1-x}$In$_x$)$_2$Se$_3$ films | 157 |
| Figure 4.21 | The schematic of the phase transition along with the phase diagram for different TIs | 158 |
| Figure 4.22 | Theoretical phase diagram and simulated thickness dependence of the surface and bulk gap developments upon TPT | 160 |
| Figure 5.1 | The generic phase diagram of quantum materials | 172 |
List of abbreviations

(TI) Topological insulator
(2DEG) 2-dimensional electron gas
(QHS) Quantum Hall state
(QSHS) Quantum spin Hall state
(TRS) Time reversal symmetry
(SOC) Spin orbit coupling
(E_F) Fermi level
(WAL) Weak anti-localization
(HLN) Hikami-Larkin-Nagoka
(TSS) Topological surface state
(WL) Weak localization
(ARPES) Angle-resolved photoemission spectroscopy
(STS) Scanning tunneling spectroscopy
(STM) Scanning tunneling microscopy
(QL) Quintuple layer
(TCI) Topological crystalline insulator
(TKI) Topological Kondo insulator
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPT</td>
<td>Topological phase transition</td>
</tr>
<tr>
<td>PLD</td>
<td>Pulsed laser deposition</td>
</tr>
<tr>
<td>MBE</td>
<td>Molecular beam epitaxy</td>
</tr>
<tr>
<td>CVD</td>
<td>Chemical vapor deposition</td>
</tr>
<tr>
<td>TEM</td>
<td>Transmission electron microscopy</td>
</tr>
<tr>
<td>ALD</td>
<td>Atomic layer deposition</td>
</tr>
<tr>
<td>MOCVD</td>
<td>Metal-organic chemical vapor deposition (MOCVD)</td>
</tr>
<tr>
<td>LPCVD</td>
<td>Low pressure chemical vapor deposition</td>
</tr>
<tr>
<td>UHV-CVD</td>
<td>Ultra-high vacuum chemical vapor deposition,</td>
</tr>
<tr>
<td>PECVD</td>
<td>Plasma-enhanced chemical vapor deposition</td>
</tr>
<tr>
<td>RHEED</td>
<td>Reflection high energy electron diffraction</td>
</tr>
<tr>
<td>UHV</td>
<td>Ultrahigh vacuum</td>
</tr>
<tr>
<td>CF</td>
<td>Conflat flange</td>
</tr>
<tr>
<td>PBN</td>
<td>Pyrolytic boron nitride</td>
</tr>
<tr>
<td>QCM</td>
<td>Quartz crystal microbalance</td>
</tr>
<tr>
<td>RBS</td>
<td>Rutherford backscattering</td>
</tr>
<tr>
<td>AFM</td>
<td>Atomic force microscopy</td>
</tr>
<tr>
<td>XRD</td>
<td>X-ray diffraction</td>
</tr>
<tr>
<td>STO</td>
<td>SrTiO₃</td>
</tr>
<tr>
<td>IPA</td>
<td>Isopropyl alcohol</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>PMMA</td>
<td>Polymethyl methacrylate</td>
</tr>
<tr>
<td>BOE</td>
<td>Buffer oxide etchant</td>
</tr>
<tr>
<td>DC</td>
<td>Direct current</td>
</tr>
<tr>
<td>NHMFL</td>
<td>National high magnetic field lab</td>
</tr>
<tr>
<td>YBCO</td>
<td>Yttrium Barium copper oxide</td>
</tr>
<tr>
<td>2DEG</td>
<td>2-dimensional electron gas</td>
</tr>
<tr>
<td>LED</td>
<td>Light emitting diode</td>
</tr>
<tr>
<td>2DFG</td>
<td>2-dimensional Fermi gas</td>
</tr>
<tr>
<td>CB</td>
<td>Conduction band</td>
</tr>
<tr>
<td>BL</td>
<td>Buffer layer</td>
</tr>
<tr>
<td>BIS-BL</td>
<td>In$_2$Se$_3$/BiInSe$_3$ buffer-layers</td>
</tr>
<tr>
<td>HAADF-STEM</td>
<td>High-angle annular dark-field scanning transmission electron microscopy</td>
</tr>
<tr>
<td>EDS</td>
<td>Energy-dispersive X-ray spectroscopy</td>
</tr>
<tr>
<td>TDTS</td>
<td>Time-domain terahertz spectroscopy</td>
</tr>
<tr>
<td>TDTMS</td>
<td>Time-domain magneto-terahertz spectroscopy</td>
</tr>
<tr>
<td>CR</td>
<td>Cyclotron resonance</td>
</tr>
<tr>
<td>FR</td>
<td>Faraday rotation</td>
</tr>
<tr>
<td>QHE</td>
<td>Quantum Hall effect</td>
</tr>
<tr>
<td>LL</td>
<td>Landau level</td>
</tr>
<tr>
<td>ZLL</td>
<td>Zeroth Landau level</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Term</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>CNP</td>
<td>Charge neutrality point</td>
</tr>
<tr>
<td>QIT</td>
<td>Quantum to insulator transition</td>
</tr>
<tr>
<td>MIT</td>
<td>Metal-insulator transition</td>
</tr>
<tr>
<td>TNT</td>
<td>Topological to normal transition</td>
</tr>
<tr>
<td>NM</td>
<td>Normal metal</td>
</tr>
<tr>
<td>QAHE</td>
<td>Quantum anomalous Hall effect</td>
</tr>
<tr>
<td>FM</td>
<td>Ferromagnetic</td>
</tr>
<tr>
<td>AHE</td>
<td>Anomalous Hall effect</td>
</tr>
<tr>
<td>MTI</td>
<td>Magnetic topological insulator</td>
</tr>
<tr>
<td>YIG</td>
<td>$Y_3Fe_5O_{12}$</td>
</tr>
<tr>
<td>TIG</td>
<td>$Tm_3Fe_5O_{12}$</td>
</tr>
<tr>
<td>SGGG</td>
<td>Substituted gadolinium gallium garnet</td>
</tr>
<tr>
<td>TSC</td>
<td>Topological superconductor</td>
</tr>
<tr>
<td>CMEMs</td>
<td>Chiral Majorana edge modes</td>
</tr>
</tbody>
</table>
Chapter 1

Prologue

1.1 Topology and the classification of materials electronic structure

The classification of solids in terms of their electronic properties, the band theory, is one of the earliest applications of the quantum mechanics to the solid-state physics. According to the band theory, solids are either conductor where the highest energy band (i.e. the conduction band) is partially filled or insulator (semiconductor) where some of the energy bands are fully filled and the rest are empty, and an energy gap separates the valence band and the conduction band. However, later by discovery of topological insulator (TI)\(^1\)\(^\text{2-3}\), it turned out that this classification is incomplete. Indeed, insulators can be divided into two types: trivial and topological. Both types are characterized by an insulating bulk with finite energy gap, while the topological phase has an additional property that its boundary (edge in 2-dimensional and surface in 3-dimensional) states are conducting. These conducting boundary states are intrinsic and related to topological properties of bulk wave functions.

The word “topological” is used to emphasize the fact that the topological phase cannot be reached by inducing slight (adiabatic) changes into the trivial and a phase transition must take place in order to go from a trivial phase into a topological phase. In essence, different insulating phases are very similar to geometric shapes that can be categorized into different topological classes based on a “genus” number \(g\) defined by Gauss and Bonnet, which can be obtained by integrating the Gaussian curvature \(K\) of a 3D object over its entire 2D surface manifold, yielding \(X_M = \frac{1}{2\pi} \oint_M K ds = 2(1 - g)^4\). The genus number is quantized to integer numbers and can be thought of as the number of holes in a geometric shape (for example: no hole \(\rightarrow g = 0\) and one hole \(\rightarrow g = 1\)) and is insensitive to smooth deformations. Based on “rubber geometry”\(^5\) the properties are
not destroyed under continuous deformations, and two objects with the same number of holes can smoothly be transformed to each other. However, a dramatic change is required, for instance, to turn a sphere (with no holes) to a torus (with a single hole). The notion of topology can be extended to physics where the geometric shapes can now be replaced by the wave function of electrons and the ordering of the energy bands. Similar to the genus number, a topological invariant number \( \nu \) (e.g. a Chern number in 2D\(^{b,c}\)) can be defined by integrating the Berry curvature (derived from the properties of Bloch wave functions) of all filled bands over the entire Brillouin zone, which can then be used to classify materials electronic structure into different topological classes. Under this classification, for example, Al\(_2\)O\(_3\) or In\(_2\)Se\(_3\) falls into a trivial insulator category with \( \nu = 0 \), whereas Bi\(_2\)Se\(_3\) is a topological insulator with \( \nu = 1 \).

### 1.2 Topological phases in 2 and 3 dimensions

For simplicity, let us start by discussing the topological phases in two dimensions. One way to understand the nature of boundary states is by using an analogy to the quantum Hall effect, a 2D electron gas (2DEG) system of electrons subject to a strong magnetic field at low temperatures where all the electrons become localized in the bulk, leading to an insulating phase with vanishing longitudinal conductance\(^8\). Under a magnetic field \( (B) \), the band structure transforms into discrete energy levels known as the Landau levels with energies \( E_n = \hbar \omega_c (N + \frac{1}{2}) \), where \( e \) is electron charge, cyclotron frequency \( \omega_c = \frac{eB}{m^*} \) (\( m^* \) is the electron effective mass), and \( N = 0,1,2, \ldots \) is the LL indices. Whenever the Fermi level sits between two consecutive \( N \) and \( N+1 \) LLs, the bulk of the 2DEG becomes insulating with \( N+1 \) number of conducting channels developing at edges. The existence of such edge modes can also be understood qualitatively in the following picture: a moving particle is deflected by the magnetic field (as a result of Lorentz force) towards the edge of the sample, where it cannot complete the cyclotron orbit, and hence is repeatedly bounced off the edge till it gets to the other end of the sample. In other words, these edge modes provide one-way
paths for electrons where the path direction is dictated by the direction of the magnetic field (upward/downward).

The insulating bulk of this system, also known as the quantum Hall state (QHS), with its Fermi level located in the gap of two adjacent LLs is indistinguishable from a trivial insulator with its Fermi level sitting in the bulk gap –one can think of flattening the bulk bands of the trivial insulator to get the LLs of the QHS. However, because of the propagating chiral edge modes, the QHS is in fact topologically non-trivial insulator. The QHS features quantized Hall conductance plateaus at multiple integer of conductance quantum $\frac{e^2}{h}$ which gives $\sigma_{xy} = \nu \frac{e^2}{h}$, where $\nu$ is the topological invariant or the Chern number and corresponds to the number of filled LLs. In fact, the reason that these metallic edge states emerge is because two materials with different topological invariant are interfaced: in this case the topologically non-trivial QHS with broken time-reversal symmetry is interfaced with the trivial insulator vacuum; this is known as bulk-boundary correspondence (Figure 1.1).

![Quantum Hall and Quantum Spin Hall](image)

**Figure 1.1** The quantum Hall state (QHS) versus the quantum spin Hall state (QSHS). **a,** The QHS has a single chiral edge mode that moves in a specific direction imposed by the direction of the magnetic field. In this case the time-reversal symmetry is broken. **b,** The QSHS, however, has helical spin-polarized edge modes that move in both directions. In this case the time-reversal symmetry is preserved, and the edge modes are time-reversed partners. The QSHS can be thought as two copies of the QHS. Adapted from Reference 11.

One interesting question that now comes to mind is whether one can realize these edge modes in a crystalline material without resorting to an external magnetic field and without breaking the time-
reversal symmetry (TRS) and only by utilizing/engineering the intrinsic properties of the material. Remarkably, such materials do exist and are generally known as topological insulators. Among various crystal fields experienced by electrons in a crystal, the spin-orbit coupling (SOC) is the closest analog to the magnetic field. The SOC couples the electron spin to its orbital motion through a relativistic mechanism where the nuclei electrostatic potential can be viewed as an effective magnetic field (Figure 1.2). This effective magnetic field has opposite signs for opposite spin directions. Therefore, the electrons in materials with sufficiently strong spin-orbit coupling (heavy elements with heavy nuclei) can be decomposed into two copies of quantum Hall state (with $\nu = 1$) with opposite magnetic fields corresponding to two electron spins, and hence it is named the quantum spin Hall state (QSHS)\(^1\) which is basically a two-dimensional TI (See Figure 1.1).

![Figure 1.2 Schematic of the spin-orbit coupling effect. a, An electron moving in the reference frame of the ions. b, From the electron's point of view, the positively-charged ions appear be moving. The ions' motion generates an effective magnetic field that couples to the electron's spin. The heavier the ions (nuclei) the larger the effective magnetic field is. In real solids, the spin–orbit coupling is more complex, but the essence of the interaction is quite the same as it is shown here.](image)

In other words, the quantum spin Hall state has two counter-propagating helical edge modes (two-way path) at each edge in presence of time-reversal symmetry as opposed to the quantum Hall
effect (Figure 1.2). The quantum spin Hall effect was observed in CdHgTe quantum wells\(^5\). The emergence of a pair of edge modes (instead of one) can also be understood as a result of Kramers’ theorem which implies two-fold degeneracy in time-reversal symmetric points, which means that at the points \(k = 0\) and \(\pm \pi/a\), where \(- k \equiv k\), the edge bands cross. Since strong SOC acts like a magnetic field, it splits the bands based on the spin of the electrons away from the TRS symmetric points.

Figure 1.3 Surface state dispersion and the Kramer’s pairs at time-reversal symmetric points \((k = 0\) and \(\pm \pi/a\)) of the Brillouin zone. a, A trivial insulator with even number of Fermi level crossing, where the surface states are not protected and b, a strong TI with odd number of Fermi level crossings, where the surface states are protected and guaranteed to exist. Adapted from Reference 1.

To see how TRS protects the edge modes, consider two band structures with an even and odd number of Fermi level crossings in the energy gap (Figure 1.3). In the former case (Figure 1.3a), the bands can be shifted into bulk bands by adding perturbations to the system; this adiabatically deforms the system into a trivial insulator and results in zero Fermi level crossing. On the other hand, in the latter case (Figure 1.3b), regardless of how these energy bands are deformed and shifted there will always be at least one, but more generally an odd number of Fermi level crossings, which preserves the gapless edge modes. Note that all the even number of crossings can eventually be adiabatically changed to zero crossing and all the odd number of Fermi level crossings can be
adiabatically deformed to one crossing; thus, if we allow for smooth deformations of the system, there are only two topologically distinct bands. That is to say the classification is $\mathbb{Z}_2$ where the topological invariant takes only values 0 or 1.

**Figure 1.4** The surface band structure of a 3D topological insulator is a Dirac cone. **a**, The schematic of a Dirac cone in a 3D TI. Since in reality the materials are far from being ideal, the Fermi level ($E_F$) does not necessarily cross the Dirac point. **b**, Electrons have opposite spin 'orientation' at states with opposite wavevector $k$ and $-k$. Adapted from Reference 12. **c**, This can also be thought as spin-up electrons move forward and spin-down electrons move backward for each direction on the surface of a 3D TI sample.

Furthermore, it is indeed natural to think about possible outcomes of strong SOC in 3D materials. The analog of two-way edge modes in 2D TI would be a Dirac cone surface states in 3D TI (Figure 1.4a). As in the case of a 2D TI where the spin (up/down) determines the direction of the edge mode, in the case of 3D TI, there is a continuum of surface modes where for each momentum direction the spin is pointed in a particular direction (see Figure 1.4b) and the spin vector rotates continuously as we circularly move around the origin. This is called the spin-momentum locking (or helical spin pattern) and is a direct manifestation of topological Dirac states on the surface of a 3D TI. The surface states of a 3D TI can be written in terms of a linearly-dispersed 2D massless
Figure 1.5 Band inversion in Bi$_2$Se$_3$. a, The order of energy bands is determined by bonding as well as crystal field, however strong spin-orbit coupling of Bi can locally change the order of the bands. Adapted from Reference 13. b, The simplified band structure of Bi$_2$Se$_3$ which qualitatively shows the band structure close to and away from the $\Gamma$ point. The p-orbital content of Bi and Se are embedded as blue and red, respectively. In the vicinity of the point the bands are inverted, and the top of the valence band has more Bi characteristic and the bottom of the conduction band has more Se characteristic. Away from the $\Gamma$ point the bands are not affected by spin-orbit coupling. The surface bands are 50% contribution of Bi and Se p-orbitals. Adapted from Reference 14.

Dirac fermion Hamiltonian $H_{surface} = \hbar v_F \sigma \cdot \vec{k} = \hbar v_F (k_x \sigma_x + k_y \sigma_y)$, where $v_F$ is the Fermi velocity and $\sigma$ is the Pauli matrix acting on the spin degree of freedom. The second equality in terms of two wave vectors ($k_x$ and $k_y$) and the two Pauli matrices show the 2D nature of TSSs. This Hamiltonian manifestly implies a helical spin structure of the TSS. It is worth noting that one may not conclude that any insulator with a strong spin-orbit coupling is going to be a topological insulator. In fact, SOC must be strong enough to invert (exchange) the valence and conduction bands in the bulk to allow for a change in the topological invariant (i.e. Chern number). In addition to the bulk band inversion, other crystal properties such as lattice symmetry groups must be taken into account to determine whether a crystal is topological or not.
In other words, SOC is a necessary factor but not sufficient for making a material topological. Once the strong SOC changes the order of energy bands and inverts the bulk gap (locally in the momentum space), we would get non-trivial bands. The statement of bulk-boundary correspondence is that as we move from the non-trivial bulk towards an interface with vacuum (i.e., trivial bands), energy bands have to unwind – become gapless and reopen/de-invert again – upon crossing the interface. Figure 1.5 shows the band inversion in an archetypal Tl, Bi₂Se₃, where the blue and red colors in Figure 1.5b corresponds to Bi and Se p-orbital content, respectively. It is clear that the strong spin-orbit coupling changes the order of the energy bands. Thus, in the vicinity of the Γ point, the top of the valence band has more Bi characteristic and the bottom of the conduction band has more Se characteristic, whereas away from the Γ point the order of the bands is not affected by SOC. The metallic surface states are 50% Bi and 50% Se contribution.

1.3 Spin-momentum locking and weak anti-localization

Due to spin-momentum locking, the backscattering off non-magnetic defects is suppressed for TSS carriers. A simple way to understand this is by noting that the probability of backward scattering vanishes due to destructive interference. Consider the electron backward path upon which the electron momentum and hence spin rotate by 180°, shown schematically in Figure 1.6. There are two possible paths for this backscattering and total probability is obtained by summing the amplitude for both clockwise (180°) and counter-clockwise (-180°) backscattering events. This leads to a destructive interference and thus the probability of backscattering vanishes. The reason is because the spin rotates by π in opposite directions through these two paths, effectively there is a +π − (−π) = 2π relative angle between the final states of the spin in the two paths. A peculiar feature of spin ½ is that 2π rotation of spin ½ returns the wave function to its original state up to a minus sign ψ to −ψ. In other words, for a fermion with spin ½, a 4π rotation is needed so that the spin goes back to its original state without a phase factor. Therefore, this fact implies that the two
backward paths interfere destructively, i.e., $\psi = \psi_{ccw} + \psi_{cw} = 0$. That is to say, the electrical resistance (as compared to its classical value) goes to zero. The robustness of surface states against non-magnetic defects can make them promising candidates for applications in the quantum computation and spintronics.

Upon application of a magnetic field to a TI, this destructive interference is impaired and backscattering protection is weakened, and as a result a small decrease in electrical conductance (or increase in electrical resistance) of a TI sample is observed. The quantum correction to the classical conductance value in disordered systems with SOC is known as weak anti-localization (WAL) effect\textsuperscript{15} which shows up as a cusp in resistance at zero field in magnetoresistance measurements. WAL diminishes quickly in an applied magnetic field and can be used as a quantitative tool to extract useful information about the transport properties of a TI material\textsuperscript{16}. While WAL is conventionally associated with strong spin-orbit coupling, which is present in TIs, a more powerful mechanism is responsible for the WAL observed in 3D TIs. The Dirac surface states of 3D topological insulators have a $\pi$ Berry phase, so time-reversed paths always destructively interfere, and WAL is always observed. With the introduction of magnetic field, the phase coherence diminishes rapidly, and this shows up as a sharp increase in resistance with magnetic field. In a two-dimensional system (i.e. thickness < $l$, $l$ is the dephasing length), this qualitative description can be made quantitative by fitting the magneto-conductance data with Hikami-Larkin-Nagaoka (HLN) formula\textsuperscript{16}: $\Delta G(B) = \frac{\tilde{A}e^2}{2\pi h} \left[ \ln \left( \frac{B \phi}{B} \right) - \Psi \left( \frac{1}{2} + \frac{B \phi}{B} \right) \right]$, where $\tilde{A}$ corresponds to the number of independent conducting channels in the system (i.e. $\tilde{A}$ is 1 if there is a single channel), $e$ is the electron charge, $h$ is the Planck’s constant, $B \phi$ is the de-phasing field, and $\Psi(x)$ is the digamma function.

As it will be explained in more details in later chapters, in practice the grown TI materials are far from being an ideal TI because the parasitic bulk conduction obscures the topological surface
Figure 1.6 Forbidden backscattering of topological surface states. Due to spin-momentum locking in TIs, if an electron backscatter off a non-magnetic impurity clockwise (top) the spin rotates by $\pi$ and if it goes counter clockwise (bottom) the spin rotates by $-\pi$. Adding the contributions shows that the amplitude vanishes for 180° backscattering. Thus, the difference between these two paths is $2\pi$ and for a half-integer spin this is associated with a phase factor $-1 (\psi \rightarrow -\psi)$ which eventually leads to destructive interference and thus the suppressed backscattering.

states (TSS) response. In TIs with conducting bulk, the usual scenario is that electrons/holes on the surface can couple to the bulk states, and therefore, the whole system acts like a single coherently coupled channel, resulting in $\tilde{A} = 1^{17-22}$ which is far from the ideal value $\tilde{A} = 2$ for true TI films with two decoupled surfaces. If, on the other hand, the bulk is insulating or if one of the TSS is decoupled from bulk, then this results in $\tilde{A} = 2$. For more details regarding WAL theory in TIs, see Reference 23. Additionally, in thin films, tunneling between the top and bottom surface states hybridizes their wave functions, opening a gap around the Dirac point and changing the Berry phase of the surface state bands. When the Fermi level is far from the gap, the system acts like a single WAL channel, resulting in $\tilde{A} = 1$. As the $E_F$ approaches the gap, a crossover of the quantum corrections to weak localization (WL) has been observed$^{24}$. A crossover from WAL to WL also occurs when time reversal symmetry is broken through the addition of magnetic dopant/layer$^{25-26}$. 
Only recently, $\tilde{\alpha}= 2$ has been reported in ungated Cu-doped Bi$_2$Se$_3$ films grown on sapphire. Chen et al. and Steinberg et al. showed that upon depleting the bulk states near one surface using electrostatic gating, $\tilde{\alpha}$ gradually increases from 1 to 2. Also, $\tilde{\alpha}=2$ has been achieved by growing a tunable non-TI (Bi$_{1-x}$In$_x$)$_2$Se$_3$ layer between two TI Bi$_2$Se$_3$ layers. In this heterostructure, the non-TI layer with sufficient thickness, and if it is properly tuned to an insulating phase by varying In concentration, fully decouples the top and bottom TI Bi$_2$Se$_3$ layers (each acting as a single channel), resulting in total 2 conducting channels. More interestingly, it has been shown that $\tilde{\alpha}$ scales nearly linearly with the number of interface pairs in MBE-grown superlattices of the TI Bi$_2$Se$_3$/normal insulator In$_2$Se$_3$.

1.4 Experimental realization

2D TIs were theoretically predicted and experimentally observed in CdTe/HgTe/CdTe quantum-well heterostructures. As shown in Figure 1.7, König et al., reported that if the thickness of HgTe layer exceeds the critical value 6.3 nm, then the bands invert due to quantum confinement effects (Hg is a heavy element), which led to the observation of nearly quantized $\frac{2e^2}{h}$ conductance originating helical 1D edge channels when the Fermi level is gated into the bulk gap. The observation of this quantized conductance is unique to 2D TI with 1D helical edge channels (Landauer Büttiker formalism) and is not expected to be seen in the transport measurement of a 3D TI due to existence of Dirac cones. Note that unlike QHS with broken time-reversal symmetry, there is no Hall conductance in a time-reversal symmetric 2D TI due to helical edge states.

Before proceeding with the experimental realization of 3D TIs, let us briefly review the angle-resolved photoemission spectroscopy (ARPES) technique which has played a central role in discovering and determining the surface electronic properties of various novel electronic structures, including high temperature superconductors, 2-dimensional materials, and of course topological materials. ARPES is a surface sensitive technique where typically X-rays with energy range of...
Figure 1.7 Experimental observation of 2D TI. a, A schematic of a 2D TI that is formed in a quantum well of HgTe of thickness $d$ sandwiched between two CdTe layers. b, When $d > 6.3$ nm, the bulk bands invert and the system becomes topological with spin-polarized edge modes. c, Transport measurements show a quantized conductance/resistance for samples thicker than 6.3 nm (samples II-IV with 7.3 nm thickness). The system is insulator for thickness below 6.3 nm (sample I with 5.3 nm thickness). Adapted from References 1 and 34.

ultraviolet (5 eV to 100 eV) or soft X-ray (100 eV to 1000 eV) is used to photo-excite electrons out of the material$^{35-36}$. The photo-excited electrons pass through a detector where they are sorted by their energy and angle, which can then be used to map out the energy-momentum dispersion of the material. For an incoming beam of photons with energy $h\omega$, based on the conservation of energy, we can write: $E_{kinetic} = h\omega - E_{Binding} - \phi$, where $E_{kinetic}$ is the kinetic energy of the photo-excited electron after leaving the material, $E_{Binding}$ is the binding energy, and $\phi$ is the work function of the electron which is material-dependent. Now based on the conservation of momentum, we can write: $h\kappa_{initial,\parallel} = h\kappa_{final,\parallel} = \sqrt{2mE_{kinetic} \sin \theta}$ and for the perpendicular direction $h\kappa_\perp = \sqrt{2m(E_{kinetic} \cos^2 \theta + V_0)}$, where $V_0$ depends on $\phi$, and can be found by looking
at the electrons emitted normal to the surface. By measuring $\theta$, the angle that electrons are emitted at, as well as their final kinetic energy, their initial energies and momentums in the crystal prior to photo-excitation can be traced back and the $E$-$k$ dispersion relation can be constructed. Furthermore, $k_\perp$, also referred to as $k_z$, determines the state’s dimensionality; for example, a 3D state should disperse with $k_z$, but a 2D state, due to confinement along the $z$-direction, does not change with $k_z$.

In 2008, the first 3D topological insulator was discovered in the Bi$_{1-x}$Sb$_x$ alloy system using ARPES and scanning tunneling spectroscopy (STS)\textsuperscript{38-42}. This TI had a small bulk band-gap and a complicated TSS band structure with five Dirac cones per surface (Figure 1.8a). The focus of TI research has since shifted to pnictogen chalcogenide compounds (Figure 1.9). The family of pnictogen chalcogenide compounds Bi$_2$Se$_3$, Bi$_2$Te$_3$, and Sb$_2$Te$_3$ were theoretically predicted\textsuperscript{13, 43} and subsequently observed by ARPES\textsuperscript{37, 44-47} (Figure 1.8b) and scanning tunneling spectroscopy/microscopy (STS/ STM)\textsuperscript{48-52} to be TIs. Bi, Sb, and Te are heavy elements and contribute to SOC in these TI materials. These heavy pnictide chalcogenides materials, in the form
Figure 1.9 Pnictogen (group V) chalcogen (group VI)-based TIs. **a**, Bi$_2$Se$_3$, Bi$_2$Te$_3$, and Sb$_2$Te$_3$ are topological insulators. Sb$_2$Se$_3$ is a trivial insulator. The pictures of elements are taken from http://periodictable.com. **b**, The rhombohedral unit cell of Bi$_2$Se$_3$ with the marked quintuple layer. **c**, The top view shows the hexagonal arrangement of atoms. **d**, The side view shows a quintuple layer, which is composed of five consecutive layers of Se-Bi-Se-Bi-Se. The bonding is strong within the QL and is van der Waals between the QLs. Bi$_2$Te$_3$, and Sb$_2$Te$_3$ have the same crystal structure. Adapted from Reference 13.

of A$_2$X$_3$ which are also known for being good thermo-electrics, are layered compounds with a rhombohedral structure. The layers are arranged along the c-axis direction in the 5 atomic layer-thick sequence X–A–X–A–X, which is called a quintuple layer (QL; 1 QL ≈ 1 nm; one unit-cell is characterized by three QLs) where Van der Waals forces bond adjoining QLs (Figure 1.9). These 3D TI materials possess simple TSS band structure with single Dirac cone centered at $\Gamma$ point of the Brillouin zone. Bi$_2$Se$_3$(lattice constant of 4.14Å) and Sb$_2$Te$_3$ (lattice constant of 4.25Å) have their Dirac point within the bulk gap, while the Dirac point in Bi$_2$Te$_3$ (with lattice constant of 4.38Å and band-gap of ~160 meV) lies beneath the top of bulk valence band, which makes it intrinsically difficult to probe physics near Dirac point in this material$^{44}$. Additionally, the TSS in Bi$_2$Te$_3$ show significant hexagonal warping. All these materials have sizeable bulk bandgap in the range of few hundred meVs, with Bi$_2$Se$_3$ having the largest gap of ~300 meV, making it a promising candidate for room-temperature applications$^{53-54}$. In the next chapters, the properties of this class of TIs will be explained in more detail.
1.5 Other topological materials

**Topological crystalline insulators (TCI):** TCIs are another type of topological materials, and unlike the aforementioned TI materials where TRS protects the odd number of Dirac cones, TCIs are protected by the crystalline symmetries such as reflection or point group rotation. In principle, crystalline symmetries may protect even number of Dirac cones on symmetry preserving faces in contrast to TRS topological insulators. For instance, the mirror symmetry in a cubic lattice stabilizes an even number of Dirac states on the faces normal to the mirror planes (Figure 1.10). TCIs were first reported in 2011, and then in 2012 it was predicted that the SnTe, PbTe, and PbSe class of materials are TCI. Shortly after, in late 2012, it was experimentally confirmed that for SnTe, Pb\textsubscript{1-x}Sn\textsubscript{x}Te, and Pb\textsubscript{1-x}Sn\textsubscript{x}Se\textsuperscript{59-60} electronic states evolve into a topological insulator upon decreasing temperature, consistent with theory.

![Figure 1.10 Topological crystalline insulator](image)

*Figure 1.10 Topological crystalline insulator.* The even number of Dirac cones (four) is shown for topological crystalline PbSe. Adapted from Reference 60.

**Topological Kondo insulator (TKI):** TKI is a strongly correlated insulator which emerges through the interplay of spin-orbit interactions and strong electron correlation. The story of TKI began almost 50 years ago with the observation of unusual behavior in the resistance versus temperature
Figure 1.11 The band structure and transport of SmB$_6$. a, The schematic of the band structure of SmB$_6$, which shows 2 weakly dispersing f-bands and a strongly dispersing d-band, as well as the TSS. Inset shows the bulk and surface Brillouin zone. Adapted from Reference 61. b, Resistance of a single crystal of SmB$_6$ as a function of thickness (normalized to its value at 20 K). The plateau grows as the sample is made thinner, suggesting that the plateau is originated from TSS. Adapted from Reference 62.

of SmB$_6$ where, upon cooling down, the resistance keeps increasing until a point that it starts to level off and plateau at low temperatures around 4 or 5 K, indicating a metallic state. The insulator part is described by Kondo insulator, the plateau however has remained puzzling for a long time.

If a magnetic impurity, say Fe, is placed inside a gold crystal. then the resistance vs. temperature of the gold features an upturn in low temperatures, unlike the pure gold itself which show a metallic temperature-dependence at low temperature. This is because the magnetic moment of Fe attracts a screening cloud of electrons around it, becoming a larger scattering center for the itinerant electrons. Now, in a crystal like SmB$_6$, every single Sm atom acts like a magnetic impurity and this forms a Kondo lattice where every Sm attracts a screening cloud which effectively opens up a gap (on the order of tens of meV) in the density of states at the Fermi level, and this eventually increases the resistance as the itinerant electrons hybridize with the f-electrons that are localized on the Sm$^{64-65}$. In 2010, it was theoretically shown that the strong correlations could indeed bring about a topological phase with a non-trivial Z$_2$ index$^{66}$ where eventually, as a result of bulk-boundary correspondence, metallic surface states emerge that short out the Kondo insulating bulk state at low temperatures, which shows up as a plateau in the resistance versus temperature. This motivated
experimentalist to study SmB$_6$ for signatures of the TSS $^{61-62,67-71}$ (Figure 1.11). Because the in-gap states of SmB$_6$ are quite flat, it is more challenging to resolve them as compared to the previously mentioned TIs (Figure 1.11).

**3D Dirac and Weyl semi-metal:** The bulk state of these semi-metals has a linear dispersion similar to the Dirac cone of a TI. The Dirac semi-metal is sometimes referred to as a 3D analogue of graphene$^{72}$ and can be thought to emerge right at the critical point of topological phase transition (TPT) from TI to trivial insulator$^{14,73-75}$ where the bulk gap closes and the conduction and the valence band touch prior to the band (de)inversion (TPT in (Bi$_{1-x}$In$_x$)$_2$Se$_3$ films will be discussed in depth in chapter 4).

![Figure 1.12 Experimental observation of 3D Dirac semi-metal by angle-resolved photoemission spectroscopy. a, ARPES dispersion map and Fermi surface of Dirac semi-metal Na$_3$Bi along with its Fermi arcs. Adapted from Reference 76. b, ARPES dispersion for 3D Dirac cone of Cd$_3$As$_2$ which clearly shows linear dispersion in all three $k$-directions Adapted from Reference 77.](image)

Observing such Dirac semimetal could be extremely challenging because experimentally it is almost impossible to precisely tune the chemical composition to reach the exact point where the band inversion occurs. However, Dirac semi-metals have been experimentally observed in certain
compounds where crystal symmetry protects and stabilizes the Dirac point. Na$_3$Bi and Cd$_3$As$_2$ are two examples of Dirac semi-metal where ARPES measurements confirmed the linear dispersion of $E$ with respect to $k_x$, $k_y$, and $k_z$ (Figure 1.12). Additionally, upon breaking the inversion symmetry or time-reversal symmetry, the 3D Dirac semi-metal converts to another type of 3D semi-metal known as Weyl semi-metal where each Dirac point of the 3D semi-metal splits into two Weyl points. The Weyl semi-metal has experimentally been observed in inversion symmetry breaking TaAs and NbAs (aka mono-arsenide) by ARPES. Another example of Weyl semi-metal as a result of time reversal-breaking is the superlattice of magnetic TI/ normal insulator with Weyl points along the growth direction.

1.6 Scope of present work

The focus of the present work is pnictogen chalcogenide TI thin-films. In chapter 2, after a brief overview about the basics of thin-film growth, we will discuss the MBE growth of this class of TI thin-films as well as the tools and techniques that can be used to characterize the quality of the grown films. Chapter 3 seeks to understand the role of defects in transport properties of these TI thin-films and introduces the effective growth methodologies to suppress the inadvertent defects and to reach true bulk-insulating TIs. In chapter 4, it will be discussed how the observation of the previously inaccessible topological quantum effects became possible in our defect- and interface-engineered TI films. Finally, chapter 5 concludes this thesis and provides a brief outlook on the future of TI research.
Chapter 2

Thin-film growth and characterization

2.1 Short History of thin-film growth

Condensed matter physics seeks to understand the properties of the crystalline materials. Crystalline materials can be categorized to 3-dimensional bulk crystals, 2-dimensional thin-films (confined in one dimension), 1-dimensional nanorods (confined in two dimensions), and 0-dimensional nanodots confined in all three dimensions. Here, we focus on thin-films.

Thin-film fabrication dates back to more than 4500 years ago where a thin layer of a malleable metal like gold was gilded on surfaces, statues and artifacts for decoration and protection purpose. One of the earliest practical examples of thin-film deposition is perhaps coating a glass with a thin layer of malleable metal to form a mirror. It was the invention of vacuum pumps and DC power which gave birth to more modern deposition techniques, such as sputtering in 1800s that was later evolved to more advanced magnetron sputtering systems in 1930s. Invention of the pulsed laser deposition (PLD) in 1960s followed by the molecular beam epitaxy (MBE) system by Arthur and Cho hit another milestone in thin-film deposition technique. Additionally, the development of more advanced vacuum technologies along with an availability of bake-able stainless-steel chambers as well as the use of electron diffraction, which was first shown by the 1937 Nobel laureates Davisson and Thomson by impinging electrons with ~100 eV of energy on Ni, resulted in substantial progress in thin-film growth. With these advancements, the epitaxial growth no longer remained limited to metal thin-films and it was further extended to epitaxial growth of high-quality semiconductors, such as GaAs by Cho and Arthur.
Generally, thin-film growth involves a template layer called a substrate upon which the thin layer of a material is grown, and its underlying crystal structure can be followed by the film even though the naturally occurring growth structure of the material is different. A major part of this dissertation sheds light on the effect of substrate on final properties of the grown topological insulator films. In this chapter, we will review the basics of thin-film growth and explain the growth of topological insulator thin-films using molecular beam epitaxy. We will also discuss the characterization techniques and tools that can be used to analyze the quality of the grown films.

### 2.2 Thin-film growth process

Thin-film growth can be thought as a chemical reaction between the substrate and adatoms (atoms/molecules adsorbed on the surface of the substrate), and thus it is a product of the complex interplay between thermodynamics, which determines the favorable equilibrium states, and kinetics, which determines how fast the reaction proceeds and what equilibrium configurations are allowed. Therefore, controlling and engineering these aspects play a leading role in achieving films with desired qualities. In principle, thin-film growth consists of three following steps:

1. **Generation of particulate material**, which can be created by thermally heating the source material (like in MBE), by impinging the target with a laser beam (like PLD) to eject particles out of the target material, by applying electric current (electroplating), or by simply introducing the gaseous precursor containing the constituents of the intended film (like in chemical vapor deposition). The particulates can be in form of atoms, molecules, or ions.

2. **Transportation of particles towards the substrate**, which depends on the growth technique. For example, in MBE and sputtering particle naturally diffuse toward the substrate due to vacuum. Specifically, in MBE the atomic/molecular particulate ballistically travel towards the substrate in the form of a directional beam because the chance of scattering is low in ultrahigh-vacuum. On the other hand, in sputtering the particulate travel in rather a
diffusive way toward the substrate and experience multiple scattering along the path. In electroplating, an electrolyte is used to transport the ions to the substrate. In chemical vapor deposition (CVD), the gaseous precursors carry the films constituents.

3. Condensation on the substrate to form a thin-film, which occurs in a complex manner and involves chemical reaction processes at the surface of the substrate that is briefly explained in the following section.

2.2.1 Physisorption and chemisorption

Once the particulates (adatoms) make their way to the surface of the substrate, under optimized circumstances, the film can be deposited. In principle, two mechanisms play a huge role in deposition of a thin-film: physical adsorption or in short physisorption and chemical adsorption or in short chemisorption (Figure 2.1).

At the surface of the substrate, van der Waals interaction creates a weak potential-well (usually modelled by Lennard-Jones formula). If the kinetic energy of the arriving particulates is low enough, then they weakly get adsorbed to the surface of the substrate by this van der Waals force. This process is called physisorption. However, if their energy is high they immediately get desorbed off the surface. Although the binding energy of physisorption process is small and is in the range of 10 to 200 meV, it plays a crucial role in thin-film deposition process. Basically, when the adatoms physically adsorb on the surface, they exchange energy with the surface to lower their energy. Their energy is inelastically transferred to the substrate through a mechanism called thermalization where they get thermally adapted to the substrate temperature. If the substrate temperature is too high the adatoms simply desorb off the substrate, however if the substrate’s temperature is low enough they remained adsorbed to the substrate. If the substrate is at the right temperature, then it can provide enough activation energy so that the physisorption gives way to chemisorption, which is a stronger adsorption with energy on the order of 1 to 10 eV. The right temperature gives the adsorbed adatoms sufficient mobility to move around and find the optimal
sites with strong confinement potential, such as a step edges and get chemically adsorb by reacting with the substrate and/or each other. As the number of incoming adatoms increases they cluster at nucleation centers. The film then starts to grow and cover the surface from these nucleation points. For more information regarding physical and chemical adsorptions we refer to Reference 92.

![Figure 2.1 Diagram of physisorption and chemisorption process.](image)

**Figure 2.1 Diagram of physisorption and chemisorption process.** If the adatoms have a high energy they stick on the surface of the substrate and can move around on the substrate with an optimal temperature due to thermal energy. If the kinetic energy of adatoms or the substrate’s temperature is too high, then the atoms will immediately desorb off the surface. If the substrate’s temperature is high enough it will provide sufficient activation energy so that the physisorption gives way to chemisorption and the adatoms chemically get adsorbed to optimal sites with deep potential.

### 2.2.2 Different arrangements and growth modes

Depending on various growth parameters, the arrangement of atoms on the substrate can occur in three different forms which are shown in Figure 2.2: 1. single crystal (monocrystalline), where the crystal structure is the same and continuous across the entire substrate (long-range ordering) and is made of a single domain with no grain boundaries, 2. Polycrystalline, where the film is made up of many crystalline domains (usually with the size of ~10 to 100 nm) that are randomly oriented with
Figure 2.2 The possible arrangements of adsorbed atoms on a substrate. a, Single crystal with long-range ordering crystal structure, where a single domain is continuously grown across the substrate. b, Polycrystalline with small domains randomly oriented with respect to each other. c, Amorphous with no ordering, where randomness is extended down to atomic scale.

respect to each other, and 3. Amorphous, where the film has no order and the randomness extends down to atomic scale with grain size on the same order as the unit cell. In principle, single crystalline phase forms under optimized growth conditions, where the deposition rate and substrate’s temperature are all optimized. The substrate’s temperature should be high enough to maximize the surface mobility but also low enough to prevent desorption. The growth rate also should be slow enough to give the adatoms on the substrate’s surface sufficient time to find their optimal crystal location before the next layer deposits. Additionally, the structural and chemical compatibility of the substrate and the intended film is crucial. On the other hand, polycrystalline and amorphous films are usually formed when the substrate’s temperature is too low, and/or the deposition rate is too fast so the adatoms do not have enough mobility on the substrate’s surface to find their optimal crystal positions. Non-compatible substrate can also lead to polycrystalline films, like the growth of Bi₂Se₃ with 6-fold symmetry on a 4-fold symmetric (001) SrTiO₃ substrate. In₂Se₃ growth on Al₂O₃ is another example of polycrystalline growth, where In₂Se₃ forms a polymorphic film with in-plane randomness. The attempt in this work is to find optimal growth conditions as well as compatible substrates for the purpose of achieving high-quality single crystal bulk-insulating topological insulators thin-films.
The process of growth of single crystal film on another crystalline layer (substrate) refers to as epitaxy which was first introduced by Royer\textsuperscript{93} and is composed of two Greek word “epi” meaning upon or akin and “taxi” meaning order or arrangement. The epitaxial growth of a thin-film on a substrate with identical composition is known as homoeptaxy, such as growing GaAs thin-film on a GaAs substrate. In contrast, if the substrate is different from the growing film, then the epitaxy is called heteroepitaxy\textsuperscript{93-94} like growing Bi\textsubscript{2}Se\textsubscript{3} on Al\textsubscript{2}O\textsubscript{3}\textsuperscript{17}.

Furthermore, the epitaxial growth can proceed in different modes which strongly depends on the compatibility of the film and the substrate as well as the growth conditions and is mostly determined by the competitions between various surface energies to minimize the surface energy of the system. These modes include: 1. The Volmer-Weber growth mode, where the adsorbed atoms aggregate around nucleation centers and form 3-dimensional islands; these islands will eventually merge together to fully cover the surface of the substrate, 2. The Frank-van der Merwe growth mode or the layer-by-layer growth mode, where the adsorbed atoms complete a uniform single layer prior to starting the next layer, and finally 3. The in-between mode of Stranski-Krastanov growth mode which is a combination of the two previous growth modes where the uniform coverage is achieved but with some amount of roughness.

Let us define the surface tension (i.e. Gibbs free energy) between the substrate and vapor thin-film and vapor and the substrate and the film (aka interfacial energy) as $\gamma_{sv}$, $\gamma_{fv}$, and $\gamma_{sf}$, respectively, with $\theta$ being the angle between the nucleus and the surface (Figure 2.3). Then, based on Young’s equation, we can write $\gamma_{sv} = \gamma_{sf} + \gamma_{fv} \cos \theta$. If $\theta \approx 0$ (complete wetting occurs) then $\gamma_{sv} \geq \gamma_{sf} + \gamma_{fv}$ which means that layer-by-layer or Frank-van der Merwe growth mode is desirable. In other words, the substrate is not willing to have any exposed surface and the film completely spreads out on the substrate in form of a uniform layer. On the other hand, if $\theta > 0$ (i.e. wetting doesn’t occur) then $\gamma_{sv} < \gamma_{sf} + \gamma_{fv}$ and as a result the island growth or Volemr-Weber growth is favored (this resembles oil droplets forming on the surface of water).
Figure 2.3 A schematic of possible epitaxial growth modes. **a,** The competition between various surface energies including $\gamma_{sv}$ (surface tension between the substrate and vapor), $\gamma_{fv}$ (surface tension between the thin-film and vapor), and $\gamma_{sf}$ (surface tension between the substrate and the film) determines the growth mode. **b,** The Volmer-Weber island growth, where atoms form clumps on the surface. **c,** The Frank-Merwe growth mode, where the growth proceeds in a layer-by-layer manner and each layer is fully formed before the next one starts. **d,** The Stranski Krastanov growth mode, which is a combination of island and layer-by-layer growth and is closer to what happens in practice.

In the layer-by-layer + island growth mode, the growth usually starts with layer-by-layer growth mode but then changes to island growth mode due to competition between $\gamma_{sf}$ (to be more precise now being between the film and film $\gamma_{ff}$) and $\gamma_{fv}$ which is the result of strain relaxation effect. This is what generally happens in heteroepitaxial growth, especially after a critical thickness. The strain in heteroepitaxial growth is caused by lattice mismatch between the substrate and the film.

The level of strain is quantified as $\frac{a_f-a_s}{a_s}$, where $a_f$ is the in-plane lattice constant of the film, and $a_s$ is the substrate’s lattice constant. If the lattice mismatch is small (usually below 3% is considered small), then the strain would be small, and the film possibly grows in layer-by-layer mode. If, however, the lattice mismatch is too large, then the strain energy will dominate over the surface
energy required for the layer-by-layer growth, and as a result the growth would be in form of island growth. Finally, in some cases, film is initially strained to match the lattice spacing and the growth proceeds in layer-by-layer fashion up to a critical thickness after which it becomes energetically too costly for the system to maintain the strain, thus the strain is relaxed by typically forming defects (generally edge dislocations).

### 2.2.3 Defects in crystals

In reality, no material is immune to defects and disorders. Defects play an important role in determining the physical, electrical, optical, and magnetic properties of materials. These properties can be modified or improved sometimes by inducing and sometimes by suppressing/minimizing the defects in the material. Generally, defects come in three categories: point defects, line defects, and planar defects. Point defects are localized (0-dimensional) and usually occur around a single point as a result of the absence of an atom from its lattice position (known as vacancy), a foreign atom substituting a native atom (known as substitution), or a native or foreign atom sitting interstitially in between the lattice sites. During the thin-film growth, such defects are created either thermodynamically or due to the presence of impurity species. Different point defects are shown in Figure 2.4a.

Line defects (Figure 2.4b) are 1-dimensional defects and occur along a line (along a crystallographic direction of a crystal) to alleviate the strain. It could happen when an extra row of atoms is inserted into the lattice, forming extra half-plane of atoms in a lattice, known as edge dislocation. Another example of line defect is when a screw dislocation happens, usually as a result of shear stress where the crystal plane moves perpendicular to the direction of applied stress and grows helically along a linear defect – looks like crystal grows around and on itself. Figure 2.4c shows a screw dislocation in Bi$_2$Se$_3$ captured by scanning tunneling microscopy.
In fact, it is because of dislocations that deformation occurs at much lower stress than in a perfect crystal. 2-dimensional planar defects (Figure 2.4d) occur when the stacking order of a crystal is interrupted. For example, when a planar shift interrupts the normal lowest energy stacking order of …ABCABCABC… (close-packed structure), leading to a stacking fault with formation of…ABCABCBABC…. Stacking faults can be seen in transmission electron microscopy (TEM) images of some of our topological insulator thin-films/heterostructure. Another example of planar defects is a twin boundary, where a plane of mirror symmetry is introduced in the ordering of crystal structure. There are other examples of defects as well, such as bulk defects which are 3-dimensional macroscopic defects like cracks, voids (clusters of vacancies), and precipitates (clustering of impurities), resulting in a locally different phase as compared to the rest of the crystal.
2.3 Different thin-film growth techniques

There exist several techniques to deposit thin-films and generally are categorized as either chemical or physical deposition techniques. In physical deposition technique or physical vapor deposition (PVD), the material that is going to be deposited is physically ejected from a source through thermal evaporation or by using laser bombardment. Example of this technique include molecular beam epitaxy (MBE), sputtering, pulsed laser deposition (PLD), thermal evaporation method, electrohydrodynamic deposition, and cathodic arc deposition. Chemical deposition technique employs a carrier gas that contains the constituents of the intended films. Examples of this technique include chemical vapor deposition (CVD), atomic layer deposition (ALD), and sol-gel method. Below, some of these techniques are briefly reviewed with specific focus on MBE which is the system used in this work for the growth of TI thin-films.

2.3.1 Chemical vapor deposition

Chemical vapor deposition (CVD) and atomic layer deposition (ALD) are some of the widely used chemical deposition technique. This technique uses precursor gases which carry the constituents of the film to be deposited. Through a chemical reaction which takes place in the vicinity and on the surface of a (usually heated) substrate, the desired elements are freed and deposited as a film on the substrate with some gaseous byproducts remaining which are then pumped away. There are several types of CVD process, including atmospheric pressure chemical vapor deposition, metal-organic chemical vapor deposition (MOCVD), low pressure (mTorr range) chemical vapor deposition (LPCVD), ultra-high vacuum chemical vapor deposition (UHV-CVD), plasma-enhanced chemical vapor deposition (PECVD) for enhancing the reactivity of the precursors. ALD is special type of CVD, where the gas precursors are introduced in an alternating fashion, which makes this technique highly controllable when it comes to growing ultra-thin-films where a precise thickness is desired. ALD is extensively used for the growth of Al₂O₃ and HfO₂. Due to fast growth
rate and high-quality large area growth, CVD and ALD are heavily used in industry for mass production of Si for microelectronic devices (where silane gas, SiH$_4$, is used), high-$\kappa$ dielectric materials, some semiconductors like GaAs to use in light emitting diodes, as well as metal thin-films. CVD can be used for the growth of graphene$^{95}$ and even topological insulators$^{96}$.

2.3.2 Sputtering

Sputtering is a fast and inexpensive form of PVD method that can be used to deposit single crystal, polycrystalline, as well as amorphous thin-films$^{97,98}$. In sputtering, the substrate and the target material are placed in a chamber which is initially evacuated with a turbo pump to remove the unwanted molecules. By bombarding the target material by ionized atoms (in form of a plasma), the particulates are ejected out of the target. The plasma is usually created from a noble gas, like argon, which is introduced in the evacuated chamber in small amount. A negative potential is applied to the target (cathode) and the anode is the chamber body (or substrate) which is grounded. This causes free electrons to accelerate away from the target and collide with Ar atoms, thereby stripping the gas atoms of an electron and thus positively ionizing them. This initially starts with the ionization of a few argon ions; as they accelerate they collide with other argon atoms, which in turn are ionized and also begin to accelerate. This is known as cascade effect. The positively charged Ar$^+$ ions are accelerated toward the target carrying energy on the order of $\sim$1 keV and eventually sputter/knock off some of the target material. The energy of the ejected particulates is relatively high in the range of 1-10 eV which then travel toward and impinge onto the substrate and eventually deposit as a thin-film.

The glow from the plasma is from the ions recombining with free electrons. Another form of sputtering is reactive sputtering in which the plasma is composed of a noble gas mixed with a reactive gas, such as oxygen or nitrogen that can react with the film and alter its chemical composition. Since it is feasible to maintain the plasma/continuous sputtering for metallic targets
Figure 2.5 A schematic of magnetron sputtering system. Argon gas is introduced in the chamber and gets ionized. The argon ions impinge on the target materials and eject particulates out of the target. The particles then fly towards the substrate and deposit as a film.

under a direct current bias and for insulating targets under an alternating current (typically in the radio frequency range), nearly any material can be ablated by sputtering method which makes sputtering a unique system compared to other PVD techniques. For both DC and RF sputtering, the deposition rate can be controlled by varying the applied power to the target. This way low deposition rates and thus epitaxial growth is achievable. Modern sputtering systems (aka magnetron sputtering systems) use a magnetic field, typically created by strong neodymium magnets, to confine charged plasma near the target. These systems have higher ionization yield and substantially (almost by an order of magnitude compared to conventional sputtering systems) minimizes the required Ar pressure for creating the plasma (usually around 1 to 3 mTorr). This lower working pressure prevents the particulates en route to the substrate from being deflected due to colliding/scattering with the background gas. The relatively high pressure of sputtering chamber leads to thin-films having significant impurities compared to other PVD systems operating at much
lower pressure. Additionally, the high energy of impinging particulates on the substrate could itself knock off, or in other words re-sputter, the already deposited thin-film and eventually impair the stoichiometry and/or the quality of the film. The re-sputtering process can be minimized if the targets are placed at nearly 90° relative to the surface of the substrate.

2.3.3 Pulsed laser deposition

Similar to sputtering, PLD\textsuperscript{99-100} is a PVD method which uses series of intense laser pulses, usually in ultra-violet to the visible frequency range, to bombard and evaporate the target material. The laser heats and ablates the target which eventually ejects the material in form of a plasma plume, sending the ejected particulates of the target material towards the substrate where, under right conditions, epitaxial growth of thin-film takes place. The repetition rate of the laser is in the range of 10 to 100 Hz, and each pulse is around $10^{-12}$ to $10^{-9}$ seconds long followed by ~0.1 to 10 millisecond breaks. The delay in repeated pulses is crucial to achieve the optimal growth. Due to intense laser heating and because PLD allows adjusting the laser’s intensity and wavelength to maximize the optical absorption, nearly any material can be ablated in PLD. Like sputtering, PLD is particularly useful for depositing multi-element compounds because it allows the use of stoichiometric targets that ensures stoichiometric films, circumventing the challenging task of manually adjusting the stoichiometry of the constituents which is usually the case in other techniques like MBE.

Unlike sputtering, PLD can be operated without or with a background gas which is particularly useful for complex oxide films\textsuperscript{101}. The lower working pressure in PLD allows real-time and \textit{in situ} monitoring by reflection high energy electron diffraction (RHEED) tool, which helps achieve optimal growth conditions. One of the drawbacks of PLD is re-sputtering and inter-layer mixing during the growth due to high kinetic energy of the particulates (in the range of 1 to 10 eV), which eventually leads to non-stoichiometric films and broadened interfaces in heterostructures.
2.3.4 Molecular Beam epitaxy

Molecular beam epitaxy (MBE) is another and perhaps the most controllable example of PVD technique and uses sources (containing elements) that are heated to thermally evaporate/sublimate the elements, which eventually epitaxially react and grow as a thin-film on a substrate. MBE was first developed in the 1960’s and 70’s for the growth of III-V semiconductor compounds. The GaAs growth, which was first attempted by Cho and Arthur using an ultrahigh vacuum MBE system equipped with RHEED, is one of the most prominent and successful examples of semiconductor growth by MBE that showcases the power of epitaxial growth by MBE where the
Figure 2.7 A schematic of an MBE system. The effusion cells contain different elements in a crucible. The crucible is surrounded by a heater which heats the element and generates atomic/molecular beam that travels towards the heated substrate and deposits as a film on the surface of the substrate. The beams are quite directional because of ultrahigh vacuum. The pneumatic shutter is open when an element is in use (for example target A and B shutters) and closed when an element is not being used (for example: shutter of target C).

Mobility of the grown GaAs samples continuously increased over 2 decades from $\sim 10^4$ to $\sim 3 \times 10^7$ cm$^2$V$^{-1}$s$^{-1}$ (highest ever mobility in any material) via defect and interface engineering. Unlike sputtering and PLD, MBE allows a controllable thermal evaporation of elements by adjusting the temperature of each elemental source, which in turn enables atomic-scale precision of thickness, deposition rate and composition of growing films. MBE is extremely efficient for the growth of heterostructures and superlattices where sharp interfaces with precise thickness and doping concentration in each layer are required. Additionally, unlike PLD and sputtering, the lower kinetic energy ($\sim 50$ meV) of the thermally evaporated particulates used in MBE avoids the re-sputtering
process or introduction of unwanted defects. Furthermore, the composition can be adjusted manually by individually evaporating the elements, unlike PLD and sputtering where the composition is mostly determined by the target. MBE performs under ultrahigh vacuum and the pressure of growth chamber in MBE is usually in the range of \( \sim 10^{-10} \) Torr. This ultrahigh vacuum minimizes the probability of particles getting scattered which produces directional molecular/atomic beam of elements with large mean free path – which is estimated by \( l \) (mean free path in cm) \( \approx 5 \times 10^{-3}/P \) (\( P \) is chamber’s pressure) to be on the order of tens of km for \( 10^{-9} \) Torr pressure\(^\text{102}\). The UHV also substantially minimizes the incorporation of any unwanted impurities during the growth process. Furthermore, the ultrahigh vacuum and the low energy of particulates yield reliable, stable and reproducible flux. Combined with RHEED, it allows the real time monitoring of the growing film which helps efficiently optimize the growth parameters in a rather short time.

Almost all elements can be used in MBE by thermal evaporation or with the use of electron beam evaporators for elements that are not easy to be thermally evaporated like Pt, Rh, B, C, etc., and refractory elements including Zr, Hf, Nb, Ta, Mo, W, Re, Ru, Ir, Os, etc. Gaseous elements like oxygen can simply be introduced in the MBE chamber. For the growth of oxide films where more energetic oxygen is required for full oxidation, plasma oxygen or ozone can be used (both can be installed on the MBE chamber). Due to the production of higher energy particulates, the flux generated by an e-beam is not as stable as the one generated by a thermal cell.

### 2.4 MBE system at Rutgers University

The MBE system in Prof. Seongshik Oh’s lab at Rutgers University was designed and built by SVT Associates. Although this system was initially designed for oxide films, it eventually got dedicated to the growth of chalcogenide-based topological insulator films and was used to grow all the TI films that are explained in this dissertation. Because it is heavily contaminated with Se and Te, this
system is not suitable for oxide films growth. Thus, recently, another MBE system was installed for the growth of complex oxide or other interesting non-oxide chalcogenide-free films. The new system is connected to the first system through a tunnel-like connection, that is kept under vacuum, which allows transferring samples from one system to the other, providing a unique opportunity for fabricating hybrid oxide/TI heterostructures and exploring their properties (Figure 2.8). The MBE chambers are made with 300 series stainless steel sheets with ~0.1-inch thickness and are welded together (on the vacuum side) or onto ConFlat (CF) flanges for attaching other vacuum components.

The MBE system is composed of three individual vacuum chambers separated from each other by pneumatic or manual UHV gate-valves: 1. The load-lock chamber for loading substrates or removing them from the MBE system, and thus it is the only chamber that is regularly vented by nitrogen gas and opened to atmosphere. 2. Isolated by a gate-valve and sitting vertically below the load-lock chamber is the preparation (in short prep) chamber (with a built-in sputtering system) where the substrates are transferred to the sputtering stage or into the growth chamber using a transfer arm. 3. Connected to the side of the prep chamber is the growth chamber with a growth stage, where the thin-film growth takes place. Both the prep and growth chambers are pumped by cryo-pumps that trap the gaseous particles which condense on a cold head. Helium compressors are used to compress and cool helium gas, which is then sent to the MBE, where it is used to cool the cryo-pumps to ~15-20 K.

The typical base pressures of the growth and the prep chambers are ~ $10^{-10}$ Torr and ~ $10^{-9}$ Torr, respectively; the pressure of the prep chamber is higher due to the exposure to the higher pressure of the load-lock during substrate transfer. The pressure is measured by anion gauge which ionizes the residual gases and measures the ion current between the
Figure 2.8 Schematic of chalcogenide and oxide-based molecular beam epitaxy systems in Rutgers University. These two systems are connected by a tunnel-like connection.

anode and cathode of the gauge. The ion current is proportional to the pressure of residual gases which then with a proper factory-set calibration the chamber pressure can be found.

Now, let us explain how we use these units to start a growth in this MBE system. Let us begin with the step prior to loading a new substrate. At this point, the whole MBE system is under vacuum and the gate-valves are all in closed mode. The substrate cassette is in the load-lock chamber which is being pumped by a turbo pump and is under $10^{-6}$ Torr vacuum. Next, after preparing a new substrate (explained in later sections), we need to vent the load-lock chamber in order to load the new substrate. To do so, we first close the UHV gate-valve between the load-lock and the turbo pump and then open valve to the high-purity $N_2$ gas tank, which fills up the load-lock. When the load-lock is fully vented, the door can be opened, and the new substrate can be inserted into the cassette. For any short delay during this stage, the $N_2$ gas tank should be kept open to maintain a constant positive pressure $N_2$ flow in the load-lock chamber to minimize the flow of atmospheric species into the system. Once the new substrate is put into the cassette, the cassette will be put back
in the load-lock chamber where it sits on a retractable holder. Next, after closing the N\textsubscript{2} gas, the load-lock chamber is first pumped down to \(\sim 10^{-2}\) Torr by an oil-free scroll pump and then down to \(10^{-6}\) Torr by a turbo pump that itself is back-pumped by a diaphragm pump (before opening the turbo pump gate-valve, we close the scroll pump’s valve). After waiting for a while and making sure the load-lock is properly evacuated, we open the UHV gate-valve between the load-lock and the prep chamber and manually raise the elevator from the prep to the load-lock until the cassette is pushed up above the retractable holder, which is then retracted, and the cassette can be transferred down into the prep chamber, after which the UHV gate-valve gets closed. As mentioned before, the pressure in the prep chamber is \(\sim 10^{-9}\) Torr, however during the cassette transfer it can jump by two orders of magnitude due to exposure to the lower vacuum of load-lock chamber. Next, the substrate is removed from the cassette using the transfer arm, and after that the cassette is lowered below the transfer arm.

Prior to opening the UHV gate-valve between the prep and the growth chambers, it is better to wait for some time so that the base pressure of the prep chamber is recovered. Once the base pressure of the prep chamber is attained, the gate-valve that separates the growth and the prep chambers is opened and the substrate is then transferred from the prep to the growth chamber where growth stage is lowered (the C-shape sample holder) to accept the plate with the mounted new substrate. After properly placing the new substrate in the growth stage and making sure that it is sitting flat on the C-shape holder of the stage, the stage is raised, the arm is slowly retracted, and the prep-growth gate-valve is closed. The substrate is now ready for the growth. A reverse process is used when taking a grown sample out of the MBE system. For convenience, the chambers are equipped with several sealed glass windows for sight into the chambers.

Most growth processes are controlled remotely by a computer, which not only makes the operation easier but also ensures consistency during the multi-stepped growth process, especially during the growth of multi-element films where multiple shutters need to be opened together at the
Figure 2.9 MBE system in Oh’s lab at Rutgers. Major MBE parts are marked: 1. The load-lock, 2. The prep/sputtering chamber, 3. The growth chamber, 4. A part of the transfer arm is visible, 5. Turbo pneumatic gate-valve which separates the turbo and the load-lock during the vent process; the turbo is almost always kept on, 6. The turbo pump which itself is backed by a diaphragm pump, 7. The manual gate-valve between the load-lock and the prep chambers, 8. The manual gate-valve between the prep and the growth chambers, 9. The rotatable growth stage, 10. One of the effusion cells, 11. The RHEED system, 12. The QCM system, 13. The growth cryopump; there is another cryopump for the prep chamber, 14. One of the shutters; for opening and closing purpose, shutters are attached to compressed air (the blue and grey tubes attached to the shutter).

same time. Furthermore, we use highly purified elements that are available commercially with purities > 99.999%. The elements are loaded into crucibles and placed in effusion cells. Typically, the crucibles are made of Al₂O₃, pyrolytic boron nitride (PBN), graphite, vitreous carbon, or tantalum. Prior to loading the elements, the empty crucibles are cleaned by out gassing to high temperatures in vacuum to remove any possible contamination. The source material is then loaded into the crucible, where the amount depends on the crucible size and level/frequency of use. The compatibility of the element and the crucible must be checked beforehand to avoid chemical reaction during heating. To protect the crucibles against cracking/deformation during melting and solidification process, a very slow ramp rate is used near the melting point.
For the MBE system, we have several low temperature cells ($< 1000^\circ C$) which can be used for elements like Bi, Sb, In and Se, Te, some medium temperature cells ($< 1400^\circ C$) that can be used for elements like Cu and Sn, as well as high temperature cells ($< 1800^\circ C$) suitable for elements like Fe, Ti, V, and Cr, etc. Additionally, a mixture of glycol and deionized water is used as coolant, which is less corrosive than pure deionized water. The coolant is circulated through a home-built 150,000 British thermal unit heat exchanger that thermally connects the coolant lines to the Rutgers physics building coolant system at $\sim 10^\circ C$. Figure 2.9 shows all the major parts of Rutgers MBE system.

### 2.5 Atomic flux measurement techniques

Accurate measurement of atomic flux, the number of atoms arriving at the surface of the substrate in a given amount of time, holds a great importance in achieving precise atomic-scale thin-films and is crucial for determining the deposition rate as well as the composition of the growing films. There are *in situ* and *ex situ* techniques available to precisely measure the atomic flux for each individual element. The major *in situ* measurement techniques are a quartz crystal microbalance (QCM), atomic absorption, and ion gauges. Among all, QCM is the only method which directly measures the flux. However, it cannot measure the flux in real time during the growth since the substrate has to be moved and the QCM head has to be inserted at the substrate position. Atomic absorption functions based on the change in intensity of a light beam that it is passing through a molecular beam. It needs to be calibrated to measure the accurate flux at the surface of the sample, however it can measure the flux in real time. Lastly, ion gauge measures the vapor pressure at the substrate location, but it requires calibration against another method and the substrate has to be moved. The most accurate method to measure the composition and thus the flux *ex situ* is Rutherford backscattering spectroscopy (RBS), which measures the energy of the backscattered ions from a high energy ion beam impinging on the sample. In this work, we used QCM in conjunction with RBS to measure the flux of each element (together provide $\lesssim 1\%$ accuracy).
2.5.1 Quartz crystal microbalance

QCM is an accurate and efficient tool for measuring tiny masses in situ and is connected to the MBE chamber using a retractable bellow system which can be inserted into the growth chamber and placed at the substrate location after raising the growth stage. The QCM measures the change in the resonance frequency of a thin quartz crystal as the molecular beam impinges on its surface. Quartz is a piezoelectric material, where an application of mechanical stress on the material generates electric field or vice versa. QCM is a wand-like tool with a head that holds a disc-like quartz crystal sandwiched between gold electrodes. By applying an alternating electric voltage across the crystal using the gold electrodes, the crystal can oscillate at its resonance frequency. As material is deposited on the quartz crystal, the frequency decreases with accrual of deposited mass on the crystal.

For a quartz crystal with the thickness $t_q$, and shear wave velocity $v_s = \sqrt{\mu / \rho_q}$ (where $\mu$ is the shear stiffness and $\rho_q$ is the quartz density), the resonance frequency is $f = \frac{v_s}{2 t_q}$. If a thin-film is deposited on the surface of the crystal, then the thickness increases to $t_q + \Delta t$, where $\Delta t = \frac{\Delta m}{A \rho}$ with $\Delta m$ being the added mass, $A$ the film area, and $\rho$ the film mass density. For small thickness limit, the thickness $\Delta t$ of a film can be replaced by the equivalent thickness $\Delta t_q$ of quartz; then the change in frequency would be $\Delta f = \frac{v_s}{2(t_q + \Delta t)} - \frac{v_s}{2t_q} \approx \frac{v_s}{2(t_q + \Delta t_q)} - \frac{v_s}{2t_q}$; which gives the Sauerbrey equation\(^{105}\) for a tiny change in thickness, $\Delta f \approx - \frac{C \Delta q v_q}{2t_q} = - \frac{f^2 \Delta m}{c \rho A}$, where $C = v_q / 2$ is the resonance frequency constant. By knowing material properties as well as the geometry of the quartz crystal and by measuring the change in the resonance frequency, the accumulated mass over a given amount of time can be obtained, which in turn can be transformed to a flux by converting the mass density to atomic number density. QCM is a sensitive method which is capable of
measuring changes as small as $\sim 1$ Hz in frequency (the resonance frequency is in MHz regime), and thus the flux can be accurately measured using this technique.

For practical QCM measurements, we use a commercially-available programmed control unit. By simply inputting the material and its mass density, as well as the $z$-ratio, which is provided for each element and is a measure of the film’s density and shear modulus relative to those of the quartz (the value is usually close to 1), the flux at a given temperature can be measured. However, in practice the QCM measurement is more complex because it is susceptible to temperature changes of the crystal during the deposition as a result of radiation from the heated cells. This non-controllable heating process could potentially shift the resonance frequency and appear as an unreal accumulation of mass on the quartz crystal. To minimize this effect the QCM system is continuously cooled by water. To account for any uncertainty/artifact, we can compare the QCM results with the RBS ones and attain a calibration factor. For this, we start with heating the desired elements to target temperatures, after waiting some time for heated elements to stabilize, we raise the growth stage, insert the QCM head into the chamber and place it at the substrate’s position (Figure 2.10). Next, we open the source shutter and continuously measure the flux of each element using the QCM program for at least 5 minutes (usually the flux is recorded at minutes 5, 10, 15, and 20). We then grow a test film with the thickness of less than 30 nm for RBS measurement; this thickness regime gives the best quantitative result for RBS\textsuperscript{92}. The QCM flux values is then compared to the true values obtained by RBS; the ratio of these two number can be used as calibration factor. For example, for Bi, the calibration factor between QCM and RBS is $\sim 84\%$ (QCM flux recorded at 5 minutes after opening the shutter), meaning that the true values of the Bi flux is $\sim 84\%$ that of the value measured by QCM. The QCM reading is usually accurate within $\sim 1\%$, and thus averaging many measurements should yield a flux within about $\lesssim 1\%$. 
Figure 2.10 QCM system on the MBE in Prof. Oh’s Lab. a, The QCM head with the mounted quartz crystal which is cooled by water. Electrical and cooling lines are marked. Pictures (b) to (d) show the procedure of inserting the QCM for the flux measurement. b, The QCM is retracted whenever it is not in use. The QCM space is separated by a gate-valve from the growth chamber. c, The growth stage has to be raised from its growth position before inserting the QCM. d, The stage is raised and has to remain raised during the QCM measurement. e, After opening the gate-valve, the QCM can now be inserted into the growth chamber and placed at the substrate’s position.

2.5.2 Rutherford Back scattering

The basic principles of Rutherford backscattering spectroscopy (RBS) lies on the famous experiment by Ernest Rutherford where a few of the alpha particles (He\(^{2+}\)) impinging on a thin sheet of gold foil was deflected, which made him conclude that a tiny dense nuclear core was causing the deflections. The Laboratory for Surface Modification at Rutgers University owns a 1.7 megavolt Tandetron accelerator that is used for RBS measurements which is a tandem accelerator that accelerates the alpha particles in two steps. The source of helium gas (alpha particles) and the target are separated from each other by a positively charged high voltage terminal. By passing through an ionizer, the He gas gets negatively ionized by an electron addition, and thus accelerates
Figure 2.11 Rutherford backscattering spectroscopy system located at the Laboratory for Surface Modification at Rutgers University. a, A picture of the RBS system with all the major parts marked by an arrow. b, An image showing inside the scattering chamber. The schematic below (b) shows the incoming ions impinging on the sample and the backscattered beam travelling towards the detector at an angle $\theta$. Adapted from www.physics.rutgers.edu/~bart/627/P627_S13_RBS_Lab_21Feb2013.pdf.

toward the positive terminal. Next, the helium ions pass through an electron stripper located near the positive terminal and the net charge of ions changes from 1$^-$ to 1$^+$ or 2$^+$, which makes them further accelerate toward the target at ground potential. The ions are then impinged on the film surface (normal to the surface), and the backscattered ions are measured at $\sim 163^\circ$ (17$^\circ$) relative to the incoming beam using a silicon detector (Figure 2.11). RBS is an efficient tool to probe the chemical composition of the films. Additionally, the ions continuously lose energy upon penetrating into (the penetration depth can be more than 1 $\mu$m) and emerging out of the film. This continuous loss of energy allows RBS to probe the composition of the film with depth resolution. For this work, we used RBS as a calibration of elemental flux by determining stoichiometry and composition of grown TI films, with thicknesses $\lesssim 30$ nm where the surface-layer approximation is believed to be valid, and the spectra peaks are sharp and easier to fit. We use SIMNRA to analyze
Figure 2.12 RBS experiment for a (Bi$_{1-x}$In$_x$)$_2$Se$_3$ film. a, The spectrum of the actual data (red) along with the simulated curve (blue) generated by SIMNRA software. b, The window that contains the parameters related to the RBS experiment. c, The window that contains the inputs about the thickness and composition of the substrate as well as the other layers, which can be used to generate the simulated fit for this sample. By continuously comparing the area under the actual data minus the background with the area under the simulated curve for each element, the desired fit is obtained. d, The typical calculation steps that can be used to determine flux, deposition rate, thickness and composition. e, the hexagonal (trigonal) lattice structure with its primitive cell marked in red.

The backscattering spectra, which generates simulated curves for comparison with the actual spectra. To determine a film’s parameters using SIMNRA, some initial inputs/guesses for the elemental composition within each layer are needed. Then based on this information, SIMNRA simulates the spectra. The numerical integration function can be used to find the area under the spectra. We calculate the total background value, which is typically around 1 to 2 counts per channel and then multiply this by the number of desired channels to find the total background.

Next, we continuously adjust the areal density/composition of each element until the area under the simulated curve matches the area under the actual (experimental) curve minus the total background.

As an example, the RBS analysis for a (Bi$_{1-x}$In$_x$)$_2$Se$_3$ film is shown in Figure 2.12.
As is shown in the table of Figure 2.12d, the areal density extracted from SIMNRA is divided by the total deposition time (in seconds) to calculate the flux. The flux and the atomic density can be used together to find the growth rate as well as the film thickness. To get the atomic density of the film, we use the area of the primitive cell (parallelogram), which is \( a^2 \sin \left( \frac{\pi}{3} \right) \) shown in Figure 2.12e, and considering that there is one atom in the primitive cell, the number of atoms in 1cm\( \times \)1cm area for, for example, Bi\(_2\)Se\(_3\) with lattice constant 4.14 Å is 1/ \( (4.14^2 \times 10^{-16} \sin \left( \frac{\pi}{3} \right)) \times 1 \text{ cm}^2 = 6.74 \times 10^{14} \). Since there are 2 Bi atoms in a quintuple layer of Bi\(_2\)Se\(_3\), the total number of atoms would be: \( 2 \times 6.74 \times 10^{14} = 1.3 \times 10^{15} \) atoms in 1cm\(^2\) of one QL of Bi\(_2\)Se\(_3\).

### 2.6 Structural characterization techniques

Characterization in materials science refers to the processes by which the structure and other properties of materials are probed. Both microscopy and spectroscopy characterization techniques are used to study the microscopic properties and structure of materials. Furthermore, characterization techniques are crucial for better understanding the growth conditions affecting the films structure/properties and thus help with optimizing the growth conditions and achieving high-quality thin-films. In principle, there are two types of structural measurement techniques. The first type probes the real space morphology or crystal structure. Examples of such techniques include atomic force microscopy (AFM) which provides a broad view of surface morphology, transmission electronic microscopy (TEM) and scanning tunneling microscopy (STM) which provide atomic level information.

The second type are diffraction techniques which gives information about the underlying crystal structure by producing interference pattern when the waves pass through a crystal. These techniques include X-ray diffraction (XRD), which can be used to probe various thicknesses depending on the energy of the incoming beam\(^{106}\). In case of thin-films, electron diffraction techniques are more useful because electrons cannot penetrate deep in the film, making this...
technique more suitable for probing the surface features\textsuperscript{107}. In this section, we review the characterization techniques used to study our MBE-grown films.

2.6.1 Reflection high energy electron diffraction

Reflection high energy electron diffraction (RHEED) is perhaps one of the most important and useful characterization techniques in MBE that helps grow high-quality films. RHEED is connected to the MBE system for \textit{in situ} monitoring the crystal structure of the film’s surface in real-time which provides specific details about the morphology and crystallinity of the film. This technique makes it possible to quickly and efficiently try and test different growth conditions which eventually helps pinpoint the optimal conditions usually within few days or a week. When RHEED is absent and only other \textit{ex situ} characterization techniques are available, achieving high-quality films could be much more challenging and time-consuming. In the absence of RHEED, the quality of the grown film is completely unknown until the film is taken out of the MBE at the end of the growth cycle and analyzed by various structural characterization techniques. Based on the obtained information, the next sample is grown with modified growth parameters with again no knowledge about the quality until the film is one more time taken out and studied by different characterization techniques. Until the desired growth is achieved, this procedure has to be repeated for every growth which could consume considerable amount of time and resource; this signifies the importance of the RHEED.

RHEED uses a 5-50 keV electron beam generated by an electron gun (housed inside the MBE) which impinges on the sample at a glancing angle of $\sim$1°. The diffracted beam from the sample impinges and fluoresces on a phosphorous screen of the detector mounted on a CF flange with a window. The image is then captured by a CCD camera that is housed in a black box mounted on the outside of the MBE/flange (Figure 2.9 and 2.13). The image is then sent to a computer and can be analyzed using a commercially-available RHEED software (in our case we use k-Space...
Figure 2.13 Reflection high energy electron diffraction. Left panel: a schematic of the RHEED where the electron gun (connected to the MBE system) generates a high energy beam of electrons that impinges on the sample and interacts with the sample. The diffraction pattern is then formed on the detector. Right panel: the real RHEED system connected to the Rutgers MBE system. The CCD camera captures the image the diffraction pattern formed on the phosphorous detector. The image is then sent to a computer for further analysis.

Associates, Inc.). For our system, a fixed-energy beam of 10 keV is used during the film growth which is helpful for a consistent and reliable comparison between different samples. The width of the beam is around 1 mm, and the sample is constantly being rotated during the growth for achieving a uniform growth across the substrate, thus the diffraction pattern is from the total area that the beam is sweeping which is a circle with a diameter of ~2 mm. RHEED is a surface sensitive technique due the short mean free path of electrons in materials. The mean free path of electrons in solids is nearly material independent and follows a universal curve shown in Figure 2.14. From this curve, the minimum mean free path of ~5-10Å at an energy ~50 eV, meaning that only the electrons that penetrate up to ~5 Å are elastically scattered.

Because RHEED is a diffraction process only elastically scattered electrons contribute to the interference phenomenon, and thus the diffraction pattern is from depths less than the mean free path of the electrons. Therefore, the image is formed by the surface electrons if the energy is targeted at the minimum of the curve. Because of its low scattering angle, the kinetic energy in the perpendicular direction for an incoming beam of 10 keV with a glancing angle of ~1° is about $10 \text{ keV} \times \sin(1\degree) \sim 174 \text{ eV}$, which based on the universal curve yields penetration depths of less
than 10 Å – this is less than one QL of Bi$_2$Se$_3$ film. Parallel to the surface, the mean free path is much longer which makes RHEED a sensitive tool to the surface morphology as well as surface roughness. From a RHEED image, a 2D flat surface can be easily distinguished from a rough surface with 3D features/islands. This can be seen from the geometric Ewald-construction which governs RHEED similar to other diffraction phenomena and is based on the constructive and destructive interference between diffracted waves. When an incoming beam of electrons (or photons), which can be described as a plane wave with wave vector $k_i$, is shined on a crystal lattice, the electrons diffracts from the lattice positions along certain directions. Because the diffractions are elastic, we have $|k_i| = |k_f|$ (although $k_f$ could take any angle). Along certain crystal directions the difference in the path lengths traveled by the electrons are an integer multiple of the wavelength, and, therefore, the interference is constructive (aka Bragg condition). P. Ewald presented a geometric method to easily determine the possible directions along which diffraction

Figure 2.14 Universal curve for electrons mean free path within a solid as a function of the energy of the impinging beam. Adapted from Reference 92.
peaks could occur\textsuperscript{108}. In Figure 2.15, the reciprocal lattice of a crystal in $k$-space and an incoming beam of electrons with wave vector $k_i$ are shown. Now, if the end of the incoming wave vector is graphically placed on a lattice position, and a circle with radius $|k_i|$ is drawn (aka the Ewald circle or Ewald sphere in 3D lattice), then where this circle intercepts a lattice point corresponds to a diffraction peak. As it was mentioned before, because of elastic diffractions the magnitude of the momentum is conserved and thus the outgoing wave vector also lies on the Ewald circle. The Bragg’s condition is satisfied whenever the outgoing wave vector ends at another lattice point and the difference between the incoming and the outgoing wave vectors becomes an integer multiple of the lattice vector $G_i$, $\Delta k = k_i - k_f = G_i$ (aka Lau’s equation). The diffraction peak occurs along this direction. This can be extended to 3D but with one major difference: the reciprocal transformation of a 3D crystal leads to another 3D lattice with its lattice points localized in 3D. In contrast, for a 2D lattice, the third dimension is localized in real space while delocalized (elongated)
in the reciprocal space. This can be understood by the fact that a 2D lattice that is localized in the $z$ direction is approximated by a delta-function, and upon transitioning to the $k$-space the Fourier transform yields a constant value, leading to elongation along the third dimension of the reciprocal space. This means that reciprocal lattice of 2D lattice can be thought of as a series of rods (reciprocal rods) with their axes along the $z$-direction. This difference between the reciprocal lattice of a 2D surface versus a 3D crystal lattice will show up in the RHEED diffraction patterns (Figure 2.16).

If the surface of the crystal is 2D and flat, then the beam coming in a shallow angle only interacts with the topmost layer of atoms. The resulting diffraction pattern is then formed along the directions where the reciprocal rods cross the Ewald sphere with its peaks located on a ring. Because the 2D reciprocal lattice is composed of reciprocal rods, upon rotating the sample, the peaks move continuously along the vertical direction on the phosphorous screen. Thus, the diffraction pattern shows up as sharp streaks. In contrast, when the surface of the sample is not atomically smooth and has 3D features/islands, the electrons penetrate these structures and construct a transmission pattern. Since the reciprocal lattice of the 3D crystal is composed of points, upon rotating the sample, the lattice points do not change in a continuous fashion and instead blink-on/off. Therefore, from a RHEED image, it can be easily determined whether the surface of a thin-film is atomically flat or composed of 3D feature. Additionally, the crystal structure and its overall quality can be determined by examining the RHEED patterns. Figures 2.16a and b shows RHEED images for some of the high-quality commercially-available single crystal substrates, such as $\text{Al}_2\text{O}_3 (0001)$ and Si (111) which are mostly used in our lab for the growth of epitaxial films. In addition to main peaks, there are diagonal lines that are visible in the RHEED patterns of these single crystals which are known as Kikuchi lines. Kikuchi lines can form when inelastically scattered electrons fulfill the Bragg condition along certain crystal directions. This type of scattering requires long mean free paths, and thus the presence of the Kikuchi lines is an indication of high crystal quality of the sample\textsuperscript{107}. 
The RHEED images of 2D surface of commercially-available single crystalline a, sapphire (Al₂O₃) substrate and b, Si (111) substrate (with visible 7 × 7 surface reconstruction). RHEED images of 2D surface of Bi₂Se₃ films grown on c, sapphire and d, Si substrates. The grown film on sapphire has sharper peaks compared to the one grown on Si. The RHEED images of the surfaces of e, In₂Te₃ film grown on Si which forms 3D clusters and f, an InAs film grown on GaAs substrate which forms 3D islands above a certain thickness (adapted from Reference 109). g, Diffraction from a flat smooth 2D surface (with reciprocal rods in k-space) forms a RHEED pattern with sharp peaks/streaks which is the case for a-d. In contrast, the 3D transmission RHEED patterns form when the surface has 3D structures (with discrete point in k-space) which is the case for (e) and (f).

In principle, the diffraction peaks for a smooth 2D surface are localized and sharp. If the surface becomes rough, then the peaks elongate and become less sharp and more diffuse. For example, as shown in Figure 2.16, the 2D surfaces of Al₂O₃ and Si substrates with flat large terraces appear as sharp highly localized diffraction peaks. Additionally, the improvement in RHEED patterns during Bi₂Se₃ growth is also visible as the films becomes more crystalline. Initially, when a thin layer of Bi₂Se₃ is deposited at low temperatures as a seed layer, the RHEED pattern is structurally poor. Upon heating the Bi₂Se₃ film to higher temperatures followed by depositing more Bi₂Se₃ layers, the peaks in the RHEED image become sharper and the Kikuchi lines start to show up. Additionally,
upon rotating the Bi$_2$Se$_3$ sample, the pattern repeats itself every 60° which indicates the hexagonal 6-fold symmetry of the crystal’s surface structure.

Apart from morphology, quantitative information about the surface structure as well as the growth rate of the film can be extracted from RHEED patterns. The distance between the main peaks in the RHEED image gives information about the in-plane lattice constant of the film. The diffraction pattern formed in a RHEED image, which is in $k$-space, measures the inverse of the real space lattice constant. By knowing the lattice constant of a substrate that shares the same surface lattice symmetry with the film, the lattice constant of the film can be obtained by taking the ratio of the film’s peak spacing and the one of the substrates. Using this ratio, which is equal to the inverse of the ratio of their corresponding real space lattice constants, and by knowing the lattice constant of the substrate, the lattice constant of the film can be calculated. This also helps with real-time measurement of the change in lattice constant during a deposition. Finally, the intensity of a peak in RHEED during the growth of a film can sometimes be used to determine the deposition rate and thus the flux. For the case of layer-by-layer growth, there is a change in intensity of the peak which oscillates with a frequency that is equal to the inverse growth rate. During the deposition of the next layer (when the surface is not fully covered), the surface becomes rougher which leads to more diffuse scattering, and thus intensity of the diffraction peak decreases. Once the full coverage is achieved, the diffuse scattering decreases and eventually the intensity of the peak recovers. The time for this recovery can be used to determine the deposition rate and flux. The results from the RHEED oscillations can be compared to the results attained from QCM and RBS for confirmation.

### 2.6.2 Other characterization tools

As a complement to RHEED, X-ray diffraction (XRD) provides further information regarding the crystal properties and especially about the out of plane structure. XRD is mostly useful for studying the bulk of the materials due to the large penetration depth of X-rays in most materials. Because
our TI films always grow along the same crystallographic direction, we rarely use XRD for these films. XRD is governed by the Bragg’s law where, for an X-ray with wavelength $\lambda$, the condition for coherent diffraction is $n\lambda = 2d \sin(\theta_n)$, where $d$ is the lattice spacing and $n$ is an integer corresponding to a peak at angle $\theta_n$. Usually, in thin-films the strongest peaks corresponding to the unit cell periodicities that are usually along the growth direction (in our films along the c-axis). XRD data on (Bi$_{1-x}$In$_x$)$_2$Se$_3$ films with $x \approx 0, 0.2, \text{and } 0.8$ shows the main periodicity of (0,0,3$n$) from the c-axis lattice periodicity of the quintuple layers (QL) in V-VI TI films (Figure 2.17).

Figure 2.17 X-ray diffraction on (Bi$_{1-x}$In$_x$)$_2$Se$_3$ films. The XRD data for (Bi$_{1-x}$In$_x$)$_2$Se$_3$ films with $x \approx 0, 0.2, \text{and } 0.8$ grown on Al$_2$O$_3$ substrates show the (0,0,3$n$) peaks. The red triangles mark the peaks along the c-axis for bulk crystals. Adapted from Reference 75.

Additionally, in conjunction with RHEED, atomic force microscopy was used to probe the surface morphology and roughness of our films. AFM relies on a force that atoms in two objects feel when they are brought in proximity. In AFM, when a tip, that is mounted on a cantilever, is brought close to the sample’s surface it feels a force which deflects the cantilever and strongly depends on the distance between the tip and the surface. In practice, the surface of sample is not completely smooth and has some roughness to it, thus as the tip moves laterally across the sample’s surface, the force on the tip changes with any change in the height of the surface. In AFM, piezo electric-actuators
Figure 2.18 AFM images of Bi$_2$Se$_3$ grown on various substrates reveal different surface morphologies. 100QL and 32 QL thick Bi$_2$Se$_3$ grown on a, Si (111) (adapted from Reference 110) and b, Al$_2$O$_3$ (0001) (courtesy of Eliav EDrey), respectively. Both of these show coherently oriented triangular features, characteristic of Bi$_2$Se$_3$ 6-fold symmetry crystal structure. 32QL thick Bi$_2$Se$_3$ grown on c, amorphous SiO$_2$ which is void of structure and d, SrTiO$_3$ (001) with cubic surface symmetry. The step-edge heights in all the images correspond to 1 QL which is about 1 nm in height, indicative of 2D and flat surfaces.

are used to move the sample in the vertical direction in a way that it maintains a fixed distance between the tip and the sample and thus a constant deflection on the cantilever. The signal from the actuators can then be used to construct a 3D map of the height of the surface as a function of the lateral directions. To avoid dragging the tip across the surface which could damage the soft surface of Bi$_2$Se$_3$, the tapping mode was used to scan the films used in this work. AFM images for Bi$_2$Se$_3$ films grown on different substrates are shown in Figure 2.18. Although RHEED shows more or less the same diffraction pattern of 2D surface of these films, their surface morphologies appear differently in AFM.

In addition to previous techniques, STM can be used to measure the topography of the surface of a film on atomic scale. This is achieved by placing an atomically sharp metallic tip few angstroms away from the sample surface and by biasing voltage and measuring the tunneling
current between them. This current depends on the separation between the sample’s surface and
the tip as well as the local density of states of the sample. To measure the topography, the tip is
moved across the sample surface while a feed-back control loop adjusts the separation between the
tip and the sample’s surface to maintain a constant current through the tip. The resulting change
can then be mapped as change in sample height. Due to the tunneling nature of STM measurement,
conducting samples with extremely clean surfaces are required. STM is performed under ultra-high vacuum. STM chamber and MBE chambers can be connected so that the sample can be measured
without breaking the vacuum and being exposed to the air. A vacuum suitcase can also be used to
remove the sample from MBE (without breaking the vacuum) and then to carry it (under the vacuum) to STM. For performing STM measurements on our films, Se or Te were deposited in situ
on the films as a capping layer to protect the films; these capping layers were then removed by
heating in the STM chamber to restore the surface. It has been shown that a quick ion milling prior
to heating yields an ultraclean surface to work with\textsuperscript{111}. The capping/de-capping process can be
performed for other surface-sensitive techniques like ARPES.

2.7 Growth of pnictogen chalcogenide TIs: Van der Waals epitaxy

In this section we explain the substrate preparation for the growth as well as the details of the
growth procedure of pnictogen-chalcogenide TI films on various high-quality commercially-
available substrates in our MBE system. We specifically focus on the growth of Bi\textsubscript{2}Se\textsubscript{3} in this
section, however this procedure can be extended to Sb\textsubscript{2}Te\textsubscript{3} and Bi\textsubscript{2}Te\textsubscript{3} with minor modifications
(for example in growth temperatures).

2.7.1 Substrate preparation

The substrates that are used for our TI growth are commercially-available and can be purchased in
different sizes: 5mm × 5mm, 1cm × 1cm, or a 3-inch wafer. We use molybdenum and Inconel
plates, which can endure high temperatures, specifically machined to hold these substrates. The
substrates have to be prepared for the growth and below we explain the *ex situ* substrate preparation steps:

1. All tools that come into contact with substrates are cleaned with isopropyl alcohol (IPA) and then blown with dry high-purity N\textsubscript{2} gas to remove any dust particles. These tools include the tweezers used for substrate handling, the scissors used for opening the substrate bags, and the Teflon microscope slide used for checking the cleanliness of the surface of the substrate under the microscope (step 1 in Figure 2.19).

2. The substrate bag is cut open by scissors, and the substrate is removed by grasping the edge of the substrate. Then the substrate is blown off with high-purity N\textsubscript{2} gas to remove any possible dust attracted to the surface. If the substrate of use is a wafer, it also has to be blown off with N\textsubscript{2} gas after removing from the box (steps 2 and 3 in Figure 2.19).

3. After this, the substrate is placed on the already cleaned Teflon microscope slide and the surface of the substrate is checked for cleanliness under an Olympus reflected light microscopy. A small number of dust particles do not affect the growth because they can be removed during subsequent cleaning steps; however, if any scratch, significant number of dust particles or residual of the polishing fluid (from the factory processing) is observed then the substrate should be discarded (steps 4 and 5 in Figure 2.19).

4. Once the substrate’s cleanliness is confirmed by the optical microscope, then it can be attached to the molybdenum or Inconel plate. The substrates can be placed within a window that is machined on the plates for this purpose. The substrate is held in place by two thin tungsten-rhenium alloyed wires that are threaded through the plate and are bent to apply force on the substrate to keep it tight in its place. For the wafer, the plate is removed, and the wafer can directly be mounted on the ring’s groove (step 6 in Figure 2.19). Finally, the surfaces of the substrate/wafer and the plate are thoroughly blown with N\textsubscript{2} gas.
5. The plate with the substrate attached to it (or the ring with the mounted wafer) is then cleaned \textit{ex situ} for 5 minutes in UV-generated ozone gas, which removes organic adsorbents on the surface (step 7 in Figure 2.19). The plate is then placed on the ring and (after a quick blow with N\textsubscript{2} gas) in the substrate cassette (step 8 in Figure 2.19). The cassette is then loaded in the load-lock chamber (step 9 in Figure 2.19) and then in the growth chamber where \textit{in situ} cleaning and the growth will take place.

![Figure 2.19 Substrate preparation and cleaning steps.](image)

2.7.2 Growth on different substrates

In this section, we review the growth of Bi\textsubscript{2}Se\textsubscript{3} on Al\textsubscript{2}O\textsubscript{3}, Si, SiO\textsubscript{2}, and SrTiO\textsubscript{3}. Before starting the actual growth process, we need to make sure the surface of the substrate is entirely cleaned. After cleaning the substrate by blowing with dry N\textsubscript{2} and then \textit{ex situ} in the ozone system for about 5 minutes, we need to clean them \textit{in situ} as well. This process is done in the growth chamber where substrates are heated to a high temperature to further clean the surface from any contamination. This is called outgassing step. This process varies depending on the substrate which is explained
below. After explaining the outgassing process for different substrates, we discuss the actual thin-film growth procedure on each substrate.

1. **Outgassing the Al₂O₃ substrates:** The Al₂O₃ (0001) substrate is heated to 750 °C with a ramp rate of 35 °C/min. This outgassing is done under molecular oxygen with pressure of 10⁻⁶ Torr which is introduced into the chamber when the substrate temperature is around 300 °C. At this point, the sample shutter is opened so that the surface of the substrate gets exposed to oxygen. Oxygen cleans the surface of the substrate by burning the carbon-based organic contaminates. The Al₂O₃ substrate is then left at 750 °C for about 10 minutes under molecular oxygen gas for further cleaning, and then is cooled to the growth temperature. At this stage, the O₂ flow is stopped and elemental sources can be heated to the target temperatures. We try to always heat the sources after stopping the O₂ flow to protect elements against oxidization. We close the sample shutter once the temperature is around 300 °C to maintain the cleanliness of substrate’s surface for the thin-film growth.

2. **Outgassing the Si substrates:** the Si (111) substrate is heated to a high temperature to remove the thin native silicon oxide on top. The substrate is usually heated to 830 °C with a ramp rate of 35 °C/min, however the outgassing temperature could slightly vary from substrate to substrate. Since Si is a reactive substrate it is important to cool down the substrate as soon as the native oxide gives way to the underneath crystalline Si layer which is visible by looking at the RHEED patterns. Prior to the heating, the RHEED pattern initially shows a cloudy amorphous structure corresponding to the thin layer of native oxide on the surface of the Si substrate. When heated to the right temperature the sharp streaks corresponding to single-crystalline Si will show up. As soon as the amorphous native oxide disappears, and the streaks start to show up, the substrate has to be cooled down to avoid any subsequent reactions. Upon cooling down and at lower temperatures, the RHEED starts to show 7 × 7 Si surface reconstruction. The appearance of this reconstruction is an indication of a clean surface because in the presence of even a monolayer of contamination on the substrate the Si surface reconstruction will not form; in this case the substrate is not suitable for the thin-film growth and should be discarded.
Furthermore, Si surface must not be exposed to O\textsubscript{2} during the outgassing step since amorphous SiO\textsubscript{2} could form again. Because of Si high reactivity, prior to the loading of an Si substrate, we place a dummy substrate in the growth stage and outgas the stage to 900 °C and leave it there for about 30 minutes (with the sample shutter closed at all time) in order to remove Se (or other contaminants that may react with the Si substrate) from the stage and the shutter. This outgassing step substantially minimizes the chance of adsorption of species to the surface and their reaction with the clean substrate.

3. **Outgassing the SiO\textsubscript{2} coated Si substrate:** Because the thermally deposited SiO\textsubscript{2} on Si is thicker and more stable than the native oxide on the Si substrate, it cannot be removed and the same outgassing steps for Al\textsubscript{2}O\textsubscript{3} can be applied to the SiO\textsubscript{2} substrate. Also, the reactivity of the Si substrate with Se is not an issue here because the underlying Si crystal is not exposed.

4. **Outgassing the SrTiO\textsubscript{3} (STO) substrate:** The STO substrate should not be overheated during the outgassing process, because overheating creates oxygen vacancies in the substrate and makes it metallic. Thus, we only heat the substrate to 600 °C in oxygen environment. Compared to Al\textsubscript{2}O\textsubscript{3}, the oxygen is introduced in the chamber at lower temperatures (around 200 °C) and gets closed later upon cooling down (at around 200 °C). The sample shutter is closed after closing the oxygen to prevent any unwanted material from depositing on the surface. The sources can be heated to their target temperatures after stopping the oxygen flow.

Now that the outgassing procedure for each substrate is explained, we can move on to the growth procedure of Bi\textsubscript{2}Se\textsubscript{3} thin-films on these substrates. The first MBE-growth of Bi\textsubscript{2}Se\textsubscript{3} was on Bi-terminated Si(111) substrate\textsuperscript{112,113} which was carried out by depositing a monolayer of a β-phase Bi to terminate the dangling bonds of the Si surface, which eventually serves as a buffer-layer for Bi\textsubscript{2}Se\textsubscript{3} growth. This was inspired by an earlier work by Wan et al.\textsuperscript{114} in 1991 where depositing a monolayer of Bi on Si (111) yielded either a full (β-phase) or 1/3 monolayer (α-phase) coverage. An alternative way (developed by Prof. Oh’s group) is terminating the Si dangling bonds with Se, which results in higher quality Bi\textsubscript{2}Se\textsubscript{3} thin-films with a clean interface (Figure 2.20). This can be
Figure 2.20 RHEED and TEM images of Bi$_2$Se$_3$ grown on Si (111) using the two-step growth technique. a, RHEED image evolution during the deposition of Bi$_2$Se$_3$. i. The $7 \times 7$ Si (111) surface reconstruction disappear upon exposure to Se. The image shows the Si lattice spacing. The RHEED pattern after depositing ii. 0.5QL, iii. 1QL and iv. 2QL of Bi$_2$Se$_3$ at 110 °C. The spacing between the peaks indicates that the lattice constant relaxes to Bi$_2$Se$_3$ lattice constant within 1QL. v. The RHEED pattern of 2 QL Bi$_2$Se$_3$ films after being annealed to 220 °C. vi. The final RHEED pattern of 16 QL Bi$_2$Se$_3$ which shows sharp and bright streaks. b, TEM images of a Bi$_2$Se$_3$ film grown on Si with the two-step method using the Se passivation technique which shows the transition from Si to Bi$_2$Se$_3$ within about 1 nm, with clean interface. The Van der Waals layered structure of Bi$_2$Se$_3$ is also visible. c, The TEM image of a Bi$_2$Se$_3$ film grown on Si by Bi passivation technique shows an amorphous interface. From these two TEM images, it is clear that Se passivation yield higher quality films. (a) and (b) are adapted from Reference 110 and (c) is adapted from Reference 112.

done by opening the Se shutter few seconds prior to Bi-shutter and has to be performed at an optimal substrate temperature (~110°C); not too high so Se reacts with Si and not too low so too much Se sticks to the Si surface. It has to be performed at the right temperature so that only a monolayer of Se deposits on the Si surface to form strong passivation bonds with Si atoms, and the excess Se desorbs from the surface (self-limiting growth). The Se flux is usually set to be around 10 times larger than Bi flux (Se flux $\sim 25 \times 10^{13}$ cm$^{-2}$s$^{-1}$) to ensure the environment is Se rich which in turn minimizes the chance of formation of Se vacancies in the final films. Additionally, upon Se exposure and as soon as the surface is passivated by Se, the RHEED pattern of $7 \times 7$ surface reconstruction disappears and instead streaks corresponding to single-crystalline Si will show up (Figure 2.20). Bi$_2$Se$_3$ film can then be grown on this Se-treated surface with a two-step growth method where a thin seed-layer of Bi$_2$Se$_3$ (3QL) is grown at a lower temperature that serves as a
**Figure 2.21 RHEED images of Bi₂Se₃ grown at different temperatures on Si (111) substrate.**

- **a,** The RHEED pattern of the Si 7 × 7 surface reconstruction starts to appear upon cooling down (after removing the native oxide at high temperatures). The RHEED pattern of **b,** 3QL and **c,** 10 QL Bi₂Se₃ deposited at 110 °C. The RHEED pattern for 3QL Bi₂Se₃ grown at **d,** 190 °C, **e,** 250 °C, and **f,** 350 °C. The RHEED pattern for Bi₂Se₃ with an initial 3 QL deposition at 110°C and a further deposition at **g,** 190 °C (the arrows shows the 30°-rotated twin structure), **h,** 230 °C, and **i,** 300 °C. From these images, it is clear that step **b** followed by step **h** yield the best RHEED pattern. Adapted from Reference 110.

The RHEED pattern of the Si 7 × 7 surface reconstruction starts to appear upon cooling down (after removing the native oxide at high temperatures). The RHEED pattern of **b,** 3QL and **c,** 10 QL Bi₂Se₃ deposited at 110 °C. The RHEED pattern for 3QL Bi₂Se₃ grown at **d,** 190 °C, **e,** 250 °C, and **f,** 350 °C. The RHEED pattern for Bi₂Se₃ with an initial 3 QL deposition at 110°C and a further deposition at **g,** 190 °C (the arrows shows the 30°-rotated twin structure), **h,** 230 °C, and **i,** 300 °C. From these images, it is clear that step **b** followed by step **h** yield the best RHEED pattern. Adapted from Reference 110.

- **template for the subsequent layers to be deposited at a higher temperature.** Upon opening the Bi shutter and as soon as the first QL is deposited, the peak spacing in the RHEED changes to Bi₂Se₃ lattice constant (~4.14 Å). The RHEED pattern further improves by depositing up to 3QL at 110 °C, however the RHEED pattern starts to degrade by depositing more layers at such low temperature which is most likely because of insufficient surface mobility that is usually provided by growing at higher temperatures. Then, by heating the sample to higher temperatures (~220 °C), the crystallinity of this thin seed-layer continuously improves and can serve as a template layer for the following Bi₂Se₃ layers deposited at 220 °C (Figure 2.21). This growth method yields high quality films with 2D flat surfaces, confirmed by RHEED and AFM. Deposition at higher
temperatures is not recommended due to reactivity of the Si substrate which could lead to formation of second-phase structures visible in the RHEED pattern (Figure 2.21). The two-step growth can be performed on other substrates as well and is extremely beneficial for substrate with high chemical reactivity. For more information regarding the Bi$_2$Se$_3$ growth on Si see Reference 110.

For the growth of Bi$_2$Se$_3$ films on Al$_2$O$_3$, we can also use the two-step growth method by depositing 3QL Bi$_2$Se$_3$ film at 135 °C followed by heating it to 300 °C and depositing the rest of the layers at 300 °C. The initial 3QL seed layer serves as a template for the subsequent layers to be deposited at a higher temperature. In contrast to Si, Al$_2$O$_3$ substrate is an inert substrate and this is why the temperatures for the two-step method can be higher. The reason that the two-step method is used for the growth on Al$_2$O$_3$ is because of the higher sticking coefficient of Bi and Se at lower temperatures. Bi and Se do not stick above a certain temperature (above 190 °C). The 3QL deposition at a lower temperature guarantees a higher sticking and a full coverage. Usually, the sticking is higher at lower temperature; however, higher temperature provides better surface mobility. From the RHEED pattern, it is visible that the peaks get sharper and brighter upon heating as the crystallinity of the initial 3QL layers improves which can now serve as a template for the following layers to be deposited. The final RHEED pattern of the grown Bi$_2$Se$_3$ film in Figure 2.22 shows sharp peaks and the Kikuchi lines, indicative of a 2D and high-quality surface.

The inertness of the Al$_2$O$_3$ yields films with larger terraces in the AFM image as compared to the ones grown on Si (See Figure 2.18). This is because nucleation can take place at various points in case of a reactive Si substrate which could result in a rougher surface. In contrast in Al$_2$O$_3$ case, the only possible sites for the film to nucleate are step edges, and since the growth proceeds from there, the size of the terraces eventually depends on the size of the step edges. XRD results of the films also confirm the presence of highly ordered layers (QL) along the growth direction (c-axis). For more details regarding the Bi$_2$Se$_3$ growth on Al$_2$O$_3$ substrate see Reference 17. The growth of reasonably crystalline Bi$_2$Se$_3$ film on an amorphous SiO$_2$ substrate or a non-symmetry
Figure 2.22 RHEED images of Bi$_2$Se$_3$ growth procedure on Al$_2$O$_3$ substrate. 

a, The RHEED image of a sapphire substrate prior to heating. 
b, The RHEED image of sapphire after being heated to 750 °C and annealed for 10 minutes under molecular O$_2$ with 10$^{-6}$ Torr pressure. 
c, The RHEED image of deposited 3 QL of Bi$_2$Se$_3$ film at 135 °C. 
d, The RHEED peaks of the 3QL Bi$_2$Se$_3$ get sharper upon heating to 300 °C, showing that its crystallinity improves upon heating. 
e, The RHEED pattern gets sharper and brighter after depositing the following 29 QL layers at 300 °C. 
f, The final RHEED pattern of the 32 QL-thick film after being annealed at 300 °C for an hour in Se environment. Adapted from Reference 17.

matched STO substrate is possible because of van der Waals epitaxial growth which is not sensitive to the substrate lattice matching. This is in contrast to GaAs growth, where nearly perfect lattice matched substrate is required so strong bonds between the substrate and the growing film form and as a result a high-quality growth takes place. The RHEED pattern for the film grown on these two substrates shows long streaks that do not change upon rotating the film which indicates that the film is aligned in the c-axis but polycrystalline in-plane. This is due to the lack of a compatible underlying crystalline template to coherently orient the film, therefore during the deposition of the first few layers of Bi$_2$Se$_3$, each nucleation center forms a crystallite with random in-plane
Figure 2.23 RHEED images of Bi$_2$Se$_3$ growth procedure on amorphous SiO$_2$ substrates. a, The RHEED image of a bare amorphous SiO$_2$ substrate. Even with heating the structure remains amorphous. b, The RHEED pattern of 3QL Bi$_2$Se$_3$ deposited at 110 °C, where a ring-like polycrystalline pattern is formed. c, The polycrystalline pattern gives way to streaky crystalline pattern upon heating the film to 220 °C. d, The RHEED pattern becomes slightly more streaky and sharper (with no ring-like pattern visible) upon depositing 3QL more at 220 °C. e, Final RHEED pattern for a 50QL-thick film. f-g, TEM shows that the film is orientated along the c-axis but is randomly oriented in-plane. The stacking fault and the grain boundary are also clear in these images. Adapted from Reference 115.

orientation. This is also visible in the AFM images where, in contrast to the films grown on Al$_2$O$_3$ and Si substrates, no coherently orientated triangular features are seen. Additionally, TEM images show properly formed QL structure along the c-axis but random in-plane rotations with visible planar defects and grain boundaries (Figure 2.23).

It is worth noting that, the Bi$_2$Se$_3$ films grown on Al$_2$O$_3$ or on SiO$_2$ substrate can be transferred to any arbitrary substrate through an etching process. This can be done by spin-coating Polymethyl methacrylate (PMMA) on the film and by placing the coated sample in a batch filled
Figure 2.24 The Bi$_2$Se$_3$ film etching and transfer process to an arbitrary substrate. Adapted from Reference 116.

with buffer oxide etch solution (BOE) or potassium hydroxide (KOH). The PMMA gives the sample sufficient buoyancy for floating on BOE. After some time, BOE etches the interface between the film and the substrate and peels off the film. Upon peeling, the PMMA/Bi$_2$Se$_3$ film remains floated on the surface of the BOE while the substrate sinks. The film can then gently be transferred to a batch of water where it floats again (PMMA-side facing up). After washing the film thoroughly in multiple water baths and making sure the film is free from any etchant residue, the film can be scooped by any desired substrate (Figure 2.24). For more information about the film transfer process and the related transport properties see Reference 116.

2.8 Electrical transport measurements

So far, we have discussed the growth procedure of TI samples as well as the in situ and ex situ characterization tools to study the quality of the grown TI films. However, to study the electronic properties, magneto-transport measurements have to be performed. Magneto-transport studies the
electrons flow through a material in response to an applied electric field/magnetic field. In fact, the
most important information about TIs and TSSs-related physics comes from low-temperature
magneto-resistance measurements. Additionally, transport measurement is the most sensitive
technique to measure the concentration of defects which could ultimately be used as a measure for
the quality of the grown films.

2.8.1 Drude transport

In 1900 and almost three years after the discovery of electron by Thompson, Paul Drude proposed
a model, based on classical theory of gases, to explain transport properties of electrons in metals\textsuperscript{117-118}. Drude model provides a basic framework for analyzing and understanding transport
processes. Despite its simple approximations, with little modification, Drude’s model
works for quantum mechanical behavior as well.

The flow of charge in a material is caused in response to an applied field. If an external
field $\vec{E}$ is applied to a charge carrier in a solid, then the force that the charge experiences is $\vec{F} = 
\frac{d\vec{p}}{dt} = q\vec{E}$, where $\vec{p}$ is momentum, $t$ is time, and $q$ is the electronic charge. This equation may naively
imply that the charge keeps on accelerating, however we know from experiments that when an
electric field is applied to a material, the current - and the drift velocity explained below - will
initially increase for some time but then quickly reach a steady state and the material dissipate heat,
indicating that the charge carrier dissipates the excess energy into the system through dissipative
processes that are caused by inelastic scattering (collisions with impurity potentials and phonons).
Thus, the above equation must be modified which is explained below.

Consider the momentum of a charge at time $t$ is $\vec{p}(t)$. Now, let us assume in the next
infinitesimal time interval $dt$, the charge, with the probability of $\frac{dt}{\tau}$, undergoes collision and loses
Figure 2.25 A schematic of the scattering processes of a charge carrier in a material under an applied electric field. In the absence of an electric field, the charge scatters randomly and on average it goes nowhere. Under the electric field, however, the charge drifts, and its path slightly curves between the collisions. Scattering could be caused by ions, phonons (lattice vibrations) and defects/impurities (foreign atoms or vacancies).

its momentum (a random momentum + \( \vec{F}(t) dt \)). There is also \( (1 - \frac{dt}{\tau}) \) probability that collision does not occur, and the momentum will be \( \vec{p}(t) + \vec{F}(t) dt \). Then the total contribution of these two events is:

\[
\vec{p}(t + dt) = \left(1 - \frac{dt}{\tau}\right)(\vec{p}(t) + \vec{F}(t) dt) + \text{random momentum} + \vec{F} \frac{dt^2}{\tau}
\]

If we neglect the higher order terms of \( dt \) then the equation simplifies as:

\[
\vec{p}(t + dt) = \vec{p}(t) - \frac{dt}{\tau} \vec{p}(t) + \vec{F}(t) dt
\]

\[
\frac{d\vec{p}}{dt} = - \frac{\vec{p}(t)}{\tau} + \vec{F}(t) = - \frac{\vec{p}(t)}{\tau} + q\vec{E}
\]
This means that collision acts like friction, dampening the motion of the charge carrier. In fact, another way to write this equation is based on what is known as the drift velocity \( v_d \). This is because the charge carrier drifts under the electric field and its path slightly curves between the collisions (Figure 2.25). By using \( \vec{p} = m\vec{v}_d \) (\( m \) is the mass of the free electron) and dividing both sides by \( m \), the last equation can be rewritten as \( \frac{d\vec{v}_d}{dt} = -\frac{\vec{v}_d}{\tau} + \frac{q}{m}\vec{E} \), where \( \tau \) is the average time between collisions and it is known as the “scattering time” or “relaxation time” (\( \frac{1}{\tau} \) is the scattering probability per unit time). Also, the average distance travelled between collisions is called “mean free path” (\( l = \tau v_d \)).

In a steady state, the left-hand side of the equation is zero and thus \( p = q\tau E \). On the other hand, \( p = mv \), and thus for all the \( N \) number of electrons passing through the unit area of the system, we get \( p = Nqm = Neq^2\tau E \rightarrow Nqv = Nq^2\tau E/m \). \( Nev \) describes the current density \( (j) \), and thus \( j = \frac{Nq^2\tau}{m}E \). The pre-factor \( \frac{Nq^2\tau}{m} \) is known as the electrical conductivity \( \sigma \), inverse of the electrical resistivity \( (\rho) \). In other words, \( j = \frac{Nq^2\tau}{m}E \equiv \sigma E \). \( \sigma \) can be further written as \( \sigma = \frac{1}{\rho} = \frac{Nq^2\tau}{m} \equiv Nq\mu \), where \( \mu \) is the mobility \((\mu = \sigma/Nq)\) which can be thought of as how much conductivity each charge carries. If we have two carrier species (i.e. negative/electron: \( n \) and positive/hole: \( p \); \( N = N_n + N_p \)), then \( \sigma = N_nq_n\mu_n + N_pq_p\mu_p \).

Note that the mass \( m \) that has been used in all the above equations is the mass of a free electron. However, for an electron in a crystal, the effect of crystal potential must be taken into account and the mass has to be renormalized to what is known as an effective mass \( m^* \). Doing so makes the Drude’s model suitable to use for analyzing the transport properties of many materials. This mass can be orders of magnitude larger than the free electron’s mass, like in heavy fermions, or an order of magnitude smaller in semiconductors. For more details see References 119-121.
2.8.2 Hall effect: Determination of type and number of charge carrier density

During his PhD at the Johns Hopkins University in the Fall of 1879, Edwin Herbert Hall discovered an effect that now bears his name, the Hall effect. The Hall effect (aka an ordinary Hall effect) measures the voltage difference transverse to the electric current in a conductor placed in a magnetic field perpendicular to the current. The voltage is produced because of the Lorentz response of charged particle to an applied magnetic field. Imagine we place a sample with length \( l \), width \( w \), and thickness \( t \) in a uniform magnetic field along the \( z \) direction and apply current along the \( x \) direction (Figure 2.26a). Then the charge (in the figure an electron is shown), according

![Image of Hall effect](image)

**Figure 2.26 Schematics of a Hall effect measurement.** a, A Schematic of the Hall effect on a sample with length \( l \), thickness \( t \), and width \( w \). The current flows along the \(-x\) direction in the sample, and the magnetic field is along the \(+z\) direction. The Hall voltage \( V_{Hall} \) is measured across the width of the sample in the \( y \) direction, perpendicular to the current. b, The Hall effect for holes (left) and electrons (right) under an out-plane magnetic field. They accumulate in the same side, but the direction of the electric field (from positive to negative) is different in each case.
to the right-hand rule is deflected toward the positive $y$ direction. As more electrons deflect to the +$y$, the imbalance between the charges generates an electric field $E_y$. As it is shown in Figure 2.26b both electrons and holes eventually deflect toward the same side, the generated electric field direction is what is different in these two cases; basically, in case of holes, that specific side becomes more positive whereas in case of electrons, the same becomes more negative. The Lorentz force that the charge carriers feel is given by $F_L = qv_dB$, and in steady state conditions the Lorentz force is equilibrated by $E_y$, yielding $qE_y = qv_dB$. $E_y$ can be indirectly found by measuring the generated voltage across the width of the sample and transverse to the applied current, also known as $V_{Hall}$ which is defined as $V_{Hall} = \int_0^w E_y \, dy = E_y w = w v_d B$. The Hall resistance is then found by dividing the Hall voltage by the applied current, which can in turn be written in terms of the current density $I = JA = J w t = qN v_d w t$, where $A = wt$ cross-sectional area of the sample.

Hence, $R_{Hall} = \frac{V_{Hall}}{I} = \frac{w v_d B}{qN v_d w t} = \frac{B}{qN_t} = \frac{B}{q n}$, where $n$, which is the number of charges per unit volume($N$) times the thickness ($t$), gives the total number of carriers per unit area of the sample (aka the areal carrier density, also referred to as $n_{2d}$ or $n_{sheet}$). Therefore, the number of carriers can be determined by the Hall effect measurement. Additionally, the sign of the slope of $R_{Hall}$ gives information about the type/sign of the charge carrier: negative slope for n-type/electrons and positive slope for p-type/holes.

### 2.9 Experimental details of transport measurement

In this section, we explain the experimental setup for our transport measurements.

#### 2.9.1 Geometry: van der Pauw and Hall-bar

The two main geometries used in our transport measurements are van der Pauw and Hall-bar (Figure 2.27). The two measurement techniques used in this work were 4-point (mostly used) and 2-point measurements. In a 4-point measurement technique, the current is sourced along two leads
Figure 2.27 A schematic of different geometries used in transport measurement. van der Pauw geometry for measuring a, the sheet resistance in longitudinal configuration and b, the Hall resistance in transverse configuration. c, A schematic of the Hall-bar geometry with length $L$ and width $W$; where the current is sourced between leads 1 and 2, and the transverse and longitudinal voltage drops are measured.

and voltage is measured between two other leads. In contrast, a 2-point measurement technique sources current and senses voltage along the same leads. This way the resistance of the sample is measured in series with the resistance of contacts/leads, making it impossible to determine resistance of the material of interest. However, 4-point measurement techniques eliminate the contribution of lead and contacts resistance, especially because of high input impedance of standard source meters.

In the late 1950s, van der Pauw showed that the resistivity ($\rho$) can be obtained through an averaging procedure with the leads in various locations on the sample. In thin-films, typically the sheet resistance ($R_{\text{sheet}}$) is reported which is related to the bulk resistivity by $\rho = R_{\text{sheet}}t$. To use the van der Pauw configuration, the following conditions must be satisfied: 1. the sample must be
flat with uniform thickness and the sample’s thickness should be much smaller than the width and
the length of the sample. 2. The sample must be connected with no isolated holes. 3. The sample
must be isotropic. 4. The leads must be located at the edges of the sample. 5. The area of contacts
should be much smaller than the total area of the film.

The sample leads in Figure 2.27a are labeled counter-clockwise as 1, 2, 3, and 4. If the current runs
from lead 1 to 2 ($I_{12}$) and the voltage drop is measured between leads 4 and 3, then the resistance
is defined as $R_{12,43} = \frac{V_{43}}{I_{12}}$ (call this $R_{xx}$). The second measurement can be done by sourcing current
and measuring voltage horizontally $R_{14,23} = \frac{V_{23}}{I_{14}}$ (call this $R_{yy}$). Technically for a more precise
measurement, another set of resistance values with the leads rotated 180°, $R_{43,12}$ and $R_{23,14}$, should
also be measured, but based on the reciprocity theorem $R_{12,43}(B) = R_{43,12}(-B)$. For zero-field
measurements ($B = 0$), $R_{12,43}$ and $R_{43,12}$ should be the same, and for measurements in the presence
of magnetic field this averaging ($\frac{R_{12,43} + R_{43,12}}{2}$) is known as symmetrization procedure which will
be explained in the next section. Van de Pauw showed that:

$$e^{-\frac{\pi}{R_{\text{sheet}}}} + e^{-\frac{\pi}{R_{\text{sheet}}}} = 1$$

Thus $R_{\text{sheet}} \approx \frac{\pi}{\ln(2)} R_{\text{long}}$, where $R_{\text{long}} = \frac{R_{xx} + R_{yy}}{2}$ and $\frac{\pi}{\ln(2)}$ is a geometric factor. Additionally,
Hall resistance with van der Pauw configuration is obtained by running the current between 1 and
3 and measuring voltage across 2 and 4 (Figure 2.27b). Hall resistance does not depend on the
sample geometry, and thus the carrier density can be obtained from the measured resistance value
without any geometric factor multiplication. In the Hall-bar geometry of Figure 2.27b, $R_{12,34}$ is
longitudinal resistance $R_{12,45}$ is the Hall resistance. For a Hall-bar with length of L, distance
between center to center of voltage leads) and width of W, $R_{\text{sheet}} = R_{12,34} \times \frac{W}{L}$. From this, the
sheet carrier density and mobility can be extracted.
Van der Pauw geometry can be set up easily and quickly which is particularly advantageous when it is intended to minimize the ambient exposure of the film, especially when the sample is not capped. On the other hand, Hall-bar allows for a more precise measurement of resistances and substantially minimizes the mixing of longitudinal and Hall resistances. However, making Hall-bar patterns requires fabrication process. The time and fabrication/chemical processes involved in making Hall-bars can cause severe degradation of films properties. In this regard, if making small Hall-bars is not necessary, manual scratching of Hall-bars is much faster – still slower than van der Pauw. Moreover, this method is free of any chemical processing and thus does not degrade the sample as much as lithography patterning. We mostly performed manual Hall-bar scratching on our samples. The samples are usually grown on a 1 cm$^2$ substrates and then cut into 4 pieces using a diamond tip pen cutter (Figure 2.28). This way, more devices can be made from one sample. Each individual piece can then be scratched into a Hall-bar.

### 2.9.2 Preparing the sample

Hall-bars used in this work were either hand-patterned by placing a Hall-bar designed shadow mask on the sample and scratching with a sharp metal tip (Figure 2.28) or fabricated using contact photolithographic patterning. 4-point measurement was the main technique used for this work and both van der Pauw and Hall-bar geometries are perfectly suitable for this purpose. For mounting the sample, we use a costume-made 15-pin sample holder. For attaching the sample to the sample holder, we use double-sided copper tape (both sides are coated with a thermally and electronically conducting adhesive layer). One side of the of the copper tape attaches to the sample holder and the other side is used for attaching the sample to the holder. The sample must sit tight and completely flat on the sample holder. Next, we use indium wires for making contacts. To do so, we start with commercially-available In beads/shots which are soft and ductile and thus can be flattened in form of a foil on a glass slide with the use of a smooth copper roller. Next, with a use of razorblade we cut thin wires from the In foil. After grabbing the wire with fine tweezers, one
Figure 2.28 Sample cutting and Hall-bar scratch procedure. a, A sample holder with a square-shaped window for attaching the sample (as large as 1 cm²) by copper tape. b, With the use of a glass slide and a diamond tip cutter, the sample can be cut to several pieces. This procedure produces some dust which is harmless (if samples handled with care) for our transport measurements, especially because the samples are capped. To avoid having dust directly on the sample’s surface, the sample can be spin-coated with PMMA prior to cutting, and then the PMMA – along with all dust and debris from the cutting process stick to it – can be washed away with acetone/isopropyl alcohol after cutting. However, this process can slightly degrade the sample properties, and thus we often avoid the PMMA spin-coating process. c, For scratching a Hall-bar, a shadow mask with a Hall-bar pattern is placed on top of the sample and is fixed in place using copper tape so that it does not move during scratching process. With the use of a sharp metal tip, a Hall-bar pattern can be manually scratched on the sample. Because the sample sits inside the square-shaped window, there always exists a small gap between the surface of the sample and the mask/glass slide, and thus they are not in contact with the sample’s surface. Two layers of copper tape also helps elevate the stencil mask even further, reassuring no contact between the mask and the sample’s surface during scratching.

end is attached to the sample holder by manually pressing with a toothpick. Using the toothpick again, the other end is directed toward the sample and is then gently pressed on the sample. The sturdiness of the contact can be tested by gently prodding the In wire with the toothpick. We repeat this for all the other leads (See Figure 2.29a and b). We then measure 2-probe resistance between different leads with a multimeter to ensure that the resistance values between all leads are relatively similar (for TI films the resistance value is typically from few to tens of kΩ depending on the sample). Next, the sample is loaded in the cryostat and now the sample is ready to be measured. For performing direct current (DC) electrical measurements in our lab, we use a source meter (a single channel Keithley 2400 or a dual channel Keithley 2636) in conjunction with a switch matrix.
Figure 2.29 Sample preparation and mounting on the Advance Research Systems cryostat and GMW magnet. a, Using a copper roller, the indium beads can be flattened into a foil on a glass slide. Indium wires can be cut from the foil by a razorblade. b, The sample attached to the 15-pin sample holder by copper tape, with indium wires attached by gently pressing them with toothpicks. c, The sample is mounted on the cryostat (rotatable cold head). d, The system is connected to the helium compressor (the blue instrument on the left). The cold head is then covered by two shrouds and rotated down to the bore of the GMW electromagnet, where e, it is made sure that it is leveled (top panel). The bores of the magnet shown in bottom panel are separated by 50.8 mm.

system (Keithley 7001), which automatically change the lead configuration in about ~0.1 seconds, as well as a temperature controller. The instruments are connected to the computer with GPIB lines and the measurement is controlled via a LabView-built program. As shown in Figure 2.30, the switch matrix system consists of rows and columns. The program offers several options for measurements of longitudinal ($R_{xx}$, $R_{yy}$) and Hall resistance ($R_{xy}$, $R_{yx}$) versus magnetic field, temperature, or gate voltage. The current and the voltage leads are connected to the rows (4 rows
for $I^+, I^-, V^+, \text{ and } V^-$) which is eventually connected to the source meter and the leads, which are attached to the sample, are connected to the columns of the switch matrix. Inside the switch matrix, rows and columns are electrically attached through electro-mechanical relays (on/off switches). When the switch matrix receives a signal from the computer to close a specific row to a specific column (see Figure 2.30), it shorts the relay between that row and column. This is particularly advantageous for van der Pauw and Hall-bar geometries because various connections can be quickly and automatically changed between the different configurations, thereby measuring $R_{xx}$, $R_{xy}$, with only a short delay. So, the measurement process can be summarized below:

1. Switch matrix closes in $R_{xx}$ configuration.

2. Depending on what measurement is being performed, the program measures the magnetic field, temperature, or gate voltage (see Figure 2.31).
3. The current is sourced in the positive direction (+I) and V is sensed. Thus \( R^+ = \frac{V}{+I} \) can be measured. The voltage is measured after a short delay when the current is stabilized.

4. After a short delay, the current is sourced, but this time in the opposite direction (−I) and then the second V is sensed, thus \( R^- = \frac{V}{-I} \) can be measured. These delays are particularly important, especially when the sample is insulating. Usually very short delay is fine for metallic samples, however insulating samples need longer delays (around 1 second or so) so that the delay time remains longer than the resistance-capacitance (RC) time constant for the circuit.

5. \( R^+ \) and \( R^- \) are averaged, stored in a file, and plotted in the program in real time. In general, this averaging is crucial and eliminates the noise and spurious effects, typically originating from thermoelectric effects within the sample at low temperatures being connected to the electronics at room temperature.

6. Repeat steps 1 to 5 for \( R_{xy} \), \( R_{yx} \), and \( R_{yx} \).

![Figure 2.31 The front interface of the custom-built program using LabView. With this program, resistance versus temperature, magnetic field, or gate voltage, can be measured. On the right, there are various parameters that can be adjusted for each measurement.](image)
2.9.3 Reverse field reciprocity and symmetrization

Reverse field reciprocity, a very useful method used in transport measurements which was proposed by H. H. Sample et al.\cite{Sample}, states that if the resistance is measured at magnetic field (+B), then the resistance for the reversed field (-B) can simply be obtained by switching the current and voltage leads without actually changing the magnetic field polarity. Let us assume for an arbitrary shaped sample with leads connected at arbitrary positions 1, 2, 3, and 4, the current flows from lead 1 to 2 and voltage drop is measured between leads 3 and 4, then the measured resistance at magnetic field +B is \( R_{12,34}(+B) \), where \( R_{12,34} = \frac{V_{34}}{I_{12}} \). Now, based on the reciprocity theorem instead of measuring \( R_{12,43}(-B) \), one can simply measure \( R_{34,12}(+B) \). This is when the automatic switch matrix system becomes particularly advantageous, especially because lead reversal is generally easier than field reversal. Additionally, the reverse-field reciprocity relation is particularly useful when the carrier density versus temperature needs to be measured (we call this sample-sample method). In the van der Pauw method, as the temperature is lowered, we measure \( R_{xy}(+B) \) and \( R_{yx}(+B) \) at a fixed magnetic field (without ever actually reversing the magnetic field polarity).

The carrier density can then be extracted from \( n_{2d} = \frac{1}{q(R_{xy}(+B)-R_{yx}(+B))} \) at each temperature. Figure 2.32 shows an example of the sample-sample method for a Sb\(_2\)Te\(_3\) film. This method is heavily used for finding the effective capping layer for our low-carrier-density TI films where the effectiveness of the capping was assessed by measuring the amount of change in the carrier density concentration of the capped film over a long period of time which is explained in more details in chapter 3. Moreover, in practice and no matter how carefully the leads are aligned, there is always a misalignment. Specifically, in our case where the contacts are manually placed on the sample the misalignment is always a concern.

Fortunately, this mixing can be removed due to the special properties of longitudinal and Hall resistances through what is known as (anti)symmetrization process. Because the longitudinal
An example of sample-sample method. An 8QL buffer layer-based Sb$_2$Te$_3$ film whose carrier type changes from p to n upon cooling down at around 119 K. The measurement was done using the van der Pauw geometry at a constant field while $R_{xy}$ and $R_{xx}$ were being recorded as the temperature was lowered from room temperature all the way down to ~6 K.

Resistance ($R_{xx}$ or $R_{yy}$) is an even function and the Hall resistance ($R_{xy}$) is an odd function of magnetic field we can write the following:

$$R_{xx\text{-measured}} (+B) = \alpha R_{xx} + \beta R_{xy}$$

And for $-B$ we can write:

$$R_{xx\text{-measured}} (-B) = \alpha R_{xx} - \beta R_{xy}$$

Summing up the two equations and assuming $\alpha \approx 1$ yields:

$$R_{xx} = \frac{R_{xx\text{-measured}} (+B) + R_{xx\text{-measured}} (-B)}{2}$$

This can be repeated for $R_{xy}$:

$$R_{xy\text{-measured}} (+B) = \gamma R_{xx} + \delta R_{xy}$$

And for the opposite polarity of the field:

$$R_{xy\text{-measured}} (-B) = \gamma R_{xx} - \delta R_{xy}$$
Subtracting the two equations and assuming \( \delta \approx 1 \) gives:

\[
R_{xy} = \frac{R_{xy\text{-measured}} (+B) - R_{xy\text{-measured}} (-B)}{2}
\]

### 2.9.4 Magnets

In Oh’s lab, there are two magnet systems with different base temperatures and magnetic fields. The cryogen-free system that is widely used (almost every day) in our group is designed and manufactured by Advanced Research Systems and operates by compressing, cooling, and re-expanding the helium gas using a helium compressor system. This cryogenic system is combined with a GMW electromagnet, providing a base temperature of 4 to 6 K and a magnetic field up to 0.62 T. The GMW system uses power supply that controls current. The current to magnetic field conversion is controlled via our homebuilt LabView program.

The sample space, which is covered by a vacuum shroud and a nickel coated copper radiation shield, is pumped down to \( \sim 10^{-6} \) torr regime, using a turbo pump, prior to cooling. To avoid any backflow from the pump to the cold sample space, the valve between the pump and the sample space is kept closed for temperatures below 150 K – the turbo pump is still running and pumping the connected lines and the samples space, though isolated from the pump, is still under vacuum during the measurement. The system is mounted on a rotatable gimbal system where the sample can be loaded and rotated inside the bore of the magnet (See figure 2.29c and d).

The second system is designed and manufactured by the American Magnetics Inc. and it uses liquefied nitrogen and helium (\(^4\)He) to reach temperatures as low as 1.5 K (Figure 2.33). The system looks like a Dewar which consists of, from outer most to the inner most, a vacuum space to minimize thermal leakage from outside at room temperature to the cryogenic liquids within, liquid nitrogen belly (LN\(_2\) at 77K) that is used to thermally insulate the liquid helium belly (LHe at 4 K), sitting below the LN\(_2\) belly. Both LN\(_2\) and LHe can be bought pre-liquefied in supply Dewars. The recently installed LHe recycling system at Rutgers located in Physics Serin building collects He
Figure 2.33 Photos of the American Magnetics, Inc. liquid helium cryostat. 

a, a cut-away of the inner structure of the AMI magnet (taken from American Magnetics, Inc. user’s manual). b, The image of the Dewar. The AMI system provides base temperature of 1.5K and magnetic field of 9T. c, The meter for the LHe recovery line records the amount of recovered He gas in ft$^3$ which is then going to be liquified in the LHe recycling system and can eventually be recollected as LHe in a Dewar for the next time use (top panel). The bottom panel shows the variable temperature insert with the sample holder and a sample attached to it.

gas from several labs and liquifies it, substantially reducing the cost (one 100-liter Dewar of LHe costs almost $1000). Then each lab can refill their Dewars based on their returned share of He gas. The LN$_2$ and LHe are both manually transferred into our AMI system. The 9 T superconducting magnet, made from low-temperature superconductor niobium-titanium alloy, is situated inside the LHe reservoir with the sample space at the center of the bore of the magnet. This magnet system uses factory installed power supply which controls the current as well as the ramping of the magnetic field. The sample space has a vacuum space that isolates the sample tube. We always pump down the space that separates the sample tube from the LHe belly with a turbo pump, at least
a night before starting the AMI cooling down, because this space must be under vacuum to minimize heat leakage from the sample tube into the LHe belly while the sample is heated to room temperature. Lastly, the sample is attached to a variable temperature insert that is removable for sample attachment (then inserted into the sample tube for measurement).

After the sample is inserted at room temperature, upon opening a small valve, also called needle valve, the LHe flows from the LHe belly into the sample space which cools down the sample to 4.2K. By constantly pumping the sample space during this stage, the sample can be cooled down faster. Further cooling can be achieved by valving off the sample tube, and fully opening the needle valve, which allows a significant amount of LHe to flow into the sample space. Then, we close the needle valve and reopen the vacuum line to the sample tube to begin the evaporation of the LHe by pumping which allows cooling down to 1.5 K at which measurements are carried out.

As mentioned before, the weak anti-localization effect is of particular interest to topological materials, appearing as a cusp in $R_{xx}$ versus magnetic field which features a sharp increase in the resistance (decrease in the conductance) at low fields. Therefore, we use a very slow ramp rate only around $B = 0$ T to capture this feature with high resolution, and after that we increase the ramp rate—the maximum allowed ramp rate is 0.024 T/s, and with using a higher ramp rate than this may cause magnet quenching (transitioning into a non-superconducting resistive state). Both the target magnetic field and the ramp rate can be controlled in another home-built LabView program.

### 2.10 Measurements at the National High Magnetic Field Laboratory

A major part of the measurements for this work was carried out at the National High Magnetic Field Laboratory (NHMFL) in Tallahassee, Florida which is the largest and highest-powered
Figure 2.34 Images of the 45 T hybrid magnet at the National High Magnetic Field Laboratory. a, A cut-away of inner structure of the 45 T hybrid magnet which is composed of a superconducting magnet, providing 11.5 T, and a resistive magnet, providing additional 33.5 T (courtesy of NHMFL). b, The illustration in the top panel shows that the energy a 45 T magnet consumes in a week of measurement is $\sim 10^{12}$ Joules. Bottom panel shows that the amount of water is needed to remove the heat generated by the 45 T magnet is $\sim 4000$ gallons per minute (courtesy of NHMFL). c, An 8-pin socket sample holder (top panel). The socket needs trimming to accommodate a 5 mm$^2$ sample and to fit in the probe. Bottom panel shows a trimmed socket with a copper tape wrapping a couple of times around the area where the sample is supposed to be attached. The holes on the pins are filled by squishing small In chunks so that attaching In wires to the pins (and then to the sample) would be easier. The last panel shows the holder with a mounted sample.

magnet laboratory in the world. This facility accommodates 14 resistive magnets connected to a 56-megawatt DC power supply with 15000 square feet of cooling equipment for removing the generated heat. Additionally, there are several superconducting magnets including the 45T hybrid magnet which provides the strongest continuous field as well as a recently-built engineering feat all-superconducting 32T magnet which features 6 miles of the state-of-the-art high-temperature superconducting tape of yttrium barium copper oxide (YBCO) which does not need LHe to be in a superconducting state.

For our measurements, we used the 35T resistive magnet with 32mm bore size and the 45T hybrid magnet which combines a resistive magnet with a superconducting one (Figure 2.34). The
resistive magnet is made of conventional metal conductors, such as copper or silver and the superconducting magnet is made of low-temperature superconductor niobium-tin and therefore needs LHe to enter a superconducting state. In this system, the superconducting magnet is initially ramped up from 0 to 11.5 T with slow ramp to prevent overheating—which takes about an hour—and the rest of the field, 11.5 T to 45 T, is provided by the resistive magnet—with maximum allowed ramp rate of 7 T/min. Also, for the 45 T magnet, due to slow ramping of the superconducting magnet usually only one magnet polarization is chosen during a given measurement day, whereas sweeping through both positive and negative polarizations is possible for the 35 T resistive magnet on the measurement day.

An 8-pin socket was used to load the sample using copper tape. Indium wires were used for van der Pauw and manually-scratched Hall-bar patterns. The contacts for the gated devices were made by wire bonding tool. The lowest temperature used for our measurements was 300 mK can be achieved via condensation/evaporation cycle of $^3$He and $^4$He. NHMFL provides portable dilution which can be installed in the magnet cell, and measurements can be done at even lower temperatures (as low as 25 mK). Furthermore, the measurement programs were provided by NHMFL. Additionally, to avoid any measurement artifact, the samples with high resistance at higher fields were coupled with a high sensitivity nanovoltmeter and biased with a DC current using Keithly 6221 current source, whose polarity alternated approximately every 1.5 s (the delay time can be controlled in the program).
Chapter 3

Defect and interface engineering of topological insulators thin-fims\(^1\)

Ideally, TIs must have an insulating bulk and metallic surface states. However, in practice TIs are far from ideal and plagued with defects and vacancies which make the bulk conducting and push the Fermi level deep into the conduction (valence) band. Thus, a major focus in TI research as well as an important step towards fabricating TI-based devices is to minimize the defects and lower the carrier density by bringing the Fermi level into the bulk energy gap and as close as possible to the Dirac point, such that the physics is governed by the topological surface states.

In fact, defect engineering has brought about major materials advancements, leading both to the observation of numerous novel physics phenomena and realization of technological applications. Defect engineering and tailoring materials properties can be performed through different growth methodologies, such as impurity addition/alloying, intercalation, (counter) doping (adding minute amounts of impurities), remote/modulation doping (to avoid impurity-induced scattering), and interface-engineering. Successful examples range from realizing exotic states of matter, such as the fractional quantum Hall system, to creating the building blocks of the technological devices, such as transistors. The fractional quantum Hall effect, awarded the 1998 Nobel prize, was first observed in a 2-dimensional electron gas (2DEG) formed with interface-engineered and modulated-doped GaAs films. The CMOS transistor, billions of which are crammed into modern computing chips, function based on gating a combination of n- and p-doped Si. The long sought-after blue light emitting diode (LED), which was required for solid-state white lighting

\(^1\)Part of this chapter is written based on a review paper Salehi, M.; et al. From classical to quantum topological surface states: evolution of thin film topological insulators with defect engineering, which is in preparation at the time of writing this dissertation.
and was awarded the 2014 Physics Nobel prize, uses both interface engineering and doping. Blue LED uses an undoped GaN template layer and involves an InGaN active layer sandwiched between properly p-doped and n-doped GaN/AlGaN.

Likewise, defect engineering strategies were extended to topological insulator systems to suppress the bulk and surface defects that have plagued both thin-films and bulk crystals. Efforts include field effect modulation\textsuperscript{19,123-131}, growth of thin-films with high surface to volume ratio on various substrates\textsuperscript{17, 19, 110, 115, 132}, compensation doping as well as intercalation\textsuperscript{27, 133}, growth of ternary\textsuperscript{124, 134-140} and quaternary\textsuperscript{141-143} compounds derived from parent materials, and more recently, growth of TI thin-films on an optimized buffer-layer through interface engineering\textsuperscript{144}. Major advances have been made with these techniques, leading to the achievement of TSS-dominated transport. In turn, a plethora of novel physics phenomena has been observed in TSS, including the quantum anomalous Hall effect\textsuperscript{145-149} and its transition to an insulator\textsuperscript{150-151}, the quantum hall effect\textsuperscript{144, 152,153-154} and its transition to an insulator\textsuperscript{155}, the axion insulator\textsuperscript{156-158}, quantized Faraday and Kerr rotation\textsuperscript{159-160}, a finite-size driven topological phase transition and metal-insulator transition\textsuperscript{14}, and the possible observation of chiral Majorana modes\textsuperscript{161} which are proposed as a promising platform for quantum computation\textsuperscript{162}.

In this chapter, we will discuss how defects have been affecting the properties of TIs and explain the means to overcome them. It will be discussed how suppressing defects in V-VI TI films through a proper interface engineering led to achieving unprecedented low carrier density TIs. The next chapter will discuss how these new generation low-carrier-density films revealed the heretofore inaccessible topological quantum aspects of TIs.

3.1 The origin of Bulk Defects\textsuperscript{2}

In practice, all materials have some level of defects and disorder. It is known that V-VI TIs are heavily doped due to crystal defects, such as vacancies and anti-site defects. For example, Bi\textsubscript{2}Se\textsubscript{3}.

is usually found to be intrinsically n-type due to Se vacancies (which leave behind two electrons in the crystal). Sb$_2$Te$_3$ is usually found to be p-type due to anti-site defects (i.e. Sb atom in Te position). Bi$_2$Te$_3$ can be either p-type or n-type depending on the growth conditions and defect types (vacancies or anti-sites). Some of these defects are thermodynamically induced in the crystal, however they can also be created due to extrinsic effects. For example, defects can be created at the interface of the film and substrate during the film growth or due to environmental exposure. These defects lead to formation of quantum confinement-induced bulk-derived two-dimensional Fermi gas (2DFG) near the surface due to band-bending effects which are briefly explained in the next section$^{163-167}$. These 2DFGs along with bulk form a parallel conduction channel, resulting in decreased contribution from TSS conduction. In fact, it has been frequently found that TSS conduction in these materials is substantially masked by the bulk conduction and it only contributes less than few percent of the total conduction, making detection of TSS through transport extremely difficult. As mentioned before, many efforts have been made to eliminate this bulk channel (bulk conduction plus any bulk related 2DFGs) in both bulk crystals and thin-films$^{139,168-172}$.

3.1.1 the Mott criterion

The defects that are embedded in the crystal can be considered as individual atoms with atomic-like bound states which rescales the Coulomb potential, yielding an effective Bohr radius $a_B = \epsilon \frac{m}{m^*} a_0$, where $\epsilon$, $m$, $m^*$ and $a_0$ are the dielectric constant of crystal, free electron mass, effective mass of bulk electrons in the crystal, and free space Bohr radius ($\approx 0.05$ nm), respectively. For example in case of Bi$_2$Se$_3$, where $\epsilon \sim 110$, $m^* \approx 0.15$ $^{173}$, $a_B$ is almost 37 nm, which corresponds to the size of the Se vacancy. Now, as the number of vacancies increases, they begin to overlap, and the electrons bound to vacancies become mobile. When this number reaches the critical dopant density ($N_C$), where the mean spacing between the dopants becomes of the same order as the effective $a_B$, a metal-to-insulator transition occurs$^{174}$. This critical value was quantified by Sir
Neville Mott in 1960s as $N_C^{-\frac{1}{3}} \approx 4a_B$, which yields $N_C \sim 3 \times 10^{14}$ cm$^{-3}$ for Bi$_2$Se$_3$. Based on this, it is not surprising that all grown bulk crystals have metallic bulk; even the lowest ever reported volume carrier density of $\sim 10^{16}$ cm$^{-3}$ in Bi$_2$Se$_3$ bulk crystals$^{175}$, which is already two orders of magnitude higher than $N_C$, has a conducting bulk. The Mott criterion gives critical information about the position of the Fermi level ($E_F$) deep inside the bulk. Since all the experimental transport data gives a bulk carrier density that exceeds the Mott criterion by at least a few orders of magnitude, it is then implied that the $E_F$, deep in the bulk, must be at the conduction band minimum. On the other hand, the value $N_C \sim 10^{14}$ cm$^{-3}$ is rather small when it is compared to the ones in, for instance Si and GaAs with $N_C \sim 10^{18}$ cm$^{-3}$, suggesting that it may be thermodynamically impossible to reach such low defect density in bulk crystals of Bi-based TI materials because of their much weaker chemical bonding. Fortunately, for TI thin-films the story can be different, and as we explain below the Fermi level can in fact be lowered into the bulk band gap by properly engineering the band-bending phenomenon as well as the dimensionality.

3.1.2 The Surface versus bulk defects: Band-bending effect

As mentioned in the previous section, the Mott criterion pins the bulk $E_F$ to the bottom of the conduction band (CB). Now, if the surface $E_F$ is different than the bulk $E_F$, then the charges keep flowing until the Fermi level is aligned everywhere in the material, which results in band-bending near the surface and thus a spatial charge imbalance near the surface – band-bending was first introduced by Schottky and Mott$^{176-177}$. If initially the surface $E_F$ is also at CB minimum, then no band-bending is required, and the bands remain flat. We know that the Fermi wave vector $k_F$ for a 2-dimensional system is $k_F = (2\pi n_{2D})^{\frac{1}{2}}$, where $n_{2D}$, which is the 2-dimensional carrier density for two topological surface states of a 3D TI, is expected to be $n_{2D} = 2 \times n_{SS}$ (if we assume each surface has similar carrier density of $n_{SS}$). Then, based on the dispersion relation for a Dirac state, the energy of TSS is $E_{SS} = \hbar v_F (4\pi n_{SS})^{\frac{1}{2}}$, where $v_F$ is the Fermi velocity. If we calculate the $n_{SS}$ for
Bi$_2$Se$_3$ with its conduction band minimum at almost 200 meV above the Dirac point and $v_F \approx 4 \times 10^5$ ms$^{-1}$ (extracted from the ARPES data), then for the surface $E_F$ to be located at the CB minimum to fulfill the flat band scenario, the $n_{TSS}$ needs to be $\approx 5 \times 10^{12}$ cm$^{-2}$ (or $1 \times 10^{13}$ cm$^{-2}$ for both surfaces). However, if the $n_{SS} > 5 \times 10^{12}$ cm$^{-2}$ (or $1 \times 10^{13}$ cm$^{-2}$) and initially surface $E_F >$ bulk $E_F$, then electrons flow from the surface towards the nearby bulk until the $E_F$ is aligned everywhere. This gives the surface a net positive charge and the nearby bulk a net negative charge, and thus causes the bands bend downward near the surface, creating an accumulation region near the surface. The downward band-bending gives rise to a non-topological 2DFG which is also observable in ARPES. If, on the other hand, initially surface $E_F <$ bulk $E_F$ where $n_{SS} < 5 \times 10^{12}$ cm$^{-2}$, then electrons flow from the nearby bulk to the surface which results in upward band-bending, and thus a depletion region near the surface. The accumulation region caused by the downward band-bending can be treated like a typical metal, and thus Thomas-Fermi approximation can be used to estimate the screening length scale $l_s \approx \left(\frac{\varepsilon_0 \pi^2 \hbar^2}{\kappa_F m^* e^2}\right)^{\frac{1}{2}}$ to be less than 1 nm, where $e$ is the electron charge, $\varepsilon_0$ is the free space dielectric constant and the Fermi wave vector $k_F \approx 0.07$ Å as well as the effective mass $m^* \approx 0.15 m_e$ are extracted from ARPES measurements. However, the story is completely different for the depletion region in the upward band-bending scenario, where Poisson equation should be used to find the depletion screening length, which is usually much larger than the length for downward band-banding. Although, downward band-bending is what it often occurs in TIs, Brahlek et al. 163 have shown that if upward band-bending is achieved for TI thin-films, then the Mott criterion, which is responsible for the conducting bulk, can be circumvented once the films are made thinner than approximately twice the depletion region. This can qualitatively be explained by the fact that the upward band-bending is the result of charge transfer from the nearby bulk to the surface, so if the film thickness is thinner than depletion region then there are not enough charges left in the nearby bulk to be transferred to the surface to equilibrate the $E_F$. Thus, the only way for
the system to do so is to allow the bulk $E_F$ to fall below the CB minimum, and hence true bulk-insulating TIs$^{27,144,155}$.

**Figure 3.1** Band-bending dependence on the TSS carrier density and thickness. The $n_{SS}$ is the total carrier density for the top and bottom topological surface states. The Mott-criterion pins the bulk $E_F$ to the conduction band (CB) minimum if the bulk defect density is larger than $\sim 3 \times 10^{14}$ cm$^{-3}$. a, $n_{SS} = 1.0 \times 10^{13}$ cm$^{-2}$ locates the surface $E_F$ at the CB minimum, and thus the bands remain flat. b, $n_{SS} > 1.0 \times 10^{13}$ cm$^{-2}$ gives rise to downward band-bending which is typically the case for conventionally grown TIs. c, If $n_{SS} < 1.0 \times 10^{13}$ cm$^{-2}$, then band-bending is upward, which is the case for optimally Cu-doped Bi$_2$Se$_3$ and the buffer layer-based Bi$_2$Se$_3$ or (Ti-doped) Sb$_2$Te$_3$ films. d, in the bulk limit, where the TI’s thickness is much larger than the depletion length ($t \gg z_d$), the Fermi level remains at the CB minimum ($t$ is film thickness). e, Band-bending is still significant in the intermediate regime ($t \approx 2z_d$). f, Once $t < 2z_d$ the Fermi level falls below the bottom of the conduction band and a true bulk-insulating TI can be achieved. Adapted from Reference 163.

Using Poisson equation $\nabla^2 V(z) = -\frac{e^2N_B}{\varepsilon \varepsilon_0}$, where $V$ is the potential energy as a function of distance from the surface ($z$), $e$ is the electron charge, and $N_B$ is the bulk dopant density (assumed to be uniformly distributed), it can be shown that $\Delta V = e^2z_d^2N_B/(2\varepsilon \varepsilon_0)$ for the boundary conditions of $V(z = z_d) = 0$, and $V(z = 0) = \Delta V$, where $\Delta V$ is the energy difference between the bands deep in the bulk and at the surface, and $z_d$ is the depletion region length (no electric field beyond
Then, by using reasonable parameters\textsuperscript{17, 27} in Bi\textsubscript{2}Se\textsubscript{3}, Brahlek \textit{et al.} have estimated the depletion region to be \(\sim 50-100\) nm. Films thinner than these length scales still preserve their topological nature, as long as the thickness remains above the critical thickness where the top and bottom surface start to hybridize\textsuperscript{178-181}, and thus a true TI with insulating bulk and the \(E_F\) in the bulk gap can be achieved in these thin-films with upward bend-bending near their surface. Note that the bulk cannot be made insulating by thinning the films if the band-bending is downward.

Additionally, surface/interface defects directly determine the position of the surface \(E_F\), and eventually are crucial for satisfying the necessary condition (surface \(E_F < \) bulk \(E_F\)) for upward band-bending, thus minimizing surface defects holds a great importance in this regard. Since TI thin-films are grown on substrates, the choice of a proper substrate plays a critical role in the level of interfacial defects and eventually in the final properties of the grown TI thin-films. Furthermore, a proper capping layer is crucial to protect the TI against atmospheric species which could potentially introduce additional defects into the TI system.

Moreover, impurity bands in doped materials can also be considered as a source of defects. Usually \(R\) vs. \(T\) measurements of TI thin-films are accompanied by a hump in the data at lower temperatures, which could be attributed to defects freeze-out at lower temperatures. In addition to the band-bending scenario where the freezing of the surface defects could be responsible for observing the hump in the \(R\) vs. \(T\) data, the impurity bands formed in a doped material, specifically the ones formed in the middle of the bulk band gap, should be considered as well. If, based on the Ioffe-Regel criterion, these bands become insulating upon cooling down, then their carriers freeze out at low temperature which could potentially give rise to the hump in the \(R\) vs. \(T\) data.
3.2 Conventionally grown TI films with no buffer layer: substrate matters

Due to their layered structures, a significant number of pnictogen-chalcogenide TI thin-films have been grown by MBE technique on various substrates including magnetic substrates\textsuperscript{182-196}. For instance, Bi\textsubscript{2}Se\textsubscript{3}, the widely studied 3D TI, has been grown directly on several substrates, such as Al\textsubscript{2}O\textsubscript{3} (lattice mismatch of 14\%)\textsuperscript{17, 197-198}, Si (lattice mismatch of -7.3\%)\textsuperscript{21, 110}, GaAs (lattice mismatch of -3.4\%)\textsuperscript{199-202}, amorphous SiO\textsubscript{2} (lattice mismatch of 18.6\%)\textsuperscript{115, 203-204}, SiC (lattice mismatch of -26\%)\textsuperscript{205}, graphene (lattice mismatch of -40.6\%)\textsuperscript{132, 206-207}, SrTiO\textsubscript{3} (lattice mismatch of 5.7\%)\textsuperscript{19, 22}, MoS\textsubscript{2} (lattice mismatch of 25\%), etc. In the previous chapter, we specifically explained the growth procedure of Bi\textsubscript{2}Se\textsubscript{3} on widely used commercially-available Al\textsubscript{2}O\textsubscript{3}, Si, and SiO\textsubscript{2} and SrTiO\textsubscript{3} substrates.

However, almost all of these grown films tend to have low carrier mobility as well as high carrier density, where even the application of gate-voltage remains inapplicable. We refer to the films with high level of defect density that are grown directly on commercially-available substrates as “conventionally-grown films”. While various factors, such as growth temperature, pressure, and cation-to-anion flux ratio play a role in determining the film quality, the choice of substrate itself seems to be the most important factor. For example, in our Bi\textsubscript{2}Se\textsubscript{3} grown on Al\textsubscript{2}O\textsubscript{3}(0001), high and almost thickness independent carrier-density ($n_{sheet} \sim 3 - 4\times10^{13}\text{cm}^{-2}$)\textsuperscript{17} is observed in a wide range of film thicknesses, suggesting two-dimensional conduction channel. Such tendency could be because of high interfacial disorder due to large lattice mismatch between Bi\textsubscript{2}Se\textsubscript{3} and Al\textsubscript{2}O\textsubscript{3} (~14\%), but lower bulk defects due to chemical inertness of Al\textsubscript{2}O\textsubscript{3}. Furthermore, the carrier density is larger than $n_{TSS} \approx 2\times(5\times10^{12})\text{cm}^{-2}$, suggesting that the band-bending is downward near the surface and both 2DFGs and TSSs contribute to the transport.
Figure 3.2 Transport properties of the Bi$_2$Se$_3$ films grown on Al$_2$O$_3$ and Si substrates. The left panel shows the sheet carrier density and the right panel shows the mobility of Bi$_2$Se$_3$ films with various thicknesses grown on Al$_2$O$_3$ (0001) and Si (111) substrates. For the films on Si, $n_{\text{sheet}}$ shows a clear thickness dependence, suggesting both bulk and interfacial defects contribute to transport, while for films grown on Al$_2$O$_3$, $n_{\text{sheet}}$ is nearly thickness independent, indicating that interfacial defects are dominant. Data taken from measurements done by Namrata Bansal and Y.S. Kim in Prof. Oh’s lab.

In contrast, the carrier-density of the films grown on Si was found to scale with film thickness ($n_{\text{sheet}} \sim t^{0.5}$), suggesting a combination of interfacial as well as bulk defects dominate the electrical properties of these films. Considering that the Si substrate has a smaller lattice mismatching with Bi$_2$Se$_3$ as compared to the Al$_2$O$_3$ substrate, it is very likely that the different thickness-dependence of the carrier-density stems from the highly reactive nature of Si (We discussed the chemical reactivity of Si substrate in the previous chapter as well). Figure 3.2 shows the transport properties of Bi$_2$Se$_3$ films grown on Si (111) and Al$_2$O$_3$ (0001).

In conventional systems like GaAs, lattice matching of the substrate is the most stringent constraint on the film growth due to strong bonding between the film and the substrate. However, the layered nature of pnictogen-chalcogen TIs and the van der Waals epitaxy allows high-quality thin-film growth on substrates with big lattice mismatch, where the film relaxes itself within the first QL. For the growth of these films, the chemical reactivity of the substrate seems to play a more
important role than the lattice matching in determination of the film quality. Recent experiment on Bi$_2$Se$_3$ films grown on amorphous SiO$_2$ which shows lower carrier density and higher carrier mobility than films grown on Si bears out this point$^{115}$; in other words, less reactivity of SiO$_2$ than Si leads to enhanced transport properties. Furthermore, lack of significant improvement in mobility and carrier-density of Bi$_2$Se$_3$ films grown on nearly perfect lattice matched but reactive InP or CdS substrates (with only 0.2% and -0.2% lattice mismatch, respectively) also substantiates the importance of chemical inertness of the substrate. In this case, the strong chemical reaction between In and Se at the film and substrate interface is believed to be the main culprit for degradation of the films electrical properties$^{209}$. In another attempt, instead of growing the film directly on InP, Chen et al.$^{210}$ incorporated Be-doped InGaAs (to suppress the n-type conductivity) and ZnCdSe buffer-layer on InP(111) substrate as a growth template for Bi$_2$Se$_3$ and they were able to obtain lower carrier-density. This enhancement is another confirmation for the important role of chemical inertness of the substrate in achieving high-quality low-carrier-density films. Therefore, in order to improve the film’s quality, substrates with both good lattice matching and chemical compatibility are required. In fact, the drop in carrier density that has been observed after peeling the Bi$_2$Se$_3$ films grown on oxide substrates by etching the interface using buffer oxide wet etching (explained in the previous chapter), and transferring them on a Si or STO substrate confirms that the interface is the major source of defects$^{116}$. Therefore, interface-engineering is an effective approach to substantially suppress defects in TI films and a major leap toward achieving true bulk-insulating TIs.

In later sections of this chapter, we will discuss the recent achievements of high-mobility and ultralow-carrier-density Bi$_2$Se$_3$ and Sb$_2$Te$_3$ thin-films grown on chemically- and structurally-matched In$_2$Se$_3$-based buffer-layer.
3.3 Defect suppression via interface-engineering: new generation TI films with a buffer layer$^3$

As it was discussed in the previous section, TI thin-films were successfully grown on various substrates. However, these thin-films harbor high level of defect density mostly originating from the interface of the film and the commercially-available substrates. Therefore, a structurally- and chemically-compatible insulating template (aka buffer-layer) has to be properly designed and virtually grown.

3.3.1 Bi$_2$Se$_3$ films on In$_2$Se$_3$/BiInSe$_3$ buffer layer: growth and transport properties

In 2015, the growth of record low-carrier-density and high-mobility Bi$_2$Se$_3$ thin-films using In$_2$Se$_3$/BiInSe$_3$ buffer-layers (BIS-BL) was reported$^{144}$. Both In$_2$Se$_3$ and BiInSe$_3$ are insulating, chemically inert, and share similar structure with Bi$_2$Se$_3$ with 3.3% and 1.6% lattice mismatch, respectively, which is much smaller than most commercially-available substrates. The buffer-layer is grown on a Bi$_2$Se$_3$ seed-layer which is the key to obtaining these high-quality samples. The growth procedure of buffer layer-based Bi$_2$Se$_3$ films is shown in Figure 3.3. The films are all grown in our custom-built SVTA MOS-V-2 MBE system with base pressure of low $10^{-10}$ Torr (the details of this system were explained in the previous chapter). After cleaning the 1 cm$^2$ sapphire (Al$_2$O$_3$) (0001) substrate ex situ by a 5-minute exposure to UV-generated ozone, the substrate was heated in situ to 750 °C under oxygen pressure of $1 \times 10^{-6}$ Torr and was kept at 750 °C for about ten minutes (the outgassing process was explained in the previous chapter). The substrate was then cooled down to 135 °C at which a thin (~3QL) seed-layer of Bi$_2$Se$_3$ is grown that is intended to serve as a template.

$^3$Section 3.3 uses some of the parts in a review paper (Salehi, M.; et al. From classical to quantum topological surface states: evolution of thin film topological insulators with defect engineering) that is in preparation at the time of writing this thesis. Section 3.3.1 is reproduced in part with permission from Koirala, N.; Brahlek, M.; Salehi, M.; et al. Nano Lett., 2015, 15 (12), 8245–8249. ©2015 American Chemical Society with permission. Also, section 3.3.2 is based on work Salehi, M.; et al. arXiv:1903.00489, 2019.
for the next 20QL In$_2$Se$_3$ layer to be deposited at 300 °C. Next, the whole heterostructure 3QL Bi$_2$Se$_3$/20QL In$_2$Se$_3$ is heated to 600 °C. Upon heating, the thin Bi$_2$Se$_3$ layer evaporates and diffuses out of the In$_2$Se$_3$ layer, leaving behind a highly insulating In$_2$Se$_3$ layer for the following layer of 20QL BiInSe$_3$, a 50% solid solution of Bi$_2$Se$_3$ and In$_2$Se$_3$, which is deposited at 275 °C. Next, the Bi$_2$Se$_3$ with a desired thickness is grown on top of the 20QL In$_2$Se$_3$/20QL BiInSe$_3$ buffer layer at 275 °C. For growing these films, 99.999% pure Bi, In and Se sources were used and thermally evaporated using Knudsen effusion cells. Elemental fluxes were measured in situ by QCM and ex situ with RBS. The Se to Bi/In flux ratio was kept above 10:1 at all time to maintain the rich Se environment. For the initial Bi$_2$Se$_3$ growth, the selenium shutter was opened 3 to 4 seconds prior to Bi shutter, and after that the Se shutter remained open at all time during the growth. The Se shutter was eventually closed upon cooling down at 135 °C to avoid Se sticking to the film. For the solid solution (Bi$_{0.5}$In$_{0.5}$)$_2$Se$_3$ growth, bismuth and indium were co-evaporated by opening both shutters at the same time.

Additionally, RHEED patterns show that the surface is 2D and smooth at every stage of the buffer layer and the film growth. Further, as is shown in Figure 3.3, unlike the growth on commercially-available substrates, high-angle annular dark-field scanning transmission electron microscopy (HAADF-STEM) shows sharp interfaces between each layer (without haziness) which is indicative of suppressed interfacial defects.

In contrast to Bi$_2$Se$_3$, In$_2$Se$_3$ comes in five different phases, where only the α-phase has a layered structure$^{211}$. Growing α-In$_2$Se$_3$ is not as straightforward as Bi$_2$Se$_3$ on sapphire, possibly due to large lattice mismatch (~16%) which makes it energetically more favorable for the In$_2$Se$_3$ film to grow in a polymorphic fashion and even form 3D structures. This can be seen in the RHEED patterns, where the growth of In$_2$Se$_3$ directly on Al$_2$O$_3$ results in a rough surface with elongated streaks which do not change upon rotating the sample, indicating that the film has many different phases and is randomly oriented in-plane. Because of similar structure, Bi$_2$Se$_3$ which grows well
Figure 3.3 The growth procedure and characterization of the Bi$_2$Se$_3$ films grown on the In$_2$Se$_3$/BiInSe$_3$ buffer layer. a, The schematic of the growth procedure of Bi$_2$Se$_3$ on the 20QL In$_2$Se$_3$/20QL BiInSe$_3$ buffer layer along with the growth temperature ($T$), sheet resistance ($R$), and the RHEED pattern at each stage of the growth. The RHEED images show a high-quality 2D surface at each stage. b, The STEM on the buffer layer-based Bi$_2$Se$_3$ shows sharp interfaces, in contrast to the conventionally grown films with hazy interface. c, The interface between the Bi$_2$Se$_3$ and BiInSe$_3$ layers is sharp, as well. d, The STEM of the Bi$_2$Se$_3$ films grown on Al$_2$O$_3$ (0001) substrate and e, the TEM of the Bi$_2$Se$_3$ grown on Si (111) substrate show hazy and disordered interfaces. Adapted from Reference 144.

on sapphire can serve as a compatible template layer for In$_2$Se$_3$, and thus high-quality layered In$_2$Se$_3$ film can be achieved once it is grown on the Bi$_2$Se$_3$ template layer. Additionally, the almost complete evaporation of Bi$_2$Se$_3$ was confirmed by RBS. For the RBS experiment, a film of 10QL In$_2$Se$_3$ on a rather thick 50QL Bi$_2$Se$_3$ was grown and then heated to 600 °C. After analyzing the RBS data (as explained in the previous chapter) we found that the film structure was composed of ~10 QL (Bi$_{0.02}$In$_{0.98}$)$_2$Se$_3$ (with 1% error bar). This shows that the 50QL Bi$_2$Se$_3$ layer almost completely evaporated away and diffused out of In$_2$Se$_3$ upon annealing the heterostructure to 600 °C, leaving behind an almost pure insulating In$_2$Se$_3$ layer (Figure 3.4a). This is consistent with the HAADF-STEM results on our films which indicates that the conducting 3QL Bi$_2$Se$_3$ seed layer evaporates almost entirely and diffuses out of the 20QL In$_2$Se$_3$ layer upon heating to 600 °C. This suggests that the 3QL Bi$_2$Se$_3$ serves as a temporary template layer for the high-quality growth of
**Figure 3.4 The RBS and STM on Bi$_2$Se$_3$ films grown on the In$_2$Se$_3$/BiInSe$_3$ buffer layer.** a, the RBS data for a 50 QL Bi$_2$Se$_3$/10 QL In$_2$Se$_3$ heterostructure heated to 600 °C. The data analysis yields a ~10 QL (Bi$_{0.02}$In$_{0.98}$)$_2$Se$_3$, which indicates that Bi$_2$Se$_3$ evaporates away almost entirely upon heating the heterostructure to 600 °C. The inset is a schematic of the film before and after the annealing to 600 °C. b, The STM images show various types of defects, Se$_{vac}$, In$_{top}$, and In$_{bot}$, on the topmost layer of 30 QL (left panel) and 5 QL (right panel) Bi$_2$Se$_3$. Adapted from Reference 144.

In$_2$Se$_3$ and then disappears. Even such a thin layer of Bi$_2$Se$_3$ (grown directly on sapphire) is conducting and because it is desired for a buffer layer to be as insulating as possible, the evaporation of 3QL is an extremely important step for obtaining an efficient highly insulting buffer layer.

Moreover, the study of In diffusion into the top Bi$_2$Se$_3$ layer holds a significant importance because it is known that the solid solution (Bi$_{1-x}$In$_x$)$_2$Se$_3$ undergoes a topological phase transition from a TI to a trivial insulator (SOC-strength weakens upon adding a lighter element In) for a critical In concentration range (usually for 4 to 7% In). The topological phase transition in (Bi$_{1-x}$In$_x$)$_2$Se$_3$ has extensively been studied in previous works. Thus, minimizing/eliminating the In diffusion into the top Bi$_2$Se$_3$ layer protects the topological nature of the Bi$_2$Se$_3$. Depositing BiInSe$_3$ between the In$_2$Se$_3$ and Bi$_2$Se$_3$ layers as well as lowering the growth temperature to 275 °C allowed achieving high-quality Bi$_2$Se$_3$ films with negligible In diffusion which was confirmed by scanning tunneling microscopy shown in Figure 3.4b. Additionally, the
Figure 3.5 ARPES confirming the topological nature of Bi$_2$Se$_3$ grown on buffer layer. ARPES on a, a 30QL Bi$_2$Se$_3$ film grown on BIS-BL with the Fermi level ~ 0.17 eV above the Dirac point and b, a 50 QL Bi$_2$Se$_3$ film grown on Al$_2$O$_3$ with its Fermi level ~ 0.33 eV above the Dirac point. It clearly shows the topological surface states in both films. It also shows the 2DFGs for the film grown on Al$_2$O$_3$ as a result of downward band-bending near the surface. Adapted from Reference 144. ARPES measurements were done by Prof. Daniel Dessau's group at the University of Colorado Boulder.

Insulating BiInSe$_3$ layer has a smaller lattice mismatch with Bi$_2$Se$_3$, which makes it an efficient template layer. The STM images of a 30QL (left panel) and a 5QL (right panel) Bi$_2$Se$_3$ film grown on the In$_2$Se$_3$/BiInSe$_3$ buffer layer show three different defects: Se vacancy (Se$_{vac}$), In$_{top}$ and In$_{bot}$. By counting In$_{top}$ and In$_{bot}$ on the surface and using the atomic areal density of Bi ($\rho_B$), we can estimate the In diffusion (In%) in the top Bi$_2$Se$_3$. It was explained in the previous chapter that the atomic density for Bi$_2$Se$_3$ is $\rho_B \approx 6.74 \times 10^{14}$ cm$^{-2}$. From the STM image, we can count In$_{top} \sim 11$ and In$_{bot} \sim 9$ in a 30 nm $\times$ 30 nm area (area). This gives In(%) = \frac{In_{top} + In_{bot}}{area \times 2 \rho_B} \times 100\% on the topmost layer of Bi$_2$Se$_3$ to be \sim 0.2% for both 5 QL and 30QL Bi$_2$Se$_3$. The factor 2 in the denominator is because there are two Bi layers in a QL. Additionally, ARPES confirms the topological nature of the films grown on BIS-BL. Figure 3.5 compares the ARPES data for a 30QL buffer layer-based Bi$_2$Se$_3$ and a 50QL Bi$_2$Se$_3$ grown on a sapphire substrate, where topological
surface states are visibly clear for both films which is a direct indication of the TI nature of these films. The film grown on sapphire, however, shows the 2DFGs which is the result of the downward band-bending. From these images, we can also compare the location of the surface Fermi levels ($E_F$) in both films. The $E_F \approx 0.33$ eV above the Dirac point for the film grown on Al$_2$O$_3$(0001). In contrast, for the buffer layer-based Bi$_2$Se$_3$ film, $E_F \approx 0.17$ eV above the Dirac point, suggesting that only the topological surface state bands are occupied. This also shows that the band-bending is upward and the 2DFGs are absent in this film, and thus a true bulk-insulating TI is achieved.

The significant elimination of the interfacial defects in these films can also be seen in transport measurement which is the most sensitive technique to measure the concentration of defects. The Hall resistance data, including the sheet carrier-density ($n_{\text{sheet}}$) and carrier mobility ($\mu$), of an 8 QL-thick Bi$_2$Se$_3$ film grown on BIS-BL are compared to the conventionally grown films on Al$_2$O$_3$ (0001) and Si (111). Figure 3.6 exhibits $n_{\text{sheet}} \approx 1 - 3 \times 10^{12}$ cm$^{-2}$ for the entire thickness range of...
5 to 60 QL (i.e. no significant thickness-dependence) which is about an order of magnitude lower than the \( n_{\text{sheet}} \) of the films grown on Al\(_2\)O\(_3\)(0001) or Si (111). The mobility of the buffer layer-based Bi\(_2\)Se\(_3\) film is about an order of magnitude higher than any previously obtained values, reaching as high as \( \mu \approx 16,000 \text{ cm}^2/\text{Vs} \) for a 50 QL-thick film (Figure 3.6).

Moreover, it is crucial to stabilize the properties of these TI thin-films with a proper capping layer which will be discussed in more detail later in the chapter. In the next chapter, we will explain how these films with the use of MoO\(_3\)/Se capping layer allowed the observation of the quantum Hall effect for the first time in a pure binary Bi\(_2\)Se\(_3\) system.

### 3.3.2 Sb\(_2\)Te\(_3\) films on In\(_2\)Se\(_3\)/Sb\(_{0.65}\)In\(_{0.35}\)Te\(_3\) buffer layer: growth and transport properties

Similar buffer layer scheme can be extended to a Te-based TI system, like Sb\(_2\)Te\(_3\), which resulted in lowest carrier-density of all time in a TI system. Sb\(_2\)Te\(_3\) has not been studied as extensively as Bi\(_2\)Se\(_3\) and was mostly incorporated inside Bi\(_2\)Se\(_3\) or Bi\(_2\)Te\(_3\) to lower the carrier density which will be discussed in the next section. Sb\(_2\)Te\(_3\) does not grow as straightforward as Bi\(_2\)Se\(_3\) and requires a slow deposition rate\(^{215}\). Similar to Bi\(_2\)Se\(_3\), the films grown directly on commercially-available substrates harbor a high level of carrier density. Therefore, it is particularly advantageous to incorporate a buffer layer between the substrate and Sb\(_2\)Te\(_3\) to grow high quality, low-carrier-density Sb\(_2\)Te\(_3\) films. Interestingly, by growing titanium- doped Sb\(_2\)Te\(_3\) on In\(_2\)Se\(_3\)/(Sb\(_{0.65}\)In\(_{0.35}\))\(_2\)Te\(_3\) buffer layer and with the use of a (Sb\(_{0.65}\)In\(_{0.35}\))\(_2\)Te\(_3\) capping layer ultra-low carrier TI films (as low as \( 1.0 \times 10^{11} \text{ cm}^{-2} \)) can be achieved.

Figure 3.7 shows a schematic of the growth procedure of the buffer layer-based Sb\(_2\)Te\(_3\) film. After properly cleaning and outgassing the 10 mm \( \times \) 10 mm sapphire (Al\(_2\)O\(_3\)) (0001) substrate (as was explained before), the substrate was cooled down to 170 °C for deposition of 3 quintuple layer (QL, 1QL \( \approx \) 1 nm) Bi\(_2\)Se\(_3\) which serves as a template for the following layer of 20QL In\(_2\)Se\(_3\) to be deposited at 300 °C. The whole Bi\(_2\)Se\(_3\)/In\(_2\)Se\(_3\) heterostructure was then heated to 600 °C so that
Bi$_2$Se$_3$ evaporated and diffused out of the In$_2$Se$_3$ layer, leaving behind a highly insulating In$_2$Se$_3$ layer. Up to this step, the procedure is similar to the buffer layer-based Bi$_2$Se$_3$ growth. To obtain ultralow defect density Sb$_2$Te$_3$ thin-films, we next grew a more closely lattice-matched buffer layer of (Sb$_{1-x}$In$_x$)$_2$Te$_3$, a solid solution of Sb$_2$Te$_3$ and In$_2$Te$_3$. However, since In$_2$Te$_3$ has a different crystal structure (defect zinc blende lattice with a lattice constant of 6.15 Å) than Sb$_2$Te$_3$, the growth structure degrades above a certain In concentration$^{216}$. Therefore, an optimized amount of In has to be added so that the structure remains crystalline and yet far from the metallic Sb$_2$Te$_3$ and as insulating as possible. We noticed that above 40% In, the structure starts to degrade with visible formation of 3D spots in the RHEED image. Therefore, due to sample to sample variation and to be safe, we decided to use Sb$_{0.65}$In$_{0.35}$Te$_3$ for the buffer layer in all samples to stay slightly lower than 40% where the structure is on the verge of becoming bad. Thus, a 15QL-thick Sb$_{0.65}$In$_{0.35}$Te$_3$ layer was deposited at 260 °C followed by depositing the (Ti-doped) Sb$_2$Te$_3$ layer at the same temperature. We also deposited another 15QL Sb$_{0.65}$In$_{0.35}$Te$_3$ at 260 °C as a capping layer for protection against air and preserving the symmetry of top and bottom surfaces. The first panel in Figure 3.7d shows HAADF-STEM of a 20QL-In$_2$Se$_3$/5QL-(Sb$_{0.65}$In$_{0.35}$)$_2$Te$_3$/5QL-Sb$_2$Te$_3$/5QL-(Sb$_{0.65}$In$_{0.35}$)$_2$Te$_3$ film (sections of this heterostructure are indicated by yellow dashed lines). The following four panels show false-color elemental mappings from energy-dispersive X-ray spectroscopy (EDS), demonstrating the presence/absence of In, Sb, Te, and Se elements in each layer. For the film growth, 99.999% pure elemental Bi, In, Se, Sb, Te, and Ti (for doping) sources were thermally evaporated using Knudsen cells. Source fluxes were measured in situ with QCM. The ratio of selenium flux to bismuth/indium flux and tellurium flux to antimony flux were maintained at above 10:1 and 5:1, respectively, as determined by QCM. For calculating the Ti flux at different temperatures, we first start by fitting the vapor pressure data as a function of temperature (Table 3.1), with a polynomial $ln(P) = a_1 + a_2T + a_3T^2$. The fitting parameters $a_1 = -59.6$, ...
Figure 3.7 Growth procedure and characterization of buffer layer-based \(\text{Sb}_2\text{Te}_3\) films. **a**, Schematic of the growth procedure of a \(\text{Sb}_2\text{Te}_3\) film on the \(\text{In}_2\text{Se}_3/(\text{Sb}_{0.65}\text{In}_{0.35})_2\text{Te}_3\) buffer layer. **b**, Streaky RHEED images for \(\text{Sb}_2\text{Te}_3\) film grown on the buffer layer before it gets capped (bottom panel) and for the capping layer which is grown at the same growth temperature (top panel) show a flat 2D growth of high-quality film and capping. SIT refers to \((\text{Sb}_{0.65}\text{In}_{0.35})_2\text{Te}_3\). **c**, Atomic force microscopy on both \(\text{Sb}_2\text{Te}_3\) film with no capping (bottom panel) and on 2 nm-thick \((\text{Sb}_{0.65}\text{In}_{0.35})_2\text{Te}_3\) capping (top panel) shows triangular features indicative of three-fold symmetry. **d**, The first panel shows HAADF-STEM for 20QL-\(\text{In}_2\text{Se}_3\)/5QL-(\(\text{Sb}_{0.65}\text{In}_{0.35})_2\text{Te}_3$/5QL-$\text{Sb}_2\text{Te}_3$/5QL-(\(\text{Sb}_{0.65}\text{In}_{0.35})_2\text{Te}_3$/ with an additional 10 nm Te capping to protect the whole super-lattice against the STEM sample preparation processes. Yellow dashed lines drawn at each interface are guides to the eye. The rightward four panels show the elemental mapping EDS images for In (light blue), Sb (green), Te (dark blue), and Se (red). If a specific element is present in a layer, then the corresponding color appears bright in that section; in its absence, the layer appears dark. For instance, \(\text{Sb}_2\text{Te}_3\) contains Te and thus appears as bright blue in the fourth panel, but it lacks Se and thus appears as dark in the fifth panel. Adapted from Reference 155.

\[
a_2 = 0.051, \quad a_3 = -1.11 \times 10^{-5}
\]
can thus be extracted. Additionally for an ideal gas, it can be shown that
\[
\Phi_i = \Phi_j \cdot \frac{P_i}{P_j} \cdot \frac{e^{a_2 T_i + a_3 T_i^2}}{e^{a_2 T_j + a_3 T_j^2}} \cdot \frac{T_j}{T_i},
\]
where \(\Phi_i\) and \(\Phi_j\) are source fluxes (atoms/cm\(^2\)/s), \(P_i\) and \(P_j\) are vapor pressures of the source (Torr) for the source temperatures, \(T_i\) and \(T_j\) (K), respectively. Therefore, by knowing the flux \(\Phi_j\) at a specific temperature (usually it is measured at higher temperatures where you have reasonably measurable flux by the QCM system,
Φ_i can be calculated. Here, for each doped sample, we systematically vary Ti flux (temperature) while keeping Sb at a constant flux/temperature. The doping level is based on \( \frac{\Phi_{\text{Ti}}}{\Phi_{\text{Ti}} + \Phi_{\text{Sb}}} \times 100\% \).

The \( \Phi_{\text{Sb}} \approx 1.2 \times 10^{13} \text{ cm}^{-2} \text{s}^{-1} \) at 395 °C and the \( \Phi_{\text{Ti}} \approx 6 \times 10^{12} \text{ cm}^{-2} \text{s}^{-1} \) at 1550 °C (Table 3.2). As always, Se shutter was opened few seconds prior to the Bi shutter for the growth of the initial Bi\(_2\)Se\(_3\) layer and eventually was closed before starting the Sb\(_{0.65}\)In\(_{0.35}\)Te\(_3\) deposition, after which the Te shutter was opened immediately. The In and Sb shutters were opened at the same time 3 to 4 seconds after Te shutter opening. Te shutter remained open all the time during the Te-based layers until it got closed at around 240 °C upon cooling down.

### Table 3.1 Ti vapor pressure as a function of temperature. The values in the table can be taken from Veeco or on line. Adapted from Reference 155.

<table>
<thead>
<tr>
<th>Pressure (Torr)</th>
<th>10(^{-8})</th>
<th>10(^{-7})</th>
<th>10(^{-6})</th>
<th>10(^{-5})</th>
<th>10(^{-4})</th>
<th>10(^{-3})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature (°C)</td>
<td>1062</td>
<td>1137</td>
<td>1227</td>
<td>1327</td>
<td>1442</td>
<td>1577</td>
</tr>
</tbody>
</table>

### Table 3.2 Some of the Ti fluxes at lower temperatures along with corresponding vapor pressure and calculated doping used for the growth of ultralow carrier density Sb\(_2\)Te\(_3\) films. Adapted from Reference 155.

<table>
<thead>
<tr>
<th>Ti Temperature (°C)</th>
<th>Vapor pressure (μTorr)</th>
<th>Flux (10(^{10}) cm(^{-2}) s(^{-1}))</th>
<th>Doping level (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1325</td>
<td>9.26</td>
<td>8.95</td>
<td>0.7</td>
</tr>
<tr>
<td>1340</td>
<td>12.7</td>
<td>12.3</td>
<td>1.0</td>
</tr>
<tr>
<td>1344</td>
<td>13.9</td>
<td>13.3</td>
<td>1.1</td>
</tr>
<tr>
<td>1350</td>
<td>15.7</td>
<td>15.1</td>
<td>1.3</td>
</tr>
<tr>
<td>1360</td>
<td>19.4</td>
<td>18.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

With the use of the buffer layer and Ti doping, we were able to develop TI Sb\(_2\)Te\(_3\) thin-films with carrier densities \( n_{\text{sheet}} \) as low as \( 1.0 \times 10^{11} \text{ cm}^2 \), the lowest reported in any TI system. As-grown Sb\(_2\)Te\(_3\) films on sapphire have high density of interfacial defects and typically show a p-type sheet carrier density of \( \sim 4.0 \times 10^{13} \text{ cm}^2 \). Grown on the optimally-designed buffer-layer, their carrier density immediately dropped to \( n_{\text{sheet}} \approx 7 \times 10^{12} \text{ cm}^2 \) (p-type). When we introduced Ti as an n-
Figure 3.8 Transport Properties of various Sb$_2$Te$_3$ films. **a**, Comparison of the lowest achieved carrier densities (solid diamonds connected by a dark red dashed line as a guide to the eye) with corresponding mobilities on the right axis (hollow diamonds connected by a blue dashed line) for Sb$_2$Te$_3$ films grown directly on sapphire (ST), buffer layer (BL)-based ST (BL-ST), and Ti-doped BL-based Sb$_2$Te$_3$ (p-type: BL Ti-ST (P) and n-type: BL Ti-ST (N)). Thickness of the film is 8QL. **b**, Sheet resistance as a function of temperature for a 30QL (Sb$_{0.65}$In$_{0.35}$)$_2$Te$_3$ sample (buffer + capping only) has insulating temperature-dependent behavior where it is not fully insulating at room temperature and becomes insulating at low temperatures. It is worth noting that a thin layer (∼10 to 20 nm) of Te capping which has been used in some of the previous studies$^{135,146}$ is not completely insulating at room temperature either but becomes highly resistive at low temperatures. **c**, Sheet resistance as a function of temperature (from room temperature down to 6 K) for four Ti-doped samples: P8-1%, P8-1.1%, N8-1.3%, and N8-1.5%. **d**, Transport properties ($n_{\text{sheet}}$ and $\mu$) of the same four samples. By adding Ti, $n_{\text{sheet}}$ decreases to ultra-low $p$ ($1.8 \times 10^{11} \text{ cm}^{-2}$) in sample P8-1%. As more Ti is added and as the Fermi level gets closer to the Dirac point (mixed n- and p-zone), the $n_{\text{sheet}}$ artificially increases to $1.3 \times 10^{12} \text{ cm}^{-2}$ in P8-1.1% and eventually additional Ti leads to an n-type sample, N8-1.3% with $n_{\text{sheet}} = -3.8 \times 10^{11} \text{ cm}^{-2}$, and even to a lower carrier density sample, N8-1.5% with $n_{\text{sheet}} = -1.4 \times 10^{11} \text{ cm}^{-2}$. Adding Ti decreases the mobility value on the p-type side, but $\mu$ increases again on the n-type side. Beyond a certain point, adding more Ti makes the system insulating. Adapted from Reference 155.
type compensation dopant (Ti$^{4+}$ replacing Sb$^{3+}$), $n_{\text{sheet}}$ further dropped to as low as p-type of $1.0 \times 10^{11}$ cm$^{-2}$ and n-type of $1.4 \times 10^{11}$ cm$^{-2}$ (Figure 3.8). These are over ten times lower than the previous best values observed in (Bi$_{1-x}$Sb$_x$)$_2$Te$_3$ thin-films$^{135}$, which showed p-type of $1.4 \times 10^{12}$ cm$^{-2}$ and n-type of $1.5 \times 10^{12}$ cm$^{-2}$, and also much lower than the best Bi$_2$Se$_3$ thin-films$^{217}$ grown on optimal buffer layers, which showed p-type of $2 \times 10^{12}$ cm$^{-2}$ and n-type of $6 \times 10^{11}$ cm$^{-2}$. Furthermore, Ti-doping to convert Sb$_2$Te$_3$ thin-films all the way into n-type is on its own surprising considering that all such attempts have previously failed for bulk crystals of Sb$_2$Te$_3$.$^{218-220}$ All the films were in situ capped with several layers of (Sb$_{0.65}$In$_{0.35}$)$_2$Te$_3$ to protect against aging and preserve symmetry between the top and bottom surfaces of the Ti film.

Figure 3.8 shows the transport properties of ungated BL-based 8QL-thick Sb$_2$Te$_3$ samples having Ti counter-dopant concentrations of 1.0% (ultra-low p-type; for ease we name it P8-1.0%) where P refers to p-type sample, 8 is the thickness in quintuple layers, and 1.0% is the Ti concentration), 1.1% (p-type but closer to the n-p mixed regime (near the Dirac point); named P8-1.1%), 1.3% (very low n-type; named N8-1.3%), and finally 1.5% (ultra-low n-type; named N8-1.5%).

### 3.4 Topological surface states-originated response$^4$

In this section, we discuss the various measurements that were carried out to investigate the TSS-originated conduction. This section is based on the data from Bi$_2$Se$_3$ films grown on BIS-BL. The Hall Effect of these films shows non-linearity at fields above $\sim0.5$ T, which usually is an evidence for existence of multiple conduction channels with different mobilities. Figure 3.9a shows the data from the Hall measurements on 5, 25 and 60 QL films up to 9 T. Other than the 5 QL film, which shows a weak non-linearity, all the other films show clear non-linearity like the 25QL and 60 QL

$^4$Section 3.4 is reproduced (with some changes) with permission from Koirala, N.; Brahlek, M.; Salehi, M.; et al. *Nano Lett.*, **2015**, *15*(12), 8245–8249. ©2015 American Chemical Society with permission.
Figure 3.9 The two-carrier model fit to the nonlinear Hall effect for Bi$_2$Se$_3$ films grown BIS-BL. a, The Hall effect shows non-linearity for the 5, 25 and 50 QL Bi$_2$Se$_3$ films. The inset shows the two-carrier model fit for the 25 QL sample. b, The sheet carrier densities and c, the mobilities obtained from the two-carrier fit as a function of the film thickness. The low-field Hall effect data is also shown for comparison. The inset demonstrates the zero-field sheet resistance versus film thickness. Adapted from Reference 144.

Films. The sheet carrier density extracted from the low-field Hall slope yields a mobility-weighted-average of different carrier species. In order to identify the sheet carrier density and mobility of each species we can use the two-carrier model (Equation 3.1) to fit the Hall Effect data:

$$R_{H_{Hall}}(B) = -\frac{B}{e} \left[ \frac{(n_1 \mu_1^2 + n_2 \mu_2^2) + B^2 \mu_1^3 \mu_2^3 (n_1 + n_2)}{(n_1 \mu_1 + n_2 \mu_2)^2 + B^2 \mu_1^2 \mu_2^2 (n_1 + n_2)^2} \right]$$  (3.1)

where $R_{H_{Hall}}(B)$ is the Hall resistance, $B$ is the applied magnetic field, $e$ is the electronic charge and $n_i$ and $\mu_i$ are the sheet carrier density and mobility, respectively, of the $i^{th}$ species with $i = 1, 2$. $n_i$ and $\mu_i$ are the fitting parameters. By fixing $\frac{R_{H_{Hall}}(0)}{B}$ to the linear low-field Hall slope, and by using the zero-field sheet resistance ($R_{sheet} = \frac{1}{e(n_1 \mu_1 + n_2 \mu_2)}$), the number of independent fitting parameters are reduce to two parameters.

Figure 3.9b shows the sheet carrier densities ($n_1$ and $n_2$) obtained from the fit as well as the sheet carrier density ($n_{LOW}$) obtained from the low-field Hall effect. $n_1$ and $n_{LOW}$ show a
negligible thickness dependence. \( n_2 \) shows sample to sample variation but remains within the range of \(-2 \times 4 \times 10^{12} \) cm\(^{-2}\) for films thicker than 15 QL with no clear thickness dependence. Figure 3.9c shows the corresponding mobilities \( \mu_1 \) and \( \mu_2 \) for the two channels as well as the low field mobility \( \mu_{LOW} \). \( \mu_1 \) is comparable but consistently higher than \( \mu_{LOW} \) which is expected for \( \mu_{LOW} \) that gives the weighted mobility of both channels. The fact that \( n_{LOW} \) and \( \mu_{LOW} \) are close to \( n_1 \) and \( \mu_1 \) implies that the conduction is dominated by this high mobility channel.

### 3.4.1 ARPES measurements

As mentioned in the previous part, the low-field carrier density \( (n_{LOW}) \) found to be smaller than \(-2 \times 10^{12} \) cm\(^{-2}\) for the entire thickness range. Two-carrier model fitting also yields a total sheet carrier density \( (n_{tot} = n_1 + n_2) \) to be at most \(-5 \times 10^{12} \) cm\(^{-2}\) which is smaller than the critical \( n_{ss} \sim 1 \times 10^{13} \) cm\(^{-2}\), suggesting that the band-bending for films grown on BIS-BL is upward with a depletion region being formed. Because of the upward band-bending, 2DFGs cannot be present. Therefore, it can be concluded that the observed channels with thickness independent sheet carrier density are originated from the TSSs. This can also be seen in the ARPES data (Figure 3.5), where the surface \( E_F \) lies in the bulk band-gap, and in contrast to the Bi\(_2\)Se\(_3\) film grown directly on Al\(_2\)O\(_3\), no 2DFGs are observed. The absence of 2DFG states can also be validated by estimating the expected sheet carrier density of TSS based on the assumption that either \( n_1 \) or \( n_2 \) originates from 2DFGs. If \( n_1 \approx 1.8 \times 10^{12} \) cm\(^{-2}\) is due to 2DFG carriers, then the Fermi wave vector for 2DFG \( k_{F,2DFG} = \sqrt{2\pi n_1} \) is estimated to be 0.034 Å. Using the ARPES data for the Bi\(_2\)Se\(_3\) film grown on Al\(_2\)O\(_3\), the Fermi wave vector of the corresponding TSS \( (k_{F,TSS}) \) can be extrapolated at this \( k_{F,2DFG} \), which yields \( k_{F,TSS} \approx 0.088 \) Å. From \( k_{F,TSS} \), the \( n_{sheet,TSS} = \frac{k_{F,TSS}^2}{4\pi} \) can be calculated to be \(-6.1 \times 10^{12} \) cm\(^{-2}\) for the corresponding TSS. The same calculations can be done for \( n_2 \approx 3 \times 10^{12} \) cm\(^{-2}\) by assuming it is 2DFG-originated, which gives \( n_{sheet,TSS} \approx 6.8 \times 10^{12} \) cm\(^{-2}\). This means that the combined sheet
carrier density from TSS and 2DFG is $\sim 7.9 \times 10^{12} \text{ cm}^{-2} \sim (9.8 \times 10^{12} \text{ cm}^{-2})$ from a single surface if $n_1 (n_2)$ is assumed to be originating from 2DFG.

For simplicity, if we assume the same carrier density for the other surface, then the total carrier density will be larger than $\sim 10^{13} \text{ cm}^{-2}$. This is more than three times higher than the observed carrier density from the Hall effect which confirms the absence of 2DFGs. Therefore, the two conducting channels can be associated to the top and bottom TSSs, respectively. Also, the fact that the MoO$_3$ and Se capping layers, which are deposited on the top surface, drastically reduce the mobilities from the uncapped value could suggest that the higher mobility channel comes from the top TSS. Additionally, we can compare the $E_F$ obtained from the Hall effect to the $E_F \approx 0.17 \text{ eV}$ obtained from the ARPES of the 30 QL Bi$_2$Se$_3$ grown on BIS-BL with the corresponding Fermi wave-vector $k_F \approx 0.052 \text{ Å}^{-1}$. The $E_F$ from the Hall effect can be calculated using $k_{F,\text{Hall}} = \sqrt{4\pi n_1}$, where $n_1 \approx 1.8 \times 10^{12}$ is obtained from the two-carrier model fit, yielding $k_{F,\text{Hall}} \approx 0.0475 \text{ Å}^{-1}$. Using the ARPES data, with this $k_F$, the $E_F \approx 0.14 \text{ eV}$ which is slightly smaller than $E_F \approx 0.17 \text{ eV}$ from
ARPES data. Such a difference could be due to the sample-to-sample variation or because of the Se de-capping process at ~250 °C required for doing ARPES measurement which could potentially create more Se vacancies, thereby pushing the Fermi level higher.

### 3.4.2 THz measurements

Time-domain terahertz spectroscopy (TDTS) is a non-destructive measurement technique that uses a transmission geometry to measure the complex transmission function, from which the complex conductivity of a material can be measured. From these measurements, data such as carrier density and mobility of the material can be extracted, and thus TDTS can be used as a complement to DC transport measurements. TDTS provides useful information about the electrodynamics of TIs. It can be exploited to study the quantum phase transition between a topological insulator and a normal insulator. More importantly, time-domain magneto-terahertz spectroscopy (TDTMS) can be used to distinguish the bulk/2DEG and topological surface state contributions and to extract cyclotron resonances from topological surface states in bulk-insulating TIs. In the previous parts, we showed that both transport measurement as well as the ARPES data confirm the TSS-originated conduction, measuring the cyclotron mass is, however, the smoking gun for showing that the response is TSS-originated.

Cyclotron resonance (CR) can be extracted from the complex Faraday rotation measurements in TDMTS. For measuring the complex Faraday rotation (FR), a phase modulation technique is used to accurately measure the polarization states which allows the measurement of $E_{xx}(t)$ and $E_{xy}(t)$ simultaneously in a single scan. Faraday rotation can be obtained by $\theta_F = \arctan(E_{xy}(\omega)/E_{xx}(\omega)) = \theta F' + i \theta F''$ after Fourier transforming into the frequency domain. Sapphire as well as the MoO$_3$ and Se capping layers do not show FR. The dip position in the imaginary part corresponds to the CR. Field dependent imaginary parts of the complex FR are shown in Figure 3.11 for Se-capped and MoO$_3$-capped films. The negative Faraday
rotation means that the carriers are electrons in both films. The data can be fit by the Drude-Lorentz model with a Drude term, a phonon term and a term for the background dielectric constant ($\varepsilon_\infty$) coming from higher energy absorptions.

Then, the conductance in magnetic field is:

$$G_\pm = -i\varepsilon_o \omega d \left[ \frac{\omega_p^2}{-\omega^2 - i\Gamma_p \omega + \omega_c \omega} - \frac{i\omega\omega_p^2}{\omega_{DL}^2 - \omega^2 - i\omega\Gamma_{DL}} + \varepsilon_\infty - 1 \right], \quad (3.2)$$

where $\omega_p$ and $\Gamma_p$ are Drude plasma frequency and scattering rate, $\omega_{pDL}$, $\omega_{DL}$, and $\Gamma_{DL}$ are, respectively, the plasma and center frequencies, and scattering rate of optical phonons, $d$ is the film thickness and the $\pm$ sign corresponds to the right/left circularly polarized light, respectively. The parameters of the phonon and the high-frequency terms are constrained by those extracted from the zero-field conductance and only the cyclotron frequency ($\omega_c$) and the scattering rate are allowed to vary. From $G_\pm$, the complex transmission for the right and left circularly polarized light $t_\pm$ can be calculated. Then we can calculate the complex FR by $\tan(\theta) = -i(t_+ - t_-)/(t_+ + t_-)$. The zero-field real conductance data is also fit by an oscillator model with a Drude term describing free electron-like motion, a Drude-Lorentz term modeling the phonon and a lattice polarizability ($\varepsilon_\infty$) term that originates from absorptions outside the spectral range.

$$G(\omega) = \left[ -\frac{\omega_p^2}{i\omega \Gamma_p} - \frac{i\omega\omega_p^2}{\omega_{DL}^2 - \omega^2 - i\omega\Gamma_{DL}} - i(\varepsilon_\infty - 1)\omega \right] \varepsilon_o d, \quad (3.3)$$

Figure 3.11 shows the real and imaginary part of Faraday rotation for a 16QL film with 20nm Se-capping and another 16QL film with 50nm MoO$_3$-capping. When thinner films with MoO$_3$ capping were measured at higher fields, they showed the quantized Faraday rotation which will be discussed in the next chapter.

The samples were measured at the Johns Hopkins University one day after the growth at Rutgers University. Sharp CR features were observed in measurements of the complex Faraday
rotation angle. The inflection point of real part of the FR corresponds to the CR and shifts to higher frequencies as \(B\) increases. As shown in Figure 3.11 for both samples, fitting the FR data gives the CR frequency (\(\omega_c\)) for each \(B\), based on which the effective mass \(m^*\) is obtained by using a linear fit \(\omega_c = \frac{eB}{2\pi m^*}\). To get the total carrier density \(n_{tot}\) and the mobility \(\mu_{optical}\), the scattering rate (\(\Gamma_D\)) and the spectral weight (\(\omega_p^2 d\)) need to be extracted from the Drude-Lorentz fit of the zero-field real conductance (\(\omega_p^2 d\) is proportional to the total area under the real conductance at zero-field \(G_{D1}\)) shown in Figure 3.11. We can then get \(\mu_{optical} = \frac{e}{2\pi \Gamma_D m^*}\) and \(n_{tot}\) as well as the Fermi wave-vector (\(k_F\)) if it is assumed the carriers populate the TSS from the following equation:

\[
\frac{2}{\pi \varepsilon_o} \int G_{D1} d \omega = \omega_p^2 d = \frac{n_{sheet} e^2}{m^* \varepsilon_o} = \frac{k_F (A + 2Bk_F)e^2}{2\pi \hbar^2 \varepsilon_o}
\]

(3.4)

where \(e\) is the electronic charge, \(\varepsilon_o\) is the free-space permittivity, and \(A = 2.02\) eV\(\cdot\)Å and \(B = 10.44\) eV\(\cdot\)Å\(^2\) are the TSS band parameters up to quadratic term (i.e. \(E_{TSS} = Ak + Bk^2\)).

The carrier density of each topological surface state (TSS) can be calculated based on
\(n_{TSS} = \frac{k_F^2}{2\pi}\). An effective mass \(m_{TSS} = \hbar k_F / v_F\) can be defined for massless Dirac fermions where the Fermi velocity of TSS is determined by \(v_F = \partial E_F / \hbar \partial k\). Table 1 lists \(\omega_p^2 d\), \(m^*\), \(m_{TSS}\), \(n_{tot}\), \(n_{TSS}\), \(\mu_{optical}\), Hall carrier density (\(n_{sheet}\)) and Hall mobility (\(\mu_{Hall}\)) for the two samples. In addition, the Hall data for a fresh uncapped 15 QL film is also shown for comparison. For calculation of \(\mu_{optical}\), \(\Gamma_D\) of \(\sim 0.5\) THz and \(\sim 1\) THz were used for Se-capped and MoO\(_3\)-capped films, respectively. The capping layers introduce additional surface-scattering sites, and thus the mobilities for the capped samples are lower as compared to uncapped ones (in the case of the Se capping, additional carriers as well).

From Table 3.3 we can see a reduced carrier density for MoO\(_3\)-capped film as compared to the Se-capped one because MoO\(_3\) depletes (n-type) carriers from Bi\(_2\)Se\(_3\) due to its higher electron affinity. The smaller \(m^* \approx 0.073m_e\) in MoO\(_3\)-capped sample compared to \(m^* \approx 0.12m_e\) for Se-
capped sample is an evidence that CR comes from the TSS carriers because the effective masses of bulk or 2DFG carriers are carrier density independent (0.11 ~ 0.13m_e, m_e is the bare electron mass)\textsuperscript{166, 173}, while the m* of Dirac-like carriers of TSS scales with carrier density (m_{TSS} \propto k_F \propto \sqrt{n_{TSS}}).

Figure 3.11 Faraday rotation at different magnetic fields, Cyclotron resonance, and zero-field real conductance for two 16 QL thick Bi_2Se_3 films grown on BIS-BL and capped by 20 nm Se and 50 nm MoO_3. Real part of the Faraday rotation (FR) at different magnetic fields for a, Se-capped and b, MoO_3-capped film. The dip in the imaginary part of the FR that shifts to higher frequencies as magnetic field is increasing, indicating the cyclotron resonance in c, 20 nm Se-capped film and d, 50 nm MoO_3-capped film. The dashed arrows in a-d show the direction of increasing magnetic field with 0.5 T step. e, The CR frequencies at different magnetic fields for same films. Solid lines are linear fit of \omega_c = \frac{eB}{2\pi m^*}. f, The zero-field real conductance as a function of frequency along with Drude-Lorentz fit for both films. The peaks near 1.9 THz correspond to bulk phonon mode of Bi_2Se_3. Adapted from Reference 144.

The agreement between the measured m* and the calculated m_{TSS} shows that CR is originating from TSS carriers. Additionally, the agreement between n_{tot}, n_{TSS}, and n_{sheet} confirms that total observed carriers in these films are TSS-originated.

### Table 3.3 Optical and DC transport parameters of two 16 QL thick Bi$_2$Se$_3$ films grown on BIS-BL with Se and MoO$_3$ capping layers along with a 15 QL uncapped film. Adapted from Reference 144.

<table>
<thead>
<tr>
<th>film</th>
<th>$\omega_p d$ (THz nm)</th>
<th>$m^*/m_e$</th>
<th>$m_{FSS}/m_e$</th>
<th>$n_{tot}$ $(10^{12}$ cm$^{-2}$)</th>
<th>$n_{TSS}$ $(10^{12}$ cm$^{-2}$)</th>
<th>$\mu_{optical}$ (cm$^2$ V$^{-1}$ s$^{-1}$)</th>
<th>$\mu_{Hall}$ (cm$^2$ V$^{-1}$ s$^{-1}$)</th>
<th>$n_{sheet}$ $(10^{12}$ cm$^{-2}$)</th>
<th>$\mu_{Hall}$ (cm$^2$ V$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Se capped</td>
<td>2.3 ± 0.1 x 10$^4$</td>
<td>0.12 ± 0.01</td>
<td>0.12 ± 0.01</td>
<td>3.5 ± 0.2</td>
<td>3.4 ± 0.2</td>
<td>4,700</td>
<td>3.1</td>
<td>3,600</td>
<td></td>
</tr>
<tr>
<td>MoO$_3$-capped</td>
<td>1.2 ± 0.1 x 10$^4$</td>
<td>0.073 ± 0.007</td>
<td>0.080 ± 0.005</td>
<td>1.1 ± 0.2</td>
<td>1.2 ± 0.2</td>
<td>3,800</td>
<td>1.1</td>
<td>3,300</td>
<td></td>
</tr>
<tr>
<td>Uncapped</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.3</td>
<td>11,400</td>
<td></td>
</tr>
</tbody>
</table>

3.5 Defect suppression via impurity addition$^5$

Besides interface-engineering, other growth techniques, such as compensation-doping and alloying have been implemented in the growth of both TI thin-films and bulk crystals to resolve the bulk conduction issue and suppress the defects, which is the point of discussion.

3.5.1 Compensation-Doping

One of the ways to eliminate defects is through compensation-doping$^{143, 228-231}$. In bulk Bi$_2$Se$_3$ crystals, compensation doping with group II elements, such as Ca has been used to lower the carrier-density and to even tune the sample from n- to p-type$^{143, 230-231}$. However, such doping has been difficult to implement in MBE-grown Bi$_2$Se$_3$ thin-films. Only recently with the use of the thin-film and a proper capping layer, a systematic and reliable carrier tuning has successfully been achieved, and finally p-type Bi$_2$Se$_3$ thin-films were obtained$^{217}$. Previous failures of counter-doping for Bi$_2$Se$_3$
Figure 3.12 Transport properties of Ca-doped Bi$_2$Se$_3$ films. a, The table summarizes the effectiveness of the Ca counter-doping for different thicknesses and various growth structures in the presence and absence of a buffer layer/capping layer. b, The left panel shows the carrier densities and the right panel shows the corresponding mobilities as a function of Ca concentration (%) for films grown on the BIS-BL and capped by MoO$_3$/Se with a fixed thickness of 20QL. Upon adding more Ca, the carrier changes from n- to p-type and finally back to n-type. The mobility decreases first sharply right around the n-to-p crossing and then gradually as the Ca concentration further increases. The structure degrades upon adding higher level of Ca, where the carrier type changes back to n-type. Adapted from Reference 217.

Thin-films were because of 1. neutralization/oxidation of the counter-dopants due to air exposure which we resolve with the aid of a proper capping layer, and 2. high density of interfacial defects which can be suppressed by incorporating a compatible buffer layer. Moon et al.\textsuperscript{217} showed that, as the Ca doping level in the buffer layer-based Bi$_2$Se$_3$ films capped with MoO$_3$/Se increases, it compensates for the intrinsic n-type carriers and the (n-type) sheet carrier-density gradually decreases (increasing negative Hall slope). Upon adding more Ca, the slope changes from negative (n-type) to positive (p-type), passing through a non-linear n-p mixed regime. As the films get thinner, higher level of Ca doping is required to reach the n-p mixed or p regimes, suggesting higher ratio of interfacial to bulk defects even in these interface-engineered films. Nonetheless, p-type Bi$_2$Se$_3$ films are achieved for all thickness range of 50 QL down to 6 QL. For thick films, as long as they are capped, p-type samples can be achieved with or without the buffer-layer (Figure 3.12).
Figure 3.13 Transport properties of the Cu doped Bi$_2$Se$_3$ grown on Al$_2$O$_3$ (0001). a, The sheet carrier density (top panel), extracted from low-field Hall measurement, as a function of Cu doping concentration ($x$ in %) show a minimum at $x \approx 2.5\%$ to $4.0\%$. The mobility (bottom panel) versus $x$ (%) has the overall monotonic decrease behavior as the doping is increased (the black guide line) but shows a peak near the optimal copper doping which is marked by a red guide line. b, The left panel shows the effective number of 2D conducting channel $\tilde{A}$ versus the composition ($x = \frac{\text{Cu}}{\text{Bi}} \times 100\%$) for 20QL-thick films, and the right panel shows $\tilde{A}$ as a function of thickness for a fixed composition ($x \approx 2.5$ to $4.0\%$). Adapted from Reference 27.

The films become n-type again with further increase in Ca doping, suggesting that, beyond a certain concentration, the compensation dopants start to act as n instead of p-type dopants, except for the 6 QL-thick sample which degrades and becomes insulating probably due to a disorder-driven topological phase transition$^{232}$. For more information regarding the Ca-doped Bi$_2$Se$_3$ films see Reference 217. Before this work, with the use of ex situ Ca ion implantation in Bi$_2$Se$_3$ films followed by an annealing process, Sharma et. al$^{233}$ were able to change the carrier type, although the sheet carrier density always remained high most likely due to Se vacancies induced by the ion beam.

Furthermore, pure p-regime shows higher carrier-density than the pure n-regime (lowest p-type carrier $\sim 1.5 \times 10^{12}$ cm$^{-2}$ vs. lowest n-type carrier: $\sim 6.0 \times 10^{11}$ cm$^{-2}$). Also, the mobility sharply decreases as soon as the majority carrier type changes from n to p-type which is very likely due to
the nature of Bi$_2$Se$_3$ band structure where the Dirac point is very close to the bulk valence band and the surface band broadens noticeably on the p-side (Figure 3.12).

Interestingly, Cu doping was found to work as a compensation dopant for Bi$_2$Se$_3$ thin-films (grown on sapphire substrate). Cu doping had previously been shown to induce superconductivity in bulk Bi$_2$Se$_3$ where the increase of the carrier density upon Cu doping at low temperatures was explained as one of the possible reasons for superconducting transition at 3K.\textsuperscript{234-235}

However, in thin-films, Cu addition functioned more like a p-type dopant, where an optimal Cu addition of \( \sim 2\% \) lowered the sheet carrier-density from \( \sim 3 \times 10^{13} \text{cm}^{-2} \) to \( \sim 5 \times 10^{12} \text{cm}^{-2} \), indicating much lower \( E_F \) compared to pure Bi$_2$Se$_3$ (Figure 3.13). Subsequently, the WAL effect, for the first time, resulted in two conducting channels (Figure 3.13) in these films (when thicker than \( \sim 10-20QL \)).\textsuperscript{27} Furthermore, a combination of the Hall effect, WAL, SdH and ARPES\textsuperscript{27} and terahertz measurements\textsuperscript{223} showed that the conduction arose from decoupled TSSs, confirming that a true bulk-insulating TI was indeed realized. In these films, majority of Cu dopants were identified to be electrically neutral, and therefore, the decrease in carrier-density could be attributed to Cu reducing the Se vacancy defects and/or somehow alleviating the interfacial defects. For more information regarding Cu-doped Bi$_2$Se$_3$ films see Reference 27.

### 3.5.2 Ternary and quaternary compounds and isovalent alloying

Another way to suppress defects in topological materials is to grow ternary (\( A_xB_{2-x}C_3 \) or \( A_2C_{3-x}D_x \), where A and B are pnictogens and C and D are chalcogens)\textsuperscript{135, 139, 141} or quaternary compounds (\( A_2B_xC_3D_{3-y} \)).\textsuperscript{141, 143, 236-237} Such isovalent alloys, derived from Bi$_2$Se$_3$, Sb$_2$Te$_3$ and Bi$_2$Te$_3$, have resulted in major advances in the quality of topological materials. As Sb$_2$Se$_3$ is topologically trivial, care must be taken to not destroy the topological nature of the material so that the surface states remain intact. These compounds are useful in tuning the defect chemistry of the material, which can lead to a decrease in carrier density and a suppressed bulk conduction. Additionally, application
of electrostatic-gating helps with further depletion of carriers and lowering the $E_F$ towards the Dirac point. Early on, by alloying Bi$_2$Se$_3$ with Sb, bulk crystals with low carrier-density of $n_{3d} \approx 2$ to $3 \times 10^{16}$ cm$^{-3}$ were obtained leading to detection of quantum oscillations originating from TSS. Although Sb does not directly deplete the n-type carriers due to its iso-valency with Bi, it helps lower the carrier-density by reducing the Se vacancies.

Additionally, growth of these compounds in some cases have led to observation of bulk-insulating phase. For example, in Bi$_x$Sb$_{2-x}$Te$_3^{135}$, Sb addition to Bi$_2$Te$_3$ pulls the Dirac point from beneath the bulk valence band into the bulk band-gap. Because Bi$_2$Te$_3$ is usually n-type doped and Sb$_2$Te$_3$ is intrinsically p-type, the $E_F$ in the solid solution can be tuned into the bulk band-gap for the optimal range of $0.75 < x < 0.96$ with the lowest sheet carrier-density of $n_{2d} \approx 1 \times 10^{12}$ cm$^{-2}$ achieved for 0.96 Sb concentration$^{135}$. The carrier sign change has been reported in other works on Bi$_x$Sb$_{2-x}$Te$_3^{172, 238-240}$. This material has been grown in bulk crystal form$^{124}$ as well.

### 3.6 Stabilizing the properties of TI films: capping layer matters$^6$

As mentioned in earlier section, the surface Fermi level of the conventionally-grown Bi$_2$Se$_3$ thin-films resides above the bottom of bulk conduction band and their sheet carrier densities are generally much higher than $\sim 1 \times 10^{13}$ cm$^{-2}$ with the bulk being metallic. Due to the initial high carrier densities, the carrier densities remain relatively stable in air with the change in carrier densities limited only to a few tens of percent$^{165, 212}$. The relative stability of these high-carrier-density samples has helped attain numerous consistent results across different probing techniques even without any special care to protect the surface layers$^{17, 241}$.

In contrast, recently developed low-carrier-density bulk-insulating Bi$_2$Se$_3$ films$^{27, 144}$ exhibit quite significant aging effect in air. The as-grown sheet carrier densities of these samples

---

$^6$Section 3.6 is based on work Salehi, M.; et al. *APL Materials* 2015, 3, 091101.
are about an order of magnitude lower than the previous generation samples. This order of magnitude decrease in carrier density implies that the few tens of percent change in carrier density of the previous generation films in air would amount to a few hundreds of percent in these new films, if we assume that similar absolute level of aging effect occurs for both types of samples. Moreover, in order to ensure the bulk insulating character of these low-carrier-density samples, it is critical to keep the sheet carrier densities substantially below \(1 \times 10^{13}\) cm\(^{-2}\) to maintain the upward band-bending near the surfaces\(^{163}\). Below, we show that in order to stabilize the low-carrier-density Bi\(_2\)Se\(_3\) films, it is essential to use proper capping layers. Among the three capping layers we studied, in situ-deposited In\(_2\)Se\(_3\), in situ-deposited Se and ex situ-deposited PMMA, only Se and PMMA exhibited good stabilization effect.

The low-carrier-density Bi\(_2\)Se\(_3\) films were grown on (Bi\(_{1-x}\)In\(_x\))\(_2\)Se\(_3\) buffer layers on sapphire substrate using MBE. As for capping layers, 100 nm amorphous Se and 50 nm amorphous In\(_2\)Se\(_3\) were deposited in situ at room temperature, whereas ~200 nm PMMA was spin-coated ex situ, followed by baking process at different temperatures. Both capped and uncapped films were measured for over 280 days. Standard Hall effect measurements were carried out at room temperature in air in magnetic field up to 0.6 T with the standard van der Pauw geometry. In order to study the evolution of the transport properties over time, \(n_{sheet}\) and \(\mu\) were measured continuously for more than 36 hours immediately after the sample was taken out of the MBE chamber. After that, they were measured almost every day for 7 days (using the sample-sample method explained in the previous chapter), and then intermittently at couple-of-day intervals for the following ~9 months.

The continuous measurement of 30 QL (1QL \(\approx\) 1 nm) uncapped sample for 40 hours (Figure 3.14) shows a telling decrease in \(n_{sheet}\) (~30\% drop from the initial value) and a simultaneous increase in \(\mu\) at around hour 16, which is also observable in other uncapped samples. Similar short-
Aging effect on room temperature transport properties of 30 QL Bi$_2$Se$_3$ films with and without various capping layers. 

**Figure 3.14**

- **a**, $n_{\text{sheet}}$ and $\mu$ for an uncapped Bi$_2$Se$_3$ film over 40 hours.
- **b**, $n_{\text{sheet}}$ and $\mu$ vs. time for Bi$_2$Se$_3$ films with and without different capping layers over ~9 months.
- **d**, Absolute change in $n_{\text{sheet}}$ vs. time for Bi$_2$Se$_3$ films with different thicknesses of Se-capping over several days. Dashed curves are $C(1 - e^{-DT})$ the fittings. The inset shows the fitting parameter $C$ and $D$ for uncapped (open) and Se-capped (filled) films. Adapted from Reference 242.

Time decrease in carrier density was observed in our previous study of high-carrier-density Bi$_2$Se$_3$ samples$^{165}$. The trigger point of this decrease depends on ambient conditions, such as relative humidity and temperature. This shows that there are competing processes such that there is a short-time window in which the sample properties improve, and after that it gets worse rapidly until the carrier density reaches twice the initial value after a day. This drastic degradation necessitates an effective capping layer for stabilizing the properties of these thin-films.
Figure 3.14b summarizes the change in carrier density for different capping layers in comparison with the uncapped sample for over 280 days, indicating different short-term and long-term behaviors in the films. Clearly, the uncapped sample shows a rapid degradation for ~80 days with 280% increase in $n_{\text{sheet}}$. However, after that, film properties recover to some extent presumably because the charge exchange process between the film and air balances out. This recovery happens for all capping layers at some point, except that for the uncapped sample it is more noticeable since it recovers after a huge change in $n_{\text{sheet}}$. Another comparison in Figure 3.14c shows the change in $\mu$ for all the capping layers along with the uncapped sample. It is clear that the mobility of the uncapped sample has a sharp drop in short time, and it gets slowly recovered after ~100 days. This is also true for capped samples, but the change in uncapped sample is more conspicuous.

In$_2$Se$_3$-capping layer makes the electrical properties of the film significantly worse immediately (Figures 3.14b and c). The In$_2$Se$_3$-capped sample has initially almost two times higher $n_{\text{sheet}}$ compared to the uncapped sample. For depositing In$_2$Se$_3$ on Bi$_2$Se$_3$, selenium shutter was open for a few seconds prior to opening indium shutter to ensure Se rich environment in the vicinity of the substrate surface before indium atoms arrive; we tried various other schemes of depositing In$_2$Se$_3$, but all resulted in similarly degraded transport properties of Bi$_2$Se$_3$ films. One possible explanation for this drastic degradation is that indium tends to bond with selenium from the Bi$_2$Se$_3$ rather than the surrounding selenium, hence this mechanism leaves behind more selenium vacancies in the film. The result from In$_2$Se$_3$-capping layer shows that although superficially this capping layer may seem compatible with Bi$_2$Se$_3$ because of its similar structure, it does not function as an effective capping layer.

PMMA-capping helps to suppress the drastic aging effect. Moreover, the coating and removing of this commonly used polymer is straightforward. Due to ex situ deposition, the transport results before and after PMMA coating could be compared. The films were exposed to
Figure 3.15 Baking temperature dependence of aging effect for room temperature transport properties of 30 QL Bi₂Se₃ films with PMMA-capping. a, \( n_{\text{sheet}} \) and b, \( \mu \) vs. time for Bi₂Se₃ films with different baking temperatures of PMMA-capping. The open symbols show \( n_{\text{sheet}} \) and \( \mu \) values for as-grown samples before PMMA is coated and baked for each baking temperature, which are shifted to the left for clarity. The inset in (a) shows the percentage change in carrier densities right after PMMA-capping compared to before. The data points in (a) overlap due to similar carrier densities of the as-grown films. Adapted from Reference 242.

air less than 10 minutes prior to PMMA-capping. Figure 3.15 summarizes \( n_{\text{sheet}} \) and \( \mu \) for different temperatures of PMMA baking. The initial values before PMMA-capping are also marked in the plots. There is a sudden increase in \( n_{\text{sheet}} \) right after PMMA coating compared to the value before capping, which might be due to the delay in capping and/or curing of the PMMA layer. However sometime after coating (Figure 3.15a), there is always a decreasing trend in \( n_{\text{sheet}} \) until it reaches a minimum close to the original \( n_{\text{sheet}} \). This is probably due to the fact that the solvent in the PMMA dries out over time. The inset graph in Figure 3.15a shows the percentage change in \( n_{\text{sheet}} \) right after PMMA-capping for each baking temperature compared to the original value before capping. This shows that for higher temperature baking, the amount of sudden change in the value of \( n_{\text{sheet}} \) is lower, and this is probably due to the fact that harder baking reduces residual solvent and the density of pores in the PMMA-capping. Although compared to the uncapped sample, PMMA-capping makes the aging much slower, PMMA-capped sample goes through degradation
before recovering again. This degradation is likely due to ambient species eventually reaching the film through PMMA layer, or due to change in PMMA properties over time. The start and peak of degradation process (Figure 3.15a) depends on PMMA baking temperature, which presumably corresponds to the amount of leftover solvent in PMMA. For the case of change in $\mu$, the mechanism is more complicated, but it is clear that the mobility of the sample measured right after PMMA coating is always lower than the value before capping, which might be due to delay in PMMA-capping or heat from the baking process. The mobility values eventually recover after a decreasing trend, which is consistent with our previous discussion.

Among the capping layers, selenium gives the best result in terms of stability of $n_{\text{sheet}}$ for both short and long-time compared to the uncapped sample. Our comprehensive study on Se-capped sample for more than 280 days using DC transport measurements demonstrates the remarkable effectiveness of this capping. It should be noted that the in situ deposition of this capping guarantees no direct exposure of the film’s surface to air. Figure 3.14a shows that the change in $n_{\text{sheet}}$ is less than 27% over the course of 280 days, compared to uncapped sample, which confirms that the Se-capping efficiently protects the sample against environmental doping. Figure 3.14d shows the absolute change in daily $n_{\text{sheet}}$ value compared to the initial one for different Se-capping thickness for several days. For a better comparison, the values of the uncapped sample are plotted, too. The change in $n_{\text{sheet}}$ approximately follows an exponentially-saturating behavior: $C(1 - e^{-Dt})$, where $t$ is time. The long-time saturation value $C$ and the exponential rate $D$ are shown in the inset of Figure 3.14d for both uncapped and Se-capped samples. The fact that $D$ is almost constant for all selenium thicknesses and is quite different for the uncapped sample, implies that $D$ is a material constant, independent of the capping thickness. On the other hand, the value for $C$ depends on the thickness of capping layer, which suggests that the thickness of capping layer determines the final level of contamination that reaches the underlying $\text{Bi}_2\text{Se}_3$ layer. Moreover, Se-capping was found to be very effective for surface sensitive techniques such as STM.
Figure 3.16 Aging of Bi$_2$Se$_3$ films measured by time domain THz spectroscopy at 5 K. Real conductance vs. THz frequency for 16 QL Bi$_2$Se$_3$ a, with no capping, b with PMMA-capping (baked at 100 °C), and c, with 20 nm Se-capping, measured at different times after growth. The dashed lines are the Drude-Lorentz fits and the peak at ~1.9 THz is the bulk phonon mode of Bi$_2$Se$_3$. d, Spectral weight, e, Scattering rate, f, $n_{\text{sheet}}$ and g, $\mu$ vs. time for different capping layers: the latter two are derived from the former two quantities. Adapted from Reference 242.

or ARPES; it has been shown by STM that a quick ion milling, prior to desorbing the Se-capping by heat, leaves behind an ultra-clean Bi$_2$Se$_3$ surface to work with$^{111}$. Time-domain terahertz spectroscopy results for uncapped, PMMA- and Se-capped samples support the DC transport data. The TDTS measurements were performed at the Johns Hopkins University. The uncapped samples were measured a day (after one-day shipping) and a week after the growth day. A sharp change in the real conductance as a function of THz frequency (Figure 3.16a) is noticeable in this short time. The real conductance for PMMA- and Se-capped samples was measured a day, a month and two months after the growth day (Figure 3.16b and c, respectively). The small change in the real conductance spectra for 20 nm Se-capping after 2 months aging confirms the effectiveness of this capping layer in stabilizing the sample’s properties.

Figures 3.16d-g show the change in spectral weight and scattering rate as well as the extracted $n_{\text{sheet}}$ and $\mu$ for fresh and aged samples and confirm that Se-capping has the most
stabilizing property. It is also worth noting that with higher thickness for Se-capping and higher baking temperature for PMMA-capping, the stability would have been even better.

3.7 Conclusion

This chapter sheds light on the journey that TIs made to evolve from the first-generation films with high level defect density to the new-generation films with ultralow defect density. In this chapter, we explained the origin of defects in TIs as well as the various defect and interface engineering that have been developed and implemented to suppress the inadvertent bulk and surface defects and to lower the Fermi level within the bulk gap and as close as possible to the Dirac point. It was discussed that substrate’s crystal structure and chemical reactivity play a leading role in determining the final properties of TI films. Thus, finding an optimal substrate holds a great importance, however such optimal substrate does not exist commercially and has to be grown virtually. We showed that how the growth of Bi$_2$Se$_3$ on the optimally-designed In$_2$Se$_3$/BiInSe$_3$ buffer layer led to films with an order of magnitude lower carrier density and an order of magnitude higher carrier mobility as compared to the first-generation films. ARPES, STM, and transport consistently confirm the TSS-dominated conduction in these films. Further, growing Sb$_2$Te$_3$ films on the In$_2$Se$_3$/Sb$_{0.65}$In$_{0.35}$Te$_3$ buffer layer along with a systematic Ti doping yield films with two orders of magnitude lower carrier density. Finally, we explained the crucial role of an effective capping in protecting these low-Fermi-level films against aging. We showed that both Se and PMMA capping layers help stabilizing the properties of Bi$_2$Se$_3$ films; they have additional advantage of easy removability for follow-up experiments. Moreover, charge-depleting capping layers with high electron affinity, such as MoO$_3$ can be used not only as protection layer but as a natural gate to further tune the carrier density. It was shown that the combination of optimally designed structurally/chemically-compatible buffer and capping layers brought down the Fermi level from more than 300 meV above the Dirac point, which was the case for the first-generation
TI s, to only few tens of meV above the Dirac point. These techniques also resolved the previously challenging task of carrier-type tunability in TI films and led to low-carrier-density p-type Bi$_2$Se$_3$ and ultralow n-type Sb$_2$Te$_3$ films for the first time. More importantly, the combination of effective capping layers and low-carrier-density TI films can open routes to previously unexplored regime of topological insulators which is the topic of the next chapter.
Chapter 4

Unveiling the topological quantum effects

4.1 Quantum Hall effect for topological surface states

Under a perpendicular magnetic field, two-dimensional electronic systems, such as semiconductor two-dimensional electron gases, graphene, and topological surfaces of a TI, exhibit the quantum Hall effect (QHE). When a 2DEG is placed in a perpendicular magnetic field, the mobile electrons deflect due to Lorentz force and eventually move in circular paths known as cyclotron orbits. As the magnetic field increases, these orbits get smaller and the electrons become more bound, leading to an insulating phase in the bulk. However, these orbits cannot be completed near the edges, which leads to skipping orbitals and eventually to integer number of chiral conducting channels at the edges of the sample. This is a simplified explanation for how QHE occurs. Under a magnetic field, the parabolic dispersion (\(E \sim k^2\)) band structure of a 2DEG changes into discrete energy levels, known as Landau levels (LLs), with energies \(E_n = \hbar \omega_c (N + \frac{1}{2})\), where the cyclotron frequency \(\omega_c = \frac{eB}{m^*}\) and \(N = 0,1,2,3, \ldots \) is the LL index. As a result of LLs formation, when the \(E_F\) locates between two consecutive LLs (\(N\) and \(N + 1\)), the QHE can be observed, where longitudinal conductance (resistance) vanishes and Hall conductance (resistance) plateaus at integer multiples of conductance quantum \(\frac{e^2}{h}\) (resistance quantum \(\frac{\hbar}{e^2}\)), yielding \(\sigma_{xy} = \nu \frac{e^2}{h}\) with \(\nu = N + 1 = 1,2,3, \ldots \). Basically, \(\nu\) is the number of conducting edge states and it is related to a bulk topological invariant known as a Chern number. In principle, one can understand QHE with the edge-state picture. Any real sample is finite, and thus there is always a confinement potential near the edges to keep the electrons within the sample which makes bulk LLs diverge near the edges, giving rise to edge states. If the Fermi level is placed between two successive LLs, then it crosses these edge states. The number of edge states that the Fermi level crosses defines the Chern
**Figure 4.1 The edge-state picture of the QHE.** a, Schematic energy spectrum of bulk LLs along with edge states in a conventional 2DEG system. b, The propagating edge states. The Fermi level crosses 2 edge states corresponding to $\nu = 2$ plateau in the Hall conductance measurement. d, The number of conducting edge channels for four different Fermi level crossings (marked in c with dashed lines), which is the algebraic sum of edge states crossing the Fermi level with positive and negative velocity ($N = n_+ - n_-$). The perpendicular magnetic field is in plane for the side surfaces, thus the Dirac cone of the side surfaces is preserved. e, The side surface Dirac cone is clearly visible in thicker regime.

The chirality of the edge channel depends on the sign of $\frac{dE}{dk}$ or its group velocity (Figure 4.1).

In contrast, when Dirac electrons with a linear dispersion $E = v_F k$ are subject to a strong magnetic field, unequally spaced LLs with energy $E_n = sgn(N)v_F\sqrt{2\hbar|N|B}$ are formed, where $v_F$ is the Fermi velocity and $N = 0, \pm 1, \pm 2, \ldots$ From this, it is clear that the N=0 LL of this system is at zero energy, unlike 2DEG case where the zeroth LL starts at $\frac{\hbar eB}{2m^*}$. This then leads to
the observation of TSS- originated QHE which are quantized as $\sigma_{xy} = \nu \frac{e^2}{h}$ (for each surface), where $\nu = N + \frac{1}{2} = \pm \frac{1}{2}, \pm \frac{3}{2}, \pm \frac{5}{2}$, etc. Thus, unlike conventional 2DEGs, where $\sigma_{xy}$ is quantized at integer values, $\sigma_{xy}$ is quantized at half-integer values for each surface of a TI. Figure 4.1 c-e show the edge state picture of the TSS-originated QHE in TIS. In transport measurements, both surfaces of a TI are measured together, thus the total Hall conductance ($\sigma_{xy,\text{total}}$) is the sum of Hall conductance from top and bottom surfaces, which gives $\sigma_{xy,\text{total}} = (N_T + \frac{1}{2} + N_B + \frac{1}{2}) = (N_T + N_B + 1) \frac{e^2}{h}$ where $N_T$ and $N_B$ are the top and bottom surface LL indices, respectively. Therefore, in practice, the measured $\sigma_{xy,\text{total}}$ plateaus at integer multiples of $\frac{e^2}{h}$. In fact, since Dirac electrons always appear in pairs in lattice models (no-go theorem), half-integer values cannot be realized. If the top and bottom surfaces of a TI are identical and degenerate ($N_T = N_B = M$), then $(2M + 1)$ is an odd integer, and thus the plateaus are observed at odd multiples of $\frac{e^2}{h}$. However, in practice, the top and bottom surfaces of TI samples are not degenerate and $\sigma_{xy,\text{total}}$ plateaus at both even and odd values. This usually happens because the top and bottom surfaces could have different carrier densities due to different growth conditions or structural asymmetry between the top and bottom interfaces. Application of gate-voltage on both surfaces could make the surfaces degenerate. It is worth noting that the Hall conductance plateaus for graphene are observed at $\sigma_{xy} = 4\times (N + \frac{1}{2}) \frac{e^2}{h}$ due to four-fold spin-valley degeneracy.

In the next two sections, we discuss the experimental observation of QHE in TI materials and we specifically focus on the QHE in new generation low-carrier-density Bi$_2$Se$_3$ and Sb$_2$Te$_3$ films. This is followed by a discussion on how the new generation films led to the observation of quantized Faraday and Kerr rotation. Finally, in the last section, we discuss how these low-Fermi-level films allowed a detailed study of topological phase transition in finite-size regime and the observation of metal-to-insulator transition in transport measurements.
4.1.1 TSS-originated QHE in low-carrier-density Bi$_2$Se$_3$ films$^7$

Although Landau levels were seen with scanning tunneling spectroscopy$^{49,178}$, the quantum Hall effect was not observed in transport in the early TI systems due to high level of defect density. As mentioned previously, one effective tool to suppress carrier density in V-VI TI systems is using a ternary or quaternary alloy$^{139,172}$ (using a Bi and Sb mixture and/or a Se and Te mixture), which lowers the $E_F$ to where electrostatic-gating can further tune the chemical potential to the Dirac point. Finally in 2014, using this method, the first observation of TSS-originated quantum Hall effect (QHE) in exfoliated Bi$_{2-x}$Sb$_x$Te$_{3-y}$Se$_y$ (BSTS) single crystals via back-gating$^{152}$ and shortly after in gated MBE-grown (Bi$_{1-x}$Sb$_x$)$_2$Te$_3$ (BST) thin-films$^{153}$ was reported. To access both top and bottom surfaces individually, dual gating (top- and back-gating) was also performed on BSTS single-crystal and QHE was studied in more detail$^{154}$. Xu et al. observed a series of ambipolar two-component half-integer Dirac quantum Hall states as well as the zeroth Landau level (ZLL), with total filling factor $\nu_{\text{total}} = \nu_{\text{top}} + \nu_{\text{bottom}} = 0$ (top $\nu_{\text{top}}$ and bottom $\nu_{\text{bottom}}$ surface filling factors are half-integers with opposite signs). For the states with $\nu = 0$, a zero $\sigma_{xy}$ plateau and a large $\rho_{xx}$ maximum but a finite $\sigma_{xx}$ minimum were observed ($\sigma_{xx} = \frac{\rho_{xx}}{\rho_{xx}^2 + \rho_{xy}^2}$ and $\sigma_{xy} = -\frac{\rho_{xy}}{\rho_{xx}^2 + \rho_{xy}^2}$).

The finite residual $\sigma_{xx}$ as well as the large non-local resistance at $\nu = 0$ was attributed to dissipative edge channels$^{247}$. It has been predicted that the TSS-originated $\nu = 0$ state is a promising platform to observe novel physics phenomena, such as topological magneto-electric effects$^{248,249}$ or excitonic superfluidity$^{250}$ in TI systems.

As was discussed in the previous section, with special interfacial engineering, MBE-grown Bi$_2$Se$_3$ films with low-carrier-density and high-mobility were achieved. On this platform, the TSS-originated QHE in pure binary Bi$_2$Se$_3$ TI films was observed for the first time in 2015.

$^7$Section 4.1.1 is reproduced (with some changes) with permission from Koirala, N.; Brahlek, M.; Salehi, M.; et al. Nano Lett., 2015, 15 (12), 8245–8249. ©2015 American Chemical Society with permission.
Figure 4.2 Surface quantum Hall Effect in an 8 QL thick MoO$_3$-capped Bi$_2$Se$_3$ film grown on In$_2$Se$_3$/BiInSe$_3$ buffer layer. 

a, The Hall resistance in magnetic fields up to ~35 T at different temperatures, which quantizes to $\frac{h}{e^2}$ at the lowest temperature 0.3 K, indicating a well-defined QHE. The inset shows the hand-scratched Hall-bar structure of the measured film. 
b, The longitudinal resistance at different temperatures for fields up to ~35 T, which goes to 0.0 ± 0.5 Ω at T = 0.3 K when Hall resistance quantizes to $\frac{h}{e^2}$, indicating dissipationless transport. The vertical arrows show the direction towards higher temperatures. The signature of the QHE persists up to higher temperatures, as high as 70 K. Adapted from Reference 144.

Using the 35T magnet system in the NHMFL, TSS-originated QHE was observed on a large-area hand-patterned Hall-bar of an ungated 8QL-thick film of pure Bi$_2$Se$_3$ (without any impurity addition or electrostatic gating) capped in situ by MoO$_3$/Se with the sheet carrier density $n_{\text{sheet}} \approx 7.0 \times 10^{11}$ cm$^{-2}$ and mobility $\mu \approx 4000$ cm$^2$V$^{-1}$s$^{-1}$ (measured from low-field Hall slope). This means that the Fermi level is about 40 to 50 meV above the Dirac point. Although a capping layer induces scattering and lowers the mobility, the mobility still remains in acceptable range to observe most of the TSS-related phenomena. The Se and MoO$_3$ were thermally evaporated from the Se and MoO$_3$ sources at room temperature. Figure 4.2 shows the completely vanishing longitudinal resistance ($R_{\text{sheet}} \sim 0.0 \pm 0.5$ Ω) above 30 T, which was not observed in previous studies, and the perfect Hall quantization at quantum resistance ($R_{\text{Hall}} = 1.00000 \pm 0.00004 \times \frac{h}{e^2} = 25813 \pm$...
1 \Omega \) above 29 T, indicating dissipationless transport with a well-defined surface QHE at 0.3 K in this bulk-insulating low-carrier-density sample. Furthermore, the signature of QHE persists even up to 70 K\textsuperscript{144}. Additionally, in the previous chapter, it was shown that these films are dominated by the TSS carriers, and thus the observed QHE is TSS-originated (for more details see Reference 144). However, all the previously studied QHEs along with the one on the new generation Bi\textsubscript{2}Se\textsubscript{3} films required high magnetic fields. This motivated us to develop a new platform with even lower Fermi level/carrier density to be able to observe the QHE at much lower fields. Thus, we extended the buffer layer scheme to the telluride-based TI system Sb\textsubscript{2}Te\textsubscript{3}, and interestingly films with even lower sheet carrier density were achieved.

4.1.2 QHE and QH-to-insulator transition in low-carrier-density Sb\textsubscript{2}Te\textsubscript{3} films\textsuperscript{8}

As also mentioned in the previous section, with the use of a compatible buffer and capping layers along with Ti doping, low-carrier-density (as low as \sim 1.0\times10\textsuperscript{11} cm\textsuperscript{-2}) thin-films of Sb\textsubscript{2}Te\textsubscript{3} can be achieved. It was shown that, with a proper and systematic Ti doping, the carrier type can be tuned from p-type (intrinsic to Sb\textsubscript{2}Te\textsubscript{3}) to n-type\textsuperscript{155}. With these films, without any electrostatic-gating, it was possible to substantially reduce the required magnetic field for observing the QHE at 5.5 T. This field is much smaller than the 24 T which was required to observe QHE in ungated BL-based Bi\textsubscript{2}Se\textsubscript{3} films\textsuperscript{144} or the 20 T for BL-based Ca-doped Bi\textsubscript{2}Se\textsubscript{3} films\textsuperscript{217}. This is so far the lowest magnetic field at which the QHE has been observed in a TI system. Figure 4.3 shows the magnetic field-dependence of $\rho_{xx}$ and $\rho_{xy}$ in ungated buffer layer-based 8QL-thick Sb\textsubscript{2}Te\textsubscript{3} films with various Ti counter-dopant concentrations of 1.0% (ultra-low p-type; we named it in the previous chapter P8-1.0%, where P refers to p-type sample, 8 is the thickness in quintuple layers, and 1.0% is the Ti concentration), 1.1% (p-type but closer to the n-p mixed regime (near the Dirac point); named P8-1.1%), 1.3% (very low n-type; named N8-1.3%), and finally 1.5% (ultra-low n-type; named N8-1.5%) for magnetic fields from 0 to 45 T at $T = 300$ mK.
Figure 4.3 Transport Properties of 8QL films with various Ti-doping levels. 

(a) $\rho_{xx}$ for 8QL Sb$_2$Te$_3$ films with 1.0%, 1.1%, 1.3% and 1.5% of Ti-doping as a function of magnetic field from 0 to 45 T. Inset shows a photo of an ungated device with a hand-scratched Hall-bar and manually-pressed indium wires as contact leads. The left panel is for $0 < B < 11$ T. 

(b) $\rho_{xy}$ as a function of magnetic field from 0 to 45 T. The left panel is for $0 < B < 11$ T. The inset in (b) right panel shows the anti-symmetrized Hall resistivity in N8-1.5%. Adapted from Reference 155.

$v = 1$ QH plateaus are seen in both n- and p-type samples. This is in contrast with Bi$_2$Se$_3$ films where QHE was observed only in n-type films. The absence of Landau levels on the p-side was also confirmed in scanning tunneling spectroscopy of Bi$_2$Se$_3$ films $^{49, 52, 217}$ (See Figure 4.4). This may be related to the nature of Bi$_2$Se$_3$ band structure and could be due to the closeness of its Dirac point to the bulk valence band and/or the broadness of surface band on the p-side.
The films properties remain almost the same even one year after the growth, indicating that Sb$_2$Te$_3$ films are robust against aging in air and the capping layer is effective. This significant enhancement in carrier-density, and the fact that Te-based systems seem less susceptible to aging and downward band-bending near the surface make Sb$_2$Te$_3$ a promising TI system for further study. Although with such low carrier density films we were able to observe the QHE in the 9 T superconducting AMI magnet system in our lab, it was interesting to find what else could be seen in these films if measured at much higher fields. This urged us to measure some of the low-carrier-density films in the 35 T and 45 T magnet systems in the NHMFL. Remarkably, we observed a magnetic field-driven transition from QH to insulator phase with a gigantic magnetoresistance ratio $\text{MR} = \frac{\rho_{\text{insul-phase}}(45T)}{\rho_{\text{QH-phase}}} - \frac{\rho_{\text{QH-phase}}}{\rho_{\text{QH-phase}}}$, where $\rho_{\text{QH-phase}}$ is the minimum $\rho_{\text{xx}}$ in the QH-regime, as large as $5 \times 10^5\%$ and $8 \times 10^6\%$ at 300 mK for N8-1.3% and N8-1.5%, respectively. Although there have been two reports for QAH-to-insulator in TI thin-films doped with magnetic impurities$^{150-151}$, this
is the first experimental observation of a magnetically-driven QH-to-insulator transition (QIT) in a TI system.

The observation of insulating phase at higher fields is surprising on its own. This is because the zeroth Landau levels (ZLLs)\(^1,41,251\) of TSS in a TI are two-fold degenerate and the degeneracy, in the presence of structural symmetry (crystal inversion and top/bottom symmetry) and in the absence of tunneling between top and bottom surface states, is robust and cannot be lifted by the Zeeman energy (indeed they are merely shifted by Zeeman field)\(^252\). Neglecting interactions, increasing the magnetic field such that the filling fraction approaches \(\nu = 0\) causes the ZLLs to be half-filled, producing a compressible (metallic) state. This result holds even in the presence of non-magnetic random disorder which on average preserves the structural symmetry. In contrast, conventional two-dimensional electron gases (2DEGs) are insulating near \(\nu = 0\), i.e., near the bottom of the lowest LL, due to the presence of localized states at the tail of the lowest LL. To investigate whether and how the QIT could occur for the surface states of TIs in the presence of ZLLs, ultralow-carrier-density samples with their Fermi level close to the Dirac point are required. Thus, these ultralow-carrier-density Sb\(_2\)Te\(_3\) films are a suitable test bed to study the physics of ZLL in TIs which was inaccessible in previous TI samples either because of high defect densities pushing the \(E_F\) away from the Dirac point or the significant level of electron-hole puddles obscuring the energy landscapes of ZLL. Even in the new generation low-carrier-density Bi\(_2\)Se\(_3\) system, the ZLL does not seem to be clearly accessible most likely because of proximity of the bulk valence band to the Dirac point\(^217\).

The transitions to QH and then to insulator can be clearly shown in form of parametric flow lines of the conductivity tensor \((\sigma_{xx} \text{ versus } \sigma_{xy})\), where \(\sigma_{xx} = \frac{\rho_{xx}}{\rho_{xx}^2 + \rho_{xy}^2}\) and \(\sigma_{xy} = \frac{\rho_{xy}}{\rho_{xx}^2 + \rho_{xy}^2}\)

for sample N8-1.5% as the applied field is tuned over \(0 < |B| < 35\) T (Figure 4.5) The trajectory includes several features: 1. Two outside curves with a cusp at low fields (from weak anti-
Figure 4.5 The conductivity tensor flow ($\sigma_{xx}$ vs. $\sigma_{xy}$) of N8-1.5% for $|B| \leq 35$ T. The field is incorporated as a color map in the plot. The points ($\pm \frac{e^2}{h}, 0$), (0,0), and ($\pm 0.5 \frac{e^2}{h}, 0.5 \frac{e^2}{h}$) correspond to the QH state, the insulator phase, and the transition between these two phases, respectively. The cusp around zero field indicates weak anti-localization effect. Adapted from Reference 155.

localization), which approach $\nu = \pm 1$ plateaus ($\pm \frac{e^2}{h}, 0$) with increasing field as the sample enters the QH regime. 2. At higher fields, semicircles corresponding to the transition between the $\nu = \pm 1$ QH plateau at ($\pm \frac{e^2}{h}, 0$) and the insulating state at (0, 0). The insulating phase at $|B| > 24$T is characterized by a diverging $\rho_{xx}$ and a finite $\rho_{xy}$ (after anti-symmetrization) upon increasing field and decreasing temperature$^{253}$ (see Figure 4.6). The points ($\pm 0.5 \frac{e^2}{h}, 0.5 \frac{e^2}{h}$) mark the topological transition from the $\nu = \pm 1$ QH state to the insulating phase.

In fact, for the semi-circle part we can write $\sigma_{xx}^2 + \left(\sigma_{xy} - \frac{e^2}{2h}\right)^2 = r^2$, where $\sigma_{xx} = \frac{\rho_{xx}}{\rho_{xx}^2 + \rho_{xy}^2}$, $\sigma_{xy} = \frac{\rho_{xy}}{\rho_{xx}^2 + \rho_{xy}^2}$, and $r$ is the radius of the semi-circle. Then expanding the left-hand side of the equation gives

$$\frac{\rho_{xx}^2}{(\rho_{xx}^2 + \rho_{xy}^2)^2} + \frac{\rho_{xy}^2}{(\rho_{xx}^2 + \rho_{xy}^2)^2} + \left(\frac{e^2}{2h}\right)^2 - \frac{(e^2/\pi)^2}{\rho_{xx}^2 + \rho_{xy}^2} = \left(\frac{e^2}{2h}\right)^2 + \frac{1-(e^2/\pi)^2}{\rho_{xx}^2 + \rho_{xy}^2}$$

which implies
Figure 4.6 A possible scenario for the insulating behavior could be a (quantum) Hall insulator. 

(a) The longitudinal (red) and Hall (blue) resistivities of N8.1.5%, at high fields at 300 mK, shown after (anti-)symmetrization between positive and negative signed fields. A dashed line indicates the value of the resistance quantum $\frac{h}{e^2}$. A $\nu = 1$ quantum Hall plateau is observed at applied fields between approximately 5 and 20 T. The quantum Hall to insulator transition occurs at $B_B = 24$ T. For clarity, the sign of the Hall resistivity has been reversed in this figure. 

(b) The longitudinal (red) and Hall (blue) resistivities (after (anti-)symmetrization) of the gated device at 30 T (dashed lines) and 35 T (solid lines) as a function of gate voltage at 300 mK. A $\nu = 1$ quantum Hall plateau is observed, centered around $V_g = 11.5$ V (at 30 T) and $V_g = 11.3$ V (at 35 T). The $\rho_{xx}$ diverges at higher gate voltages as carriers are depleted, while the $\rho_{xy}$ remains roughly quantized at $\frac{h}{e^2}$ up to around $V_g = 18$ V, at which $\rho_{xx} = 1.2$ M$\Omega$. This observation suggests the system may be a quantized Hall insulator in approximately the range $V_g = 14$ V to 18 V. Adapted from Reference 155.

that if $\rho_{xy}$ remains quantized and/or $\rho_{xx}$ diverges to large values we get a decent clean semi-circle trajectory with a radius $\frac{e^2}{2h}$ in the conductance tensor flow lines.

Additionally, two possibilities are proposed for the measured insulating behavior at high fields (near $\nu = 0$): 1. a coherent excitonic superfluid state between the degenerate ZLLs $^{250}$ 2. a Hall insulator state $^{253,254}$. Coulomb interactions are essential for realizing the former candidate phase, while the latter phase can be stabilized in thin-films by disorder in strong magnetic fields even in the absence of interaction. Regarding the first possibility, our samples are well within the range (thickness < $3\ell_B$, where the magnetic length $\ell_B \approx \frac{26nm}{\sqrt{B}}$) where the inter-surface Coulomb
Figure 4.7 The schematic of energy spectrum of LLs along with the edge modes for different 2D systems in a magnetic field with and without considering the Zeeman coupling. Schematic of energy spectrum of a, TI thin-film b, graphene, c, 2DEG in the absence of magnetic field (B = 0; the first column), in a small magnetic field without the Zeeman-coupling (B = small; the second column) and in a strong magnetic field with the Zeeman-coupling (B = large; the third column). The zero energy is marked by dashed lines. Because of spin-momentum locking in TI surface Hamiltonian, the Zeeman field does not lift the LL degeneracies; instead, it shifts all LLs. Particularly, the Zeeman term moves the top and bottom LLs up or down depending on magnetic field direction. The energy shift is more pronounced for the zeroth LL and becomes smaller for higher LLs. Note that for a TI thin-film with a finite-size induced tunneling gap $\Delta_t$, the degeneracy of zeroth LLs is lifted with a gap of $\Delta_t$. Although, in practice this gap is smeared by disorder, applying sufficiently large magnetic field could resolve this gap to the clean-limit. In contrast, in 2DEG and in graphene the Zeeman field lifts the spin degeneracy (here spin-split 2DEG is shown). The band crossing of the counter-propagating edge modes can be gapped by impurity scattering, effectively driving the system to an insulating regime. Adapted from Reference 155.

interaction (between top and bottom surfaces) dominates over intra-surface Coulomb interactions (on top or bottom surfaces). Based on Reference 250, we might expect a coherent superfluid phase with a charge gap of ~16 meV. However, there is a competing process at strong magnetic fields due to quadratic corrections to the surface dispersion relation near the $\Gamma'$ point. These corrections,
Figure 4.8 Surface LLs and edge modes on the side surfaces. a, a thick TI sample without (top panel) and with (bottom panel) Zeeman field and b, a hybridized TI with a finite-size tunneling gap of $\Delta_t$, which lifts the degeneracy of the zeroth LLs. Although, in practice this gap is smeared by disorder and could be effectively zero, applying sufficiently large magnetic field could resolve this gap to the clean-limit. Higher magnetic fields could increase this gap further. The Zeeman energy shift of LLs is also shown in the bottom panel of b. Adapted from Reference 155.

which are more pronounced in thin-films, split the zeroth LLs, favoring a Hall insulating phase (See Figure 4.6). We theoretically estimate an upper bound for the magnetic field-induced gap to be $\sim 10$ meV. This gap is not produced by the Zeeman spin splitting but due to the quadratic band-structure correction: see Figure 4.7, where the energy spectra of a TI thin-film, graphene, and a 2DEG system both in the absence of and in the presence of magnetic field with and without the Zeeman effect are schematically presented. Furthermore, there always exists a tiny mass gap $\Delta_t$ at the Dirac point in TI thin-films due to tunneling between the top and bottom surface (even if the sample is thicker than the limit where the top and bottom surfaces are measured to be hybridize$^{178-181}$), which leads to insulating behavior at low magnetic field when gated to the Dirac point (aka
Figure 4.9 The longitudinal conductivity of a gated device at zero field, gate-tuned to the charge neutrality point, shown on a log scale as a function of inverse temperature. The data is fit to an Arrhenius model $\sigma_{xx} \approx \exp\left(-\frac{\Delta_p}{k_B T}\right) + \sigma_{xx}^0$, finding $\Delta_p = 161 \mu eV$. The gap at charge neutrality is understood as a consequence of hybridization between the top and bottom surface states. We attribute the constant offset $\sigma_{xx}^0$ to Joule heating as its value was observed to decrease with decreasing current bias. This measurement used a 1 nA current bias for temperatures $T < 1.5$ K. This measurement included data from two separate cooldowns (one in a He-3/He-4 dilution refrigerator and one in a He-4 system). To account for offsets in gating between the two cooldowns, the conductivity shown is the minimum value of $\sigma_{xx}$ (as a function of gate-voltage) at each temperature. Adapted from Reference 155.

charge neutrality point (CNP)). At high magnetic fields, the ZLL are then split by $\Delta_e$; a Hall insulator phase could be present between the two zeroth Landau levels (See Figure 4.8). However, in practice, this finite-size gap in the surface can be smeared by disorder and electron-hole puddles, which reduces $\Delta_e$ from its clean-limit value. This could be why we measure a rather small gap on the order of 0.16 meV in our gated devices (Figure 4.9). Theoretically, disorder always drives the (finite-size gapped) system toward a metal$^{255}$. Similar physics applies to the low magnetic field regimes, where LLs are mixed by disorder. In fact, this could explain why we do not observe higher LLs in our transport measurements, especially because the spacings between the higher LLs are smaller, thus they become more susceptible to disorder and electron-hole puddles. In contrast, the first LL is rather robust against disorder. Theoretically, the effect of disorder on LLs can be modeled and the results are summarized in phase diagrams of Figure 4.10 (See the Supporting Information of Reference 155 for theoretical details). Upon increasing the magnetic field, LLs become sharper
Theoretical modeling of the effect of random impurity potential on LLs in the absence (a and b) and the presence (c and d) of tunneling-induced surface gap. The competition between the magnetic field and random disorder potential can be characterized in terms of a dimensionless parameter $\frac{\Gamma}{\hbar \omega_c}$, which, roughly speaking, compares the LL broadening and the LL spacing. a, Hall conductance of gapless TSSs (inset: density of states). Upon increasing the disorder (larger $\frac{\Gamma}{\hbar \omega_c}$), the higher $\sigma_{xy}$ plateaus disappear except $\nu = \pm 1$, showing how robust the $\nu = 1$ plateau is even in the strong disorder limit. b, The phase diagram of disordered two decoupled Dirac Landau levels (no tunneling gap). Each area is denoted by its corresponding value of the Hall conductance $\sigma_{xy}$. The orange line indicates the critical line between the $\nu = 1$ and $\nu = -1$ plateaus. c, Hall conductance of gapped TSS (inset: density of states) for various disorder strengths again shows how robust the $\nu = \pm 1$ plateaus are. d, Phase diagram of disordered two tunnel-coupled Dirac Landau levels. In particular, the dark and light orange regions indicate the critical metallic region (driven by disorder) and the Hall insulator (formed in the strong magnetic field limit), respectively. Adapted from Reference 155. For detailed theoretical calculations see the same reference.
and at large fields such that \( \omega_c \tau \gtrsim 1 \), where \( \omega_c = \sqrt{2 \frac{v_F}{\ell_B}} \), the smearing of LLs is greatly reduced. Hence, a sufficiently large magnetic field may resolve the clean-limit tunneling gap, which could lead to an insulating if the Fermi level of a sample is extremely low and locates in this gap – which can be the case in these ultralow-carrier-density Sb\(_2\)Te\(_3\) films. The clean-limit gap can theoretically be calculated. The low energy band structure of Sb\(_2\)Te\(_3\) can be approximated by a four band model\(^{251}\). It can then be shown that the surface Hamiltonian can be written as
\[
H_{\text{sur}} = v_F \eta_3 \otimes (k_x \sigma_2 - k_y \sigma_1) + \Delta_t (k_\perp) \eta_1 \otimes \mathbb{I}_2,
\]
where \( \eta_i \) and \( \sigma_i \) are Pauli matrices in top (bottom) surface and spin, respectively, \( \Delta_t (k_\perp) = \Delta_0 + \Delta_2 k_\perp^2 \) denotes a tunneling amplitude between top and bottom surfaces and \( k_\perp^2 = k_x^2 + k_y^2 \). For a thick sample, the surface states decay exponentially into the bulk where the decay length \( \xi \) is inversely proportional to the bulk energy gap.

The inter-plane tunneling for thin samples comes from the overlap of these decaying functions on top and bottom surfaces which results in:
\[
\Delta_0 \approx M_1 \frac{L}{\xi^3} e^{-\frac{L}{\xi}} \quad \text{and} \quad \Delta_2 \approx M_2 \frac{L}{\xi^5} e^{-\frac{L}{\xi}}
\]
where \( L \) is the sample thickness and \( M_1 \) and \( M_2 \) are some dimensionful band structure parameters. Using the parameter values obtained in Liu, et. al \(^{251}\), we obtain \( \Delta_0 \approx 6 \) meV and \( \Delta_2 \approx 15 \) eVÅ\(^2\) for a sample of thickness of \( L = 8 \) nm. However, this gap is different from the experimentally measured \( \Delta_t \approx 0.16 \) meV which could be small due to smearing. In addition to the tunneling gap, as was explained earlier, strong magnetic fields have a secondary effect of pushing the surface zeroth Landau levels away from each other because of the quadratic term in the surface spectrum, which further stabilizes the insulating phase. This additional gap can roughly be calculated from theory as well, which is on the order of 5 meV (See the Supporting Information of Reference 155 for more details and theoretical discussions). Figure 4.11a (top panel) shows \( \rho_{xx} \) of the sample N8-1.5% as
Figure 4.11 Scaling-invariant plots for an 8QL Sb$_2$Te$_3$ film with the lowest carrier density. a, $\rho_{xx}$ as a function of magnetic field for N8-1.5% at temperatures from 300 mK to 10 K (See Figure 4.13c for the flow of conductivity tensor at different temperatures). The critical magnetic field ($B_c \approx 23.9$ T), marked by a dashed line in the inset, is where all the curves pass through. The bottom panel is the corresponding scale-invariant plot for temperatures, yielding $\kappa = 0.20 \pm 0.02$. b, Top panel shows I-V curves for the 1.5%Ti-doped sample at a constant field ranging from 21 T to 27 T with 0.25 T step. Bottom panel is the corresponding scale invariant plot and the inset shows $\rho_{xx}$ as a function of $V_{xx}$. $\kappa' = 0.14$ is extracted from the fit. c, $\rho_{xx}$ as a function of magnetic field for the 1.5%Ti-doped sample for different currents. Like the temperature-dependent curves in (A), the curves corresponding to different currents pass through a critical magnetic field ~24.2 T (the small shift results from sample aging). The bottom panel shows the corresponding scale-invariant plot for current. Adapted from Reference 155.

a function of high magnetic fields from 18 T to 45 T for different temperatures from 300 mK to 10 K. The resistivity curves at different temperatures pass through a critical magnetic field ($B_c$) of 23.9T with $\rho_{xx} (B_c) = 0.75 \frac{h}{e^2}$ and the critical filling fraction $\nu_c = 0.25$ (See Figure 4.12, also see Figures 4.13a and b for corresponding $\sigma_{xx}$ and $\sigma_{xy}$ and Figure 4.13c for flow lines of conductivity tensor for different temperatures). Such critical field, resistivity, and filling fraction are together a hallmark of a scale-invariant critical point. To better understand the nature of quantum criticality
near the QIT, the transition’s critical exponents can be extracted through a systematic scaling analysis. Near the critical magnetic field at zero temperature, the system is characterized by
power law divergence of the localization length $\xi \propto |B - B_c|^{-\nu}$. The localization length is related to the size of electron wave functions. At finite temperatures, a new length-scale $L_\phi \propto T^{-\frac{1}{2z}}$ emerges, where $z$ is the dynamical critical exponent. This length is associated with inelastic scattering. Physically, $z$ describes the dispersion relation $\varepsilon \propto k^2$ where $\varepsilon$ is energy and $k$ is momentum. Thus, the relation $L_\phi \propto T^{-\frac{1}{2}}$ can be deduced from the dimensional analysis that $T \sim \varepsilon$ and $L_\phi \sim k^{-1}$. Hence, we heuristically use the parameter $\delta = \frac{L_\phi}{\xi}$ to define two possible phases: a metallic phase when $\delta \leq 1$ and an insulating phase when $\delta \gg 1$ which in our case would be the quantum Hall plateau and the strong field insulator. Near the critical point $\delta \to 0$ and the magnetoresistance tensor $\rho_{\alpha\beta}$ (or other thermodynamic properties) may assume a polynomial series $\rho_{\alpha\beta}(B, T) = r_{\alpha\beta}(\delta) + \cdots$ which is only a function of a single parameter $\delta$. For instance, the longitudinal resistivity obeys a scaling law $\rho_{xx}(B, T) = \frac{h}{e^2} f_1(\frac{L_\phi}{\xi})$ where $f_1(\delta)$ is a smooth function.

This is called the single parameter scaling ansatz.

From the scaling hypothesis mentioned above, we have $\rho_{xx}(B, T) \propto |B - B_c|T^{-\kappa}$ near $B_c$, where $\kappa = \frac{1}{2z}$. The collapse of data as a function of the scaling-invariant parameter $|B - B_c|T^{-\kappa}$ up to 10 K gives $\kappa = 0.20 \pm 0.02$ in Figure 4.11a (bottom panel). The same $\kappa$ is extracted for N8-1.3%, indicating the universality of the scaling behavior (Figure 4.12). To evaluate $\nu$ and $z$ individually, however, we need to probe the dynamics of the system, which is done using the differential resistivity measurement. Similar to finite temperature, finite voltage (equivalently, electric field $E$) induces another length-scale $L_E \propto E^{-\frac{1}{z+1}}$ and the differential resistivity $\tilde{\rho}_{xx} = \frac{dV}{dl}$ takes the scaling form $\tilde{\rho}_{xx}(B, E) = \frac{h}{e^2} f_2(\frac{L_E}{\xi})$, implying that $\tilde{\rho}_{xx}(B, E) \propto |B - B_c|V^{-\kappa'}$ near $B_c$ where $\kappa' = \frac{1}{(z+1)\nu}$.

The differential resistivity evolves from $\frac{d^2|V|}{dl^2} > 0$ in the QH state, to $\frac{d^2|V|}{dl^2} = 0$ at $B = B_c$, to $\frac{d^2|V|}{dl^2} < 0$ in the insulating state (Figure 4.11b, top panel). The critical scaling (Figure 4.11b, bottom panel)
Figure 4.13 More details on the scaling analysis for N8-1.5%. Magnetic field dependence of \( \sigma_{xx} \) and \( \sigma_{xy} \) of N8-1.5% sample for different temperatures. \( \sigma_{xx} \)'s peak and \( \sigma_{xy} \)'s cross at the critical magnetic field \( (B_c) \). The \( \nu = 0 \) and \( \nu = -1 \) plateaus are marked by dashed lines in the \( \sigma_{xy} \) plot. c. The flow lines of conductivity tensor \( (\sigma_{xx} \ vs. \ \sigma_{xy}) \) in the same sample, where the results at 3 different temperatures and from 11 T to 45 T collapse on a semicircle-like trajectory extending from \((-1,0)\), corresponding to the QH phase, to \((0,0)\) for the insulating phase. \((-0.5,0.5)\) represents the transition point between these two phases. The signature of quantum Hall-to-insulator transition can be observed even at higher temperatures (as high as 20 K). d, Magnetic field dependence of \( \rho_{xx} \) of N8-1.5% sample for different currents (also shown in Figure 4.11c). The curves corresponding to different currents cross at a critical magnetic field \( \sim 24.2 \) T (the small shift of \( B_c \) compared to the one in the temperature plot of Figure 4.11a could be due to sample aging). The inset shows the corresponding current scale-invariant plot. The underlying assumption is that the dissipated energy in the system effectively heats up the electrons via \( k_B T_e \sim eE L_\phi(T_e) \), which implies that the electron temperature \( T_e \) is related to the current \( I \) as \( T_e \propto I^{2/(1+z)} \). Therefore, the current-scaling plot admits the scaling form \( \rho_{xx} \propto |B - B_c|^{-b} \), where \( b = \frac{1}{\nu(1+z)} \). Combining \( b \) and \( k' \), we found that \( z = 1.9 \pm 0.3 \) and \( \nu = 2.7 \pm 0.5 \). Adapted from Reference 155.

yields \( k' = 0.14 \pm 0.01 \). Using the values of \( k' \) and \( \kappa \), we extract \( z = 2.1 \pm 0.2 \) and \( \nu = 2.4 \pm 0.3 \). Additionally, we measured \( \rho_{xx} \) as a function of magnetic field for different currents.
Figure 4.14 Transport properties of a gated 8QL film with a 2nm (Sb$_{0.65}$In$_{0.35}$)$_2$Te$_3$ capping. a, $\rho_{xx}$ as a function of gate-voltage (-10 V $\leq V_g \leq$ 28 V) for different fields 0 T, 5 T, 10 T, 15 T, 20 T, 25 T, 30 T, and 35 T at 300 mK. Although vanishing $\rho_{xx}$ was not observed, the QH regime is indicated by a dip near $V_g = 12$ V at B > 20 T, reaching 2.4 kΩ/sq at 35 T. The peak (~119 kΩ/sq at zero field and ~2.9 MΩ/sq at 35T) corresponds to the CNP. Curves at B > 10 T pass through a critical point $V_g^c = 15.4$ V. At higher gate-voltages, an incipient n-type QHE is observed. Inset shows a photo of the fabricated device. b, Anti-symmetrized $\rho_{xy}$ as a function of gating from -10 V to 28 V for the same fields. The $\pm \frac{h}{e^2}$ plateaus are marked with dashed lines on the plot. The noisiness at high fields comes from incomplete elimination of the geometrically-mixed large $\rho_{xx}$ component during the anti-symmetrization process. c, The parametric flow of the conductivity tensor ($\sigma_{xx}$ vs. $\sigma_{xy}$) after (anti-)symmetrization. Each line represents a gate sweep from $V_g = -10$ V to 28 V at a constant field. Adapted from Reference 155.

at a constant temperature (Figure 4.9c), again finding a critical point at $B_c = 24.2$ T. Similar exponents $z = 1.9 \pm 0.3$ and $\nu = 2.7 \pm 0.5$ are extracted from the current scaling plot$^{260}$ in Figure 4.11c (See Figure 4.13 for more detailed calculations of current scaling). The scaling behavior of QITs$^{259, 261-263}$ and plateau-to-plateau transitions$^{260, 264}$ in conventional 2DEGs always lead to $z \approx 1$. Although there is no fundamental reason why QITs should belong to the same universality class, there have been so far no exceptions in terms of the dynamical exponent.

Our observation of $z \approx 2$ suggests that QIT in TSS belongs to a different universality class than QITs in conventional 2DEGs. The value $z = 1$ in 2DEGs is generally attributed to unscreened Coulomb interactions$^{265-266}$ and so our observation of $z = 2$ could indicate enhanced screening.

Interestingly, another study on the low-carrier-density bulk crystal TI system Sn-$\text{Bi}_{1.1}$Sb$_{0.9}$Te$_2$S has reported $z = 2$ for plateau-plateau transition from $\nu = 2$ to $\nu = 1$ $^{267}$. Whether the difference in
Figure 4.15 The raw transport data of the gated 8QL Sb$_2$Te$_3$ film. a, Raw data for $\rho_{xx}$ (Figure 4.14) as a function of gate-voltage ($-10 \text{ V} \leq V_g \leq 28 \text{ V}$) for different fields of 0 T, 5 T, 10 T, 15 T, 20 T, 25 T, 30 T, and 35 T at 300 mK. b, Raw data for $\rho_{xy}$ as a function of gate-voltage ($-10 \text{ V} \leq V_g \leq 28 \text{ V}$) for the same fields with $h/e^2$ and $-h/e^2$ plateaus corresponding to the QHE on p and n sides, respectively. The large peaks in $\rho_{xx}$ are due to mixing with large $\rho_{xx}$ around the charge neutrality point. The anti-symmetrized result for $\rho_{xy}$ is shown in Figure 4.14b. c, $\sigma_{xx}$ as a function of gate-voltage $-10 \text{ V} \leq V_g \leq 28 \text{ V}$ (bottom axis) and sheet carrier density (top axis) for the different fields. d, $\sigma_{xy}$ as a function of gate-voltage $-10 \text{ V} \leq V_g \leq 28 \text{ V}$ and sheet carrier density for the same fields, where $\nu = 0$ and $\nu = 1$ plateaus are visible. Adapted from Reference 155.

screening results from the different band structures (linear in TIs vs. quadratic in 2DEGs) or excitonic superfluid interactions or other higher order effects is an open question for future studies.

While the above QITs are driven by increasing magnetic field at fixed (small) carrier densities, Figure 4.14 presents QITs driven by decreasing carrier density at fixed (large) magnetic fields. The film is a top-gated 8QL-thick Sb$_2$Te$_3$. By applying a positive gate voltage, we tuned the majority
carrier from p- to n-type, realizing a well-defined QHE on the p-side and an incipient QHE on the n-side. Figure 4.14a and b show $\rho_{xx}$ and $\rho_{xy}$, respectively, as a function of gate-voltage at different magnetic fields: $\rho_{xy}$ is anti-symmetrized to eliminate mixing with large longitudinal between gate-voltage and $n_{\text{sheet}}$, are shown in Fig 4.15 and 4.16, respectively. Figure 4.14c shows the parametric conductivity as the carrier density is tuned by gating: each line represents a gate sweep at a constant magnetic field. At high fields, the parametric conductivity follows semicircles with radius $0.5 \frac{e^2}{h}$, indicating the transition from the $\nu = 1$ QH state to an insulating state, just as in Figure 4.5. However, as the carrier density changes sign through the charge neutrality point, the flow line exits the insulating phase and moves towards the opposite quantum Hall phase, which never occurs in the magnetic field-driven QIT in Figure 4.5.

Figure 4.16 $n_{\text{sheet}}$ as a function of gate-voltage the gated 8QL Sb$_2$Te$_3$ film. The black line is the raw data for $n_{\text{sheet}}$ as a function of gate-voltage, where the top gate of the gated device was swept from -10 V to 28 V at fields between $B = -1$ T to 1 T. At each gate voltage, the Hall resistance was found by linear fitting, and was used to find the carrier density. The jumps in the raw data is due to mixed carrier density near CNP which artificially gives a flat Hall slope. The Hall resistance fit quality near charge neutrality, however, is poor due to mixing between the Hall and longitudinal resistivities. Therefore, to get a sensible relationship between the carrier density and gate voltage we exclude the bad part of the data and use the remaining data (green curve) with a polynomial fit. The red curve is a 9th degree polynomial fit to the data. Adapted from Reference 155.
4.2 Quantized Faraday and Kerr rotation

With the aid of high-precision time-domain terahertz polarimetry technique as well as low-Fermi-level bulk insulating TIs, the topological magneto-electric effects can be probed. In principle, a true TI is considered as a bulk magnetolectric material whose magnetoelectric response is a quantized coefficient and its value is set by the fine-structure constant .

Magnetoelectrics are materials in which magnetic field could induce polarization or electric field could create magnetization.

The new generation interface-engineered bulk-insulating Bi$_2$Se$_3$ films capped by MoO$_3$/Se (capping layers are transparent in THz regime) with low surface $E_F$ ($\sim$ 30 to 60 meV above the Dirac point for ungated samples) in conjunction with time-domain terahertz polarimetry led to the first-time observation of axion electrodynamics in a 3D topological insulator and quantized Faraday and Kerr rotation. Both MoO$_3$ and Se are transparent in THz regime, and thus they do not contribute to the Faraday/Kerr rotation. The consequences of axion electrodynamics are the additional source and current terms in the modified Gauss’s and Ampère’s laws which are responsible for a half-integer TSS-originated QHE. It can be shown that the Faraday rotation in the quantum regime is given by

$$\tan(\phi_F) = \frac{2\alpha}{1+n} \left( N_t + \frac{1}{2} + N_b + \frac{1}{2} \right),$$

where $\alpha$ is fine structure constant, $n \sim 3.1$ is index of refraction of the sapphire substrate in THz range, and $N_b$ and $N_t$ are the highest fully filled LL of the bottom and top surfaces of the film and depends on the Fermi level and magnetic field. As shown in Figure 4.17, when THz light transmits through a film and substrate, the substrate itself performs as an optical resonator, yielding a series of pulses with different histories of interaction with the film. The first peak transmits through the film and undergoes a Faraday rotation. The second peak undergoes an additional reflection as well as Kerr rotation (Figure 4.17c).

---

$^9$Section 4.2 is based on Wu, L.; Salehi, M.; et al., Science 2016, 354 (6316), 1124-1127. Reprinted with permission from AAAS.
After subtracting the Faraday rotation, the Kerr rotation can be separately measured. In the quantized regime, it can be shown that the Kerr rotation in the quantum regime is: \[ \tan(\phi_K) = \frac{4n_e \alpha}{n^2 - 1} \left( N_e + \frac{1}{2} + N_p + \frac{1}{2} \right) \]. The Faraday rotation is a complex quantity, where the real part is the rotation of the ellipse from its major axis and the imaginary part corresponds to the ellipticity (Figure 4.17). When measured at low fields, the Faraday rotation demonstrates semiclassical cyclotron resonance, yielding to the shifting of the real part’s inflection point and the imaginary part’s minimum/dip with the field. As shown in Figure 4.18, when the low-Fermi-level Bi$_2$Se$_3$ samples with different thicknesses are measured at high fields, they feature quantized Faraday (Kerr) rotation where the low-frequency tail in the real part becomes flat. Thicker sample has slightly higher carrier-density and as a result a higher filling factor. Magnetic field of 7T is large...
Figure 4.18 Quantized Faraday and Kerr rotations. a, Quantized Kerr rotation and b, Quantized Faraday rotation in buffer-layer-based Bi$_2$Se$_3$ films with MoO$_3$ and Se capping layers for different thicknesses. Dashed black lines are theoretical expectation values assuming certain values for the filling factor of the surface states. Adapted from Reference 159. Reprinted with permission from AAAS.

enough to observe the quantization, which is much lower than the ~24 T field required for observing QHE in Bi$_2$Se$_3$ samples in DC transport. This is most likely because THz measurement only detects TSS-originated response and does not probe the edges of the sample, and thus it is not affected by non-chiral/imperfect edge states. At the time of writing this thesis, the new generation Sb$_2$Te$_3$ films are being measured by Peter Armitage’s group in the Johns Hopkins University and by Liang Wu’s group in the University of Pennsylvania. The required field for observation of quantized Faraday/Kerr rotation is even smaller than 7T for new generation BL-based Sb$_2$Te$_3$ featuring lower carrier-density than BL-based Bi$_2$Se$_3$.

Moreover, by combining the quantized Faraday and Kerr rotation equations, the dependence on the substrate’s index of refraction can be eliminated and the fine-structure constant given by $\alpha_{measured} = \left(\frac{1}{N_L+N_h+1}\right)\left(\frac{\tan(\phi_F)^2-\tan(\phi_F)\tan(\phi_K)}{\tan(\phi_K)-2\tan(\phi_F)}\right)$ can directly be measured. In fact, the value Wu, et al. obtained from their measurements is close to the accepted value $\alpha \approx 7.3$ mrad. For more information about this work, see Reference 159.
4.3 Finite-size driven topological and metal-insulator transition in (Bi$_{1-x}$In$_x$)$_2$Se$_3$ thin-films

As mentioned in chapter 1, the emergence of TSS can be understood from the fact that TIs and non-TIs belong to two different topological classes and hence cannot be adiabatically connected. If we start from a bulk sample of archetypal 3D TI Bi$_2$Se$_3$ (0.3eV bulk gap) and move toward its topologically trivial insulator twin In$_2$Se$_3$ (1.3eV bulk gap with the same $D_{3d}$ crystal structure) by varying In concentration in solid solution (Bi$_{1-x}$In$_x$)$_2$Se$_3$, there exists a critical $x$ at which the bulk gap closes and topological phase transition from TI to non-TI occurs and beyond which the bulk gap reopens and bulk bands de-invert to their atomic limit ordering. Upon adding In and as the bulk gap becomes smaller due to weakening of spin-orbit coupling, the penetration depth of TSS keeps growing until it diverges at the critical point where the surface states eventually disappear by merging into the bulk states. Therefore, as the sample is made thinner, the diverging thickness of the top and bottom surface states becomes comparable with the film thickness, and hence TSSs hybridize and open a gap at the Dirac point. This process effectively converts the system into a trivial insulator before In concentration reaches the bulk critical point. In other words, TPT in thin samples is driven not only by the composition, which solely determines the bulk transition, but also by the sample thickness. Remarkably, such finite-size effects on TPT have not been investigated carefully in the literature neither experimentally nor theoretically.

In ideal materials, TIs must be metallic (which we name it topological-metallic) on the topological side of the TPT and must be insulating on the non-TI side (which we name it normal-insulating), and therefore, simple transport measurements should be able to determine the TPT transition. However, such a simple and ideal TPT has never been observed before, and instead the system remained metallic (instead of turning into a normal-insulating it remains normal-
metallic) on both sides of the transition. This is because in all previous studies, the Fermi level always remained high in the bulk conduction band, leading to a metallic bulk throughout the transition. Consequently, TPT has never been observed in the transport studies and only weak signatures of increased scattering around the transition have been observed\textsuperscript{212}. By utilizing a new generation of bulk insulating (Bi\textsubscript{1-x}In\textsubscript{x})\textsubscript{2}Se\textsubscript{3} thin-films we demonstrate, for the first time, the finite-size driven TPT, relying exclusively on transport measurements which is also strongly supported by a theoretical model. This has led to a two-dimensional (thickness and composition) TPT phase diagram, shedding light on how TPT occurs in finite-size TI systems.

Thin-films with thicknesses of 5, 6, 7, 8, 10, 12, 20, and 40 QL (1QL\textasciitilde1nm) (Bi\textsubscript{1-x}In\textsubscript{x})\textsubscript{2}Se\textsubscript{3} (0\textless{}x\textless{}0.1) were grown on an insulating (Bi\textsubscript{1-x}In\textsubscript{x})\textsubscript{2}Se\textsubscript{3} buffer layer and on a sapphire substrate, the details of the growth can be found in Reference 144. In this new generation (Bi\textsubscript{1-x}In\textsubscript{x})\textsubscript{2}Se\textsubscript{3} films, the surface Fermi level is already below the bottom of the conduction band with sheet carriers ($n_{\text{sheet}} \approx 1 – 3 \times 10^{12} \text{ cm}^{-2}$) for the entire thickness range of 5 to 60 QL. To lower the surface Fermi level even further, MoO\textsubscript{3}-capping was deposited \textit{in situ} near room temperature. As mentioned before MoO\textsubscript{3}-capping depletes electrons from the Bi\textsubscript{2}Se\textsubscript{3} layer which lowers the carrier density and pushes the surface Fermi level closer to the Dirac point.

Figure 4.19a summarizes the sheet carrier density ($n_{\text{sheet}}$) and mobility ($\mu$) for samples (before they become insulating) with different thicknesses and In concentrations. Since the surface Fermi level of these samples is in the bulk gap, we can calculate the 2D surface Fermi level if we consider up to quadratic dispersion for surface states $E_F = Ak_F + Bk_F^2$ ($A = 2.02$ eV\textcdot\text{Å} and $B = 10.44$ eV\textcdot\text{Å}^2), where 2D fermi vector $k_F = (2\pi n_{\text{sheet}})^{1/2}$ (assuming similar carrier density for two TSSs)\textsuperscript{223,225}. Based on these equations, Fermi level is estimated to be between 45 to 70 meV for the entire thickness range (with no indium). Thus, on this platform we were able to investigate topological to normal transition in finite-size samples.
Figure 4.19 Transport properties of (Bi$_{1-x}$In$_x$)$_2$Se$_3$ thin-films. a, $n_{\text{sheet}}$ and $\mu$ for different thicknesses as a function of In concentration. Longitudinal resistance versus temperature of (Bi$_{1-x}$In$_x$)$_2$Se$_3$ films grown on the buffer layer for different indium concentration b, 5QL with 0% indium c, 6QL with 0% indium d, 7QL with 0% and 3% indium e, 8QL with 0%, 2%, 3% and 4% indium f, 10QL with 0%, 2%, 4%, 5%, 6% and 8% indium g, 12QL with 6%, 7% and 8% indium h, 20QL with 7%, 8% and 10% indium i, 40QL with 9% indium. Adapted from Reference 14.

Figures 4.19b-i show the sheet resistance as a function of temperature for each thickness and with different In concentration. To be more precise, let us expound on how thickness and In concentration modify TSSs. Each surface Dirac point is protected by time reversal symmetry, thus the only way to open up a gap at the Dirac point, without breaking TRS, is to couple one surface to another. The tunneling rate between top and bottom surfaces is proportional to $e^{-d/2\lambda_{ss}}$ (based on Fermi golden rule), where $d$ is the film thickness and surface state penetration $\lambda_{ss} = \frac{\hbar v_F}{E_g} \frac{\rho}{2\pi}$, substitution yields $e^{-\frac{E_g \lambda_{ss}}{\hbar v_F}}$ where $E_g$ is the bulk gap and $v_F$ is the Fermi velocity of the Dirac band.

Once the penetration depth becomes comparable with the thickness, the hybridization of top and bottom surfaces leads to gap opening. Therefore, changing thickness and/or the bulk gap (for instance, by adding indium in this case) are two knobs to manipulate the tunneling rate, and hence the surface gap opening. Furthermore, once the surface gap is sufficiently large, the $\pi$ Berry phase is no longer effective to prohibit the back scattering and the mobility edge region emerges at the bottom of the gapped surface band; as a result, the system is more susceptible to localization. If the surface Fermi level is sufficiently lowered, in this case with the use of low-carrier-density thin-
films and MoO$_3$ capping, it will be eventually placed inside the surface gap (or the mobility edge region) and the material becomes insulating.

ARPES studies for Bi$_2$Se$_3$ grown on 6H-SiC (0001) and Al$_2$O$_3$(0001) show that the surface hybridization occurs for thicknesses below 6QL$^{179-180}$. Thus, for the 5QL-thick sample in which the surface states are already hybridized and the surface gap is open, since the surface Fermi level is low (either in the surface gap or below the mobility edge) the material is insulating (Figure 4.19b) with no need of indium doping.

Also, the 6QL-thick sample with 0% In, which is on the verge of surface gap opening, is weakly insulating. This is while a conventionally-grown 3QL-thick Bi$_2$Se$_3$ thin-film on sapphire shows metallic temperature dependence due to high surface Fermi level. As shown in Figure 4.19, the critical indium concentrations to make 7QL, 8QL and 10QL samples insulating are 3%, 4%, and 5% (±1%) respectively. Basically, MoO$_3$-capping can be thought of as a natural gate-dielectric whose gating effectiveness persists up to a certain thickness. The MoO$_3$-capping is less effective for 12 and 20 QLs, which needs more indium to become insulating, and eventually becomes neutral for 40 QL thick sample, which remains metallic even with 9% indium. The Ioffe-Regel criterion can be used to identify the metal- insulator transition (MIT)$^{271}$ in this system with spin-orbit coupling. Based on this criterion, a material is a metal if $k_F l > 1$, or an insulator if $k_F l \leq 1$, where $k_F$ is the 2D Fermi vector and the mean-free path $l = \left(\frac{\hbar}{e}\right) \left(2\pi n_{\text{sheet}}\right)^{\frac{1}{3}}$ for 2-D regime.

Figure 4.20a shows the $k_F l$ as a function of indium concentration for each thickness of Bi$_2$Se$_3$ sample. The intersection of dashed line $k_F l = 1$ and each curve is estimated to be the critical indium (within 1% error bar) concentration required for MIT. In our new generation samples, we were able to see the transition in low indium concentration and interestingly in the range where surface states are expected to be gapped based on previous ARPES and TDTS$^{30,212}$. We attribute this MIT to the location of the Fermi level that is in the gap (or below the mobility edge). To further establish the
Figure 4.20 Metal-insulator transition and Weak anti-localization in buffer layer-based \((\text{Bi}_{1-x}\text{In}_x)\text{Se}_3\) films. \(a\), \(k_Fl\) for buffer layer-based \((\text{Bi}_{1-x}\text{In}_x)\text{Se}_3\) films with different thicknesses and different indium concentrations. The intersection of each curve with the dashed line at 1 is where the metal to insulator transition occurs. \(b\), Magneto-conductance for the 10QL sample with different indium concentration. \(c\), \(\tilde{\Lambda}\) for different thicknesses with different indium concentration. Adapted from Reference 14.

was carried out based on HLN formula for magneto-conductance \(\Delta G(B) = \frac{\tilde{\Lambda}e^2}{2\pi\hbar} [\ln\left(\frac{B\phi}{B}\right) - \Psi\left(\frac{1}{2}\right) + \ldots]\), where \(\hbar\) is Planck’s constant, \(\tilde{\Lambda}\) is a parameter related to the number of conducting 2D channels (\(\tilde{\Lambda} = 1\) for each channel), \(B\phi\) is the de-phasing field, and \(\Psi(x)\) is the digamma function.

The change in magneto-conductance for a 10QL sample with different In concentration (Figure 4.20b) shows the gradual suppression of the cusp corresponding to WAL. Further, the transition of \(\tilde{\Lambda}\) from 1 to 0 (Figure 4.20c) indicates the transition from metallic to insulator state.

Figure 4.21a shows the evolution of band structure as a function of SOC for different thicknesses which is computed using the low-energy theory based on the four-band massive Dirac model\(^1\). In these band structures, the band color is calculated based on its orbital content, where the reference colors are: blue (even parity, Bi-like), red (odd parity, Se-like), and yellow (50-50 mixture of Bi and Se). The color of an arbitrary mixed energy eigenstate is calculated by a linear combination according to the probability amplitude for each orbital type. Interestingly, this color scheme reveals that in the TI phase there is a band inversion near the \(\Gamma\) point (i.e. the conduction band has more Se
Figure 4.21 The schematic of the phase transition along with the phase diagram for different TIs. a, A Schematic of the TPT process for different thicknesses and as a function of SOC weakening. b, Phase diagram of the buffer layer-based \((Bi_{1-x}In_x)_2Se_3\) films (the dashed line is a guide to eye for the MIT boundary). Blue triangles correspond to topological-metallic data points where \(k_F l > 1\); green circles are normal-insulating data points with \(k_F l < 1\); brown triangles show normal-metallic data points with \(k_F l > 1\), red squares with \(k_F l = 1\) taken from Figure 4.20a, and the MIT line goes through these points. In the normal insulating region, \(E_{F1}\) and \(E_{F2}\) corresponds to two possible locations of the surface Fermi level, the former is below the mobility edge and the latter is inside the surface hybridization gap. \(E_{mob}\) represents the mobility edge. c, Phase diagram of the ideal TI with the Fermi level at the Dirac point. d, Phase diagram when the surface Fermi level is high and above the bottom of the conduction band, which corresponds to the previous generation \((Bi_{1-x}In_x)_2Se_3\) samples. Adapted from Reference 14.

characteristic, whereas the valence band has more Bi characteristic), while the bands are not affected by the band inversion away from the \(\Gamma\) point. Most importantly, the TSSs (shown in yellow) are 50% mixture of Bi and Se p orbitals (going from even to odd parity). Note that for ultra-thin samples, such as 3QL-thick sample the gap is already open for pure Bi\(_2\)Se\(_3\) (Figure 4.21a, first panel) and the gap gets smaller (for example, it is smaller in 5QL-thick sample in Figure 4.21a, second panel) until it eventually becomes gapless for higher thicknesses. For 10QL-thick sample (Figure 4.21a, third panel) at the critical indium concentration, the gap opens at the Dirac point and
the TSSs will gradually merge into the bulk and the bulk gap increases beyond this point along with
the full bulk band de-inversion.

Let us now explain the qualitative differences of TPT in thin and thick regimes. In the thin
regime, the bulk gap keeps decreasing upon weakening SOC (by means of adding lighter element
indium) which causes TSSs gradually hybridize (mostly at the Dirac point) and start to merge into
the bulk states (this is particularly clear in the picture for 10QL sample that yellow turns to blue
and red right after gap opening). In this process, the bulk gap never completely closes, and this is
evident from the fact that minimum energy of an electron always remains finite due to the spatial
confinement in thin-films. However, in the thick regime, TSSs remain gapless until the critical
transition point where the bulk conduction and balance band meet at the \( \Gamma \) point and TSSs merge
into the bulk states (see Figure 4.21a, fourth panel). After this point, the bulk gap increases and the
bulk bands de-invert.

Figure 4.22 compares the surface and bulk gap opening in the thick and thin regimes. We
emphasize that the bulk gap never closes across the TPT point in the thin regime, in contrast to
what is commonly expected. For further clarification, we compare our phase diagram with the ones
for two extreme cases: 1) the ideal case when the Fermi level is exactly at the Dirac point (Figure
4.22c), and 2) when the Fermi level is high and resides in the conduction band (Figure 4.22d). In
the first case, the entire normal region beyond TNT becomes insulating and TNT and MIT lines
fall on top of each other; whereas in the second case the Fermi level never locates inside the surface
gap/mobility edge, and thus the normal region remains metallic. The phase diagram for higher
Fermi level (case 2) is more similar to the phase diagram shown by Wu et al. for TDTS on
conventionally grown \((\text{Bi}_{1-x}\text{In}_x)\text{Se}_3\) samples on sapphire\(^{212}\).

The phase diagram in Figure 4.21b summarizes the main result. It consists of three different
regions: topological-metallic, normal-metallic, and normal-insulating. This phase diagram focuses
Figure 4.22 Theoretical phase diagram and simulated thickness dependence of the surface and bulk gap developments upon TPT. 

a, Theoretical phase diagram of TPT for the thin regime. The phase boundaries are determined using the four-band Dirac model. The right panel demonstrates different phases as well as the transitions between them. The mass parameter \( m \) on the horizontal axis represents the SOC strength.

b, The critical point for TNT in thick samples is sharp and well-defined but is rather broad and not indisputably definable in thin samples. The vertical dashed line \( (m = 0) \) is the phase boundary for the thick regime. Both axes are normalized by the Dirac velocity. Adapted from Reference 14.

on the thickness \( \leq 20QL \) and indium concentration <10%. The black solid curve defines the boundary of topological to normal transition (TNT), which can qualitatively be thought as twice the penetration depth. It starts from 6QL where the sample is assumed to be on the verge of surface gap opening at 0% In concentration and approaches asymptotically to 7% In concentration which is the critical value of \( x \) for an infinitely thick sample. In conventional samples due to high Fermi level, the whole normal region which is beyond the TNT boundary remained metallic (normal-metallic). However, with the use of low-carrier-density samples grown in this work, we are able to probe this region more carefully. The dashed line in the normal region separating the metallic and insulating sub-regions (MIT boundary) is defined by data points corresponding to In concentration that gives \( k_F l = 1 \) for each thickness. Below \( k_F l = 1 \), the system is normal-insulating, while above \( k_F l = 1 \), the system is normal-metallic. The location of this MIT boundary solely depends on where the Fermi level is located and in the extreme case when the Fermi level is at the Dirac point, it falls on the TNT boundary. We can consider TNT and MIT combined as TPT process. Note that the presence of the mobility edge can move the MIT boundary. In our theoretical model, with the mobility edge, the insulating behavior is observed at lower percentage of In (the metallic
part shrinks) and also at higher thicknesses (Figure 4.22a). Our case is rather more interesting as we are able to split the normal region into two sub-regions of insulating and metallic showing that the location of Fermi level matters for how the system behaves; nonetheless, if the Fermi level can be lowered further (close to the ideal case) for example by gating or further growth engineering, the insulating regime grows, and the metallic regime shrinks. Further study with external gating would allow us to study this regime in more details. For more information about this work see Reference 14 and its Supporting Information.

4.4 Conclusion

All in all, the observation of topological quantum effects shows the importance of defect suppression and lowering the Fermi level in the bulk gap and as close as possible to the Dirac point where the physics is governed by topological surface states. The low-carrier-density interface-engineered Bi$_2$Se$_3$ films with a high electron affinity MoO$_3$-capping allowed the observation of QHE for the first time in this binary TI system. These films enabled the observation of quantized Faraday/Kerr rotation for the first time in any TI system. With the use of low-Fermi-level buffer layer-based ($Bi_{1-x}In_x$)$_2$Se$_3$ thin-films, we were able to study the topological phase transition in finite-size samples and observe metal-to-insulator transition for the first time in transport measurements. These results introduce a new version of topological phase transition notion in finite-size samples and shed light on the role of spatial confinement as an extra tuning parameter controlling the topological critical point and these samples can be used as a promising platform to study novel topological states, such as Dirac and Weyl semimetals.

Furthermore, the ultralow-carrier-density Sb$_2$Te$_3$ films allowed us to reach an extreme quantum limit of the topological surface state and explore the zeroth LL of TSS that was previously inaccessible due to large electron-hole puddles. In particular, we showed that the zeroth LL harbors a new insulating phase with a distinct dynamical critical exponent, implying that the origin of this
insulating phase is fundamentally different from that of non-topological systems. This study opens a new door for testing various other proposals, including surface exotic phases\textsuperscript{250, 272} and topological quantum computing\textsuperscript{273}. In addition, considering that this bismuth-free TI system precludes any possible incorporation of $\text{Bi}_2\text{Te}_3$, whose Dirac point is buried in the valence band, it provides a promising new platform for high-temperature quantum anomalous Hall and other magneto-topological effects\textsuperscript{150-151, 156, 161}, when combined with proper magnetic coupling schemes.
Chapter 5

Epilogue

After the prediction and the subsequent discovery of TIs with the use of highly surface-sensitive probes ARPES and STM about a decade ago, many efforts were made to grow high-quality TI thin-films which eventually gave birth to the first-generation of TI materials. However, it soon was found that the grown films harbor high level of defect densities with parasitic bulk conduction coupling the top and bottom surfaces and masking the TSS response. Such high level of defect density was a major obstacle to utilizing the unique properties of topological surface states in TIs. After that, the goal was to develop true bulk-insulating TI systems and to lower the Fermi level as close as possible to the Dirac point where the physics is governed by TSS. In this dissertation, we discussed the effective growth methodologies and various defect engineering schemes that have been developed and implemented over the past number of years to realize this goal and to overcome the bulk conduction problem. Finally, the advanced defect and interface engineering techniques led to new generation bulk-insulating TI films with extremely low level of defect density and with their Fermi level close to the Dirac point, which revealed a series of topological quantum effects, such as quantum Hall effect, quantum anomalous Hall effect, quantized Faraday/Kerr rotations, axion insulators, topological quantum phase transitions, zeroth-Landau-level phases, etc. This once again emphasize the critical role of defect and interface engineering in tailoring the materials properties and achieving high-quality films.

5.1 Summary of the results

In chapter 3, we discussed the origin of defects in TIs and introduced various growth techniques that can be deployed to suppress these defects. By providing several experimental examples, we explained that even though the layered TIs grow in van der Waals fashion, the substrate still plays
a crucial role in determining the final properties of TI films. We showed that how it was made possible to achieve bulk-insulating TI films with unprecedented low carrier density by growing them on artificially-designed and virtually-grown structurally/chemically-compatible substrates. We expound on the growth and transport properties of the low-carrier-density Bi$_2$Se$_3$ thin-films grown on the In$_2$Se$_3$/BiInSe$_3$ buffer layer as well as the ultralow-carrier-density (Ti-doped) Sb$_2$Te$_3$ grown on the In$_2$Se$_3$/(Sb$_{0.65}$In$_{0.35}$)$_2$Te$_3$ buffer layer. Furthermore, both ARPES and STM confirmed that the films are topological insulators. Transport in conjunction with cyclotron resonance measurements showed that the response in these films was dominated by TSS carriers. We also emphasized the role of a proper capping layer in stabilizing the properties of these low-Fermi-level films, and some effective capping layers, such as Se, MoO$_3$, and (Sb$_{0.65}$In$_{0.35}$)$_2$Te$_3$.

Moreover, the observation of the QHE is strongly dependent on how low the carrier-density of the system is. The absence of low carrier-density TI systems prevented an immediate observation of the QHE in transport in the early TI systems, despite QH physics having been observed in 2DEG systems for several decades$^8$. As a matter of fact, QHE in TIs was measured even after the observation of the previously unobserved phenomenon of the quantum anomalous Hall effect, which implies that obtaining low-carrier-density TI systems remained challenging for the researchers for quite a long time after TIs discovery. In chapter 4, we explained how the new-generation low-carrier density TI films allowed the observation of the TSS-originated QHE. With the use of interfaced-engineered low-Fermi-level Bi$_2$Se$_3$ films, the TSS-originated QHE was observed for the first time in pure binary Bi$_2$Se$_3$ films without any impurity addition or gate-voltage application. It was also shown that the required magnetic field for observing the QHE was substantially reduced in ultralow carrier-density buffer layer-based Ti-doped Sb$_2$Te$_3$ films, which exhibit QHE at fields as low as 5.5T. Additionally, these Sb$_2$Te$_3$ TI films demonstrated quantum Hall to insulator transition at high magnetic fields for the first time in a TI system. The insulating behavior at high fields was attributed to either a coherent superfluid state or a magnetic field-driven Hall-insulator. Although QIT implies that the topological (Chern) number changes from 1 to most
cases) to 0, this does not *ipso facto* imply that every QIT should belong to the same universality class. In fact, the extracted dynamical critical exponent \( z \approx 2 \) in TI Sb\(_2\)Te\(_3\) films is clearly different from \( z \approx 1^{260-261} \) of 2DEGs, suggesting that QITs in these two systems should belong to different universality classes.

Furthermore, the low carrier density samples enabled the observation of quantized Faraday and Kerr rotation. It also allowed the study of the topological phase transition in the finite-size regime, and for the first time, the metal to insulator transition was observed in transport when the system undergoes TPT.

### 5.2 Outlook and future direction

Along the QIT direction, an interesting future research is to investigate the underlying mechanism responsible for the discrepancy in the critical exponents of this transition for TIS and 2DEGs. Whether this is due to the difference in zero-field band structures of TIs and 2DEGs, or due to some other higher order effects is an open question for future studies. Further studies of QITs, specifically extracting \( z \) and \( \nu \), in other Dirac systems such as graphene could help resolve this question.

Moreover, as explained before, the gapless topological surface states of TIs are protected by TRS. Interesting phenomena are predicted when TRS is broken, opening a gap around the Dirac point\(^{274-277}\). TRS can be broken either by proximitized magnetism (by growing a TI on a magnetic layer or vice versa) or by inducing intrinsic magnetism (by doping TI films with magnetic impurities). The strong SOC of TIs, which leads to band inversion and thus the existence of TSS, combined with a magnetically-induced exchange gap (around the Dirac point) could give rise to the quantum anomalous Hall effect (QAHE). Several ARPES and STM/STS studies investigated

\(^{11}\text{Part of section 5.2 is based on a review paper Salehi, M.; et al. From classical to quantum topological surface states: evolution of thin film topological insulators with defect engineering, which is in preparation at the time of writing this dissertation.}\)
the magnetism-induced gap and the QAHE in TIs\textsuperscript{278-281}. Like the $\nu = 1$ QH system, the QAHE system features an insulating two-dimensional bulk and a single chiral edge mode that conducts along the one-dimensional boundary of the system, resulting in quantized longitudinal and Hall conductivities $\sigma_{xx} = 0$ and $\sigma_{xy} = \frac{e^2}{h}$ (or, equivalently, $\rho_{xx} = 0$ and $\rho_{xy} = \frac{\hbar}{e^2}$). Unlike the QHE, the QAHE requires no external magnetic field, but only magnetism. Inspired by Duncan Haldane’s 1988 proposal for the QHE without Landau levels (LL)\textsuperscript{282}, which shared the 2016 Nobel prize, the QAHE was predicted in 2008 in doped 2D TI HgTe quantum wells\textsuperscript{274}. It was soon found, however, that magnetically doped HgTe samples exhibit paramagnetism, rather than ferromagnetism\textsuperscript{283}.

It was subsequently predicted that ferromagnetic (FM) order can be induced by the van Vleck mechanism in thin-films of pnictogen chalcogenide TIs when doped with a proper transition metal element, either Cr or Fe\textsuperscript{275}. It was predicted that, in thin magnetic TI films (having hybridized surface states), the QAHE is realized when the FM exchange energy exceeds the hybridization gap. Initially, Bi$_2$Se$_3$ seemed a promising platform to realize the QAHE, as the Dirac point of its surface state is within its large bulk band-gap of 0.3 eV. However, the QAHE was never realized in magnetically-doped Bi$_2$Se$_3$ thin-films, despite much effort. Only a small anomalous Hall effect (AHE) of a few Ohms was seen in transport measurements in V-doped Bi$_2$Se$_3$ samples\textsuperscript{284} and, more recently, in low-carrier-density buffer-layer-based Bi$_2$Se$_3$ thin-films with Cr modulation-doping\textsuperscript{285}. The lack of QAHE in Bi$_2$Se$_3$ may be understood as a consequence of only the Bi atoms contributing to SOC: upon magnetic impurity substitution into the Bi sites, SOC weakens and the compound becomes a trivial insulator\textsuperscript{286-287}. In contrast, the cation and anion atoms both contribute to SOC in Te-based TIs. Since topological non-triviality is more robust to impurity doping in Te-based materials than in Se-based materials\textsuperscript{287-288}, the search for a quantum anomalous Hall insulator should originate from Te-based topological insulators.
The QAHE was first realized in 2013\textsuperscript{145}, 132 years after the discovery of AH effect\textsuperscript{289}. Using the ternary compound Bi\textsubscript{x}Sb\textsubscript{2-x}Te\textsubscript{3} magnetically doped with Cr, the $E_F$ was lowered by adjusting the Bi/Sb ratio. The films were grown by MBE on a SrTiO\textsubscript{3} dielectric substrate, and the chemical potential was fine-tuned into the magnetic exchange-induced gap by electrostatic-gating (the exchange gap has not been resolved by ARPES due to the small bandwidth). At 25 mK, a quantized Hall resistance ($\frac{h}{e^2} \approx 25.8k\Omega$) at zero field, concurrent with small longitudinal resistance ($0.098\frac{h}{e^2}$), was observed at zero field. In 2015, A modulation doping scheme, where the magnetic dopants are concentrated in 1 nm-thick Cr-rich layers, has also allowed the QAHE to be seen at higher temperatures. In a penta-layer superlattice of (1nm (Bi\textsubscript{0.22}Sb\textsubscript{0.78})\textsubscript{2}Te\textsubscript{3} / 1nm Cr\textsubscript{0.46}(Bi\textsubscript{0.22}Sb\textsubscript{0.78})\textsubscript{1.54}Te\textsubscript{3} / 4nm (Bi\textsubscript{0.22}Sb\textsubscript{0.78})\textsubscript{2}Te\textsubscript{3} / 1nm Cr\textsubscript{0.46}(Bi\textsubscript{0.22}Sb\textsubscript{0.78})\textsubscript{1.54}Te\textsubscript{3} / 1nm (Bi\textsubscript{0.22}Sb\textsubscript{0.78})\textsubscript{2}Te\textsubscript{3}) on a InP(111) substrate the QAHE was observed at temperatures up to almost 2 K (the signature of QAH remains up to 4.2k)\textsuperscript{147}. Shortly after, the QAHE was, surprisingly, also realized in V-doped Bi\textsubscript{x}Sb\textsubscript{2-x}Te\textsubscript{3} films\textsuperscript{51}. V-doped films were not initially considered a candidate QAH insulator because simulations predicted the formation, upon introduction of substitutional V dopants, of $d$-orbital impurity bands at the Fermi energy\textsuperscript{275}. Nevertheless, a V\textsubscript{0.11}(Bi\textsubscript{0.29}Sb\textsubscript{0.71})\textsubscript{1.89}Te\textsubscript{3} film exhibited the QAHE with Hall conductivity $(0.9998 \pm 0.0006)\frac{e^2}{h}$ and nearly vanishing longitudinal resistivity $(3.35 \pm 1.76)\Omega$ at 25 mK and zero field\textsuperscript{181}.

Co-doping has been proposed as a step towards higher temperature QAH systems\textsuperscript{290}. Recently, using Cr and V co-doping, a 5QL-thick MBE-grown (Cr\textsubscript{0.16}V\textsubscript{0.84})\textsubscript{0.19}(Bi\textsubscript{x}Sb\textsubscript{1-x})\textsubscript{1.81}Te\textsubscript{3} film demonstrated $\rho_{xy} = \frac{h}{e^2}$ and $\rho_{xx} = 0.009\frac{h}{e^2}$ at 300 mK\textsuperscript{291} where $T_c = 25$ K. In contrast, modulation doped films with a Cr-rich layer and a V-rich layer reach an altogether different state. Because Cr-doped and V-doped TIs have different coercivities, the magnetization of the Cr-rich layer and the V-rich layer flip at different external fields. Between the two coercive fields, the magnetization of the two layers points oppositely\textsuperscript{157,292}. The resulting state, known as an axion insulator, features an
everywhere-gapped surface yet no edge states. In the axion insulator state, the surfaces of the TI are predicted to exhibit a topological magnetoelectric effect, wherein an applied magnetic field produces an electrical polarization and vice versa \(^{248,249}\).

The extremely low temperature (\(\leq 2\) K) required to observe perfect QAHE\(^{145-148,151}\) remains a major barrier for applicability of this exotic effect. Yet, it is believed that achieving a higher Curie temperature, along with stronger long-range ferromagnetic coupling, could result in the observation of the QAHE at even room temperature. To this end, MBE-grown TI films seem more promising: the Curie temperature \((T_c)\) was increased from 20 K for bulk material to 177 K for MBE-grown Cr-doped Sb\(_2\)Te\(_3\) thin-films, and from 24 K for bulk material to 190 K for MBE-grown V-doped Sb\(_2\)Te\(_3\) thin-films\(^{293-294}\). The higher Curie temperatures may result from increased magnetic dopant solubility at the lower temperatures used for MBE growth. While increased magnetic doping may raise \(T_c\) and widen the magnetic exchange gap, doing so degrades the sample’s crystallinity and eventually destroys the TSS, when either the bulk becomes topologically trivial or the impurity concentration exceeds the solid solubility of the host material. In this regard, the low-defect-density films could accommodate more magnetic dopants, and hence the \(T_c\) can be raised by further doping the material with magnetic impurity without destroying the crystal structure or topological nature of the system.

An alternative to magnetic doping is to introduce magnetism into the TI by proximity. In this scheme, a nonmagnetic TI is grown on top of an insulating magnetic layer (or vice versa). Inducing magnetism by proximity allows independent optimization of the magnetic and electronic properties of the system. In particular, a high \(T_c\) magnetic insulator may be selected without concern of dopant solubility or dopant-induced disorder and scattering in the TI. Much work has sought a high \(T_c\), out-of-plane magnet capable of proximitizing an exchange gap in a TI (although, in a special case, an in-plane magnet could also induce a QAHE\(^{295}\)). Candidate magnetic layers include the ferromagnetic insulator EuS\(^{184-187}\), Ferrimagnet insulator Y\(_3\)Fe\(_5\)O\(_{12}\)(YIG)\(^{188-191,296}\) with in-plane
magnetization and Tm$_3$Fe$_5$O$_{12}$ (TIG) with out-of-plane magnetization$^{192}$, antiferromagnetic conductor CrSb$^{193,297}$ with Néel temperature ($T_N$) $\sim$ 700 K, and magnetic insulator BaFe$_{12}$O$_{19}$ with $T_c \sim 723$K$^{194}$. Although a QAHE has yet to be observed, recently an anomalous Hall effect was observed up to 400 K in a bilayer film of Bi$_x$Sb$_{2-x}$Te$_3$ and the ferrimagnetic insulator TIG, grown on a (111)-oriented substituted gadolinium gallium garnet (SGGG) substrate$^{192}$. TIG has $T_c = 560$ K.

When coupled with Bi$_2$Te$_3$, the $T_c$ of the out-of-plane magnet Cr$_2$Ge$_2$Te$_6$ is enhanced from 61 K to 108 K, accompanied by an anomalous Hall effect$^{195}$. Even more dramatically, Bi$_2$Se$_3$/EuS bilayers have exhibited interfacial magnetism at even room temperature, despite bulk EuS having $T_c = 17$ K$^{298}$. This implies that electronic interactions between the Eu atoms and the TI surface states enhances magnetic order in both materials, causing high temperature magnetization in the 2QL-thick TI. Furthermore, proximity to a TI changes the magnetic anisotropy of EuS. While EuS films favor in-plane magnetization, when coupled to a thin TI, the SOC leads to an out-of-plane magnetic moment in the TI surface. Despite the interesting progress in TI/magnet heterostructures, materials are still far from a sufficient exchange coupling to realize the QAHE. Improved interface-engineering may enhance the exchange coupling between TIs and FM insulators.

Although extremely low temperature QAH temperature is an issue that has to be resolved for application purposes, realization of the QAH and dissipationless edge states requires neither external magnetic field nor high mobility samples which could pave a path for developing low-power-consumption, topological quantum electronic and spintronic devices. While a highly quantized QAHE has been established at milli-kelvin temperatures, the dream of a high temperature QAH system for technological applications requires improved magnetic topological insulators (MTI) materials platforms. To this end, we have seen the importance of both strengthening the exchange coupling and suppressing disorder. Regarding the former point, magnetic order in topological insulators needs to be better understood. In pursuit of MTIs with less disorder, we
nominate interface-engineered Sb$_2$Te$_3$ platforms, featuring ultra-low carrier-density and suppressed interfacial disorder. In these materials, suppressed disorder would prevent smearing of the exchange-induced bandgap (as well as charged impurity sites that could produce hopping conduction), preserving the clean-limit exchange gap. Furthermore, developing such Bi-free system precludes any possible incorporation of Bi$_2$Te$_3$, whose Dirac point is buried far beneath the valence band, in the system, yielding a more well-defined and cleaner Dirac point.

Furthermore, because of the spin-momentum locking in TIs, a properly-designed ferromagnetic/TI heterostructure is believed to be a promising platform for realizing practical spintronic and topotronic devices.

Additionally, combination of superconductors and TIs$^{299-300}$ leads to exotic physics phenomena. For instance, it is believed that proximity effect between a TI and an s-wave superconductor would lead to spinless p-wave superconductivity at the interface which supports Majorana bound states at magnetic flux vortices. Additionally, in principle, a superconductor in good electrical contact with a QAH insulator should proximitize a superconducting gap in the QAH insulator, forming a topological superconductor (TSC) with chiral Majorana edge modes (CMEMs) propagating along the boundary of the TSC. Under an appropriate applied field, it is predicted that only a single CMEM propagates, which leads to a half-quantized two-terminal conductance$^{301}$. Such a half-quantized conductance plateau was recently reported$^{161}$; it has been noted, however, that a half-quantized conductance plateau could instead result from dissipative effects in the QAH insulator, rather than the presence of CMEMs$^{302-303}$. All in all, there are still plenty unexplored areas in this branch of TI-research that can be studied with potential applications in quantum computers; and in this regard the new generation TI samples with low defect density could be a promising platform.

Last but not least, a compelling future direction is developing new topological materials beyond the layered pnictogen chalcogen-based TIs, such as complex oxides. Along this direction,
it would be useful to look at the phase diagram of quantum materials in Figure 5.1, provided by Reference 304, where the horizontal axis is the SOC strength and the vertical axis is the electron-electron interactions (on-site coulomb interaction). Most of the materials we are familiar with, like band insulators and metals, reside in the bottom left of this diagram, where both correlation and the SOC are weak. The materials explained in this dissertation belong to the strong SOC but weakly interacting region of the diagram, marked as “topological insulator or semimetal”, where the band structure picture is still valid. Moving along the vertical axis and towards stronger correlation where the band structure picture breaks down, sits the Mott insulator which is known to be an antiferromagnet. Now, it would be interesting to develop and study the novel topological materials that emerge as a result of the cooperative interplay of SOC and correlation. For example, assume we start with the region with strong SOC and move vertically towards stronger correlation, where interaction could stabilize exotic phases of matter, such as axion insulators, Weyl semimetals and topological Mott insulators, which could be very stable even at higher temperatures due to the presence of the interaction. This can be done experimentally by, for example, substituting elements with more localized orbitals.
Figure 5.1 The generic phase diagram of quantum materials. Phase diagram is in terms of spin-orbit coupling strength \((\lambda/t)\) on the horizontal axis and the electron-electron interaction \((U/t)\) on the vertical axis. Adapted from Reference 304.

Another direction is to start from a strongly correlated Mott insulator which is considered an ordered phase and move horizontally towards stronger SOC by, say, doping the system with heavier elements we enter a regime where spin-spin interaction is direction dependent that could eventually lead to frustration, an essential ingredient for spin liquid phase to emerge. Spin liquids are exotic class of paramagnets which do not form any ordered phase even at absolute zero temperature. A particular set of spin liquids are called topological spin liquids (e.g. Kitaev spin liquids) which are of great interest for both fundamental physics (macroscopic realization of entanglement) and application purposes (quantum computing). Iridates and double perovskites are some of the proposed platforms to explore this phase diagram, especially if they can be grown epitaxially. Epitaxial growth, and in particular MBE, gives a wide range of tunability over dimensionality and strain, by growing the film on substrates with smaller or larger lattice constant, and it allows fine-tuning the chemical potential via systematic doping process. MBE also enables the growth of materials that are not stable in the bulk form which adds even more to the many merits of this technique.
References


effect in the ternary topological insulator (Bi$_{x}$Sb$_{1-x}$)$_{3}$Te$_{3}$ by composition tuning. *Nature Nanotechnol.* **2011**, *6* (11), 705-709.


183


140. Xiong, J.; Petersen, A. C.; Qu, D.; Hor, Y. S.; Cava, R. J.; Ong, N. P., Quantum oscillations in a topological insulator Bi₂Te₂Se with large bulk resistivity (6 Ω cm). *Physica E: Low-dimensional Systems andNanostructures* 2012, 44 (5), 917-920.


185


182. Lee, Y. F.; Kumar, R.; Hunte, F.; Narayan, J.; Schwartz, J., Microstructure and transport properties of epitaxial topological insulator Bi$_2$Se$_3$ thin films grown on MgO (100), Cr$_2$O$_3$ (0001), and Al$_2$O$_3$ (0001) templates. *J. Appl. Phys.* 2015, 118 (12), 125309.


214. Liu, J.; Vanderbilt, D., Topological phase transitions in (Bi\(_{1-x}\)In\(_x\))\(_2\)Se\(_3\) and (Bi\(_{1-x}\)Sb\(_x\))\(_2\)Se\(_3\). Phys. Rev. B 2013, 88 (22), 224202.


237. Xia, B.; Ren, P.; Sulaev, A.; Liu, P.; Shen, S.-Q.; Wang, L., Indications of surface-dominated transport in single crystalline nanoflake devices of topological insulator Bi$_{1.5}$Sb$_{0.5}$Te$_{1.8}$Se$_{1.2}$. *Phys. Rev. B* **2013**, *87* (8), 085442.


