RICE UNIVERSITY

Experiments on quantum phases in InAs/GaSb bilayers: Topological insulator and exciton condensation

by

Lingjie Du

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Rui-Rui Du, Chair
Professor of Physics and Astronomy

Matthew Foster
Assistant Professor of Physics and Astronomy

Jun Lou
Professor of Materials Science and Nano Engineering

HOUSTON, TEXAS
October 2016
Experiments on quantum phases in InAs/GaSb bilayers: Topological insulator and exciton condensation

by

Lingjie Du

A THESIS Submitted IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Rui-Rui Du, Chair
Professor of Physics and Astronomy

Matthew Foster
Assistant Professor of Physics and Astronomy

Jun Lou
Professor of Materials Science and Nano Engineering

HOUSTON, TEXAS
October 2016
ABSTRACT

Experiments on quantum phases in InAs/GaSb bilayers: Topological insulator and exciton condensation

by

Lingjie Du

Recent developments in Quantum Spin Hall (QSH) effect have triggered much attention in inverted InAs/GaSb Quantum wells (QWs), which are the leading material in QSH systems. Inverted InAs/GaSb QWs are a type II heterostructure with the broken gap, where two dimensional (2D) electrons and holes are confined in spatially separated QWs. From the 1970s until now, the ground state of this structure has been discussed between two candidates: exciton insulator (BCS type exciton condensation) and hybridization gap. The QSH effect was theoretically proposed in the bulk hybridization gap. Although pioneer works about QSH effect have been performed, the conductive hybridization gap limits further exploration. For example, the existence of the QSH effect in this system is still not conclusive. In this thesis, through double-gate modulation, we investigated the whole phase in the inverted band of this structure. We observed two distinct quantum phases: time reversal symmetry (TRS) QSH insulator in the deeply inverted regime and exciton insulator in the shallowly inverted regime. In the deeply inverted regime, with the strain effect in InGaSb QW, we realized the insulating hybridization gap for the first time, which gave us the opportunity to observe TRS QSH effect in this system for the first time. With the largest bulk gap in known QSH systems, we observed the helical
edges had the longest coherence length (nearly 13µm) and were more stable against temperature, compared with previous results, which paved the way to construct the room temperature topological circuit. In the shallowly inverted regime, the quantized plateau of QSH effect was observed in mesoscopic devices for the first time. Surprisingly, this helical edge mode was robust under the high magnetic field, demonstrating the first TRS broken QSH insulator. This novel quantum phase could not be understood in the single particle topological theory. Further studies showed that the bulk gap was dominated by exciton gap instead of hybridization gap. We performed the low temperature transport and Terahertz transmission measurement on the bulk exciton gap, and observed the solid evidence for the existence of BCS-like exciton condensation, which was under search for more than fifty years. Furthermore, we performed one dimensional Coulomb drag experiments in the topological circuit. We observed positive and negative drag results dependent on the temperature, indicating the charge symmetry and many-body correlation.
Acknowledgments

First, I would very much like to thank my advisor Professor Rui-Rui Du for bringing me into the field of experimental physics, and giving me a great opportunity to work in the cutting-edge field of topological physics. I am deeply indebted to his invaluable guidance and support through my projects. He, playing the role of not only the advisor but also avuncular mentor, leaded me into the world of quantum transport, and guided the road to become a responsible and earnest physicist. He taught me the experiments hand by hand. I can still recall the details of his guides in my first experiment in He3 system, in National high magnetic field laboratory at Tallahassee, and in 35T high magnetic field at Tallahassee.

I would also like to thank to my other committee members Professor Matthew Foster and Professor Jun Lou for reading the thesis and serving on the thesis committee. Professor Matthew Foster was not only on my PhD and Master committees but also gave numerous discussions in the Coulomb drag project. I am also very grateful to Professor Junichiro Kono for his insightful discussion and generous support in the exciton insulator project. Thanks also go to Professor Anthony Chan, who served on my master thesis committee.

High quality InAs/GaSb quantum wells were grown by molecular beam epitaxy by Dr. Gerard Sullivan and Amal Lkhlassi at Teledyne. Without their continuous supports, my graduate work is hard to finish. Thank you!
I would also like thank my labmates, Dr. Knez Ivan, who impartd the initial cleanroom fabrication skill in InAs/GaSb to me, Dr. Yanhua Dai, who helped me in Scan Electron Microscopy, Hongyu Xiong, Dr. Ruiyuan Liu, and Jie Zhang for their assistance on experimental processes.

Thanks also go to our collaborators Dr. Kai Chang and Dr. Wenkai Lou from Institute of Semiconductors, Chinese Academy of Sciences, for their continues helps in theoretical calculations.

I appreciated the generous help and expert technical assistance from Dr. Tim Murphy, Dr. Ju-Hyun Park and Glover Jones at the National High Magnetic Field laboratory. Dr. Ju-Hyun Park helps me a lot in the experiments of high magnetic field.

I am grateful to Dr. Tim Gilheart and Dr. Kelley Bradley for training me and keeping the clean room running, at Rice University and University of Houston, respectively. Thanks also go to Dr. Gang Liang for training and helping me in the fabrication. Without their efforts, this work could not be finished.

I am also indebted to Xinwei Li from Prof. Kono’s group who worked together with me in the optical experiments searching for exciton insulator, and Dr. Xuan Wang from Prof. Kono’s group who helped me a lot in the Labview and many other things. Thanks also go to Dr. Jian Lin from Prof. Tour’s group for the helps in the cleanroom, Dr. Fangfang Wen from Prof. Halas’ group, Dr. Yajing Li and Dr. Heng Ji from Prof. Natelson’s group who got me started with e-beam lithography, as well as to Loah Stevens and Geetanjali Vengurlekar for the careful proofreading of this thesis.
I acknowledge the Rice University Graduate Research Fellowship and Robert A. Welch Predoctoral Fellowship for providing financial support. This work has been supported by the National Science Foundation, the Department of energy, and the Welch foundation.

Specially, I appreciated the department of Physics for giving me the opportunity of the PhD program. Five years ago, my PhD application was rejected by another program at Rice due to a poor TOEFL score, but the department of Physics gave me a special consideration and another opportunity to pursue the graduate study at Rice. Thank Umbe Cantu, Barbara Braun, and others who helped me very much, for your trust and efforts!

Finally, I thank and dedicate this thesis to my wife Minjie for her love, the support and the trust in the past years, following me across the world in my choice to pursue a carrier in science.
# Contents

Abstract ii

Acknowledgments iv

List of Figures x

1. Introduction
   1.1 Topological system 1
   1.2 Inverted InAs/GaSb heterostructure 4
   1.3 Exciton insulator 5
   1.4 About this thesis 7

2. Background
   2.1 6.1Å family of III-V semiconductor 9
   2.2 Inverted regime 14
   2.3 Hybridization 17
   2.4 Quantum Spin Hall effect 19
   2.5 Landauer-Büttiker formula 22
   2.6 Exciton 27
   2.7 BCS like exciton condensation 34

3. Experiment basic
   3.1 Fabrication 42
   3.2 Conductance measurement 46
   3.3 Capacitance measurement 48
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4</td>
<td>Charge neutral point density</td>
<td>52</td>
</tr>
<tr>
<td>3.5</td>
<td>Previous experiments</td>
<td>56</td>
</tr>
<tr>
<td>4.</td>
<td>Quantum Spin Hall effect in strained InAs/InGaSb quantum wells</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>History</td>
<td>59</td>
</tr>
<tr>
<td>4.2</td>
<td>Strain effect in InAs/InGaSb</td>
<td>61</td>
</tr>
<tr>
<td>4.3</td>
<td>Transport properties of bulk states in strained-layer InAs/InGaSb Quantum wells</td>
<td>66</td>
</tr>
<tr>
<td>4.4</td>
<td>Long coherent helical edge</td>
<td>71</td>
</tr>
<tr>
<td>4.5</td>
<td>In-situ electrostatic manipulation of helical edge</td>
<td>77</td>
</tr>
<tr>
<td>4.6</td>
<td>Time Reversal Symmetry Quantum Spin Hall insulator</td>
<td>82</td>
</tr>
<tr>
<td>4.7</td>
<td>Conclusion</td>
<td>86</td>
</tr>
<tr>
<td>5.</td>
<td>Time Reversal Symmetry broken Quantum Spin Hall effect</td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>Structure detail and wafer characterization</td>
<td>90</td>
</tr>
<tr>
<td>5.2</td>
<td>Shallowly inverted band</td>
<td>92</td>
</tr>
<tr>
<td>5.3</td>
<td>Insulating bulk gap</td>
<td>96</td>
</tr>
<tr>
<td>5.4</td>
<td>Quantized plateau</td>
<td>98</td>
</tr>
<tr>
<td>5.5</td>
<td>Time reversal symmetry broken helical edge state</td>
<td>104</td>
</tr>
<tr>
<td>5.6</td>
<td>Conclusion</td>
<td>110</td>
</tr>
<tr>
<td>6.</td>
<td>BCS like exciton condensation</td>
<td></td>
</tr>
<tr>
<td>6.1</td>
<td>Motivation</td>
<td>112</td>
</tr>
<tr>
<td>6.2</td>
<td>Tuning into dilute limit of carrier densities</td>
<td>115</td>
</tr>
<tr>
<td>6.3</td>
<td>Origin of hard gap</td>
<td>120</td>
</tr>
<tr>
<td>6.4</td>
<td>Formation of excitonic ground state in shallowly inverted band</td>
<td>130</td>
</tr>
</tbody>
</table>
6.5 BCS like exciton condensation 136

6.6 Quantum phase transition driven by high magnetic field 147

7. 1D Coulomb Drag in topological insulator

7.1 Introduction 158

7.2 Device fabrication 161

7.3 Characterization 166

7.4 Coulomb drag 169

8. Conclusion 181

Reference 186

Appendix 193
List of Figures

Figure 2.1 Band lineups in the 6.1Å family material and the lattice constants of each material. The shaded parts show the energy gaps with all energies in eV. The figure is adopted from Ref. (29). ................................................................. 10

Figure 2.2 a) shows the band diagram of inverted InAs/GaSb QWs. The electrons are confined in InAs layer and the holes are confined in GaSb layer. The red dashed line indicates the top of hole band while the blue dashed line indicates the bottom of electron band. b) presents the energy dispersion of electron band E and hole band H. The dashed line shows the crossing of the uncoupled E and H. The anti-crossing opens due to the band hybridization. The front gate and back gate shown in a) can modulate the band separation and the Fermi energy. ................................................................. 11

Figure 2.3 Schematic of inverted and noninverted band in InAs/GaSb QWs.... 12

Figure 2.4 Energy band dispersion of 17nm InAs and 5nm GaSb QWs under various perpendicular electric fields. The figure is adopted from Ref. (31). ... 16

Figure 2.5 Band dispersion based on eight band self consistent calculation(10). a) shows the trivial noninverted band without helical edge states, where the InAs layer is 8.1nm and GaSb layer is 10nm. b) shows the nontrivial inverted band with helical edge states, where the InAs layer is 10nm and GaSb layer is 10nm. Figures are adopted from (10). ................................................................. 22

Figure 2.6 Six-terminal Hall bar device with QSH edge states ......................... 25

Figure 2.7(left) Four-terminal π bar device and (right) Four-terminal H bar....... 26

Figure 2.8 (left) The Frenkel exction in the crystal cell. The electron and hole are bound to the same cell unit with a small Bohr diameter. (right) The Wannier exciton in the semiconductor. The Bohr diameter is much larger than the lattice constant. .................................................................................................................. 28

Figure 2.9 The band diagram of semiconductor under the light radiation. An electron is excited by one photon from the valence band to the conduction band, and a hole is left in the valence band. The electron and hole have opposite group velocities............................................................................................................. 30
Figure 2.10 The energy dispersion of exciton gas. The n=1 and n=2 state correspond to the 1s and 2p energy level, respectively. The 1s level is the exciton ground state which is less than the band gap $E_g$. ................................................................. 31

Figure 2.11 Exciton phase transition and crossover. In the dilute limit of exciton density, exciton gas exists in higher temperature as a excited state. At lower temperature, excitons start to condensate following Bose-Einstein distribution. As the exciton density increases, the exciton gas starts to dissove into the electron-hole plasma. In this case, at lower temperature, the plasma has a quantum phase transtion to BCS-like exciton condensation, which is also called exciton insulator .................................................................................................................................................. 34

Figure 2.12 The crossover from BEC of exciton to BCS like exciton condensation. ................................................................................................................................................................. 36

Figure 2.13 Order-parameter $\Delta_k$, and the pair-excitation spectrum $E(k)$, as a function of k at four different densities. This figure is adopted from (52) ....... 39

Figure 3.1 shows a typical InAs/GaSb wafer structure. 2DEG is confined in the InAs layer while 2DHG is confined in the GaSb layer. ................................................. 43

Figure 3.2 Sketch of the Corbino measurement setup. The outer conductor was grounded, and a voltage was applied in the inner conductor; we measured the current flowing across the ring. $V_f$ and $V_b$ were applied to modulate the carriers in the ring. .................................................................................................................................................. 47

Figure 3.3 Diagram for capacitance measurement in QWs................. 48

Figure 3.4 Coaxes in the 4K regime of the probe .................................... 51

Figure 3.5 4274 LCR meter ...................................................................... 51

Figure 3.6 Capacitance model for the double gate devices. ...................... 52

Figure 3.7 a, Illustration for the equilibrium density at CNP. The red dotted line is for magneto transport of unsymmetrical macro-Hallbar. In the linear part of the trace near 2V, the trace represents the electron density. The blue line is for the capacitance integration over $V_f$ which represents the absolute density. The green dashed line is for the electron density changing with the saturated rate. $n_0$ could be read from the crossing point of vertical dashed black line and green dashed line. b, Electron density increment per $V_f$ in high electron density regime for different $V_b$. ......................................................................................................................... 54
Figure 3.8 The comparison between the approximation we used and the linear approximation for high $n_0$ and low $n_0$ .............................................................. 56

Figure 4.1 Phase diagram for different front and back gate voltages. The red regime represents the inverted band and the blue regime indicates the noninverted band. The figure is adopted from Ref. (10) ...................................................... 60

Figure 4.2 Contour map of the Fermi surface of deeply inverted InAs and GaSb bands. The figure is adopted from (57) ................................................................. 63

Figure 4.3 (a) Energy spectrum of InAs(11nm)/GaSb(7nm) quantum well with an in-plane magnetic field $B = 0$T with selfconsistent calculation by eight-band Kane model (Solid line); (b), (c), (d) and (e) are the components of conduction band 2(CB2), conduction band 1(CB1), valence band 1(VB1) and valence band 2(VB2), respectively. The dash dotted line in (a) is the reduced single band model. The data is calculated by Dr. Kai Chang ......................................................... 63

Figure 4.4 Wafer structures of the strained-layer InAs/Ga0.75In0.25Sb QWs used for experiments ............................................................................................................. 64

Figure 4.5 A TEM photograph of a strained wafer; blue and red lines are guide for eyes. The data is taken by Zhongdong Han ......................................................... 65

Figure 4.6 Calculated band structure of the 9.5 nm InAs/4nm Ga0.75In0.25Sb QWs, CB1, VB1 and CB2, VB2 represent for different spin components. The band is calculated by Dr. Kai Chang ................................................................. 65

Figure 4.7 Magneto-transport data of the 30 μm ×10 μm Hall bar in a, electron-dominant regime and b, hole-dominant regime. The data is taken by Tingxin Li .............................................................................................................. 66

Figure 4.8 $B/eR_{xy}V_f$ trace of a 50×50 μm Hall bar under different back gate voltages. The top trace corresponds to 4V backgate voltage and the bottom one corresponds to 0V backgate voltage .............................................................................. 66

Figure 4.9 The CNP density vs back gate voltage traces ................................................................................................................................. 67

Figure 4.10 G-V$_f$ traces under 0 T, 4 T, 6 T, 8 T, 12 T, and 18 T in-plane magnetic field at (A) $V_b = 0$ V and (B) $V_b = 4$ V .............................................................................................................. 68
Figure 4.11 (left) the in-plane dispersion relations of electrons in InAs and holes in GaSb and (right) the in-plane dispersions under a parallel magnetic field. .........................................................................................................................69

Figure 4.12 Arrhenius plots for strained-layer InAs/InGaSb QWs (black squares), and unstrained 12.5 nm InAs/10 nm GaSb QWs (red cycles). Energy gaps can be deduced by fitting $G \propto \exp(-\Delta/2kT)$, as shown by straight dash lines in the plot. Data with black squares were taken by Tingxin Li. ...............70

Figure 4.13 Longitudinal resistance vs front gate voltage in a 1×2 μm π bar at 300 mK. ................................................................................................................................................................72

Figure 4.14 Longitudinal resistance vs front gate voltage in a 5×10 μm Hall bar at 300 mK. ................................................................................................................................................................72

Figure 4.15 Longitudinal resistance vs front gate voltage in a 50 ×100 μm Hall bar at 300 mK. ................................................................................................................................................................73

Figure 4.16 Longitudinal resistance ratioed by the predicted quantized resistance in 1×2 μm π bar, 5×10 μm, 25×50 μm and 50×100 μm Hall bars. ..................................................................................................................................................73

Figure 4.17 Temperature dependence of $R_{xx}$ at the CNP in 1×2 μm π bar. .....76

Figure 4.18 Hybridization gap energy calculated by 8-band selfconsistent calculation by Dr. Kai Chang, as a function of InAs layer thickness, InGaSb layer thickness and component. ........................................................................................................77

Figure 4.19 $R_{xx}$-$V_f$ traces measured from a 10 μm ×5 μm Hall bar device at T~20 mK with $V_b = 0$ V, 1 V, 2 V, 3 V, and 4 V. ........................................................................................................................................78

Figure 4.20 $R_{xx}$-$V_f$ traces measured from a 100 μm ×50 μm Hall bar device at T~20 mK with $V_b = 0$ V, 1 V, 2 V, 3 V, and 4 V. The edge coherence length increases with decreasing $V_b$........................................................................................................79

Figure 4.21 The coherence length as a function of the back gate voltage.......79

Figure 4.22 G-$V_f$ traces under 0 T, 1 T, 2 T, 3 T, and 4 T perpendicular magnetic field at upper panel with $V_b = 0$ V and at lower panel with $V_b = 4$ V.82

Figure 4.23 $R_{xx}$-$V_f$ traces of the 100×50 μm Hall bar under different perpendicular magnetic fields at $V_b =$0V.................................................................83
Figure 4.24 $R_{xx}V_f$ traces of the 100×50 μm Hall bar under different perpendicular magnetic fields at $V_b = 4V$. ................................................................. 83

Figure 4.25 Magnetic field dependence of coherence length under $V_b = 0V$ (circle) and $V_b = 4V$ (square). ................................................................. 84

Figure 5.1 Phase diagram for different front and back gate voltages. The red reimge represents the inverted band and the blue regime indicates the noninverted band. .............................................................................. 89

Figure 5.2 InAs/GaSb wafer structure where Si doping sheet is placed at the interface between the GaSb layer and the InAs layer. .............................................. 91

Figure 5.3 A two-dimensional topological insulator is engineered from two common semiconductors, InAs and GaSb, which hosts a robust quantum spin Hall effect. Show schematically the band structure of a InAs/GaSb bilayer, and the potential fluctuations induced by Si dopants at the interface. .............................................. 91

Figure 5.4 Magnetoresistance and Hall resistance traces measured in a Si-doped quantum wells. ...................................................................................... 92

Figure 5.5 Left panel and right panel show the back gate voltage and front gate voltage dependence of $R_{xx}$ in a 1×1μm meso hallbar, respectively. In the left panel, traces under different fixed front gate voltage are taken. ...................... 94

Figure 5.6 $R_{xx}$ vs $V_f$ traces in a 50x50μm macro hallbar for fixed back gate voltages.............................................................................................................. 94

Figure 5.7 $R_{xx}$ vs $V_b$ traces in a 50x50μm macro hallbar under zero front gate voltage.............................................................................................................. 95

Figure 5.8 The temperature-dependent conductance traces measured in a Corbino disk are displayed............................................................................. 96

Figure 5.9 The Arrhenius plot shows that the conductance vanishes exponentially with T ........................................................................................................ 97

Figure 5.10 Wide conductance plateaus quantized to $2e^2/h$ and $4e^2/h$, respectively for two device configurations shown in inset; both have length 2 μm and width 1 μm. (B) Plateau persists to 4K, and conductance increase at higher temperature. ...... 98

Figure 5.11 (left) Schematic layout of a four-terminal H bar device and (right) Nonlocal four-terminal resistance measured on the H bar device...................... 101
Figure 5.12 (A) Electrical charge transport in large devices is due to edge channels. (B) The resistance scales linearly with the edge length, indicating a phase coherence length of 4.4 \( \mu \text{m} \); the coherence length is independent of temperature between 20mK and 4K.

Figure 5.13 Longitudinal resistance ratioed by the predicted quantized resistance in 1x1\( \mu \text{m} \) junction, 1x1\( \mu \text{m} \) pi bar, 1x2\( \mu \text{m} \) pi bar, 1x2\( \mu \text{m} \), 5x10\( \mu \text{m} \), and 10x20\( \mu \text{m} \) Hall bars.

Figure 5.14 The conductance measured in Corbino disk at \( T = 300 \text{ mK} \) are shown, respectively, for magnetic field applied in the plane in upper panel, or perpendicular to the plane in lower panel. In either case, there is no evidence for gap closing at increasing magnetic field; a continuous magnetic field sweep shows that 2D bulk is always completely insulating from 0 to 8T.

Figure 5.15 The hard-gap energy is shown to increase with applied perpendicular magnetic field.

Figure 5.16 The edge helical liquid in the InAs/GaSb bilayer retains its transport characteristics in strong external magnetic fields, here examined up to 12T. Upper panel shows plateau values measured for four different devices with in-plane magnetic field applied parallel (open circles) or perpendicular (open triangles) to the edge axis. Lower panel shows the same four samples were measured (\( T = 300 \text{ mK} \)) in a field applied perpendicular to the 2D plane, with the three Hall bar devices showing increasing conductance, and the two-terminal device (blue squares) showing decreasing conductance. The device sizes are noted with “2\( \mu \text{m} \)” for 1x2\( \mu \text{m} \) pi-bar, “1\( \mu \text{m} \)” for 1x1\( \mu \text{m} \) two terminal device, “10\( \mu \text{m} \)” and “20\( \mu \text{m} \)” for 5x10\( \mu \text{m} \) and 10x20\( \mu \text{m} \) hall bar.

Figure 5.17 Shows the plateaus measured from the pi-bar device (shown in (A), red open circles) at 20mK, at different applied in-plane fields parallel to the edge.

Figure 5.18 under a perpendicular field, counter-propergating paterners in Kramers pair spatially separate, reducing the contact-coupled backscattering.

Figure 6.1 The semiconductor to semimetal transition and exciton insulator is expected to exist in the diluted semimetal.
Figure 6.2 Device layout and density measurement under double-gate control. Sketch of device layout. Holes are in GaSb layer (green) while electrons are in InAs layer (red). Front and back gate (blue) are fabricated to tune carrier densities. 113

Figure 6.3 Phase diagram of exciton with temperature and exciton density. 115

Figure 6.4 Two wafers used in works related in this chapter. The wafer in the left panel has Si doping between the interface of quantum wells. 115

Figure 6.5 CNP density $n_o$ as a function of $V_f$ and $V_b$ in units of $10^{10}/cm^2$. The density is obtained through the method introduced in Section 3.4. 116

Figure 6.6 (A) and (B) are $B/eR_{xy}$ vs $\Delta V_f$ traces of the asymmetric 50$\mu$m x 50$\mu$m Hall bar for $V_b = -6$V and 0V, respectively. Data are taken at 300mk with 1T perpendicular magnetic field. The inset in (A) is a schematic of the asymmetric Hall bar. The region in the dashed box is covered by front gate. Insets in (A) and (B) are showing band alignments corresponding to the deeply- and shallowly-inverted regime, or dense and low $n_o$, respectively. The red regime I is the electron dominating regime. The blue regime IV is the hole dominating regime. The green regime II is the electron-hole coexisting regime. The light green regime III is the soft gap. The dotted line means residual electron and hole filling in hybridization gap so there is no hard gap observed. The gold regime V is the hard gap without electrons or holes. 117

Figure 6.7 Sketch of the Corbino measurement setup. The outer conductor was grounded, and a voltage was applied in the inner conductor; we measured the current flowing across the ring. $V_f$ and $V_b$ were applied to modulate the carriers in the ring. 119

Figure 6.8 $\Delta V_f$ dependence of the conductance $\sigma_{xx}$ for device C1 from $V_b = -6$ to 0V, with a decrement of 1.5V. $n_o$ is given in units of $10^{10}/cm^2$. The red lines correspond to 0T. At $V_b = -6$V, the system forms semimetal under 35T, which is consistent with hybridization origin. As $|V_b|$ decreases, $n_o$ decreases and an insulator emerges. 120

Figure 6.9 $R_{xx}$ vs $V_f$ traces of 50x50$\mu$m macrohallbar in 300mK, as a function of fixed backgate voltages and the inplane magnetic field. The blue traces represent the transport under zero magnetic field, while the red traces show the transport under 8T inplane magnetic field. 121
Figure 6.10 Energy dispersion calculated from the 8-band self-consistent model for tunneling electrons and holes for $B_{||}=0$, 9, 18 and 35T, respectively. The data is calculated by Dr. Kai Chang. 122

Figure 6.11 Tunneling of electrons and holes in the real space. The arrow indicates electrons and holes taking part in the tunneling and the shadow regimes show the average tunneling distance. 123

Figure 6.12 $\Delta V_f$ dependence of the conductance $\sigma_{xx}$ for device C1 from $V_b = -6$ to 0V, with a decrement of 1.5V. $n_o$ is given in units of $10^{10}/cm^2$. The blue lines correspond to $B_{||}=35T$; the red lines correspond to 0T. 124

Figure 6.13 $\Delta V_f$ dependence of the conductance $\sigma_{xx}$ for device C1 in shallowly inverted band and $B_{||}$ ranging from 0T to 35T. Traces were obtained at 30mK. A region of broad zero conductance can be seen from 0 to 35 T. The color corresponds to $\sigma_{xx}$ in units of Siemens. 126

Figure 6.14 Conductance traces for different temperatures (5.8, 5, 4.5, 3.95, 3.45, 2.95, 2.6 and 2.29K) in shallowly inverted band and under 35T. Higher temperature corresponds to upper trace. 127

Figure 6.15 Dependence of dip conductance in Corbino measurement on 1/T under 35T. Solid line is guide to the eye. The data can be fit with $\sigma_{xx} \propto \exp(-\Delta/2kT)$ to obtain $\Delta$. 128

Figure 6.16 Gap energy $\Delta$ obtained for a series of $B_{||}$ (0T, 9T, 18T, 27T and 35T) in the low density case. 128

Figure 6.17 Magneto-conductance of Corbino device C1 under magnetic field. The total magnetic field is 35T. By rotating the device, the device plain has a deviation angle with magnetic field so that there is a small perpendicular magnetic field ($0T \pm 2T$). The blue solid lines are for positive perpendicular magnetic field while the red dotted lines are for negative perpendicular magnetic field. 129

Figure 6.18 the dependence of the conductance minimum on 1/T. 130

Figure 6.19 (left panel) shows dependence of the conductance minimum on 1/T for different $V_b$. Here the $\sigma_{xx}$ is normalized by $\sigma_{xx} = \sigma_{xx \ min}/(\sigma_{xx \ min} at \sim 2.5K)$. Solid lines are guides to the eye. Black circles are for $V_b=0V$, red squares are for $V_b=-1.5V$, green down-triangles are for $V_b=-3V$, blue up-triangles are for $V_b=-4.5V$, and yellow stars are for $V_b=-6V$. (right panel) shows the activated energy vs $n_o$. 132
Figure 6.20 a, CV curves with $V_b=0\text{V}(\text{red line})$ and $-6\text{V}(\text{blue line})$. At 300mK, a low frequency (100Hz) ac voltage is delivered to the frontgate with QWs grounded, and the capacitance between frontgate and QWs can be measured. b shows CV curves with $V_b=0\text{V}$ from wafer A(red line) and wafer B(blue line), respectively. In both cases, large capacitance drops exist. c, CV curves under different temperatures(0.3K, 2K, 4K, 6K and 10K). Curves are taken in C1 with $V_b=0$. The observed insulating gap disappears for the temperature up to 10K.

Figure 6.21 (a) and (b) show nonlocal measurement performed in meso-H bar from wafer A under 0T and 35T, respectively. The dotted lines indicate the expected resistance value from Landauer-Büttiker formula. (c) and (d) show nonlocal measurement performed in meso-H bar from wafer B under 0T and 35T, respectively. In the gap, the current path is shown in the inset of (a) as red and green arrows.

Figure 6.22 Gap function $\Delta(k)$ (red line) and the pair-excitation energy $E(k)$ (green line) of the exciton droplet as a function of $k$. The blue cone means the part could be accessible to THz radiation. The grey part indicates the exciton gap regime. The data is caculated by Dr. Kai Chang.

Figure 6.23 The home made Aluminium cover with black polypropylene window.

Figure 6.24 Transmission spectra at the gap with ratios of transmitted intensity to a reference spectrum (low mobility and low density electron regime) (normalized by ratioing to zero field spectra in the electron regime at the same temperature).

Figure 6.25 Transmission spectra at the gap under different temperatures.

Figure 6.26 Transmission spectra at the gap under magnetic fields at 1.4K(left panel) and 20K(right panel).

Figure 6.27 Lanlau level fun chart of a typical inverted QW structure. The figure is adopted from (7).

Figure 6.28 The schematics of the contact-buried Corbino and Hall bar devices.

Figure 6.29 A phenomenological phase diagram is shown for distinct topological states under perpendicular fields up to 35T.
Figure 6.30 Bottom panels display experimental traces of longitudinal conductance $\sigma_{xx}$ and Hall conductance $\sigma_{xy}$ measured at a fixed field of 8T, 16T, 35T, respectively, as a function of gate voltage, where wide ranges of zero-conductance can be seen punctuated by $\sigma_{xx}$ peaks. Here from right to left, the Fermi energy $E_F$ is swept from electron- towards mobility gap and into hole- regimes. The peak below the $v=1$ QH plateau marks the boundary into QSH state. Remarkably, inspecting traces from zero to 35T in top panel, we found that the bulk gap for QSH increases with perpendicular fields, instead of closing. The inset depicts a model of "canned helical state" which is deformed adabatically from the zero-filed HL. Note that the in-plane spin components remain spin-momentum locked .......................................................... 150

Figure 6.31 Hall resistance $R_{xy}$ measured under fixed fields up to 18T in device B................................................................................................................................. 154

Figure 6.32 Hall resistance $R_{xy}$ at CNP extracted from Figure 6.31 ............... 155

Figure 6.33 Longitudinal resistance $R_{xx}$ measured under fixed fields up to 18T in device B................................................................................................................................. 155

Figure 6.34 Longitudinal resistance $R_{xx}$ at CNP extracted from Figure 6.33 156

Figure 7.1 Schematic diagram of positive and negative drags. Arrows show the direction of electron flow. In the negative drag, the direction of electron flow in the drag circuit is opposite to that in the drive circuit, while in the positive drag, the direction of electron flow in the drag circuit is the same as that in the drive circuit................................................................................................................................. 160

Figure 7.2 Schematics of the fabrication process ............................................. 162

Figure 7.3 Schematics of the fabrication process ............................................. 162

Figure 7.4 Diagram of the Coulomb drag circuit (left). The current goes upward in the drive circuit following with the solid blue line. The dashed line indicates the other helical edge which does not hold the current. (right) The scanning electron microscopy of the slit regime. The white mark means 100nm. ........................................................................................................................................................................... 165

Figure 7.5 Longitudinal resistance vs the front gate voltage in a 1x2$\mu$m $\pi$ bar. Inset shows the temperature dependence of the longitudinal resistance peak. ........................................................................................................................................................................... 166
Figure 7.6 The frontgate voltage dependence of the bulk conductance in a Corbino device. .................................................................................................................. 167

Figure 7.7 The leaking current vs bias voltage in the Coulomb drag device.... 168

Figure 7.8 $R_d$ as a function of front gate voltage in the Coulomb drag device in 300mK.......................................................................................................................... 170

Figure 7.9 $R_d$ as a function of front gate voltage in the Coulomb drag device in 1K.......................................................................................................................... 171

Figure 7.10 $R_d$ as a function of the front gate voltage in the Coulomb drag device in 1.5K. .................................................................................................................. 172

Figure 7.11 $R_d$ as a function of the front gate voltage in the Coulomb drag device in 2K. .................................................................................................................. 173

Figure 7.12 $R_d$ as a function of the front gate voltage in the Coulomb drag device for negative driving current................................................................. 173

Figure 7.13 Temperature dependence of the Coulomb drag signal at the peak of the negative drag regime................................................................. 174

Figure 7.14 Temperature dependence of the Coulomb drag signal at the peak of the negative drag regime in another device........................................... 175
Chapter 1

Introduction

1.1. Topological system

The observation and classification of various quantum phases and phase transition is a key point in condensed matter physics. In 1980s, the observation of the Integer Quantum Hall (IQH) effect by Klitzing[1] in a Silicon metal–oxide–semiconductor-field-effect transistor (MOSFET) and the Fractional Quantum Hall (FQH) effect by Tsui, Stormer, and Gossard(2) in GaAs/AlGaAs quantum well (QW) opened the era of topological physics. The Quantum Hall(QH) effect with
quantum conductance plateaus in the precision of one part in billion has been universally observed in various systems with distinct material details that could not be classified into the previous paradigm but introduces the notion of topological order. The QH state is topologically distinct from all previous known states of matter, giving the first example of a topological system. In the QH effect, the bulk of the two dimensional (2D) system is insulating, and there are the edge modes that carry currents at the boundary of the sample. The current flows in one direction in the absence of dissipation, giving precise quantized hall conductance which only jumps in a series of integer values in units of quantum conductance, i.e. $e^2/h$. To describe this new quantum phase, the concept of topological invariance was introduced(3), which ignored detailed differences in materials but focused on fundamental properties. In the IQH effect, the quantized plateaus of Hall conductance could be treated as the topological invariant, also called the Chern invariant, which keep constant if the system only smoothly varies in material parameters (i.e. carrier density and QW structure). In the FQH effect, the topological order and the many-body interaction between electrons must be considered. The FQH effect is the first example of a topological nontrivial system originating from a many-body interaction. The formation of the FQH state could be understood in terms of composite fermions, where one electron and an even number of magnetic flux quanta form a new quasiparticle.

Either IQH or FQH states belong to a topological class where time reversal symmetry (TRS) is explicitly broken due to the presence of a magnetic field.
Recently, a topologically nontrivial quantum state distinct from QH state was proposed\(^{(4,5)}\), called the TRS topological insulator (TI), with TRS preserved in 2D and 3D electronic insulators, in the presence of strong spin orbit coupling. The first example of a 2D TRS TI or Quantum spin Hall (QSH) insulator is the HgTe QW, predicted by Bernevig et al\(^{(6)}\) and observed by König et al\(^{(7)}\). In this system, the bulk is insulating due to spin-orbit coupling, and on the boundary of the sample there exists 1D Dirac type helical edge state. Similar to a normal insulator, the TI has an energy gap in the bulk between the conduction band and the valence band, which contributes to the zero bulk conductance. At the interface between two insulators with opposite symmetries of conduction and valence bands, the linear dispersion relation of the gapless interface state emerges\(^{(8,9)}\), leading to a Dirac type 2D surface or 1D edge. A chiral edge in QH state propagates in one direction, protected by the Chern invariant; as a contrast, a helical edge in QSH state possesses two counterpropagating edges with opposite spins, protected by \(Z_2\) invariant. After the prediction and realization of QSH effect, splendid advances were achieved with series of 2D\(^{(7,10–13)}\) and 3D TIs\(^{(14–19,19–25)}\) proposed to exist. In inverted InAs/GaSb QWs, the experimental studies of which are the subject of this thesis, QSH effect was predicted\(^{(10)}\) and investigated\(^{(12)}\) by Knez et al in the hybridization\(^{(26)}\) bulk gap with sizable bulk conductance. However, either in HgTe QW or InAs/GaSb QWs, the evidence for the quantized plateau of helical edge state is missing, which would be the critical signature to the existence of the QSH effect. Moreover, QSH effect in an insulating hybridization gap of InAs/GaSb QWs has not
been reported. The bulk conductance mixes with the edge conduction, making the existence of helical edge in this system inconclusive. Both thrusts initiate the study of topological phases in InAs/GaSb QWs in this thesis. Especially, precisely quantized conductance plateau was observed\(^{(27)}\) under the circumstance of broken TRS which could not be explained by the existing single particle topological theory\(^{(4,5)}\) but was ascribed to the existence of exciton insulator\(^{(28)}\). The remainder of the thesis focuses on experimental studies, particularly transport and optics measurements of the bulk exciton insulator in this broken gap InAs/GaSb semiconducting material.

### 1.2. Inverted InAs/GaSb heterostructure

The InAs/GaSb QWs are composed of InAs, GaSb and AlSb, which belong to a class of lattice matched components, named the 6.1Å family\(^{(29)}\). Members of the 6.1 Å family have an approximate lattice constant of nearly 6.1Å. AlSb serves as a QW barrier to InAs and GaSb QWs. The broken gap type II band alignment between InAs and GaSb, with the conduction band of bulk InAs 150meV lower than the valence band of GaSb, enables the coexistence of spatially separated 2D electrons in InAs QW and 2D holes in GaSb QW confined by AlSb barriers, if the InAs and GaSb QW is wide enough as proposed by Esaki\(^{(30)}\).

Through tuning the width of InAs and GaSb QWs, the energy between the bottom of conduction band and the top of valence band can be continuously tuned\(^{(31)}\). Additionally, with the front and back gates, the band structure and the
Fermi energy can also be modulated, meaning the transition from noninverted band (normal insulator) to inverted band could be realized. The inverted band was originally thought to be a semimetal state that was unstable against exciton ground state near the semiconductor to semimetal transition. In 1995, Cheng(32) and Kono(33) observed magneto exciton in the presence of perpendicular magnetic field, in this double layer system.

On the other hand, Altarelli(34) pointed out that hybridization from tunneling between InAs and GaSb QW would dramatically suppress the formation of exciton. In the inverted case, as the inplane momentums and energies of electron and hole match, carriers could transfer between two QWs and a hybridization gap opens at the cross point in the otherwise semimetal band structure due to the hybridization of bands. In 1997, Yang(26) confirmed the existence of hybridization gap in the deeply inverted InAs/GaSb QWs through inplane magnetic field experiments. Moreover a smooth connection between band inside and outside of the sample is required on the boundary of the sample, which results into a gapless edge mode with linear dispersion. In this thesis, I will show the hybridization gap and exciton gap are the competing gap mechanisms, dependent on the band inversion.

1.3. **Exciton insulator**

In semiconductors, an incoming photon may excite an electron in the valence band to the conduction band, leaving a hole in the valence band. Under the Coulomb attraction, the negatively charged electron in the conduction band and positive
charged hole in the valence band are energetically favorable to form bound pairs and create quasiparticles, named as excitons, with a hydrogen like spectrum. The exciton was predicted to undergo several phase transitions dependent on the exciton density and temperature. As the temperature is sufficiently low, the exciton experiences a phase transition from exciton gas to Bose Einstein Condensation (BEC) of exciton (35–37). In this case, the excitons are separated, with their wave functions not overlapping, making the individual exciton like a bosonic particle composed of an electron and a hole. Then if the exciton density increases, as originally studied by Keldysh and Kozlov (37) with mean field treatment, the exciton wavefunctions grow bigger and bigger until they start to overlap. Then electron-hole Coulomb attraction is screened and thus effectively becomes weaker. As a result, the wavefunction between electrons and holes becomes less bound together such that it becomes less favorable for particular electrons to pair with particular holes. Such loosely correlated excitonic states opens a gap near the Fermi surface just like BCS state of Cooper pair, which is call exciton insulator or BCS like excitonic condensation. The exciton insulator is well-defined by BCS like exciton state. To explore the presence of exciton insulator, Mott and others (38,39) proposed that it could exist near the semimetal-semiconductor transition with the natural existence of electrons and holes which spontaneously leads to an excitonic instability without photoexcitation.

In inverted InAs/GaSb heterostructure, if the hybridization could be neglected, it naturally processes spatial separated electrons and holes without the photon excitation, which are proposed to form exciton ground state (35)
spontaneously with a BCS like gap opening. However, until now, relevant experiments are still lacking. The main difficulty in searching excitonic insulator is the low charge neutral point (CNP) density, where electron and hole are bound to form exciton instead of e-h plasma, and the interlayer tunneling which forms hybridization and mixes bands thereby reducing binding energy. In this thesis, I will show these two conditions can be satisfied simultaneously in a shallowly inverted band, which provides a natural platform to explore exciton insulator.

1.4. About this thesis

Works presented in this thesis are experiment studies of distinctive topological quantum phases in inverted InAs/GaSb systems. In the light of recent experiments and theoretical proposals in this system regarding topological insulating phase, the insulating hybridization gap was realized with TRS QSH effect observed, for the first time, in the InAs/GaSb system, through the strain technology. The helical edge had a coherence length more than 10μm with the largest bulk gap in known QSH systems. This result was consistent with the theoretical prediction of the QSH insulator in this system and agreed with initial experiments in the HgTe QW. Nevertheless, I found the hybridization gap and associated TRS QSH insulator dominated in the deeply inverted regime, and as the system approached to the shallowly inverted regime, precisely quantized conductance plateau of the QSH effect was observed for the first time. This was a novel topological phase, which had a helical edge at the boundary of the sample with 4.4μm coherence length.
Surprisingly, the edge mode persisted with weak dependence of the temperature and in spite of the inplane magnetic field. On the other hand, under the perpendicular magnetic field, the inner of the helical edges shrank while the outer one expanded. It indicated this topological effect was not protected by TRS, which was the first observation of the TRS broken QSH insulator. Theoretically, it was proposed that the exciton insulator would be responsible for this novel phase. Then I utilized double-gate driven transition from deeply inverted regime to shallowly inverted regime. With the CNP density tuned from the high density to the low density, I observed the continuous modulation of the bulk state from the hybridization gap to a hard gap which was then confirmed as the exciton insulator gap. The exciton insulator gap was insensitive to the inplane magnetic field and represented an incompressible property, which not only existed together with the hybridization gap but also was strengthened as the hybridization gap became weaker. Under the Terahertz radiation on the 2D bulk hard gap, we observed, for the first time, the particular absorption spectrum of the excitonic insulator, which elucidated the hard gap originated from the exciton insulator. Then I studied the helical edge state associated with the exciton insulator under the high magnetic field and observed the edge state exhibited canned helical behavior under 35T perpendicular magnetic field. Finally, I developed a new fabrication method of 1D Coulomb drag circuit, allowing us to probe fundamental properties of edge channels with additional insights regarding this novel quantum state.
2.1. 6.1 Å family of III-V semiconductor

In III-V family materials, InAs, GaSb and AlSb attract great interests due to an approximately matched lattice around 6.1 Å as shown in Figure 2.1. The energy gap in 6.1 Å family ranges from 0.36eV in InAs and 0.78eV in GaSb to 1.61eV in AlSb. Due to the similar lattice, these materials construct the heterostructure, confining electrons and holes. Normally, with AlSb working as a barrier, the heterostructure combined with a thin layer of InAs or GaSb provides great wells.
depth of 1.35eV for electrons and of 0.42eV for holes. The effective mass of InAs is nearly 0.03 in unit of free electron mass and the effective mass of GaSb is about 0.37 in unit of free electron mass, giving high carrier density in QWs up to $10^{13} cm^{-2}$.

![Figure 2.1 Band lineups in the 6.1Å family material and the lattice constants of each material. The shaded parts show the energy gaps with all energies in eV. The figure is adopted from Ref. (29).](image)

Of the particular interest is the broken-gap band lineup between InAs and GaSb, which was found in 1977 by Sakaki(30). The conduction band bottom of InAs is lower than the valence band top of GaSb with 0.15eV, triggering most of interest in 6.1 Å family as shown in Figure 2.1. Combined with the fact that the valence band top of AlSb is lower than that of GaSb with 0.41eV, the heterostructure of...
AlSb/InAs/GaSb/AlSb allows for the formation of 2D electron in InAs by the confinement from wells of AlSb and GaSb, and the existence of 2D holes in GaSb by the confinement from wells of AlSb and InAs. This band structure was called type II heterostructure. Beyond the use of binary AlSb, ternary alloy of AlGaSb is used with 20% of Ga to reduce the oxidization of Al and enhances the chemical stability of AlSb, with an inconsequential decrease in the height of electron barrier.

Figure 2.2 a) shows the band diagram of inverted InAs/GaSb QWs. The electrons are confined in InAs layer and the holes are confined in GaSb layer. The red dashed line indicates the top of hole band while the blue dashed line indicates the bottom of electron band. b) presents the energy dispersion of electron band E and hole band H. The dashed line shows the crossing of the uncoupled E and H. The anti-crossing opens due to the band hybridization. The front gate and back gate shown in a) can modulate the band separation and the Fermi energy.

The heterostructure of AlGaSb/InAs/GaSb/AlGaSb is the one I focused in this thesis. For a given width of QW as shown in Figure 2.2, energy levels in QWs are discrete with several electron and hole subbands. Here we only study the first
electron and heavy-hole subbands, with higher subbands neglected. If we simply assume the QW barrier is infinite, the energy separation between the bottom of the subband and that of QW is inversely proportional to the square of QW width. So if the QW is narrower, the bottom of the subband is lifted. In this structure, according to the widths of QWs, there are two fundamentally distinct regimes as indicated in Figure 2.3. If the widths of InAs and GaSb QWs are larger than critical values, the bottom of the conduction band is higher than the top of the valence band, which is called the inverted band. As the widths of InAs and GaSb QW are smaller than critical values, the bottom of the conductance band is higher than the top of the valence band, which is referred to the noninverted band.

![Figure 2.3 Schematic of inverted and noninverted band in InAs/GaSb QWs.](image)

In the inverted regime, 2D electron gas is confined in InAs QW while 2D hole gas is confined in GaSb QW. The carriers can be introduced or changed by mainly three ways(29): 1. Surface state; 2. Conventional shallow bulk donor; 3. Electrostatic
gate. In the deep InAs QW, the surface state on the top AlGaSb barrier is easy to become an important doping source for electrons in non-intentional doped QWs. The doping effect is dependent on the surface coverage caps with different surface state energies. There are mainly two types of coverage caps: a 3nm InAs layer and a 3nm GaSb layer. It should be mentioned these caps also play an important role in the protection of AlGaSb against the oxidation. For the thin GaSb cap, its surface state has the energy of about 0.5eV above the bottom of InAs band under the flat band approximation. In this case, electrons are transferred from the surface state to the QW until there is an energy balance between the Fermi level and the surface state energy. For a thinner top barrier, the electrons are easier to transfer with the formation of higher equilibrium electron concentration. For a 250nm AlGaSb top barrier, the electron concentration is nearly $2.5 \times 10^{11} \text{cm}^{-2}$. For a 50nm AlGaSb top barrier, the electron concentration is nearly $5 \times 10^{11} \text{cm}^{-2}$. For a 25nm AlGaSb top barrier, the electron concentration is nearly $1.5 \times 10^{12} \text{cm}^{-2}$. For the InAs cap, under the flat band condition, its surface energy is only 150meV higher than the InAs conduction band, giving the lowest electron density induced by the surface state. Besides InAs and GaSb cap, AlSb and Be-doped GaSb caps can also be applied. With multiple caps and the selective removal of the specific cap, not only the initial electron density can be modulated, but also the nano-lateral structure could be realized.

For the conventional shallow bulk donor, the high barrier allows the high electron density to be achieved by modulation doping. In modulation doping, the
donors placed within the barrier, drain electrons to the QW, dependent on the spatial separation between donors and the QW. The ionized impurities are separated from the QW, so the impurity scattering is reduced, with higher mobility. Usually, two doping recipes are used in this system, with Te as a donor and Be as an acceptor. Compared with the surface states, the bulk donors can be utilized to slightly modulate the initial density in QWs. Besides two sources above, the electrostatic gate could be applied to intentionally dope the QWs, which will be discussed in the next section.

### 2.2. Inverted regime

In the inverted regime, electrons in InAs QW exhibit a parabolic dispersion while holes in GaSb QW have a parabolic dispersion directing oppositely, with a band overlap. As the Fermi level is lower than the conduction band of GaSb QW but higher than the valence band of InAs QW, electrons and holes coexist and are spatially separated. As a result, they form an indirect band in the real space. If we do not consider any coupling between electrons and holes, this band structure is a standard semi-metal. Through tuning the chemical component in the barrier and the width of QWs, ideally we can change the band inversion of this semi-metal from “deeply inverted” to “shallowly inverted”, and further vary the system from inverted band to noninverted normal semiconductor band. Also through the electric field perpendicular to the plane in a double gate structure, this system gives us the opportunity to realize the in-situ modulation of the band structure and the tuning of
Fermi level, originally proposed by Naveh and Laikhtman (31). In their proposal, an electric field across the double QWs, changes the energy profile of the wall and effectively shifts the valence band edge of GaSb QW and the conduction band edge of InAs QW in opposite directions, leading to the band bending. As an instance, if we apply a positive electric field perpendicular to the plane, the conduction band shifts down and the valence band shifts up so that the overlap between them increases. Further in their calculations, with a two-band effective band-orbital model and the real band parameter, the self-consistent numerical solution of Poisson and Schrödinger equations manifests that the realization of this operation is experimentally possible with suitable parameters. Figure 2.4 shows the inplane band structure of 17nm InAs and 5nm GaSb QWs under various external electric fields. In the calculation, if the electric field is small, it could be approximately treated as a perturbation and proportional to the change of band overlap at \( k=0 \), giving a quantitative treatment for the small electric field. Thus, with the electric field, the conduction and valence bands can be modulated continuously.
Figure 2.4 Energy band dispersion of 17nm InAs and 5nm GaSb QWs under various perpendicular electric fields. The figure is adopted from Ref. (31).

Besides tuning the band inversion, electrostatic gates also allow us to change the Fermi level that corresponds to the carrier density. This electrostatic gate effect could be understood in a standard capacitance model. The density difference per voltage is determined by total capacitance that is dependent on the geometry capacitance and quantum capacitance. Since the density of state(DOS) of electrons or holes are extremely high, the quantum capacitance could be neglected in the total capacitance and the carrier density changes with the applied bias linearly. Due to the doping of the surface state, the Fermi level of QWs initially is pinned to a higher position than the top of the valence band. In this case, we need to use the gates to
lower the Fermi level towards the band overlap regime, where a series of phenomena including the QSH effect and exciton insulator are expected. To realize both the modulation of the band inversion and the tuning of Fermi level position, two gates are necessary to satisfy the required degrees of freedoms in the energy spectrum.

2.3. Hybridization

In the inverted regime, the electron and hole bands cross at the certain momentum value. This structure was initially treated as a semimetal\((30)\) in the spatially separated electrons and holes with instability against the exciton insulator. However, in 1983, Altarelli\((34)\) pointed out that the electron wavefunction would extend into GaSb QW and the overlap between electron and hole wavefunctions resulted into quantum mechanical coupling between electrons and holes. In other words, as the electron and hole band crossed at a finite momentum value where the carriers in two QWs had equal momentum and energy, electrons/holes tunneled forward and backward between two QWs in the real space, such that this degeneracy was lifted and a hybridization gap on the order of \(4\text{meV}\)\((12,26)\) opened at the crossing point. The tunneling could only happen between the states with the same symmetry. Here the electron state with \(1/2\) angular momentum, only couples to the heave-hole state via the mixed light-hole component with a \(1/2\) angular momentum, at the finite momentum away from the center of the band. Hence, the conservation of the angular momentum requires electrons to couple with holes
including light hole component. The light hole component is mixed in the heavy hole due to an inversion asymmetry term which is linear to the momentum. On the other hand, the electron state comes from s-orbitals while the hole state originates from spin-orbit coupled p-orbitals with $p_x + ip_y$ rotational symmetry perpendicular to the wells plane. These states have opposite parities and the parity selection rule requires the coupling of these states through an odd operator under the space inversion. Within the $kp$ approach, the tunneling coupling can be described as $w(k_x + ik_y)$ \((34,41)\) where $w$ is a constant. Hence as the band inversion becomes smaller, the momentum of the crossing point approaches to zero, and the coupling term is close to zero\((10,41,42)\), meaning smaller hybridization gap for less inverted band structure.

Near the crossing point, the band is gapped due to the existence of hybridization and a conductivity minimum can be expected. When the inplane momentum is far away from the crossing point, the tunneling has little effect to the electron and hole bands, corresponding to those of two uncoupled QWs. The conductivity could be described by the Drude formula. In the experiment\((26)\), within hybridization gap a conductance minimal was observed but the capacitance measurement represented a capacitance drop of less than 1\%, indicating the DOS in the gap was comparable with that outside of the gap, i.e. electron or hole. Further the DOS in the gap was estimated as the sum of that of electrons and holes, showing that the gap was soft. In the thesis, I will revisit this hybridization gap and use strain-layered InAs/InGaSb QW to realize the insulating bulk gap.
2.4. Quantum Spin Hall effect

Before introducing the QSH effect, we will review the first topological phenomenon in condensed matter physics: the QH effect. In the QH effect, under perpendicular magnetic field, a series of Landau levels (LL) form in the bulk with the energy gap between two adjacent LLs. As the Fermi level is placed in the middle of gap, all states inside the bulk are localized and cannot contribute to the conductivity. Due to the boundary confinement, LLs bend up at the boundary of the sample and the Fermi level crosses the bent LLs with chiral edges generated. The number of edges at the Fermi level is just that of the occupied LLs, giving the filling factor \( \nu \) as an integer. In this case, although the bulk is insulating, the boundary carries the edge current that contributes to a series of plateaus of Hall resistance. Since the edge current flows unidirectionally, backscattering could not happen between edges. In this case, there is a zero longitudinal resistance. For the case of \( \nu=1 \), there is only one chiral edge state at the boundary of the sample. If one put two copies of this system together, on one layer there exists spin up electrons with a chiral edge state and the other layer has the spin down electrons with anti-chiral edge state.\(^{(43)}\) This combined system could be analogues to the QSH insulator. It should be mentioned in this case the magnetic field is still present, making it distinct from the real QSH system.

The QSH effect was proposed in the graphene by Kane and Mele\(^{(44)}\), but due to small spin-orbit interaction this proposal could not be realized in experiments. Then Bernevig, Hughes, and Zhang searched this effect in semiconductors and
predicted(6) the QSH effect in HgTe QW which then was observed by Würzburg group(7). The QSH phase is presented by an energy gap in the bulk and topologically protected helical edge states at the boundary, under the protection of the TRS. The helical edge state carries two single mode 1D edge currents that counter propagate with opposite spins. In HgTe QW, the front gate has the same effect with the back gate that turns the Fermi level, thus the band inversion can only be changed by the width of HgTe QW. As the QW thickness is larger than the critical value, the band is in the inverted regime and the spin orbit coupling opens a topological bulk charge-excitation gap. In the bulk, the energy gap between the valence band and the conduction band is always negative, while the energy gap becomes positive outside of the bulk, such as vacuum or gate dielectric insulator. To ensure a smooth band connection between energy state outside and inside the sample, the gap must close at the boundary that introduces gapless edge state. As the Fermi level is placed at the CNP, i.e. the bulk gap, the Fermi level unavoidably crosses the energy levels at the boundary, which is similar with the generation of chiral edge state in the QH effect. The gapless edge state could also be understood by Einstein’s equivalence principle. The energy difference between the conduction and valence bands can be treated as the Einstein mass. In this case, the mass is negative inside the topological bulk due to the band inversion, and positive outside the bulk. The smooth connection of this mass between positive and negative ones determines that the mass at the boundary is zero. According to the relativity theory, a zero mass means the energy is linear with the momentum, giving a Dirac fermion with the linear
dispersion. Different from the graphene hosting 2D Dirac fermion (45), here the Dirac fermion is 1D with the single mode.

This QSH effect protected by the TRS in principle could be described by the Hamiltonian as functions of the momentum, spin and position operators. Generally, this Hamiltonian includes the kinetic energy which is proportional to $p^2$, spin orbit interaction which is proportional to $r \times p$, and a term to describe potential that should only relate with $r$. Under TRS, the momentum and spin operator change sign while the position operator keeps the same, so the Hamiltonian is invariant under the TRS. However, if we apply a magnetic field, $p$ becomes $p + eA$ where $A$ is the vector potential and does not change the sign under time reversal, such that the Hamiltonian is not time reversal invariant. For a TRS protected system, Kramer’s theorem requires each energy band to be accompanied by its Kramer’s partners at the time reversal invariant points which is the middle of the Brillouin zone ($k=0$), and hence have a pair of edge modes. To preserve the TRS, one edge mode should have opposite momentums and spins. Then the helical edge state could be generated.

In inverted InAs/GaSb QWs, when the Fermi level is located at the CNP, electrons tunneling between QWs open a bulk hybridization gap and the QSH effect would host here, as calculated by Liu (10) in a 8 band kp model. It should be mentioned that the real system has additional complexities which simple argument did not include. In InAs/GaSb QWs with the zinc-blende crystal structure, there is not inversion symmetry along QW growth direction, meaning the bulk and structure inversion symmetry is not preserved and competes with the hybridization.
Nevertheless, the numerical calculation shows these effects modify the phase transition but could not destroy the QSH phase. As shown in Figure 2.5 with 8 band self-consistent calculation, 10nm InAs/10nm GaSb QWs show gapless helical edge states while 8.1nm InAs/10nm GaSb QWs do not host edge states.

![Figure 2.5 Band dispersion based on eight band self consistent calculation(10).](image)

(a) NI regime (b) QSH regime

Figure 2.5 Band dispersion based on eight band self consistent calculation(10). a) shows the trivial noninverted band without helical edge states, where the InAs layer is 8.1nm and GaSb layer is 10nm. b) shows the nontrivial inverted band with helical edge states, where the InAs layer is 10nm and GaSb layer is 10nm. Figures are adopted from (10).

2.5. Landauer-Büttiker formula

1D helical edge state is able to be detected in transport measurement. For the mesoscopic quantum transport, the ballistic transport of edge state is described by Landauer-Büttiker (LB) formula that defines electron transport in multi-terminal devices. In case that LB formula is valid, there are two constraints. One is that quantum coherence should be kept between neighboring contacts, which means
single wave function is defined from one contact to the neighboring one. The other is that there is not current flowing into the voltage probe. If these constraints are satisfied, the current and the probe voltages can be described\(^{(46)}\) by

\[ I_{qp} = \frac{e^2}{h} \sum_q \left[ T_{qp} V_p - T_{pq} V_q \right] \] (1)

where \( q \) and \( p \) are contact labels, and \( T_{qp} \) is the transmission probability from contact \( q \) to \( p \). In the QH effect which is successfully defined by LB formula, the transmission probability \( T_{qp} = 1 \) if \( p = q + 1 \). For a six-terminal Hall bar device as shown in Figure 2.6, as the current goes from terminal 1 to terminal 4 and the voltage probes between other terminals, we solve the formula of (1) and obtain,

\[
\begin{pmatrix}
I_1 \\
I_2 \\
I_3 \\
I_4 \\
I_5 \\
I_6
\end{pmatrix} = \frac{e^2}{h} \begin{pmatrix}
1 & -1 & 0 & 0 & 0 & 0 \\
0 & 1 & -1 & 0 & 0 & 0 \\
0 & 0 & 1 & -1 & 0 & 0 \\
0 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 \\
-1 & 0 & 0 & 0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
V_1 \\
V_2 \\
V_3 \\
V_4 \\
V_5 \\
V_6
\end{pmatrix}
\]

Since the current goes into terminal 1 and gets out from terminal 6, we get \( I_1 = -I_6 = I \) and \( V_6 = 0 \). Terminal 2, 3, 4, and 5 are the voltage probe contacts, so

\( I_2 = I_3 = I_4 = I_5 = 0 \), and we have

\[ R_{23,16} = R_{45,16} = 0, R_{24,16} = R_{35,16} = \frac{h}{e^2}, \]

where \( R_{ij, nm} = (V_i - V_l)/I \). Now let us see the case of the QSH effect. As discussed, we could treat the QSH state as two opposite \( v = 1 \) QH states. In other words, it could
be seen as one layer with $v=1$ QH state of 2D electron gas and the neighboring layer with $v=1$ QH state of 2D hole gas, with the contacts connecting both layers. In this case, according to LB formula, the relationship between the terminal currents and voltages is: the transmission probability is 1 for neighboring contacts, i.e. $T_{pq} = 1$ if $q$ and $p$ are neighbors, which means $T_{pq} = 1$ if $p = q + 1$ and $T_{pq} = -1$ if $p = q - 1$ with other transmission probability zero. Thus, with this new rule, for a six terminal Hall bar with the current going from terminal 1 to terminal 6, the transmission matrix $T_{pq}$ is

$$
T_{pq} = \begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 0 & 0 & 1 & 0
\end{pmatrix}
$$

which gives the linear equations as

$$
\begin{pmatrix}
I_1 \\
I_2 \\
I_3 \\
I_4 \\
I_5 \\
I_6
\end{pmatrix}
= \frac{e^2}{\hbar}
\begin{pmatrix}
2 & -1 & 0 & 0 & 0 & -1 \\
-1 & 2 & -1 & 0 & 0 & 0 \\
0 & -1 & 2 & -1 & 0 & 0 \\
0 & 0 & -1 & 2 & -1 & 0 \\
0 & 0 & 0 & -1 & 2 & -1 \\
-1 & 0 & 0 & 0 & -1 & 2
\end{pmatrix}
\begin{pmatrix}
V_1 \\
V_2 \\
V_3 \\
V_4 \\
V_5 \\
V_6
\end{pmatrix}
$$

Similarly, we have $I_2 = I_3 = I_4 = I_5 = 0$, $I_1 = -I_6 = I$ and $V_6 = 0V$ with contacts 2, 3, 5 and 4 voltage contacts, so the formula gives $R_{23,16} = R_{45,16} = \frac{h}{2e^2}$. On the other
hand, for a four-terminal device in Figure 2.7, such as π bar and H bar, when the sample is in the QSH regime, the transmission matrix $T_{pq}$ is

$$
T_{pq} = \begin{pmatrix}
0 & 1 & 0 & 1 \\
1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 \\
1 & 0 & 1 & 0 \\
\end{pmatrix},
$$

and the linear equations become

$$
\begin{pmatrix}
I_1 \\
I_2 \\
I_3 \\
I_4 \\
\end{pmatrix} = \frac{e^2}{h} \begin{pmatrix}
2 & -1 & 0 & -1 \\
-1 & 2 & -1 & 0 \\
0 & -1 & 2 & -1 \\
-1 & 0 & -1 & 2 \\
\end{pmatrix} \begin{pmatrix}
V_1 \\
V_2 \\
V_3 \\
V_4 \\
\end{pmatrix}.
$$

Figure 2.6 Six-terminal Hall bar device with QSH edge states
If we pass the current from contacts 1 to 4 and measure the voltage drop between contacts 2 and 3, we have $V_4 = 0$, $I_1 = -I_4$, and $I_2 = I_3 = 0$. Thus we obtain $V_2 - V_3 = I_1 \frac{h}{4e^2}$ and $G_{14,23} = \frac{4e^2}{h}$. In this case, for the edge transport, the edge channel exists at the boundary of the sample and the current goes along the edge channel from one contact to neighboring one, for both $\pi$-bar and H bar devices, as shown in Figure 2.7.

Especially, the measurement on H bar device represents the nonlocal characteristic of helical edge modes. In a $\pi$ bar device, as the bulk is conductive, the current going from contact 1 and 4 pass through the contacts 2 and 3, so the voltage signal $V_{14,23}$ in $\pi$ bar device gives the resistance of nearly several k$\Omega$. If this resist is close to the quantized value, we could not distinguish it from the quantized helical edge resistance. On a contrary, in H bar device, the current going from contacts 1 and 4 is not supposed to locally pass through the contacts 2 and 3, so the voltage signal $V_{14,23}$ approaches to zero (but from Poisson's equation, one could expect a signal of less than 100$\Omega$). As the device goes into the QSH regime, the helical edge

---

**Figure 2.7** (left) Four-terminal $\pi$ bar device and (right) Four-terminal H bar.
goes around the boundary, including the voltage contacts. For instance, the edge current flows from contact 1 to contact 2, then to contact 3, and finally goes back to contact 4. According to the LB formula, the nonlocal resistance in H bar gives
\[ R_{14,23} = \frac{h}{4e^2} = 6.45k\Omega. \]

For a macroscopic device, where the distance between neighboring contacts is longer than the phase coherent length, the original LB formula is invalid. Since the contact in the QSH effect works as the phase breaking scatterer, we can also treat the phase breaking scatterers as the virtual contacts. As the length is longer than the coherence length, we could estimate the conductance by inserting several virtual contacts between contacts, due to the lack of phase coherence. As a result, for a device length \( L \), we put \( N = \frac{L}{\lambda_\varphi} - 1 \) phase breaking probes between neighboring contacts. In a macroscopic Hall bar, we obtain \( G_{14,23} = \frac{\lambda_\varphi e^2}{L \hbar} \) with the virtual contacts. Thus, in order to observe the quantized conductance, the size of device must be smaller than the coherence length. On the other hand, in the macroscopic device, although the quantized conductance cannot be observed, we could use the deviating conductance to obtain the coherence length of the helical edge.

2.6. Exciton

Inverted InAs/GaSb QWs initially were proposed as an artificial broken gap structure in 1980s to search exciton ground state. An exciton includes one electron and one hole, analogue to a hydrogen where one electron is bound to one proton.
Here, similarly, the electron and hole that are attracted to each other by Coulomb interaction are bound to form an exciton, which is a boson. The concept of the exciton was proposed by Frenkel in 1931(47) in crystal, where electrons and holes belonging to the same crystal cell were correlated by Coulomb attraction, with the binding energy of 100meV to 1eV, as shown in Figure 2.8.

![Figure 2.8](image)

**Figure 2.8** (left) The Frenkel exciton in the crystal cell. The electron and hole are bound to the same cell unit with a small Bohr diameter. (right) The Wannier exciton in the semiconductor. The Bohr diameter is much larger than the lattice constant.

Later, in the end of 1930s, Wannier and Mott proposed the concept of exciton in semiconductors. In semiconductors, the large dielectric constant makes the screening effect dominating, reducing Coulomb interaction. Due to the small effective mass of electrons and holes in semiconductors, the radius of the exciton is much larger than the lattice constant as shown in Figure 2.8, and the binding energy is relatively small, just a few meV. This kind of exciton is called Wannier Excitons(48).
For Wannier excitons, the effective mass approximation is used to treat the periodic crystal potential and describes electrons and holes as free particles with parabolic dispersion. In a semiconductor insulator, the valence band is fully occupied by electrons and there are not electrons in the conductance band. In this case, if there is a photon with the energy larger than the band gap, one electron would be excited from the valence band to the conduction band and leaves a hole in the valence band. The group velocity of electrons and holes at a finite inplane momentum is proportional to the band slope, given by $v = \frac{1}{h} \frac{\partial E}{\partial k}$. Thus for the example in Figure 2.9, the electrons and holes have opposite group velocities, which are also required by the momentum conservation. The excited electrons and holes are attracted by each other through Coulomb attraction, resulting into the exciton state. An essential requirement for the exciton formation is that the electron and hole should have the same group velocity, where the electrons and holes move with same speed and form a bound state. Thus, optically excited exciton only exist near $k=0$. 
Figure 2.9 The band diagram of a semiconductor under light radiation. An electron is excited by one photon from the valence band to the conduction band, and a hole is left in the valence band. The electron and hole have opposite group velocities.

In a bulk semiconductor, if we consider the exciton with a dielectric constant $\varepsilon$, the binding energy of the ground exciton state is

$$E_B = \frac{\mu e^4}{2\hbar^2 \varepsilon^2} = \frac{\hbar^2}{2\mu a_B^2},$$

where the reduced mass $\mu = \frac{m_e m_h}{m_e + m_h}$ and the Bohr radius $a_B = \frac{\hbar^2}{\mu e^2}$. These energy levels are located a little below the bottom of the conduction band with a series of hydrogen-like energy levels as shown in Figure 2.10.
Figure 2.10 The energy dispersion of exciton gas. The $n=1$ and $n=2$ state correspond to the 1s and 2p energy level, respectively. The 1s level is the exciton ground state which is less than the band gap $E_g$.

Exciton wave functions can be described by the relative and center-of-mass coordinates. The relative coordinate is fixed by the exciton Bohr radius, which is the average distance between the electron and hole. Thus we use the center-of-mass momentum to define the exciton state. As shown in Figure 2.10, where the energy dispersion for discrete exciton states are noted, $E(n,0)=0$ means the vacuum state in which the valence band is fully occupied and the conduction band is completely unoccupied. The $n=1$ exciton state is 1s state with respect to the continuum state by the exciton binding energy. As one photon is absorbed, due to the momentum conservation, only the exciton state near $k=0$ could be accessed.

In 1980s, the development of the film growth technology gave the opportunity to study Wannier excitons in 2D heterostructure by creating artificially potential wells and barriers for electrons and holes. In this case, the QW allowed for
the confinement of electrons and holes to the 2D plane, in layered semiconductor structure. In the QW, ideally the binding energy of exciton increased by a factor of 4 than the bulk exciton binding energy:

\[ E_{B}^{2D} = 4E_{B}. \]

Thus, 2D exciton with higher binding energy is more stable, and more suitable in the further search for exciton condensation, than 3D exciton. Since excitons are approximated as bosons and obey bosonic statistics, they are able to condense into a degenerate quantum state, i.e. Bose-Einstein condensation. For electrons and holes existing in the same QW, or called the direct exciton, the excited electrons in the conduction band quickly hop back into the valence band by emitting a photon with their excess energy. This electron-hole recombination time is pretty small, making exciton condensation difficult to realize in the direct exciton.

In order to increase the electron-hole recombination time, a new design of heterostructure was applied, where there were two parallel QWs with electrons in one well and holes in the other. Under the optical radiation, electrons and holes were generated similarly with the direct exciton. Then an electric field was applied perpendicular to the wall, which moved one type of carriers into adjacent well. The formed electron-hole bound state was called indirect exciton with the electron in one well and holes in the neighboring well. The indirect exciton was proposed by Lozovik and Yudson(49) as well as Shevchenko(50). Under this situation, the separation between electrons and holes reduced the recombination rate with the
lifetime increasing exponentially with the applied electric field, thus the spatially indirect excitons stood for a long lifetime with an increase by $10^3$ to $10^6$. Moreover, the spatially separated QWs have another advantage: the separation makes excitons similar to electric dipoles through the repulsive interaction that prevents electrons and holes from forming electron-hole plasma.

In the optical excitation that generates the electron and hole pairs, the short lifetime of optical excited non-equilibrium exciton is still a major difficulty to observe BEC of exciton. Additionally, the high exciton density and the low temperature are hard to realize in such non-equilibrium exciton. In this case, Datta(51) proposed an equilibrium exciton which was in the ground state in the spatially separated QWs. In his proposal, through varying the energy separation between the bottom of the conduction band in one well and the top of the valence band in the adjacent well, i.e. indirect band gap, one made the indirect band gap smaller than the indirect exciton binding energy. In this case, the exciton energy $E_x = E_g - E_B$ was negative and a spontaneous excitonic instability without recombination could be expected. This proposal is difficult to achieve in the direct exciton. In the direct exciton, because the dielectric constant that increases with the decrease of band gap causes a smaller binding energy, the band gap is always larger than the binding energy. By contrast, in the indirect exciton, the direct band gap can still be kept large with a small dielectric constant while the indirect band gap is tuned to be small. By tuning indirect band gap, we expect to realize the modulation of the exciton density with the electrostatic gate. Since optical excitation is not
necessary, this kind of experiment can be performed in the low temperature which is essential for the realization of exciton ground state. With the broken gap, the type II heterostructure-inverted InAs/GaSb is the nature candidate for this equilibrium exciton system.

2.7. BCS like exciton condensation

![Diagram](image)

Figure 2.11 Exciton phase transition and crossover. In the dilute limit of exciton density, exciton gas exists in higher temperature as an excited state. At lower temperature, excitons start to condensate following Bose-Einstein distribution. As the exciton density increases, the exciton gas starts to dissolve into the electron-hole plasma. In this case, at lower temperature, the plasma has a quantum phase transition to BCS-like exciton condensation, which is also called exciton insulator.

A similar example with excitons in semiconductors is paired electron-electron cooper pairs in superconductors. Naturally people expect the collection of excitons to experience similar macroscopic coherence with superconducting pairs under certain conditions. Different from cooper pairs, the internal relative coordinator of exciton can be significantly modulated with regard to the
neighboring exciton distance, affecting the property of a collection of excitons and
triggering quantum phase transitions as shown in Figure 2.11. Nevertheless, the
quantum coherent states of excitons are also sensitive to the temperature, similar
with the superconducting cooper pairs. At the low exciton density and under
extremely low temperature, if the Bohr radius of exciton is much smaller than the
adjacent exciton distance, excitons are well separated and can be treated as a real
boson, with the internal structure inert. In this case, the collective mode of excitons
induces BEC in center-of-mass momentum space, meaning all excitons are located in
zero center-of-mass momentum, in much the same way as the weakly interacting
bosons. As the exciton density increases, the neighboring exciton distance decreases;
at the same time, the screening effect of electron-hole Coulomb interaction is
strengthened, making the exciton binding energy also decreases. The binding energy
is related with the Bohr radius, in a relation of $a_B = \frac{\hbar}{\sqrt{2mE_B}}$, where $m$ is the
reduced effective mass. In this case, the exciton wavefunction becomes less bound
together and grows bigger and bigger. The neighboring exciton wavefunctions
overlap with each other, changing from single exciton binding to collective electron-
hole pairing. As a result, free fermions due to weak pairing have the loosely
correlated wavefunction in much the same way with that of the weak-coupling BCS
cooper pairs in superconductor. Here excitons cannot be treated as hydrogen-like
bosons, but form correlated pairing states. That is to say, the wavefunction overlap
makes it less clear which electron pairs with which hole, in a sharp contrast with
BEC of exciton where one electron is strongly bound to one hole. Such state in the
dense electron-hole system is called excitonic insulator or BCS like exciton condensation as shown in Figure 2.12. Here, the electrons and holes play the role of electrons with different spins in BCS superconductor. As the density increases, there is a continuous phase transition from BEC to BCS like exciton condensation. For \( nd^2 \ll 1 \), where \( n \) is the exciton density, the condensed excitonic phase is BEC of exciton. In the other limit, for \( nd^2 > 1 \), \( n \) means density of cooper pairs with zero momentum, and the excitonic phase is BCS like condensation.

![Figure 2.12](image)

**Figure 2.12 The crossover from BEC of exciton to BCS like exciton condensation.**

As the temperature is higher than the critical one, in the dilute limit, the condensed excitons in zero center-of-mass momentum become hydrogen-like boson gas. As the exciton density increases, the hydrogen-like excitons dissociate and become free electron and hole plasmas. The exciton gas and the free electron hole plasma is the excited state, compared with BEC of exciton and BCS-like exciton condensation which are the ground states.

It is interesting to compare BEC of exciton, BCS superconductor and BCS like exciton condensation. For BEC of exciton, it is strongly bound electrons and holes (bosons) that condense; for superconductors, it is weakly correlated electrons
which are condensed fermions; for BCS-like exciton condensation, it is weakly correlated electron and hole that are also condensed fermions. In 1968, Keldysh and Kozlov(37) pointed out that the BCS wavefunction could define all of these three quantum states.

We start from an isotropic, parabolic, and indirect gap two band double layer semiconductor QWs. Electrons exist in the conduction band of one layer and holes exist in the valence band of the other layer. The Hamiltonian of this system can be characterized(52) as

\[
\hat{h}_{e-h} = \sum_i \frac{p_{i,e}^2}{2m_{i,e}^*} + \sum_i \frac{p_{i,h}^2}{2m_{i,h}^*} + \sum_{i<j} \frac{e^2}{4\pi\epsilon |\vec{r}_{i,e} - \vec{r}_{j,e}|} + \sum_{i<j} \frac{e^2}{4\pi\epsilon |\vec{r}_{i,h} - \vec{r}_{j,h}|} \\
+ \sum_{i,j} \frac{e^2}{4\pi\epsilon \sqrt{|\vec{r}_{i,e} - \vec{r}_{j,h}|^2 + d^2}}
\]

where \( \epsilon \) is the average dielectric constant, and \( d \) is the separation between the electron layer and the hole layer. The first two terms define the kinetic energy of the electron and hole, respectively. The third and fourth terms mean intralayer Coulomb interaction between electron/hole. The last term defines the interlayer Coulomb interaction. The interlayer electron hole tunneling term is neglected, since in our experiment its magnitude is found to be negligible. The Hamiltonian in the second quantization form can be expressed as:
\[ \hat{h}_{e-h} = \sum_k E_k^e a_k^+ a_k + \sum_k E_k^h b_k^+ b_k + \frac{1}{2} \sum_{k,k',q} V_{q}^{ee} (a_{k+q}^+ a_{k'-q}^+ a_k a_k) \]

\[ + \frac{1}{2} \sum_{k,k',q} V_{q}^{hh} (b_{k+q}^+ b_{k'-q}^+ b_k b_k) - \sum_{k,k',q} V_{q}^{eh} (a_{k+q}^+ b_{k'-q}^+ b_k a_k) \]

This Hamiltonian can be rewritten as \( \hat{h}_{e-h} = \hat{h}_0 + \hat{h}_{coul} \), where \( \hat{h}_0 = \)
\[ \sum_k E_k^e a_k^+ a_k + \sum_k E_k^h b_k^+ b_k \]
and \( \hat{h}_{coul} = \frac{1}{2} \sum_k (V_{q}^{ee} \rho_{q}^e \rho_{-q}^e + V_{q}^{hh} \rho_{q}^h \rho_{-q}^h - 2V_{q}^{eh} \rho_{q}^e \rho_{-q}^h). a_k^+ \)

(\( a_k \) and \( b_k^+ \) (\( b_k \)) are creation (annihilation) operators for electrons in the conduction and valence bands. The density operators are \( \rho_{q}^e = \sum_k a_{k+q}^+ a_k \), \( \rho_{q}^h = \sum_k b_k b_{k+q}^+ \). The parameters \( V_{q}^{ee} = V_{q}^{ee} = \frac{e^2}{2\varepsilon q} \) and \( V_{q}^{eh} = \frac{e^2}{2\varepsilon q} e^{-qd}. \)

With mean field theory, we have the following ground state wavefunction of interacting electrons and holes, in much the same form of the BCS wavefunction in superconductor

\[ |\Psi_0\rangle = \prod_{k_l} (u_{k_l} + v_{k_l} a_{k_l}^+ b_{k_l}^+) |0\rangle \]

where \( |0\rangle \) is the vacuum state corresponding to the case where the valence band is completely full and the conduction band is totally empty. \( u_{k_l}^2 \) (\( v_{k_l}^2 \)) represent occupation (annihilation) probability and \( u_{k_l}^2 + v_{k_l}^2 = 1 \). In the BCS theory, the solutions are obtained by variation principle minimizing the free energy at a given chemical potential \( \mu \): \( f = \langle \hat{h}_{e-h} \rangle - \mu \langle n \rangle \), where \( \langle n \rangle = \sum_k v_k^2 \). Then we obtain a couple of self-consistent nonlinear equations
\[ \Delta(k) = \sum_{k'} V_{k-k'}^{eh} \frac{\Delta(k')}{E(k')} \]

\[ \xi(k) = (E^h_k + E^h_k - \mu) - \sum_{k'} V_{k-k'}^{hh}(1 - \frac{\xi(k')}{E(k')}) \]

\[ E(k)^2 = \Delta(k)^2 + \xi(k)^2 \]

\( \Delta(k) \) is the gap-equation and also the order-parameter, which defines the exciton gap. \( E(k) \) can be understood as the pair-breaking excitation spectrum: it is the energy cost of taking one-pair out of the condensate and placing them in plane-wave states of momentum \( k \), which corresponds to the energy of absorbed photon.

![Figure 2.13](image)

**Figure 2.13** Order-parameter \( \Delta(k) \), and the pair-excitation spectrum \( E(k) \), as a function of \( k \) at four different densities. This figure is adopted from (52)

If we solve the above equations with the parameters of GaAs QW, we obtain a series of results for different exciton densities as shown in Figure 2.13. The dotted
line is \( r_s = 2.11 \), where \( r_s \) is defined by \( \pi r_s a_B^2 = 1/n \); the solid line is \( r_s = 3.69 \); the dashed line is \( r_s = 4.72 \); the dotted-dashed line is \( r_s = 9.56 \). For each \( r_s \), the upper curve describes \( E(k) \), and the lower curve is \( \Delta(k) \). The dotted line represents the energy dispersion of BCS like exciton condensation. As \( r_s \) increases, the exciton density decreases with the crossover from BCS like exciton condensation to BEC of exciton. The dotted-dashed line can be identified as the case of BEC of exciton. This crossover is a smooth transition. In the following chapters, we will revisit this formula with the realistic case of inverted InAs/GaSb QWs.

Under the search for BEC of exciton, there have been significant achievements in experiment studies of optically excited exciton, such as the enhanced exciton mobility, increased radiative decay rate, and the photoluminescence noise\(^{(53,54)}\). All of these results suggest the collective and coherent behavior of many-body effect. However, in the other limit, i.e. BCS-like exciton condensation, there is not much advance in the non-equilibrium exciton. One main reason is that the optical excitation could not generate high density exciton under the low temperature. On the other hand, the host system for the equilibrium exciton, just like the inverted InAs/GaSb QWs, is very rare. Moreover without the gates, it is hard to modulate the CNP density in the inverted InAs/GaSb QWs to a suitable value for exciton ground state, while the fabrication of the electrostatic gates is always a challenge. To perform optical absorption experiments, the semitransparent gate need to be millimeter by millimeter, which is a giant challenge for the fabrication, since even the gate with the area of micrometer by
micrometer has already been difficult. One main problem is the fast oxidization on the surface, inducing the gate leakage and hysteresis. There is another problem in this system, as we have discussed: just after the invention of this inverted InAs/GaSb structure, it was pointed out that electron-hole tunneling naturally suppressed Coulomb interaction and was the competitive ground state with regard to exciton insulator. Since the 1980s, there were lots of debates about the existence of exciton insulator in inverted InAs/GaSb QWs. In 1995, there were several far infrared magneto experiments (32, 33) in the ungated systems and 1s-2p transition was observed under magnetic field. However, the similar result was obtained in another group (55) but they did not find the activated behavior corresponding to the dissociation of exciton. Due to the lack of gates, these experiments were not performed at the CNP, and therefore there was not a conclusive result.

In this thesis, we will present the observation of exciton insulator in the inverted InAs/GaSb QWs. We will show the exciton insulator is a competing ground state with hybridization gap and in the certain condition the exciton insulator dominates.
Chapter 3

Experiment basic

3.1. Fabrication

Semiconductor wafers of InAs/GaSb QWs were grown by molecular beam epitaxy (MBE) technique. Typical wafer structure as shown in Figure 3.1 contains 1μm thick AlGaSb barrier layer, 12.5nm InAs/10nm GaSb QWs for wafer A and 11nm InAs/7nm GaSb QWs for wafer B, 50nm AlGaSb upbarrier, a 3nm GaSb cap layer, and a 3nm InAs cap layer. In wafer A, the interface between GaSb and InAs QW is doped with a sheet of Si in a dilute concentration of $\sim 10^{11}/cm^2$ while there is not doping in wafer B. Doped Si would be a donor for electrons and an acceptor for
holes, resulting into a random potential fluctuation with residual carriers localized.

For wafer C, Indium alloy is applied to GaSb QW with a strained effect.

![Diagram of wafer structure]

**Figure 3.1 shows a typical InAs/GaSb wafer structure.** 2DEG is confined in the InAs layer while 2DHG is confined in the GaSb layer.

The doped GaAs substrate serves as the global back gate. In the flip chip process, this substrate can be etched and one could deposit the metal gate on the local area to define the local back gate. This kind of local gate is necessary for the multiple back gate controls, such as Coulomb drag experiment. In a normal fabrication process, we use photolithography to define the mesa and the wet etching or dry etching is followed. For the wet etching, the solution used is $H_3PO_4: H_2O_2: H_2O: Citric acid = 3: 5: 220: 55$. Usually 200nm etching depth is enough. For the dry etch, $BCl_3$ mixed with $N_2/Ar$ is used. Compared with the wet etching, the dry etching has a more vertical profile. Sometimes, $N_2/Ar$ can also be used for milling if the mask is hard baked. After the mesa is etched, we use ebeam lithography to define the submicro structure with the following wet or dry etching.
As a next step, alloy contacts are defined with photolithography and made with ebeam evaporator. The deposited film includes Ge, Pd and Au with a layer thickness ratio of 43:30:87. If there is not a high temperature process following, one need to anneal the sample at 300°C in a forming gas (30% H₂ and 70% N₂) for a few minutes or we can directly use Indium to make contacts without annealing. Then we would have the structure ready for the gate fabrication. There are several options to choose. One way is to deposit SiN film with Plasma enhance chemical vapor deposition. With this method, the surface of the sample is automatically cleaned with hydrogen plasma at the beginning of the film deposition. This clean surface is helpful to reduce the hysteresis of gate sweeping. Also this step is performed under 300°C so that the metal is automatically annealed. You can also deposit SiO₂ or Al₂O₃ with physical vapor deposition. This deposition method requires a low pressure environment, usually 10⁻⁶ torr, in the deposition process. Another deposition method is to grow Al₂O₃/HfO₂ through Atomic layer deposition, which is in principle a chemical method and different from previous two methods. After the dielectric layer is deposited, we need to use the photolithography to make vias. One can select wet etching or dry etching. In the wet etching, 5% diluted HF solution is used to etch Al₂O₃ and the commercial buffered HF is used to etch SiO₂, while in the dry etching, CF₄/SF₆ and O₂ are used to etch SiN. Finally, we use ebeam lithography or photo lithography to define the gate pattern which is followed with the metal gate deposition and lift off. The above is a normal fabrication process.
In this system, there is alternative method about the gate fabrication. In this method, the 50nm AlGaSb upper barrier is used as a Schottky barrier. The principle is to utilize citric solution to etch InAs QW on the boundary of the sample. In the lateral etching profile, there will be an undercut structure which could protect the InAs and GaSb QWs away from the gate metal. To optimize the gate modulation, the aluminum is a suitable gate material. As a result, the alloy metallization, dielectric layer deposition and via etch steps could be neglected. This method is more convenient than the conventional method. However, it should be careful to use this method here. If the undercut distant is small (tens of nanometer), there will be screening on the helical edge from the metal gate. In this case, the measured helical edge state may not be treated as an isolating edge. In a contrast, in the conventional method, the edge is decoupled from the metal gate by a thick dielectric material. Ideally we hope the metal gate is separated from the helical edge in a very large distance, which is usually 100-200nm for safety.

Under the perpendicular magnetic field, the normal fabrication may not work. This is because the metal contacts are connected to the gated mesa with the ungated 2DEG that becomes insulating under the certain magnetic field. Thus in order to perform measurement under ultrahigh magnetic field, such as 35T, the metal contacts need to be underneath the electrostatic gate.
3.2. Conductance measurement

In the conductance-voltage measurement, transport results from several kinds of devices were presented. The first kind was the mesoscopic devices that included 1x2\(\mu\)m Hall bar, 0.5x1\(\mu\)m Hall bar, 1x1\(\mu\)m Hall bar, 1x2\(\mu\)m \(\pi\) bar, 0.5x1\(\mu\)m \(\pi\) bar, 1x3\(\mu\)m H bar, and 1x1\(\mu\)m two-terminal junction. In these devices, quantized plateaus of helical edge modes were observed. The second kind was macroscopic Hall bars with typical sizes of 5x10\(\mu\)m, 10x20\(\mu\)m and 25x50\(\mu\)m, giving the information of the helical edge coherence length. The third kind of devices was Corbino device. In the Hall bar measurement, the signal was from the edge and the potential bulk channel. Only the Hall device measurement could not tell where this signal was from. Usually we had an assumption that the bulk was insulating so that what we measured was from the helical edges, which however needed to be verified. In this case, the Corbino device provided the unique and direct measurement to the bulk property. The Corbino device included a metal contact of 1mm inner diameter, a concentric metal contact of 2mm outer diameter, and the mesa that is between them. In this geometry, the edge was shunted by metal contacts and would not contribute to the measured conductance. With the Corbino device, we not only could determine whether the bulk was insulating, but also can obtain the direct information about the bulk gap. The fourth kind of devices was unsymmetrical Hall bar. In this geometry, there were two contacts on one side, while there was only one contact on the other side. The purpose of this geometry was to detect the CNP position. The contact works as a phase breaking center for helical edge state. In this
case, the unsymmetrical distribution of contacts brings different segments of helical edges. For example, on the side of two contacts, there were three segments while on the other side, there were two segments. As a result, the Hall contacts measured a greatly increased voltage drop in the presence of helical edge state at the CNP. Note that if the device was dominated by 2DEG, there was not any difference between this unsymmetrical Hall bar and symmetrical Hall bar. By this way, we could know the position of the CNP, which was necessary to obtain the CNP density. The devices were measured at the low temperature from 20mK to 4.2K. The low temperature measurements were carried out in a He3 refrigerator combined with a 8T superconducting magnet or in a He3/He4 dilution refrigerator with a 18T/35T magnet at the National High Magnetic field laboratory.

Figure 3.2 Sketch of the Corbino measurement setup. The outer conductor was grounded, and a voltage was applied in the inner conductor; we measured the current flowing across the ring. V_f and V_b were applied to modulate the carriers in the ring.
3.3. Capacitance measurement

![Diagram for capacitance measurement in QWs.](image)

**Figure 3.3 Diagram for capacitance measurement in QWs.**

Although the conductance measurement with Corbino geometry was employed, the DOS information was not involved. Through the conductance measurement, a real gap was hard to be distinguished from Anderson localization, since both of them had an insulating behavior. However, the reduced DOS would happen in a real gap but not in the localization, becoming the prominent difference between the localization and a real gap. Regardless, the capacitance measurement provided the direct information of DOS about the bulk, which was essential in the exploration of unpredicted insulating state. Here, as shown in Figure 3.3, we performed Capacitance-voltage (CV) measurement with the front and back gates to study the bulk property. Besides dc front and back gate voltages, a small ac voltage with 100Hz low frequency $f$ was delivered to the front gate. By measuring the charge deviation per voltage, we obtained the capacitance between the front gate
and QWs, \( C_m = \frac{dQ}{dv} \). There were four components contributing to the measured capacitance per unit area \( C_m \) in our device structure,

\[
\frac{1}{C_m} = \frac{1}{C_d} + \frac{d}{\varepsilon_{AlGaSb}} + \frac{<z>}{\varepsilon_{QW}} + \frac{1}{e^2D(E)}
\]

where \( C_d \) was the capacitance from the dielectric layer per unit area, \( \varepsilon_{AlGaSb} \) and \( d \) were the dielectric constant and the thickness of AlGaSb, respectively, \( \varepsilon_{QW} \) and \( <z> \) were the average dielectric constant and average distance from the carrier to QW, and \( D(E) \) was the DOS. The first three terms represented the geometry capacitance \( C_g \) and the last term showed the quantum capacitance per unit area. The geometry capacitance was nearly constant, but changed a little dependent on the third term. In the structure where InAs QW was on the top of GaSb QW, as the electron dominated the system, \( <z> \) represented nearly the half thickness of InAs QW. When the system was dominated by holes, \( <z> \) means the sum of the thickness of InAs QW and half thickness of GaSb QW. In experiments, this difference was nearly 1%, so it could be neglected(26). In the electron or hole regime, \( C_g \) was much smaller than the quantum capacitance of 2DEG, therefore \( C_g \) was nearly consistent with \( C_m \). If the system was in a soft gap filled with carriers, the DOS was the same with that in the electron or hole regime. If the system was in a hard gap where the DOS was near zero, the quantum capacitance decreased and was much smaller than \( C_g \), so \( C_m \) became related with the DOS. Experimentally as the system went from the electron regime to a hard gap, we could expect to see a capacitance drop within the hard gap.
The normal enamelled wires could not be used in the CV measurement, due to the large capacitance coupling between them. The measured capacitance is the sum of the capacitance between wires and the device capacitance. Therefore to perform CV measurement, the capacitance between wires should be eliminated and the coax is used. Different from the enamelled wires, the coax has additional outside metal shell that screens the ac signal, making CV measurement possible. To perform CV measurement, I modified a He3 probe by adding four coaxes. Since the coax has small resistance connecting the room temperature and 300mK, it is a good thermal conductor which could increase the evaporation of He3 liquid. Therefore to decrease the thermal power from the room temperature, the coax was intensively coiled to the 4K regime of the probe with GE varnish as shown in Figure 3.4. In this way, even after four coaxes were added, there was little effect to the stability of He3 system. Then the capacitance was measured with a LCR meter as shown in Figure 3.5. 100Hz is the lowest frequency provided by this meter.
Figure 3.4 Coaxes in the 4K regime of the probe

Figure 3.5 4274 LCR meter
3.4. Charge neutral point density

In the single carrier regime, the electron/hole would change with the geometry capacitance. However, in the two-carrier regime, owing to the imperfect screening of 2DEG, the front gate could simultaneously tune the electron and hole densities. On the other side, if we modulate the back gate, as the hole appears, the back gate only tunes the hole density.

The equivalent circuit mode is shown in Figure 3.6, where $c_f/c_b$ means the geometry capacitance between InAs QW/GaSb QW and the front/back gate, $c_s$ means the capacitance between two QWs, $c_e = \frac{e^2 m_e}{\pi h^2}$ is the quantum capacitance of electrons, $c_h = \frac{e^2 m_h}{\pi h^2}$ is the quantum capacitance of holes.

Under a fixed small perpendicular field, the density of the single carrier is obtained by $R_{xy} = -\frac{B}{ne}$ for electrons, and $R_{xy} = \frac{B}{pe}$ for holes. This density is linear with the gate voltage, giving a constant carrier density increment per gate voltage.
In our system, if the Fermi level is lifted above the top of valence band, there is only electron in QWs, which can be achieved by tuning front gate and back gate. Figure 3.7a shows for $V_f$ sweeping from 2V, the electron density is high and linear with $V_f$ so we can extract electron density increment per $V_f$. As we decrease $V_b$, the top of the valence band is lifted and may be raised above the Fermi level, where the hole is induced. At $V_b=-1.5V$, the electron density increment per $V_f$ still keeps the same with that at $V_b=0V$. It indicates for either case the Fermi level is located above the top of the valence band. Then at more negative $V_b$, the electron density rate starts to decrease, because the electron cannot fully screen the electric field. Then the electron density rate decreases with more negative $V_b$. Finally at $V_b=-7.5V$, where the hole density is comparable with the electron’s, the electron density rate keeps saturated and the saturated rate is $1.8 \times 10^{-11}/cm^2V$ as presented in Figure 3.7b.

On the other hand, we obtain the capacitance per unit area from CV measurement. The capacitance per unit area $C_m$ means the absolute density($|n|+|p|$) increment per $V_f$. So if we integrate the capacitance over $V_f$ and know the initial absolute density, we can obtain the absolute density at a certain $V_f$. For the high electron density, where the Fermi level is located above the top of the valence band, the hole density can be neglected so the absolute density is just the electron density which could be obtained from SdH oscillation. The absolute density versus $V_f$ is shown in the blue line of Figure 3.7a. The absolute density decreases with a constant rate until it gets to nearly zero, where the CNP is nearly reached(vertical dashed black line).
Under 1T perpendicular magnetic field, the magneto transport trace is shown as the red dotted line in Figure 3.7a, where $V_f$ is normalized to $\Delta V_f = V_f - V_{cnp}$ ($V_{cnp}$ is the front gate voltage at the CNP). In the linear regime, the electron density extracted from the magneto transport agrees with the net density. For $\Delta V_f=1V$, the trace bends up and doesn’t correspond to the electron density, indicating the hole emerges. In this case, although the absolute density still goes along the blue trace, the electron density should decrease more slowly due to the imperfect screening.

Then we make an approximation that as hole appears the electron changes with the saturated rate (green dashed line) until the CNP is reached (Figure 3.7. Finally we can get the equilibrium density $n_0$ (the circle marker).

---

**Figure 3.7 a,** Illustration for the equilibrium density at CNP. The red dotted line is for magneto transport of unsymmetrical macro-Hallbar. In the linear part of the trace near 2V, the trace represents the electron density. The blue line is for the capacitance integration over $V_f$ which represents the absolute density. The green dashed line is for the electron density changing with the saturated rate. $n_0$ could be read from the crossing point of vertical dashed black line and green dashed line. **b,** Electron density increment per $V_f$ in high electron density regime for different $V_b$. 
The key of this method is to make the linear approximation in the two-carrier regime, instead of the single-carrier regime. In the simple approximation used in previous literatures\(^{(56)}\), people normally has the assumption: the modulation of gate voltage to the electron or hole remains the same until the carrier is depleted as shown in the dashed blue line of Figure 3.8. The two approximations have no differences in the deeply inverted band\((n_0 \text{ is large, i.e. } \sim 2 \times 10^{11} \text{cm}^{-2})\), where two carriers always dominate. However, in the shallowly inverted band\((n_0 \text{ is small, i.e. } <10^{11} \text{cm}^{-2})\), under the simple linear approximation, the electron density would follow the blue line while the hole density goes along the extended line of hole hall signal. It not only violates the charge conservation which means the total density variation(electron density variation plus the hole density variation) per voltage is more than what the gate could provide, but also neglects the screening effect. As a result, it gives a low-estimated density or even negative density(there is not overlap between the electron density and hole density traces, misleadingly indicating the band is noninverted). In a contrast, after the imperfect screening is considered with the model in Figure 3.6, the approximation we used not only shows that the band is “inverted” but also gives the accurate \(n_0\).

It should be mentioned according to the results in Figure 3.7 we could obtain electrons(holes) change with \(1.8 \times 10^{-11}/\text{cm}^2\text{V}(0.66 \times 10^{-11}/\text{cm}^2\text{V})\) in the two-carrier regime. Combined this result with the model in Figure 3.6, we can extract the effective electron mass which is \(0.055m_o\).
Figure 3.8 The comparison between the approximation we used and the linear approximation for high $n_0$ and low $n_0$.

3.5. Previous experiments

Before we delve into experiment findings, it is helpful to quickly review previous experiments in inverted InAs/GaSb QWs. The hybridization gap was firstly confirmed through the inplane magneto transport and capacitance measurements(26,57), where QWs were tuned into the deeply inverted band. Hybridization gap was found to be conductive with the high DOS, indicating a soft gap. Later, transport experiments by Cooper (58) and Knez (59) in a double gate structure showed the strong resistance peak corresponding to hybridization gap. Then in the conductive gap, through sizable bulk conductance, Knez deducted the bulk contribution and revealed helical edge modes in mesoscopic devices(12). Moreover Knez probed helical edge modes in a superconductor-semiconductor heterostructure via Andreev reflection(13). These experiments have indicated the existence of helical edge modes in inverted InAs/GaSb QWs, but due to the
conductive hybridization gap the TRS protection was not demonstrated. On the other hand, in the initial study of the QSH effect in HgTe QW, with the insulating bulk, the local and nonlocal measurement observed the conductance of helical edge state that was close to the quantized value, as well as the protection from the TRS. However, the quantized plateau of the QSH effect has never been observed in any QSH system, making the existence of the QSH effect not fully conclusive.

Furthermore, in both HgTe QW and InAs/GaSb QWs, the edge coherence length was limited to 1-2 µm and the bulk gaps were nearly 2-4meV, which restricted the potential application of QSH insulator.

On the other hand, in the search for exciton ground state, pioneer works were made by Cheng(32) and Kono(33) near 1995 in magneto-far infrared experiments with ungated inverted structure, where an additional absorption peak above the cyclotron resonance was observed and interpreted as 1s-2p magneto exciton transition. These experiments were made out of the CNP, which could not provide clear evidence to the existence of exciton insulator(26). Later several magneto far infrared experiments were performed and the additional peak was reported and argued to be from nonparabolicity(55) or come from LL hybridization(60–62). The further progresses to solve these arguments faced two critical problems: one is technical and the other is physical. Physically, electron-hole tunneling naturally suppressed Coulomb interaction and eliminated exciton ground state(51). Datta(51) proposed to insert a AlSb barrier between electron and hole layers to eliminate hybridization effect. However, the barrier also decreased the
binding energy of exciton, which may be responsible to the semimetal behavior reported in the inverted structure with AlSb barrier (58). Then technically, one hopes Fermi level could be placed at the CNP by electrostatic gating. However, the fabrication of semitransparent gate in this material is challenging, due to fast oxidation on the surface. Another factor is the gate area need to be millimeter by millimeter, required by the long wavelength of far-infrared light. The gate leakage possibility exponentially increases with the gate area. Both difficulties nearly stopped the further search for exciton equilibrium state in this system for nearly 20 years.

In this thesis, the problems mentioned above will be solved.
Chapter 4

Quantum Spin Hall effect in stained InAs/InGaSb quantum wells

4.1. History

In 2008, Liu (10) predicted that the TRS QSH insulator could exist within the bulk hybridization gap of inverted InAs/GaSb QWs. In 2011, Knez (12) found that the helical edge still existed, although the bulk with the hybridization gap was conductive. These results motivated us to realize the insulating hybridization gap, in order to confirm the existence of helical edge modes. Then two methods were proposed: one was to dope the dilute silicon between InAs and GaSb QW; the other
was to alloy InSb into the original GaSb QW. Chronologically, the bulk insulator was realized with the first method in 2012 (appear in arxiv.org in 2013 and publish in 2015(27)) and with the second one in 2014 (appear in arxiv.org in 2016(64)). However, careful investigations(65) showed in the first method it was exciton insulator not hybridization gap that played the major role in the insulating bulk. To keep a continuous logic flow, I will talk about the second method first in this chapter. In the phase diagram(Figure 4.1) introduced by Liu(10), as shown in the gold circle mark(Figure 4.1), this chapter will be focused on the deeply inverted regime.

![Figure 4.1 Phase diagram for different front and back gate voltages. The red regime represents the inverted band and the blue regime indicates the noninverted band. The figure is adopted from Ref. (10).](image-url)
4.2. Strain effect in InAs/InGaSb

The idea of InGaSb alloy comes from the strain effect in InAs/GaSb type-II superlattice infrared detectors. InAs/GaSb infrared detectors\(^{(66)}\) have been known to possess two prominent advantages over other materials: (1) InAs and GaSb have approximately the same lattice constant of 6.1 Å, thus high quality superlattice structure could be grown by MBE; (2) Due to the broken-gap band alignment, the bulk gap of InAs/GaSb QWs can be well controlled by adjusting the QW width. To reach the long-wavelength sensitivity (wave length>10μm), thicker layer and smaller band gap are necessary, however in this situation, optical matrix elements that are determined by the overlap between wavefunctions of electrons and holes are small, leading to a low optical absorption efficiency. An alternative way to solve this problem is to grow strained-layer InAs/InGaSb superlattice by alloying GaSb with InSb, proposed by Smith and Maihlot in 1987\(^{(67)}\). InSb has the lattice constant of nearly 6.4 Å, thus InGaSb still has the similar lattice constant with InAs. Because of the compression strain in the growth plane, the bottom of the conduction band in InAs QW shifts downward while the energy level of valence band in InGaSb splits into heavy hole and light hole levels, respectively, where the top of heavy hole band is higher than the original top valence band of GaSb. As a result, to reach a fixed band inversion, the InAs/InGaSb QWs are narrower than InAs/GaSb QWs thereby increasing the optical absorption efficiency. Such strain-engineering has led to the invention of high-performance long-wavelength superlattice infrared detectors. Similar physics idea may guide the construction of a large-gap QSH insulator.
The strain could eliminate the mismatch between the Fermi surfaces of the conduction and valance bands. In inverted InAs/GaSb QWs, the heavy hole band is mixed with the light hole band, and thus the Fermi surface of the valence band is not round, which mismatches with the Fermi surface of the electron band, as shown in Figure 4.2. A detailed calculation based on the self-consistent 8 band calculation is presented in the left panel of Figure 4.3. The right panel of Figure 4.3 indicates the components of the conduction band and the valence band with respect of the heavy hole, the light hole and the electron. In this calculation, in the small plane momentum, the valence band is dominated by the heavy hole. Nevertheless, in the large plane momentum, the valence band is the mixture of heavy hole and light hole, resulting into a nearly square Fermi surface. As we have discussed, in order to obtain stronger hybridization, the cross point should locate at a large momentum with stronger light hole coupling, which has to result into a square Fermi surface of the hole band at the CNP. However, by alloying GaSb with InSb, the heavy hole-light hole splitting is enhanced, making the heavy hole separated from the light hole in InGaSb QW. As a result, the Fermi surface of electrons has a better match with that of holes, reducing the residual non-hybridized carriers.
Figure 4.2 Contour map of the Fermi surface of deeply inverted InAs and GaSb bands. The figure is adopted from (57).

Figure 4.3 (a) Energy spectrum of InAs(11nm)/GaSb(7nm) quantum well with an in-plane magnetic field B =0T with selfconsistent calculation by eight-band Kane model (Solid line); (b), (c), (d) and (e) are the components of conduction band 2(CB2), conduction band 1(CB1), valence band 1(VB1) and valence band 2(VB2), respectively. The dash dotted line in (a) is the reduced single band model. The data is calculated by Dr. Kai Chang.
On the other hand, to achieve the inverted band structure that is essential for the QSH insulating phase, InAs/GaSb QWs should not be too narrow. Based on the strain effect described above, we could reach the same inverted band structure with narrower strained-layer InAs/InGaSb QWs. The tunneling between electrons and holes in narrower QWs becomes stronger therefore the hybridization-induced gap should increase. In the MBE-wafer we used for present transport measurements, the thickness of InAs QW is 9.5 nm while the thickness of InGaSb QWs is 4 nm, as shown in Figure 4.4. Transmission electron microscope photography in Figure 4.5 shows crystalline structure remains coherent across the heterostructure interfaces regardless of in-plane strain. Figure 4.6 shows the calculated band structure by 8-band Kane model, with a hybridization gap \( \sim 12 \) meV, which is about three-fold increases from the value \( \sim 4 \) meV in unstrained InAs/GaSb QWs(26). Well resolved SdH oscillations in both electron regime and hole regime are presented in Figure 4.7.

![Figure 4.4 Wafer structures of the strained-layer InAs/Ga0.75In0.25Sb QWs used for experiments.](image)

Figure 4.5 A TEM photograph of a strained wafer; blue and red lines are guide for eyes. The data is taken by Zhongdong Han.

Figure 4.6 Calculated band structure of the 9.5 nm InAs/ 4nm Ga0.75In0.25Sb QWs, CB1, VB1 and CB2, VB2 represent for different spin components. The band is calculated by Dr. Kai Chang.
Figure 4.7 Magneto-transport data of the 30 μm ×10 μm Hall bar in a, electron-dominant regime and b, hole-dominant regime. The data is taken by Tingxin Li.

4.3. Transport properties of bulk states in strained-layer

InAs/InGaSb Quantum wells

Figure 4.8 $B/eR_{xy}-V_f$ trace of a 50×50 μm Hall bar under different back gate voltages. The top trace corresponds to 4V backgate voltage and the bottom one corresponds to 0V backgate voltage.
I fabricated a 50\times 50\mu m Hall bar device with the front and back gates. Under 1T perpendicular magnetic field and several fixed back gate voltages, I swept the front gate voltage and monitored $B/eR_{xy}$, as shown in Figure 4.8. As the electron density was relatively high, i.e. $5 \times 10^{11}/cm^2$, $B/eR_{xy}$ was linear with the front gate voltage, indicating an electron dominant regime. Near the CNP, the trace became divergent, exhibiting the gap regime was accessed. According to the linear approximation, I could extract the CNP density. Then I raised the back gate voltage, and the CNP density increased as shown in Figure 4.9, indicating a more inverted band.

![Figure 4.9 The CNP density vs back gate voltage traces.](image)
To directly measure the bulk conductance, I fabricated double-gate Corbino devices. In this case, the edge conductance was shunted and had no contribution to the signals. Figure 4.10 showed the traces of conductance per square versus $V_f$ at $V_b = 0$ V and 4 V, respectively. Electron-hole hybridization was most favored at the CNP where electrons and holes with the same energy and momentum resonantly tunneled. At the CNP, the conductance dip dropped to zero, showing the entrance into an energy gap. At more positive $V_b$, the bulk band became more inverted, with a less insulating bulk. Nevertheless, the bulk conductance was negligible at a low temperature, about 100 M$\Omega$ per square at 20 mK for $V_b = 0$ V, and about 25 M$\Omega$ per square at 20 mK for $V_b = 4$ V.

To verify the origin of the observed gap, the in-plane magnetic field $B_{\parallel}$ was applied. Under $B_{\parallel}$ applied along x axis of the sample, Lorenz force gave tunneling
carriers additional momentum along y axis, resulting in a relative shift of band dispersions \( k_y = -eB(z)/\hbar \) (tunneling distance \( z \) was limited by one-half thickness of QWs). With the momentum shift, the electrons and holes at two crossing points would have different energies (Figure 4.11). If the momentum shift was large enough, the Fermi level could not totally locate in the gap with the gap disappearing. Consequently, the hybridization was suppressed due to the momentum mismatch, and the conductance in the gap significantly increased.

![Figure 4.11](image)

**Figure 4.11** (left) the in-plane dispersion relations of electrons in InAs and holes in GaSb and (right) the in-plane dispersions under a parallel magnetic field.

This method (26) was used to distinguish whether the gap in inverted InAs/GaSb QWs came from the hybridization. As shown in Figure 4.10, the insulating state at the CNP was weakened and gradually disappeared at increasing
B\textsubscript{∥}, agreeing with the behavior of the hybridization gap. It is clear at larger B\textsubscript{∥} than 12T the gap was totally gone, presenting an insulator to semimetal transition.

![Figure 4.12 Arrhenius plots for strained-layer InAs/InGaSb QWs (black squares), and unstrained 12.5 nm InAs/10 nm GaSb QWs (red cycles). Energy gaps can be deduced by fitting G\textasciitilde exp(−Δ/2kB\textsubscript{T}), as shown by straight dash lines in the plot. Data with black squares were taken by Tingxin Li.](image)

Figure 4.12 showed the Arrhenius plots of a Corbino device made by strained-layer InAs/InGaSb QWs and another Corbino device made by InAs/GaSb QWs. The normal InAs/GaSb QWs showed an insulating gap ~ 26K at lower temperatures, which was associated with the thermal activated behavior of exciton insulator (Chapter 6). In the strained-layer InAs/InGaSb QWs, since the Fermi surfaces of the electron and hole bands were perfectly matched, the hybridization was greatly strengthened with the exciton effect the minor. Although the bulk conductance was suppressed at the low temperature for InAs/InGaSb QWs, it lacked
the exponential dependence characterizing exciton insulator. This explanation was further confirmed by the activated behavior at higher temperature, where hybridization gap energy could be roughly estimated, which was ~ 130 K for InAs/InGaSb QWs and much larger than 66 K for InAs/GaSb QWs. Overall, a larger hybridization gap has been achieved by strain-engineering, in the reasonable agreement with the calculations shown in Figure 4.6. It should be mentioned that it was the first realization of insulating hybridization gap, since this structure was invented. Also the observed gap was the largest bulk gap in the known QSH systems, which would promote the realization of the promising room temperature TI.

**4.4. Long coherent helical edge**

In previous works, due to the conductive bulk, the existence of helical edge modes was deduced by subtracting the bulk conductance\(^{(12)}\). Corbino experiment under \(B_{//}\) gave the direct observation to the insulating hybridization gap and paved a solid way to demonstrate the TRS QSH effect. To investigate helical edge property at the boundary of devices, several mesoscopic devices were fabricated with the size of \(1 \times 2 \mu\text{m} \ \pi \text{bar}, 5 \times 10 \mu\text{m} \ \text{Hall bar}, 25 \times 50 \mu\text{m} \ \text{Hall bar}, \) and \(50 \times 100 \mu\text{m} \ \text{Hall bar}.\)
Figure 4.13 Longitudinal resistance vs front gate voltage in a 1× 2μm π bar at 300mK.

Figure 4.14 Longitudinal resistance vs front gate voltage in a 5× 10μm Hall bar at 300mK.
Figure 4.15 Longitudinal resistance vs front gate voltage in a $50 \times 100\mu m$ Hall bar at 300mK.

Figure 4.16 Longitudinal resistance ratioed by the predicted quantized resistance in $1 \times 2\mu m$, $5 \times 10\mu m$, $25 \times 50\mu m$ and $50 \times 100\mu m$ Hall bars.
For a 1x2μm π bar device, it was initially in the electron dominating regime. Then as the front gate swept from the electron regime to the hole regime, at the CNP where hybridization gap dominated, the resistance increased to a quantized value of 6.45kΩ as demonstrated in Figure 4.13. This resistance agreed with the theoretical value for helical edge modes. Then we further measured 5x10μm Hall bar device, as shown in Figure 4.14, and the peak resistance at the CNP was close to the quantized value of 12.9 kΩ. The measurement results in these devices agreed with those obtained with LB formula. It demonstrated if one added two contacts outside of local measuring contacts the resistance would change to corresponding quantized value. This nonlocal measurement confirmed the existence of the QSH effect in inverted InAs/InGaSb QWs.

In order to explore the coherence length of helical edges in this system, I measured 25x50μm and 50x100μm Hall bar devices and found the peak resistance deviated from the predicted quantized value. For 50x100μm Hall bar device, $R_{xx}$ resistance at the CNP was nearly 100 kΩ(Figure 4.15). As shown in Corbino measurement, the bulk gap had more than 10MΩ so this resistance came from the edge states. This result also indicated the coherence length of nearly 12μm. This coherence length agreed with results(Figure 4.13 and Figure 4.14) in mesoscopic devices where the coherence between neighboring contacts was kept within 10μm distance.

As shown in Figure 4.16, I collected all CNP resistance from devices with different lengths and normalized them to the quantized value of corresponding
geometries. As the length decreased from 100μm, the resistance decreased linearly. When the length was below 12μm, the resistance kept constant to the quantized value. This result agreed with the meaning of topological helical edge and exclusively confirmed the existence of nontrivial edges in InAs/InGaSb QWs. By contrast, the trivial edge should be linear with the length as shown in (68). Hence the quantized saturation behavior distinguished the observed helical edge state from reported trivial edges in InAs/GaSb QWs(68).

It showed the coherence could be further preserved in a longer length than 10μm which was much longer than previous reported one (2μm) in this system. This is also the longest coherent helical edge in known QSH systems, which was essential for the application in the microengineering. This enhancement of coherence length can be attributed to the large hybridization gap in InAs/InGaSb QWs, which would be validated in the in-situ edge-bulk gap modulation. We could expect longer and even macroscopic coherence length if the gap further increased to stand for room temperature.
Theoretically, the large bulk gap should be helpful for the thermal stability of helical edges. For 1×2μm π bar device, as shown in Figure 4.17, the longitudinal resistance trace presented the predicted quantized value at the CNP. Then I increased the temperature up to 37K and found the CNP resistant always near quantized value as presented in Figure 4.17. In inverted InAs/GaSb QWs with 4meV bulk gap, the quantized value disappeared at 4K. Here this thermal stability was stronger than that in InAs/GaSb QWs, due to larger hybridization gap. If hybridization gap was further strengthened, 77K and even room temperature helical edge mode could be expected, which was important for both fundamental and practical interests.
4.5. In-situ electrostatic manipulation of helical edge

In the strained-layer InAs/InGaSb QWs, hybridization gap became much larger with longer helical edge. Additionally, double gates gave us the opportunity to turn the band inversion and edge velocity. The exploration of relation between the gap energy and the coherence length gave us an opportunity to perform the effective manipulation to the single mode edge.

![Graph](image)

Figure 4.18 Hybridization gap energy calculated by 8-band selfconsistent calculation by Dr. Kai Chang, as a function of InAs layer thickness, InGaSb layer thickness and component.

As shown in Figure 4.18, for InGaSb QWs with different thicknesses, the hybridization gap energy had a nonlinear relation with the thickness of InAs QW. For the parameters of QWs used here, the InGaSb QW thickness was 4nm and the InAs QW thickness was 9.5nm, corresponding to the gap energy of 14meV. For more
inverted band which meant thicker InAs QW or InGaSb QW, the gap energy continued to decrease. This calculation agreed with our experimental observation in Figure 4.10. In our experiments, at \( V_b = 4 \text{V} \) where the band was more inverted than that at \( V_b = 0 \text{V} \), the bulk at the CNP became more conductive, indicating a smaller bulk hybridization gap. Remarkably, according to the CNP density and the gap energy, we obtained the Fermi velocity of the edge \( v_F = \frac{\Delta}{2k_{\text{cross}}} \sim 10^5 \text{m/s} \) that was comparable with the one in HgTe QW\((v_F \sim 5.5 \times 10^5 \text{m/s})\). In experiments, the energy gaps at \( V_b = 4 \text{V} \) and \( V_b = 0 \text{V} \) were still comparable, although the gap at \( V_b = 0 \text{V} \) was a little larger. This was also supported by the calculation as shown in Figure 4.18, where the gap energy only decreased by 10\% if InAs QW thickness increased by 2nm. Similarly, we can approximately get the Fermi velocity of the edge state \( v_F \sim 8 \times 10^4 \text{m/s} \), at \( V_b = 4 \text{V} \).

![Image](image_url)

**Figure 4.19** \( R_{xx} - V_f \) traces measured from a 10 \( \mu \text{m} \times 5 \mu \text{m} \) Hall bar device at \( T \sim 20 \text{mK} \) with \( V_b = 0 \text{ V}, 1 \text{ V}, 2 \text{ V}, 3 \text{ V}, \) and \( 4 \text{ V} \).
Figure 4.20 $R_{xx}$-$V_f$ traces measured from a 100 μm ×50 μm Hall bar device at $T \sim 20$ mK with $V_b = 0$ V, 1V, 2 V, 3 V, and 4 V. The edge coherence length increases with decreasing $V_b$.

Figure 4.21 The coherence length as a function of the back gate voltage.
In topological physics, one expects a bulk-edge corresponding relation, motivating us to further investigate the helical edge coherence length under different bulk gap energies. We now turned to the helical edge properties of strained-layer InAs/InGaSb QWs, showing not only the band inversion but also associated helical edges can be tuned effectively. To perform this experiment, I used 5x10μm and 50x100μm Hall bar devices. Figure 4.19 and Figure 4.20 showed the longitudinal resistance $R_{xx}$ vs $V_f$ traces of two Hall bar devices at different back gate voltages, respectively. The measured $R_{xx}$ can be considered from helical edges, since the bulk was fully insulating at such low T. At $V_b = 0$ V, the resistance peak was about 115 kΩ in 50x100μm Hall bar, corresponding to a coherence length of 12μm; while the resistance peak was quantized in 5x10μm Hall bar. Then I increased $V_b$ and found for both devices the resistance gradually increased, meaning shorter coherence length. On the other hand, the trivial edge was from the unclean fabrication and would not change with gating, so the increased $R_{xx}$ with higher $V_b$ confirmed again what we observed was not trivial edge but nontrivial edge that had the correspondence with the bulk topological gap. It also can be seen from Figure 4.21 that the coherence length decreased from ~ 12μm at $V_b = 0$ V to ~ 6.5μm at $V_b = 4$ V. Thus the effective tuning of helical edge mode and the in-situ electrostatic manipulation to helical edge were realized for the first time. A plausible explanation for above data was related to the interaction effects in the helical edges (69–71). At higher $V_b$, the bulk band became more inverted, with hence a larger $k_{cross}$ and a slowly changing $\Delta$. Overall, this would lead to a smaller Fermi velocity of the helical
edges, resulting in more prominent interaction effects for edge electrons. Under $V_b=0\text{V}$, the edge scattering time $\tau = \frac{\lambda}{v_F} = \frac{2\lambda k_{cross}}{\Delta} \sim 100\text{ps}$; while under $V_b=4\text{V}$, the scattering time was nearly 70ps. Less edge scattering time meant there were more backscattering within a unit time. The backscattering possibility increased when the electron-electron interactions were stronger, and thus the helical edge states exhibited a shorter coherence length with shorter scattering time in more inverted band. This experiment result demonstrated larger bulk gap had longer coherence length, confirming the result we got in Section 4.4.
Figure 4.22 $G-V_f$ traces under 0 T, 1 T, 2 T, 3 T, and 4 T perpendicular magnetic field at upper panel with $V_b = 0$ V and at lower panel with $V_b = 4$ V
Figure 4.23 $R_{xx}$-$V_f$ traces of the 100×50 μm Hall bar under different perpendicular magnetic fields at $V_b = 0V$.

Figure 4.24 $R_{xx}$-$V_f$ traces of the 100×50 μm Hall bar under different perpendicular magnetic fields at $V_b = 4V$. 
I have demonstrated hybridization gap and associated helical edge modes. To examine the TRS property of helical edges, magnetic field dependent experiments were required. In general, applied magnetic field broke the TRS and opened a gap in helical edges. Under the magnetic field, dissipationless transport of 1D massless Dirac fermion could be tuned into the dissipative transport of 1D massive fermion and thereby the helical edge resistance increased.

In previous results, the bulk was conductive under 0T such that the helical edge was shunted by the bulk conductance. Then magnetic field not only may break TRS but also drove the bulk more insulating. Both factors increased longitudinal resistance $R_{xx}$, hence it was hard to ascribe increased $R_{xx}$ to the broken TRS.
To exclude possible bulk contribution, I applied perpendicular magnetic field to Corbino device under the low and high CNP densities. As shown in Figure 4.22, under magnetic field the bulk still kept insulating. The insulating bulk made the investigation of broken TRS possible. Then I applied magnetic field to the Hall bar devices. Remarkably, for all devices made by strained-layer InAs/InGaSb QWs, the helical edge conductance showed clear magnetic field dependence. As shown in Figure 4.23 and Figure 4.24, for the 50×100 μm Hall bar devices with the low and high CNP densities, the peak resistance increased with the magnetic field, agreeing with magneto transport result in HgTe QW(7). Helical edge states were only protected against non-magnetic back-scattering, while magnetic scattering broke down helical edge and counterpropagating spin-up and spin-down channels equilibrated. These results confirmed that the helical edge was protected by the TRS. If we extracted the coherence length from 50x100 μm Hall bar at the low and high densities, as shown in Figure 4.25, it was clear that the coherence length decreased in the presence of magnetic field. The tuning of the coherence length can be understood as the control of the edge gap opening in the 1D massive fermion. The observations exclusively provided the evidence for the existence of the TRS QSH effect in the hybridized InAs/GaSb system, for the first time. Remarkably, under perpendicular magnetic field larger than 4T, for both macroscopic and mesoscopic devices, the peak/plateau resistance started to decrease, indicating the edge states undergoing a transition from TRS broken helical edge to chiral like edge. More results about this transition can be found in section 5.5.
It is interesting to compare electrostatic control and magnetic control. In the way of electrostatic control, the band inversion and gap energy was changed to modulate the coherence length, while in the way of magnetic control, the edge coherence length was manipulated through breaking TRS under magnetic field. Under broken TRS, a small gap opened in helical edge modes, making the edge transport dissipative with the resistance deviating away from the initial value. As a further study, one could introduce the nanostructure of magnet insulator near the helical edge to locally open a trivial gap at edge states.

4.7. Conclusion

Compared with previous works, by strain-engineering, the insulating hybridization gap was realized, for the first time. This gap was the largest bulk gap in the known QSH systems. Also the longest coherent helical edge (13μm) was realized. I showed through enlarging the bulk gap the coherent helical edge became longer and more stable against the temperature. Here TRS protected QSH insulator was observed for the first time in InAs/GaSb systems, which confirmed the observation in HgTe QW(7). Moreover for the first time, the in-situ electrostatic manipulation of helical edge modes was demonstrated. Further it was confirmed that the edge states can be gapped out by applying magnetic fields to break TRS. Two ways to manipulate the helical edge were presented, and the way to realize room temperature QSH insulator was experimentally demonstrated, with one step
closer to the device and circuit applications of QSH insulator based on semiconductor technology.
Chapter 5

Time Reversal Symmetry broken Quantum Spin Hall effect

As presented in Chapter 4, the TRS symmetry QSH effect has been reported in the InAs/InGaSb QWs, where the band was deeply inverted. When the band was less inverted, the bulk became more insulating with longer coherence length. If one further turned the band to “the shallowly inverted”, an insulating bulk could be expected in the conventional InAs/GaSb QWs. In this chapter, I would focus on the shallowly inverted band as marked in Figure 5.1.

In the shallowly inverted InAs/GaSb QWs, the gap was observed to become insulating and robust helical edge states with wide conductance plateaus were
precisely quantized to $2e^2/h$ in a broad temperature regime. Here I presented data of robust helical edge states in engineered semiconductor systems that were immune to disordered bulk. However, the scattering-time of the edge states was found unexpectedly long.

![Figure 5.1 Phase diagram for different front and back gate voltages. The red regime represents the inverted band and the blue regime indicates the noninverted band.](image)

On the other hand, the TRS has been widely believed from experimental and theoretical aspects to be a necessary ingredient for the emergence of the QSH insulating phase, commonly characterized via the $Z_2$ topological invariant(4,5). Applying a magnetic field breaks TRS and takes off the topological protection of the helical liquid from backscattering. This argument is supported by strong magnetic
field dependence in the QSH phase in HgTe QW(7) and the strained layer InAs/InGaSb QWs(64).

Remarkably, under the external magnetic field, even in the presence of disorder, the helical edge state was still robust with the quantized plateaus persisting to 10T in-plane field. In a perpendicular field, the broken TRS led to a spatial separation of the movers in the Kramers pair and consequently the intra-pair backscattering phase space vanished, i.e., the conductance increased from $2e^2/h$ in strong fields manifesting chiral edge transport. It is the first observation of TRS broken QSH insulator. Deeper and exotic origin, such as the interplay of exciton condensation and the QSH phase, are now being examined(28).

5.1. Structure detail and wafer characterization

The semiconductor wafers of InAs/GaSb QWs were grown by MBE technique. As shown in Figure 5.2, the wafer structure contained a N+ GaAs (001) substrate, 1 μm thick insulating buffer layer, InAs/GaSb QWs with barriers made of AlSb, and a thin GaSb cap layer. The interface between GaSb and InAs was doped by a sheet of Si, with a concentration of ~ $10^{11}/cm^2$ (Figure 5.3). Due to Si impurity scattering, the carrier mobility measured at 300 mK outside of the CNP showed some reduction. Typical mobility of electrons was $40000cm^2/Vs$ at a density of $5 \times 10^{11}/cm^2$, and the mobility of holes was about one order of magnitude lower at similar densities. Magneto-transport showed well resolved quantum oscillations in Figure 5.4.
Because the edge states are topological in nature, the disorder should have little effect to their existence and transport properties.

**Figure 5.2** InAs/GaSb wafer structure where Si doping sheet is placed at the interface between the GaSb layer and the InAs layer.

**Figure 5.3** A two-dimensional topological insulator is engineered from two common semiconductors, InAs and GaSb, which hosts a robust quantum spin Hall effect.
Show schematically the band structure of a InAs/GaSb bilayer, and the potential fluctuations induced by Si dopants at the interface.

Figure 5.4 Magnetoresistance and Hall resistance traces measured in a Si-doped quantum wells.

5.2. Shallowly inverted band

The devices had the front and back gates with different geometries. With double gate tuning, the band inversion could be regulated from “deeply inverted” to “shallowly inverted”.

It should be noted that in previous experiments(12,26,59) the wafer design accepted the one used in Yang’s experiment(26), where GaSb QW was on top of InAs QW. On the contrary, for the wafer as shown in Figure 5.2, InAs QW was above GaSb QW. In the case that InAs QW was on top, the front gate was close to the electrons with holes partially screened; while in the case that GaSb QW was on top, the front
gate was near the holes. In other words, if InAs QW was on top, its front gate/back
gate had the same effect with the back gate/front gate in the case that GaSb QW was
on top. However, in experiments, the tuning ability of the back gate was limited due
to the Schottky barrier; therefore we always had to rely on the front gate, which
made two growth sequences different.

Let us review the front gate sweeping with different growth sequences.
When the devices were cooled down to the low temperature, regardless of the
growth sequence, the initial situation should be similar, with a medium electron
density, such as $5 \times 10^{11}/cm^2$. If one swept the front gate in GaSb-on-top sequence,
Fermi level touched the top of the hole band at the certain electron density, i.e.$2 \times 10^{11}/cm^2$, with the hole emerging. The holes were in close to the front gate,
screening the electric field, therefore the hole density increased with the gate
voltage and the electron density kept nearly unchanged. In this case, at the CNP the
electron and hole densities became equally high, with the band deeply inverted. On
the other hand, if one swept the front gate in the InAs-on-top sequence, as the holes
emerged, the electrons were in close to the front gate, and due to relatively small
effective mass, the electric field is still partially applied to the holes whose density
increased slowly. At the CNP, the electron and hole densities became equally low
with the band shallowly inverted. Therefore, in order to achieve the shallowly
inverted band, for the GaSb-on-top case, one should sweep the back gate, while for
the InAs-on-top case, one should sweep the front gate. For the wafer used here, even
without the back gate, as the front gate tuned Fermi level to the CNP, the band was
still shallowly inverted. In this case, if one decreased the back gate voltage, holes can be induced initially and the band inversion could be tuned to “deeply inverted”.

Figure 5.5 Left panel and right panel show the back gate voltage and front gate voltage dependence of $R_{xx}$ in a 1 x1μm meso hallbar, respectively. In the left panel, traces under different fixed front gate voltage are taken.

Figure 5.6 $R_{xx}$ vs $V_f$ traces in a 50x50μm macro hallbar for fixed back gate voltages.
Figure 5.7 $R_{xx}$ vs $V_b$ traces in a 50x50μm macro hallbar under zero front gate voltage.

To validate such argument, I fabricated 1x2μm meso Hall bar and 50x50μm macro Hall bar with the presence of both the front and back gates. As shown in Chapter 4, a more conductive bulk corresponded to a more inverted band. Firstly, for the 1x2μm Hall bar, at zero back gate voltage, I swept $V_f$ and found the quantized plateau appeared at the CNP indicating an insulating bulk and shallowly inverted band, as shown in the right panel of Figure 5.5. On the contrary, at $V_f=0V$, I swept $V_b$ and the resistance at the CNP was nearly 6 kΩ, suggesting 10kΩ bulk resistance. Highly conductive bulk could be treated as a result of the deeply inverted band. It should be mentioned that in the InAs-on-top case the conductive bulk under $V_b$ sweeping was close to the reported one under $V_f$ sweeping in the GaSb-on-top case(12). Then, I decreased the fixed $V_f$, meaning lower band inversion. As shown in Figure 5.5, at lower $V_f$, the resistance at the CNP increased and approached to the
quantized value with more insulating bulk, indicating less inverted band. For 50x50μm macro Hall bar, similarly $V_f$ sweeping resulted into a 200 kΩ resistance peak at the CNP (Figure 5.6), while $V_b$ sweeping leaded to a several kΩ peak (Figure 5.7). Then, I decreased the fixed $V_b$ for $V_f$ sweeping, where the band should be more inverted. As shown in Figure 5.6, the peak resistance dropped by half of magnitude as $V_b$ was lowered to -6V.

Thus, with two devices and various gate tuning ways, I have demonstrated that the InAs-on-top growth sequence could result a shallowly inverted band.

5.3. Insulating bulk gap

Figure 5.8 The temperature-dependent conductance traces measured in a Corbino disk are displayed.
Figure 5.9 The Arrhenius plot shows that the conductance vanishes exponentially with T.

To quantitatively investigate the bulk gap at the CNP, Corbino device was fabricated for transport measurement. In this geometry, the edge transport was shunted via concentric contacts, and hence conductance measurements probed bulk properties exclusively. As illustrated in Figure 5.8, in the electron regime, the bulk conductivity was pretty high, and then it dropped quickly with a conductance dip coming to zero at the CNP. The corresponding resistance was 10MΩ, limited by the lockin equipment, with the actual resistance much larger than 10MΩ.

Then as we increased the temperature, the zero conductance dip became narrow and then lifted. Analysis of Arrhenius plot (Figure 5.9) was followed by a standard procedure in quantum transport to deduce the energy gap: 

\[ G \propto e^{-\Delta / 2k_B T} \]

where \( \Delta \) was the energy required to create a pair of electron-hole over the gap, and \( k_B \) was the Boltzmann constant. Transverse conductance suppressed to zero in the
gap, showed the exponentially activated temperature dependence and allowed the direct extraction of gap values. At higher T, the gap value $\Delta_{\text{mini}} \sim 66$K was deduced, consistent with a hybridization-induced mini-gap. As the temperature was further reduced below $\sim 10$K, the conductance continued to vanish exponentially with a different slope, indicating opening of a hard gap $\Delta$ in the energy spectrum (it actually is an exciton insulator gap and will be discussed detailed in the next chapter). In consequence, at temperatures on the order of 1K and below, the bulk was completely insulating and the transport occurred only along the edge.

5.4. Quantized plateau

![Graph showing conductance plateaus](image)

Figure 5.10 Wide conductance plateaus quantized to $2e^2/h$ and $4e^2/h$, respectively for two device configurations shown in inset; both have length 2 \(\mu\)m and width 1 \(\mu\)m. (B) Plateau persists to 4K, and conductance increase at higher temperature.
As I have presented, the bulk gap became insulating in the shallowly inverted InAs/GaSb QWs. The next step was to investigate the potential helical edge mode in this system. The difference between the nontrivial helical edge and the trivial wire was that only the former had the nonlocal signature. In the exploration of the QSH effect in HgTe QW, nonlocal experiments confirmed the nontrivial feature of helical edge modes. On the other hand, the evidence for quantized plateau in the QSH effect was still lacking. In the context of the QH physics, precisely quantized conductance plateaus (to multiples of $e^2/h$) was the smoking gun evidence for chiral edge states.

Here, in order to observe quantum plateau in the nonlocal measurements, I prepared two mesoscopic devices: one is 1x2μm π bar and the other is 1x2μm Hall bar. The difference between these devices was that there were additional two nonlocal contacts in the Hall bar. For the trivial edge, there should not be any difference for two geometries, while for the helical edge, they should present $4e^2/h$ and $2e^2/h$ conductance plateau for π bar and Hall bar, respectively.

As the Fermi energy was tuned into the hard gap via the front gate, longitudinal conductance measurements in π bar configuration as well as that in Hall bar geometry, revealed wide plateaus near perfectly quantized to $4e^2/h$ and $2e^2/h$, respectively (Figure 5.10), as expected for non-local transport in helical edge channels(7,11), based on LB analysis(46). This experimental result confirmed nonlocal contacts would affect the local measurement. In the π bar geometry, at the boundary without contacts, two counterpropagating helical edge states connected the chemical potential from the source and drain, respectively, and they were not in
equilibrium with each other because the elastic backscattering vanished between two channels. Then in the Hall bar geometry, two probe contacts necessarily leaded to the equilibration of two helical channels with the opposite spin orientations, because the voltage probes were not spin sensitive. Note that the conductance value here was quantized to better than one percent - unprecedented by any other known topologically ordered system other than IQHE and FQHE, indicating a high degree of topological protection. Due to the existence of Si doping in the bulk, it was convincing that the helical edge in InAs/GaSb QWs was robust against non-magnetic disorder scattering.

Nonlocal measurement was also performed in the meso H bar device as shown Figure 5.11(left) at 300mK. The current went from contact 1 to contact 2 while the voltage was measured between contacts 3 and 4. For the H bar, at the electron dominating regime, the voltage difference between contacts 3 and 4 tended to zero. But according to Poisson’s equation, diffuse transport gave a measured resistance less than 100Ω. At the gap regime, helical edge states would appear as demonstrated in Figure 5.11(right) and the measured voltage difference became a nonlocal quantization value.
Figure 5.11 (left) Schematic layout of a four-terminal H bar device and (right) Nonlocal four-terminal resistance measured on the H bar device

Figure 5.12 (A) Electrical charge transport in large devices is due to edge channels. (B) The resistance scales linearly with the edge length, indicating a phase coherence length of 4.4 μm; the coherence length is independent of temperature between 20mK and 4K.

Furthermore, as the edge length of Hall bar increased to macroscopic dimension, that longitudinal resistance at the CNP linearly increased with the device
length as shown in Figure 5.12. It has been previously shown (6,10) that helical edge states were protected only against non-magnetic back-scattering, while magnetic or inelastic scattering gave a coherence length $\lambda$ at which edge transport broke down and counterpropagating spin-up and spin-down channels equilibrated. In this case, approximate longitudinal resistance was obtained by series addition of $N~L/\lambda$ half-quantum resistors, giving a total resistance value of $(L/\lambda)\cdot h/2e^2$. This approximation was in excellent agreement with the data presented in Figure 5.12, giving $\lambda = 4.4 \mu m$.

I made a summary of the measured devices with the resistance normalized to the quantized values as shown in Figure 5.13. When the length was longer than 4.4μm, one coherent helical edge was broken into several segments of coherent edges with measured resistance larger than quantized resistance. As the length became shorter, the longitudinal resistance decreased until the length hit the coherence length. Then even if the length was further shorter, the longitudinal resistance kept constant to quantized value. This behavior made the measured edge state distinct from the trivial wire where the measured resistance should be always linear with length.
Remarkably, the edge scattering time, i.e., $\tau = \frac{\lambda}{v_F} = \frac{2\lambda k_{\text{cross}}}{\Delta} \approx 200\,\text{ps}$ (approaching that of the highest-mobility 2DEG in GaAs)\cite{72}, appeared to be extremely long regardless of the disordered bulk; here the Fermi velocity of the edge state $v_F \approx 1.5 \times 10^4\,\text{m/s}$ was much smaller than that of 2DEG or HgTe QW ($v_F \approx 5.5 \times 10^5\,\text{m/s}$)\cite{7} due to the fact that the gap opened at a finite wavevector instead of the zone center. In addition, the quantized plateau and the coherent length were found to be independent with temperature between 20 mK and 4K as shown in Figure 5.11 and Figure 5.13. In Ref. (73), it was proposed that the coupling between helical edge states and potential quantum dots in the bulk could be responsible for this temperature independence. Later, Ref. (28) further pointed out

Figure 5.13 Longitudinal resistance ratioed by the predicted quantized resistance in $1 \times 1\,\mu\text{m}$ junction, $1 \times 1\,\mu\text{m}$ π-bar, $1 \times 2\,\mu\text{m}$ π-bar, $1 \times 2\,\mu\text{m}$, $5 \times 10\,\mu\text{m}$, and $10 \times 20\,\mu\text{m}$ Hall bars.
that in the phase transition from nontrivial hybridization gap to trivial exciton condensation there existed an area where the helical edge state and the bulk exciton condensation interplayed; the edge robustness against temperature could be attributed to the interaction. In InAs/GaSb system, the trivial exciton condensation corresponded to the noninverted band, while hybridization gap corresponded to the deeply inverted regime. Thus, the transition area just indicated the shallowly inverted band, agreeing with the experimental parameter. In the next chapter, I will show the evidence for exciton insulator in this system.

5.5. Time reversal symmetry broken helical edge state

Figure 5.14 The conductance measured in Corbino disk at $T = 300 \text{ mK}$ are shown, respectively, for magnetic field applied in the plane in upper panel, or perpendicular...
to the plane in lower panel. In either case, there is no evidence for gap closing at increasing magnetic field; a continuous magnetic field sweep shows that 2D bulk is always completely insulating from 0 to 8T.

![Graph showing gap energy vs. magnetic field](image)

**Figure 5.15** The hard-gap energy is shown to increase with applied perpendicular magnetic field

Magnetic field was a standard tool to verify the TRS protection in the QSH insulator, just as shown in section 4.6. Here the edge transport properties under TRS broken by magnetic fields along each major axis of the device, were examined up to 12T. To ensure the measured signal under the magnetic field was from the edge, I applied the magnetic field to Corbino device. As shown in Figure 5.14, under the inplane magnetic field up to 8T, the bulk conductance at the CNP always kept zero while under the perpendicular field up to 6T, the bulk gap became more insulating with the raised gap energy as shown in Figure 5.15. In this case, under the magnetic field, any measured conductance signal was from the edge contribution. Moreover there should not be the bulk topology transition under both inplane
magnetic field and perpendicular field, since the bulk gap kept opened. According to the topological theory, once the bulk had a transition to different topology, the bulk gap would close and open again. One example was the IQH effect where the LL gap closed and reopened as the Fermi level went across two neighboring LLs. Another familiar example was just the TRS QSH effect where the bulk has to close and reopen with 1D Dirac fermion in the interface as the Fermi level went across the boundary connecting the nontrivial bulk characterized via the $\mathbb{Z}_2$ topological invariant and the trivial vacuum.

![Graph showing transport characteristics in strong magnetic fields](image)

**Figure 5.16** The edge helical liquid in the InAs/GaSb bilayer retains its transport characteristics in strong external magnetic fields, here examined up to 12T. Upper panel shows plateau values measured for four different
devices with in-plane magnetic field applied parallel (open circles) or perpendicular (open triangles) to the edge axis. Lower panel shows the same four samples were measured (T = 300 mK) in a field applied perpendicular to the 2D plane, with the three Hall bar devices showing increasing conductance, and the two-terminal device (blue squares) showing decreasing conductance. The device sizes are noted with “2μm” for 1x2μm π-bar, “1μm” for 1x1μm two-terminal device, “10μm” and “20μm” for 5x10μm and 10x20μm hall bar.

As presented in Figure 5.16, under in-plane magnetic fields applied along and perpendicular to the current flow, the conductance plateau value remained quantized for mesoscopic samples, as well as stayed constant for longer devices, for fields even close to 10T. We could look in close the detail for a meso π bar device up to 10T as shown in Figure 5.17. As far as the edge conductance was concerned, this could also be interpreted as a lacking of evidence for the gap opening in the edge spectrum across the gap. According to the topological theory, except only one specific direction, the magnetic field of other directions could break the TRS of edge states. However, now magnetic fields of both directions cannot destroy the helical edge state, indicating this system had protection other than TRS and differed from the TRS QSH insulator.
Figure 5.17 Shows the plateaus measured from the $\pi$-bar device (shown in (A). red open circles) at 20mK, at different applied in-plane fields parallel to the edge.

Finally, I examined the same four samples in a field applied perpendicular to the 2D plane(Figure 5.16). For three Hall bar devices, different from the results under inplane magnetic field, both the quantized and non-quantized conductance...
increased. Considering the bulk was still insulating, we knew that the additional conductance had to come from the edge state. This result seemed against the theoretical prediction about helical edge with LB formula, where the quantized value had already been the minimal conductance. In LB formula,

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

\(T_{pq}\) was the transmission possibility from contact \(q\) to \(p\), which was treated as 100% for the neighboring contacts. Now the conductance was smaller than the quantized value, indicating the transmission possibility along one direction became less than 100%. Moreover this transmission possibility continued to decrease under higher magnetic field. It should be mentioned that the helical edge still existed, which was demonstrated by the bulk gap opening. The decreased transmission possibility could be understood in the below picture: due to the Lorentz force, the magnetic field would push the edge modes of one chirality (say, right) outward and the opposite chirality inward (Figure 5.18), and thus the conductance measured by edge contacts should weight more on the right chirality.

To validate this idea, the same experiments were performed in the two-terminal device. According to LB formula, if the transmission possibility along one direction decreased, the two terminal resistance signal that was dominated by Hall resistance, increased with the field, instead of decreasing. With an opposite conductance variation, such an experiment could also exclude trivial material issue, such as the local conductive bulk. It should be noted that the long edge would be
broken into several segments under zero field. Now broken TRS leaded to a spatial separation of the movers in the Kramers pairs and consequently the inter-pair backscattering vanished, which indeed helped to reserve the coherence of the helical edge mode.

As a summary, under the magnetic field where the TRS was explicitly broken, the helical edge state still persisted. Under the perpendicular field, two counterpropagating helical edges state were separated in the real space with the edge along one direction shrinking into the bulk and the other edge expanding. As a result, the voltage contacts would only partially measure the expanding edge. Nevertheless, the helical edge state presented here was immune to the broken TRS, against to the TRS QSH effect in deeply inverted InAs/InGaSb QWs. It should be mentioned that the observed TRS broken QSH insulator was out of the scope of the single particle topological theory and asked for deeper understanding and additional insight regarding to the interaction.

5.6. Conclusion

Quantized plateau of the QSH state to a remarkable degree of accuracy was observed for the first time. Robust helical edge states in engineered semiconductor systems were immune to disordered bulk, as well as temperature and perturbations from external magnetic fields which broke TRS. The study presented the first observation of TRS broken QSH insulator, which was not predicted by the single particle topological theory. These unexpected observations were related with
deeper and exotic origin, such as interplay of exciton condensation and QSH phase, which would be discussed in the next chapter.
Chapter 6

BCS like Exciton condensation

6.1. Motivation

In previous two chapters, through alloying GaSb with InSb, the TRS QSH effect was observed in the hybridization gap of the deeply inverted band; in the shallowly inverted band, the TRS broken QSH effect was observed, which was extremely robust against the temperature and magnetic field. The first case validated the great success of recent developed single particle topological band theory, while the second case couldn’t be interpreted with the existing topological
theory. Later, theoretically it was proposed that such unexpected properties were attributed to the interplay between the QSH effect and the bulk exciton insulator.

Figure 6.1 The semiconductor to semimetal transition and exciton insulator is expected to exist in the diluted semimetal

Figure 6.2 Device layout and density measurement under double-gate control. Sketch of device layout. Holes are in GaSb layer (green) while electrons are in InAs layer (red). Front and back gate (blue) are fabricated to tune carrier densities.
On the other hand, exciton insulator, or BCS-like exciton condensation was proposed several decades earlier\(^{(38,74,75)}\) to be an equilibrium ground state in certain dilute semimetals. Near the semimetal to semiconductor transition, diluted semimetal has the instability against the formation of exciton ground state with the negative exciton energy\((E_g - E_b, \text{ where } E_g \text{ is negative and } E_b \text{ is positive})\), giving a spontaneously gapped spectrum without optical pumping. However, the conclusive evidence for its existence was still lacking. In 1980s, the broken gap type-II heterostructure (inverted InAs/GaSb QWs), was proposed, with a negative band gap, which allowed the coexistence of spatially-separated electrons and holes and offered a natural setting for exciton insulator\((\text{Figure 6.2})\). The major challenge to observe exciton insulator in this system was to achieve the low CNP density, or shallowly inverted band, where electrons and holes were bound to form excitons instead of collective e-h plasma\((\text{Figure 6.3})\). In the shallow band inversion, intralayer particle distance was larger than the layer separation, where interlayer Coulomb interaction should be prevented from being screened; On the other hand, the light hole component became the minor, such that Fermi surfaces of electron and hole bands at the CNP were well-matched. Also the interlayer tunneling, which was the mechanism for hybridization gap and mixed bands thereby reducing binding energy, could be neglected.
Figure 6.3 Phase diagram of exciton with temperature and exciton density.

Thus, motivated by recent progresses in the QSH effect and the long-term search for exciton ground state in inverted InAs/GaSb QWs, the bulk gap property in the shallowly inverted band were investigated from transport and optical experiments.

6.2. Tuning into dilute limit of carrier densities

Figure 6.4 Two wafers used in works related in this chapter. The wafer in the left panel has Si doping between the interface of quantum wells.
Figure 6.5 CNP density $n_o$ as a function of $V_f$ and $V_b$ in units of $10^{10}/\text{cm}^2$. The density is obtained through the method introduced in Section 3.4.

The devices were made from inverted InAs/GaSb QWs with/without Si doping (Figure 6.4); two wafers yielded similar results. Both wafers had the InAs-on-top growth sequence, so the front gate could tune the Fermi level to the CNP with the shallowly inverted band. Then the back gate controlled the hole density and realized the band transition from “shallowly inverted” to “deeply inverted”. With the front and back gates, the CNP density could be gated continuously in the range of $5 - 9 \times 10^{10}/\text{cm}^2$ (Figure 6.5). In the shallowly inverted band, the intralayer particle distance $L(\sim 45 \text{nm})$ was much larger than the layer separation $d(\sim 11 \text{nm})$. Hence Coulomb interaction which was less screened by neighboring carriers has to be considered. As shown in Figure 5.6, longitudinal resistance vs $V_f$ at different $V_b$ in a 50x50$\mu$m Hall bar was presented. Under $V_b=0\text{V}$, the bulk at the CNP became
insulating and helical edge modes appeared. Since the sample edge length was much larger than the coherence length, helical edge state had the contribution of 200 kΩ. At smaller $V_b$, more holes were induced and the bulk became more conductive, which had a parallel resistance and lowered the peak resistance. Thus smaller $V_b$ indicates more inverted band structure and higher equilibrium density. Here our studies focused on the highest and lowest equilibrium density in Figure 6.5, corresponding to $V_b=-6V$ and $V_b=0V$.

Figure 6.6 (A) and (B) are $B/eR_{xy}$ vs $\Delta V_f$ traces of the asymmetric 50μm x 50μm Hall bar for $V_b=-6V$ and 0V, respectively. Data are taken at 300mk with 1T perpendicular magnetic field. The inset in (A) is a schematic of the asymmetric Hall bar. The region in the dashed box is covered by front gate. Insets in (A) and (B) are showing band alignments corresponding to the deeply- and shallowly-
inverted regime, or dense and low \( n_o \), respectively. The red regime I is the electron dominating regime. The blue regime IV is the hole dominating regime. The green regime II is the electron-hole coexisting regime. The light green regime III is the soft gap. The dotted line means residual electron and hole filling in hybridization gap so there is no hard gap observed. The gold regime V is the hard gap without electrons or holes.

Figure 6.6 showed \( B/eR_{xy} \) versus \( \Delta V_f \) for a Hall bar device at \( V_b = -6 \text{V} \) (deeply inverted) and 0V (shallowly inverted), respectively; \( B \) was the perpendicular magnetic field and \( \Delta V_f \) was the bias increment from the CNP. At \( V_b = -6 \text{V} \), at high electron density (regime I), \( B/eR_{xy} \) was consistent with the density obtained from SdH oscillations, and linear with \( \Delta V_f \). As \( \Delta V_f \) decreased and the top of hole band was reached by the Fermi level, two-carrier transport dominated with \( R_{xy} \) traces divergent and the system reached e-h hybridized regime (regime II). Because of high DOS in hybridization gap, \( B/eR_{xy} \) was still dominated by electron-hole residue carriers even in the gap (regime III). As the Fermi level was below the bottom of the electron band, holes dominated and \( R_{xy} \) was recovered to be linear again (regime IV). At \( V_b = 0 \text{V} \) where less holes were introduced by \( V_b \), similarly the system was initially dominated by the electrons with the Fermi level higher than the top of the hole band. As the CNP was approached, the system went from regime I to regime II with the trace divergent, indicating the hole appeared and that two-carrier transport appeared. However, in the regime II, with the divergent behavior disappearing, a new plateau-like feature (regime V) in \( B/eR_{xy} \) was observed, and could be attributed to the formation of a hard gap which was absent in single-particle model.
Figure 6.7 Sketch of the Corbino measurement setup. The outer conductor was grounded, and a voltage was applied in the inner conductor; we measured the current flowing across the ring. $V_f$ and $V_b$ were applied to modulate the carriers in the ring.

To further explore the bulk at the CNP, I measured the bulk conductance $\sigma_{xx}$ in a double-gate Corbino device C1 (Figure 6.7). Since edge states were shunted by concentric metallic contacts, $\sigma_{xx}$ exclusively measured bulk properties. For $n_o = 9 \times 10^{10} / cm^2$ and 0T there was a $\sigma_{xx}$ dip (red trace) around the CNP that corresponded to hybridization gap (Figure 6.8). In subsequent panels (left to right corresponding to decreasing $n_o$), the $\sigma_{xx}$ dip decreased to zero; the width of zero conductance around CNP increased. These data obtained at zero B have shown "soft gap" (i.e., finite DOS in the gap) to "hard gap" transition. It could be confirmed quantitatively via thermal activation energy measurements. In the low density case, $\sigma_{xx}$ vs. $1/T$ can be fitted over two order of magnitude into an Arrhenius plot $\sigma_{xx} \propto \exp(-\Delta/2k_B T)$, giving $\Delta/2 = 13K$ (Figure 5.9).
Figure 6.8 $\Delta V_f$ dependence of the conductance $\sigma_{xx}$ for device C1 from $V_b = -6$ to $0\,\text{V}$, with a decrement of $1.5\,\text{V}$. $n_o$ is given in units of $10^{10}/\text{cm}^2$. The red lines correspond to $0\,\text{T}$. At $V_b = -6\,\text{V}$, the system forms semimetal under $35\,\text{T}$, which is consistent with hybridization origin. As $|V_b|$ decreases, $n_o$ decreases and an insulator emerges.

6.3. Origin of hard gap

Within the single-particle hybridization picture, electrons and holes with the same Fermi momentum tunneled between two QWs, forming a hybridization gap\cite{26,59}. Assuming that $B_{\parallel}$ applied along $y$ axis in the plane, Lorenz force gave tunneling carriers additional momentum along $x$ axis, resulting in a relative shift of band dispersions $k_x = -eB\langle z\rangle/h$, (tunneling distance $\langle z\rangle$ is limited by one-half thickness of QW). Consequently, the interlayer tunneling was suppressed due to the momentum mismatch. When $B_{\parallel}$ was large enough, two bands were separated in the
momentum space and hybridization gap was eliminated with the semimetal emerging.

Figure 6.9 $R_{xx}$ vs $V_f$ traces of 50x50µm macrohallbar in 300mK, as a function of fixed backgate voltages and the inplane magnetic field. The blue traces represent the transport under zero magnetic field, while the red traces show the transport under 8T inplane magnetic field.

To delineate the origin of the soft and hard gaps, first I applied 8T $B_\parallel$ to the 50x50µm macroscopic Hall bar as shown in Figure 5.6, under different $V_b$. At $V_b=0V$, the longitudinal resistance(red trace) near the CNP under $B_\parallel$ kept the same with the one(blue trace) without magnetic field. Then as the band became more inverted, the red trace started to deviate from the blue trace. At $V_b=-6V$, $B_\parallel$ reduced the peak resistance by 20%, demonstrating two bands had a relative momentum shift. That the peak still existed was because $B_\parallel$ was not large enough to separate two bands. As
shown by 8-band self-consistent calculation using our device parameters (Figure 6.10), as soon as $B_\parallel$ increased beyond 18T, two bands were separated in the momentum space and the system became an indirect double-layer semimetal. 35T $B_\parallel$ was large enough to separate two bands in the momentum space. In this case, 35T $B_\parallel$ effectively created additional potential barrier for tunneling, which eliminated hybridization gap and was referred to as "magnetic barrier". This magnetic barrier with zero width avoided the decrease of exciton binding energy and provided an ideal platform to explore exciton insulator.

![Figure 6.10](image)

Figure 6.10 Energy dispersion calculated from the 8-band self-consistent model for tunneling electrons and holes for $B_\parallel=0$, 9, 18 and 35T, respectively. The data is calculated by Dr. Kai Chang.
It should be mentioned that here the carrier tunneling distance we used was 10nm, which was an average distance for electrons and holes. In fact, even if $B_\parallel$ was larger than 18T, carriers still could tunnel within a smaller distance. If we further increased $B_\parallel$, the tunneling only with distance larger than $k(z) \times 18T / B_\parallel$ was forbidden. Under 35T $B_\parallel$, the tunneling allowed area still existed and was spatially limited within about 2.5nm away from interface, but these carriers only took a minor proportion so could not form a hybridization gap.

Figure 6.11 Tunneling of electrons and holes in the real space. The arrow indicates electrons and holes taking part in the tunneling and the shaddow regimes show the average tunneling distance.
Figure 6.12 $\Delta V_f$ dependence of the conductance $\sigma_{xx}$ for device C1 from $V_b = -6$ to 0V, with a decrement of 1.5V. $n_o$ is given in units of $10^{10}/\text{cm}^2$. The blue lines correspond to $B_\parallel=35\text{T}$; the red lines correspond to 0T.

With this picture in mind, I applied an in-plane magnetic field $B_\parallel$ to the same Corbino device and measured the $\sigma_{xx}$ as a function of $B_\parallel$ in Figure 6.12. The response of $\sigma_{xx}$ to $B_\parallel$ was dramatically depended on $n_o$. For $n_o = 9 \times 10^{10}/\text{cm}^2$, $\sigma_{xx}$ increased from the $B_\parallel = 0$ value by 4 fold. For $n_o = 5.6 \times 10^{10}/\text{cm}^2$, $\sigma_{xx}$ remained zero but its width decreased. Finally, for the lowest density $n_o = 5 \times 10^{10}/\text{cm}^2$, $\sigma_{xx}$ was characteristically the same as that of $B = 0$. Together with the $\sigma_{xx}$ results at zero field (red), these observations could be interpreted as what follows: the departure of $\sigma_{xx}$ (35T) from $\sigma_{xx}$ (0T) was a qualitative measure of how much
contributions from hybridization to the gap formation, using $n_o$ as a tuning parameter. It was naturally inferred that at the highest density the hybridization dominated the bulk at the CNP while at the lowest density hybridization did not play much role on the hard gap formation. On the other hand, more evidences later would show that spontaneous exciton binding emerged as a leading mechanism for forming the gap and hence exciton insulator became the ground state in the shallowly inverted double QWs.

In addition, the $B_{||}$ results in Figure 6.12 also confirmed the hard gap at $V_b=0V$ was in the inverted band. At $V_b=-4.5V$, the lift of conductance dip at the CNP after applying 35T showed hybridization gap still existed, while the CNP conductance dip under 35T presented the hard gap had gradually came out. Then at higher $V_b$, the different between conductance traces under $B_{||}=0$T and 35T confirmed hybridization gap always existed but tuned weaker, and the hard gap became stronger until it dominated the bulk under $V_b=0V$. Thus, the hard gap and hybridization gap coexisted in a large range of $V_b$. Hybridization gap had to exist in the inverted band structure, demonstrating the hard gap was also formed in the inverted band. As a contrast, in the inverted band to noninverted band phase transition, the vanishing of hybridization gap occurred before the appearance of semiconductor gap and there was not any overlap between two gaps.
Figure 6.13 $\Delta V_f$ dependence of the conductance $\sigma_{xx}$ for device C1 in shallowly inverted band and $B_{||}$ ranging from 0T to 35T. Traces were obtained at 30mK. A region of broad zero conductance can be seen from 0 to 35 T. The color corresponds to $\sigma_{xx}$ in units of Siemens.

From now on, I concentrated on the shallowly inverted band, i.e., $n_o = 5 \times 10^{10}/cm^2$. Figure 6.13 showed a plot of $\sigma_{xx}$ in a series of $\Delta V_f$-sweeps under consecutive and fixed $B_{||}$, where a broad zero conductance (between ±0.5V) could be seen from $B_{||} = 0$ to 35 T. In other words, the gap remained open continuously in spite of strong external magnetic fields. The same result was obtained for Corbino device C2 made from Wafer B. That the gap remained open was further confirmed quantitatively by thermal activation energy measurements. Shown in Figure 6.14 as an example (for device C1 at $B_{||} = 35$T), the $\sigma_{xx}$ vs. $1/T$ could be fitted over two orders of magnitude into an Arrhenius plot $\sigma_{xx} \propto \exp(-\Delta/2kT)$ (Figure 6.15), with the
gap energy $\Delta = 25$K. Following the same procedures, we obtained $\Delta = 27 \pm 1$K for $B_{||} =$ 0, 9, 18, 27, and 35T, with a peak value 28K near 9T, as shown in Figure 6.16.

Figure 6.14 Conductance traces for different temperatures (5.8, 5, 4.5, 3.95, 3.45, 2.95, 2.6 and 2.29K) in shallowly inverted band and under 35T. Higher temperature corresponds to upper trace.
Figure 6.15 Dependence of dip conductance in Corbino measurement on $1/T$ under 35T. Solid line is guide to the eye. The data can be fit with $\sigma_{xx} \propto \exp(-\Delta/2kT)$ to obtain $\Delta$.

Figure 6.16 Gap energy $\Delta$ obtained for a series of $B_{//}(0T, 9T, 18T, 27T$ and 35T) in the low density case.
It was necessary to study the gap of 35T $B_{\parallel}$ under small perpendicular field in the shallowly inverted band. By rotating the device by a small degree, 0.5T, 1T and 2T perpendicular field could be applied to device C1 with $B_{\parallel}$ still close to 35T. As shown in Figure 6.17, the conductance of electrons and holes were sensitive to the perpendicular field, while the exciton gap kept the same and independent with perpendicular field. It confirmed observed nontrivial gap at 35T did not come from perpendicular field caused by small deviation from the plane.

![Graph](image)

**Figure 6.17** Magneto-conductance of Corbino device C1 under magnetic field. The total magnetic field is 35T. By rotating the device, the device plain has a deviation angle with magnetic field so that there is a small perpendicular magnetic field (0T-±2T). The blue solid lines are for positive perpendicular magnetic field while the red dotted lines are for negative perpendicular magnetic field.
6.4. Formation of excitonic ground state in shallowly inverted band

Figure 6.18 the dependence of the conductance minimum on $1/T$.

At $V_b=0V$, the dependence of the conductance minimum on $1/T$ was measured as shown in Figure 6.18. The Arrhenius plot showed that the conductance $\sigma_{xx}$ vanished exponentially with $T$. There were two linear regimes, which were at lower and higher temperature respectively. At higher temperature, the linear regime corresponded to hybridization gap which was discussed in lots of literature. At lower temperature, it corresponded to the emergent exciton gap.

Then I measured the gap as a function of $n_o$ for a quantitative confirmation of "soft gap" to "hard gap" transition when the lowest density was approached. The data at the CNP followed well the relation $\sigma_{xx} \propto \exp(-\Delta/2kT)$ yielding a set of $\Delta$ as
plotted in Figure 6.19. The activated behavior demonstrated the dissolution of exciton in the high temperature. Most interestingly, $\Delta$ diminished steeply to nearly zero as $n_o$ increased by just a factor of 2 (Figure 6.19). Hence, $\sigma_{xx}$ was strongly correlated with $1/n_o$, presenting the low CNP density was critical in the formation of hard gap. This result, combined with $n_o$ dependent $\sigma_{xx} - \Delta V_f$ under $B_{||}$ confirmed $n_o$-dependent Coulomb interactions driven Mott transition from exciton insulator state to semi-metallic e-h plasma. This transition couldn’t be explained by the single-particle band theory (10), where hybridization gap was expected to decrease for lower $n_o$, but could be understood with exciton insulator model (35). Here exciton diameter $d=2\hbar/\sqrt{2m}\epsilon_0$ was nearly 49nm, where $m$ was exciton reduced mass, so $n_od^2 \sim 1$ satisfied the critical condition for exciton insulator. As $n_o$ increases, Coulomb interaction was screened and tunneling was strengthened; both factors dramatically reduced $\epsilon_0$ so that $n_od^2$ dramatically increased with dissolving of excitons. Through tuning $n_o$ and temperature, the observations agreed with exciton insulator phase diagram predicted as shown in Figure 6.3.
Figure 6.19 (left panel) shows dependence of the conductance minimum on 1/T for different $V_b$. Here the $\sigma_{xx}$ is normalized by $\sigma_{xx} = \sigma_{xx\, \text{min}}/\langle\sigma_{xx\, \text{min}}\rangle$ at $\sim2.5K).$ Solid lines are guides to the eye. Black circles are for $V_b=0V$, red squares are for $V_b=-1.5V$, green down-triangles are for $V_b=-3V$, blue up-triangles are for $V_b=-4.5V$, and yellow stars are for $V_b=-6V$. (right panel) shows the activated energy vs $n_o$.

CV measurement was performed at different $V_b$ in Corbino device C1 and C2(Figure 6.20). Besides dc $V_f$ and $V_b$, a small ac voltage with 100Hz low frequency $f$ was delivered to the front gate, and the capacitance between front gate and QWs could be measured. For CV measurement(26), the geometry capacitance $C_g$ and quantum capacitance $e^2 D$ contributed to the measured capacitance per unit area $C_m$,

$$\frac{1}{C_m} = \frac{1}{C_g} + \frac{1}{e^2 D}$$
where $D$ is the DOS. In electrons or holes dominating regime, $C_g$ was much smaller than quantum capacitance, so $C_g$ dominated $C_m$. When DOS decreased and quantum capacitance was smaller than $C_g$, $C_m$ dropped and quantum capacitance dominated.

![CV curves with $V_b=0V$ (red line) and -6V (blue line). At 300mK, a low frequency (100Hz) ac voltage is delivered to the frontgate with QWs grounded, and the capacitance between frontgate and QWs can be measured. b shows CV curves with $V_b=0$V from wafer A (red line) and wafer B (blue line), respectively. In both cases, large capacitance drops exist. c. CV curves under different temperatures (0.3K, 2K, 4K, 6K and 10K). Curves are taken in C1 with $V_b=0$. The observed insulating gap disappears for the temperature up to 10K.

Figure 6.20 a, CV curves with $V_b=0V$ (red line) and -6V (blue line). At 300mK, a low frequency (100Hz) ac voltage is delivered to the frontgate with QWs grounded, and the capacitance between frontgate and QWs can be measured. b shows CV curves with $V_b=0V$ from wafer A (red line) and wafer B (blue line), respectively. In both cases, large capacitance drops exist. c. CV curves under different temperatures (0.3K, 2K, 4K, 6K and 10K). Curves are taken in C1 with $V_b=0$. The observed insulating gap disappears for the temperature up to 10K.

In CV measurement, at $V_b=-6V$, as illustrated in the blue trace of Figure 6.20a, the capacitance at the CNP was the same with that in the single carrier regime, agreeing with previous CV results (26) about hybridized gap that tunneling was too weak to form a hard gap. The DOS in this gap can be treated as the mixture of electron and hole’s. At $V_b=0V$, the capacitance at the CNP had 90% drop, compared with that in electron/hole regime. The capacitance drop manifested the insulating
state at the CNP was from the gap formation, not localization, since the localization could not reduce DOS. CV measurement performed in Corbino device from wafer B with \( V_b = 0V \), also showed there was a capacitance drop at the CNP (Figure 6.20b). It should be mentioned that in wafer B there was not doping with the localization minor. Therefore CV results in wafer A and B demonstrated the hard gap in the shallowly inverted band didn’t origin from the localization or Si doping, but supported that coexisting electrons and holes would form excitons with a reduced DOS.

Next I increased the temperature. Figure 6.20c showed CV traces in shallowly inverted band at different temperatures. At higher temperature the capacitance drop began to lift at 6K and disappeared at 10K. This could be attributed to the thermally activated excitons which dissolved and increased DOS, giving \( T_c \sim 10K \) with the gap energy of 20K. With these results, transport gap \( \Delta \sim 20K-25K \) was confirmed. It also confirmed the exciton gap dominated in low temperature regime below 6K and provided the temperature measurement window for it. This result combined with that in Figure 6.19, showed hybridization gap in the high temperature regime and exciton gap in the low temperature regime.
Figure 6.21 (a) and (b) show nonlocal measurement performed in meso-H bar from wafer A under 0T and 35T, respectively. The dotted lines indicate the expected resistance value from Landauer-Büttiker formula. (c) and (d) show nonlocal measurement performed in meso-H bar from wafer B under 0T and 35T, respectively. In the gap, the current path is shown in the inset of (a) as red and green arrows.

Remarkably, the experiments had also confirmed the helical edge transport throughout the entire field range. To this end, nonlocal measurement (11) in the dilute limit was performed in mesoscopic H bar device H1. The electrical current was passed through contacts 3 and 4 and the voltage was measured between contacts 1 and 2. In an ideally bulk-insulating QSH insulator, the currents went through all contacts with the path surrounding the bulk. According to LB formula (46), $R_{12,34}$ should measure quantized resistance of $\frac{h}{4e^2} \sim 6.45k\Omega$. Under both $B_{//}=0T$ and 35T, as shown in Figure 6.21a and b, I indeed observed quantized plateau close to this value. The same observations were true for device H2 from wafer B, as shown in Figure 6.21 c and d.
Nonlocal results not only confirmed the bulk at the shallowly inverted QWs was truly insulating and revealed helical edges in this case could survive under extremely high $B_{//}$, but also should be taken as a direct evidence for topological insulator originated from excitonic ground state. Here, the intralayer particle distance $L$ increasing to $\sim 45\text{nm}$, much larger than the layer separation $d\sim 11\text{nm}$, made the interlayer interaction much stronger than intralayer one. On the other hand, observed nontrivial helical edge state required topological nature of the hard gap in the shallowly inverted band for the correlation between spatially-separated electron and hole layers. In the case that single-particle coupling wasn’t the origin of correlation between layers, the observed nontrivial gap required such correlation to be achieved by interlayer Coulomb interaction with the formation of exciton condensation(28,76).

### 6.5. BCS like exciton condensation

When the lowest CNP density was achieved($V_b=0\text{V}$), the intralayer particle distance $\sim 50\text{nm}$ was much larger than the interlayer distance $\sim 10\text{nm}$. Thus strengthened Coulomb interaction in spatially separated electrons and holes made exciton insulator expected. In this case, through BCS mean field theory(37,41,52), exciton dispersion in k space $E(k)$ and the gap function $\Delta_k$ were calculated as shown in Figure 6.22. Similar with cooper pairs, the correlated electron-hole pair in exciton insulator leaded to an unstable Fermi surface, and spontaneously opened an exciton insulator gap with $\Delta_{\text{max}}$ near $k=k_f$ observable in transport measurement.(35) As
shown in the theoretical calculation in Figure 6.22, $\Delta_{\text{max}}(k=k_f)$ of nearly 2meV was consistent with the exciton gap energy obtained in transport results. On the other hand, as a ground state, the bulk exciton insulator state absorbed incident photons and thus was excited to electrons and holes, with the dispersion $E(k)$ identified as the energy cost of taking excitons out of condensation. Together with exciton distribution, exciton dispersion gave particular absorption features: the first absorption line was near $k=k_f$ with $E(k_f)=E_{\text{min}}=\Delta_{\text{max}}$ corresponding to the exciton insulator gap; the second one appeared near $k=0$ with $E(0)\sim8\text{meV}$. Furthermore, both the $\Delta_{\text{max}}$ and $E(k)$ absorption features should dominate below a critical temperature $T_c\sim\Delta_{\text{max}}/2\sim10\text{K}$. If the temperature was higher than $T_c$, cooper pair like state would dissolve, driving exciton insulator features disappeared. Moreover owning to the weak binding nature, the intra-exciton levels became minor so the 1s-2p transition in the exciton gas should not be expected. These particular features made BCS-like exciton condensation distinct from excitonic BEC or optically-excited exciton gas, and discernible in measurement.
Figure 6.22 Gap function $\Delta(k)$ (red line) and the pair-excitation energy $E(k)$ (green line) of the exciton droplet as a function of $k$. The blue cone means the part could be accessible to THz radiation. The grey part indicates the exciton gap regime. The data is calculated by Dr. Kai Chang.

As demonstrated in transport experiments, the exciton gap with $\Delta_{\text{max}}$ existed in the shallowly inverted band. But conductance and capacitance measurements could not give the information of exciton absorption, which should be accessed through THz absorption experiment. From the device fabrication, THz experiments had different requirement from transport experiments. One exciton absorption feature had the energy of nearly 2meV, corresponding to a wavelength of 600µm. To cover this energy feature and ensure the signal strength that was proportional to the device area, the device needed to be at least 3x3mm. However, in order to tune
the band inversion and the Fermi level, we still need the front gate which should be semitransparent and had large area of 3x3mm. Although the idea to realize exciton insulator in this broken gap double layer structure was proposed in 1980s, due to the difficulty in such gate fabrication, there was little progress in optical experiments. The numbered experiments(32,33) were performed in the ungated devices which cannot be turned to CNP.

Here I successfully fabricated 5x5mm device covered with a semitransparent gate. The gate was 10nm Pd film with Al2O3 dielectric layer underneath. In optical experiments, the transmission of THz pulse with the range of 0.3-2.5THz through the device was measured, in the collaboration with Xinwei from Prof. Kono’g group. In this experiment, there were two pairs of optical windows designed in the cryogen Dewar which allowed the room temperature radiation going through the device. As we started this experiment, we had no ideal about this but found the basic transport result could not be repeated in this Dewar. After the sample was loaded in the Dewar, it had a lower initial resistance and I could not tune the system into the CNP (with $V_f$ decreasing the resistance just kept increasing). Then we tried to find the problem. First we suspected that it was from the different cooling speeds (the cooling speed in transport experiments was faster). Then when I did the cooling down with the same speed in the transport Dewar, I did not see the problem in the optical Dewar. Then we suspected the glue on the backside of the sample induced the problem through the shrinking. After I fabricated a new sample without glue, the problem was still there. Further I suspected the different ways of sample mounting
caused the problem. In transport measurement, we usually put the sample into a sample holder, but in optical measurement, the sample was mounted in a hole which was required by the light going through. Therefore in the transport Dewar, under the low temperature, the sample would shrink but in the optical Dewar would be extended. However after we mounted the sample in the sample holder and cooled down it in the optical Dewar, the problem was still there. These results have already located the problem in the optical Dewar. In these tries, we found if we opened the room light, the sample resistance dramatically decreased. With the room dark, even if we opened the door, the sample would become insulating. Then we suspected it was the problem of light doping. First we covered the optical windows with Al shells, the sample was found to be insensitive to the room light but we still could not repeat the results in the transport Dewar. Then as we covered the optical windows from the cooling down, the initial resistance was found to be higher (but still lower than that in the transport Dewar) and the sample still could not be tuned to the CNP. But good news was that the result of every cooling down in this way was repeatable. Then we suspected that maybe we did not covered the optical windows very well. So in the cooling down, we lifted the sample to make it higher than the optical window. Thus even if there were leaked lights, they could not reach the sample. We found in this way finally the results could be repeated in the optical Dewar. However an odd thing happened. After carefully sealing the optical windows, if we located the sample near optical windows, we failed to repeat again. After a long time deep thought, I found we missed a detail: the room temperature radiation. Even if we
covered the optical window with metal shells, the metal shell was in room temperature so there was still the room temperature radiation going into the sample space. This radiation induced optical doping, making the device sensitive to the room temperature environment. To eliminate this radiation, I made an Al cover (Figure 6.23) and we installed additional pair of black polypropylene films on it. The black polypropylene could filter the room temperature radiation (1µm to 10µm) effectively (Thanks to Xinwei). Finally we could repeat the transport result in the optical Dewar, even under the THz radiation. Our experience indicated it was necessary to perform the joint transport measurement with optical experiments.

Figure 6.23 The home made Aluminium cover with black polypropylene window
Through the semitransparent front gate, we modulated the system into the CNP and the low mobility electron regime separately. Electron regime we selected had a low mobility, so it had no feature in the THz range, working as an ideal reference for the exciton-induced transmission. To directly measure the low energy exciton absorption, the low temperature THz transmission spectroscopy (33,77,78) was performed. Transmission spectra at the CNP in the shallowly inverted band at 1.4K were shown in Figure 6.24. The dip in the transmission trace meant the absorption line. One main observation was the occurrence of two absorption lines in the spectra: one was a sharp line near 0.5 THz(23K) while the other ranged broadly near 1.8THz(85K). Observed line energies were consistent with the calculated $E_{\text{min}}(k=k_f)$ and $E(k=0)$, respectively. In exciton insulator formed via Coulomb attraction between spatially separated electron and hole, these two lines could be attributed to different optical transitions of exciton in $k$ space. The sharp line came from exciton absorption near Fermi level, corresponding to exciton insulator gap, while the broad line corresponded to the exciton absorption near $k=0$. It should be noted in the single carrier regime or at the CNP of deeply inverted band we did not observe any absorption line.
Figure 6.24 Transmission spectra at the gap with ratios of transmitted intensity to a reference spectrum (low mobility and low density electron regime) (normalized by ratioing to zero-field spectra in the electron regime at the same temperature)

The temperature was also crucial for absorption lines. As the temperature gradually increased (Figure 6.25), both lines shrunk quickly. For the broad line, the line amplitude decreased from 13% at 1.4K to 4% at 10K, showing a dramatic drop of absorption near 10K. It was similar for the sharp line. Eventually, at 40K, we found no features were observed in the transmission spectra, thus the system was metallic. Therefore, the transition from insulator into the metallic state was thermodynamic and $T_c \sim 10K$, agreeing with the theoretical value (79,80).

Importantly, $T_c$ was much smaller than the broad line energy $\sim 85K$, suggesting a many-body interaction in the formation of broad line.
Exciton insulator should have a magnetic stability which had already been shown in the transport result and thus was expected in the optical measurement. In the left panel of Figure 6.26, as the perpendicular magnetic field was applied at 1.4K, for both lines, the absorption amplitude became larger, which could be understood as LL concentrated the DOS. The magnetic stability was more explicit in the case of 20K. As shown in the right panel of Figure 6.26, at 20K, the lines originally were gone under 0T; by contrast, under magnetic field, both lines were strongly strengthened and re-emerged. It showed the absorption line was magnetically

\[\text{Figure 6.25 Transmission spectra at the gap under different temperatures.}\]
stabilized against a thermal transition to e-h plasma, which was consistent with the theoretical prediction of magnetic stability of exciton insulator. (80)

Figure 6.26 Transmission spectra at the gap under magnetic fields at 1.4K (left panel) and 20K (right panel)

Despite different line energies, both lines disappeared under the same Tc and reappeared simultaneously under the same magnetic field. Such association confirmed the broad line had the same origin with the sharp one. Two observed lines and line energies were in a full agreement with expectations for absorption spectrum in exciton insulator in Figure 6.22, providing direct evidence for the existence of exciton insulator, where excitons could exist without optical pumping and spontaneously form the BCS-like gapped state.
As a conclusion, BCS like exciton condensation in inverted InAs/GaSb QWs which hosted spatially-separated electrons and holes, was investigated using CNP density \((n_o \sim p_o)\) in gated-device as a tuning parameter. For the deeply inverted band, \(n_o \gg 5 \times 10^{10}/\text{cm}^2\), a soft gap \((i.e., \text{finite density of states})\) opened predominately by hybridization; For the shallowly inverted band, approaching the dilute limit \(n_o \sim 5 \times 10^{10}/\text{cm}^2\), a hard gap \(\sim 2\text{meV}\) opened leading to a true bulk insulator with quantized edges. Two gaps crossed in the intermediate equilibrium density. The hard gap was dramatically reduced as the QWs were tuned to less dilute. Moreover, the response of gaps to in-plane magnetic fields \((B_{/\!/})\) showed that for soft gap vanished and became semimetal above 10T where tunneling was eliminated, consistent with e-h hybridization origin, while for nontrivial hard gap opened continuously for \(B_{/\!/}\) as high as 35T. The data were remarkably consistent with spontaneous exciton binding bulk state in dilute InAs/GaSb bilayers, suggestive of the formation of BCS like exciton condensation. The 2meV exciton gap was further confirmed in the low temperature THz absorption experiments. In the THz experiment, the associated two absorption features were observed to vanish near 10K and became more stable under the magnetic field, which quantitatively agreed with the BCS like exciton condensation in the prediction from mean field theory. Our findings provided compelling experimental observation of spontaneous exciton ground state. Future work can be expected for bilayer superfluidity and Bose-Einstein condensation, and others, in this highly tunable e-h system.
If we combine the conclusions in Chapter 5 and Chapter 6, it is obvious that BCS-like exciton condensation is topologically nontrivial with broken TRS, in this system. Whether it is a general phenomenon, still need more inputs in the theoretical aspect and more experiments in other systems.

6.6. Quantum phase transition driven by high magnetic field

The fate of helical edge state under broken TRS driven by magnetic field in inverted InAs/GaSb QWs has been investigated in Chapter 4. Under the perpendicular magnetic field, two loops of edge currents were driven spatially separated, with one edge shrinking into the bulk and the other expanding towards the boundary. Through transport and optical measurements, the bulk gap at the CNP with the low density was studied, where exciton insulator was observed in this electron-hole spatially separated system. Naturally, one could expect the evolution of exciton gap in the magnetic field-induced topological phase transition.
Figure 6.27 Landau level fun chart of a typical inverted QW structure. The figure is adopted from (7).

At the CNP, if the bulk behaved as a semimetal, which meant balanced electrons and holes distributed in two QWs, under the magnetic field, a series of LLs were formed in electron and hole QWs, respectively and the lowest electron like LL was expected to cross the lowest hole like LL at a critical magnetic field (Figure 6.27). The cyclotron energy was $B e / m$ and Zeeman energy splitting was $\mu_b g B$. So the energy difference between lowest filling levels of electron and hole was

$$E_g - \frac{1}{2} \left( \frac{eB}{m_e} + \frac{eB}{m_h} \right) + \mu_b g B,$$

where $g$ was 12 and $E_g = 4.3 \text{meV}$, indicating the critical field was 4T. Above this field, the lowest electron like LL was higher than the lowest hole like LL and if the LL broadening was small, a bulk gap were present with a trivial insulator. However, as investigated at Chapter 5, under zero field, the
semimetal in the low density was unstable against the formation of exciton insulator. In the presence of exciton insulator, which essentially belonged to the many-body problem, the single particle band theory cannot be applied anymore. For exciton gas, there were not electrons or holes in the bulk; under the magnetic field, instead of LL, a series of exciton levels would develop similarly with the hydrogen model. However, until now, the behavior of the exciton insulator under the high magnetic field was still unclear theoretically. The experiments here would present an opportunity to learn physics beyond the simple textbook QSH theory which was routinely confirmed by experiments.

Figure 6.28 The schematics of the contact-buried Corbino and Hall bar devices
Figure 6.29 A phenomenological phase diagram is shown for distinct topological states under perpendicular fields up to 35T.

Figure 6.30 Bottom panels display experimental traces of longitudinal conductance $\sigma_{xx}$ and Hall conductance $\sigma_{xy}$ measured at a fixed field of 8T,
16T, 35T, respectively, as a function of gate voltage, where wide ranges of zero-conductance can be seen punctuated by $\sigma_{xx}$ peaks. Here from right to left, the Fermi energy $E_F$ is swept from electron- towards mobility gap and into hole- regimes. The peak below the $v=1$ QH plateau marks the boundary into QSH state. Remarkably, inspecting traces from zero to 35T in top panel, we found that the bulk gap for QSH increases with perpendicular fields, instead of closing. The inset depicts a model of "canned helical state" which is deformed adabatically from the zero-filed HL. Note that the in-plane spin components remain spin-momentum locked.

Under the high magnetic field, the ungated 2DEG connecting the QSH insulator and metal contacts would be insulating. To solve this problem, the gates need to be extended to cover metal contacts that thus would connect to the QSH insulator directly. I fabricated such contact-buried devices as shown in Figure 6.28. Figure 6.29 showed the magnetic dependence of the bulk conductance in the contact-buried Corbino device under $V_f$ sweeping. At zero magnetic field, there was a zero conductance plateau near the CNP, corresponding to the exciton bulk gap. With the magnetic field rising, several QH transitions between different LLs appeared in the electron regime with conductance peaks. For example, under 4T, the transition from the second LL to the first LL occurred near $V_f=-1.2V$. Remarkably, as the Fermi level was further swept down, a phase transition between the QH insulator and exciton insulator was observed with a conductance peak. Moreover, the zero conductance plateau corresponding to exciton insulator kept constant until 4T, indicating a similar exciton energy. On the other side, the conductance minimums indicated the first, second and third LLs; the conductance peaks of corresponding transitions were separated with the same voltage drop. This was
because the total electron density without spin degeneracy in one LL per unit was constant, i.e. $eB/h$, determining the voltage drop according to the simple capacitor model. However, the equivalent voltage drop was broken by the peak of first LL to exciton insulator transition, which could be attributed to the emergence of exciton gap in prior of electron depleting.

Then at a higher magnetic field, i.e. 6T, similarly, due to the exciton ground state near the CNP, the lowest filling factor peak shifted to a more positive voltage. Since the exciton energy was more than 2meV under magnetic field (80), the system under the magnetic field that was even a few Tesla larger than 4T should still be in the exciton regime, which was responsible to the peak shift at 6T, and not fall into the noninverted semiconducting area.

As the magnetic field was further lifted away from the crossover regime, the zero conductance plateau became wider and its right edge was linear with the front gate voltage. Under 35T, the energy of the lowest electron filling level was 30meV higher than that of the highest hole filling level. In experiments, the level broadened due to the disorder. If the disorder was large enough, there still might be electrons and holes coexisting within this 30meV energy gap. For the electron filling level, the level broadening could be estimated as $\hbar \sqrt{\frac{2eB}{\pi m_e \tau}} \approx 20meV$, where $\tau$ took 0.28ps. Another factor needed to be considered was that the disorder would effectively reduce the band gap. Also the exciton effect cannot be neglected. In this situation, if the exciton binding energy was larger than the decreased effective gap energy, the
exciton would become the ground state. As an example, in Xia’s calculation (80), the binding energy has already increased by 3meV under 3T, which would further increase under higher magnetic field. As the exciton binding energy became larger than 10meV at 35T, magneto exciton state instead of the semiconducting gap would spontaneously form. Thus, this “noninverted” band under 35T may be different from the normally recognized noninverted semiconductor band.

Further, the edge state was investigated in a 50×100μm contact-buried Hall bar device A as shown in Figure 6.30. At a fixed magnetic field, a series of topological phase transitions as a function of $E_f$ took place in the phase diagram. In the QH regime, the transition between different quantized conductance plateaus coincided with the conductance peak. The peak below $\nu=1$ QH plateau marked the boundary of exciton insulator. In the exciton regime, the Hall resistance increased with the magnetic field, consistent with the chiral-like state demonstrated in the Chapter 5.5. The perpendicular field up to 6T separated helical edges in the spatial space, and longitudinal conductance results showed a helical to chiral-like transition. It should be noted that the chiral-like state near the CNP was topologically distinct from the nearby $\nu=1$ chiral topological phase, as indicated by a conduction peak due to gap closing and reopening. At the highest field of 35T, the QSH regime showed an approximate Hall plateau of $e^2/h$, resulting from the imbalance between the left- and the right-movers and showing two movers were totally separated. On the contrary, in the nearby $\nu=1$ QH topological phase, there was only one mover with the single direction. This unequivocally demonstrated that edge transport in InAs/GaSb QWs
was robust to broken TRS, and confirmed the model of "canned helical state" that was deformed adiabatically from the zero-field helical edge.

Figure 6.31 Hall resistance $R_{xy}$ measured under fixed fields up to 18T in device B.
Figure 6.32 Hall resistance $R_{xy}$ at CNP extracted from Figure 6.31

Figure 6.31 showed the Hall resistance $R_{xy}$ vs $V_f$ under perpendicular magnetic fields up to 18T from the Hall device B with a 50x50μm size. Due to the slightly asymmetry of Hall bar geometry, the Hall resistance induced by the helical edge at the CNP was not exact zero under 0T. At higher magnetic field, the inner loop edge had a gradually smaller transmission possibility and the Hall resistance was lifted as shown in Figure 6.32. The results in device B were consistent with those in device A.

Figure 6.33 Longitudinal resistance $R_{xx}$ measured under fixed fields up to 18T in device B.
In the device B, the longitudinal resistance under the magnetic field was also measured as presented in Figure 6.33. According to LB formula, for a reduced transmission possibility of one edge in one direction, the transverse resistance should decrease with the magnetic field. Consistent with the prediction, as shown in Figure 6.34, higher magnetic field induced a lower longitudinal resistance at the CNP.

Recently, in inverted HgTe QW(81), similar result under magnetic field has been reported by the microwave impedance microscopy experiments. The critical field in their case was 3T, and thus the lowest electron LL was above the highest hole LL at 10T. Moreover, because of the small g factor, 10T was large enough to break the TRS. Surprisingly, at the magnetic field up to 10T, the helical edge channels persisted near the CNP. Moreover in the strain-layered InAs/InGaSb QWs,
at the magnetic field higher than 4T, the helical edges presented a re-emergent helical edge state, instead of vanishing helical edge mode. Combining all of these results, we found that the helical edge state might be robust to the broken TRS, regardless of the material details. The mesoscopic mechanism of robust helical edge state observed in different systems remained unclear despite the fascinating experimental results. With InAs/GaSb QWs as one of the leading QSH systems, our results should extend broadly understanding in the physics of topological material under broken TRS and motivated the relative development in theoretical topological field. Our observation represented an opportunity to understand beyond the single particle topological theory and shed light on understanding the QSH effect in real systems.
Chapter 7

1D Coulomb Drag in topological insulator

7.1. Introduction

Observed exciton ground state validated the theoretical prediction about exciton QSH effect (28). As I have discussed in Chapter 5, it had notable impacts on the transport property leading to dissipationless transport in the mesoscopic device. The remarkable property of robust helical edge states was their inherent protection against the disorder, temperature, and broken TRS with the slow Fermi velocity \( v_F \sim 1.5 \times 10^4 \text{m/s} \) and the long edge scattering time 200ps.
These characters of this nontrivial edge state intrigued great interests in inter-edge Coulomb interaction (82) realized via the topological 1D edge state. Coulomb drag as a standard toolbox could provide additional insight including fundamental microscopic properties and many-body structure regarding the novel edge state, especially in the presence of bulk electron-hole pairs. Due to the reduced screening, the interaction effect in 1D was stronger than 2D and widely explored to search the exotic quantum state with many-body interaction such as Wigner crystal (83), Luttinger liquid (84) and exciton quasicondensate (85) through Coulomb drag experiment.

In previous 1D Coulomb drag experiments (83,84,86), 1D quantum wire was a representative system that was formed by the quantum confinement in 2D electron gas, and the electron density needed to be precisely tuned by the voltage and magnetic field for the single mode property. As a contrast, the TI edges here were 1D wires naturally generated with the single mode and precise quantization conductance plateau, in the absence of a magnetic field. The combination of topological protection and electron correlations implied that a TI edge was an ideal system for exploring Coulomb drag physics.
In Coulomb drag, one electric current inducted another coupled current, purely through the long-range Coulomb interaction. Coulomb drag has two types: positive drag and negative drag (Figure 7.1). For the positive Coulomb drag, electric current in drag circuit was induced in the same direction as the current in drive circuit, while for the negative Coulomb drag, it was the opposite. Theoretically (87–93), only positive drag was expected. However, in the plane-integrated quantum wires, because the electron density in both driving and drag circuits was very low, the negative drag emerged which was enhanced at lower temperatures and attributed to the many body interaction. Later, in the vertical-integrated quantum
wire, an increase of the voltage induced alternately positive and negative drags. These fascinating experiment results attracted intensive theoretical interests.

7.2. Device fabrication

Previously, 1D Coulomb drag was realized in coupled quantum wires through two different ways. The first way was that two wires were generated and separated laterally by split gates technology including three distinct gates. A soft barrier was formed by negative voltage in the center gate, and the barrier width was limited by the depth of 2DEG in GaAs/AlGaAs QW, which was nearly 90nm. The second way was to construct a coupled circuit vertically through the epoxy-bond-and-stop-etch technology, where coupled circuits were separated vertically by 15nm by a hard barrier between QWs. In this approach, to minimize the separation between coupled wires, two wires needed to be aligned accurately, enlarging the difficulty of fabrication. In both ways, wires were created by split gates, and either soft or hard barriers had a large dielectric constant which reduced Coulomb interaction but strengthened the capacitively coupling. Here, we used 1D topological edges in circuits as drive and drag wires which were separated by the vacuum spacer of 50nm. There were two advantages compared with traditional quantum wire approaches. First, topological edges in circuits were naturally 1D with the single mode, which was immune to the elastic scattering, disorder, or localization usually existing in quantum wires. Second, the vacuum spacer had the lowest dielectric
constant, bringing the strongest Coulomb interaction and smallest capacitive tunneling.

Figure 7.2 Schematics of the fabrication process

Figure 7.3 Schematics of the fabrication process
The fabrication schematic of the topological circuits is shown in Figure 7.2 and Figure 7.3. An inverted InAs/GaSb heterostructure was used to fabricate. Mesoscopic H bar geometry was defined by ebeam lithography and photo lithography, and wet etched by phosphoric based etching solution with 600nm depth. GePdAu ohmic contacts were deposited on the structure. Then a 25nm PMGI layer was spin coated and hard baked at 250oC, followed with 45nm SiO₂ layer deposited by ebeam evaporator. 50nm PMMA was finally spin coated to form the triple-layer structure. Focus ion beam (FIB) lithography was used to define a single line with higher resolution, and the pattern was developed in the MIBK:IPA(1:3) solution. Due to the heavy gallium ion and their limited energy, FIB has negligible ion scattering and proximity effect in the PMMA, as well as the low backscattering from the substrate. The penetration depth of such ions was much smaller than that of other particles and hence only PMMA film less than 60nm was satisfactorily processed using FIB lithography. The SiO₂ layer served as a protection layer for QW to stop the penetration of Ga ion. FIB lithography defined a wire pattern with the width of nearly 30nm and the length of 2μm in the center of H bar geometry. After Reactive ion etch of CHF₃/Ar, the pattern was transferred to the SiO₂ layer. Then oxygen based RIE further transferred the pattern to the PMGI layer. With SiO₂ hard mask and BCl₃ based RIE, a slit was etched into QWs with the depth of nearly 300nm. Then a dilute phosphoric based acid and NH₄OH based wet etchant was followed, to eliminate possible pollution caused by RIE. The PMGI and SiO₂ layers were striped under PG Remover. The slit width was nearly 50nm. With the shadow evaporation,
Al₂O₃ was deposited through e beam evaporator with an angle of 45 degree. Then
Al₂O₃ film was evaporated from the other side with 45 degree, followed by Al₂O₃
film evaporated from original side with 45 degree. With such shadow evaporation
technology, an Al₂O₃ air bridge could be constructed over the slit. Finally, using
ebeam lithography, Al gate was deposited on the airbridge-split-H bar device. The
heterostructure had a 250nm thick AlGaSb up barrier. The distance between the
gate and the edge was more than 500nm, much larger than the separation between
two edges, to eliminate possible screening from the metal gate. Hence the helical
dge could be treated as a really free current edge. This distance ensured a uniform
electric field through QWs to flatten the potential roughness, and the simultaneous
appearance of topological edge in both driving and drag circuits. In this case, the
fringing field of the gate on the top of the driving circuit also was imposed on the
drag circuit, and the split gates had no difference from this global gate. Although the
split edge length was only 1μm, it was 250 times longer than the effective separation
between split edges which was 4nm (according to the dielectric constant of AlGaSb).
In measurement, the devices were immersed in the liquid He₃ which had a dielectric
constant close to vacuum, i.e. relative dielectric constant 1.05.
Figure 7.4 Diagram of the Coulomb drag circuit (left). The current goes upward in the drive circuit following with the solid blue line. The dashed line indicates the other helical edge which does not hold the current. (right) The scanning electron microscopy of the slit regime. The white mark means 100nm.

As shown in Figure 7.4, the coupled wires consisted of two helical edges with the same length forming in the center of a H bar device, and spatially separated with a 50nm distance. The current was injected into contact 3 and contact 1 was grounded. The voltage between contact 2 and 4 was measured. As the global gate tuned the split H bar into the QSH regime, the bulk became insulating and the helical edge modes went along the boundary. Consequently, the drag and drive circuits, consisting of helical edge modes, formed. As the nonlocal current went through the edge (blue line) in the drive circuit, in the drag circuit two channels were ready for dragging. H bar geometry ensured that if there was drag signal it must be induced by the nonlocal helical edge, since only the edge current went through the driving and drag wires. If the system was dominated by the bulk current, there was little current in the driving wire.
7.3. Characterization

Figure 7.5 Longitudinal resistance vs the front gate voltage in a 1x2μm π bar. Inset shows the temperature dependence of the longitudinal resistance peak.
Figure 7.6 The frontgate voltage dependence of the bulk conductance in a Corbino device.

To characterize properties of helical edges, a 1x2μm π bar and Corbino device were fabricated. Figure 7.5 and Figure 7.6 showed the transport result of the helical edge and bulk states. In Corbino device, where edges were shunted by metal contacts, only the bulk transport could be contributed to the measured conductance. There was a zero conductance plateau at the CNP indicating the bulk resistance was more than 100MΩ. Due to the insulating bulk, in the π bar, a resistance peak measured at the CNP must come from the helical edge. Moreover, the resistance peak approached to the quantized value, suggesting the single mode helical edge state with ballistic transport.
Figure 7.7 The leaking current vs bias voltage in the Coulomb drag device.

In the Coulomb drag device, the leakage current was also measured. The bias voltage was applied between contacts 1 and 3, with the current monitored as shown in Figure 7.7. The resistance was much more than 100GΩ as the voltage difference was within 0.5V, demonstrating the vacuum barrier insulates two edges. To avoid possible interedge leakage current, 100mV constant voltage, connecting with a 100MΩ resistance that was much larger than the device resistance, was applied to the drive circuit, and thus 1nA constant current was generated. The top gate defined the topological edges, including the drive/drag edge and several segments of long edges. Considering the coherence length of helical edges is 1.5-2μm, we know the in serial edge resistance was much larger than the quantized resistance, which made the direct probe of quantized resistance in split topological edge unrealistic.
7.4. Coulomb drag

In the Coulomb drag experiment, the current was injected into the drive circuit and the voltage drop $V_{drag}$ was measured in the drag circuit. To improve the noise to signal ratio, I used a low frequency (3.1 Hz) alternating current $I$ with lock-in technique, which was equivalent to the direct current measurement, since the 300 ms was much longer than the edge/electron scattering time 200 ps. The drag resistance was described as $R_d = -V_{drag}/I$. In the positive drag, $R_d$ was negative while in the negative drag, $R_d$ was positive. Typical experimental data of $R_d$ versus the global front gate voltage was shown in Figure 7.8. For the positive front gate voltage, the device was dominated by electrons, and thus, there should be little current injected into the drive circuit with the drag resistance close to zero. Then I tuned the Fermi level to the CNP and nonlocal edge transport dominated with helical edge current injecting into the drive circuit. I observed a positive drag resistance in the drag circuit at 300 mK. A resistance plateau appeared near the CNP and the maximum of drag resistance reached 800 Ω. The fluctuation on the drag resistance plateau was reproducible and did not stem from the electrical noise. It should be mentioned that the CNP ranged in a broad voltage as shown in Figure 7.6.
Figure 7.8 $R_d$ as a function of front gate voltage in the Coulomb drag device in 300mK.

Further I increased the temperature to 1K, and Figure 7.9 demonstrated the drag resistance at the CNP decreased with the alternative appearance of positive and negative Coulomb drag, dependent on the front gate voltage.
For a higher temperature (1.5K), Figure 7.10 exhibited that the drag resistance was mainly dominated by the positive drag with the negative drag resistance $R_d$. As the temperature reached 2K, the drag resistance at the CNP was nearly zero. It should be noted that at 2K the helical edge mode was still dominating the device, meaning the drive current existed, but there was not drag current in the circuit.

The rectification and ratchets mechanisms (94,95) could develop a voltage in the drag circuit which was non-symmetric with respect to the probe inversion. To examine the drag signal in Figure 7.8, I changed the current direction in the drive circuit and found the signal was independent of the drive current direction as
shown in Figure 7.12, confirming the signal was not mainly from a charge-fluctuation in asymmetric circuits (94,95).

**Figure 7.10** $R_d$ as a function of the front gate voltage in the Coulomb drag device in 1.5K.
Figure 7.11 $R_d$ as a function of the front gate voltage in the Coulomb drag device in 2K.

Figure 7.12 $R_d$ as a function of the front gate voltage in the Coulomb drag device for negative driving current.
Then I fixed the voltage at the CNP and modulated the temperature. As shown in Figure 7.13, the negative drag resistance was nearly 800Ω at 300mK which monotonously decreased with higher temperature. For temperature higher than 1K, the drag resistance became positive, and reached a maximum at 1.3K. Then, as the temperature further increased, the positive drag resistance decreased to zero. Different from the negative drag resistance, the positive drag resistance presented a non-monotonous behavior. When the temperature was higher than 1.7K, the drag resistance approached to zero. A similar result was obtained in another device with the same parameter as shown in Figure 7.14. The temperature dependent results showed that the mechanisms of positive and negative drags were competing.

![Figure 7.13](image)

*Figure 7.13 Temperature dependence of the Coulomb drag signal at the peak of the negative drag regime.*
Figure 7.14 Temperature dependence of the Coulomb drag signal at the peak of the negative drag regime in another device.

Negative drag was previously reported in split gate coupled 1D quantum wires(83) and vertically integrated quantum wires(84,86), as well as 2D coupled QWs(96). Also, current mirror effect can produce negative drag signal which was current rectification in parallel arrays of tunneling junction(97).

In the split gate coupled 1D Coulomb drag, two wires were laterally separated by electrostatic gate, and the effective barrier between 1D wires was soft and nearly 100nm. Under the low magnetic field, the positive drag dominated while under the higher magnetic field negative drag took over. Also, the negative drag required lower temperature and the low electron density. In the vertically integrated Coulomb drag where two wires were separated with a hard barrier of 15nm, the negative drag was also reported in the low electron density, i.e. less than
a full subband occupied. This could be attributed to the decrease of wire separation
which raised interwire correlation, and the magnetic field was not necessary.
Further, dependent on electron densities, positive and re-entrant negative drag
alternatively emerged. However, in our devices, magnetic field was also not
necessary to realize the negative drag. It should be attributed to the low dielectric
constant of the vacuum barrier. If we converted the split separation here to that in
GaAs, the separation would be only 4nm. Furthermore in coupled topological edges,
both positive drag and negative drag were shown and dependent on the
temperature. Helical edge modes in drive and drag wires were naturally 1D with the
single mode, required by the topological property. It corresponded to the one single
1D subband on each wire in the case of quantum wire; as a contrast, there was only
positive drag reported in that case. Thus, the results presented here were different
from previous ones in coupled quantum wires.

The capacitance between two edges was estimated to be $2 \times 10^{-18} \text{C}$, with an
impedance of $10^{16} \Omega$. Considering the input voltage is $100 \text{mV}$, we can obtain the
current induced by interedge capacitance as $10^{-17} \text{A}$ and the corresponding drag
resistance of $10^{-4} \Omega$, which was much smaller than the measured drag resistance. On
the other hand, in the case that two edges were capacitively connected through the
top gate, the drag resistance was also on the order of $10^{-4} \Omega$. These calculations
showed the observed drag current was not from the capacitance coupling.

For the current rectification processing, its IV characteristics were highly
nonlinear (with respect to $I$) and non-symmetric with respect to the probe inversion
with the monotonic temperature dependence, usually in capacitively coupled quantum dot systems. However, in this experiment, the low dielectric constant suppressed the capacitively coupling between two wires. According to the capacitively coupling, 50nm vacuum barrier was equivalent to nearly 650nm in GaAs, which was a separation large enough to isolate wires. Also, the drag resistance in the coupled topological edges was non-monotonic with temperature, and had both positive and negative drags in different temperature regimes, which could not be explained by the rectification processing.

Most of existing theories predicted a positive 1D Coulomb drag, until the negative drag was observed experimentally in two wires with the low electron density which was explained in the picture of Wigner crystallization. However, the Wigner crystal model cannot explain the high-density negative drag here. The re-entrant negative drag showed a transition to positive drag at higher temperature, qualitatively agreeing with our results. But it appeared with unbalanced electron densities, which was different from our case. Although it was predicted that the negative Coulomb drag can be induced by rectification of energy fluctuations, the drag in the model occurred near a transition of different conductance channels which were absent in our system. Also, this model could not explain the behavior of non-monotonic temperature dependence. Thus, our results could not be explained by electron-electron Coulomb drag.

Recently, it was proposed(98) that the particle-hole pairing over the coupled wires lead to the positive drag of hole current, i.e. negative drag of electron current,
which required a balanced density between electrons and holes in two wires respectively. According to the theory, if both wires were dominated by electrons of the same density, the Coulomb interaction between wires locked relative positions of electron over the wires, leading to the positive drag. On a contrary, if two wires were dominated by electrons and holes with the same density respectively, the strong attraction between electrons and holes on separated wires resulted in particle-hole pairs over the wires, preferring the positive drag of hole current. In our system, the drag circuit had counterpropagating edges with opposite spins, corresponding to 1D fully occupied subband quantum wires, while the drive circuit has a one-direction edge which also corresponded to 1D fully occupied subband wire. If we treated the edge in the drive circuit as the electron like quasiparticle flow, the edges with the distinct flowing direction in the drag circuit could be treated as the hole-like quasiparticle flow. Actually, such electron-hole asymmetry in helical edges recently was reported in bilayer graphene with bulk particle-hole symmetry(43). Thus, our results can be assigned to electron-hole asymmetry of helical edges, where the electron-like quasiparticle flowed in one edge and the hole-like quasiparticle flowed in the other edge in the same direction. On the other hand, as discussed in the Chapter 6, exciton condensation induced by bulk particle-hole symmetry was observed at the CNP of this shallowly inverted InAs/GaSb QWs, and the edge-bulk corresponding relation would also require the particle-hole symmetry in the edge. Under this picture, as the electron-like edge current flowed in the drive circuit, in the drag circuit, both electron-like and hole-like edge positive drag would
be induced. When the temperature was below 1.5K, the electron-like drag edge dominated, leading to a positive drag resistance. When the temperature was below 1.3K, the hole-like drag edge started to contribute to the net current. Then, at a lower temperature than 1K, the hole-like edge became the major effect, resulting in a negative drag resistance. The negative drag emerging in lower temperatures indicated a correlated state dominated in the negative Coulomb drag.

In the recent Coulomb drag experiment of 2D Dirac fermion in double layer graphene (99), where 2D Dirac fermions dominated in both graphene layers, the positive drag was exhibited under zero field and the strength approached zero at low temperatures. However, under small magnetic field, the drag was altered to negative with much large drag strength, which cannot be interpreted by existing theories. Authors proposed the following explanation in terms of condensation of excitons: in that system, electrons at N = 0 LL in one graphene layer can then pair with vacancy-like states at the N=0 LL in the other layer. That was to say: the small magnetic field strengthened the correlation, driving the drag from positive to negative. In the 1D case, negative drag also needed the low temperature, the low electron density, and magnetic field under a relative larger wire separation. However, as wires were in close proximity at the nanoscale, the tuning of positive to negative drag only needed the low temperature while magnetic field was not necessary. In the 2D case, although the layer separation was also nanoscale, due to the increased screening, in order to induce the tuning from positive to negative drag, the low temperature is not enough and the magnetic field is necessary. All of these
results pointed out that negative Coulomb drag was associated with a correlation that was expected to be stronger in lower dimension, closer separation, magnetic field, and lower temperature. As a result, here in Coulomb drag of 1D Dirac fermion (helical edge mode), as the correlation became larger by lowering temperature, the drag was observed to alter from positive to negative.
In the past four decades, the debates about the ground state in the broken gap InAs/GaSb QWs have not ceased. It was initially believed that the semimetal would exist, which was unstable against exciton insulator. Then, the self-consistent calculation and the following experiments demonstrated a hybridization gap owing to the interlayer tunneling, but the gap was found to be weak and its DOS was close to the semimetal. In this case, the existence of exciton insulator is still an open question. Recently, the TRS QSH effect was proposed in the hybridization gap, but the conductive hybridization gap made the TRS QSH effect inconclusive in this system, which motivated to realize insulating hybridization gap. However, the understanding of recent QSH experiments indicated the existence of exciton insulator in the bulk.
In this thesis, the quantum phases in the inverted InAs/GaSb QWs were completely investigated through the double gate tuning. We have observed that the exciton insulator and the hybridization gap were the ground states in the shallowly inverted and deeply inverted regimes, respectively. In the deeply inverted regime, the light and heavy hole band mixed and the tunneling dominated. By alloying GaSb QW with InSb, we have solved the problem of the conductive hybridization gap and observed the insulating hybridization gap for the first time. Moreover this gap had the largest gap energy in all known QSH systems, which provided the solid platform to explore the QSH effect. We have observed the exclusive evidence for the existence of the QSH effect in this insulating bulk gap via the nonlocal measurement. Owning to the large bulk gap, the observed helical edge modes had a coherence length of nearly 13μm, which was the longest in records, and was thermally stable up to 30K, much higher than the previous 4K. These results experimentally confirmed that the QSH effect with larger bulk gap could survive at higher temperature, demonstrating the way to construct a room temperature TI. We observed the QSH effect here was protected by the TRS. The conductance from helical edge state increased under the magnetic field which broke the TRS. It should be mentioned that this was the first time to observe the TRS QSH effect in InAs/GaSb system. Then, we performed in-situ electrostatic tuning to helical edge modes. As the bulk gap was tuned larger, we found the coherence length of helical edge became longer, which confirmed the observed long coherence length under the large bulk gap. If the bulk gap further
increased, we could expect that not only the helical edge could survive in room temperature but also the coherence length became macroscopic.

When the band was less inverted, the tunneling decreased due to the decoupling of light and heavy holes. In addition, the CNP density was few in the shallowly inverted band, such that the intralayer particle distance was much larger than the interlayer distance and the interlayer Coulomb interaction must be considered. In the shallowly inverted band, we observed the quantized plateau of the QSH effect for the first time. The plateau had the precision of one part in one hundred in multiple mesoscopic devices. As the edge length grew longer than the coherence length, the edges were broken into several segments. The quantized plateau and the coherence length were found to be robust under the disorder and the temperature. Surprisingly, under the inplane magnetic field breaking the TRS, we observed that the quantized plateau and the coherence length kept constant in both mesoscopic and macroscopic devices. Under the perpendicular field, the different longitudinal conductance behaviors in two terminal junction and multi-terminal devices confirmed that the helical edges were spatially separated with one edge shrinking into the bulk and the other approaching to the device boundary. These results experimentally indicated the helical edge mode in the shallowly inverted band had a protection other than the TRS and systematically presented the discovery of the TRS broken QSH insulator.

We further investigated the origin of the bulk gap in this novel quantum phase by using the CNP density in double-gate devices as a tuning parameter. We
found two distinct gap regimes: for I, \( n_o \gg 5 \times 10^{10} / \text{cm}^2 \), a soft gap opened predominately by hybridization, which closed at \( B_{//} \sim 10 \text{T} \); for II, approaching \( n_o \sim 5 \times 10^{10} / \text{cm}^2 \), a hard gap opened that leaded to a true bulk insulator with quantized edges, continuously for \( B_{//} \) up to 35T. As the CNP density decreased, the hard gap emerged and coexisted with the hybridization gap. The contribution from the hybridization gap continued to decease for lower density, and at the lowest density, the hybridization gap could be neglected. In addition, the hard gap was found to be correlated with \( 1/n \) and finally dominated the bulk gap. Capacitance measurement also indicated the dramatic drop of DOS within the hard gap, showing the formation of neutral particles. Our results confirmed that the hard gap cannot be explained by single-particle band theory but instead represented many-body correlations. The activation measurement and temperature dependent capacitance measurement indicated the hard gap energy was nearly 20-25K, which was consistent with the prediction of exciton insulator gap. Finally we performed the low temperature THz transmission experiments. Importantly, under THz photon radiation, two associated absorption lines with energy \( \sim 20 \text{K} \) and \( 85 \text{K} \) were observed in the hard gap. Lines had a critical temperature \( \sim 10 \text{K} \) and a magnetic stability against e-h plasma, agreeing with the results obtained in transport measurement. The two features simultaneously vanished below a critical temperature of \( \sim 10 \text{K} \) and reappeared under a magnetic field, confirming these two features had the same origin. The data gave a full agreement to characteristic signatures of the exciton insulator and provided the direct evidence for the exciton insulator in this system. It should be mentioned this experiment is the first observation of the exciton insulator with conclusive evidence. Our results pointed out the importance
of charge interactions in properties of the QSH effect, in addition to single-particle band theories. Our findings opened up new avenues for the exploration and control of exotic interacting TI matter.

The observed bulk exciton insulator and helical edge state on the boundary clearly show that BCS-like exciton condensation is topologically nontrivial with broken TRS, in this system. Whether it is a general phenomenon, still need more inputs in the theoretical aspect and more experiments in other systems.

Further, the natural single mode topological edge gave a great platform to perform 1D Coulomb drag, which in turn provided more insights regarding the novel edge state. Specially, we fabricated an airbridge-split-H bar device, where 50nm vacuum barrier laterally isolated adjacent topological edges. Without the dielectric material between edges, interedge capacitive coupling was greatly suppressed to the minimum, the interedge Coulomb interaction was strengthened, and the correlation effect could be expected. We observed the drag signal with the positive drag above 1K and the negative drag below 1K. This was not reported in previous 1D Coulomb drag results, but indicated a charge symmetry and potential exciton correlation.


Appendix

A. Optical lithography

Cleave a 5×5 mm sample piece and follow steps shown below to define pattern on the sample by optical lithography

1. Clean sample surface. Rinse the sample in following solvents in sequence:
   \[\text{Acetone}(5\text{min}) \rightarrow \text{Methanol}(5\text{min}) \rightarrow \text{DI water}(\text{min})\]. Every step is necessary.

2. Spin coat the sample with S1813 photoresist at 5000 rpm for 60 seconds.
   Thickness of photoresist film will be around 1μm.

3. Soft bake the sample at \textbf{95°C-115°C} for 60 seconds

4. Do exposure with 90 \textit{mJ/cm}^2

5. Immerse the sample in MF 321 for 30 seconds

6. Immerse the sample in water for 60 seconds

7. If the next step is wet etching, hard bake the sample at \textbf{120°C} for 120 seconds

8. Strip the photoresist in Acetone.

B. E beam lithography

1. Clean sample surface

2. Spin coat the sample with PMMA A4 at 5000 rpm for 60 seconds.
3, Bake the sample at 180°C for 120 seconds in hot plate

4, Do exposure with the dose 220 μC/cm² for the surface of GaSb in JOEL 6500, at 30 keV and 7 mm working distance.

5, Immerse the sample in MIBK:IPA=1:3 for 30 seconds

6, Immerse the sample in IPA for 60 seconds

7, No hard bake is required.

8, Strip the photoresist in Acetone.

C. Focus Ion beam lithography

1, Clean sample surface

2, Spin coat the sample with PMMA A2 at 5000 rpm for 60 seconds. The thickness should be below 50nm.

3, Bake the sample at 180°C for 120 seconds in hot plate

4, Make a pattern with integrated program in FEI 235, at 30keV.

5, Immerse the sample in MIBK:IPA=1:3 for 30 seconds

6, Immerse the sample in IPA for 60 seconds

7, No hard bake is required.

8, Strip the photoresist in PG-Remover.
D. Wet etching

1. For GaAs, do etching with etchant $H_3PO_4: H_2O_2: H_2O = 1:1:38$, whose etching rate is 120nm/min.

2. For InAs, do etching with etchant Citric: $H_2O_2 = 1:1$, where Citric is prepared with 1g $C_6H_8O_7$ : 1mL $H_2O$. 10 seconds are enough to etch 3nm InAs. This etchant will not etch GaSb or AlGaSb.

3. For GaSb and AlGaSb, do etching with MF 321/ $NH_4OH: H_2O = 1:5$, whose etching rate is 20nm/min. This etchant will not etch InAs.

4. For InAs, GaSb and AlGaSb, do etching with etchant $H_3PO_4: Citric: H_2O_2: H_2O = 3:55:5:220$, whose etching rate is nearly 30nm/min.

5. For $Al_2O_3$, do etching with etchant $HF: H_2O = 5:95$. 20 seconds are enough to etch 100nm.

E. Dry etching

1. For etching $Si_3N_4$ with Reactive Ion etching, Chamber pressure: 100m Torr, Power: 100W, 50 sccm $CF_4$, 6 sccm $O_2$, with an etching rate of 200nm/min. The machine used is Trion Loadlock.

2. For etching GaSb and AlGaSb with Reactive Ion etching, Chamber pressure: 100m Torr, Power: 100W, 30 sccm $BCl_3$, 15 sccm Ar, with an etching rate of 250nm/min. The machine used is Trion.
**F. Metallization**

Make contact without involving thin film deposition in GaAs.

1. Melt metals Indium (In) and Tin (Sn) with 1:1, and mix them to form alloy with solder iron tip.
2. Set the solder temperature to 300°C.
3. Scratch the sample surface with diamond tip and put InSn alloy to the scratched area with solder iron tip.
4. Set thermal station temperature $T_1 = 450$ °C and $T_2 = 460$ °C
5. Flush forming gas (85% N₂ and 15% H₂) for 10 minute.
6. Turn on heater for 20 minutes. Keep 450 °C for 15min.
7. Turn off heater and wait for cooling down with keeping forming gas flowing.

Make contact without involving thin film deposition in InAs/GaSb.

1. Set the solder temperature to 300°C.
2. Put In to the sample area with solder iron tip

    Make contacts used thin film deposition.

1. Make pattern with optical lithography
2. Deposit germanium, palladium and gold layers in an e-beam evaporator system, with layer thicknesses 43nm, 30nm, and 87nm, respectively. In order to ensure longevity of crucible liners, the e-beam power should be ramped up and down slowly.
3. Do lift-off. Immerse sample in Acetone and ultrasonic for 10 seconds so that all the metal films except for those on contacts windows are removed.

4. Anneal the device at 300°C in a forming gas for 3 minutes.

G. Gate

Dielectric evaporation

1. For growing $Si_3N_4$ with Plasma enhanced chemical vapor deposition with a rate of 30nm/min. The machine used is Trion Loadlock

   Chamber pressure: 600m Torr,
   Power: 50W,
   Temperature: 350°C,

   $SiH_4$: 12 sccm,
   $NH_3$: 10 sccm,
   $N_2$: 200 sccm.

2. For growing $Al_2O_3$ with Atomic layer deposition,

   Temperature: 150°C,
   Pump time: 20s,
   $N_2$ purge rate: 20sccm.

3. $Al_2O_3$ could also be deposited with ebeam evaporator, with the pressure under $10^{-5}$ Torr and the deposition rate of 6nm/min.
Gate metal evaporation with e beam evaporator

1. In the normal case, Al of 50nm or Au/Ti of 50/5nm are deposited.

2. In the optical using, Pd film of 10nm is deposited with the pressure under $10^{-5}$ Torr and the deposition rate of 1nm/min.

3. Liftoff. Acetone (5min) → Methanol (5min) → DI water (min).